SINGLE-CELL ANALYSIS OF MUSCLE STEM CELL HETEROGENEITY IN TISSUE REGENERATION

A Dissertation by

Andrea Joseph De Micheli

Presented to the Faculty of the Graduate School of

Cornell University

In Partial Fulfillment to the Requirements for the Degree of

Doctor of Philosophy

Ithaca, New York
May 2020

This work is licensed under a creative common license (CC BY-NC-ND 4.0)

Andrea De Micheli, 2020

Abstract

Skeletal muscle is an essential organ that supports many vital functions. Within skeletal muscle, a resident population of stem cells called *Muscle Stem Cells* (MuSCs) is responsible to coordinate muscle repair after injuries. Muscle tissue regeneration by MuSCs is temporally regulated at the cellular and tissue levels by stem cell differentiation and cell-communication programs. In this thesis, I use *single-cell RNA-sequencing* technology to map the transcriptome of the cell populations that compose muscle tissue and provide a new description of MuSC differentiation programs in homeostasis and in regeneration. By generating so-called *"transcriptomic atlases"*, I explore the cellular heterogeneity within myogenic differentiation and formulate new hypotheses on cell-interactions that direct muscle repair. Specifically, I first present a single-cell analysis of the mouse regenerating muscle tissue where, by ligand-receptor bioinformatic modeling, I identify new heterotypic communication channels signals involved in muscle repair. Second, I present a single-cell transcriptomic atlas of human donor muscle tissue where I describe the effect of ageing and disease on MuSCs. These studies also highlight differences between mouse and human muscle. Although mouse and human MuSCs share similar transcriptomic profiles, there are notable differences in surface proteins that are commonly used to mark those cells. Human MuSCs also appear more heterogenous and exist in two distinct subpopulations. Together, these two single-cell transcriptomic atlases provide reference resources to examine the role of muscle-tissue cellular diversity and communication in regeneration, aging, disease, and across species.

Biographical Sketch

Andrea J. De Micheli grew up in California and in Switzerland. In 2010, he started a Bachelor's degree at the *École Polytechnique Fédérale de Lausanne* (EPFL). During his undergraduate studies, Andrea developed an interest in biotechnologies, which motivated him to work for six months as a research assistant for *Diagnostics For All* (Cambridge, MA, USA). Following that research experience, Andrea decided to pursue a Master's degree in Bioengineering at EPFL and graduated from with honors in 2016. Andrea worked in the labs of Prof. Carlotta Guiducci and Prof. Matthias Lutolf, on research projects at the intersection between microtechnology and stem-cell bioengineering.

In 2015, Andrea moved back to the USA to join the *Meinig School of Biomedical Engineering* at *Cornell University* and started a PhD research project with Prof. Benjamin Cosgrove. During his doctoral studies, Andrea got involved in teaching activities and worked on a new teaching module for undergraduates on next-generation RNA-seq technology, while supported by a fellowship from the US Department of Education. At Cornell, he also got involved with the graduate student community and served as member of the Graduate and Professional Student Assembly (GPSA), where he organized the annual faculty awards. His doctoral studies also led him to collaborate at the *Englander Institute for Precision Medicine* (EIPM) and with Dr. Olivier Elemento in New York City (NYC).

In January 2020, Andrea started a postdoctoral research project at the *Hospital for Special Surgery* (HSS) in NYC with Dr. Christopher Mendias. His project will be focused on examining tendon ageing with single-cell technologies.

Beyond his passion for biotechnology research, Andrea enjoys spending time hiking in nature and playing the clarinet.

Acknowledgments

There are numerous people I would like to acknowledge and without whom the completion of this thesis would have not been possible.

Foremost, I would like to thank my mentor Prof. **Ben Cosgrove** for his unwavering support and for introducing me to the field of muscle stem cell biology. In particular, I am incredibly grateful that he encouraged me to attend many international conferences and workshops at which I was able to share my scientific work. It has truly been a privilege to learn from you.

I would like to also recognize the members of the Cosgrove lab that all played important and unique roles. In particular Dr. **Sharon Soueid-Baumgarten** for introducing me to the lab in my early days at Cornell and for teaching me many techniques in biology. I would like to also acknowledge the first wave of Cosgrove lab members that helped me get started back in 2015: **Francis Chen**, **Brenton Munson**, **Victor Aguilar**, and my friend **Alex Loiben**; all of you have played a significant role in my scientific and personal life. I was also fortunate to have worked with brilliant undergraduates that have supported me and believed in my ideas. In particular **Ashritha Bheemidi** for helping with the bead project and **Paula Fraczek** for her meticulous work with mice.

The projects undertaken during my PhD also led me to collaborate with other labs and scientists that were all influential for the completion of the projects in this thesis. Most importantly, I would like to thank Prof. **Olivier Elemento** at Weill Cornell Medicine and his lab for hosting me and providing me with all the resources and expertise I needed in computation biology and bioinformatics. I would also like to acknowledge the people from the Barbieri lab (**Ivana**, **Lesa**, and many others…) that made me feel welcomed at WCM when I moved to NYC. I would also like to thank Dr. **Christopher Mendias** at HSS and his lab for their support, as well as the members of my thesis committee, in particular Prof. **Iwijn De Vlaminck** for his expertise in single-cell analysis and the many other constructive conversations.

Last but not least, I would like to dedicate this work to my partner **Héléna** and to **my family** for all their love and encouragements,

Andrea

Contributions

I would like to acknowledge the work of my colleagues and friends that contributed to this thesis:

Chapter 3: Andrea De Micheli and Benjamin Cosgrove designed the study and wrote the manuscript. Andrea, Paula Fraczek, and Sharon Soueid-Baumgarten organized the mouse colony and procedures and performed the scRNA-seq study. Andrea and Hiranmayi Ravichandran performed the CyTOF study. Andrea with assistance from Ben, Iwijn De Vlaminck, and Olivier Elemento, performed the scRNA-seq and CyTOF data analysis. Emily Laurilliard and Charles Heinke performed the ligand screen experiment.

Chapter 4: Andrea and Benjamin Cosgrove designed the study and wrote the manuscript. Dr. Jason Spector obtained the human tissue samples. Andrea performed the tissue dissociations, scRNA-seq, and data analysis, with assistance from Ben and Olivier Elemento.

Appendix 1: Andrea designed the micro-bead platform with assistance from Benjamin Cosgrove. Andrea and Ashritha Bheemidi developed and optimized the platform.

Appendix 2: Andrea and Benjamin Cosgrove wrote the review paper.

Chapters 1, **2**, and **5** were written by Andrea. Figures were drawn by Héléna Macherel.

Thesis Committee

Prof. Benjamin D. Cosgrove (Chair)
Prof. Claudia Fischbach-Teschl
Prof. Jan Lammerding
Prof. Iwijn De Vlaminck

Table of Contents

List of figures and tables

List of key abbreviations

ATAC-seq – Assay for Transposase-Accessible Chromatin using sequencing

CCA – Canonical correlation analysis

CyTOF – Cytometry by time-of-flight (or mass cytometry)

ECM – Extracellular matrix

FACS – Fluorescence activated cell sorting

FAPs – Fibro/adipogenic progenitors

LR – Ligand-receptor (interaction)

mRNA – Messenger RNA

MuSC – Muscle stem cell

PCA – Principal component analysis

RBC – Red blood cell

scRNA-seq – Single-cell RNA-sequencing

Sdc – Syndecan

SNN – Shared nearest neighbor

TA – Tibialis anterior muscle

UMAP – Uniform Manifold approximation and projection

UMI – Unique molecular identifier

Introduction

This doctoral thesis in biomedical engineering studies muscle stem cells and muscle tissue regeneration using single-cell bioinformatic analysis and modeling.

Stem cells are capable to self-renew (proliferate into more stem cells) and differentiate (transform) into other cell types (Zakrzewski et al., 2019). A single embryonic stem cell can for instance produce an entire multi-cellular organism. In the adult, stem cells are rare but found throughout the body. They have the immense responsibility to maintain and regenerate tissues. Upon differentiation, stem cell progenitors replace cells that are either lost because of tissue turnover or injury. For example, the gut epithelium is restored by stem cells buried inside the tissue's crypts (Santos et al., 2018); the blood lineage is replenished by hematopoietic stem cells in the bone marrow (Pinho and Frenette, 2019); and, after injury, the structure and contractile ability of skeletal muscle fibers is restored by muscle stem cells (Almada and Wagers, 2016).

Skeletal muscle is an essential organ that support many vital functions such as locomotion, respiration, digestion, and heat production (Mukund and Subramaniam, 2019). Muscle tissue is composed of bundles of multinucleated myofiber cells with muscle stem cells (MuSCs) located between their plasma membrane and a basal lamina of extra-cellular proteins (Fig. 1). Upon muscle injury, MuSCs activate to repair the tissue. Muscle tissue regeneration by MuSCs is a complex process organized at the cellular and

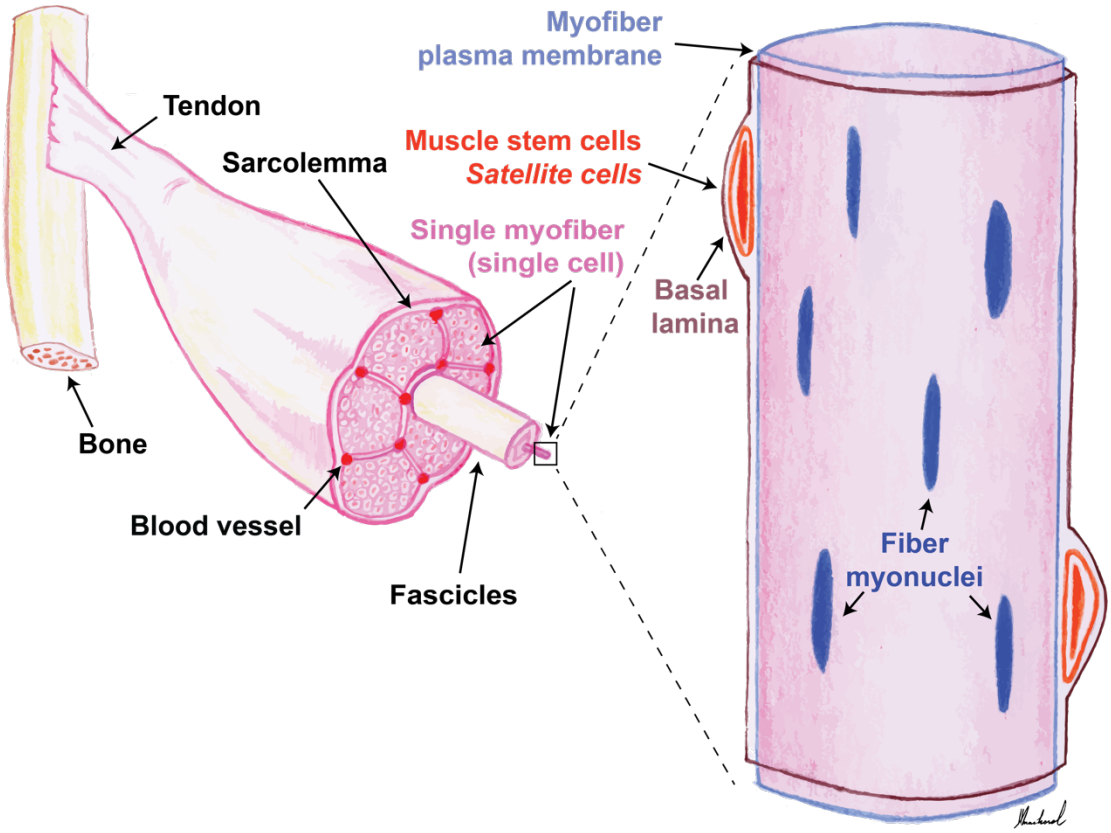


Figure 1 – Anatomy of muscle tissue and location of muscle stem cells (MuSCs). Skeletal muscle tissue is organized in successive bundles of fibers, starting with bundles of fascicles, which themselves are bundles of multinucleated myofibers. MuSCs are located at the periphery of myofibers between the myofiber plasma membrane and a basal lamina of extra-cellular matrix (ECM) proteins.

tissue levels (Yin et al., 2013). For instance, MuSCs activate specific gene regulatory programs to leave quiescence, expand in numbers, and differentiate to form new muscle fibers. This process is also stimulated by a wide panel of extrinsic signaling molecules from the local muscle microenvironment that act on MuSCs through their surface receptors (Mashinchian et al., 2018). Other types of cells, in particular immune cells and fibro/adipogenic progenitors, also contribute to the repair process in close communication with MuSCs (Bentzinger et al., 2013a). Though inherently a self-regenerating organ, a number of conditions and diseases can affect the ability of MuSCs to maintain muscle tissue, with systemic consequences on health and wellbeing. Notably, sarcopenia, the loss of muscle mass and function in the elderly can be attributed to dysfunctional MuSCs

(Ryall et al., 2008). Sarcopenia affect 10-30% of the world's population and represent an important public heath burden (Beaudart et al., 2014; Shafiee et al., 2017). Muscle stem cells have therefore been extensively investigated as a potential source of therapeutic regeneration. Nonetheless, many fundamental questions in muscle stem cell biology remain to be resolved.

For example, the molecular and genetic mechanisms that drive MuSC activation and muscle maturation are still incompletely understood, which limits our ability to control MuSCs *in vitro* and use them efficiently in cell-therapies. Elucidating the communication signals between MuSCs and its native environment will permit us to recapitulate these communication signals in bioengineered *in vitro* constructs to control desired stem cell programs (Cosgrove et al., 2009a). Another specific challenge remains the ability to expand the number of MuSCs in culture through self-renewal divisions without affecting its potential to generate new functional muscle. Furthermore, there is also a need to further refine our molecular definition of MuSCs through surface markers and associate new molecular signatures to regenerative phenotypes. New surface markers will allow us to more selectively purify highly regenerative cells for therapeutic transplantation. All in all, bioengineering approaches promise to address some of these questions in novel and powerful ways, in particular using single-cell technologies and bioinformatic models.

Single-cell technologies are enabling fine-grained and unbiased examination of biology and physiological processes at the level of the individual cell (Svensson et al.,

2018). In particular, high-throughput *single-cell RNA-sequencing* (scRNA-seq) have permitted us to profile the transcriptomes of thousands of individual cells in parallel and build *"cellular atlases"* of tissues and organs (Regev et al., 2017). These atlases are "transcriptomic datasets" that describe the gene expression profile of tissues with single-cell resolution and constitute a basis to generate, in a data-driven manner, new biological hypothesis (Rood et al., 2019).

Single-cell technology is also redefining the identity of cells from increasingly diverse molecular feature sets. Conceptually, this has permitted to place collections of cells, and consequently entire biological systems, inside a high-dimensional space of identities, states and functions. This space provides a powerful framework to model biological events such as cellular decisions, genetic regulatory switches, differentiation trajectories, and cell-cell communication networks (Hwang et al., 2018). In this thesis we use scRNA-seq to generate single-cell atlases from the mouse muscle tissue during regeneration as well as from human muscle biopsies. Powered by those datasets, we developed a bioinformatic framework to examine MuSC and muscle tissue cellular heterogeneity and suggest new molecular markers through ligand-receptor communication and myogenic trajectory modeling.

This thesis is divided into five chapters. In **Chapter 1** I review the biology of MuSCs, with a particular focus on tissue regeneration, surface markers of the myogenic lineage, and ligand-receptor interactions between MuSCs, its microenvironment, and

other cell types. In **Chapter 2** I review scRNA-seq technologies and bioinformatic methods to study muscle. In particular I discuss muscle tissue sample preparation and considerations for scRNA-seq, as well as specific data elaboration methods, and bioinformatic models to model myogenesis and ligand-receptor interactions. In **Chapter 3** I present a comprehensive scRNA-seq atlas and analysis of the mouse regenerating muscle tissue. I also present a bioinformatic ligand-receptor model from which I identify new communication signals involved in myogenic differentiation. In **Chapter 4** I apply the bioinformatic analysis framework developed in Chapter 3 to human muscle samples. I discuss specific challenges in data analysis and interpretation associated to human clinical samples. Finally, I provide in **Chapter 5** a summary and future directions for the work presented herein.

Introduction references

Almada, A.E., and Wagers, A.J. (2016). Molecular circuitry of stem cell fate in skeletal muscle regeneration, ageing and disease. Nature Reviews Molecular Cell Biology 17, 267–279.

Beaudart, C., Rizzoli, R., Bruyère, O., Reginster, J.-Y., and Biver, E. (2014). Sarcopenia: burden and challenges for public health. Arch Public Health 72, 45.

Bentzinger, C.F., Wang, Y.X., Dumont, N.A., and Rudnicki, M.A. (2013). Cellular dynamics in the muscle satellite cell niche. EMBO Reports 14, 1062–1072.

Cosgrove, B.D., Sacco, A., Gilbert, P.M., and Blau, H.M. (2009). A home away from home: Challenges and opportunities in engineering in vitro muscle satellite cell niches. Differentiation 78, 185–194.

Hwang, B., Lee, J.H., and Bang, D. (2018). Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp Mol Med 50, 96.

Mashinchian, O., Pisconti, A., Le Moal, E., and Bentzinger, C.F. (2018). The Muscle Stem Cell Niche in Health and Disease. In Current Topics in Developmental Biology, (Elsevier), pp. 23–65.

Mukund, K., and Subramaniam, S. (2019). Skeletal muscle: A review of molecular structure and function, in health and disease. WIREs Syst Biol Med.

Pinho, S., and Frenette, P.S. (2019). Haematopoietic stem cell activity and interactions with the niche. Nat Rev Mol Cell Biol 20, 303–320.

Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., et al. (2017). The Human Cell Atlas. ELife 6, e27041.

Rood, J.E., Stuart, T., Ghazanfar, S., Biancalani, T., Fisher, E., Butler, A., Hupalowska, A., Gaffney, L., Mauck, W., Eraslan, G., et al. (2019). Toward a Common Coordinate Framework for the Human Body. Cell 179, 1455–1467.

Ryall, J.G., Schertzer, J.D., and Lynch, G.S. (2008). Cellular and molecular mechanisms underlying age-related skeletal muscle wasting and weakness. Biogerontology 9, 213–228.

Santos, A.J.M., Lo, Y.-H., Mah, A.T., and Kuo, C.J. (2018). The Intestinal Stem Cell Niche: Homeostasis and Adaptations. Trends in Cell Biology 28, 1062–1078.

Shafiee, G., Keshtkar, A., Soltani, A., Ahadi, Z., Larijani, B., and Heshmat, R. (2017). Prevalence of sarcopenia in the world: a systematic review and meta- analysis of general population studies. J Diabetes Metab Disord 16, 21.

Svensson, V., Vento-Tormo, R., and Teichmann, S.A. (2018). Exponential scaling of single-cell RNA-seq in the past decade. Nature Protocols 13, 599–604.

Yin, H., Price, F., and Rudnicki, M.A. (2013). Satellite Cells and the Muscle Stem Cell Niche. Physiological Reviews 93, 23–67.

Zakrzewski, W., Dobrzyński, M., Szymonowicz, M., and Rybak, Z. (2019). Stem cells: past, present, and future. Stem Cell Res Ther 10, 68.

Chapter 1
Muscle stem cell biology

1.1 Muscle stem cells repair muscle tissue upon injuries

Skeletal muscle is a self-renewing organ that supports many vital functions such as locomotion, respiration, digestion, and metabolism (Mukund and Subramaniam, 2019). Muscle tissue is maintained by muscle stem cells (MuSCs), an essential adult population of stem cells that repair muscle fibers (or myofibers) throughout life (Dumont et al., 2015). MuSCs have been extensively investigated given their regenerative potential and promise to treat myopathies in stem cell-based therapies (Cerletti et al., 2008; Tedesco and Cossu, 2012).

Muscle stem cells were first reported by Alexander Mauro in 1961 during an electron microscope study of the peripheral region of myofibers in the frog (Mauro, 1961). Mauro described MuSCs as tiny cells, with *"sticking paucity of cytoplasm relative to its nucleus"* and *"wedged between the plasma membrane of muscle fibers and the basement membrane"* (Mauro, 1961). This membrane bound compartment will later be referred as the *muscle stem cell niche;* a microenvironment known today to play an essential role at directing the fate of MuSCs (Yin et al., 2013). Given the peripherical location of the niche with respect to the myofiber, MuSCs are also referred as *satellite cells*.

About four decades after their anatomical discovery MuSCs were functionally characterized as a self-sufficient, self-renewing, and heterogeneous source of muscle tissue regeneration (Collins et al., 2005; Kuang et al., 2007). As other adult stem cells,

MuSCs possess both the capacity to self-renew and to differentiate into mature and functional myofibers. These hallmark stem cell properties were demonstrated by a series of transplantation experiments, typically by injecting genetically labeled MuSCs into an injured mouse muscle to stimulate repair from the transplanted cells. For example, Collins et al. (2005) showed that transplanting single healthy myofibers (with as few as 7 MuSCs) into a dystrophic mouse was sufficient to repair the damaged muscle, and consequently provided key evidence that MuSCs contributed to tissue repair by a mechanism of differentiation (Collins et al., 2005). Similarly, Montarras et al. (2005) demonstrated that MuSCs isolated from muscle by fluorescence-activated cell sorting (FACS) contributed to muscle repair when engrafted (Montarras et al., 2005). Furthermore, Kuang et al. (2007) revealed that MuSCs are an heterogenous and hierarchical cell population with the capacity to asymmetrically self-renewed to produce two molecularly distinct cells: a quiescent stem cell and a progenitor committed to myogenic differentiation (Kuang et al., 2007). Finally, the capacity of MuSCs to self-renew was further demonstrated by Sacco et al. (2008) where she showed by bioluminescence imaging that a single MuSC was capable to regenerate the tissue's quiescent stem cell pool (Sacco et al., 2008). Together, these experiments are some among others that fundamentally demonstrated that Mauro's *satellite cells* are an adult stem cell population capable of repairing muscle tissue.

1.2 Myogenesis and the muscle stem cell lineage

It is now widely accepted that in the homeostatic muscle MuSCs are mostly quiescent (Snow, 1977) and characterized by the expression of the paired box protein 7 (Pax7) transcription factor (Seale et al., 2000; Kuang et al., 2006). To maintain muscle

tissue, MuSCs divide following either one of two mechanisms: symmetric or asymmetric self-renewal (Dumont et al., 2015) (Fig. 2). Symmetric divisions produce two identical *Pax7*+ MuSCs and serves to maintain the tissue's stem cell reservoir. Upon muscle injury, MuSCs are alerted (Rodgers et al., 2014) into a *Pax7*+ *Myf5*+ activated state (Beauchamp et al., 2000) and then divide in an asymmetric fashion to produce two molecularly distinct progenies: one retaining its quiescent phenotype while the other primed to proliferate into a committed *Pax7*– *Myf5*+ *MyoD*+ progenitor called *myoblast* (Cornelison and Wold, 1997; Cornelison et al., 2000). Subsequently, myoblasts differentiate to generate fusion competent Myogenin+ *myocytes*, which finally fuse into multi-nucleated myofibers (Nabeshima et al., 1993; Zammit, 2017). Some activated MuSCs also return to quiescence. Fused myofibers are characterized by expression of Myosin heavy chain (*Myh1*), Desmin (*Des*), α-actin (*Acta1*), and other contractile machinery, become surrounded by an ECM basal lamina, innervated by motor neurons at neuro-muscular junctions to provide contractile organization from the peripheral nervous system (Mukund and Subramaniam, 2019).

1.3 Surface markers of mouse and human MuSCs

Myogenic progenitors have also been characterized by a multitude of surface receptors (Fig. 2). These markers can be used to purify MuSCs by FACS for *in vitro* studies, transplantation experiments, and therapeutic applications (Maesner et al., 2016).

Surface proteins that mark MuSCs include α7-integrin (Sherwood et al., 2014), β1-integrin (Kuang et al., 2007; Cerletti et al., 2008), Cd34 (Beauchamp et al., 2000), Cxcr4 (Cerletti et al., 2008), Syndecan-3/4 (Cornelison et al., 2001), Vcam1 (Liu et al., 2015),

M-cadherin (Irintchev et al., 1994), Neural cell adhesion molecule (NCAM) (Capkovic et al., 2008), C-met (Cornelison and Wold, 1997), and Abgc2 (Tanaka et al., 2009), and have been used in various combinations to describe largely overlapping mouse myogenic populations (Maesner et al., 2016). These markers, however, have been primarily evaluated in mouse models rather than from human samples and their molecular identity between species vary. For example, unlike in mouse, human MuSCs are not CD34-positive (Péault et al., 2007). Instead, human MuSCs are β1-integrin+ (ITGB1+) CD56+ (NCAM+) EGFR+ and/or CD82+ (Pisani et al., 2010; Charville et al., 2015; Alexander et al., 2016; Uezumi et al., 2016; Wang et al., 2019b).

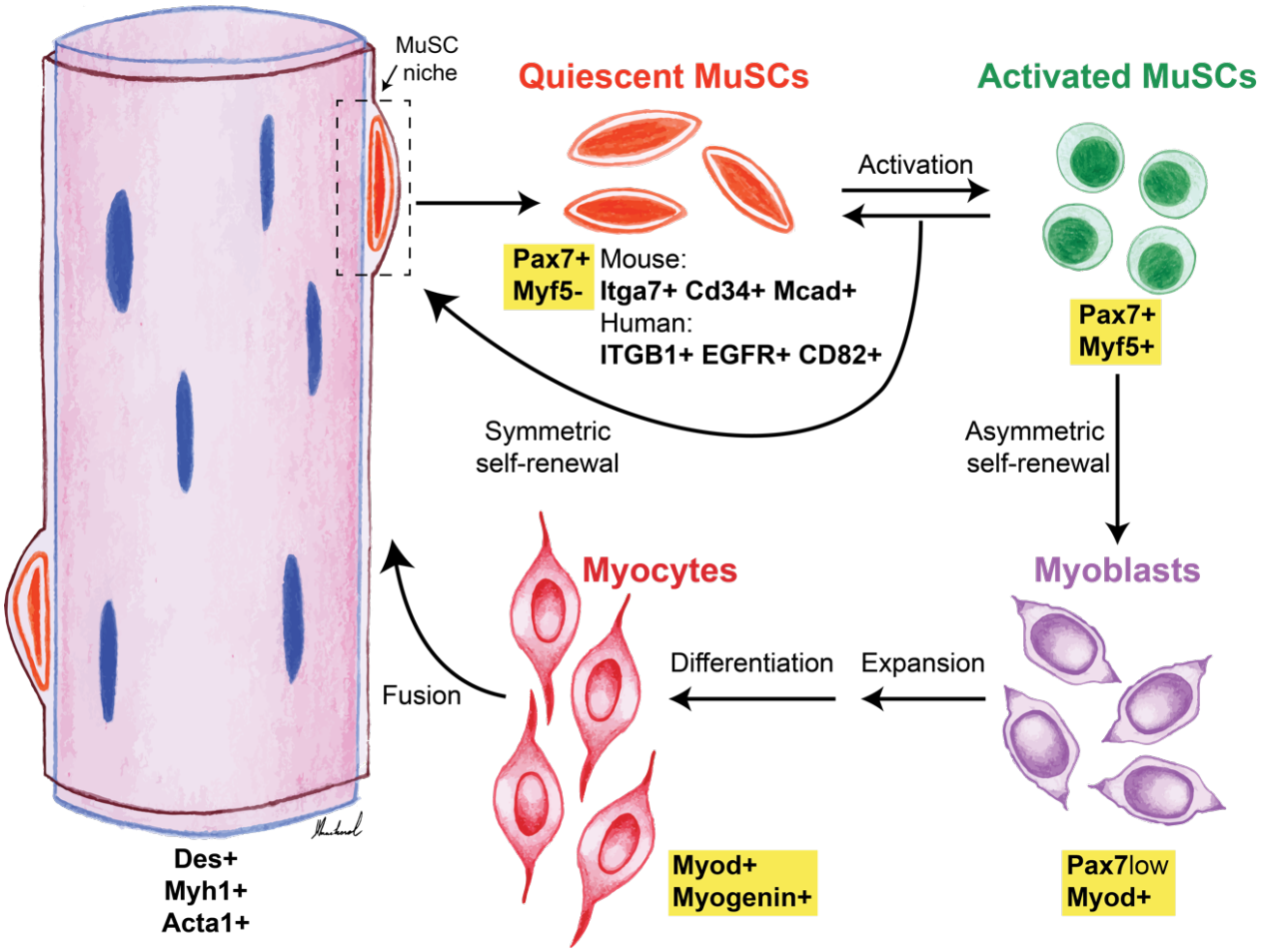


Figure 2 (*previous page)* **– Myogenesis and the muscle stem cell differentiation lineage.** After injury, quiescent MuSCs are activated and divide either by symmetric or asymmetric self-renewal. Symmetric divisions yield two identical quiescent MuSCs. Asymmetric divisions will produce a myoblast committed to myogenic differentiation. Subsequently, myoblasts will expand in number and further differentiate into myocytes. Myocytes will fuse together to form new multinucleated myofibers. A multitude of markers delineate these myogenic subpopulations such as surface receptors, cytoplasmic proteins, and transcription factors (highlighted in yellow).

Gene and protein expression studies have revealed that MuSCs are highly heterogeneous and demonstrate diverse regenerative capabilities (Dumont et al., 2015). Therefore, a finer molecular definition of this heterogeneity from MuSC surface markers will help us better characterize, purify, and study those cells. In this thesis we propose to use single-cell RNA-seq and bioinformatic analysis to identify new markers that delineate myogenic subpopulations in mouse and human.

1.4 Ligand-receptor interactions within the muscle stem cell niche

We presented in the previous section surface markers that characterize MuSCs. These markers also have important functional roles in mediating ligand-receptor interactions within the MuSC niche (Yin et al., 2013).

Muscle stem cells reside in a complex, molecularly defined, and dynamic microenvironment termed *niche* (Fig. 3). The niche exposes MuSCs to complex sets of cues, of both biochemical and biophysical nature, that regulate biological programs of quiescence, activation, self-renewal, and differentiation (Kuang et al., 2008; Cosgrove et al., 2009b; Reilly and Engler, 2010). Stem cells respond to biomolecules through surface receptors that activate specific intracellular signaling pathways regulating gene expression and fate choices. Inside the MuSC niche these biomolecular ligands are presented in a bipolar fashion (Fig. 3). On the apical side, MuSCs interact with ligands

expressed by myofibers. On the basal side, MuSCs are exposed to ECM proteins from the basal lamina and growth factors trapped within. This bipolar organization of niche ligands regulates MuSC homeostasis and division decisions (Kuang et al., 2007; Le Grand et al., 2009; Troy et al., 2012).

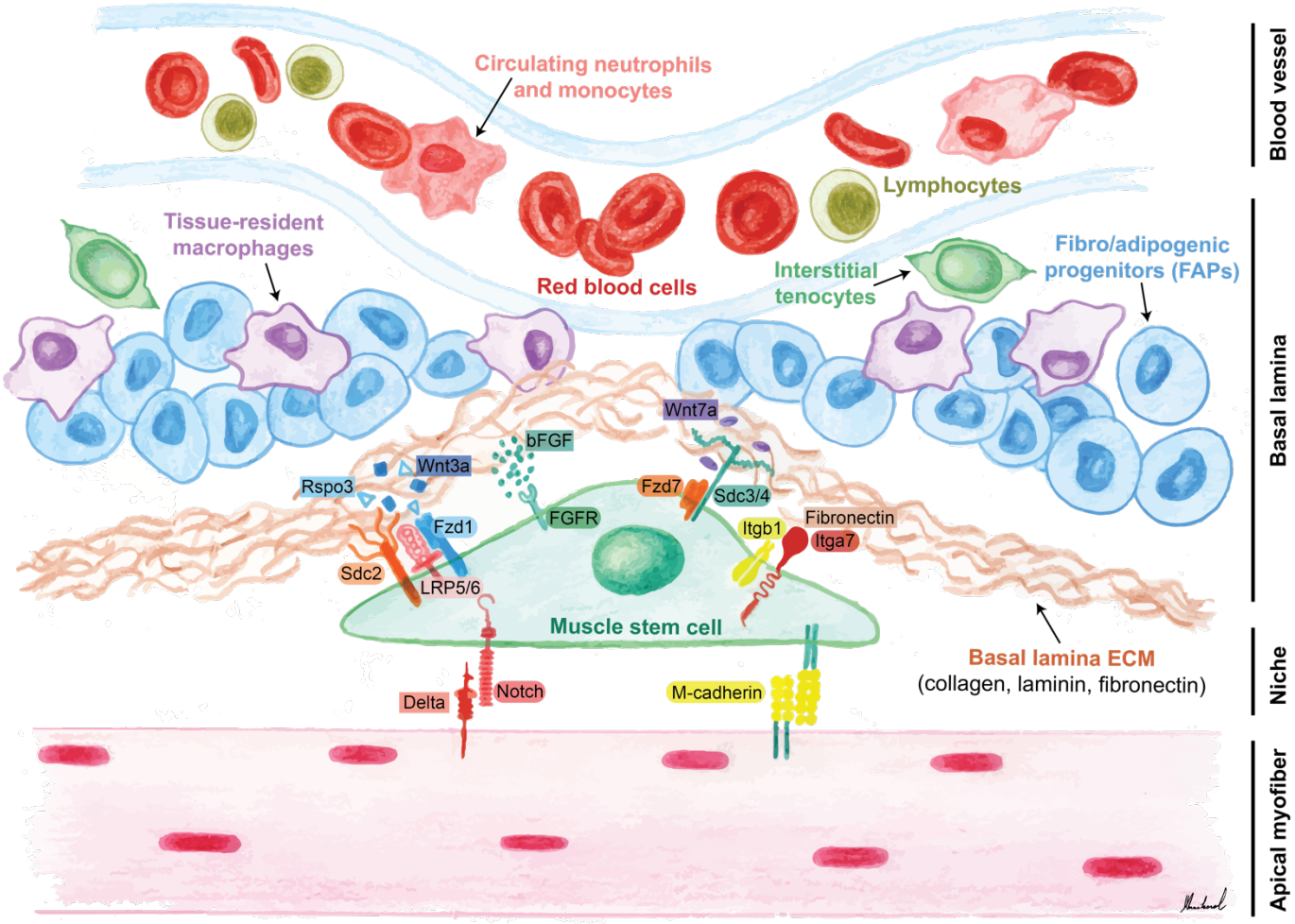


Figure 3 – The architecture of the MuSC niche and its molecular and cellular constituents. The MuSC niche is located at the periphery of the myofibers between the plasma membrane of apical myofibers and a basal lamina rich in ECM proteins. The niche provides MuSCs with signaling molecules from ECM proteins as well as ligands expressed at the surface of myofibers. The ECM also serves a reservoir for growth factors and cytokines secreted by tissue-resident (macrophages, FAPs, tenocytes) cells or from the blood circulation (neutrophils, monocytes, lymphocytes, after injury). Some ligands (squared box) and receptors (rounded box) discussed and proposed in this thesis are depicted above. Ligand-receptor interactions within the niche are very complex, dynamic, and essential for properly regulating MuSCs.

Apical myofibers produce numerous surface or secreted factors that regulate MuSC fate. For example, the surface receptor M-cadherin (*Cdh15* gene) is expressed by both MuSCs and myofibers and provides a homotypic adhesive interaction on the apical side of the niche (Irintchev et al., 1994). This interaction has been shown to regulate quiescence as its disruption, for example following to muscle injury, lead to MuSC activation (Goel et al., 2017). Most notably, quiescent MuSCs express Notch (*Notch1-4*) receptors that interact with the Delta/Jagged family of ligands (*Dll1*, *Dll4*, *Jag1*, *Jag2*) present on the surface of myofibers and possibly other cell types as well (Mourikis and Tajbakhsh, 2014). Notch signaling is require to maintain MuSC quiescence (Bjornson et al., 2012), promote self-renewal after injury (Low et al., 2018), and prevent premature differentiation (Conboy and Rando, 2002; Conboy et al., 2003; Wen et al., 2012). In addition, the Notch downstream pathway involves the RBPJ transcription factor whose targets are *Pax7*, *Hes,* and *Hey* (two repressors of *MyoD*) (Buas and Kadesch, 2010). Deletion of *RBPJ* leads to spontaneous MuSC activation and premature differentiation (Bjornson et al., 2012). Myofibers also secrete SDF-1, a ligand for the MuSC receptor CXCR4, which stimulates MuSC migration after injury (Ratajczak, 2003).

The basal side of the niche offers MuSCs adhesion and signaling cues from a multitude of ECM proteins (Sanes, 2003). For instance, MuSCs express high level of integrin receptors α7 and β1 (*Itga7*, *Itgb2*), that bind laminins, collagen IV, fibronectin, as well as other ECM proteins to support MuSC survival (Blanco-Bose et al., 2001). Particularly, integrin-laminin interactions are important to guide myoblast differentiation (Yao et al., 1996) and integrin-collagen IV interactions are necessary for MuSC self-

renewal (Urciuolo et al., 2013). MuSCs also express high levels of FGF receptor 1 and 4 (*FGFR1, FGFR4*) (Pawlikowski et al., 2017). One of its ligands, FGF2, is primarily found embedded in the basal lamina ECM of regenerating muscle tissue and acts as a mitogen, to encourage MuSC symmetric divisions while inhibiting myogenesis (Yablonka-Reuveni et al., 2015).

Other important MuSC surface receptors that we specially investigate in this thesis are Syndecans (*Sdc1-4*). Thanks to their heparan sulfate side chains, these proteoglycans attract signaling molecules to form supramolecular complexes with other receptors such as FGFR and Notch (Pisconti et al., 2012). In particular, Syndecan-3 has been shown to interact with Notch1 (Pisconti et al., 2010). Genetic deletion of *Sdc3* impairs Notch signaling and MuSC quiescence, while weakening MuSC-niche interactions. Interestingly, the *Sdc3*$^{-/-}$ mouse mutant display improved muscle regeneration after injury, though failure of MuSC to return to quiescence and home back to their niche provokes the accumulation of an interstitial Myf5+/Pax7- cell population (Pisconti et al., 2012, 2016). Syndecan-4 is also a common marker for MuSCs and interacts with FGF and HGF receptors (Cornelison, 2004). Syndecan-4 activity has been shown to be required for normal MuSC activation and proliferation. Compared to *Sdc3*$^{-/-}$ mutants, *Sdc4*$^{-/-}$ muscle display delayed regeneration after injury. Syndecan-4 also forms a co-receptor complex with Frizzled-7 (*Fzd7*) and binds fibronectin (Bentzinger et al., 2013b), another key interaction that directs the muscle repair process. In particular, loss of fibronectin from the niche affects regeneration and has been defined as a hallmarks of muscle tissue ageing (Lukjanenko et al., 2016).

Finally, the ECM also acts as a reservoir for growth factors from local or systemic sources. Notably, Wnt7a binds to the Sdc4-Fzd7 co-receptor complex and fibronectin to stimulate the symmetric expansion of MuSCs through the planar cell polarity pathway (Le Grand et al., 2009; Bentzinger et al., 2013b). Overexpression of Wnt7a increases MuSC number and enhances regeneration in dystrophic muscle (Le Grand et al., 2009). Wnt3a, on the contrary, have been shown to promoted myogenic differentiation and have been associated with fibrosis and ageing (Brack et al., 2007).

Niche ligand-receptor interactions are essential for MuSC fate decisions. Dysregulation of these interactions, for example following receptor loss of function, downstream signaling pathway disruption, or variation in niche ECM ligand content, can lead to regenerative failures and myopathies (Evano and Tajbakhsh, 2018). In this thesis we develop a ligand-receptor interaction model to unbiasedly investigate MuSC-niche and cell-cell communication signals in the regenerating muscle. Identifying new interactions through single-cell data modeling will allow us to better understand the factors that governs myogenic differentiation in health and disease.

1.5 Non-myogenic participants of muscle regeneration and their cell-signaling molecules

Non-myogenic cell types such as immune cells, fibro/adipogenic progenitors (FAPs), and endothelial cells also participate to muscle regeneration (Bentzinger et al., 2013a). While these cells have district roles during the repair process (such as debris clearing, resolution of inflammation, basal lamina and neo-vasculature restoration), they

also exhibit a potent influence on MuSCs and regulate myogenic differentiation through ligand-receptor and paracrine signaling. In this section we describe the cell types and their cell-signaling molecules that affect myogenesis.

Muscle regeneration can be divided into a pro- and anti-inflammatory phase, during which immune cells secrete specific cytokines and signaling molecules in a strictly controlled and temporal fashion (Tidball, 2017). These molecules affect each stage of myogenesis, from MuSC activation to myocyte fusion.

The pro-inflammatory phase immediately begins after myofiber damage and death. The rupture of the myofiber membrane releases inflammatory factors containing damage-associated molecular patters (DAMPs) that triggers the complement cascade, which rapidly attract Ly6G+ Cd11b+ neutrophils within two hours post-injury (Fielding et al., 1993). Neutrophils invade the damaged area and remove necrotic tissue and cellular debris by phagocytosis and by producing myeloperoxidase and free radicals. Neutrophils also secrete chemokines (MIP-1α, MCP-1) to recruit monocytes and macrophages (Scapini et al., 2000; Shireman et al., 2007). Pro-inflammatory Ly6C+ monocytes and M1 macrophages are present during the first two days of regeneration and secrete cytokines that stimulate, directly or indirectly, myogenesis. For example, during the early stages of regeneration M1-macrophages secrete the interferon gamma molecule (IFN-γ) that represses *Myog* on MuSCs through the JAK-STAT1 pathway (Londhe and Davie, 2013). This assures myoblasts remain in an undifferentiated but proliferative state. Later, immune cells secrete the tumor necrosis factor (TNF) to prime MuSCs and myoblasts for

differentiation. TNF signaling through the MAPK/p38a pathway provokes the epigenetic silencing of *Pax7* and *Notch1* genes, thereby ending MuSC quiescence and promoting proliferation (Palacios et al., 2010). Similarly, the TNF-like weak inducer of apoptosis factor (TWEAK) secreted my macrophages can also promote MuSC proliferation through the canonical NF-kB and Notch pathways as well as myocyte fusion through the non-canonical NF-kB pathway (Enwere et al., 2012, 2014).

After the inflammatory phase, immune cells transition to an anti-inflammatory phenotype (Ly6C- monocytes and M2 macrophages) and remain in the tissue for up to 10 days (Chazaud et al., 2003; Arnold et al., 2007). M2 macrophages encourage regeneration by secreting anti-inflammatory factors, such as IL-10, TGF-β, IGF-1, and VEGF, which further commit myoblasts to differentiate and monocyte to fuse. For example, IGF-1 is a strong mitogen for muscle progenitors and enhances muscle growth (Duan et al., 2010; Valdés et al., 2013). Recently, TGF-β has also been proposed to modulate myocyte fusion independently from differentiation (Girardi et al., 2019). Most of these secreted immune factors have also been evaluated in culture and promote to different extent the differentiation of myoblasts (Saclier et al., 2013).

Lymphoid cells are also present in injured muscle. In particular, Cd4+ Treg have been shown to stimulate muscle repair through secreting the growth factor amphiregulin, which directly acts on the epithelial growth factor receptor (EGFR) expressed by MuSCs (Burzyn et al., 2013).

During muscle repair, fibroblasts and fibro/adipogenic progenitors (FAPs) remodel the extracellular basal lamina and can also help macrophages clear debris (Heredia et al., 2013). Similar to MuSCs, FAPs are a *PDGFRα*+ quiescent mesenchymal stem cell population found in muscle tissue in the vicinity of blood vessels. Upon injury, FAPs activate, proliferate (peak at day 5 post-injury), and differentiate into fibroblasts or adipocytes (Joe et al., 2010; Uezumi et al., 2011; Pretheeban et al., 2012). Their main role is to deposit new ECM by secreting laminins, collagen, fibronectin, and matrix metalloproteinases, in order to restore the basal lamina. Furthermore, FAPs have also been hypothesized to closely communicate with MuSCs during regeneration, though few interactions have been fully characterized to date. In one study, FAPs have been shown to secrete the interleukin IL-6 which stimulate MuSCs to differentiate (Joe et al., 2010). In return, myogenic progenitors appear to inhibit FAP-to-adipocyte differentiation (Uezumi et al., 2010; Heredia et al., 2013). In a more recent study, genetic deletion FAPs in a transgenic PDGFRα^{CreER} mouse provoked loss of MuSCs in the homeostatic muscle (Wosczyna et al., 2019), thus suggesting that there are likely numerous cell-signaling molecules secreted by FAPs that are required for muscle maintenance and repair.

Endothelial cells also promote the proliferation of muscle progenitors by secreting VEGF (Christov et al., 2007). Myoblasts also secrete VEGF to stimulate the growth of a new vasculature (Rhoads et al., 2009). Interestingly, smooth muscle cells have been shown to be important for MuSC re-entry into quiescence (Abou-Khalil et al., 2010).

Non-myogenic cell populations participate to muscle repair through complex networks of cell-signaling molecules. In this thesis, we use scRNA-seq to propose through ligand-receptor modeling new cell-communication channels involved in muscle repair. Modeling these communications channels will help us identify cells and molecules that govern MuSC maintenance and myogenic differentiation. These factors can be used in co-culture studies to stimulate muscle repair or in therapeutic models, for example to form muscle constructs *in vitro* or to manipulate the muscle immune system *in vivo.*

1.6 Chapter references

Abou-Khalil, R., Mounier, R., and Chazaud, B. (2010). Regulation of myogenic stem cell behaviour by vessel cells: The “ménage à trois” of satellite cells, periendothelial cells and endothelial cells. Cell Cycle *9*, 892–896.

Alexander, M.S., Rozkalne, A., Colletta, A., Spinazzola, J.M., Johnson, S., Rahimov, F., Meng, H., Lawlor, M.W., Estrella, E., Kunkel, L.M., et al. (2016). CD82 Is a Marker for Prospective Isolation of Human Muscle Satellite Cells and Is Linked to Muscular Dystrophies. Cell Stem Cell *19*, 800–807.

Arnold, L., Henry, A., Poron, F., Baba-Amer, Y., van Rooijen, N., Plonquet, A., Gherardi, R.K., and Chazaud, B. (2007). Inflammatory monocytes recruited after skeletal muscle injury switch into antiinflammatory macrophages to support myogenesis. The Journal of Experimental Medicine *204*, 1057–1069.

Beauchamp, J.R., Heslop, L., Yu, D.S.W., Tajbakhsh, S., Kelly, R.G., Wernig, A., Buckingham, M.E., Partridge, T.A., and Zammit, P.S. (2000). Expression of Cd34 and Myf5 Defines the Majority of Quiescent Adult Skeletal Muscle Satellite Cells. The Journal of Cell Biology *151*, 1221–1234.

Bentzinger, C.F., Wang, Y.X., von Maltzahn, J., Soleimani, V.D., Yin, H., and Rudnicki, M.A. (2013a). Fibronectin Regulates Wnt7a Signaling and Satellite Cell Expansion. Cell Stem Cell *12*, 75–87.

Bentzinger, C.F., Wang, Y.X., Dumont, N.A., and Rudnicki, M.A. (2013b). Cellular dynamics in the muscle satellite cell niche. EMBO Reports *14*, 1062–1072.

Bjornson, C.R.R., Cheung, T.H., Liu, L., Tripathi, P.V., Steeper, K.M., and Rando, T.A. (2012). Notch Signaling Is Necessary to Maintain Quiescence in Adult Muscle Stem Cells. STEM CELLS *30*, 232–242.

Blanco-Bose, W.E., Yao, C.-C., Kramer, R.H., and Blau, H.M. (2001). Purification of Mouse Primary Myoblasts Based on α7 Integrin Expression. Experimental Cell Research *265*, 212–220.

Brack, A.S., Conboy, M.J., Roy, S., Lee, M., Kuo, C.J., Keller, C., and Rando, T.A. (2007). Increased Wnt signaling during aging alters muscle stem cell fate and increases fibrosis. Science *317*, 807–810.

Buas, M.F., and Kadesch, T. (2010). Regulation of skeletal myogenesis by Notch. Experimental Cell Research *316*, 3028–3033.

Burzyn, D., Kuswanto, W., Kolodin, D., Shadrach, J.L., Cerletti, M., Jang, Y., Sefik, E., Tan, T.G., Wagers, A.J., Benoist, C., et al. (2013). A Special Population of Regulatory T Cells Potentiates Muscle Repair. Cell *155*, 1282–1295.

Capkovic, K., Stevenson, S., Johnson, M., Thelen, J., and Cornelison, D. (2008). Neural cell adhesion molecule (NCAM) marks adult myogenic cells committed to differentiation. Experimental Cell Research *314*, 1553–1565.

Cerletti, M., Jurga, S., Witczak, C.A., Hirshman, M.F., Shadrach, J.L., Goodyear, L.J., and Wagers, A.J. (2008). Highly Efficient, Functional Engraftment of Skeletal Muscle Stem Cells in Dystrophic Muscles. Cell *134*, 37–47.

Charville, G.W., Cheung, T.H., Yoo, B., Santos, P.J., Lee, G.K., Shrager, J.B., and Rando, T.A. (2015). Ex Vivo Expansion and In Vivo Self-Renewal of Human Muscle Stem Cells. Stem Cell Reports *5*, 621–632.

Chazaud, B., Sonnet, C., Lafuste, P., Bassez, G., Rimaniol, A.-C., Poron, F., Authier, F.-J., Dreyfus, P.A., and Gherardi, R.K. (2003). Satellite cells attract monocytes and use macrophages as a support to escape apoptosis and enhance muscle growth. The Journal of Cell Biology *163*, 1133–1143.

Christov, C., Chrétien, F., Abou-Khalil, R., Bassez, G., Vallet, G., Authier, F.-J., Bassaglia, Y., Shinin, V., Tajbakhsh, S., Chazaud, B., et al. (2007). Muscle Satellite Cells and Endothelial Cells: Close Neighbors and Privileged Partners. MBoC *18*, 1397–1409.

Collins, C.A., Olsen, I., Zammit, P.S., Heslop, L., Petrie, A., Partridge, T.A., and Morgan, J.E. (2005). Stem Cell Function, Self-Renewal, and Behavioral Heterogeneity of Cells from the Adult Muscle Satellite Cell Niche. Cell *122*, 289–301.

Conboy, I.M., and Rando, T.A. (2002). The Regulation of Notch Signaling Controls Satellite Cell Activation and Cell Fate Determination in Postnatal Myogenesis. Developmental Cell 13.

Conboy, I.M., Conboy, M.J., Smythe, G.M., and Rando, T.A. (2003). Notch-Mediated Restoration of Regenerative Potential to Aged Muscle. Science *302*, 1575.

Cornelison, D.D.W. (2004). Essential and separable roles for Syndecan-3 and Syndecan-4 in skeletal muscle development and regeneration. Genes & Development *18*, 2231–2236.

Cornelison, D.D.W., and Wold, B.J. (1997). Single-Cell Analysis of Regulatory Gene Expression in Quiescent and Activated Mouse Skeletal Muscle Satellite Cells. Developmental Biology *191*, 270–283.

Cornelison, D.D.W., Olwin, B.B., Rudnicki, M.A., and Wold, B.J. (2000). MyoD–/– Satellite Cells in Single-Fiber Culture Are Differentiation Defective and MRF4 Deficient. Developmental Biology *224*, 122–137.

Cornelison, D.D.W., Filla, M.S., Stanley, H.M., Rapraeger, A.C., and Olwin, B.B. (2001). Syndecan-3 and Syndecan-4 Specifically Mark Skeletal Muscle Satellite Cells and Are Implicated in Satellite Cell Maintenance and Muscle Regeneration. Developmental Biology *239*, 79–94.

Cosgrove, B.D., Sacco, A., Gilbert, P.M., and Blau, H.M. (2009). A home away from home: Challenges and opportunities in engineering in vitro muscle satellite cell niches. Differentiation *78*, 185–194.

Duan, C., Ren, H., and Gao, S. (2010). Insulin-like growth factors (IGFs), IGF receptors, and IGF-binding proteins: Roles in skeletal muscle growth and differentiation. General and Comparative Endocrinology *167*, 344–351.

Dumont, N.A., Wang, Y.X., and Rudnicki, M.A. (2015). Intrinsic and extrinsic mechanisms regulating satellite cell function. Development *142*, 1572–1581.

Enwere, E.K., Holbrook, J., Lejmi-Mrad, R., Vineham, J., Timusk, K., Sivaraj, B., Isaac, M., Uehling, D., Al-awar, R., LaCasse, E., et al. (2012). TWEAK and cIAP1 Regulate Myoblast Fusion Through the Noncanonical NF-κB Signaling Pathway. Sci. Signal. *5*, ra75.

Enwere, E.K., LaCasse, E.C., Adam, N.J., and Korneluk, R.G. (2014). Role of the TWEAK-Fn14-cIAP1-NF-κB Signaling Axis in the Regulation of Myogenesis and Muscle Homeostasis. Front. Immunol. *5*.

Evano, B., and Tajbakhsh, S. (2018). Skeletal muscle stem cells in comfort and stress. Npj Regenerative Medicine *3*, 24.

Fielding, R.A., Manfredi, T.J., Ding, W., Fiatarone, M.A., Evans, W.J., and Cannon, J.G. (1993). Acute phase response in exercise. III. Neutrophil and IL-1 beta accumulation in skeletal muscle. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology *265*, R166–R172.

Girardi, F., Taleb, A., Giordani, L., Cadot, B., Datye, A., Ebrahimi, M., Gamage, D.G., Millay, D.P., Gilbert, P.M., and Grand, F.L. (2019). TGFβ signaling curbs cell fusion and muscle regeneration. BioRxiv 557009.

Goel, A.J., Rieder, M.-K., Arnold, H.-H., Radice, G.L., and Krauss, R.S. (2017). Niche Cadherins Control the Quiescence-to-Activation Transition in Muscle Stem Cells. Cell Rep *21*, 2236–2250.

Heredia, J.E., Mukundan, L., Chen, F.M., Mueller, A.A., Deo, R.C., Locksley, R.M., Rando, T.A., and Chawla, A. (2013). Type 2 Innate Signals Stimulate Fibro/Adipogenic Progenitors to Facilitate Muscle Regeneration. Cell *153*, 376–388.

Irintchev, A., Zeschnigk, M., Starzinski-Powitz, A., and Wernig, A. (1994). Expression pattern of M-cadherin in normal, denervated, and regenerating mouse muscles. Dev. Dyn. *199*, 326–337.

Joe, A.W.B., Yi, L., Natarajan, A., Le Grand, F., So, L., Wang, J., Rudnicki, M.A., and Rossi, F.M.V. (2010). Muscle injury activates resident fibro/adipogenic progenitors that facilitate myogenesis. Nature Cell Biology *12*, 153–163.

Kuang, S., Chargé, S.B., Seale, P., Huh, M., and Rudnicki, M.A. (2006). Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. The Journal of Cell Biology *172*, 103–113.

Kuang, S., Kuroda, K., Le Grand, F., and Rudnicki, M.A. (2007). Asymmetric Self-Renewal and Commitment of Satellite Stem Cells in Muscle. Cell *129*, 999–1010.

Kuang, S., Gillespie, M.A., and Rudnicki, M.A. (2008). Niche Regulation of Muscle Satellite Cell Self-Renewal and Differentiation. Cell Stem Cell *2*, 22–31.

Le Grand, F., Jones, A.E., Seale, V., Scimè, A., and Rudnicki, M.A. (2009). Wnt7a Activates the Planar Cell Polarity Pathway to Drive the Symmetric Expansion of Satellite Stem Cells. Cell Stem Cell *4*, 535–547.

Liu, L., Cheung, T.H., Charville, G.W., and Rando, T.A. (2015). Isolation of skeletal muscle stem cells by fluorescence-activated cell sorting. Nat Protoc *10*, 1612–1624.

Londhe, P., and Davie, J.K. (2013). Interferon-γ Resets Muscle Cell Fate by Stimulating the Sequential Recruitment of JARID2 and PRC2 to Promoters to Repress Myogenesis. Sci. Signal. *6*, ra107.

Low, S., Barnes, J.L., Zammit, P.S., and Beauchamp, J.R. (2018). Delta-Like 4 Activates Notch 3 to Regulate Self-Renewal in Skeletal Muscle Stem Cells: Regulation of Self-Renewal in Skeletal Muscle Stem Cells. Stem Cells *36*, 458–466.

Lukjanenko, L., Jung, M.J., Hegde, N., Perruisseau-Carrier, C., Migliavacca, E., Rozo, M., Karaz, S., Jacot, G., Schmidt, M., Li, L., et al. (2016). Loss of fibronectin from the aged stem cell niche affects the regenerative capacity of skeletal muscle in mice. Nature Medicine *22*, 897–905.

Maesner, C.C., Almada, A.E., and Wagers, A.J. (2016). Established cell surface markers efficiently isolate highly overlapping populations of skeletal muscle satellite cells by fluorescence-activated cell sorting. Skeletal Muscle *6*, 35.

Mauro, A. (1961). SATELLITE CELL OF SKELETAL MUSCLE FIBERS. The Journal of Biophysical and Biochemical Cytology *9*, 493–495.

Montarras, D., Morgan, J., Collins, C., Relaix, F., Zaffran, S., Cumano, A., Partridge, T., and Buckingham, M. (2005). Direct Isolation of Satellite Cells for Skeletal Muscle Regeneration. Science *309*, 2064.

Mourikis, P., and Tajbakhsh, S. (2014). Distinct contextual roles for Notch signalling in skeletal muscle stem cells. BMC Dev Biol *14*, 2.

Mukund, K., and Subramaniam, S. (2019). Skeletal muscle: A review of molecular structure and function, in health and disease. WIREs Syst Biol Med.

Nabeshima, Y., Hanaoka, K., Hayasaka, M., Esuml, E., Li, S., Nonaka, I., and Nabeshima, Y. (1993). Myogenin gene disruption results in perinatal lethality because of severe muscle defect. Nature *364*, 532–535.

Palacios, D., Mozzetta, C., Consalvi, S., Caretti, G., Saccone, V., Proserpio, V., Marquez, V.E., Valente, S., Mai, A., Forcales, S.V., et al. (2010). TNF/p38α/Polycomb Signaling to Pax7 Locus in Satellite Cells Links Inflammation to the Epigenetic Control of Muscle Regeneration. Cell Stem Cell *7*, 455–469.

Pawlikowski, B., Vogler, T.O., Gadek, K., and Olwin, B.B. (2017). Regulation of skeletal muscle stem cells by fibroblast growth factors: Regulation of Skeletal Muscle SCs by FGFs. Dev. Dyn. *246*, 359–367.

Péault, B., Rudnicki, M., Torrente, Y., Cossu, G., Tremblay, J.P., Partridge, T., Gussoni, E., Kunkel, L.M., and Huard, J. (2007). Stem and Progenitor Cells in Skeletal Muscle Development, Maintenance, and Therapy. Molecular Therapy *15*, 867–877.

Pisani, D.F., Clement, N., Loubat, A., Plaisant, M., Sacconi, S., Kurzenne, J.-Y., Desnuelle, C., Dani, C., and Dechesne, C.A. (2010). Hierarchization of Myogenic and Adipogenic Progenitors Within Human Skeletal Muscle. STEM CELLS *28*, 2182–2194.

Pisconti, A., Cornelison, D.D.W., Olguín, H.C., Antwine, T.L., and Olwin, B.B. (2010). Syndecan-3 and Notch cooperate in regulating adult myogenesis. The Journal of Cell Biology *190*, 427–441.

Pisconti, A., Bernet, J.D., and Olwin, B.B. (2012). Syndecans in skeletal muscle development, regeneration and homeostasis. Muscles Ligaments Tendons J *2*, 1–9.

Pisconti, A., Banks, G.B., Babaeijandaghi, F., Betta, N.D., Rossi, F.M.V., Chamberlain, J.S., and Olwin, B.B. (2016). Loss of niche-satellite cell interactions in syndecan-3 null mice alters muscle progenitor cell homeostasis improving muscle regeneration. Skeletal Muscle *6*, 34.

Pretheeban, T., Lemos, D.R., Paylor, B., Zhang, R.-H., and Rossi, F.M. (2012). Role of stem/progenitor cells in reparative disorders. Fibrogenesis Tissue Repair *5*, 20.

Ratajczak, M.Z. (2003). 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.

Reilly, G.C., and Engler, A.J. (2010). Intrinsic extracellular matrix properties regulate stem cell differentiation. Journal of Biomechanics *43*, 55–62.

Rhoads, R.P., Johnson, R.M., Rathbone, C.R., Liu, X., Temm-Grove, C., Sheehan, S.M., Hoying, J.B., and Allen, R.E. (2009). Satellite cell-mediated angiogenesis in vitro coincides with a functional hypoxia-inducible factor pathway. American Journal of Physiology-Cell Physiology *296*, C1321–C1328.

Rodgers, J.T., King, K.Y., Brett, J.O., Cromie, M.J., Charville, G.W., Maguire, K.K., Brunson, C., Mastey, N., Liu, L., Tsai, C.-R., et al. (2014). mTORC1 controls the adaptive transition of quiescent stem cells from G0 to GAlert. Nature *510*, 393–396.

Sacco, A., Doyonnas, R., Kraft, P., Vitorovic, S., and Blau, H.M. (2008). Self-renewal and expansion of single transplanted muscle stem cells. Nature *456*, 502–506.

Sanes, J.R. (2003). The Basement Membrane/Basal Lamina of Skeletal Muscle. Journal of Biological Chemistry *278*, 12601–12604.

Scapini, P., Lapinet-Vera, J.A., Gasperini, S., Calzetti, F., Bazzoni, F., and Cassatella, M.A. (2000). The neutrophil as a cellular source of chemokines. Immunological Reviews *177*, 195–203.

Seale, P., Sabourin, L.A., Girgis-Gabardo, A., Mansouri, A., Gruss, P., and Rudnicki, M.A. (2000). Pax7 Is Required for the Specification of Myogenic Satellite Cells. Cell *102*, 777–786.

Sherwood, R.I., Christensen, J.L., Conboy, I.M., Conboy, M.J., Rando, T.A., Weissman, I.L., and Wagers, A.J. (2014). Isolation of Adult Mouse Myogenic Progenitors: Functional Heterogeneity of Cells within and Engrafting Skeletal Muscle. 12.

Shireman, P.K., Contreras-Shannon, V., Ochoa, O., Karia, B.P., Michalek, J.E., and McManus, L.M. (2007). MCP-1 deficiency causes altered inflammation with impaired skeletal muscle regeneration. Journal of Leukocyte Biology *81*, 775–785.

Snow, M.H. (1977). Myogenic cell formation in regenerating rat skeletal muscle injured by mincing I. A fine structural study. The Anatomical Record *188*, 181–199.

Tanaka, K.K., Hall, J.K., Troy, A.A., Cornelison, D.D.W., Majka, S.M., and Olwin, B.B. (2009). Syndecan-4-Expressing Muscle Progenitor Cells in the SP Engraft as Satellite Cells during Muscle Regeneration. Cell Stem Cell *4*, 217–225.

Tedesco, F.S., and Cossu, G. (2012). Stem cell therapies for muscle disorders: Current Opinion in Neurology *25*, 597–603.

Tidball, J.G. (2017). Regulation of muscle growth and regeneration by the immune system. Nature Reviews Immunology *17*, 165–178.

Troy, A., Cadwallader, A.B., Fedorov, Y., Tyner, K., Tanaka, K.K., and Olwin, B.B. (2012). Coordination of Satellite Cell Activation and Self-Renewal by Par-Complex-Dependent Asymmetric Activation of p38α/β MAPK. Cell Stem Cell *11*, 541–553.

Uezumi, A., Fukada, S., Yamamoto, N., Takeda, S., and Tsuchida, K. (2010). Mesenchymal progenitors distinct from satellite cells contribute to ectopic fat cell formation in skeletal muscle. Nature Cell Biology *12*, 143–152.

Uezumi, A., Ito, T., Morikawa, D., Shimizu, N., Yoneda, T., Segawa, M., Yamaguchi, M., Ogawa, R., Matev, M.M., Miyagoe-Suzuki, Y., et al. (2011). Fibrosis and adipogenesis originate from a common mesenchymal progenitor in skeletal muscle. Journal of Cell Science *124*, 3654–3664.

Uezumi, A., Nakatani, M., Ikemoto-Uezumi, M., Yamamoto, N., Morita, M., Yamaguchi, A., Yamada, H., Kasai, T., Masuda, S., Narita, A., et al. (2016). Cell-Surface Protein Profiling Identifies Distinctive Markers of Progenitor Cells in Human Skeletal Muscle. Stem Cell Reports *7*, 263–278.

Urciuolo, A., Quarta, M., Morbidoni, V., Gattazzo, F., Molon, S., Grumati, P., Montemurro, F., Tedesco, F.S., Blaauw, B., Cossu, G., et al. (2013). Collagen VI regulates satellite cell self-renewal and muscle regeneration. Nature Communications *4*, 1964.

Valdés, J.A., Flores, S., Fuentes, E.N., Osorio-Fuentealba, C., Jaimovich, E., and Molina, A. (2013). IGF-1 induces IP3-dependent calcium signal involved in the regulation of myostatin gene expression mediated by NFAT during myoblast differentiation. Journal of Cellular Physiology *228*, 1452–1463.

Wang, Y.X., Feige, P., Brun, C.E., Hekmatnejad, B., Dumont, N.A., Renaud, J.-M., Faulkes, S., Guindon, D.E., and Rudnicki, M.A. (2019). EGFR-Aurka Signaling Rescues Polarity and Regeneration Defects in Dystrophin-Deficient Muscle Stem Cells by Increasing Asymmetric Divisions. Cell Stem Cell *24*, 419-432.e6.

Wen, Y., Bi, P., Liu, W., Asakura, A., Keller, C., and Kuang, S. (2012). Constitutive Notch Activation Upregulates Pax7 and Promotes the Self-Renewal of Skeletal Muscle Satellite Cells. Molecular and Cellular Biology *32*, 2300–2311.

Wosczyna, M.N., Konishi, C.T., Perez Carbajal, E.E., Wang, T.T., Walsh, R.A., Gan, Q., Wagner, M.W., and Rando, T.A. (2019). Mesenchymal Stromal Cells Are Required for

Regeneration and Homeostatic Maintenance of Skeletal Muscle. Cell Reports *27*, 2029-2035.e5.

Yablonka-Reuveni, Z., Danoviz, M.E., Phelps, M., and Stuelsatz, P. (2015). Myogenic-specific ablation of Fgfr1 impairs FGF2-mediated proliferation of satellite cells at the myofiber niche but does not abolish the capacity for muscle regeneration. Front. Aging Neurosci. *7*.

Yao, C.C., Ziober, B.L., Sutherland, A.E., Mendrick, D.L., and Kramer, R.H. (1996). Laminins promote the locomotion of skeletal myoblasts via the alpha 7 integrin receptor. J. Cell Sci. *109*, 3139.

Yin, H., Price, F., and Rudnicki, M.A. (2013). Satellite Cells and the Muscle Stem Cell Niche. Physiological Reviews *93*, 23–67.

Zammit, P.S. (2017). Function of the myogenic regulatory factors Myf5, MyoD, Myogenin and MRF4 in skeletal muscle, satellite cells and regenerative myogenesis. Seminars in Cell & Developmental Biology *72*, 19–32.

Chapter 2

Technologies and bioinformatic methods to investigate muscle tissue transcriptomes with single-cell resolution

2.1 Introduction

Tissues are highly diverse and complex biological systems with single cells as their fundamental unit. Traditionally, cells have been characterized and classified from microscopic observations of tissue sections, later enhanced by protein staining techniques such as immunohistochemistry and flow cytometry. Although these techniques provide both valuable structural and protein expression information, they cannot directly measure gene expression or mRNA levels that are both fundamental to our understanding of biology.

Single-cell RNA-sequencing (scRNA-seq) has transformed biology and medicine as a technique to measure gene expression levels from thousands of individual cells in one controlled experiment. Contrary to antibody-to-protein assays or “gene chip” microarrays, scRNA-seq allows for an unbiased examination of cell transcriptomes and facilitates the discovery of new cell types and markers. To date, scRNA-seq technology has been used in many research contexts, starting as a way to “catalogue” the types of

cells present within a tissue (Haber et al., 2017). In stem cell biology, scRNA-seq has been used to expose the cellular heterogeneity and propose new boundaries in cell-state transitions during differentiation (Fletcher et al., 2017) as well as to reconstruct developmental trajectories of entire organisms (Farrell et al., 2018; Kester and van Oudenaarden, 2018). In addition, scRNA-seq data is now powering the development of increasingly complex bioinformatic algorithms that attempt to model intricate biological processes such as lineage differentiation dynamics (Trapnell et al., 2014; Qiu et al., 2017; La Manno et al., 2018), gene regulatory network regulation (Aibar et al., 2017), or ligand-receptor interactions (Vento-Tormo et al., 2018). Today, the field of single-cell biology is accumulating enormous amounts of data akin to that of the human genome sequencing efforts of the early 2000s. These datasets are transforming the ways we generate hypotheses and conduct research in biology as well as how we deliver more precise and personalized medicine.

2.2 Single-cell RNA-seq technologies

The first transcriptome of a single cell was sequenced about 11 years ago (Tang et al., 2009). In 2009, Tang and colleagues developed a protocol to measure mRNA from a single mouse blastomere. They measured 75% more genes compared to microarray technology and showed heterogeneity across cells from a four-cell-stage embryo. Since that study, the field has seen a substantial upscaling in scRNA-seq experimental capacity (Svensson et al., 2018). Today, scRNA-seq technology can simultaneously measurement mRNAs from hundreds of thousands of different cells or nuclei in parallel. This

experimental capacity has allowed us to identify rare types of cell inside a tissue and to appreciate the inherent gene expression heterogeneity within a cell population.

Technological advancements included improvements in single-cell isolation, mRNA capture efficiency, and sequencing library preparation chemistry. Notable examples included the addition of multiplexing using DNA barcoding and plate-based assays (SmartSeq2), followed by integrated microfluidic circuits such as the Fluidigm C1, the Drop-seq, and the commercial cousin 10X Chromium platform (Picelli et al., 2013; Macosko et al., 2015; Xin et al., 2016). Recently, split-pooling with combinatorial barcoding assay permitted to survey over 150,000 cells from multiple different organs in one single scRNA-seq run (Rosenberg et al., 2018). In this thesis, SmartSeq2, Drop-seq, and the 10X platform were used to study muscle-tissue cells and MuSCs heterogeneity (Fig. 5). The working principle of the Drop-seq and 10X is described in the Fig. 4.

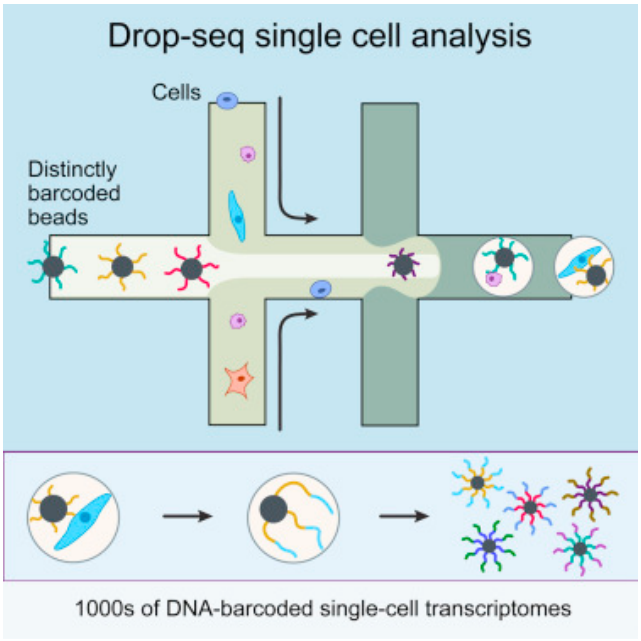


Figure 4 – Principle of the Drop-seq and 10X technologies. Drop-seq and 10X both rely on a droplet-microfluidic device as depicted above. The device is designed to encapsulate individual cells inside an aqueous droplet along with a mRNA-capture bead. The beads are covered by poly-T oligonucleotides that

are distinctively barcoded. Each molecule on the bead is also given a unique molecular identifier (or UMI) used count how many mRNA molecules are captured by a bead. Upon encapsulation, the cell is lysed, and the poly-A mRNAs hybridize to the beads. The device outputs an emulsion of droplets with over 1000s of confined single-cell transcriptomes captured by the beads. After successive rounds of reverse transcription and PCR amplification the resulting cDNA from each cell is uniquely barcoded, pooled together, and sequenced from the 3' end using a high-throughput sequencer such as Illumina. These technologies can generate thousands of single-cell transcriptomes in one run. Adapted from Macosko et al. (2015).

Each method has its own advantages and limitations; mainly in terms of sensitivity and whether the assay can capture the full-length mRNA (SmartSeq2) or only tens of bases at either the 3' or 5' end of the molecule (Drop-seq, 10X). Full length sequencing has better sensitivity for detecting low-expressed genes, offer greater depth of sequencing since more than a transcript fragment is mapped to a particular gene, and allows for splice variants identification. Full-length sequencing assay, however, cannot yet survey as many cells as Drop-seq or 10X and costs more per cell (Ziegenhain et al., 2017; Wang et al., 2019a). Balancing the number of sequenced cells with the sequencing depth for each cell is also an important factor to consider. While cells can deliver a broader overview of the cell populations within a sample, "deeper" cells increase the cell's quality as more genes are detected. For instance, though Drop-seq and 10X can profile tens of thousands of cells, the number of unique molecules (transcripts) detected per cell is less compared to SmartSeq2. An explorative survey of cell types with a tissue might initially benefit from a Drop-Seq or 10X experiment, followed by a deeper examination of specific cell populations using FACS-sorted cells and SmartSeq2. Finally, due in part to the random sampling nature of mRNA capture, scRNA-seq assays are vulnerable to "drop-out events" compared to bulk RNA-seq experiments. Drop-outs are a false negative observation of genes that are expressed by cells but not detected due to lack of sensitivity, for example due to the steric nature of the transcript or simply its low abundance.

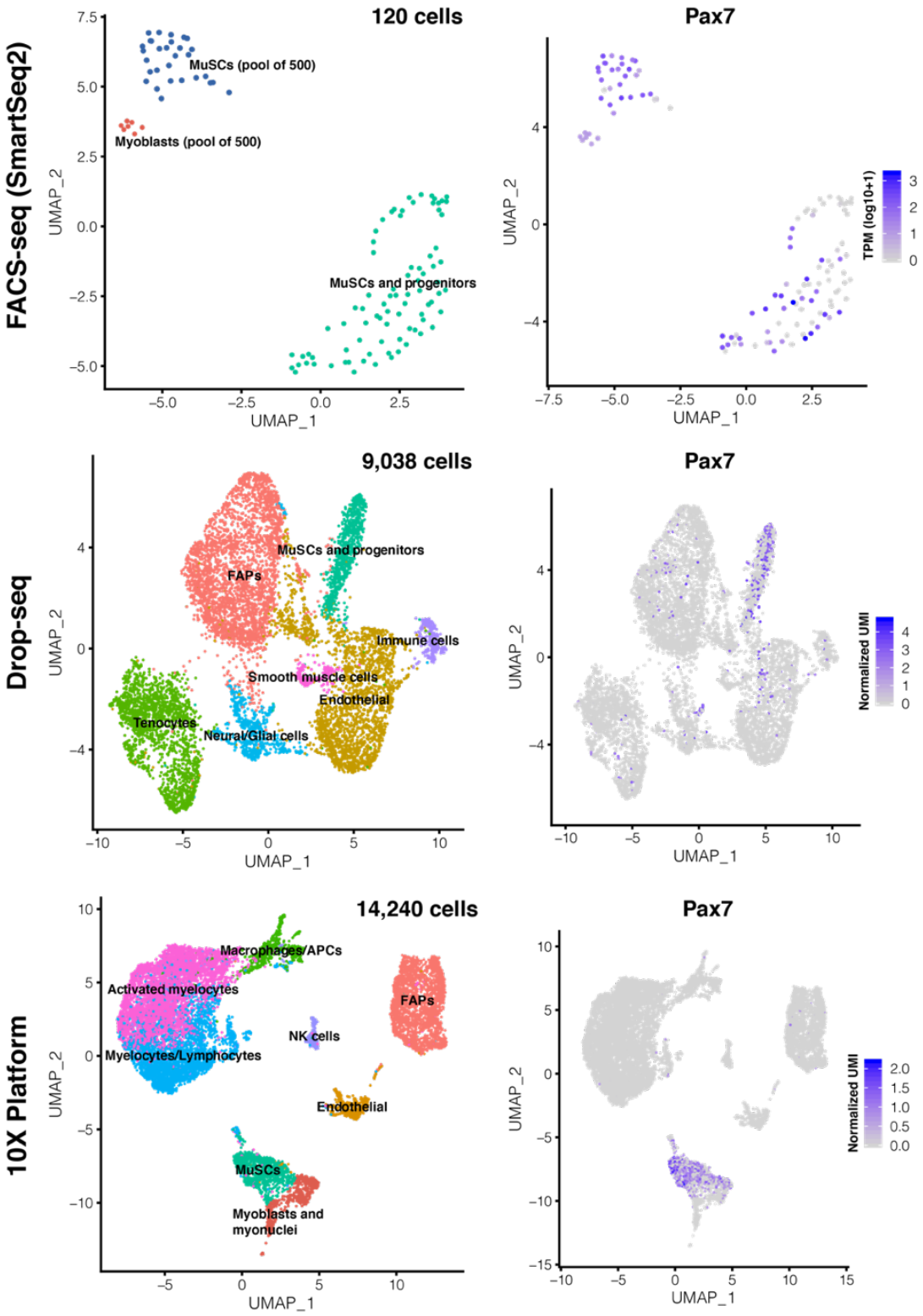


Figure 5 – Comparison between SmartSeq2, Drop-seq, and 10X experiments. Each technology can deliver different number of cells and depth (number of reads or UMIs per cell) in one experiment, as exemplified here by Pax7 expression.

2.3 Previous scRNA-seq studies in muscle biology

The first scRNA-seq study done in muscle described the molecular and functional heterogeneity of MuSCs (Cho and Doles, 2017). Using the Fluidigm C1 platform, Cho and Doles (2017) sequenced 21 single-MuSC transcriptomes that were FACS-sorted from a transgenic *Pax7-tdTomato* inducible Cre mouse. They observed highly variable expression from key muscle markers such as *Pax7* and *Cd34*. Interestingly, the most robustly expressed genes was *Sdc4* (expressed in all 21 MuSCs), followed by *Myf5* and *Vcam1*; confirming previous studies that established those genes as MuSC markers in mouse (Pisconti et al., 2012; Choo et al., 2017). Furthermore, they attempted to quantify the degree of heterogeneity by non-negative matrix factorization. They concluded that each single cell exhibited such a unique molecular identity that they were likely at different stages along the myogenic differentiation lineage. They also provided gene ontology and pathway analyses and stated that the transcriptomic heterogeneity observed reflected different functional cell states. In particular, they showed diversity in upregulated metabolic pathways, and suggested that metabolic reprograming contributes to MuSC heterogeneity. In addition, they noted that one third of the MuSC were also associated with RNA gene ontology terms and proposed that mRNA stability and translation regulation may be important checkpoints for myogenic cell-states transitions (Cho and Doles, 2017).

The study provided a first description of the MuSC molecular heterogeneity at the single-cell and RNA levels while confirmed previous observations done about metabolic reprogramming (Ryall et al., 2015). Nonetheless, the study had some important technical limitations. Most notably, they only sequenced 21 MuSCs with important dropout events

for key muscle genes. For instance, only 9 cells expressed Pax7, a surprisingly low number given MuSCs were isolated from an induced *Pax7-tdTomato* mouse. These dropout events challenged whether the described heterogeneity may not be dominated by technical artefacts. In addition, they use FACS to isolate MuSCs, which is disruptive process that likely affect the gene expression profile of the cells. Finally, they only surveyed Pax7+ cells, possibly missing other important myogenic progenitors or non-myogenic cells that may help us elucidate the cellular heterogeneity within muscle tissue.

The next two studies utilized the 10X platform and generated greater numbers of single-cell transcriptomes, thus delivering more robust statistical analyses (Giordani et al., 2019; Dell'Orso et al., 2019).

Giordani et al. (2019) used the 10X platform to classify MuSCs and the other types of cells present in the mouse muscle tissue. Collectively, Giordani sequenced about 12,000 single-cell transcriptomes and resolved 10 distinct cell types. Using mass cytometry (CyTOF), he also provided protein expression confirmation for known surface markers such as Itga7+ MuSCs, Cd31+ endothelial cells, and Pdgfra+ fibro/adipogenic progenitors (FAPs). In particular, he claimed to have identified two "understudied" muscle-tissue cell populations: *Scx*+ tenocytes and *Itga7*+ *Vcam*- "muscle-mesenchymal" cells. Using a transgenic *Scx-nGFP* mouse, Giordani showed that *Scx*+ tenocytes were located in the interstitial space between myofiber bundles and postulated that they may represent a new tendon fibroblast population that help regenerating muscle attach to the bone. Furthermore, he characterized the muscle-mesenchymal cells by isolating them by FACS from a *Pax7-nGFP* mouse and plating them in culture. Following a limiting dilution assay,

he showed that only a small (~30%) percentage of these cells were able to form myotubes, while remaining Pax7-negative, thus attributing their low clonogenicity to that of a myogenic but non-stem progenitor population. Since there are other non-stem cellular contributors to muscle regeneration, such as *Twist2*+ (Liu et al., 2017), he also performed bulk RNA-seq and showed that his muscle-mesenchymal population most resembled smooth muscle cells. Finally, he transplanted genetically labeled *Scx*+ and *Itga7*+ *Vcam*- into an injured mouse to evaluate their contribution to muscle repair. He showed that *Scx*+ were localized in the interstitium and suggested that these cells may be contributing to ECM remodeling. Likewise, by transplanting *Itga7*+ *Vcam*- cells, they showed engraftment and differentiation into new myofibers, though not as efficiently as when co-transplanted with MuSCs; suggesting that the two cell types could be synergistically interacting during muscle regeneration (Giordani et al., 2019).

His findings illustrated how high-throughput single-cell analyses can reveal poorly described cell subpopulations leading to new hypotheses on their myogenic function. Nonetheless, the methodology also raised the question about how “new” cell types and boundaries are defined – should they be at all – as cells exist in dynamic and transient continuum of states far less “digital” that what bioinformatic algorithms suggest.

Dell’Orso et al. (2019) provided a more refined single-cell description of MuSC activation and differentiation from FACS-sorted cells and a pseudotime trajectory model. They used FACS to both enrich for live cells and to specifically sort out MuSCs for scRNA-seq. From the 9 resolved muscle-tissue cell types, they noted that MuSCs formed 2 subpopulations: “close to quiescence” MuSCs enriched for genes such as *Pax7*, *Hes1*,

and the growth arrest-specific protein *Gas*; and “early activation” MuSCs enriched for *Myf5* and *Myod*, as well as many ribosomal genes, suggesting that these cells are in a mitotically activated state. Gene ontology analysis also showed that the “close to quiescence” population was enriched for cell cycle arrest. To place these two MuSC subpopulations in the context of myogenic differentiation, they included transcriptomes from myoblasts and generated a pseudotime trajectory model (Qiu et al., 2017). They showed a branched progression from “close to quiescence” and “early activated” MuSCs to either a population of cycling myoblasts or committed myocytes. The myoblasts expressed cell cycle regulators such as *Ccnd1* and eIFs (translation initiation), while the committed myocytes were enriched for *Myog* and cycling-depend kinase inhibitors such as *Cdkn1a*. To further extend their trajectory model, they added single-MuSCs from 60h post-injury muscles and described a linear progression from quiescent MuSCs to primary myoblasts with the “injured MuSCs” in between. Translation regulator were enriched in injured MuSCs as well as metabolic genes involved energy-producing pathways such as glycolysis. Lastly, they also studied the dynamics of metabolic pathway connectivity and revealed that metabolic reprogramming plays a central role in regulating MuSC state transition as first described using single-cell data by Cho and Doles (2017) (Dell’Orso et al., 2019).

Their study provided a more continuous picture of myogenic cell states, though the experimental design has some important biases and limitations. For instance, the pseudotime model was constructed by combining single-cell datasets FACS-sorted MuSCs, injured MuSCs, and cultured myoblasts. The cells were heavily manipulated and

originated from different experiments settings and possible batch technical effects were not addressed bioinformatically.

In this thesis, we posit that an injury time-course experiment, with no selection and with minimal cell manipulation, may more accurately and unbiasedly recapitulate the myogenic lineage and allow us to identify new molecular markers. In Chapter 3 we present a pseudotime model of myogenic differentiation from muscle progenitors sourced at 3 different timepoints post-injury (with multiple replicates per timepoint and no FACS selection).

2.4 Muscle sample preparation and biases in scRNA-seq experiments

Single-cell RNA-sequencing experiments involve many steps requiring the user to make decisions on the sample preparation procedure, starting with how the tissue is extracted from the animal. For muscle, multiple protocols have been developed and optimize to digest mouse and human muscle tissues into a single-cell suspension (Spinazzola and Gussoni, 2017). In this thesis we used Collagenase D and Dispase II for mouse and human muscle samples; two enzymes that breakdown collagen and other ECM proteins that hold myofibers together. The digested muscle is then mechanically dissociated with tweezers, pipettes, and syringes, and filtered with cell-strainers to remove connective tissue and myofiber fragments. Red blood cells were also lysed. The final suspension was thoroughly washed in low-serum PBS before scRNA-seq. A possible consideration is whether to use FACS to enrich for live cells and further remove debris. Though this strategy is commonly used in a number of published scRNA-seq studies (Cho and Doles,

2017; Dell'Orso et al., 2019), we noticed that FACS biases the proportions of cell types resolved, likely because not all types of cell metabolize the viability dye with the same efficiency (Fig. 8).

The steps presented above affect the cells and their transcriptome in two important ways. First, tissue dissociation can provoke cell membrane damage, which exposes mRNA to nucleases. This affects the quantity and the quality of the mRNA recovered for scRNA-seq. Second, cell manipulation can lead to gene expression biases in the transcriptomic readout; for example, the expression of stress-response genes or down-regulation of quiescence markers (van den Brink et al., 2017). This is particularly true for MuSCs that immediately downregulate *Pax7* as soon as removed from their *in vivo* niche (Zammit et al., 2006). In addition, it has recently been shown that collagenase digestion at 37^{o}C provokes heat shock and stress-response genes upregulation in solid tumor tissues (O'Flanagan et al., 2019). All of these changes will confound the cell's true *in vivo* gene expressions signature, which can lead to misinterpretations and/or obscuration of important biological information.

To address these issues, a number of data correction and sample preparation strategies have been proposed. First, the percentage of reads in a cell that is are mapped to mitochondrial genes can be used to remove structurally compromised and dead cells. High mitochondrial-to-cytoplasmic mRNA content ratio is usually a good indication of cell damage as mitochondrial RNA will remain in the cell even if the cell membrane is damaged (Ilicic et al., 2016). Nonetheless, high mitochondrial gene expression can be an important biological feature of some cell types such as in cancer (Perdas et al., 2019).

Gene expression biases, however, are difficult to correct and are better mitigated by carefully optimized sample preparation. Working fast, on ice, and fixing (Alles et al., 2017) the cells usually minimizes transcriptomic biases, though not all types of cells are affected the same way. For example, we noticed that MuSCs are particularly sensitive to tissue dissociation and significantly express stress-response genes in differential expression analysis (Fig. 17A). Another proposed strategy is *in situ* fixation protocols aimed at "freezing" the transcriptomes of cells prior to tissue dissection and dissociation. For instance, Machado et al. (2017) showed by bulk RNA-seq that fixing muscle tissue *in situ* revealed new signatures of quiescence and early activation in MuSC (Machado et al., 2017). Finally, another data correction approach for transcriptomic biases could be to include a cell-stress parameter as a covariate to the normalization models we present in the next section. This would involve deriving the weights of a gene expression stress vector for each type of cells in the tissue.

2.5 Data heterogeneity, normalization and integration methods

Single-cell RNA-sequencing data is highly variable as it captures the heterogenous landscape of cell types and states. These heterogeneities may originate from differences in cell state or identity, stochastic bursts in gene expression, but also from technical effects such as mRNA capture efficiency, sequencing depth, dropouts, noise, batch, etc. It is therefore vital to correct scRNA-seq data through normalization and integration methods in order to remove undesired sources of variability.

Compared to bulk RNA-seq, scRNA-seq data is considerably nosier as less genes are detected per cell on average and, because of dropout events, generates very sparse

gene expression matrices. The most straightforward manner to normalize scRNA-seq data is to divide each gene expression value by the total number of transcripts (or UMIs) captured per cell. This “size factor” normalization assumes that all cells have the same total number of mRNA molecule and that UMI differences between cells are only due to technical effects. However, this assumption does not always reflect the tissue’s content as some cells may be larger in size while other in a dormant state and thus making less mRNA. For example, MuSCs that are both small in size and generally quiescent.

Hafemeister and Satija (2019) recently demonstrated that size factor-based normalization affects the true expression value of low-expressed genes. In particular, cells with a low number of UMIs tend to overrepresent high-abundance genes, which dampers the variance contribution from other genes. To address this matter, they proposed a new normalization method using a regularized negative binomial model (Hafemeister and Satija, 2019). This approach normalizes gene expression values with a negative binomial distribution where the model parameters are further constrained by “regularization”. Regularization limits overfitting by sharing information across genes expressed at similar levels. Similar methods to improve parameter estimation and model robustness have been applied to differential expression testing in bulk RNA-seq such as DESeq2 (Love et al., 2014). This specific normalization method is used in Chapter 4.

During data normalization, one can also choose to regress out biological and technical covariates. In this thesis we chose to regress out the percentage of UMIs per cell mapped to a mitochondrial gene as proxy for cell stress and death. Other covariates that can be regressed out include the cell-cycle stage. However, we found that correcting

for cell-cycle confounds the different myogenic cell populations (MuSC, myoblasts, myocytes) which have strong dependencies on cell-cycle genes (Zammit, 2017). Among other technical covariates, we choose to consider the number of UMIs and gene detected per cell; which, respectively, can account for differences in cell size and quiescence.

Single-cell RNA-seq experiments often require comparing or combining independently generated datasets that may have been generated on different instruments, at different times, in different labs, or sourced from different animals (some of which subjected to a particular treatment). Each independent experiment can affect the measurement of the transcriptome or the cell's transcriptome itself. To address this issue, batch effect correction algorithms and data integration methods have been proposed. While batch effect correction is typically performed between biological and/or technical replicates sharing similar gene expression matrices, data integration can combine datasets with important differences in cell composition, transcriptomic profile, and even data structure. For instance, some integration method can also combine different single-cell data modalities such as RNA- and ATAC-seq (a measure of chromatin accessibility) along with spatial information using non-linear algorithms (Stuart et al., 2019).

In this thesis we use two different data integration methods. In Chapter 3 we integrate biological replicates and conditions (days post muscle injury) using canonical correlation analysis (CCA) (Butler et al., 2018) given samples are fairly similar (same mouse strain and muscle). This approach first calculates a correlation subspace from

genes that are highly correlated and thus shared between datasets. Then, each CCA component is corrected for global differences in scale by dynamic time warping resulting in a unified dataset that projects what is most shared between datasets and attenuates dataset-specific sources of variation. This method, however, tends to overcorrect biological variability and mask cell populations that are unique to one dataset. In Chapter 4, we instead used Scanorama (Hie et al., 2019) to integrate human muscle datasets given the important cellular diversity between sample. Scanorama is based on computer vision and can robustly integrate heterogenous single-cell data even if one dataset presents a unique cell population. We provide a comparison of two data integration methods in Fig. S8 for highly heterogenous human muscle samples. Batch effect and integration methods are continuously been perfected (Stuart et al., 2019) and proposed (Korsunsky et al., 2019) to this date.

2.6 Modelling myogenesis with single-cell resolution

2.6.1 Trajectory inference analysis (Monocle)

In response to muscle injury, myogenic progenitors transition between states and express different sets of genes, thereby producing a dynamic repertoire of surface proteins and transcription factors. However, these markers may not always selective enough to isolate desired transient subpopulations, thus making it difficult to study specific timepoints of myogenic differentiation (cf. Chapter 1.3). Moreover, in a scRNA-seq experiment, the transcriptome of MuSC and progenitor cells is measured at their respective stages within the differentiation lineage, without any cell ordering or dynamic description of gene expression (Fig 6). Trajectory inference analysis and modeling from

scRNA-seq data may therefore help us organize these cells and reveal the dynamics of surface marker expression along the differentiation continuum.

Single-cell trajectory models, such as Monocle (Trapnell et al., 2014; Qiu et al., 2017), are therefore useful tools to organize these “asynchronous” cells along a trajectory in *pseudotime*, an abstract measure of how much an individual cell will have progressed through a biological process. While pseudotime directly relates to transcriptional differences between cells, in this thesis we use pseudotime as a proxy for differentiation time and to model myogenesis. Monocle is an unsupervised machine learning problem driven by genes that are differentially expressed between myogenic subpopulations. Monocle applies a reverse graph embedding strategy to reduce the dimensionality of the data, from which it learns a trajectory (manifold learning) and orders the cells along that trajectory. Here, we use trajectory inference analysis to model myogenesis, identify new surface markers that delineate important stages of myogenic differentiation, and delineate other differentiation stages or branches as well as new myogenic subpopulations that may be relevantly describe the physiology of muscle tissue repair.

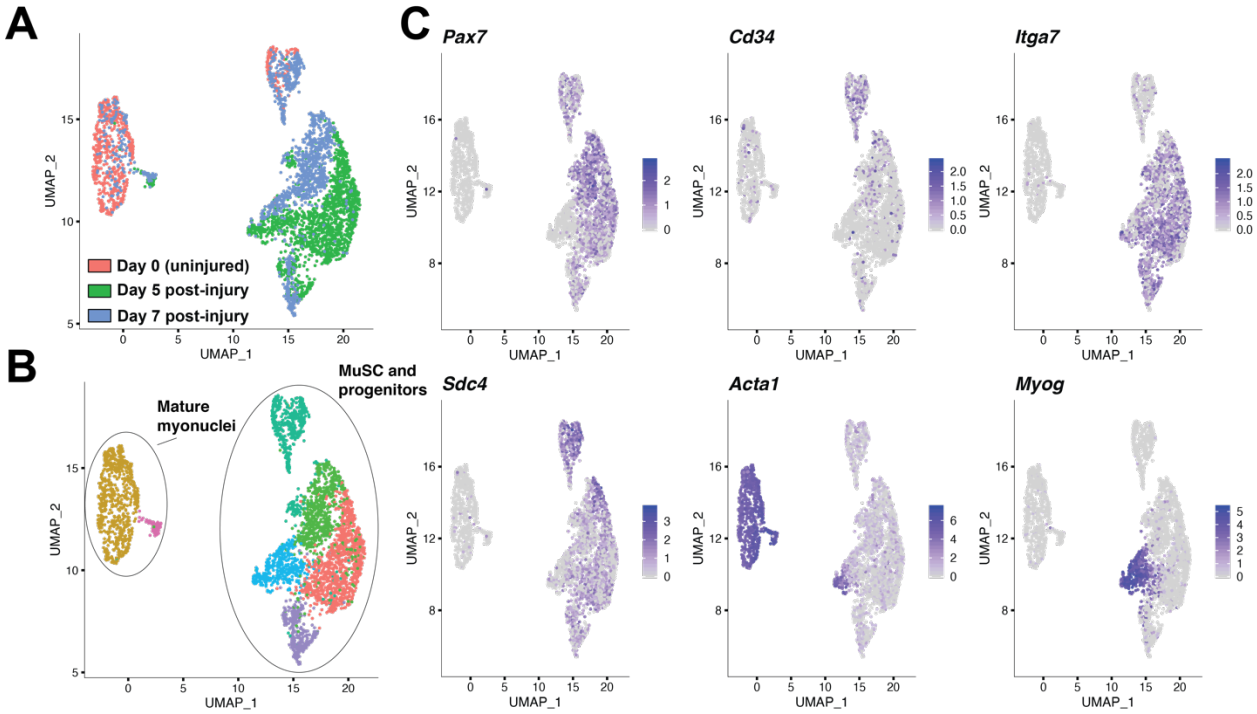


Figure 6 – Muscle progenitor cell population "asynchronous" heterogeneity. (A) Collection of single-cell transcriptomes of MuSC and progenitors from muscles at different days post-injury. These cells exist at different “times” along a myogenic differentiation lineage. **(B)** Collectively, these single-cell transcriptomes describe 7 different subpopulations of MuSCs, progenitors, and mature myonuclei, that are labeled by unsupervised clustering. Monocle will use the genes that are differentially expressed between these clustered subpopulations to organize cells in pseudotime. **(C)** Example of known muscle markers that are heterogeneously expressed (normalized expression) along the myogenic differentiation lineage.

2.6.2 Ligand-receptor interaction models

Cell-cell interactions through ligand-receptor (LR) complexes regulate muscle tissue repair. In particular, interactions between MuSCs and niche constituents direct stem cell fate decisions post injury (cf. Chapter 1.4). Ligand-receptor models from scRNA-seq data provides a unique opportunity to model cell-communication channels between MuSC surface receptors and their respective ligands to generate hypotheses on the downstream signaling pathways that regulate myogenesis. In the muscle tissue, we have seen that these ligands can be expressed on the surface of another cell or secreted (basal lamina ECM protein, cytokine for example, cf. Chapter 1.5).

The objective of a LR model is to associate ligands and receptors together from their respective gene expression values and then make predictions on the cell-to-cell specificity of the interaction by applying constraints to the model. A curated database of LR pairs can facilitate the description of known interactions (Ramilowski et al., 2015; Skelly et al., 2018), but models can also be designed to propose new interaction pairs (Vento-Tormo et al., 2018). Below we describe some of the different steps and design consideration of LR models from scRNA-seq data with some published examples given these methods are fairly novel in the field.

The first step is to decide on a rule to link two interacting cell types. For example, this can be done for example by selecting ligands and receptors genes expressed above a chosen threshold value. After the link is established, it is important to quantitatively score the “strength” and “exclusivity” of the interaction. Interaction strength can be calculated from expression values or a correlation score. To account for LR specificity, the model can also weight in factors such as known interaction kinetics since not all highly expressed receptors have a strong affinity for its ligands. Further, since different cell types can express the same ligand or receptor, it is also critical to assess whether the communication is exclusive or shared between cell types. This can be achieved using knowledge on the tissue anatomy or by differentially gene expression testing. It is also important to consider secreted ligands that may originate from paracrine signaling outside of the tissue studied by scRNA-seq. Finally, a score of significance for the predicted interaction should be provided using statistical testing to avoid calling interactions that are

due to chance. This step typically involves comparing the predicted interaction score to a null distribution, which can be generated by statistical techniques such as bootstrapping or permutation testing.

A few LR models using scRNA-seq data have been proposed. For example, Zhou et al. (2017) examined intercellular communication between different type of melanoma cancer cells by simply ruling that two populations are communicating when the ligand and receptor average expression value was above 3 standard deviation from the mean. They also ranked interactions by an “exclusiveness score of interaction” by dividing the LR expression value of one cell type by the mean across all other cell types (Zhou et al., 2017). Following a similar strategy, Skelly et al. (2018) mapped intracellular communication in the mouse heart. They linked two cell type together if 20% of the cells within a cluster had a receptor or ligand expression value greater than zero (Skelly et al., 2018). Their model did not however evaluate interaction exclusivity nor accounted for the heart compartmentalized anatomy. Cohen et al. (2018) investigated lung basophils interactions using a graph-based approach. They calculated the Spearman correlation between LR expression values and used hierarchical clustering to infer interaction modules (Cohen et al., 2018). Kummar et al. (2018) instead calculated and interaction score for each ligand, receptor, and cell type by multiplying the average receptor expression of one cell type with the average ligand expression of the other cell type. To identify significant interaction across experimental conditions, they used a one-sided Wilcoxon rank-sum test (Kumar et al., 2018). This method however does not test

significance between interactions within the same experimental group. To address this question of significance, Roser Vento-Torno et al. (2018) used empirical shuffling to evaluate which pairs were cell-type specific. This consisted in first calculating a LR score by averaging the ligand and receptor expression values for genes that are at least expressed in 10% of cell in a group. Then, by randomly shuffling 1000 times the cell-type labels, they generated a null distribution of interaction score for a given pair to finally ask whether the true interaction score is in the top 5% of the distribution (the p-value) (Vento-Tormo et al., 2018). As of December 2019, their LR analysis framework is certainly the most sophisticated, and now also accounts for heteromeric complexes (Efremova et al., 2019). Their public repository CellPhoneDB 2.0 is published online but is currently only available for human genes.

To investigate the LR and cell-cell interactions dynamics during muscle regeneration, we developed a new bioinformatic model. The code and documentation are available online on GitHub*. The model is designed to investigate the interactions between receptors that are differentially expressed in MuSCs and ligands expressed by other muscle cell types but easily be adapted to other scenarios. In summary the model:

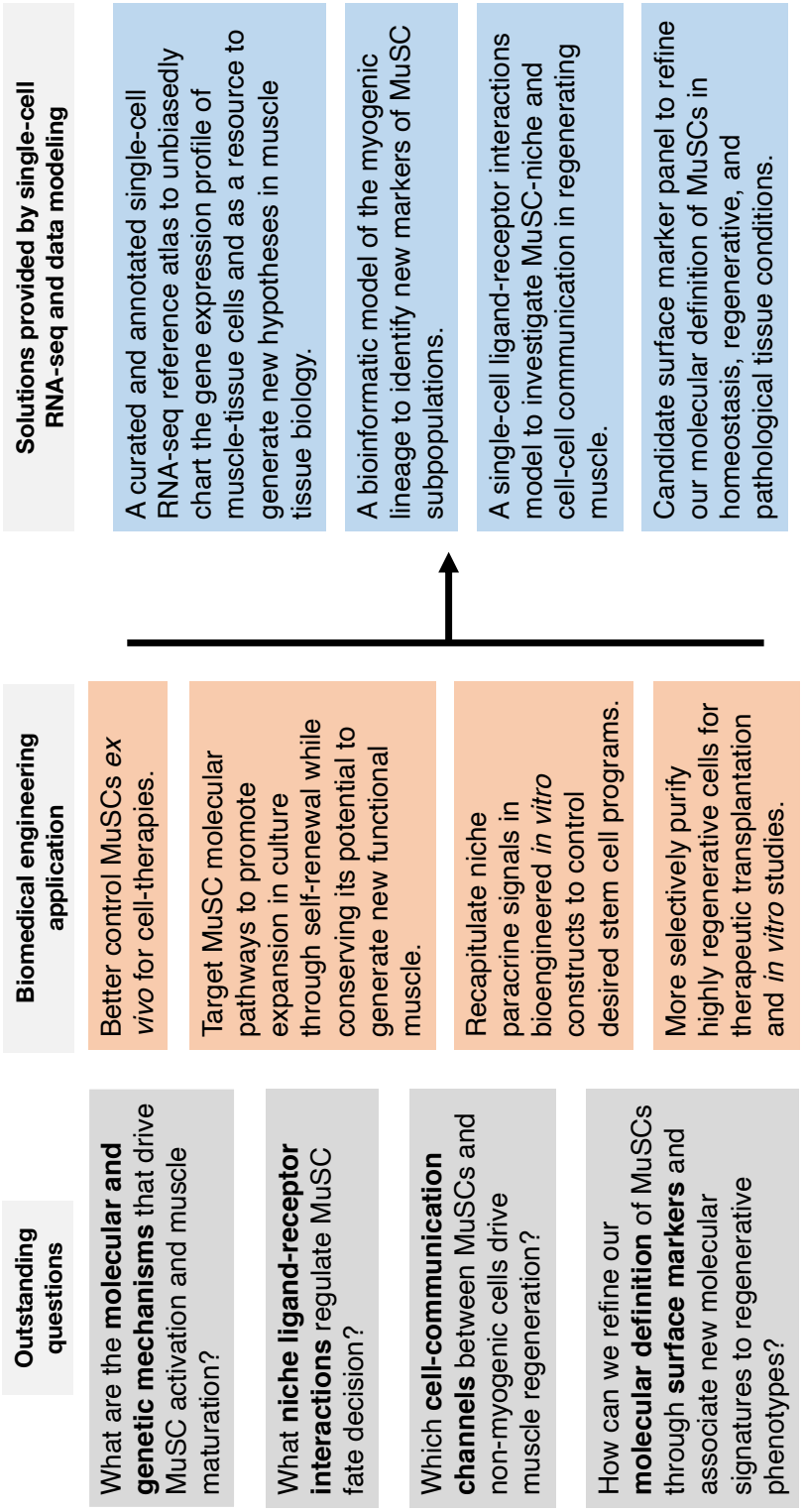
- First calculates a list of candidate receptors that are differentially expressed in the target cell population (MuSCs for instance) compared of the other cell types. To be differentially expressed, the receptor must be expressed in at least 25% of cells.

*https://github.com/andrea-de-micheli/L-R_interaction_model_for_scRNAseq

- For each population, calculates the average normalized gene expression value for all ligands and receptors.
- Removes ligands whose average expression value across all cell populations are in the bottom 5%.
- Computes a LR score matrix. For each candidate receptor, the algorithm multiplies the receptor non-zero expression value with the average normalized expression value of all its ligands. Each row of the score matrix will be a LR pair and each column a specific population (including the target population to account for autocrine interaction).
- Finally, scale the LR matrix by calculating z-scores across cell types (Fig. S6).

In this thesis I use a database of curated R-L pairs published by the Ramilowski and Pinto labs for human and mouse genes respectively (Ramilowski et al., 2015; Skelly et al., 2018).

2.7 Outstanding questions in MuSC biomedical engineering addressed by scRNA-seq



Outstanding questions	Biomedical engineering application	Solutions provided by single-cell RNA-seq and data modeling
What are the **molecular and genetic mechanisms** that drive MuSC activation and muscle maturation?	Better control MuSCs *ex vivo* for cell-therapies.	A curated and annotated single-cell RNA-seq reference atlas to unbiasedly chart the gene expression profile of muscle-tissue cells and as a resource to generate new hypotheses in muscle tissue biology.
What **niche ligand-receptor interactions** regulate MuSC fate decision?	Target MuSC molecular pathways to promote expansion in culture through self-renewal while conserving its potential to generate new functional muscle.	A bioinformatic model of the myogenic lineage to identify new markers of MuSC subpopulations.
Which **cell-communication channels** between MuSCs and non-myogenic cells drive muscle regeneration?	Recapitulate niche paracrine signals in bioengineered *in vitro* constructs to control desired stem cell programs.	A single-cell ligand-receptor interactions model to investigate MuSC-niche and cell-cell communication in regenerating muscle.
How can we refine our **molecular definition** of MuSCs through **surface markers** and associate new molecular signatures to regenerative phenotypes?	More selectively purify highly regenerative cells for therapeutic transplantation and *in vitro* studies.	Candidate surface marker panel to refine our molecular definition of MuSCs in homeostasis, regenerative, and pathological tissue conditions.

Table 1 – Outstanding question in MuSC biomedical engineering addressed in this thesis by scRNA-seq.

2.8 Chapter references

Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat Methods *14*, 1083–1086.

Alles, J., Karaiskos, N., Praktiknjo, S.D., Grosswendt, S., Wahle, P., Ruffault, P.-L., Ayoub, S., Schreyer, L., Boltengagen, A., Birchmeier, C., et al. (2017). Cell fixation and preservation for droplet-based single-cell transcriptomics. BMC Biol *15*, 44.

van den Brink, S.C., Sage, F., Vértesy, Á., Spanjaard, B., Peterson-Maduro, J., Baron, C.S., Robin, C., and van Oudenaarden, A. (2017). Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. Nature Methods *14*, 935–936.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nature Biotechnology *36*, 411–420.

Cho, D.S., and Doles, J.D. (2017). Single cell transcriptome analysis of muscle satellite cells reveals widespread transcriptional heterogeneity. Gene *636*, 54–63.

Choo, H.-J., Canner, J.P., Vest, K.E., Thompson, Z., and Pavlath, G.K. (2017). A tale of two niches: differential functions for VCAM-1 in satellite cells under basal and injured conditions. American Journal of Physiology-Cell Physiology *313*, C392–C404.

Cohen, M., Giladi, A., Gorki, A.-D., Solodkin, D.G., Zada, M., Hladik, A., Miklosi, A., Salame, T.-M., Halpern, K.B., David, E., et al. (2018). Lung Single-Cell Signaling Interaction Map Reveals Basophil Role in Macrophage Imprinting. Cell *175*, 1031-1044.e18.

Dell'Orso, S., Juan, A.H., Ko, K.-D., Naz, F., Perovanovic, J., Gutierrez-Cruz, G., Feng, X., and Sartorelli, V. (2019). Single cell analysis of adult mouse skeletal muscle stem cells in homeostatic and regenerative conditions. Development *146*, dev174177.

Efremova, M., Vento-Tormo, M., Teichmann, S.A., and Vento-Tormo, R. (2019). CellPhoneDB v2.0: Inferring cell-cell communication from combined expression of multi-subunit receptor-ligand complexes. BioRxiv 680926.

Farrell, J.A., Wang, Y., Riesenfeld, S.J., Shekhar, K., Regev, A., and Schier, A.F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. Science *360*, eaar3131.

Fletcher, R.B., Das, D., Gadye, L., Street, K.N., Baudhuin, A., Wagner, A., Cole, M.B., Flores, Q., Choi, Y.G., Yosef, N., et al. (2017). Deconstructing Olfactory Stem Cell Trajectories at Single-Cell Resolution. Cell Stem Cell 20, 817-830.e8.

Giordani, L., He, G.J., Negroni, E., Sakai, H., Law, J.Y.C., Siu, M.M., Wan, R., Corneau, A., Tajbakhsh, S., Cheung, T.H., et al. (2019). High-Dimensional Single-Cell Cartography Reveals Novel Skeletal Muscle-Resident Cell Populations. Molecular Cell *74*, 609-621.e6.

Haber, A.L., Biton, M., Rogel, N., Herbst, R.H., Shekhar, K., Smillie, C., Burgin, G., Delorey, T.M., Howitt, M.R., Katz, Y., et al. (2017). A single-cell survey of the small intestinal epithelium. Nature *551*, 333–339.

Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol *20*, 296.

Hie, B., Bryson, B., and Berger, B. (2019). Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. Nature Biotechnology *37*, 685–691.

Ilicic, T., Kim, J.K., Kolodziejczyk, A.A., Bagger, F.O., McCarthy, D.J., Marioni, J.C., and Teichmann, S.A. (2016). Classification of low quality cells from single-cell RNA-seq data. Genome Biology *17*, 29.

Kester, L., and van Oudenaarden, A. (2018). Single-Cell Transcriptomics Meets Lineage Tracing. Cell Stem Cell *23*, 166–179.

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nature Methods *16*, 1289–1296.

Kumar, M.P., Du, J., Lagoudas, G., Jiao, Y., Sawyer, A., Drummond, D.C., Lauffenburger, D.A., and Raue, A. (2018). Analysis of Single-Cell RNA-Seq Identifies Cell-Cell Communication Associated with Tumor Characteristics. Cell Reports *25*, 1458-1468.e4.

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494–498.

Liu, N., Garry, G.A., Li, S., Bezprozvannaya, S., Sanchez-Ortiz, E., Chen, B., Shelton, J.M., Jaichander, P., Bassel-Duby, R., and Olson, E.N. (2017). A Twist2-dependent progenitor cell contributes to adult skeletal muscle. Nature Cell Biology *19*, 202–213.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

Machado, L., Esteves de Lima, J., Fabre, O., Proux, C., Legendre, R., Szegedi, A., Varet, H., Ingerslev, L.R., Barrès, R., Relaix, F., et al. (2017). In Situ Fixation Redefines

Quiescence and Early Activation of Skeletal Muscle Stem Cells. Cell Reports *21*, 1982–1993.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202–1214.

O'Flanagan, C.H., Campbell, K.R., Zhang, A.W., Kabeer, F., Lim, J.L.P., Biele, J., Eirew, P., Lai, D., McPherson, A., Kong, E., et al. (2019). Dissociation of solid tumor tissues with cold active protease for single-cell RNA-seq minimizes conserved collagenase-associated stress responses. Genome Biology *20*, 210.

Perdas, E., Stawski, R., Kaczka, K., Nowak, D., and Zubrzycka, M. (2019). Altered levels of circulating nuclear and mitochondrial DNA in patients with Papillary Thyroid Cancer. Sci Rep *9*, 14438.

Picelli, S., Bjorklund, A.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Meth *10*, 1096–1098.

Pisconti, A., Bernet, J.D., and Olwin, B.B. (2012). Syndecans in skeletal muscle development, regeneration and homeostasis. Muscles Ligaments Tendons J *2*, 1–9.

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nature Methods *14*, 979–982.

Ramilowski, J.A., Goldberg, T., Harshbarger, J., Kloppmann, E., Lizio, M., Satagopam, V.P., Itoh, M., Kawaji, H., Carninci, P., Rost, B., et al. (2015). A draft network of ligand–receptor-mediated multicellular signalling in human. Nature Communications *6*, 7866.

Rosenberg, A.B., Roco, C.M., Muscat, R.A., Kuchina, A., Sample, P., Yao, Z., Graybuck, L.T., Peeler, D.J., Mukherjee, S., Chen, W., et al. (2018). Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. Science *360*, 176.

Ryall, J.G., Dell'Orso, S., Derfoul, A., Juan, A., Zare, H., Feng, X., Clermont, D., Koulnis, M., Gutierrez-Cruz, G., Fulco, M., et al. (2015). The NAD+-Dependent SIRT1 Deacetylase Translates a Metabolic Switch into Regulatory Epigenetics in Skeletal Muscle Stem Cells. Cell Stem Cell *16*, 171–183.

Skelly, D.A., Squiers, G.T., McLellan, M.A., Bolisetty, M.T., Robson, P., Rosenthal, N.A., and Pinto, A.R. (2018). Single-Cell Transcriptional Profiling Reveals Cellular Diversity and Intercommunication in the Mouse Heart. Cell Reports *22*, 600–610.

Spinazzola, J.M., and Gussoni, E. (2017). Isolation of Primary Human Skeletal Muscle Cells. Bio Protoc *7*, e2591.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., III, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888-1902.e21.

Svensson, V., Vento-Tormo, R., and Teichmann, S.A. (2018). Exponential scaling of single-cell RNA-seq in the past decade. Nature Protocols *13*, 599–604.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nature Methods *6*, 377–382.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nature Biotechnology *32*, 381–386.

Vento-Tormo, R., Efremova, M., Botting, R.A., Turco, M.Y., Vento-Tormo, M., Meyer, K.B., Park, J.-E., Stephenson, E., Polański, K., Goncalves, A., et al. (2018). Single-cell reconstruction of the early maternal–fetal interface in humans. Nature *563*, 347–353.

Wang, Y.J., Schug, J., Lin, J., Wang, Z., Kossenkov, A., the HPAP Consortium, and Kaestner, K.H. (2019). Comparative analysis of commercially available single-cell RNA sequencing platforms for their performance in complex human tissues (Genomics).

Xin, Y., Kim, J., Ni, M., Wei, Y., Okamoto, H., Lee, J., Adler, C., Cavino, K., Murphy, A.J., Yancopoulos, G.D., et al. (2016). Use of the Fluidigm C1 platform for RNA sequencing of single mouse pancreatic islet cells. Proc Natl Acad Sci USA *113*, 3293.

Zammit, P.S. (2017). Function of the myogenic regulatory factors Myf5, MyoD, Myogenin and MRF4 in skeletal muscle, satellite cells and regenerative myogenesis. Seminars in Cell & Developmental Biology *72*, 19–32.

Zammit, P.S., Relaix, F., Nagata, Y., Ruiz, A.P., Collins, C.A., Partridge, T.A., and Beauchamp, J.R. (2006). Pax7 and myogenic progression in skeletal muscle satellite cells. J. Cell Sci. *119*, 1824.

Zhou, J.X., Taramelli, R., Pedrini, E., Knijnenburg, T., and Huang, S. (2017). Extracting Intercellular Signaling Network of Cancer Tissues using Ligand-Receptor Expression Patterns from Whole-tumor and Single-cell Transcriptomes. Scientific Reports *7*, 8815.

Ziegenhain, C., Vieth, B., Parekh, S., Reinius, B., Guillaumet-Adkins, A., Smets, M., Leonhardt, H., Heyn, H., Hellmann, I., and Enard, W. (2017). Comparative Analysis of Single-Cell RNA Sequencing Methods. Molecular Cell *65*, 631-643.e4.

Chapter 3

Single-cell analysis of the muscle stem cell hierarchy identifies heterotypic communication signals involved in mouse skeletal muscle regeneration[†]

Abstract

Functionally heterogeneous subpopulations of MuSCs have been identified based on their expression of myogenic regulatory factors and surface markers. However, a unified organization of muscle stem and progenitor cells and their subpopulations remains unresolved. Here, we performed temporal analysis of skeletal muscle regeneration using single-cell RNA-sequencing (scRNA-seq) of notexin-injured adult mouse muscles. We generated over 34,000 single-cell transcriptomes spanning four muscle regeneration time-points and identified 15 distinct cell types, including a heterogeneous population of MuSCs and progenitor cells. Our analysis provides a hierarchical map of myogenic cell populations and identifies stage-specific regulatory programs that govern their contributions to muscle regeneration. In this transcriptomic atlas, we observed cell type-specific regenerative dynamics, exemplified by waves of transient amplification and

[†]This chapter is adapted from the following publication:
Single-Cell Analysis of the Muscle Stem Cell Hierarchy Identifies Heterotypic Communication Signals Involved in Skeletal Muscle Regeneration. (2020).
Andrea J. De Micheli, Emily J. Laurilliard, Charles L. Heinke, Hiranmayi Ravichandran, Paula Fraczek, Sharon Soueid-Baumgarten, Iwijn De Vlaminck, Olivier Elemento, and Benjamin D. Cosgrove
Cell Reports 30, 3583–3595. March 10, 2020. https://doi.org/10.1016/j.celrep.2020.02.067.

diversification of multiple immune cell types and, concomitantly, myogenic cells. Unbiased trajectory inference organized the myogenic cell populations within the atlas into a continuum, consisting of a hierarchy of quiescent MuSCs, cycling progenitors, committed myoblasts, and terminally differentiated myocytes. This myogenic trajectory matched prior understanding and also revealed that MuSC stages are defined by synchronous changes in regulatory factors, cell cycle-associated, and surface receptor gene expression. Lastly, we analyzed the transcriptomic atlas to identify over 100 candidate heterotypic communication signals between myogenic and non-myogenic cell populations, including many involving the fibroblast growth factor (FGF), Notch, and Syndecan receptor families and their associated ligands. Syndecan receptors were implicated in a large fraction of these cell communication interactions and were observed to exhibit transcriptional heterogeneity within the myogenic continuum. Using multiplexed mass cytometry (CyTOF), we confirmed that cycling, but not quiescent or committed, MuSCs exhibit Syndecan protein expression diversification, suggesting that Syndecan signaling interactions may coordinate stage-specific myogenic cell fate regulation. We performed ligand stimulation experiments *in vitro* and confirmed that three of the model-identified ligands (FGF2, TGFβ1, RSPO3) regulate muscle stem/progenitor cell proliferation in a Syndecan-dependent manner. This scRNA-seq reference atlas provides a hierarchical organization of myogenic subpopulations as a resource to investigate cell-cell interactions that regulate myogenic stem and progenitor cell fates in muscle regeneration.

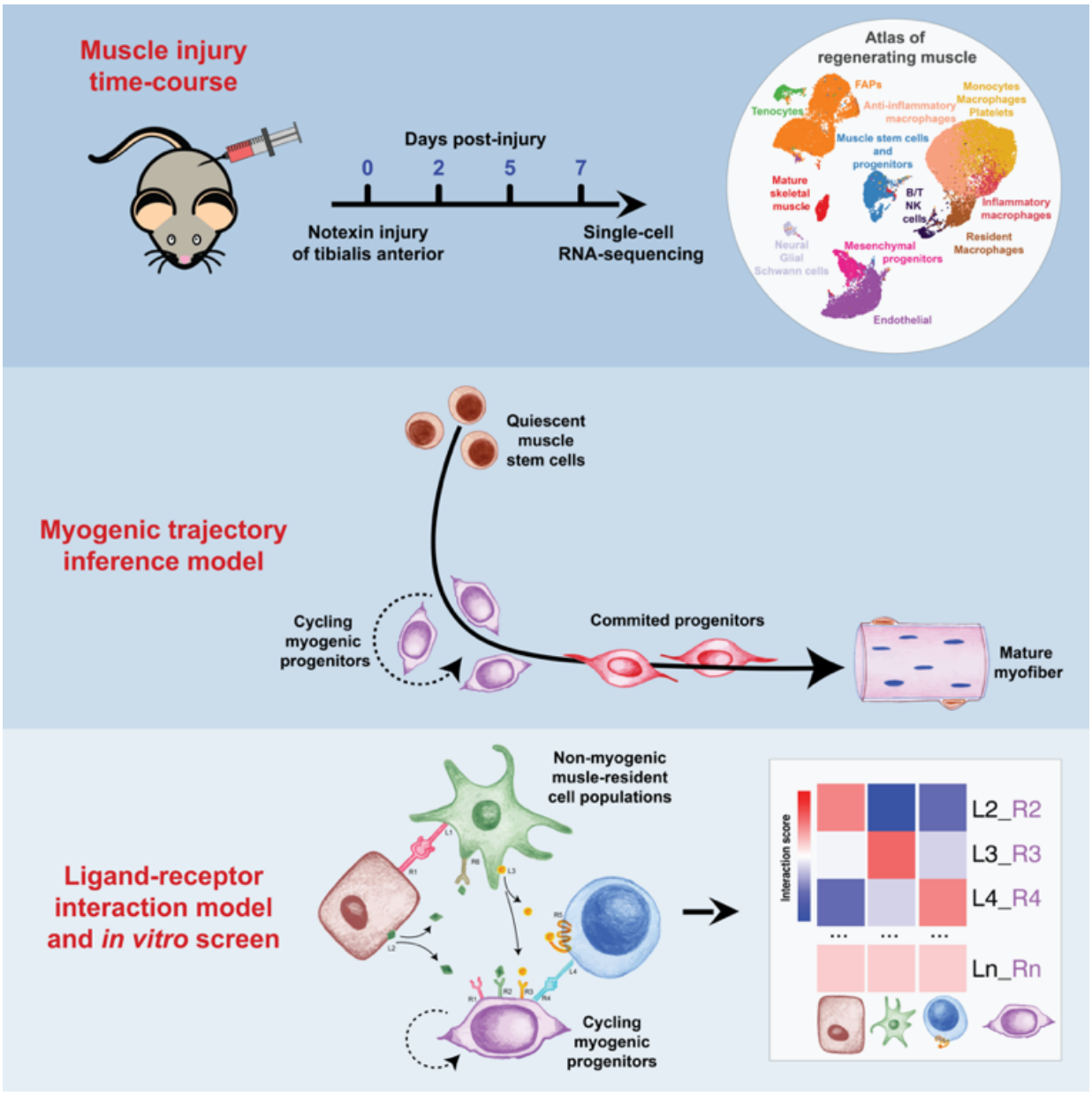


Figure 7 – Graphical abstract of the study.

3.1 Introduction

Muscle stem cells are essential for skeletal muscle homeostasis and regeneration throughout lifespan (Blau et al., 2015; Wang and Rudnicki, 2011). Muscle regenerative process is orchestrated by a network of ligand-receptor interactions with a variety of cell types including immune cells, endothelial cells, and FAPs (Wosczyna and Rando, 2018). For instance, FAPs secrete fibronectin, insulin-like growth factor-1 (IGF-1), and other matrix proteins and growth factors to coordinate muscle tissue repair through the regulation of myogenic cell fates and the clearance of cellular debris (Heredia et al., 2013; Joe et al., 2010; Lukjanenko et al., 2016).

A continuum of myogenic stem and progenitor cell populations are present in regenerating muscle (Motohashi and Asakura, 2014; Tierney and Sacco, 2016). Following injury, quiescent Pax7-expressing MuSCs progress in the cell cycle and exhibit an activated myogenic expression program characterized by genes such as *Myf5*, *Myod1*, and *Myog* (Zammit, 2017). The concept of the myogenic cell lineage was largely derived from lineage tracing and prospective isolation studies using myogenic regulatory factors and cell cycle stages to define cell states (Biressi and Rando, 2010). Myogenic stem/progenitor cells populations, enriched to high purity through surface antigen profiles and/or transgenic reporters, nonetheless exhibit substantial molecular and functional heterogeneity throughout adulthood (Chakkalakal et al., 2014; Cornelison and Wold, 1997; Cosgrove et al., 2014; Kuang et al., 2007; Porpiglia et al., 2017; Rocheteau et al., 2012; Sacco et al., 2008; Sousa-Victor et al., 2014; Tierney et al., 2018). These findings suggest that myogenic stem/progenitor cell lineage can be interpreted as a continuum of

hierarchical cell states. However, it remains unresolved how global profiles in cell cycle mediators, regulatory factors, and surface markers define this myogenic continuum.

Recent advances in single-cell analyses and algorithms provide potent new strategies to infer cell differentiation trajectories (Hwang et al., 2018; Wagner et al., 2016). Here, we generate a single-cell transcriptomic atlas from injured tissues to describe the myogenic continuum and multicellular communication networks involved in mouse muscle repair. We use single-cell RNA-sequencing (scRNA-seq) to collect a multi-cell-type transcriptomic reference time-course, spanning four timepoints post-injury and over 34,000 single-cell transcriptomes. We analyze this atlas to identify the compositional and gene expression dynamics of the cellular constituents of muscle repair in a manner not biased by prior knowledge or lineage tracing. Using trajectory inference, we organize more than 3,200 individual myogenic cell transcriptomes in pseudotime to reveal their hierarchical organization and identify regulatory factor and surface marker expression profiles unique to distinct myogenic subpopulations. Finally, we develop a ligand-receptor interaction model to identify cell- and stage-specific communication signals between non-myogenic cells and myogenic cell subpopulations involved the muscle repair process. From a panel of ligands identified in the model, we performed an *in vitro* stimulation experiment to evaluate their effect on MuSC and progenitor cell proliferation as well as their dependencies on Syndecan receptors.

3.2 Results

3.2.1 A single-cell RNA-sequencing atlas of mouse muscle regeneration

Skeletal muscle regeneration in response to local tissue damage depends on the coordinated interactions of multiple myogenic and non-myogenic cell types over a time-course of weeks (Wosczyna and Rando, 2018). To gather a comprehensive view of this process, we generated a transcriptomic atlas of adult mouse hindlimb muscle regeneration using droplet-based single-cell 3' RNA-sequencing (scRNA-seq) on the 10X Chromium platform. We collected tibialis anterior (TA) muscles of healthy adult (4-7 month) C57BL6 mice at 0 (no injury), 2, 5, and 7 days following injection of the myotoxin notexin to induce myofiber damage (n = 2-3 mice per time-point). We dissected and then enzymatically digested the TA muscles into single-cell suspensions and then filtered to remove cellular debris and, in some samples, applied red blood cell (RBC) lysis before performing scRNA-seq (Fig. 9A).

We used the Seurat package for scRNA-seq data filtering and processing (see Methods). We removed cells with fewer than 200 genes detected, fewer than 1000 UMIs, or with more than 20% of UMIs mapped to mitochondrial genes (Fig. S1A-B). Applying these filters eliminated erythroblasts to levels similar to those observed in samples where RBC lysis was performed, suggesting that the RBC step could be omitted (Fig. S1D). We also found that the use of Calcein-AM based FACS sorting to enrich for viable cells altered the cellular composition. Specifically, cell sorting based on Calcein-AM positivity decreased the incidence of MuSC and progenitor cell as well as anti-inflammatory macrophage populations (Fig. 8). This is could be due to biases introduced by viability selection, perhaps owing to reduced metabolism of Calcein-AM by quiescent cells. Based

on this observation, we omitted FACS viability sorting in the collection of these scRNA-seq datasets.

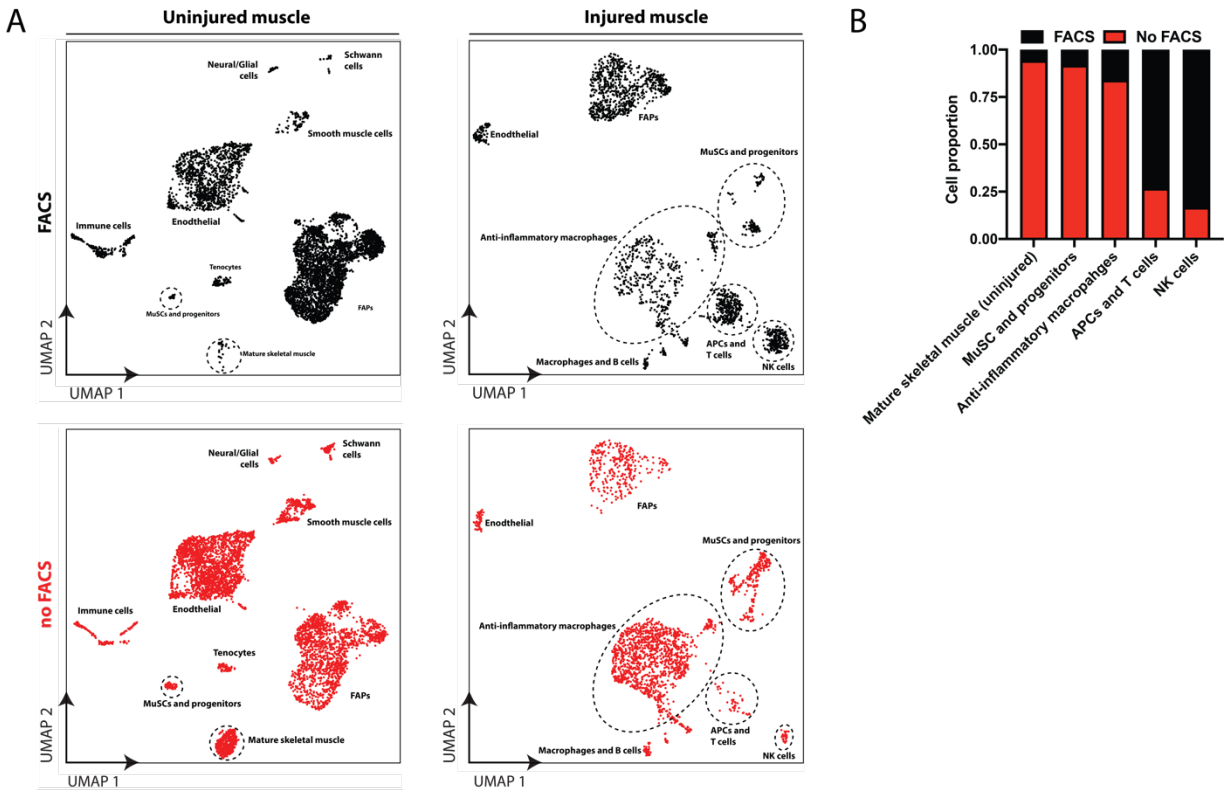
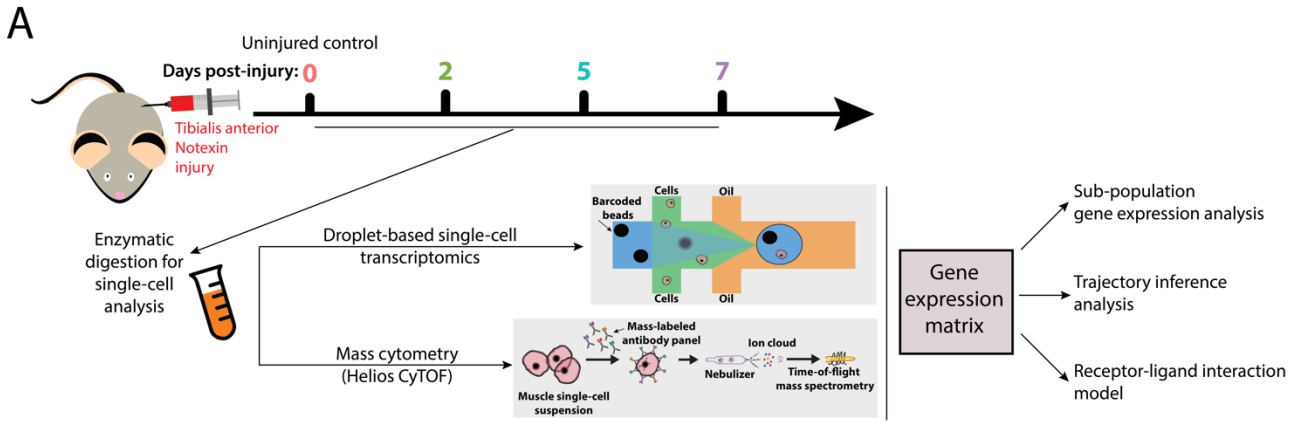


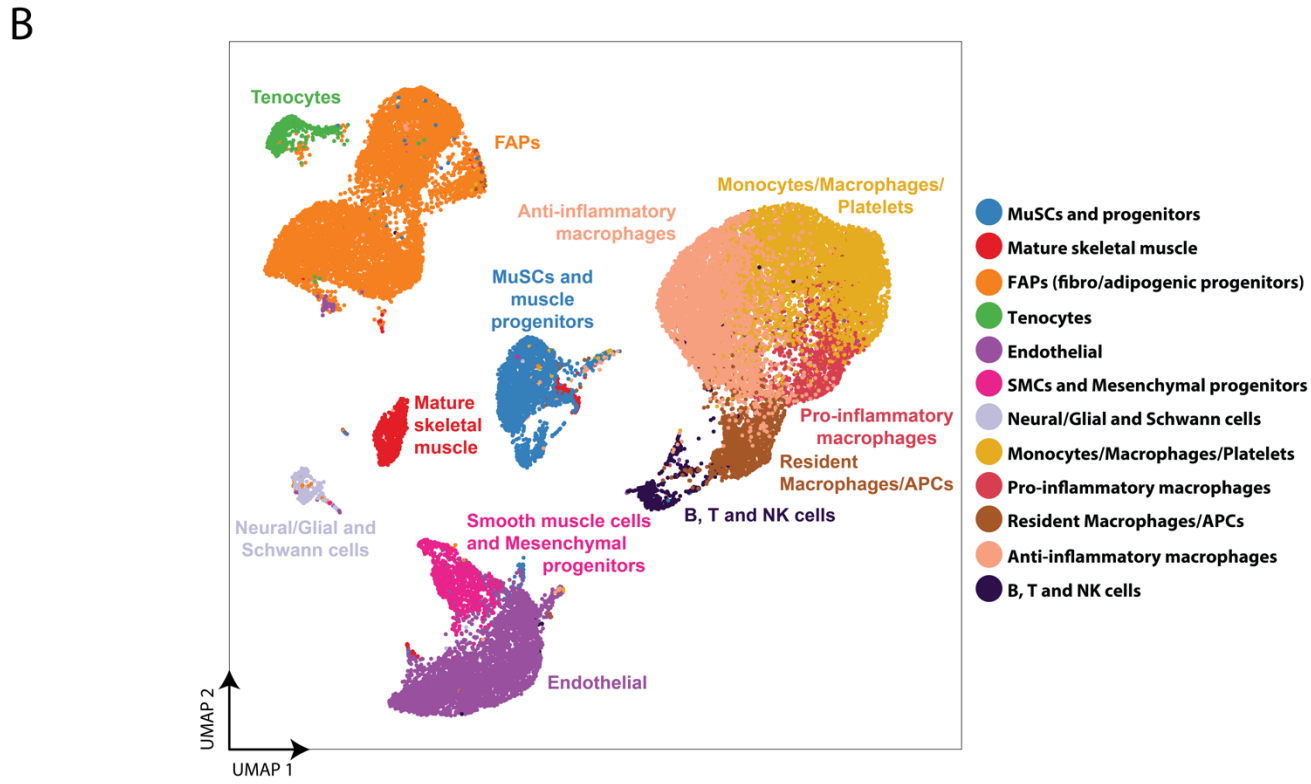
Figure 8 – Evaluation of selection biases and resulting cell type abundances introduced by FACS filtering prior to scRNA-seq. (A) UMAP projection of single-cell transcriptomes obtained by either the standard cell isolation protocol (see **Methods**) or to one that included a Calcein-AM+ enrichment of live cells and removal of debris by FACS. Each condition represents 2000 randomly chosen cells. FACS selection introduces measurable biases in cell population number for MuSCs, mature skeletal muscle in the uninjured muscle and various immune cell types in the injured muscle. **(B)** Cell proportions comparing the FACS to no FACS conditions normalized to the number of total cells in the cluster (dashed circle on UMAP plot).

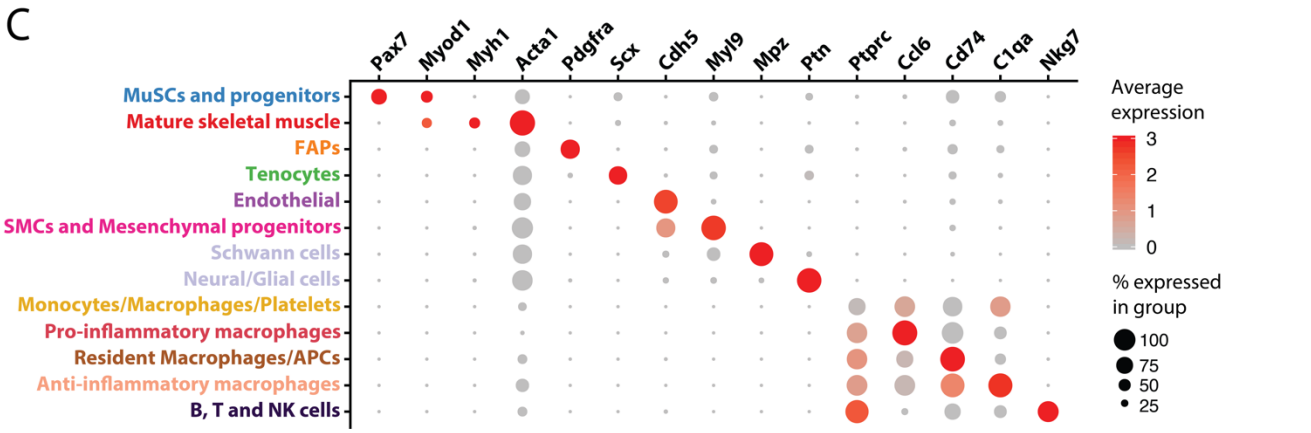
After filtering, the scRNA-seq datasets each contained on average 3,444 ± 1,286 cells. We compared time-point replicate datasets generated from different mice (n=2-3) to evaluate batch effects and mouse-to-mouse variability. We observed only minor differences, most notably a small variation in the incidence of macrophage population in day 7 post-injury samples (Fig. S1C), suggesting that these datasets contained minimal batch effects. Thus, for subsequent analyses, we combined time-point biological

replicates, without batch correction, to improve cell sample size and statistical power. Initially, the samples were assembled into a unified transcriptomic atlas containing 34,438 cells, expressing a total of 19,584 detectable genes (Figs. 9B and S1A). We next annotated the cell types present in this muscle regeneration atlas. We employed uniform manifold approximation and projection (UMAP) to visualize the individual transcriptomes of all cells in the unified dataset (Fig. 9B) (Becht et al., 2018). Independently, we performed unsupervised shared nearest neighbor (SNN) clustering, which partitioned cells into 12 groups based on their transcriptomic programs (Fig. 9B). We observed more refined groups, potentially revealing additional subtypes and cell states, when SNN clustering was applied to samples collected at individual time-points (Fig. 11). In examining the full atlas, we interpreted the clusters as broadly defining 12 different cell populations. To identify these populations, we examined the normalized expression level and frequency of canonical cell type genes and named them based on their exclusivity in these expression patterns (Fig. 9C). To discriminate more ambiguous populations, especially in the immune subpopulation, we performed differential expression analysis using a negative binomial model between cells within the cluster and all other cells in the atlas (Figs. S2-S4).

Figure 9 (*next page*) **– Assembly and curation of a scRNA-seq atlas of mouse muscle regeneration.** (**A**) Experimental design overview. Cell suspensions were collected from digested TA muscles of adult mice at various time points (0 [no injury], 2, 5, and 7 days) following notexin injury (n=2-3) and subjected to scRNA-seq and mass cytometry (CyTOF), followed by downstream analyses. (**B**) Complete 34,438 cell transcriptomic atlas assembled from all sample time-points. Data is presented as a UMAP projection to visualize variation in single-cell transcriptomes. Unsupervised SNN clustering resolved at least 12 distinct types of cells (color-coded in legend). More resolved cell type clusters, distinguishing Neural/Glial from Schwann cells, immature B from cytotoxic T cells, were evident when analyzing time-points individually (see Fig. 11A). (**C**) Identification of cell types from SNN clusters based on cluster-average expression of canonical genes. Dot size represents the percentage of cells with a non-zero expression level and color-scale represents the average expression level across all cells within cluster.


A
Uninjured control
Days post-injury: 0
2
5
7
Tibialis anterior
Notexin
injury
Enzymatic
digestion for
single-cell
analysis
Droplet-based single-cell
transcriptomics
Barcoded
beads
Cells
Oil
Cells
Oil
Mass cytometry
(Helios CyTOF)
Mass-labeled
antibody panel
Ion cloud
Nebulizer
Time-of-flight
mass spectrometry
Muscle single-cell
suspension
Gene
expression
matrix
Sub-population
gene expression analysis
Trajectory inference
analysis
Receptor-ligand interaction
model


B
Tenocytes
FAPs
Anti-inflammatory
macrophages
Monocytes/Macrophages/
Platelets
MuSCs and
muscle
progenitors
Mature
skeletal
muscle
Pro-inflammatory
macrophages
Resident
Macrophages/APCs
B, T and NK cells
Neural/Glial and
Schwann cells
Smooth muscle cells
and Mesenchymal
progenitors
Endothelial
UMAP 2
UMAP 1
MuSCs and progenitors
Mature skeletal muscle
FAPs (fibro/adipogenic progenitors)
Tenocytes
Endothelial
SMCs and Mesenchymal progenitors
Neural/Glial and Schwann cells
Monocytes/Macrophages/Platelets
Pro-inflammatory macrophages
Resident Macrophages/APCs
Anti-inflammatory macrophages
B, T and NK cells


C
Pax7
Myod1
Myh1
Acta1
Pdgfra
Scx
Cdh5
Myl9
Mpz
Ptn
Ptprc
Ccl6
Cd74
C1qa
Nkg7
MuSCs and progenitors
Mature skeletal muscle
FAPs
Tenocytes
Endothelial
SMCs and Mesenchymal progenitors
Schwann cells
Neural/Glial cells
Monocytes/Macrophages/Platelets
Pro-inflammatory macrophages
Resident Macrophages/APCs
Anti-inflammatory macrophages
B, T and NK cells
Average
expression
3
2
1
0
% expressed
in group
100
75
50
25

We observed a population of myogenic progenitors, containing MuSCs and myoblasts, which expressed the myogenic transcription factors *Pax7* and *Myod1* (Wang and Rudnicki, 2011). We detected a population of mature myocytes and/or myofibers, which expressed *Myh1* (myosin heavy chain 1) and *Acta1* (skeletal muscle alpha actin 1), both proteins involved in the contractile function of terminally differentiated skeletal muscle cells (Lyons et al., 1990). The limited incidence of this population is likely due to the size-filtering out of multinucleated myofibers during the cell isolation protocol. *Acta1* is also expressed, but at a lower level and frequency, in other cell types, which is not surprising given that it is a common component of the cytoskeleton and plays a role in cell migration. We also identified a population of *Pdgfra*-expressing fibro/adipogenic progenitor cells (FAPs) (Uezumi et al., 2011) and tenocytes that express the tenogenic transcription factor *Scx* (scleraxis), glycoprotein encoding genes *Fmod* (fibromodulin) and *Tnmd* (tenomodulin) (Fig. S2) (Docheva et al., 2005; Schweitzer et al., 2001). Two closely related *Pecam1* (cell adhesion molecule CD31)-expressing populations were identified as, first, smooth muscle cells (SMCs) and mesenchymal progenitors that exclusively express *Myl9* (myosin light chain 9) (Gaylinn et al., 1989) and, second, endothelial cells that express *Cdh5* (VE-cadherin) (Christov et al., 2007; Zordan et al., 2014). We also find a mixed group of neuro-muscular cells, which are enriched for expression of the Schwann cell marker *Mpz* (myelin protein zero) and the neuronal gene *Ptn* (pleiotrophin) (Liu et al., 2015).

The cumulatively largest and most ambiguous group are immune cells, which dynamically infiltrate muscle and mediate inflammatory regulation of tissue repair. We

observed substantial overlap in immune cell marker gene expression (Fig. 9C), which agrees with the immune cell phenotype continuum model (Novak and Koh, 2013). We therefore examined multiple gene signatures to subdivide immune cell clusters (Figs. S3 and S4) based on prior literature (Chazaud, 2016; Tidball, 2017). One immune cluster contains resident macrophages and antigen presenting cells (APCs), as identified by expression of *Cd74* and other MHCII complex encoding genes. A second immune cluster contains a general group of monocytes, macrophages, and platelets, marked by expression of *Cd68* (common immune glycoprotein) and *Pf4* (platelet factor 4). A third immune cluster contains activated (pro-inflammatory) macrophages as identified by expression of the inflammatory neutrophil and macrophage markers including *Ccl6* (chemokine C-C motif ligand 6). A fourth immune cluster contains anti-inflammatory macrophages, as identified by expression of *C1qa*, which encodes for part of the C1 complement complex. A fifth immune cluster contains immature and mature B- and T-lymphocytes and NK cells, as identified by expression of the common lymphocyte gene *Ptprc* (protein tyrosine phosphatase receptor type C, or Cd45) and *Nkg7* (natural killer cell granule protein 7). This annotated scRNA-seq dataset provides a reference atlas to examine the cell populations and gene expression dynamics during muscle regeneration.

3.2.2 Cell-type and gene expression dynamics of muscle regeneration at single-cell resolution

Next, we analyzed the transcriptomic atlas to assess how these populations are dynamically altered in incidence and gene expression following muscle injury. Overall, we discerned between 5-12 distinct cell populations via SNN clustering at each time-point for

a total of 15 unique populations (Fig. 11A-B). Uninjured muscles (day 0) were largely composed of endothelial cells, FAPs, and mature myocytes/myonuclei, with other cell populations detected infrequently. After injury, we observed a transient increase in multiple immune cell types and a concomitant decline in non-immune populations. By day 5, the immune cell population diversified while the myogenic populations recovered. By day 7, most cell populations approached their day 0 frequencies and gene expression states (Figs. 10, 11C, S4), revealing a return to homeostasis.

Pecam1-expressing endothelial cells were the most abundant cell population at day 0 (38%) and exhibited only minor changes in detection frequency and gene expression heterogeneity throughout the time-course (Christov et al., 2007). In contrast, *Pdgfra*-expressing FAPs exhibited compositional and expression variability. FAPs were a substantial cell fraction of uninjured muscles (31%) and expressed the ECM genes *Col3a1*, *Dcn*, and *Gsn* in homeostasis (Figs. 10, 11A-C, S4). At day 2 post-injury, FAPs also express *Ccl7*, *Cxcl5*, and *Cxcl1*, chemokines that attract monocytes and neutrophils. FAPs have elevated expression of *Col1a1*, *Col1a2* and *Mt* and diminished expression of *Gsn* and *Dcn*, suggesting an “activated” or “remodeling” state (Fig. 10). At day 5, FAPs have further elevated expression of *Col1a1* and *Col1a2* and also express *Postn*, *Bgn,* and *Sparc* (Fig. 11C). At day 7, the FAP expression profile resembles the uninjured samples, albeit with more heterogeneity, indicating a resolution to homeostasis (Fig. 10).

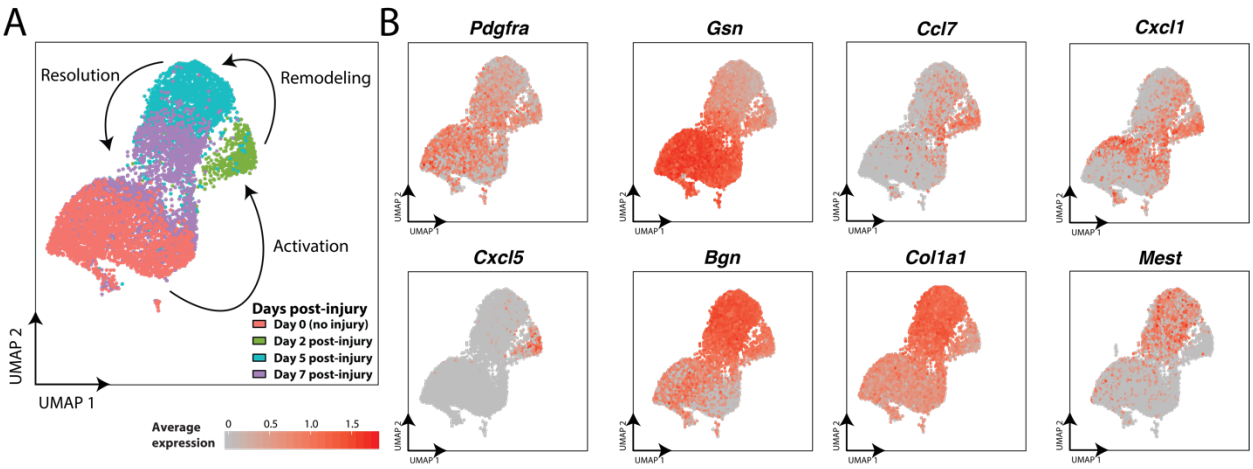
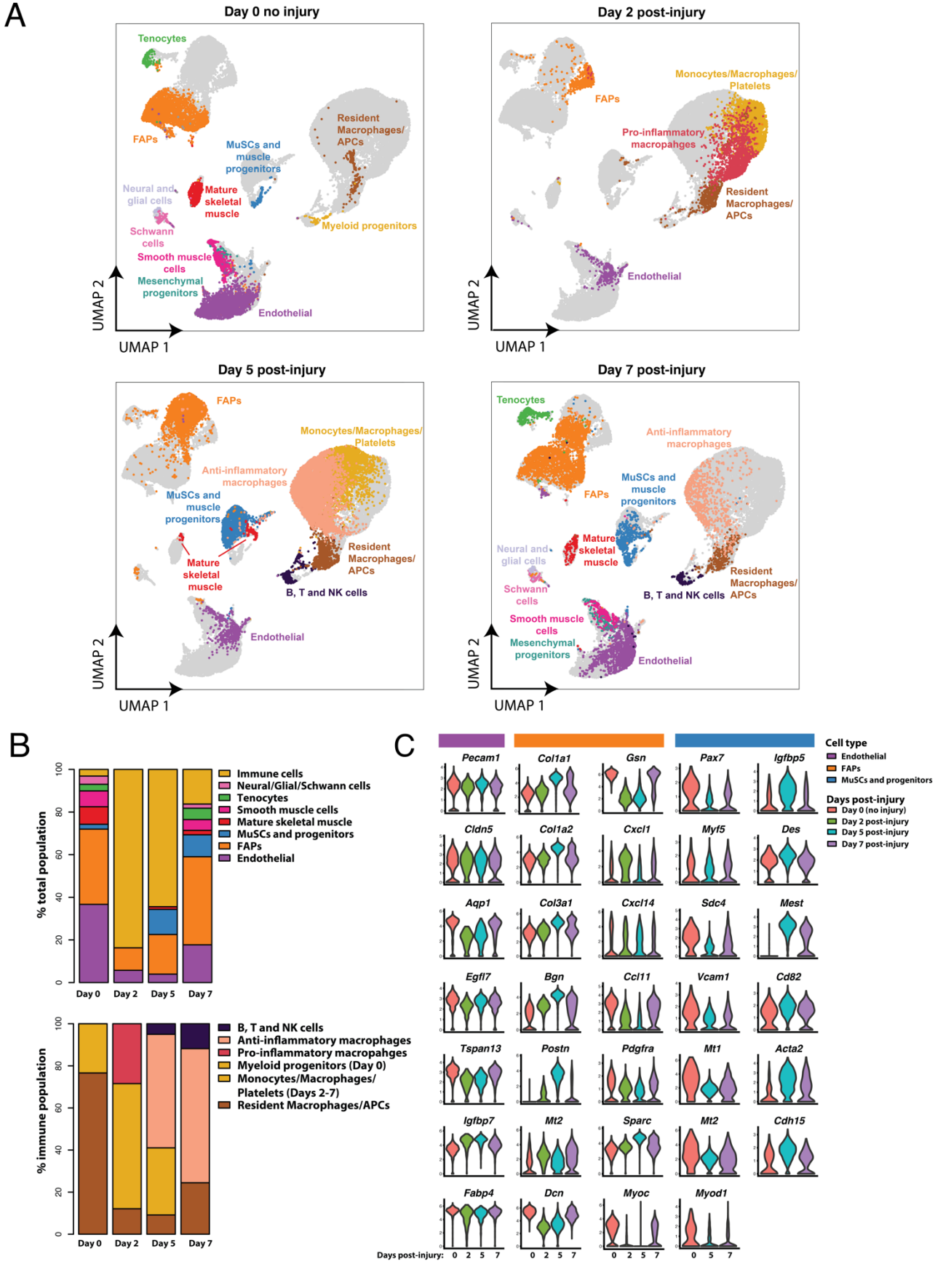


Figure 10 – Gene expression dynamics of FAPs markers post-injury. (**A**) UMAP projection of single-FAP transcriptome colored by day post-injury. (**B**) Expression level of key FAP genes that describe the shift in gene expression profile post-injury.

Figure 11 (*next page*) **– Cell composition and gene expression dynamics of muscle regeneration.** (**A**) UMAP atlases of muscle single-cell transcriptomes split by time-points post-injury containing, respectively, 7,025, 5,524, 14,240, and 7,646 cells for day 0, 2, 5, and 7 days post-notexin injury. Fifteen total cell types were identified using SNN clustering applied to each time-point. Cells from other time-points are in gray. (**B**) Compositional dynamics of cell types throughout the regeneration time course. Immune cells are grouped together (top) or separated (bottom). (**C**) Violin plots presenting the heterogeneous gene expression changes for a selection of differentially expressed genes within the endothelial, FAP, and MuSC and myogenic progenitors populations at each time-point. MuSCs and myogenic progenitor cells were too rare at day 2 to analyze.


A
Day 0 no injury
Tenocytes
FAPs
MuSCs and muscle progenitors
Resident Macrophages/ APCs
Neural and glial cells
Mature skeletal muscle
Myeloid progenitors
Schwann cells
Smooth muscle cells
Mesenchymal progenitors
Endothelial
UMAP 2
UMAP 1
Day 2 post-injury
FAPs
Monocytes/Macrophages/ Platelets
Pro-inflammatory macropahges
Resident Macrophages/ APCs
Endothelial
UMAP 2
UMAP 1
Day 5 post-injury
FAPs
Monocytes/Macrophages/ Platelets
Anti-inflammatory macrophages
MuSCs and muscle progenitors
Mature skeletal muscle
Resident Macrophages/ APCs
B, T and NK cells
Endothelial
UMAP 2
UMAP 1
Day 7 post-injury
Tenocytes
Anti-inflammatory macrophages
FAPs
MuSCs and muscle progenitors
Neural and glial cells
Mature skeletal muscle
Schwann cells
B, T and NK cells
Resident Macrophages/ APCs
Smooth muscle cells
Mesenchymal progenitors
Endothelial
UMAP 2
UMAP 1
B
% total population
Day 0
Day 2
Day 5
Day 7
Immune cells
Neural/Glial/Schwann cells
Tenocytes
Smooth muscle cells
Mature skeletal muscle
MuSCs and progenitors
FAPs
Endothelial
% immune population
B, T and NK cells
Anti-inflammatory macrophages
Pro-inflammatory macropahges
Myeloid progenitors (Day 0)
Monocytes/Macrophages/ Platelets (Days 2-7)
Resident Macrophages/APCs
C
Pecam1
Col1a1
Gsn
Pax7
Igfbp5
Cldn5
Col1a2
Cxcl1
Myf5
Des
Aqp1
Col3a1
Cxcl14
Sdc4
Mest
Egfl7
Bgn
Ccl11
Vcam1
Cd82
Tspan13
Postn
Pdgfra
Mt1
Acta2
Igfbp7
Mt2
Sparc
Mt2
Cdh15
Fabp4
Dcn
Myoc
Myod1
Days post-injury: 0 2 5 7
Cell type
Endothelial
FAPs
MuSCs and progenitors
Days post-injury
Day 0 (no injury)
Day 2 post-injury
Day 5 post-injury
Day 7 post-injury

Likewise, the immune cell populations within this atlas exhibit notable dynamics. At day 0, we detect that immune cell populations comprise 5% of the uninjured muscle and can be subdivided into a more defined group of *Cd79*$^{+}$ *Ly6d*$^{+}$ immature myeloblasts, *Cd3*$^{+}$ T-cells, and a group of resident macrophages and APCs expressing *Lyz2, Cd74*, and *Ccl6* (Fig. S2). At day 2, the inflammatory response initiates, as reflected by a dramatic increase in the frequency (84%) of immune cells detected (Fig. 11A). This early-response immune compartment can be divided into three populations. First, we observed APCs characterized by expression of MHC class II proteins such as *Cd74* and the *H2* family. Within this group we distinguish small populations of *Cd7*$^{+}$ mature T-cells, *Klrd1*$^{+}$ NK cells, and *Cd209a*$^{+}$ dendritic cells. Second, we identified pro-inflammatory macrophages (including activated M1-like macrophages) that express *Ccl9* (a chemokine that attracts *Cd11b*$^{+}$ *Ccr1*$^{+}$ dendritic cells), *Ccr2* (a chemokine involved in monocyte chemotaxis), and *Ly6c2.* Third, we observed a less defined group of cells that express a wide variety of markers including *Cd68*$^{+}$ monocytes and *Pf4*$^{+}$ platelets (Figs. S3, S4). At day 5 post-injury, immune cells remain prevalent (64%) (Fig. 11A). We observed however a shift in the macrophage population from a pro-inflammatory to an anti-inflammatory phenotype. We found that these macrophages express *C1q* complement genes as well as *Apoe*, characteristic of the M2-like anti-inflammatory phenotype (Baitsch et al., 2011; Ho et al., 2016). We detect some *Aif1*$^{+}$ macrophages and/or dendritic cells that have been reported to modulate muscle repair (Kuschel et al., 2000). We found elevated frequencies of APCs, *Lsp1*$^{+}$ *Ccr7*$^{+}$ B-cells, *Ccl5*$^{+}$ *Xcl1*$^{+}$ T-cells and *Nkg7*$^{+}$ *Klrd1*$^{+}$ NK cells (Fig. S4). At day 7, immune cells were more infrequent (17%) and started to

resemble their day 0 cell type composition, with some M2-like macrophages, T-lymphocyte and NK-cells remaining.

The myogenic cell populations exhibited a temporal profile inverted in abundance relative to the immune cells and transcriptionally heterogeneous. At day 0, we detected a small population (2%) of MuSCs and progenitor cells expressing *Pax7*, *Sdc4*, *Vcam1*, and *Myod1*, and also detected *Acta1*$^{+}$ *Myl1*$^{+}$ mature myocytes and/or myofiber nuclei (9%). At day 2, we did not detect any MuSCs or progenitors though they are reported to be in an activated state following notexin injury. This surprising finding is likely due the relative rarity of myogenic cells compared to immune cells, providing a sampling challenge at this time-point. At day 5, we observed a large expansion of the MuSC and progenitor population (12%) and decline of mature skeletal muscle (or terminally differentiated myocytes) (1%) as expected. This population heterogeneously expresses lower levels of *Pax7*, *Vcam1,* and *Sdc4*, higher levels of *Myod1, Des*, *Mest*, and the cell cycle inhibitor *Cdkn1c*, suggesting they include cells in varied activation and/or cycling states (Fig. 11C). This population also induces expression of *Igfbp5,* known to regulate myogenic differentiation (Ren et al., 2008). At day 7, the myogenic cell population remains in a heterogeneously activated and differentiation state with some recovery of *Pax7*$^{+}$ MuSC population (Fig. 11C).

Lastly, some cell populations were almost exclusively found at days 0 and 7. We observed a small (3%) population of tenocytes that express the canonical marker *Scx* (scleraxis), *Tnmd* (tenomodulin), *Fmod* (fibromodulin), and *Thbs4* (thrombospondin) (Giordani et al., 2019) (Fig. S2). We also observed two closely related populations of

mesenchymal progenitors and smooth muscle cells, which were distinguished based on their expression of *Myh11* and *Acta2.* In a similar manner, Schwann cells (enriched for *Mpz*, *Mbp*, *Fxyd3*, and *Prx*) were distinguished from other related neural and glial cells (enriched for *Ptn*, *Postn*, *Cadm1*, *Lyz2* and *Col20a1*) (Fig. S2).

3.2.3 Single-cell trajectory inference organizes a myogenic cell continuum involved in muscle regeneration

The consensus model of adult muscle regeneration states that a subset of MuSCs leave quiescence after tissue damage and enter an activated cycling state to generate progeny through a combination of asymmetric and symmetric division events (Fig. 2) (Wang and Rudnicki, 2011). We asked whether unbiased analyses could reconstruct this consensus model and provide new insights into the continuum of myogenic cell-states.

First, we explored the cellular heterogeneity within the cumulative myogenic cell population by selecting the MuSCs, progenitors, and mature myocytes from the unified transcriptomic atlas. Unbiased SNN clustering revealed five sub-populations of myogenic cells (Fig. S3A). These subpopulations clarify the myogenic heterogeneity in two ways: across post-injury time-points and within the stage of differentiation (Fig. S5B). Cluster 5 is comprised of terminally differentiated myocytes expressing elevated levels of Acta1 and Myh1 from all time-points, whereas clusters 1-4 represent a heterogeneous population of MuSCs and their progeny. Cluster 1 contains cells from both days 0 and 7 post-injury and is largely enriched for Pax7, Sdc4, Vcam1, and Cd34 transcripts, therefore are likely quiescent MuSCs (Fig. S5C). Clusters 2, 3, and 4 contain cells from days 5 and 7 post-injury and that heterogeneously express the myogenic activation and commitment

markers Myf5, Myod1, and Myog and have reduced expression of MuSC markers, suggesting these each contain a mixture of activated MuSC and myoblasts. Notably, clusters 2-4 likely represent a varied but sequential transition towards myogenic commitment, which is exemplified by stepwise elevation in Myog expression (Fig. S5C). Therefore, we conclude that cluster 4 is comprised of committed myoblasts whereas clusters 2 and 3 primarily contain activated MuSCs. These clusters describe an organized partitioning of the myogenic cell population.

Next, given that SNN clustering lacks hierarchical structure, we sought to organize these subpopulations using a trajectory inference model to delineate their interrelatedness. We applied Monocle reverse graph embedding (Qiu et al., 2017) to the cumulative myogenic cell population within the atlas to infer a hierarchical trajectory (Fig. 12A). The Monocle analysis focused on differentially expressed genes within one of these five myogenic clusters and aligns cells into a one-dimensional “pseudotime” axis. The pseudotime trajectory presented an organized, branched progression of cells from quiescent MuSCs to cycling and differentiating progenitors to terminally differentiated myocytes, which can be seen by labelling individual cells using the cell population annotations from the unified atlas (Fig. 9B). Both the beginning and ending branches within the trajectory are composed of cells from days 0 and 7 post-injury, indicated they consist of Pax7hi quiescent MuSCs and Acta1hi terminally differentiated myocytes, respectively, which are both absent at intermediate time-points (Fig. 12A). A subset of cells diverts at the central node into a third branch comprised of day 5 and 7 post-injury cells enriched for Cdk1 and Cdc20, indicating that they are actively cycling myoblasts.

Near the central node, we observed day 5 and 7 post-injury cells that expressed myoblast and myocyte markers such as Myog and Myod1 (not shown). We interpret this branch structure as a bidirectional trajectory in which activated and cycling progenitors (predominantly from day 5 post-injury) can bifurcate either towards further commitment into terminal myocytes needed for myofiber repair or towards a return to quiescence.

Then, we performed a differential expression analysis along the pseudotime axis in order to identify genes that explain this myogenic cell progression (Fig. 12B). The top 86 differentially expressed genes partition in four clusters that are distinguished by their pseudo-temporal gene expression patterns (see Methods). The “early” gene cluster contains quiescence-associated MuSC genes, including Pax7, Id3 (a direct target of Pax7 and inhibitor of myogenic activation differentiation), Mt2, Klf4 (a potent inhibitor of smooth muscle differentiation), and cell cycle inhibitor Btg2 (which has been shown to interact with Id3 to regulate neural progenitor cell differentiation) (Farioli-Vecchioli et al., 2009; Kumar et al., 2009). The first “intermediate” gene cluster group contains activation- and cycling-associated genes, including Myod1, Hmgb2 (regulates MuSC differentiation through IGF-2), as well as multiple mitotic genes such as the cyclin-dependent kinases Cks2 and Cdk1, Smc4 (essential to condense chromatin), and the anti-apoptotic gene Birc5. The second “intermediate” gene cluster contains commitment-associated genes such as the cyclin-dependent kinase inhibitor Cdkn1c, the myogenic differentiation regulator Igfbp5 (Ren et al., 2008), and genes encoding transmembrane proteins Cdh15, Itm2a, Cd82 and Cd63. The “late” gene cluster contains terminal differentiation-

associated genes such as Myl1, the troponin family complex members (e.g., Tnni2), and metabolic enzymes Gyg and Ak1 (Janssen et al., 2000).

Last, we analyzed the trajectory of the myogenic stem/progenitor cell population (other cell types excluded from this analysis, Figs. 12C-D and S5D). This refined trajectory model identifies three branches of immature myogenic cells within the transcriptomic atlas (Fig. 12C-D), distinguished by their quiescent ("Qu"), cycling ("Cy"), and non-cycling committed ("Co") signatures. The "Qu" branch is enriched for Btg2 and Id3 expression, whereas the "Cy" branch is enriched for G2-state mitotic genes Cdk1, Cdc20, and Ccnb2, and the "Co" branch, is enriched for the cell cycle inhibitory genes Cdkn1c and Myog (Figs. 12D and S5E). When visualized with respect to the pseudotime progression axis, these genes help distinguish between cells likely to project into "cycling" and "committed" cell fates (Fig. 12D). In conclusion, Monocle trajectory inference analysis of myogenic single-cells transcriptomes during tissue regeneration confirms the generally accepted model of adult myogenesis and provides a hierarchy of intermediate myogenic populations based on cell cycle and commitment gene expression patterns. The clustering and trajectory model together suggest that Sdc4, Id3, and Btg2 are transcriptional markers that aid in decomposing MuSC heterogeneity and may regulate distinct MuSC states.

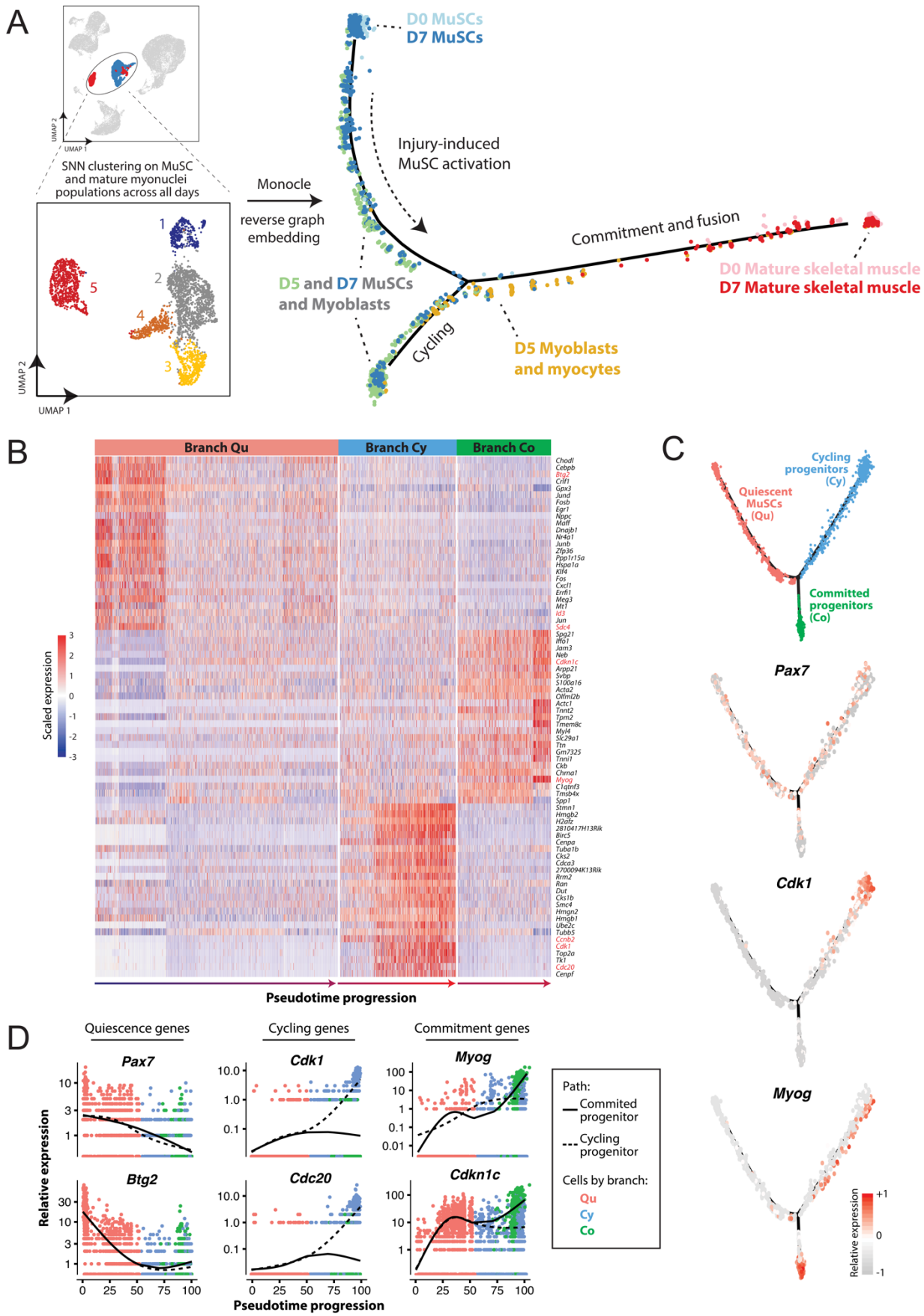

A
UMAP 2
UMAP 1
SNN clustering on MuSC and mature myonuclei populations across all days
1
2
3
4
5
Monocle
reverse graph embedding
D0 MuSCs
D7 MuSCs
Injury-induced MuSC activation
Commitment and fusion
D0 Mature skeletal muscle
D7 Mature skeletal muscle
D5 and D7 MuSCs and Myoblasts
Cycling
D5 Myoblasts and myocytes
B
Branch Qu
Branch Cy
Branch Co
Scaled expression
Pseudotime progression
C
Cycling progenitors (Cy)
Quiescent MuSCs (Qu)
Committed progenitors (Co)
Pax7
Cdk1
Myog
Relative expression
D
Quiescence genes
Cycling genes
Commitment genes
Pax7
Cdk1
Myog
Btg2
Cdc20
Cdkn1c
Relative expression
Pseudotime progression
Path:
Commited progenitor
Cycling progenitor
Cells by branch:
Qu
Cy
Co

Figure 12 (*previous page*) **– Inferring a muscle stem/progenitor cell hierarchy using Monocle pseudotime model.** (**A**) All cells within the muscle stem/progenitor and mature myocyte cell clusters (3,276 total cells) from day 0, 5, and 7 post-injury (*top left*) were selected and re-analyzed with SNN/UMAP (*bottom left*) and Monocle reverse graph embedding (*right*). Graph embedding results are presented with cells color-coded by day and labeled with cluster identities. (**B**) Heatmap showing three group of "pseudo-synchronous" genes that are differentially expressed along the Monocle pseudotime axis (Fig. 12C). Mitotic genes are labeled in red. (**C**) A refined analysis of muscle stem and progenitor sub-populations, after removal of mature myocytes, by Monocle trajectory inference. Monocle feature plots are presented showing three branch groups (*Qu, Cy, Co*) connected by a learned manifold (black lines). Same colors are used to associate individual cells with branch groups in (C, *top left*) and (D). The abundance of *Pax7*, *Myog*, and *Cdk1* transcripts are plotted for individual cells using a Z-score normalized color-scale. (**D**) Pseudotime ordered single-cell expression trajectories for genes enriched in the quiescence (Qu) cluster (*Pax7*, *Btg2*), in the cycling (Cy) cluster (*Cdk1*, *Cdc20*), and in the commitment (Co) cluster (*Myog*, *Cdkn1c*). Overlaid lines correspond to inferred cell trajectories associated with ending in the cycling (hatched) and commitment (solid) clusters.

3.2.4 Diversification of Syndecan receptor expression in myogenic stem and progenitors provides stage-specific heterotypic cell communication channels

Cell communication signals, acting through secreted ligands binding to receptors on muscle stem and progenitor cells, govern a multitude of cell-fate regulation mechanisms critical for muscle homeostasis and regeneration (Yin et al., 2013). To explore the dynamic cell communication network governing muscle repair, we generated a model that scores for interactions between receptors expressed by non-mature myogenic cells and ligands expressed by other cell types. We reasoned that this model could identify ligand-receptor co-expression pairs that provide "insulated" heterotypic cell-cell interactions potentially influencing cell-fate outcomes in myogenic cells but no other cell types in muscle. First, we considered possible ligand-receptor pairs from a database containing 2,009 mouse intercellular interaction signals (Skelly et al., 2018). Second, we identified receptor genes from this list that are differentially expressed in the MuSC and progenitor cell populations relative to all other cells within the transcriptomic atlas at any time-point. Third, we calculated interaction scores by multiplying the average transcript expression value of each differentially expressed receptor gene in the myogenic

stem/progenitor cell populations by the expression value of each cognate ligand gene (averaged over all cells within each other annotated cell population). We note this model does not consider spatial proximity between cell types, whether proteins are expressed, interaction kinetics, or whether the putative interaction pairs are documented specifically within myogenic cells.

Our model identifies 63 and 158 ligand-receptor pairs for the uninjured and injured (days 5 and 7 post-injury, combined) muscles, respectively, and 87 of these pairs were unique to injured samples (Fig. S6). We only select interactions where the receptor is differentially expressed in the MuSC and progenitor population. Moreover, for each pair, we consider the interaction significant when the score is greater than the 50th percentile when compared to all cell types. We represent significant interactions by a pairwise chord plot (Fig. 13A). In uninjured muscles, the majority of these pairs involve myogenic cell expression of receptor genes Fgfr1, Fgfr4, Sdc4, Tgfbr3, Cd63 or Cd82, consistent with findings that MuSCs express diverse members of the FGFR, Syndecan (Sdc), TGF-β and tetraspanin families (Pawlikowski et al., 2017). Notably, Fgfr1 and Fgfr4 interactions are mediated by a diverse set of 15 FGF ligand genes expressed across multiple cell types (mature skeletal muscle, FAPs, tenocytes, and neural/glial/Schwann cells), suggesting broad redundancy in FGFR signaling interactions in uninjured muscles (Fig. S6). This ligand-receptor promiscuity is also exemplified by the pairs involving Sdc4, the only Sdc gene differentially expressed by myogenic cells within uninjured muscles, which involve multiple ligand genes from various cell types including Ccl5 (myeloid progenitors), Thbs1 and Tnc (tenocytes), Fgf2, Tfpi, Mdk, and Rspo3 (FAPs), Tgm2 and Cxcl12 (SMCs), and

Adam12 (neural/glial cells), and Fgf6 (mature myocytes). In the injured muscle we observed a diversification in the ligand-receptor interaction landscape, highlighted by inclusion of additional myogenic cell receptor genes (Sdc1, Sdc2, Notch1, Notch3, and Cd151) within the interaction pairs.

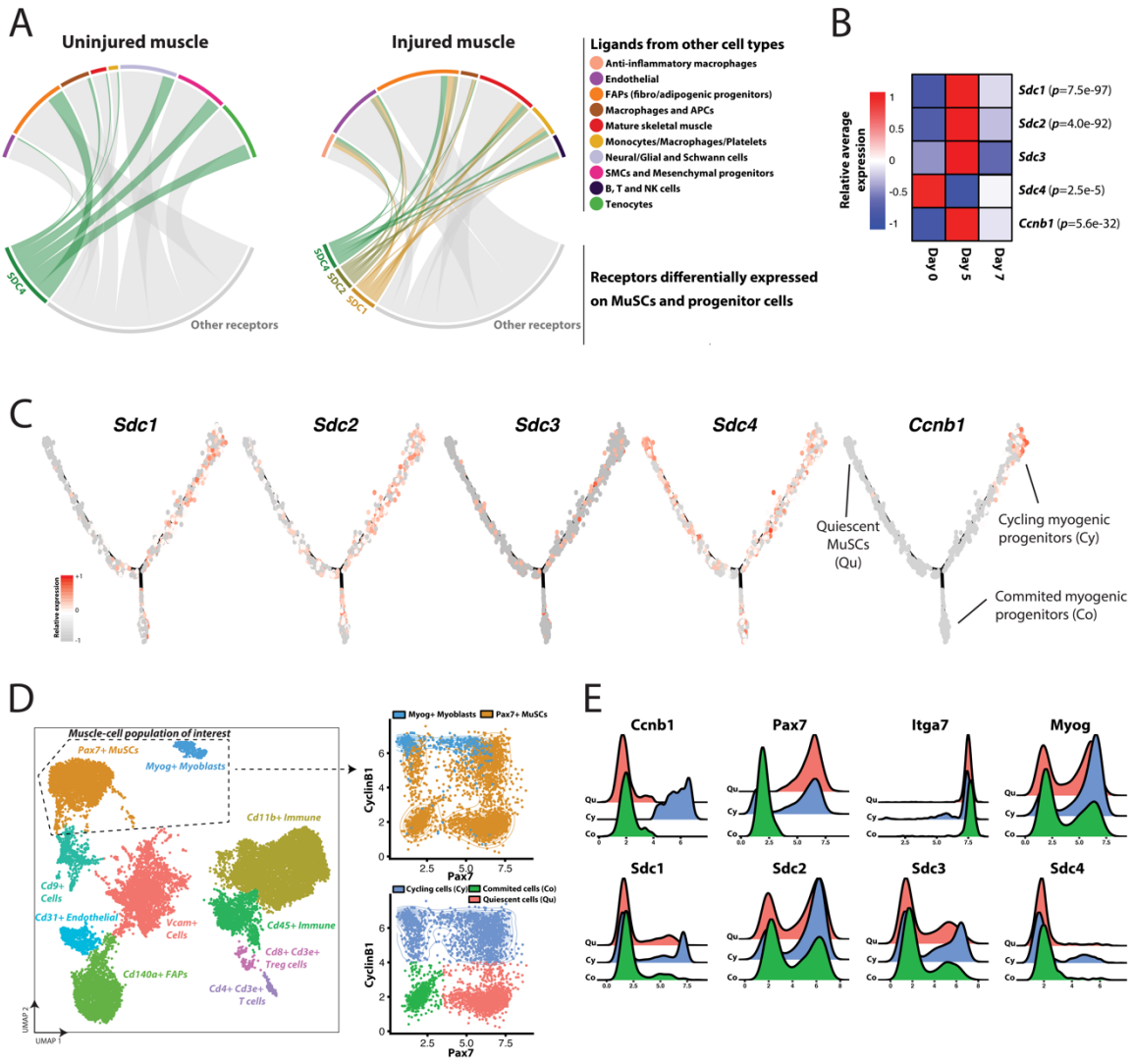


Figure 13 – Ligand-receptor model reveals diversification of communication signals linked to heterogeneously expressed Syndecan family receptors during muscle regeneration. (**A**) Chord plot summarizing the significant pairwise interactions between receptor genes differentially expressed in MuSC and progenitor cell population and ligand genes expressed by other cell types within the transcriptomic atlas. Left, uninjured (day 0) samples. Right, injured (day 5 and 7 post-injury) samples. Differentially expressed receptor genes outside of the Syndecan family are in grey. For a given ligand-receptor pair, we only represent interactions whose score (Fig. S6) is greater than the 50th percentile across all cell types. (**B**) *Sdc1/2/3/4* and *Ccnb1* (Cyclin-B1) transcript averages across all non-mature myogenic cells within the transcriptomic atlas, split by days post-injury. *p*-values listed if differentially expressed across time-points

when modeled using a negative binomial distribution. (**D**, left) CyTOF atlas, consisting of 19,028 cells collected from regenerating (day 5 post-injury) muscles and stained with a panel of 35 antibodies (see Table S1) including Syndecan-1/2/3/4 and Cyclin-B1. UMAP and unsupervised SNN clustering identified 11 populations including a population of $Pax7^+$ MuSCs (orange) and $Myog^+$ myogenic progenitors (blue). These two myogenic clusters were grouped for further analysis. (**D**, right) Cyclin-B1 versus Pax7 scatter plots. Top coded using CyTOF SNN cluster identifiers. Bottom, coded by sub-population gates: Cyclin-$B1^-$ $Pax7^+$ quiescent cells (Qu; pink), Cyclin-$B1^+$ cycling progenitors (Cy; blue), and Cyclin-$B1^-$ $Pax7^-$ committed myocytes (Co; green). (**E**) expression histograms for Syndecan-1/2/3/4 and other myogenic markers for the three subpopulations identified in (**D**).

Given the observed frequent and diverse involvement of Sdc receptor genes in these co-expressed ligand-receptor scores and their documented role in MuSC regulation (Pisconti et al., 2012), we asked whether Sdc genes exhibit stage-specific expression patterns within the myogenic compartment of the transcriptomic atlas. We performed differential expression testing on the non-mature myogenic cell populations within the transcriptomic atlas between the day 0, 5, and 7 samples, and found that Sdc1 and Sdc2 were elevated at day 5 post-injury (along with the cycling MuSC gene Ccnb1), and Sdc4 was elevated in the uninjured muscle (Fig. 13B). Further, Sdc1 and Sdc2 have enhanced expression in the Ccnb1hi “Cy” branch of the Monocle trajectory (Fig. 13C).

We validated Syndecan protein expression variation at different stages of the myogenic hierarchy through an independent analysis based on mass cytometry (CyTOF). We immunostained a single-cell suspension from digested muscles at 5 days post-notexin injury using a panel of 35 antibodies (including the Sdc1-4) to label myogenic and other cell types. After gating for live cells and removing debris (Fig. 14), the CyTOF data was analyzed using a similar bioinformatic pipeline as for scRNA-seq datasets. Unsupervised SNN clustering revealed 11 subpopulations of cells including endothelial cells, FAPs, macrophages, lymphocytes, mesenchymal progenitors, an unresolved group of $Cd9^+$ cells, and two myogenic populations (Fig. 13D). One myogenic population

consisted of Pax7^{+} MuSCs whereas the other consisted of Myog^{+} myoblasts and myocytes (Fig. 13D). These myogenic populations were merged, and sub-populations were gated based on their cycling and Pax7 expression status into quiescent cells (Cyclin-B1^{-} Pax7^{+}), cycling progenitors (Cyclin-B1^{+}), and committed myocytes (Cyclin-B1^{-} Pax7^{-}) (Fig. 13E). We observed that these three sub-populations uniformly express the myogenic surface marker Integrin-α7 but have heterogeneous expression of Sdcs. In particular, sub-fractions of cells highly expressing Sdc-1/2/3/4 were all enriched in the cycling progenitors compared to other myogenic sub-populations. These observations provide confirmation of Syndecan protein expression heterogeneity, matching the *Sdc* transcript heterogeneity observed in the hierarchically organized muscle cell atlas. Further, the transient induction of Syndecans in cycling (Cyclin-B1^{+}) muscle progenitors may allow for diversified signaling responses to secreted ligands in regenerating, but not uninjured, muscles.

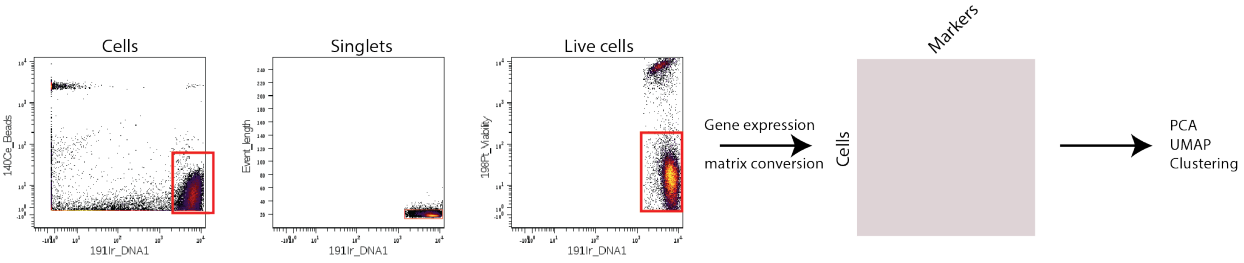


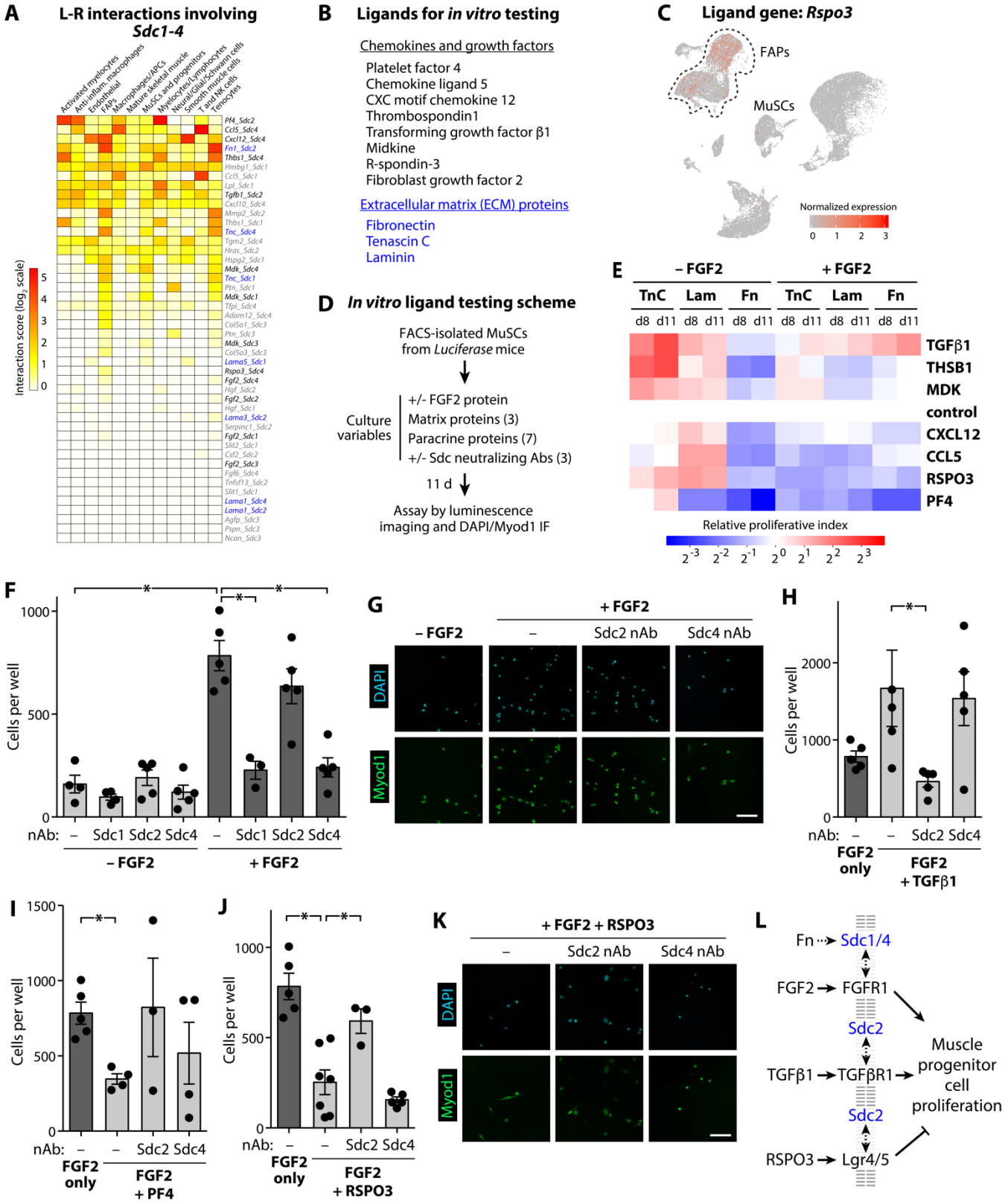
Figure 14 – Mass cytometry (CyTOF) gating strategy for day-5 injured muscle-tissue cells. Gating strategy before generation of a gene expression matrix. Cells were discriminated from debris using 140Ce beads, then from doublets by the event length, and finally from dead cells using a viability marker (198Pt-negative). The resulting gated FCS file was transformed into a gene expression matrix scaling by an inverse hyperbolic sine transformation.

3.2.5 Ligand-receptor interaction model identifies paracrine communication factors influencing MuSC proliferation *in vitro* with differential dependence on Syndecans

To examine the role of the ligand-receptor interaction pairs on MuSC function, we performed in vitro recombinant protein treatment tests on cultured MuSCs and focused on Sdc interactions given their prevalence in the transcriptomic data set (Fig. 13A). First, we inspected a rank-ordered list of interaction scores involving Sdc genes and selected ligands to test based on the highest score values and the availability of commercial recombinant proteins (Fig. 15A-B). We chose eight secreted proteins, including FGF2, TGFβ1, and CXCL12/SDF1 which have well-documented influences on muscle stem/progenitor cell fate outcomes and six others (PF4/CXCL4, CCL5, THBS1, MDK, and RSPO3) which have not been reported to play a role in MuSC function. We also chose three ECM proteins, Fibronectin (Fn), Tenascin C (TnC), and Laminin (Lam), which were involved in multiple Sdc interaction pairs. Some of these exhibit cell type-specific gene expression patterns, exemplified by Rspo3 which is almost exclusively transcribed by FAPs, but most are expressed by multiple muscle cell types (Fig. 15C). We performed an in vitro ligand treatment screen on MuSCs isolated from young-adult Luciferase transgenic mice by FACS sorting. In this screen, we tested seven ligand candidates in the context of each ECM protein (pre-coated onto tissue culture plastic) and with or without FGF2 co-treatment (Fig. 15D-E). We observed MuSC proliferation enhancement or suppression by specific ligand treatments in a context-dependent manner. Notably, TGFβ1 increased cell number whereas RSPO3 and PF4 decreased cell number across many co-treatment conditions. These trends were most consistent on Fn substrates with FGF2 mitogen co-treatment, so we used these conditions to examine the dependence of

Sdc activity by applying neutralizing antibodies to Sdc1, Sdc2 or Sdc4 (Fig. 15F-J). We observed no changes of MuSC survival and proliferation due to Sdc neutralizing antibodies in the absence of FGF, but we found that both Sdc1 and Sdc4 neutralizing antibody treatments reduced FGF2-stimulated MuSC proliferation (Fig. 15G-F). Given the interaction model results, we tested Sdc2 and Sdc4 neutralization for the other ligands in the context of Fn/FGF2 cultures (Fig. 15H-K). We observed that cell proliferation enhanced by TGFβ1 and suppressed by RSPO3 both depend on Sdc2 but not Sdc4. Across all treatments, cells were observed to uniformly exhibit Myod1 positivity indicating a mixed population of proliferating MuSCs and myoblast progenitors at day 11 of culture (Fig. 15G-K). These results suggest that Syndecan co-receptors exhibit diverse, context-dependent roles in governing muscle stem/progenitor cell fates and that specific ligand-receptor complexes may differentially engage with Sdc1/4 or Sdc2 to potentiate their signaling effects in myogenic cells (Fig. 15L).

Figure 15 *(next page)* **– Syndecan family proteins differentially mediate paracrine ligand-induced muscle stem/progenitor cell proliferation.** (**A**-**B**) Transcriptomic interaction scores for any LR pairs involving *Sdc* genes. Interaction scores calculated by averaging across all cells within the same annotated cell-type collectively between the day 0, 5 and 7 samples. Scores were summed across all cell-types and then rank-ordered in the heatmap. LR candidate pairs selected for subsequent analysis are indicated in black (for soluble factors) or blue (for matrix factors), with other pairs indicated in gray. (**C**) UMAP feature plot of the full atlas showing the scaled expression level of ligand gene *Rspo3*. (**D**) Scheme for *in vitro* testing candidate ligands. MuSCs were isolated from 3-month *Luciferase* transgenic mice by $(PI/CD45/CD11b/CD31/Sca1)^{-}$ $CD34^{+}$ Integrin-$\alpha7^{+}$ sorting. (**E**) Initial screen of candidate ligands. Proliferative index of *Luciferase*-expressing MuSCs in culture on either Tenascin C, Laminin, or Fibronectin-coated tissue culture plastic, with or without FGF2, and treated with seven different candidate ligands. Data presented as mean of n=4-8 replicates and normalized by time-point (day 8 or 11) and treatment condition to the non-ligand controls. (**F**-**K**) MuSC cultures on Fibronectin-coated plastic, treated with FGF2, TGFβ1, PF4 and/or RSPO3, and neutralizing antibodies (nAb) to Syndecans 1, 2 or 4. Proliferation assayed by nuclei counts (DAPI) and differentiation by anti-Myod1 immunofluorescence at 11 d. Data in **F**, **H**, **I**, and **J** presented as mean ± s.e.m. of n = 3-6 replicates. * indicates comparisons with $p < 0.05$ by unpaired T-test. Scale bar in **G** and **K**, 100 μm. (**L**) Summary of ligand-induced cell proliferation findings.


A
L-R interactions involving Sdc1-4
Activated myolocytes
Anti-inflam. macrophages
Endothelial
FAPs
Macrophages/APCs
Mature skeletal muscle
MuSCs and progenitors
Myelocytes/Lymphocytes
Neural/Glial/Schwann cells
Smooth muscle cells
T and NK cells
Tenocytes
Interaction score (log2 scale)
B
Ligands for in vitro testing
Chemokines and growth factors
Platelet factor 4
Chemokine ligand 5
CXC motif chemokine 12
Thrombospondin1
Transforming growth factor β1
Midkine
R-spondin-3
Fibroblast growth factor 2
Extracellular matrix (ECM) proteins
Fibronectin
Tenascin C
Laminin
C
Ligand gene: Rspo3
FAPs
MuSCs
Normalized expression
D
In vitro ligand testing scheme
FACS-isolated MuSCs from Luciferase mice
Culture variables
+/- FGF2 protein
Matrix proteins (3)
Paracrine proteins (7)
+/- Sdc neutralizing Abs (3)
11 d
Assay by luminescence imaging and DAPI/Myod1 IF
E
– FGF2
+ FGF2
TnC
Lam
Fn
d8 d11
TGFβ1
THSB1
MDK
control
CXCL12
CCL5
RSPO3
PF4
Relative proliferative index
F
Cells per well
nAb: – Sdc1 Sdc2 Sdc4
– FGF2
+ FGF2
G
+ FGF2
– FGF2
Sdc2 nAb
Sdc4 nAb
DAPI
Myod1
H
Cells per well
nAb: – Sdc2 Sdc4
FGF2 only
FGF2 + TGFβ1
I
Cells per well
nAb: – Sdc2 Sdc4
FGF2 only
FGF2 + PF4
J
Cells per well
nAb: – Sdc2 Sdc4
FGF2 only
FGF2 + RSPO3
K
+ FGF2 + RSPO3
Sdc2 nAb
Sdc4 nAb
DAPI
Myod1
L
Fn
Sdc1/4
FGF2
FGFR1
Sdc2
TGFβ1
TGFβR1
Sdc2
RSPO3
Lgr4/5
Muscle progenitor cell proliferation

3.3 Discussion

Combined with the development of increasingly complex computational methods, scRNA-seq has emerged as a powerful tool to profile the transcriptome of thousands of individual cells in one experiment (Stuart and Satija, 2019). scRNA-seq analysis permits an unbiased survey of cellular complexity and heterogeneity with substantial experimental scope. Here, we leverage these recent developments to build a comprehensive temporal atlas of muscle tissue repair with over 34,000 single-cell transcriptomes, adding to the growing repository of single-cell skeletal muscle datasets.

To date, a handful of muscle-focused scRNA-seq projects have been reported. The first study was performed on FACS-sorted *Pax7*-tdTomato^{+} MuSCs using the Fluidigm C1 system (Cho and Doles, 2017). Though limited to 21 single MuSCs, it provided a transcriptomic view of MuSC heterogeneity, highlighted enriched levels of *Sdc4* transcripts. The study also highlighted some technical limitations, low recovery of *Pax7* transcripts, which motivates the need to generate greater numbers of single-cell transcriptomes to allow more robust statistical analyses. Recently, Giordani et al. presented a transcriptomic atlas of 12,441 muscle-resident cells (Giordani et al., 2019). Their study identified 10 distinct types of cells, including a population of Itga7^{+} Vcam1^{-} smooth muscle and mesenchymal cells that enhance MuSC engraftment when co-transplanted in mice. Their findings illustrated the potential of high-throughput single-cell analysis to reveal poorly described populations and to generate new hypotheses. Whereas the Giordani et al. study focused on homeostatic mouse muscle tissue, another recent scRNA-seq study focused on regenerating muscle. Dell'Orso et al. presented a transcriptomic atlas of about 3,500 FACS-sorted MuSCs and progenitor cells from

homeostatic and notexin-injured muscles (Dell'Orso et al., 2019). They identified two subpopulations of MuSCs with distinct yet overlapping gene expression profiles corresponding to a quiescent and activated state. In addition, they aligned in pseudotime injured, uninjured MuSCs, and primary myoblasts to reveal 7 classes of genes, including of mitochondrial and glycolytic origin, from which they inferred dynamics in metabolic reprogramming activity. Though the first report to describe the dynamics of MuSCs activation and differentiation by scRNA-seq, their analysis was limited to FACS-sorted cells and a single time-point post-injury (60 hours), which might omit the cellular and temporal complexity of muscle regeneration. Finally, using a CyTOF dataset composed of 23 markers, Porpiglia et al. built a trajectory model using the X-shift algorithm of myogenic differentiation post-injury (Porpiglia et al., 2017). They identified two new surface markers, CD9 and CD104, that were used to describe two subpopulations of muscle progenitors, demonstrating how trajectory models from single-cell data can be used to discover new combinations of surface markers for the prospective isolation of MuSCs and their progeny.

Here, we present a unified and annotated single-cell transcriptomic reference atlas of muscle regeneration in adult mice. Our scRNA-seq and CyTOF analyses confirm prior consensus regarding the cell populations involved in the temporal response to muscle injury, and provide a deeper annotation of additional cell types, sub-populations and states with more resolved dynamics, compared to prior scRNA-seq studies (Chapter 2.3) (Cho and Doles, 2017; Porpiglia et al., 2017; Giordani et al., 2019; Dell'Orso et al., 2019). We present comprehensive scRNA-seq dataset compendium describing a total of 34,438 cells (15 different cell types via SNN clustering) including 3,276 from MuSCs and mature

muscle cells (Figs. 9 and 11). The complexity of this transcriptomic atlas powered development of a hierarchical continuum model of myogenic cell populations and ligand-receptor cell communication analysis (Figs. 12-13). We also presented a CyTOF dataset composed of 35 markers to provide an orthogonal validation of myogenic sub-populations and their surface receptor expression variability. Trajectory analysis allowed us to parse the myogenic differentiation lineage post-injury in four distinct groups: quiescent MuSCs, cycling progenitors, committed progenitors, and mature skeletal muscle, with distinct gene expression signatures (Fig. 12). Here we resolved that *Id3* and *Btg2* are both enriched in quiescent MuSCs with their expression decreases following injury-induced MuSC activation and differentiation. Id3 is a DNA binding protein that has been found to be a direct target of Pax7 (Kumar et al., 2009). Id3 is robustly expressed in quiescent MuSCs and blocks differentiation either by directly blocking the activity of pro-myogenic transcription factors such as Myf5, MyoD or by maintaining high levels of Hes1 (Kumar et al., 2009). The role of Btg2 has not yet been previously described in myogenic cells. In hematopoietic and neural lineages, Btg2 promotes differentiation by inhibiting both Id3 and cyclin D1 to restrict cell cycle progression (Yuniati et al., 2019). Though we did not identify the mechanism of Btg2 regulation of myogenic differentiation, these data suggest *Btg2* and *Id3* are distinct transcriptional markers of quiescent MuSCs.

We developed a cell-communication model that allowed us to map interactions between MuSC receptors and ligands expressed by other cell types during muscle repair, which highlighted the complex role of Syndecan receptors in coordinating muscle progenitor heterogeneity. Syndecans (Sdcs) are transmembrane heparan sulfate

proteoglycans that have been characterized as regulators of muscle development, homeostasis, and regeneration (Pisconti et al., 2012). Previous studies have found that *Sdc1* is uniquely expressed in developing muscle, while the other muscle Sdcs (*Sdc2*, *Sdc3*, *Sdc4*) are expressed in MuSCs with differing post-injury expression dynamics specific to each receptor (Pisconti et al., 2012). *Sdc3* plays a role in maintaining MuSC quiescence, as *Sdc3*$^{-/-}$ muscles exhibit a loss in MuSC number and have a homeostasis defect (Pisconti et al., 2016; Pisconti et al., 2010). *Sdc4*, in contrast, seems to play a role in MuSCs activation through regulating FGF and HGF signaling (Cornelison et al., 2004). Both our scRNA-seq and CyTOF data suggest that these four Sdcs are expressed heterogeneously, and in a stage-specific manner, at both the transcript and protein level within quiescent, cycling, and committed myogenic stem/progenitor cells (Fig. 13). *Sdc1*, though previously thought not to be expressed in postnatal muscle, was detected both at the transcription and protein level, by scRNA-seq and CyTOF respectively, in activated cycling muscle progenitors and committed cells but not in MuSCs. Moreover, *Sdc2* was expressed in some quiescent MuSCs and activated progenitors, *Sdc3* expression was restricted to cycling MuSCs and progenitors, and *Sdc4* expressed in quiescent MuSCs and cycling progenitors alike. Together, these Sdcs were involved in a notable fraction of the heterotypic cell communication interactome, suggesting that temporal heterogeneity in Sdc expression may enable myogenic stage-specific fate regulation to a shared set of Sdc-binding ligands.

The mechanistic role of Syndecans in regulating MuSC signal transduction remains poorly understood. Sdc3 cooperates with Notch to promote MuSC cycling and

self-renewal (Pisconti et al., 2010). Sdc4 serves as a co-receptor with Fzd7 to promote Wnt7a-induced MuSC symmetric expansion (Bentzinger et al., 2013). These reports suggest that Syndecans may serve as co-receptor proteins that interplay with numerous ligand-receptor systems involved in myogenic cell fate regulatory pathways. We confirmed the role of Syndecans in mediating ligand-stimulated myogenic cell fates through neutralizing antibody perturbration studies in MuSC cultures treated with ligands suggested by the heterotypic communication model (Fig. 15). Sdc1 and Sdc4 were both necessary for FGF2/Fn-induced MuSC proliferation (Figs. 15F-H). Sdc2 (but not Sdc4) was necessary for both TGFβ1-augmented and RSPO3-suppressed MuSC proliferation (Figs. 15H-K). These findings demonstrated the utility of a cell communication-factor based analysis of muscle regeneration in identifying new heterotypic ligand-receptor interactions involved in MuSC fate regulation. Further, these observations argue that Syndecans serve as diverse co-receptor proteins to potentiate ligand-receptor stimulated muscle stem/progenitor cell fates and individual Syndecan family members may mediate both quiescence and self-renewal/proliferation signaling (Fig. 15L).

Though clustering and visualization approaches presented here and by others can be used to identify new cell types and biomarkers, we urge caution on how cell types are defined from these data. First, single-cell data is not immune to technical imperfections such as doublets or sequencing reads that misalign (Stuart and Satija, 2019). Second, single-cell data is sensitive to sample preparation methods, especially from the solid tissue sources. For example, we observed significant variability in the inclusion of tenocytes based on muscle-tendon resection precision (data not shown). We also found

other biases introduced in the scRNA-seq composition introduced by FACS sorting can select for metabolically active immune cells. Furthermore, since scRNA-seq is based on RNA detection, some quiescent cell types such as MuSCs are challenging to unambiguously detect and will be benefited from *in vivo* transcript recovery advances (Machado et al., 2017; van Velthoven et al., 2017). Even with these limitations, our scRNA-seq study provides a view that discards some prior assumptions on the boundaries between myogenic cell types and suggests a myogenic continuum that is endowed with a plasticity of cell-cycle and commitment states. This annotated temporal scRNA-seq atlas of muscle regeneration may provide reference resource to examine the role of cellular diversity and communication in aging, disease, and across species.

3.4 Supplementary Material

A

Sample name	Average reads per cell	Median genes per cell	Median UMIs per cell	Total genes detected	Number of cells after QC filtering
D0_A	95,810	1,262	2,690	17,613	2,139
D0_B	66,817	1,186	2,555	17,963	3,440
D0_C	26,987	342	1,854	18,825	1,449
D2_A	39,157	2,270	7,972	18,048	3,654
D2_B	129,868	2,597	10,366	17,361	1,870
D5_A	86,761	2,960	11,800	19,564	4,146
D5_B	46,859	2,534	9,004	19,365	4,665
D5_C	39,322	2,577	9,148	19,584	5,429
D7_A	45,165	1,694	4,185	18,813	4,303
D7_B	73,277	2,281	7,405	19,116	3,343
				19,584 different genes detected	34,438 cells analysed

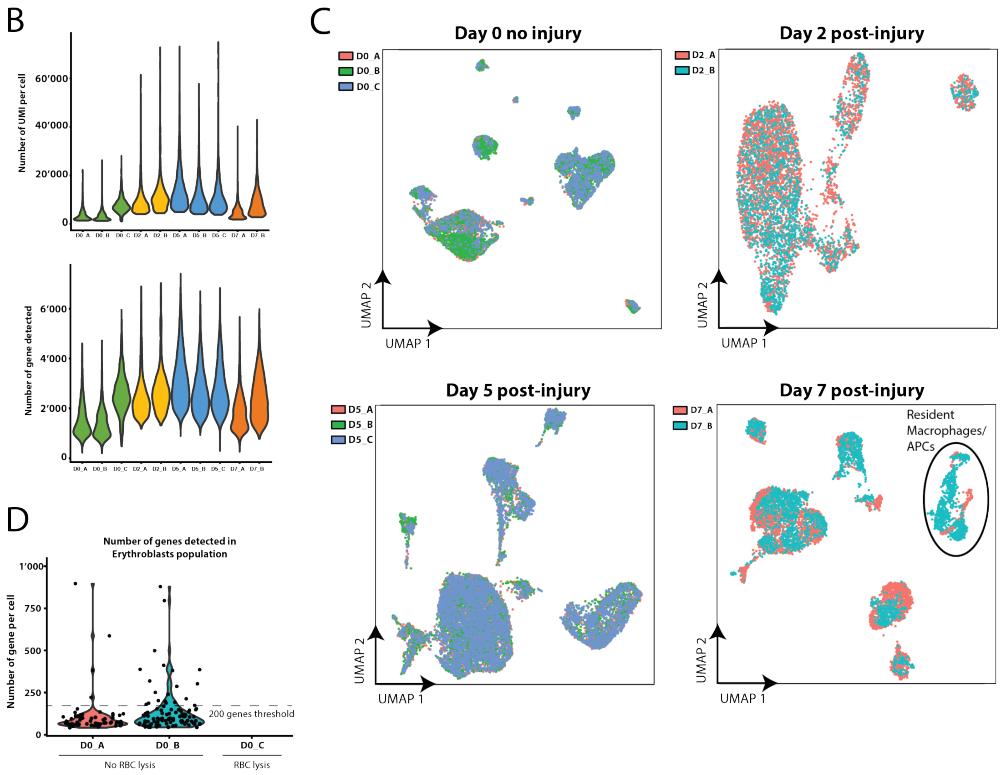


Figure S1 – Technical and quality control measures for scRNA-seq datasets. (**A**) Sequencing reads processing statistics from the 10X Cell Ranger pipeline. The samples originated from 9 individual female mice (4-6 months of age). Samples are color-coded by time-point. Differences in sample quality were addressed by applying a quality control (QC) filter in Seurat. The last column represents the number of cells that were used for the analysis after QC. A total number of 34,438 cells expressing 19,584 different genes were used for the downstream analysis. (**B**) Number of unique molecular identifiers (UMIs) and genes per cell per sample after QC filtering. On average cells have more UMIs and genes detected for injured samples (day 2 and 5 post-injury). The samples are color-coded as in (A). (**C**) Technical effect and differences in population number across samples from the same timepoint. Single-cell transcriptomes were projected on a UMAP plot using cross-correlation analysis (CCA) scores between samples. Some divergences in cell number can be observed for the immune cluster between D7_A and D7_B. (**D**) Distribution of the number of genes per Erythroblast cell type with or without red blood cell (RBC) lysis during the single-cell suspension preparation. Erythroblasts express a small number of genes and were removed from the analysis after applying a 200-gene minimum filter. RBC lysis also removed all detectable Erythroblasts. RBC lysis was applied to all samples in this study expect D0_A and D0_B.

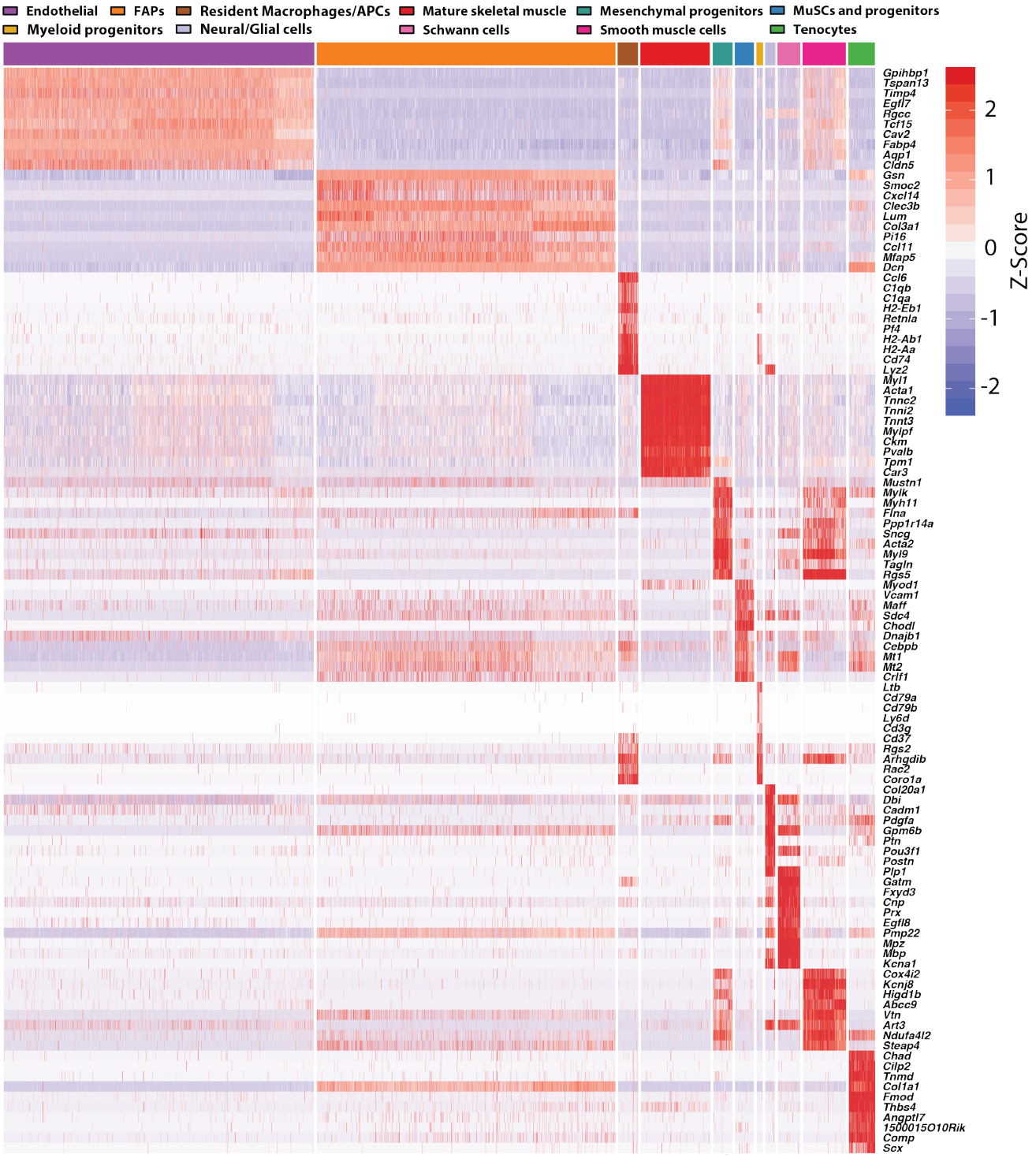


Figure S2 – Top differentially expressed genes per cluster in the uninjured (day 0) muscle. Normalized expression (Z-score) heatmap for top differentially expressed genes in the 11 subpopulations identified by SNN at the uninjured (day 0) time-point. The columns represent cells and are organized by cell type as color-coded.

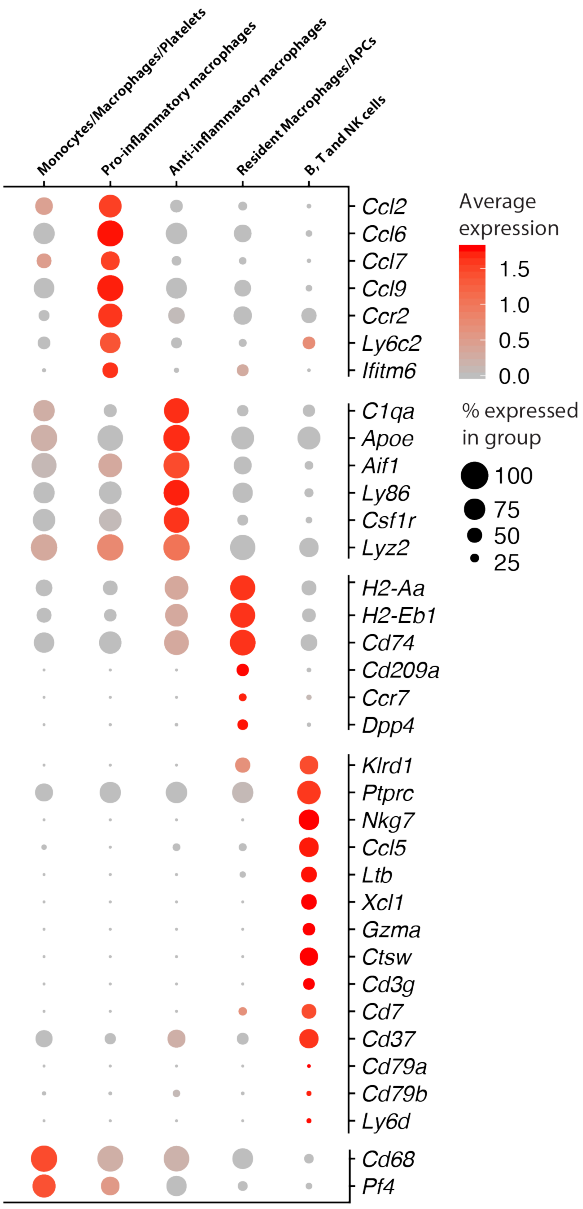


Figure S3 – Expression of immune markers identified by differential expression analysis. Average expression of immune genes identified by differential expression analysis in the combined dataset (Fig. 9B). These markers allowed delineation of 5 different immune cell populations pre and post-injury. The dot size represents the percentage of cells within a group with an expression level > 0.

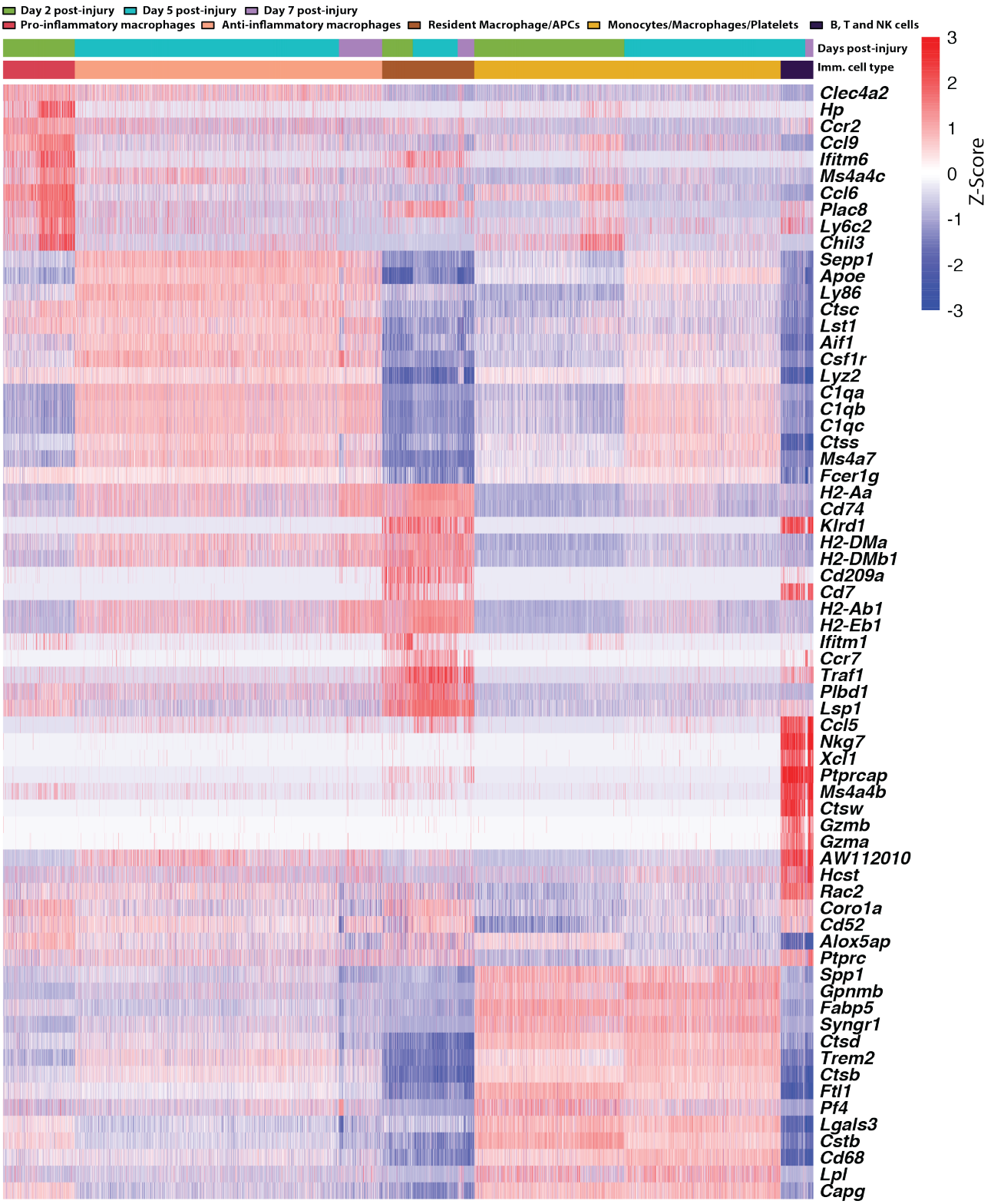
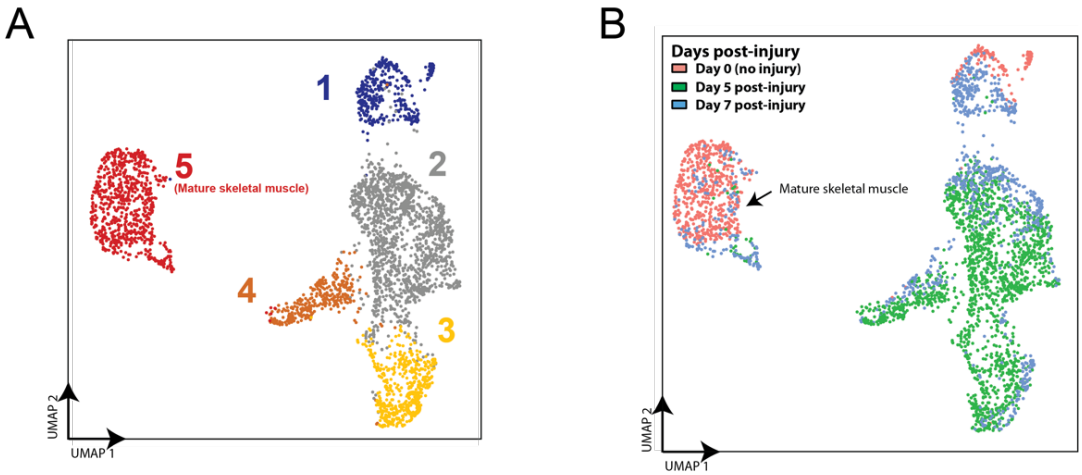
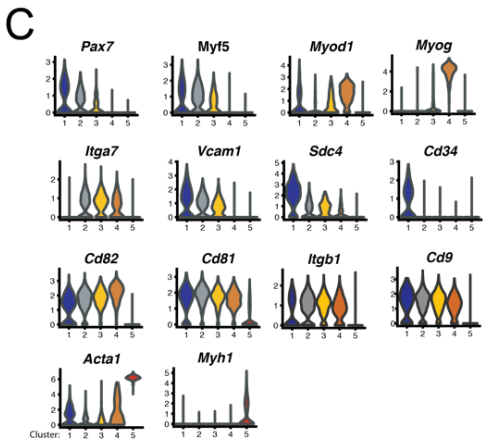
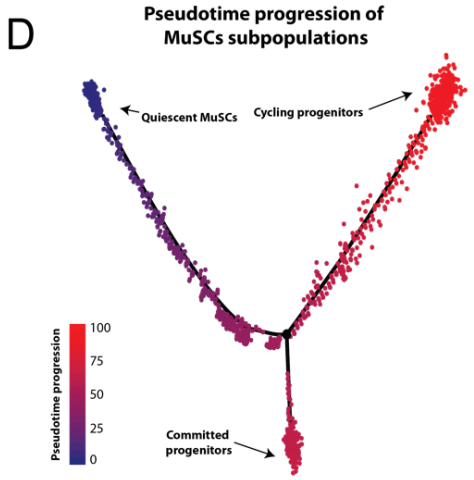


Figure S4 – Top differentially expressed genes in immune subpopulations across days post-injury. Normalized expression (Z-Score) heatmap for top differentially expressed genes in 5 subpopulations of immune (imm.) cells identified by SNN post-injury. The columns represent cells and are organized first by cell type and then by day post-injury as color-coded.


A
1
2
3
4
5
(Mature skeletal muscle)
UMAP 2
UMAP 1
B
Days post-injury
Day 0 (no injury)
Day 5 post-injury
Day 7 post-injury
Mature skeletal muscle
UMAP 2
UMAP 1


C
Pax7
Myf5
Myod1
Myog
Itga7
Vcam1
Sdc4
Cd34
Cd82
Cd81
Itgb1
Cd9
Acta1
Myh1
Cluster: 1 2 3 4 5


D
Pseudotime progression of
MuSCs subpopulations
Quiescent MuSCs
Cycling progenitors
Committed
progenitors
Pseudotime progression
100
75
50
25
0

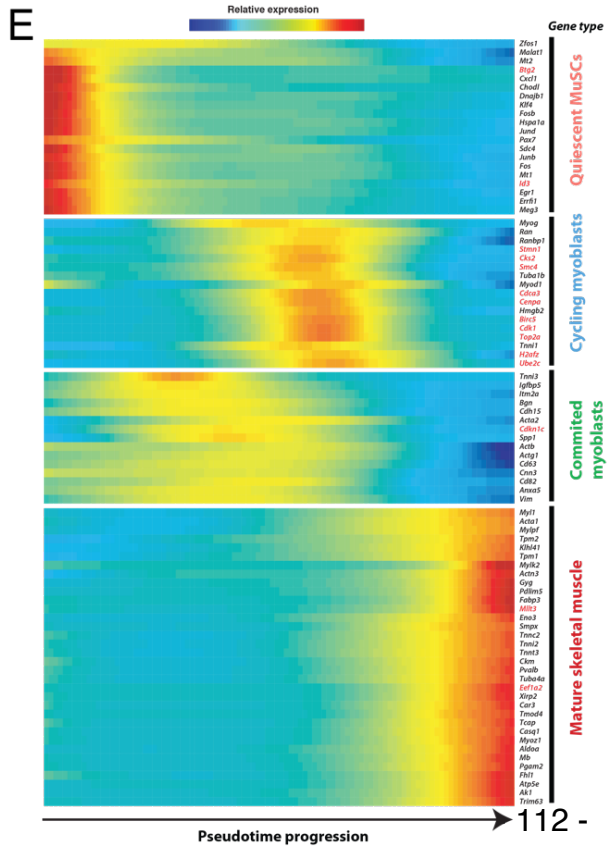

E
Relative expression
Gene type
Quiescent MuSCs
Cycling myoblasts
Commited
myoblasts
Mature skeletal muscle
Pseudotime progression

Figure S5 (*previous page*) **– MuSCs heterogeneity in gene expression pre- and post-injury.** (**A**) UMAP projection of MuSC and mature skeletal muscle single-cell transcriptomes (3,276 total cells from day 0, 5, 7) labeled by unsupervised SNN clustering. Clustering reveals 5 subpopulations of cells, including mature skeletal muscle cells ("5") and MuSCs ("1"-"4") that are composed of cells from different days post-injury. (**B**) UMAP projection of cells in (A) labeled by day post-injury. (**C**) Violin plots representing the expression level of key muscle genes in the subpopulations identified by SNN clustering (A). (**D**) Single-cell feature plot visualization of pseudotime progression model of MuSC activation and commitment. (**E**) Top 75 genes differentially expressed in the three branches of the MuSC trajectory identified by Monocle (Fig. 12B) and ordered by branch and then pseudotime progression.

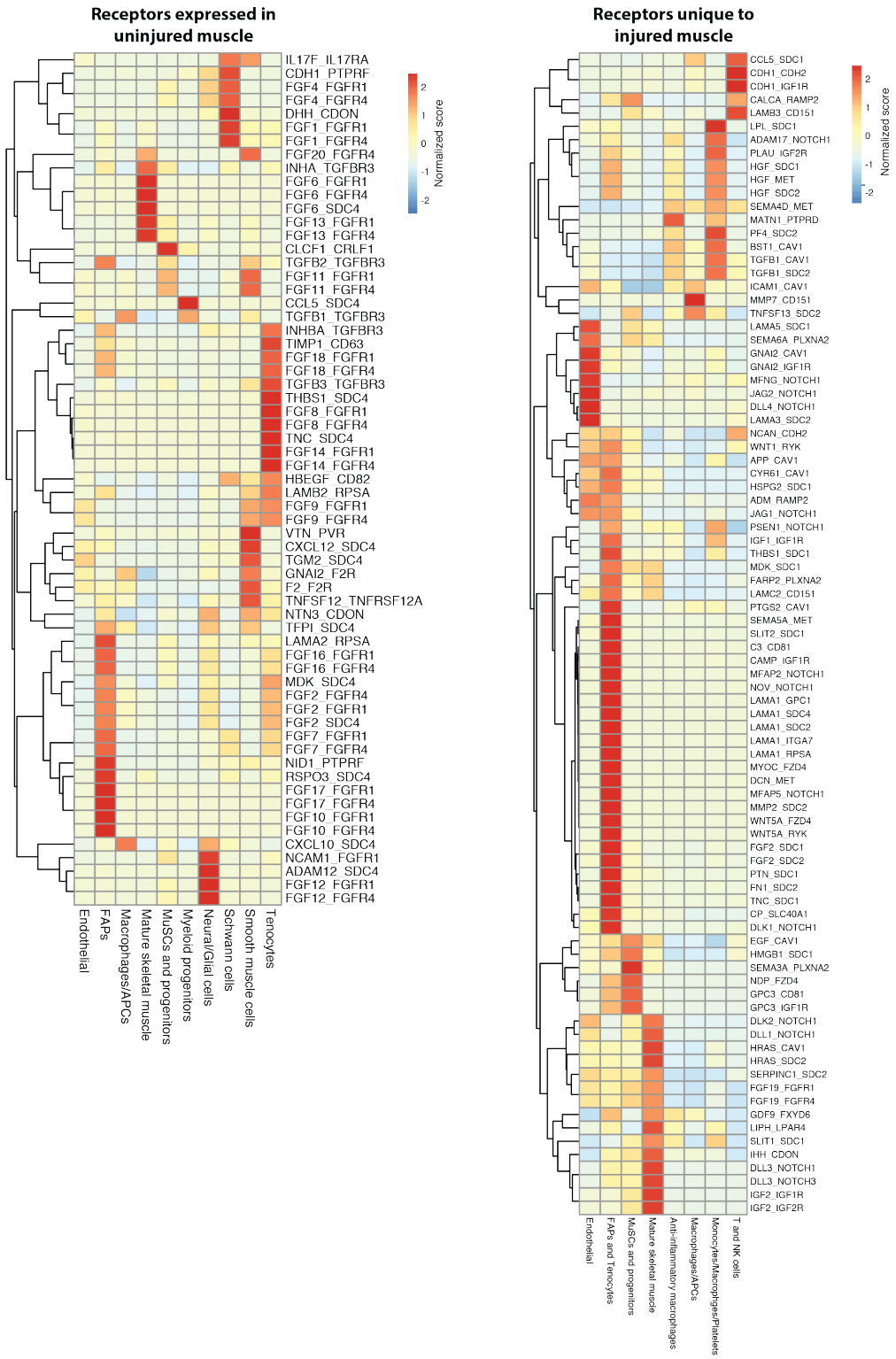


Figure S6 – Ligand-receptor interaction score heatmap. The heatmap represents score between receptors differentially expressed in MuSCs and progenitor cells and ligands on other cell types. Rows represent the ligand-receptor pair in the format LIGAND_RECEPTOR. Columns represent the cell type expressing the ligand and includes MuSCs and progenitors for autocrine interactions. The score for each pair has been normalized across ligand cell types with a positive value indicating that the pair has a high score for a particular ligand and cell type compared to other cell types.

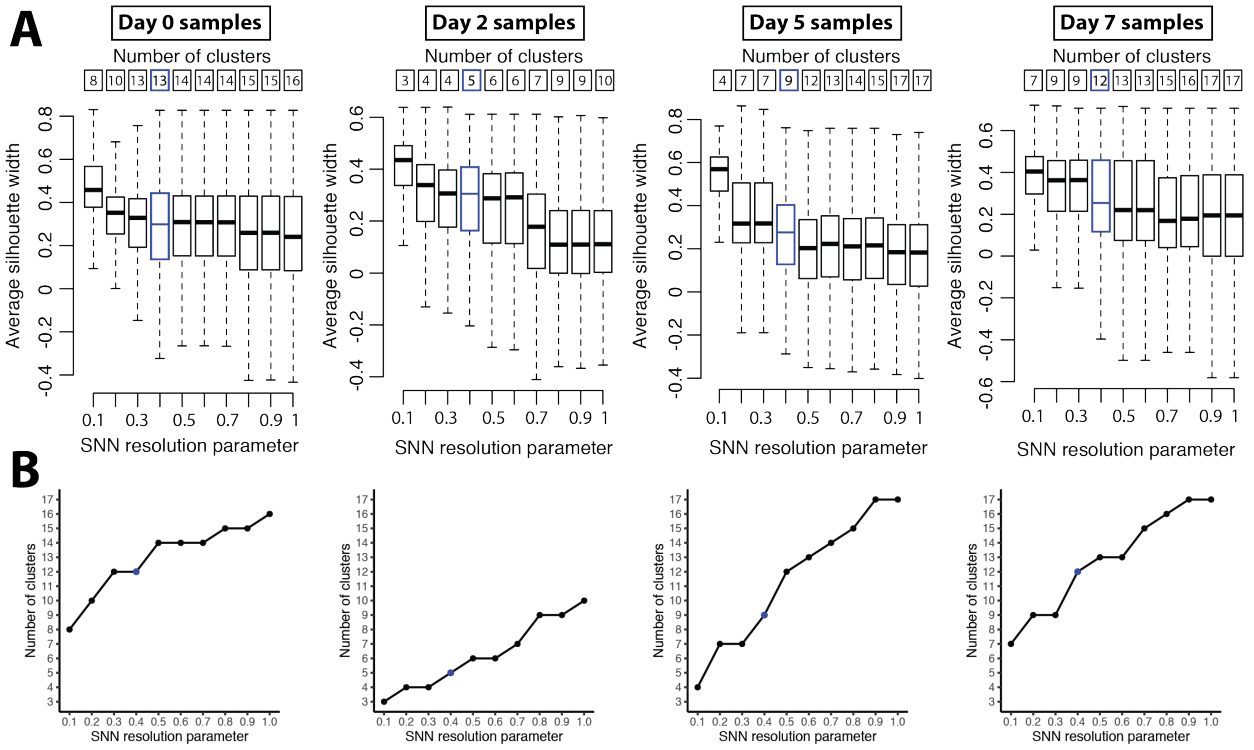


Figure S7 – Silhouette analysis of SNN clustering. (A) Boxplots per day samples representing the average silhouette width as a function of SNN resolution parameters. The blue box plot corresponds to the SNN parameter value chosen for cluster annotation and downstream analysis. **(B)** Number of clusters identified as a function of the resolution parameter.

3.5 Experimental section‡

Mice and single-cell isolation. The Cornell University Institutional Animal Care and Use Committee (IACUC) approved all animal protocols and experiments were performed in compliance with its institutional guidelines. Adult C57BL/6J mice were obtained from Jackson Laboratories (#000664; Bar Harbor, ME) and were used at 4-6 months of age. To induce muscle injury, one or both tibialis anterior muscles received a single 10 µL intramuscular injection of notexin (10 µg/ml; Latoxan, France). For scRNA-seq and CyTOF analyses, one tibialis anterior muscle per mouse was collected to generate single-cell suspensions at either 0, 2, 5, or 7 days post-notexin injury. For cell culture studies, both quadriceps, tibialis anterior, and gastrocnemius muscles were collected without prior injury. Muscles were digested with 2.5 mg/ml Collagenase D (Roche) and 0.04 U/ml Dispase II (Roche) followed by manual dissociation, and filtration to generate a single-cell suspension. For some samples, red blood cell lysis was performed after filtration.

MuSC isolation by FACS. For some scRNA-seq studies (see Fig. 8) we enriched the single-cell suspension for Calcein-AM+ propidium iodide– (PI) viable cells using a BD FACS Aria III instrument (BD Biosciences, San Jose, CA). The single-cell suspensions were stained with Calcein-AM (1 µg per 100,000 cells) and PI before sorting. Live cells were sorted selecting for Calcein-AM+ and debris were removed using a standard FSC/SSC gating strategy. For cell culture studies, MuSCs were prospectively isolated from single-cell suspensions following magnetic depletion, as previously described

‡Complete information about reagents and resource is available in the published manuscript (De Micheli et al., 2020) and in Appendix 3.

(Cosgrove et al., 2014; Sacco et al., 2008). We isolated MuSCs using a (PI/CD45/CD11b/CD31/Sca1)– CD34+ α7-integrin+ cell sorting gate. See Reagents and Resources for specific antibody clones and fluorophores.

Muscle stem cell culture and ligand stimulation. All culture experiments were performed in 96-well, black-walled, tissue-culture treated plates pre-coated with Laminin, Fibronectin, or Tenascin-C. FACS-isolated MuSCs were seeded at ~300 cells/well in primary myoblast growth medium with or without recombinant FGF2. The following day and then every three days afterward, GM with or without FGF2 was changed and supplemented with either recombinant paracrine ligand proteins and/or Syndecan neutralizing antibodies. See Reagents and Resources for product information and concentrations.

Luminescence imaging analysis. For luminescence imaging, cells were incubated with fresh GM containing 150 ug/mL D-luciferin (Goldbio #LUCNA1G) for 10 min at 37°C. Following D-luciferin incubation, bioluminescence was measured using the Xenogen IVIS-200 Imaging system (Perkin Elmer). Luminescence was quantified in each culture well to determine the relative abundance of Luciferase-expressing muscle cells using the Living Image Software (Perkin Elmer) to perform per-well ROI segmentation. After imaging, cell medium was replaced with GM containing ligands and/or neutralizing antibodies.

Immunofluorescence imaging and cell counting. Cultured cells were fixed with formaldehyde and permeabilized with Triton X-100 and then were blocked with 15% donkey serum in 0.01% Triton X-100. Cells were stained with anti-Myod1 (Santa Cruz # sc-377460) and then an AlexaFluor488-conjugated secondary antibody (Thermo Fisher # A32766). DAPI was used for nuclear counter-staining. Fluorescence imaging was performed with a Nikon Eclipse-Ti with an 10x/0.25NA objective and SpectraX light engine source (Lumencor). Nuclei were counted using ImageJ for segmentation with application of the BaSiC plug-in for background removal.

Single-cell RNA-sequencing. After digestion, single-cell suspensions were washed and resuspended in 0.04% BSA in PBS at a concentration of at least 400 cells/µL. Cells were counted manually with a hemocytometer to determine their concentration. Single-cell RNA-sequencing libraries were then prepared using the Chromium Single Cell 3' reagent kit v2 (10X Genomics) in accordance with the manufacturer's protocol. Briefly, the cells were diluted into the Chromium Single Cell A Chip as to yield a recovery of ~6,000 single-cell transcriptomes with <5% doublet rate. Following the library preparation, the libraries were sequenced in multiplex (n=2 per sequencing run) on the NextSeq 500 sequencer (Illumina) to produce between 200 and 250 million reads per library and on average a minimum of 30,000 reads per single cell.

Analysis of scRNA-seq data. Sequencing reads were processed with the Cell Ranger version 3.0.1 (10X Genomics) using the mouse reference transcriptome mm10. From the

gene expression matrix, the downstream analysis was carried out with R version 3.5.1 (2018-07-02). Quality control, filtering, data clustering and visualization, and the differential expression analysis was carried out using Seurat version 2.3.4 R package (Butler et al., 2018) with some custom modifications to the standard pipeline. Each of the 9 datasets were first analyzed independently before combining datasets from the same timepoint together for an integrated analysis. For each individual dataset, genes expressed in less that 3 cells as well as cells <1000 UMIs and <200 genes were removed from the gene expression matrix. In addition, we removed any single-cell with >20% UMIs mapped to mitochondrial genes, as well as obvious outliers in number of UMIs (cell doublets). After log-normalizing the data, the expression of each gene was scaled regressing out the number of UMI and the percent mitochondrial gene expressed in each cell. We performed PCA on the gene expression matrix and used the first 15 principal components for clustering and visualization. Unsupervised shared nearest neighbor (SNN) clustering was performed with a resolution of 0.4 and visualization was done using uniform manifold approximation and projection (UMAP) (Becht et al., 2018). We performed a silhouette analysis (R cluster package) on combined day samples (Fig. S7) to choose the most optimal SNN resolution parameter, balancing the number of expected clusters (given known marker expression) with a maximal average silhouette width. Finally, differential expression analysis was achieved using Seurat's "FindAllMarkers" function using a likelihood ratio test that assumes the data follows a negative binomial distribution and only considering genes with >log2(0.25) fold-change and expressed in at least 25% of cells in the cluster.

To analyze datasets merged from the same timepoint, we performed canonical correlation analysis (CCA) between datasets followed by data alignment using nonlinear dynamic time warping. We used the CCA matrix instead of PCA for visualization and clustering, and consequently, the differential expression analysis between clustered cells. The CCA space allows us to highlight shared patterns in gene expression profile between datasets. In addition, we did not observe significant differences in results between working in the CCA or PCA space, thus indicating little batch effect between samples from the same timepoint. We finally combined the 9 datasets into a uniform UMAP atlas using the CCA scores calculated from all the datasets. However, we retain the labels identified independently at each timepoint and the data was not re-clustered.

Monocle trajectory analysis. We used the Monocle version 2.8.0 R package (Qiu et al., 2017) to organize cells in pseudotime and infer new trajectories of MuSCs subpopulations post-injury. First, we subsetted the cells labeled as MuSCs and progenitor cells and mature skeletal muscle cells from the Seurat dataset and across all timepoints and samples. Second, we performed unsupervised SNN clustering in order to identify new subpopulations in the data, from which we then used the Seurat “FindAllClusters” function (as described above) to find differentially expressed genes that characterize the subpopulations. We then selected the top 150 genes based on fold-change expression with a minimum of log2(0.8) and adjusted p-value of 0.01. This list genes of differentially expressed genes is then used by Monocle for clustering and ordering cells using the DDRTree method and reverse graph embedding. To identify genes that are differentially

expressed across Monocle branches (States), we transferred the labels back to the Seurat dataset and performed differential expression analysis as described above.

Ligand-receptor cell communication model. The model aims at scoring potential ligand-receptor interactions between MuSCs (receptor) and other cell types (ligand). We used the ligand-receptor interaction database from Skelly et al. (Skelly et al., 2018). To calculate the score for a given ligand-receptor pair, we multiply the average receptor expression in MuSCs with the average ligand expression per other cell type (including MuSCs to consider for autocrine interactions). We only considered receptors that are differentially expressed in MuSCs at any given time-point post-injury.

CyTOF sample preparation and staining. Most of the cell surface markers were obtained from Fluidigm and few others were conjugated in-house using the Maxpar X8 Multimetal Labeling Kit (see Table S1 for complete antibody list). The cells were stained with 50 µl of the antibody cocktail (100 µl of total staining volume) for 30 minutes at RT with intermittent vortexing. Following 2 washes (300g, 5 mins, RT), the cells were fixed with freshly prepared 1.6% PFA for 10 minutes. Thereafter, the cells were incubated for 30 minutes in 1 mL of the Nuclear Antigen Staining Buffer working solution. Then the cells were washed with 2 ml of Nuclear Antigen Staining Perm at 800g, 5 minutes, RT. 50 µl of nuclear antigen antibody cocktail was added to 50 µl of cell pellet solution and incubated for 45 minutes at RT. Following antibody staining, the cells were also stained with Cell-ID Intercalator-Ir-125 µM diluted to 1:1000 with MaxPar Fix and Perm buffer for

1 hour at RT. In addition, cells were also stained with 10 µl of Cisplatin for viability in 1 mL of pre-warmed serum free medium for 5 minutes at RT. Finally, cells were also stained with IdU (5-Iodo-2-deoxyuridine) to label the S-phase at a concentration of 50 µM for 30 minutes at 37oC.

CyTOF data acquisition and analysis. Cells were washed twice with staining buffer and then with MilliQ water. After the final wash, the cells were adjusted to a concentration of 106 cells/mL with 1:10 EQ beads to MilliQ water solution. Prior to the acquisition, the instrument was tuned and calibrated using the EQ standard beads. The acquisition speed of the sample was maintained within 400 events/seconds to avoid doublets and ion cloud fusion errors in the data. The output FCS files were normalized using the Fluidigm normalizer algorithm that is embedded within the CyTOF software (Version 6.7.1014). The CyTOF data was first gated using the Cytobank software in order to exclude debris, dead cells and doublets (Kotecha et al., 2010). The resulting FCS file was then converted into a gene expression matrix using the Cytofkit R package using an inverse hyperbolic sine transformation (Chen et al., 2016). The gene expression matrix was then analyzed with a Seurat-based custom pipeline, which allowed for SNN clustering and UMAP visualization.

Data and code availability

The mouse single-cell RNA-sequencing data are available at the GEO repository under accession numbers GSE143435 (FACS-sorted samples) and GSE143437 (non-FACS

sorted samples). The code and documentation for the ligand-receptor cell communication model is available online on GitHub: https://github.com/andrea-de-micheli/L-R_interaction_model_for_scRNAseq.

3.6 Chapter references

Baitsch, D., Bock, H.H., Engel, T., Telgmann, R., Muller-Tidow, C., Varga, G., Bot, M., Herz, J., Robenek, H., von Eckardstein, A., et al. (2011). Apolipoprotein E induces antiinflammatory phenotype in macrophages. Arterioscler Thromb Vasc Biol 31, 1160-1168.

Becht, E., McInnes, L., Healy, J., Dutertre, C.A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol.

Biressi, S., and Rando, T.A. (2010). Heterogeneity in the muscle satellite cell population. Semin Cell Dev Biol 21, 845-854.

Blau, H.M., Cosgrove, B.D., and Ho, A.T. (2015). The central role of muscle stem cells in regenerative failure with aging. Nat Med 21, 854-862.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 36, 411-420.

Chakkalakal, J.V., Christensen, J., Xiang, W., Tierney, M.T., Boscolo, F.S., Sacco, A., and Brack, A.S. (2014). Early forming label-retaining muscle stem cells require p27kip1 for maintenance of the primitive state. Development 141, 1649-1659.

Chazaud, B. (2016). Inflammation during skeletal muscle regeneration and tissue remodeling: application to exercise-induced muscle damage management. Immunol Cell Biol 94, 140-145.

Chen, H., Lau, M.C., Wong, M.T., Newell, E.W., Poidinger, M., and Chen, J. (2016). Cytofkit: A Bioconductor Package for an Integrated Mass Cytometry Data Analysis Pipeline. PLoS Comput Biol 12, e1005112.

Cho, D.S., and Doles, J.D. (2017). Single cell transcriptome analysis of muscle satellite cells reveals widespread transcriptional heterogeneity. Gene 636, 54-63.

Christov, C., Chretien, F., Abou-Khalil, R., Bassez, G., Vallet, G., Authier, F.J., Bassaglia, Y., Shinin, V., Tajbakhsh, S., Chazaud, B., et al. (2007). Muscle satellite cells and endothelial cells: close neighbors and privileged partners. Mol Biol Cell 18, 1397-1409.

Cornelison, D.D., Wilcox-Adelman, S.A., Goetinck, P.F., Rauvala, H., Rapraeger, A.C., and Olwin, B.B. (2004). Essential and separable roles for Syndecan-3 and Syndecan-4 in skeletal muscle development and regeneration. Genes Dev 18, 2231-2236.

Cornelison, D.D., and Wold, B.J. (1997). Single-cell analysis of regulatory gene expression in quiescent and activated mouse skeletal muscle satellite cells. Dev Biol 191, 270-283.

Cosgrove, B.D., Gilbert, P.M., Porpiglia, E., Mourkioti, F., Lee, S.P., Corbel, S.Y., Llewellyn, M.E., Delp, S.L., and Blau, H.M. (2014). Rejuvenation of the muscle stem cell population restores strength to injured aged muscles. Nat Med 20, 255-264.

De Micheli, A.J., Laurilliard, E.J., Heinke, C.L., Ravichandran, H., Fraczek, P., Soueid-Baumgarten, S., De Vlaminck, I., Elemento, O., and Cosgrove, B.D. (2020). Single-Cell Analysis of the Muscle Stem Cell Hierarchy Identifies Heterotypic Communication Signals Involved in Skeletal Muscle Regeneration. Cell Reports 30, 3583-3595.e5.

Dell'Orso, S., Juan, A.H., Ko, K.D., Naz, F., Perovanovic, J., Gutierrez-Cruz, G., Feng, X., and Sartorelli, V. (2019). Single cell analysis of adult mouse skeletal muscle stem cells in homeostatic and regenerative conditions. Development 146.

Docheva, D., Hunziker, E.B., Fassler, R., and Brandau, O. (2005). Tenomodulin is necessary for tenocyte proliferation and tendon maturation. Mol Cell Biol 25, 699-705.

Farioli-Vecchioli, S., Saraulli, D., Costanzi, M., Leonardi, L., Cina, I., Micheli, L., Nutini, M., Longone, P., Oh, S.P., Cestari, V., et al. (2009). Impaired terminal differentiation of hippocampal granule neurons and defective contextual memory in PC3/Tis21 knockout mice. PloS one 4, e8339.

Gaylinn, B.D., Eddinger, T.J., Martino, P.A., Monical, P.L., Hunt, D.F., and Murphy, R.A. (1989). Expression of nonmuscle myosin heavy and light chains in smooth muscle. Am J Physiol 257, C997-1004.

Giordani, L., He, G.J., Negroni, E., Sakai, H., Law, J.Y.C., Siu, M.M., Wan, R., Corneau, A., Tajbakhsh, S., Cheung, T.H., et al. (2019). High-Dimensional Single-Cell Cartography Reveals Novel Skeletal Muscle-Resident Cell Populations. Mol Cell 74, 609-621 e606.

Heredia, J.E., Mukundan, L., Chen, F.M., Mueller, A.A., Deo, R.C., Locksley, R.M., Rando, T.A., and Chawla, A. (2013). Type 2 innate signals stimulate fibro/adipogenic progenitors to facilitate muscle regeneration. Cell 153, 376-388.

Ho, M.M., Manughian-Peter, A., Spivia, W.R., Taylor, A., and Fraser, D.A. (2016). Macrophage molecular signaling and inflammatory responses during ingestion of atherogenic lipoproteins are modulated by complement protein C1q. Atherosclerosis 253, 38-46.

Hwang, B., Lee, J.H., and Bang, D. (2018). Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp Mol Med 50, 96.

Janssen, E., Dzeja, P.P., Oerlemans, F., Simonetti, A.W., Heerschap, A., de Haan, A., Rush, P.S., Terjung, R.R., Wieringa, B., and Terzic, A. (2000). Adenylate kinase 1 gene deletion disrupts muscle energetic economy despite metabolic rearrangement. EMBO J 19, 6371-6381.

Joe, A.W., Yi, L., Natarajan, A., Le Grand, F., So, L., Wang, J., Rudnicki, M.A., and Rossi, F.M. (2010). Muscle injury activates resident fibro/adipogenic progenitors that facilitate myogenesis. Nat Cell Biol 12, 153-163.

Kotecha, N., Krutzik, P.O., and Irish, J.M. (2010). Web-based analysis and publication of flow cytometry experiments. Curr Protoc Cytom Chapter 10, Unit10 17.

Kuang, S., Kuroda, K., Le Grand, F., and Rudnicki, M.A. (2007). Asymmetric self-renewal and commitment of satellite stem cells in muscle. Cell 129, 999-1010.

Kumar, D., Shadrach, J.L., Wagers, A.J., and Lassar, A.B. (2009). Id3 is a direct transcriptional target of Pax7 in quiescent satellite cells. Mol Biol Cell 20, 3170-3177.

Kuschel, R., Deininger, M.H., Meyermann, R., Bornemann, A., Yablonka-Reuveni, Z., and Schluesener, H.J. (2000). Allograft inflammatory factor-1 is expressed by macrophages in injured skeletal muscle and abrogates proliferation and differentiation of satellite cells. J Neuropathol Exp Neurol 59, 323-332.

Liu, Z., Jin, Y.Q., Chen, L., Wang, Y., Yang, X., Cheng, J., Wu, W., Qi, Z., and Shen, Z. (2015). Specific marker expression and cell state of Schwann cells during culture in vitro. PloS one 10, e0123278.

Lukjanenko, L., Jung, M.J., Hegde, N., Perruisseau-Carrier, C., Migliavacca, E., Rozo, M., Karaz, S., Jacot, G., Schmidt, M., Li, L., et al. (2016). Loss of fibronectin from the aged stem cell niche affects the regenerative capacity of skeletal muscle in mice. Nat Med 22, 897-905.

Lyons, G.E., Ontell, M., Cox, R., Sassoon, D., and Buckingham, M. (1990). The expression of myosin genes in developing skeletal muscle in the mouse embryo. J Cell Biol 111, 1465-1476.

Machado, L., Esteves de Lima, J., Fabre, O., Proux, C., Legendre, R., Szegedi, A., Varet, H., Ingerslev, L.R., Barres, R., Relaix, F., et al. (2017). In Situ Fixation Redefines Quiescence and Early Activation of Skeletal Muscle Stem Cells. Cell reports 21, 1982-1993.

Motohashi, N., and Asakura, A. (2014). Muscle satellite cell heterogeneity and self-renewal. Frontiers in cell and developmental biology 2.

Novak, M.L., and Koh, T.J. (2013). Macrophage phenotypes during tissue repair. J Leukoc Biol 93, 875-881.

Pawlikowski, B., Orion Vogler, T., Gadek, K., and Olwin, B. (2017). Regulation of Skeletal Muscle Stem Cells by Fibroblast Growth Factors. Dev Dyn.

Pisconti, A., Banks, G.B., Babaeijandaghi, F., Betta, N.D., Rossi, F.M., Chamberlain, J.S., and Olwin, B.B. (2016). Loss of niche-satellite cell interactions in syndecan-3 null mice alters muscle progenitor cell homeostasis improving muscle regeneration. Skelet Muscle 6, 34.

Pisconti, A., Bernet, J.D., and Olwin, B.B. (2012). Syndecans in skeletal muscle development, regeneration and homeostasis. Muscles Ligaments Tendons J 2, 1-9.

Pisconti, A., Cornelison, D.D., Olguin, H.C., Antwine, T.L., and Olwin, B.B. (2010). Syndecan-3 and Notch cooperate in regulating adult myogenesis. J Cell Biol 190, 427-441.

Porpiglia, E., Samusik, N., Ho, A.T., Cosgrove, B.D., Mai, T., Davis, K.L., Jager, A., Nolan, G.P., Bendall, S.C., Fantl, W.J., et al. (2017). High-resolution myogenic lineage mapping by single-cell mass cytometry. Nat Cell Biol 19, 558-567.

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nature methods 14, 979-982.

Ren, H., Yin, P., and Duan, C. (2008). IGFBP-5 regulates muscle cell differentiation by binding to IGF-II and switching on the IGF-II auto-regulation loop. J Cell Biol 182, 979-991.

Rocheteau, P., Gayraud-Morel, B., Siegl-Cachedenier, I., Blasco, M.A., and Tajbakhsh, S. (2012). A subpopulation of adult skeletal muscle stem cells retains all template DNA strands after cell division. Cell 148, 112-125.

Sacco, A., Doyonnas, R., Kraft, P., Vitorovic, S., and Blau, H.M. (2008). Self-renewal and expansion of single transplanted muscle stem cells. Nature 456, 502-506.

Schweitzer, R., Chyung, J.H., Murtaugh, L.C., Brent, A.E., Rosen, V., Olson, E.N., Lassar, A., and Tabin, C.J. (2001). Analysis of the tendon cell fate using Scleraxis, a specific marker for tendons and ligaments. Development 128, 3855-3866.

Skelly, D.A., Squiers, G.T., McLellan, M.A., Bolisetty, M.T., Robson, P., Rosenthal, N.A., and Pinto, A.R. (2018). Single-Cell Transcriptional Profiling Reveals Cellular Diversity and Intercommunication in the Mouse Heart. Cell reports 22, 600-610.

Sousa-Victor, P., Gutarra, S., Garcia-Prat, L., Rodriguez-Ubreva, J., Ortet, L., Ruiz-Bonilla, V., Jardi, M., Ballestar, E., Gonzalez, S., Serrano, A.L., et al. (2014). Geriatric muscle stem cells switch reversible quiescence into senescence. Nature 506, 316-321.
Stuart, T., and Satija, R. (2019). Integrative single-cell analysis. Nat Rev Genet 20, 257-272.

Tidball, J.G. (2017). Regulation of muscle growth and regeneration by the immune system. Nat Rev Immunol 17, 165-178.

Tierney, M.T., and Sacco, A. (2016). Satellite Cell Heterogeneity in Skeletal Muscle Homeostasis. Trends Cell Biol 26, 434-444.

Tierney, M.T., Stec, M.J., Rulands, S., Simons, B.D., and Sacco, A. (2018). Muscle Stem Cells Exhibit Distinct Clonal Dynamics in Response to Tissue Repair and Homeostatic Aging. Cell Stem Cell 22, 119-127 e113.

Uezumi, A., Ito, T., Morikawa, D., Shimizu, N., Yoneda, T., Segawa, M., Yamaguchi, M., Ogawa, R., Matev, M.M., Miyagoe-Suzuki, Y., et al. (2011). Fibrosis and adipogenesis originate from a common mesenchymal progenitor in skeletal muscle. J Cell Sci 124, 3654-3664.

van Velthoven, C.T.J., de Morree, A., Egner, I.M., Brett, J.O., and Rando, T.A. (2017). Transcriptional Profiling of Quiescent Muscle Stem Cells In Vivo. Cell reports 21, 1994-2004.

Wagner, A., Regev, A., and Yosef, N. (2016). Revealing the vectors of cellular identity with single-cell genomics. Nat Biotechnol 34, 1145-1160.

Wang, Y.X., and Rudnicki, M.A. (2011). Satellite cells, the engines of muscle repair. Nat Rev Mol Cell Biol 13, 127-133.

Wosczyna, M.N., and Rando, T.A. (2018). A Muscle Stem Cell Support Group: Coordinated Cellular Responses in Muscle Regeneration. Dev Cell 46, 135-143.

Yin, H., Price, F., and Rudnicki, M.A. (2013). Satellite cells and the muscle stem cell niche. Physiol Rev 93, 23-67.

Yuniati, L., Scheijen, B., van der Meer, L.T., and van Leeuwen, F.N. (2019). Tumor suppressors BTG1 and BTG2: Beyond growth control. J Cell Physiol 234, 5379-5389.

Zammit, P.S. (2017). Function of the myogenic regulatory factors Myf5, MyoD, Myogenin and MRF4 in skeletal muscle, satellite cells and regenerative myogenesis. Seminars in Cell & Developmental Biology 72, 19–32.

Zordan, P., Rigamonti, E., Freudenberg, K., Conti, V., Azzoni, E., Rovere-Querini, P., and Brunelli, S. (2014). Macrophages commit postnatal endothelium-derived progenitors to angiogenesis and restrict endothelial to mesenchymal transition during muscle regeneration. Cell Death Dis 5, e1031.

Chapter 4

A single-cell transcriptomic atlas of human skeletal muscle tissue reveals bifurcated MuSC subpopulations[§]

Abstract

Single-cell RNA-sequencing (scRNA-seq) facilitates the unbiased reconstruction of multicellular tissue systems in health and disease. Here, we present a curated scRNA-seq dataset of human muscle samples from 10 adult donors with diverse anatomical locations. We integrated ~22,000 single-cell transcriptomes using Scanorama to account for technical and biological variation and resolved 16 distinct populations of muscle-resident cells using unsupervised clustering of the data compendium. These cell populations included muscle stem/progenitor cells (MuSCs), which bifurcated into discrete “quiescent” and “early-activated” MuSC subpopulations. Differential expression analysis identified transcriptional profiles altered in the activated MuSCs including genes associated with ageing, obesity, diabetes, and impaired muscle regeneration, as well as long non-coding RNAs previously undescribed in human myogenic cells. Further, we modeled ligand-receptor interactions and observed enrichment of the TWEAK-FN14 pathway in activated MuSCs, a characteristic signature of muscle wasting diseases. In

[§]This chapter is adapted from the following publication and submitted to BMC’s Skeletal Muscle journal:
A reference single-cell transcriptomic atlas of human skeletal muscle tissue reveals bifurcated muscle stem cell populations. (2020).
Andrea J. De Micheli, Jason A. Spector, Olivier Elemento, and Benjamin D Cosgrove.
bioRxiv 2020.01.21.914713. https://doi.org/10.1101/2020.01.21.914713.

contrast, the quiescent MuSCs have enhanced expression of the EGFR receptor, a recognized human MuSC marker. This work provides a new technical resource to examine human muscle tissue heterogeneity and identify potential targets in MuSC diversity and dysregulation in disease contexts.

4.1 Introduction

Skeletal muscles are essential to daily functions such as locomotion, respiration, and metabolism. Upon damage, resident muscle stem cells (MuSCs) repair the tissue in coordination with supporting non-myogenic cell types such as immune cells, fibroblasts, and endothelial cells (Bentzinger et al., 2013). However, with age and disease, the repair capacity of MuSCs declines, leading to complications such as fibrotic scarring, reduced muscle mass and strength (Blau et al., 2015; Järvinen et al., 2014), fat accumulation and decreased insulin sensitivity (Addison et al., 2014), all of which severely affect mobility and quality of life (Larsson et al., 2018).

Human MuSCs are defined by the expression of the paired box family transcription factor PAX7 and can be isolated using various surface marker proteins including β1-integrin (CD29), NCAM (CD56), EGFR, and CD82 to varying purities (Pisani et al., 2010; Charville et al., 2015; Alexander et al., 2016; Uezumi et al., 2016; Wang et al., 2019). With ageing, human MuSCs exhibit a heterogeneous expression of the senescence marker $p16^{Ink4a}$ and accumulate other cell-intrinsic alterations in myogenic gene expression programs, cell cycle control, and metabolic regulation (Sousa-Victor, et al., 2014; Blau, et al., 2015). However, given their varied molecular and functional states, our understanding of MuSCs in adult human muscle tissue remains incompletely defined. In

addition, cellular coordination in the regulation of human muscle homeostasis and regeneration remains poorly understood due to the lack of experimentally tractable models with multiple human muscle cell types. Given these challenges we posited that an unbiased single-cell reference atlas of skeletal muscle could provide a useful framework to explore MuSC variability and communication in adult humans.

Here, we deeply profiled the transcriptome of thousands of individual MuSCs and muscle-resident cells from diverse adult human muscle samples using single-cell RNA-sequencing (scRNA-seq). After integrating these donor datasets to conserve biological information and overcome technical variation, we resolved two subpopulations of MuSCs with distinct gene expression signatures. Using differential gene expression analysis and ligand-receptor interaction modeling, we extend the known repertoire of human MuSC gene expression programs, suggesting new regulatory programs that may be associated with human MuSC activation, as well as features of human muscle aging and disease.

4.2 Results

4.2.1 Collection and integration of diverse human scRNA-seq datasets

We used scRNA-seq to collect and annotate a single-cell transcriptomic dataset of diverse adult human muscle samples. The muscle samples were from surgically discarded tissue from n=10 donors (range: 41 to 81 years old) undergoing reconstructive procedures and originating from a wide variety of anatomical sites in otherwise healthy patients (Fig. 16A). Each sample was ~50 mg after removal of extraneous fat and connective tissue. Muscle samples were enzymatically digested into single-cell suspensions and independently loaded into the 10X Chromium system. All together, we

collected over 22,000 human muscle single-cell transcriptomes (2206 ± 1961 cells per dataset) into a single data compendium. Using unsupervised clustering, we resolved 16 types of cells of immune, vascular, and stromal origin, as well as two distinct subpopulations of MuSCs and some myofiber myonuclei (Fig. 16B).

Given important differences in anatomical site, donor health history, age, sex, and surgical procedures, the muscle samples were highly heterogeneous in terms of cell-type diversity and underlying gene expression profiles. Comparing the resulting scRNA-seq datasets is therefore a challenge that we addressed using recently developed bioinformatic integration methods (Stuart et al., 2019ab; Hie et al., 2019). Our goal was to assemble a unified dataset of human muscle tissue that faithfully conserved sources of biological variability such as donor, anatomical location, and cell composition heterogeneity, while accounting for technical biases. We tested four different scRNA-seq data integration methods (Fig. S8) and found that Scanorama (Hie et al., 2019) followed by scaling the output by regressing against the library chemistry technical variable ("10X chemistry") and the number of genes detected per single-cell best satisfied this goal. Detailed information on our methodology is provided in Fig. S8. After integrating the 10 datasets, we noted remarkable consistency amid cell types across donors (Fig. 16C-E); owing to the robustness of scRNA-seq technology, the bioinformatic method chosen, and our sample preparation protocol. Differential gene expression analysis between the 16 distinct subpopulations identified an extensive set of unique markers that we grouped into 4 categories (Fig. 16D).

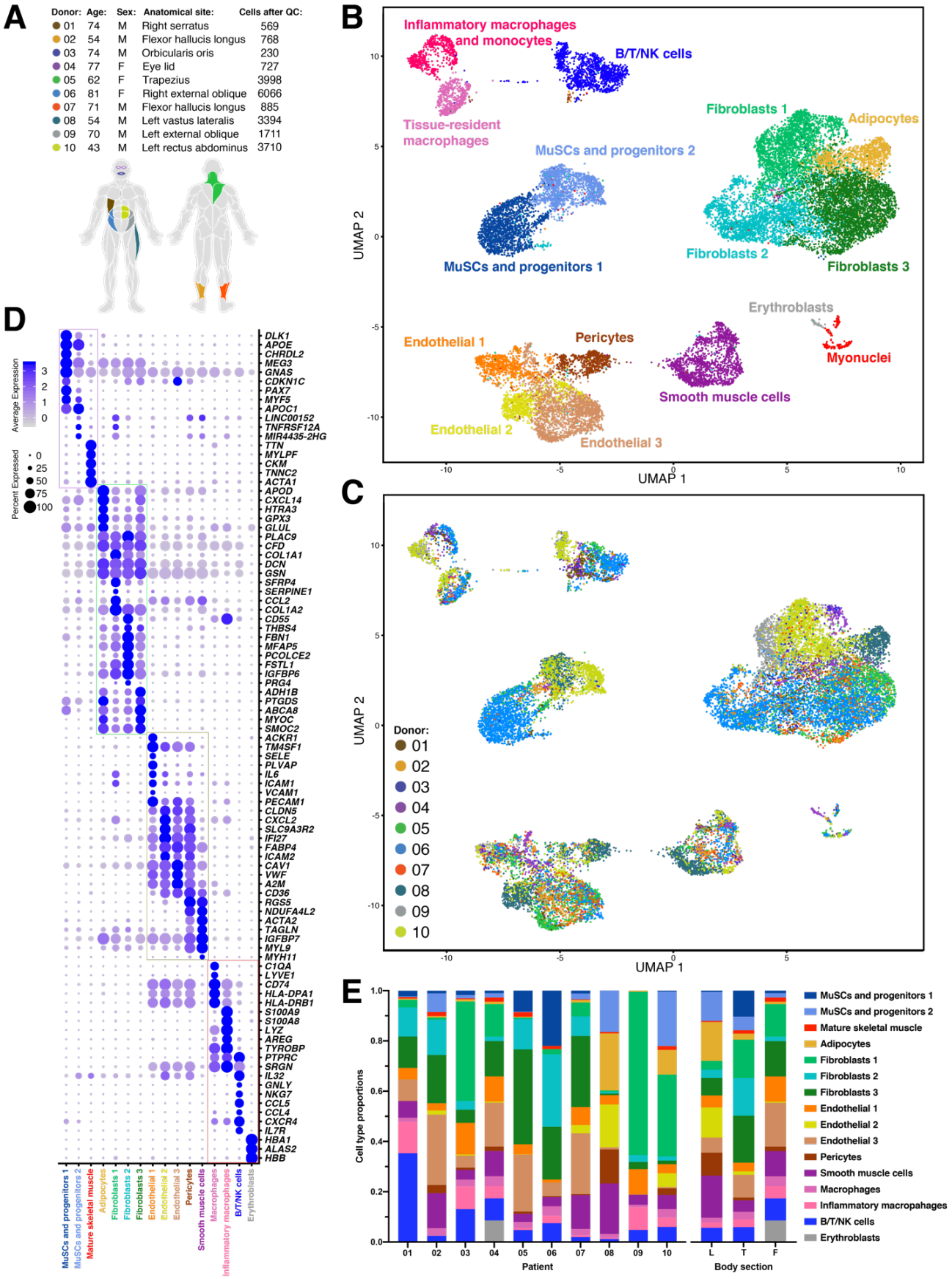

A
Donor: Age: Sex: Anatomical site: Cells after QC:
01 74 M Right serratus 569
02 54 M Flexor hallucis longus 768
03 74 M Orbicularis oris 230
04 77 F Eye lid 727
05 62 F Trapezius 3998
06 81 F Right external oblique 6066
07 71 M Flexor hallucis longus 885
08 54 M Left vastus lateralis 3394
09 70 M Left external oblique 1711
10 43 M Left rectus abdominus 3710
B
Inflammatory macrophages and monocytes
B/T/NK cells
Tissue-resident macrophages
Fibroblasts 1
Adipocytes
MuSCs and progenitors 2
MuSCs and progenitors 1
Fibroblasts 2
Fibroblasts 3
Erythroblasts
Endothelial 1
Pericytes
Myonuclei
Smooth muscle cells
Endothelial 2
Endothelial 3
UMAP 1
UMAP 2
C
Donor:
01
02
03
04
05
06
07
08
09
10
UMAP 1
UMAP 2
D
Average Expression
Percent Expressed
DLK1 APOE CHRDL2 MEG3 GNAS CDKN1C PAX7 MYF5 APOC1 LINC00152 TNFRSF12A MIR4435-2HG TTN MYLPF CKM TNNC2 ACTA1 APOD CXCL14 HTRA3 GPX3 GLUL PLAC9 CFD COL1A1 DCN GSN SFRP4 SERPINE1 CCL2 COL1A2 CD55 THBS4 FBN1 MFAP5 PCOLCE2 FSTL1 IGFBP6 PRG4 ADH1B PTGDS ABCA8 MYOC SMOC2 ACKR1 TM4SF1 SELE PLVAP IL6 ICAM1 VCAM1 PECAM1 CLDN5 CXCL2 SLC9A3R2 IFI27 FABP4 ICAM2 CAV1 VWF A2M CD36 RGS5 NDUFA4L2 ACTA2 TAGLN IGFBP7 MYL9 MYH11 C1QA LYVE1 CD74 HLA-DPA1 HLA-DRB1 S100A9 S100A8 LYZ AREG TYROBP PTPRC SRGN IL32 GNLY NKG7 CCL5 CCL4 CXCR4 IL7R HBA1 ALAS2 HBB
MuSCs and progenitors 1
MuSCs and progenitors 2
Mature skeletal muscle
Adipocytes
Fibroblasts 1
Fibroblasts 2
Fibroblasts 3
Endothelial 1
Endothelial 2
Endothelial 3
Pericytes
Smooth muscle cells
Macrophages
Inflammatory macrophages
B/T/NK cells
Erythroblasts
E
Cell type proportions
Patient
Body section
MuSCs and progenitors 1
MuSCs and progenitors 2
Mature skeletal muscle
Adipocytes
Fibroblasts 1
Fibroblasts 2
Fibroblasts 3
Endothelial 1
Endothelial 2
Endothelial 3
Pericytes
Smooth muscle cells
Macrophages
Inflammatory macropahages
B/T/NK cells
Erythroblasts

Figure 16 *(previous page)***. Single-cell transcriptomic map of human muscle tissue biopsies. (A)** Metadata (sex, age, anatomical site, and the number of single-cell transcriptomes after quality control (QC) filtering) from n=10 donors. Colors indicate sample anatomical sites. **(B)** Scanorama-integrated and batch-effect corrected transcriptomic atlas revealing a consensus description of 16 distinct muscle-tissue cell populations. **(C)** Transcriptomic atlas colored by donor and anatomical location. **(D)** Dot-plot showing differentially expressed genes that distinguish the cell populations. Grouped in four compartments: muscle, endothelial/vascular, stromal, and immune. **(E)** Cell type proportions as annotated in (B) across the 10 donors and grouped by body sections. L = leg (donors 02, 07, 08), T = trunk (donors 01, 05, 06, 09, 10), F = face (donors 03, 04).

4.2.2 ScRNA-seq resolves the cellular diversity of human muscle and novel markers

We annotated and interpretation the consensus cell atlas (Fig. 16B,D) into cell type sub-populations as follows. We identify four types of stromal cells starting with adipocytes found to be expressing apolipoprotein D (*APOD*) (Muffat et al., 2010), the brown fat tissue adipokine *CXCL14* (Cereijo et al., 2018), *GPX3*, and *GLUL*. Among the 3 other subpopulations of fibroblast-like cells, Fibroblasts 1 express high levels of collagen 1 (*COL1A1*), *SFRP4*, *SERPINE1*, and *CCL2*; Fibroblasts 2 express fibronectin (*FBN1*), the microfibril-associated glycoprotein *MFAP5*, and *CD55* known to be expressed by synoviocytes (Karpus et al., 2015); and Fibroblast 3 is mainly characterized by *SMOC2*, a marker of tendon fibroblasts (De Micheli et al., 2020c).

We also identify 5 types of vascular cells, including 3 endothelial subpopulations, and a subpopulation of pericytes and smooth muscle cells (SMCs). Pericytes and SMCs express the canonical markers *RGS5* and *MYH11*. Endothelial 1 express E-selectin (*SELE*), *IL6*, *ICAM1*, *VCAM1*. These genes are upregulated at sites of inflammation to facilitate immune cell recruitment, suggesting this Endothelial 1 cell population may be involved in homeostatic muscle tissue remodeling (Watson et al., 1996; Goncharov et al., 2017). Endothelial 2 cells are distinguished by expressing high levels of claudin-5

(*CLDN5*), *ICAM2*, and the chemokine *CXCL2*. Endothelial 3 express high levels of the platelet-recruiting Von Willebrand Factor (*VWF*) and caveolin-1 (*CAV1*), a protein known to regulate cholesterol metabolism, atherosclerosis progression, as well as MuSC activation (Fernández-Hernando et al., 2010, Volonte et al., 2004).

We also noted two types myeloid immune cells. First, tissue-resident and anti-inflammatory macrophages which express *CD74* and histocompatibility complex *HLA* proteins. Second, activated macrophages and monocytes that express inflammatory markers such as *S100A9* (calgranulin) and *LYZ* (lysozyme). Moreover, *S100A9* transcript abundance levels have been shown to be a feature in ageing and chronic inflammation (Swindell et al., 2013). We also identified a pool of T/B lymphocytes and NK cells characterized by *IL7R* and *NKG7,* respectively, as well as a small subset of *HBA1*$^{+}$ erythroblasts.

Finally, we identified two subpopulations of MuSCs (henceforth called “MuSC1” and “MuSC2”). MuSC1 highly expressed the canonical myogenic transcription factor *PAX7* (Kuang et al., 2006), as well as chordin-like protein 2 (*CHRDL2*) and Delta-like non-canonical Notch ligand 1 (*DLK1*). *CHRDL2* has been shown been previously shown to be expressed in freshly isolated quiescent human MuSCs (Charville et al., 2015), though its function is still to be understood. DLK1 is an inhibitor of adipogenesis whose role in muscle has mainly been recognized in the embryo but remains controversial in adult muscle regeneration (Waddell et al., 2010; Andersen et al., 2013; Zhang et al., 2019). In contrast to MuSC1, MuSC2 expressed lower levels of *PAX7* but maintain expression of *MYF5* (a marker of activated MuSCs) and *APOC1* (Fig. 17B). Interestingly, the MuSC2

population also had elevated expression of two long non-coding RNAs (lncRNAs), *LINC00152* and *MIR4435-2HG*. LncRNAs are involved in regulating myogenesis (Hagan et al., 2017). Surprisingly, we detected low expression of the myogenic commitment factors *MYOD1* and *MYOG* (Fig. 17B), in contrast to scRNA-seq analyses of adult mouse muscle (Dell'Orso et al., 2019; De Micheli et al., 2020). These observations suggest that the MuSC1 and MuSC2 populations are both comprised largely of muscle stem cells, not committed myogenic progenitors. In addition, we noted that "Myonuclei" population (Fig. 17B) was enriched for myosin light chain (*MYLFP*), skeletal alpha-actin (*ATCA1*), and troponin C (*TNNC2*), proteins involved in muscle contraction. This multiple-donor scRNA-seq atlas highlights the cellular diversity of human muscle tissue and revealed two distinct MuSC subpopulations along with specific myogenic expression programs.

4.2.3 Homeostatic human muscle contains two distinct MuSC subpopulations

Next we examined genes that were differentially expressed between the MuSC1 and MuSC2 subpopulations and the biological processes that characterize them (Fig. 17A-B). The MuSC1 subpopulation was enriched for *PAX7*, *DLK1*, and *CHRDL2*, as well as for the cyclin-dependent kinase inhibitor *CKDN1C* (encoding $P57^{KIP2}$), suggesting that these cells are quiescent and not cycling. In addition, this subpopulation expresses the transcription factor *BTG2*, which was identified in mouse to be enriched in quiescent MuSCs (De Micheli et al., 2020). We also note that the MuSC1 subpopulation expressed elevated levels of mitochondrial genes as well as *FOS*, *JUN*, and *ERG1*. Upregulation of these genes has been shown to be potential artefacts of the enzymatic digestion during the sample preparation (van den Brink et al., 2017).

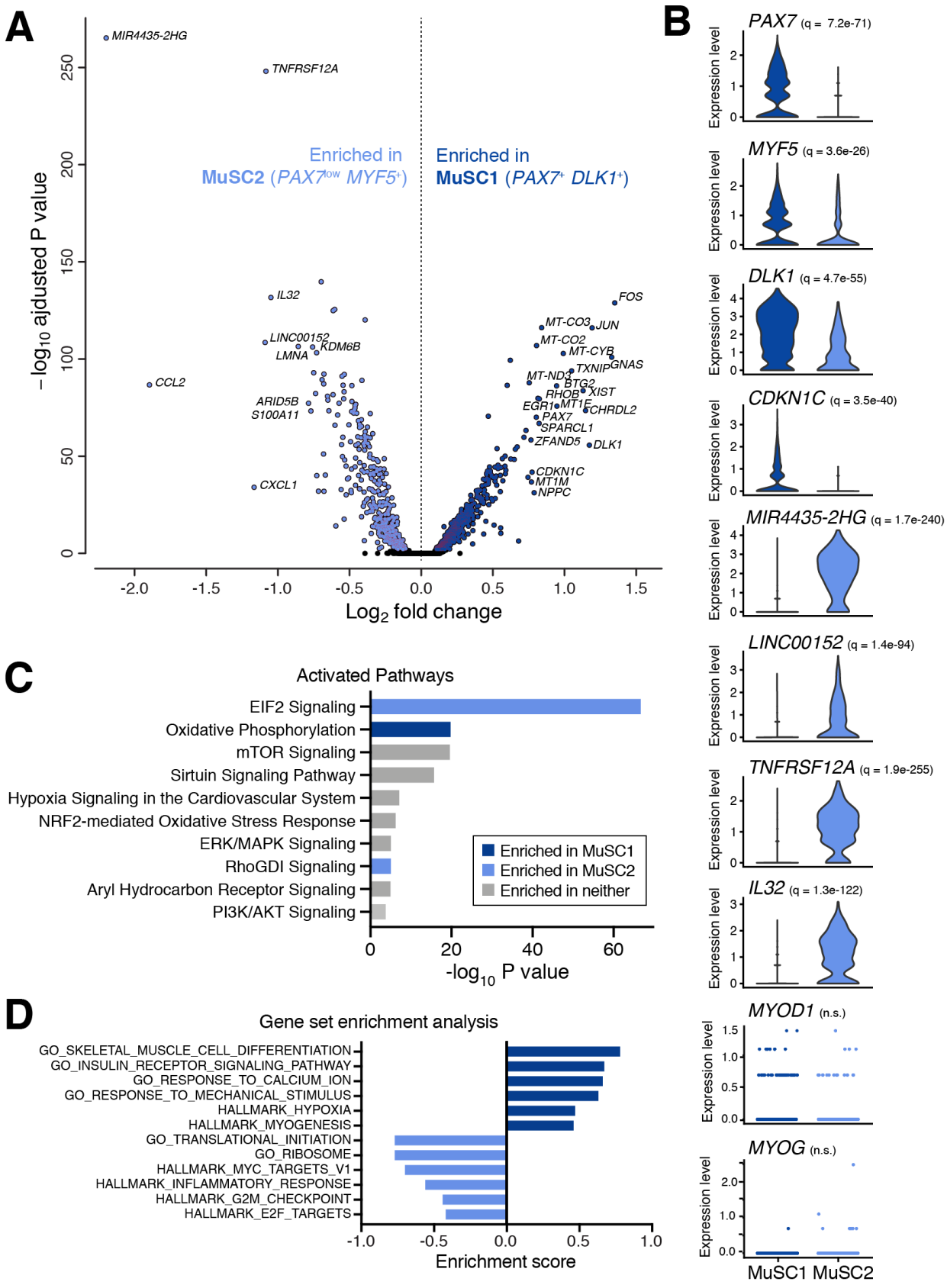

A
MIR4435-2HG
TNFRSF12A
Enriched in
MuSC2 (PAX7low MYF5+)
Enriched in
MuSC1 (PAX7+ DLK1+)
IL32
LINC00152
KDM6B
LMNA
CCL2
ARID5B
S100A11
CXCL1
FOS
MT-CO3
JUN
MT-CO2
MT-CYB
GNAS
TXNIP
MT-ND1
BTG2
RHOB
XIST
EGR1
MT1E
CHRDL2
PAX7
SPARCL1
ZFAND5
DLK1
CDKN1C
MT1M
NPPC
−log10 ajdusted P value
Log2 fold change
B
PAX7 (q = 7.2e-71)
MYF5 (q = 3.6e-26)
DLK1 (q = 4.7e-55)
CDKN1C (q = 3.5e-40)
MIR4435-2HG (q = 1.7e-240)
LINC00152 (q = 1.4e-94)
TNFRSF12A (q = 1.9e-255)
IL32 (q = 1.3e-122)
MYOD1 (n.s.)
MYOG (n.s.)
Expression level
MuSC1 MuSC2
C
Activated Pathways
EIF2 Signaling
Oxidative Phosphorylation
mTOR Signaling
Sirtuin Signaling Pathway
Hypoxia Signaling in the Cardiovascular System
NRF2-mediated Oxidative Stress Response
ERK/MAPK Signaling
RhoGDI Signaling
Aryl Hydrocarbon Receptor Signaling
PI3K/AKT Signaling
Enriched in MuSC1
Enriched in MuSC2
Enriched in neither
-log10 P value
D
Gene set enrichment analysis
GO_SKELETAL_MUSCLE_CELL_DIFFERENTIATION
GO_INSULIN_RECEPTOR_SIGNALING_PATHWAY
GO_RESPONSE_TO_CALCIUM_ION
GO_RESPONSE_TO_MECHANICAL_STIMULUS
HALLMARK_HYPOXIA
HALLMARK_MYOGENESIS
GO_TRANSLATIONAL_INITIATION
GO_RIBOSOME
HALLMARK_MYC_TARGETS_V1
HALLMARK_INFLAMMATORY_RESPONSE
HALLMARK_G2M_CHECKPOINT
HALLMARK_E2F_TARGETS
Enrichment score

Figure 17 (*previous page*) **Gene expression and pathway analysis comparison between two MuSC subpopulations. (A)** Volcano plot from comparing transcript levels between all cells within "MuSC1" and "MuSC2" subpopulations. Log_2 fold-change in normalized gene expression versus -log_{10}-adjusted p-value plotted. Differentially expressed genes (adjusted *p*-value < 0.05) are colored dark or light blue (based on their enrichment in MuSC1 or MuSC2). Genes with log_2 fold-change > 0.75 are labeled. **(B)** Normalized expression values of select differentially expressed genes. *q*-values reported in inset. **(C)** Top activated canonical pathways by Ingenuity Pathway Analysis based on differentially expressed genes and ranked by *p*-value. Pathways significantly enriched in either population with |z-score| > 1 are indicated in blue. **(D)** Select gene ontology (GO) terms and hallmark pathways enriched between the MuSC subpopulations as identified by Gene Set Enrichment Analysis (GSEA) and ranked by enrichment score (ES).

The MuSC2 subpopulation was enriched for multiple markers of inflammation including *CCL2*, *CXCL1*, *IL32*, and surface receptor *TNFRSF12*/*FN14*. In particular, CCL2 and CXCL1 are inflammatory cytokines known to be upregulated in muscle repair, exercise, and fat metabolism (Harmon et al., 1985; Pedersen et al., 2012). In addition, IL32 has been shown to have inflammatory properties in human obesity (Catalán et al., 2017) and have a negative impact on insulin sensitivity and myogenesis (Davegårdh et al., 2017), while TNFRSF12/FN14 has been implicated in various muscle wasting diseases (Mittal et al., 2010; Enwere et al., 2014) and metabolic dysfunction (Sato et al., 2014). Furthermore, the MuSC2 population is enriched for ribosomal gene expression (e.g. *RPLP1* and *RPS6*; data not shown), indicating that these cells may have elevated translational mechanisms. Lastly, the MuSC1 population has enriched expression of the myogenic gene *PAX7* and, to a lesser extent, *MYF5*, compared the MuSC2 population. These observations suggest that MuSC1 is comprised of quiescent MuSCs and MuSC2 is comprised of an early-activated MuSCs.

We performed Ingenuity Pathway Analysis (IPA) to compare biological processes differentially activated between the MuSC1 and MuSC2 populations. The IPA gene group "Oxidative Phosphorylation" is enriched in MuSC1 (Ryall et al., 2015), while "EIF2 Signaling", associated with protein translation processes, is enriched in MuSC2 (Fig.

17C). Furthermore, Gene Set Enrichment Analysis (GSEA) also found MuSC1 to be enriched for “myogenesis”, “muscle cell differentiation”, “hypoxia”, and “response to mechanical stimulus” gene sets, consistent with the observation that these cells are both less differentiation and may have stress-associated gene induction due to tissue dissociation (van den Brink et al., 2017) (Fig. 17D). MuSC2 cells are enriched for “ribosome and translational initiation”, “MYC targets” and “E2F (cell proliferation)”, “G2M checkpoint (cell division)”, and “inflammation” gene sets, further supporting the interpretation that these cells may be in an early activated or partially differentiated state within an inflammatory environment (Fig. 17D). Taken together, these observations suggest that the MuSC1 population is comprised of quiescent MuSCs, while the MuSC2 population is comprised of active, proliferating, and/or dysregulated MuSCs, with expression alterations associated with inflammation, ageing, and muscle wasting. Differentially expressed genes such as *IL32*, *CXCL1*, *CCL2*, and *TNFRSF12/FN14* may constitute a marker set for MuSC variation in chronic muscle inflammation in various pathologies.

4.2.4 Ligand-receptor interaction model identifies new surface markers and cell-signaling channels in human skeletal muscle homeostasis

Next, we used a ligand-receptor (LR) interaction model and a database of LR pairs (Ramilowski et al., 2015) to map cell signaling communication channels in human muscle and uncover differences between MuSC1 and MuSC2 subpopulations. The model also identifies interacting ligand(s) and is restricted to receptor genes differentially expressed by a specific cell type within the consensus human muscle cell atlas (Fig. 16B). For each

LR pair, the model calculates an interaction score from differentially expressed receptors on a given cell population (e.g. “MuSC1”) relative to all other population and ligands expressed by other cell types. Both the MuSC1 and MuSC2 subpopulations are involved in numerous LR interactions, as both ligand- and receptor-expressing cells (Fig. 18A), though a majority of all LR interactions instead involve other cell types suggesting that a small subset of potential paracrine interactions in human MuSCs may include MuSCs.

Given the distinct expression profiles between the MuSC1 and MuSC2 subpopulations, we sought to identify genes that could facilitate surface antigen-based separation of these two human MuSC populations for prospective isolation strategies. We identified surface receptor genes that were differentially expressed between the MuSC1 and MuSC2 populations, using a database of 542 human surface “receptor” genes (Ramilowski et al., 2015) (Fig. 18B). MuSC1 were exhibit elevated expression of EGFR, ITGB1, FGFR4, SDC2, as well as the three tetraspanins CD81, CD82, and CD151. EGFR is a recently established human MuSC marker and is required for basal-apical asymmetric cell division (Charville et al., 2015; Wang et al., 2019). The tetraspanin CD82 is also a recently recognized human MuSC maker (Alexander et al., 2016), while CD9 and CD81 have been identified to control muscle myoblast fusion (Charrin et al., 2013). Furthermore, Syndecans (SDCs) have been identified in mouse to be heterogeneously expressed on MuSCs and myoblasts during muscle repair (De Micheli et al., 2020) and have been shown to form co-receptor complexes with integrin β1 (ITGB1) and FGFR4 upstream of signaling pathways regulating myogenesis (Pawlikowski et al., 2017). Only SDC4 and SDC3 have yet been identified to mark adult mouse MuSCs (Pisconti et al., 2012). In

comparison, the MuSC2 population has elevated expression of CD44 and TNFRSF12/FN14 as previously noted. The CD44 receptor has been shown to regulate myoblast migration and fusion in mouse, but also mark MuSCs in osteoarthritis patients (Mylona et al., 2006; Scimeca et al., 2015).

Next, we focused the LR analysis on the MuSC1 and MuSC2 populations. We identified 73 and 6 significant LR interactions for the MuSC1 and MuSC2 populations, respectively (Fig. 18C). Over one third of all interactions in the MuSC1 population involve the *EGFR* receptor, which has recently been shown to play a critical role in directing MuSC asymmetric division in regenerating muscle (Wang et al., 2019). A limited number of EGFR ligands have been identified in muscle repair. For example, amphiregulin (AREG) is secreted by T_{reg} cells (Burzyn et al., 2013). According to our model findings, EGFR may also interact with ligands expressed by immune cells, such as with TGF-α (*TGFA*), heparin-biding EGF (*HBEGF*), amphiregulin (*AREG*), and epiregulin (*EREG*). Other EGFR ligands include brevican (*BCAN*), and betacellulin (*BTC*) produced by endothelial cells, ECM proteins fibulin 3 (*EFEMP1*), decorin (*DCN*), and tenascin C (*TNC*) expressed by fibroblasts, and FGF13, AHM, NRG4 and EGF, expressed by mature skeletal myofibers. We also detect seven interactions involving NOTCH3 with a variety of ligands. Notch3 signaling is involved in maintaining MuSC quiescence, in particular through interaction with DLL4 (Low et al. 2018), which we found differently expressed by endothelial cells along with JAG2. In addition, NOTCH3 also interacts with the ECM protein thrombospondin-2 (THBS2).

Only two receptors, TNFRSF12/FN14 and RPSA, were found differentially expressed in MuSC2 compared to other cell types. The first, TNFRSF12/FN14, interacts

with the TWEAK cytokine ligand. While typically recognized to be expressed by macrophages and other immune cells (Tajrishi et al., 2014), our model suggests TWEAK is also expressed by the Fibroblasts 2 and Pericyte cell populations, though not in a statistically significant manner. The second, RPSA, is surface ribosomal protein that interacts with laminins (LAM), a dual-specificity phosphatase 18 (DUSP18), and prion protein 2 (PRND), which taken together may suggest various pathological processes such as prion diseases and cancer (Pampeno et al., 2014; Wu et al., 2019). Together, this ligand-receptor analysis identifies a broad set of surface markers that could refine the molecular definition of human MuSCs and their subpopulations, as well as candidate cell-communication channels differentially involved in healthy and diseased muscle tissues.

Lastly, we performed a comparative analysis of receptor gene expression between mouse and human MuSCs. We integrated the human scRNA-seq datasets described in Fig. 16 and an adult mouse muscle injury-response scRNA-seq time-course previously reported (De Micheli et al., 2020) by converting mouse genes to their corresponding human ortholog. The multi-species scRNA-seq atlas was integrated with Scanorama and corrected with Harmony (Fig. S9). From this integrated atlas, we annotated all clusters as in Fig. 1. We identified two MuSC clusters which both contained cells from both mouse and human samples. We then performed differential expression analysis between species comparing aggregated human MuSC1 and MuSC2 cells to mouse MuSCs from the uninjured timepoint (Fig. S9). We found that EGFR and CD99 were most differentially expressed by human MuSCs and, conversely, CRLF1 and SDC4 were most enriched in mouse MuSCs. These findings suggest mouse and human MuSC exhibit species-specific receptor expression signatures.

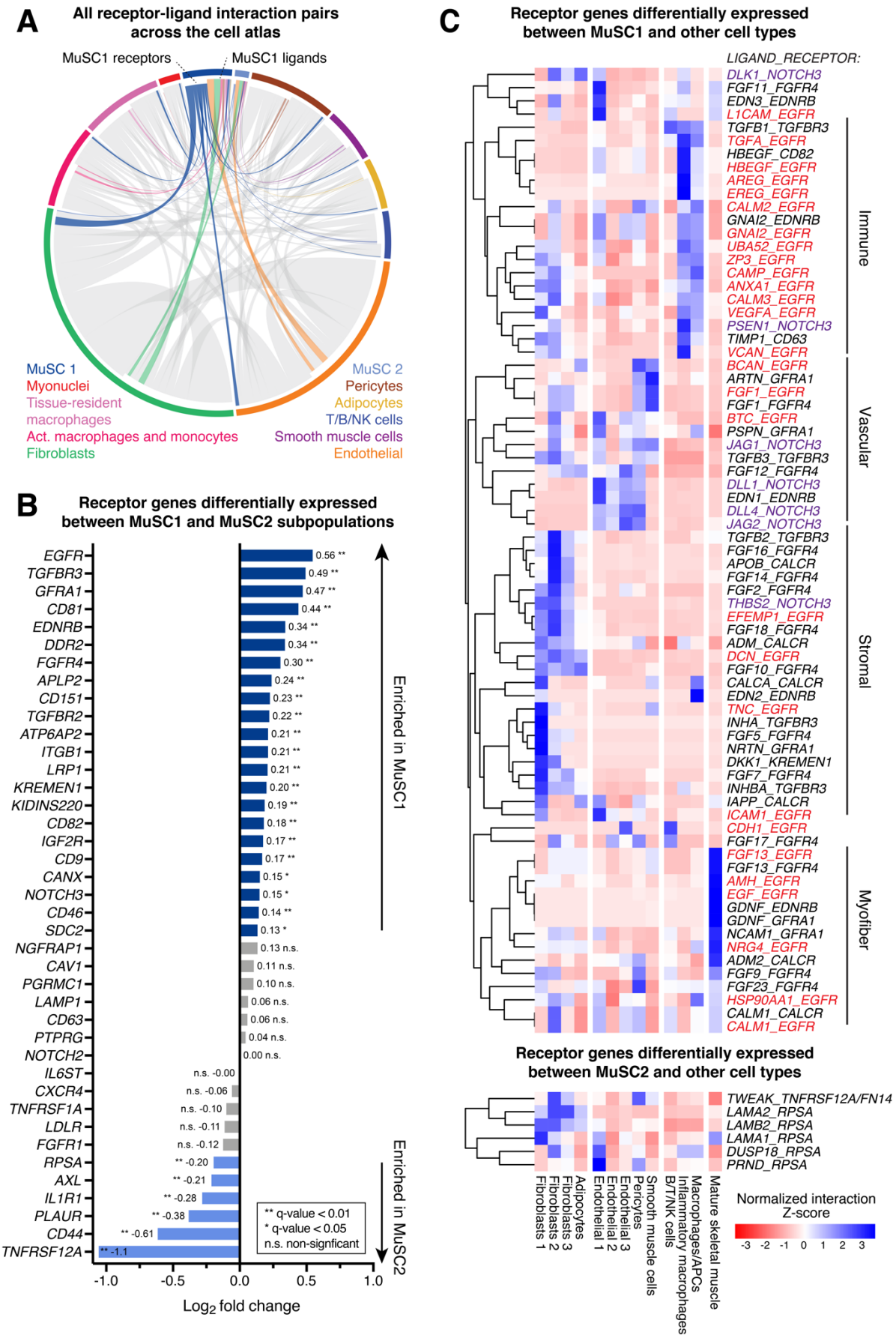

A
All receptor-ligand interaction pairs across the cell atlas
MuSC1 receptors
MuSC1 ligands
MuSC 1
Myonuclei
Tissue-resident macrophages
Act. macrophages and monocytes
Fibroblasts
MuSC 2
Pericytes
Adipocytes
T/B/NK cells
Smooth muscle cells
Endothelial
B
Receptor genes differentially expressed between MuSC1 and MuSC2 subpopulations
EGFR 0.56 **
TGFBR3 0.49 **
GFRA1 0.47 **
CD81 0.44 **
EDNRB 0.34 **
DDR2 0.34 **
FGFR4 0.30 **
APLP2 0.24 **
CD151 0.23 **
TGFBR2 0.22 **
ATP6AP2 0.21 **
ITGB1 0.21 **
LRP1 0.21 **
KREMEN1 0.20 **
KIDINS220 0.19 **
CD82 0.18 **
IGF2R 0.17 **
CD9 0.17 **
CANX 0.15 *
NOTCH3 0.15 *
CD46 0.14 **
SDC2 0.13 *
NGFRAP1 0.13 n.s.
CAV1 0.11 n.s.
PGRMC1 0.10 n.s.
LAMP1 0.06 n.s.
CD63 0.06 n.s.
PTPRG 0.04 n.s.
NOTCH2 0.00 n.s.
IL6ST n.s. -0.00
CXCR4 n.s. -0.06
TNFRSF1A n.s. -0.10
LDLR n.s. -0.11
FGFR1 n.s. -0.12
RPSA ** -0.20
AXL ** -0.21
IL1R1 ** -0.28
PLAUR ** -0.38
CD44 ** -0.61
TNFRSF12A ** -1.1
Enriched in MuSC1
Enriched in MuSC2
** q-value < 0.01
* q-value < 0.05
n.s. non-signficant
-1.0
-0.5
0.0
0.5
1.0
Log2 fold change
C
Receptor genes differentially expressed between MuSC1 and other cell types
LIGAND_RECEPTOR:
DLK1_NOTCH3
FGF11_FGFR4
EDN3_EDNRB
L1CAM_EGFR
TGFB1_TGFBR3
TGFA_EGFR
HBEGF_CD82
HBEGF_EGFR
AREG_EGFR
EREG_EGFR
CALM2_EGFR
GNAI2_EDNRB
GNAI2_EGFR
UBA52_EGFR
ZP3_EGFR
CAMP_EGFR
ANXA1_EGFR
CALM3_EGFR
VEGFA_EGFR
PSEN1_NOTCH3
TIMP1_CD63
VCAN_EGFR
BCAN_EGFR
ARTN_GFRA1
FGF1_EGFR
FGF1_FGFR4
BTC_EGFR
PSPN_GFRA1
JAG1_NOTCH3
TGFB3_TGFBR3
FGF12_FGFR4
DLL1_NOTCH3
EDN1_EDNRB
DLL4_NOTCH3
JAG2_NOTCH3
TGFB2_TGFBR3
FGF16_FGFR4
APOB_CALCR
FGF14_FGFR4
FGF2_FGFR4
THBS2_NOTCH3
EFEMP1_EGFR
FGF18_FGFR4
ADM_CALCR
DCN_EGFR
FGF10_FGFR4
CALCA_CALCR
EDN2_EDNRB
TNC_EGFR
INHA_TGFBR3
FGF5_FGFR4
NRTN_GFRA1
DKK1_KREMEN1
FGF7_FGFR4
INHBA_TGFBR3
IAPP_CALCR
ICAM1_EGFR
CDH1_EGFR
FGF17_FGFR4
FGF13_EGFR
FGF13_FGFR4
AMH_EGFR
EGF_EGFR
GDNF_EDNRB
GDNF_GFRA1
NCAM1_GFRA1
NRG4_EGFR
ADM2_CALCR
FGF9_FGFR4
FGF23_FGFR4
HSP90AA1_EGFR
CALM1_CALCR
CALM1_EGFR
Immune
Vascular
Stromal
Myofiber
Receptor genes differentially expressed between MuSC2 and other cell types
TWEAK_TNFRSF12A/FN14
LAMA2_RPSA
LAMB2_RPSA
LAMA1_RPSA
DUSP18_RPSA
PRND_RPSA
Fibroblasts 1
Fibroblasts 2
Fibroblasts 3
Adipocytes
Endothelial 1
Endothelial 2
Endothelial 3
Pericytes
Smooth muscle cells
B/T/NK cells
Inflammatory macrophages
Macrophages/APCs
Mature skeletal muscle
Normalized interaction Z-score
-3 -2 -1 0 1 2 3

Figure 18 (*previous page*)**. Differentially expressed receptors and LR interaction between cell types.** **(A)** Chord plot of all LR interactions across cell types within the consensus atlas based on co-expression. Each cell type is color-coded with its receptor genes more displaced from the perimeter than its ligand genes. All interactions not involving MuSC1 or MuSC2 are in gray. **(B)** List of differentially expressed receptors between MuSC1 and MuSC2 ranked by $\log_2$ fold-change in expression. Positive average values correspond to genes that are upregulated in MuSC1, whereas negative values are upregulated in MuSC2. Receptors that are statistically significant (FDR-corrected *q*-value < 0.05) are in blue while those not statistically significant are in grey. **(C)** Heatmap representing row-normalized (Z-score) LR interaction scores. Rows represent ligand-receptor interaction pairs in the format LIGAND_RECEPTOR, where the receptor is either differentially expressed in the MuSC1 or MuSC2 populations compared to all the other cell types. Columns identify cell types expressing the ligand. Asterisks after the pair name also indicates that the ligand is differentially expressed by the other cell type and that interaction is likely cell-type specific. Red pairs involve the EGFR receptor, purple pairs the NOTCH3 receptor. A positive value indicates that the interaction has a high score for a particular ligand and cell type compared to other cell types.

4.3 Discussion

Here we present an annotated multi-donor dataset consisting of 22,000 single-cell transcriptomes from 10 different donors and unique anatomical sites, some of which difficult to access outside of reconstructive surgeries. We performed single-cell RNA sequencing and the bioinformatic integration method Scanorama to examine the cellular heterogeneity across diverse adult human muscle tissue samples. We observed that Scanorama performed data integration more successfully than other approaches (Fig. S8). We describe the muscle tissue cellular heterogeneity and provide a comprehensive analysis of differentially expressed genes for 16 resolved cell subpopulations (Fig. 16). This analysis suggests new gene markers for muscle FAPs and vascular endothelial cells that may provide unique perspective to human muscle physiology.

Most notably, this analysis suggests that human muscle may contain two distinct MuSC subpopulations (Fig. 17), adding to a growing literature documenting human muscle cell transcriptional diversity (Riddle et al., 2018; Rubenstein et al., 2020). Given the broad donor age range in this study, these two subpopulations may constitute a healthy and an aged/diseased MuSC pool. We conclude that the “MuSC1” subpopulation

to be largely comprised of “quiescent” MuCSs, owning high levels of *PAX7*, the mitotic inhibitor *CDKN1C*, and *DLK1*. Interestingly, *DLK1* may be an important regulator for human MuSC maintenance and a marker of healthy tissue given its role in inhibiting adipogenesis (Andersen et al., 2013). Conversely, we identified in the “MuSC2” population signatures of inflammation and increased fat metabolism (*CCL2* and *CXCL1*), reduced insulin sensitivity (*IL32*), cell cycle (EIF2 Signaling terms), and muscle wasting (*TNFRSF12*/*FN14*), thereby suggesting that these cells may constitute an “early-activated” and possibly dysfunctional MuSC pool. These markers are consistent with prior observations that excessive fat accumulation in muscle can be attributed to obesity, diabetes, and ageing (Addison et al., 2014). In addition, we identify two upregulated lncRNAs that warrant further investigation as candidate non-coding regulators of myogenesis (Hagan et al., 2017). Moreover, the finding of two human MuSC subpopulations mirrors similar observations made from mouse muscle scRNA-seq analyses (De Micheli et al., 2020; Dell’Orso et al., 2019) and agrees with the general conceptual framework that muscle stem cells transition between quiescent, activated and cycling states. Future studies should compare these MuSC subpopulations across species.

Ligand-receptor interaction models from scRNA-seq data can help formulate new hypotheses about cell-communication channels that regulate muscle function (De Micheli et al., 2020). Identifying new MuSCs surface receptors will also help us refine MuSC purification protocols for prospective isolation studies used for *in vitro* and transplantation models. Our LR model revealed a set of 40 surface receptor genes that are distinctly

expressed between MuSC1 and MuSC2, confirming some prior reports but also providing new candidate surface antigens for human MuSC subpopulation fractionation (Fig. 14). For example, we identify that SDC2 may mark “quiescent” MuSCs while CD44, TNFRSF12, and RPSA “early-activated” MuSCs in ageing and disease contexts. In addition, our model proposed 79 cell-communication signals that may act between MuSCs and other cell types; in particular with fibroblasts, myofibers and immune cells through the EGFR receptor, and with vascular cells through the NOTCH3 receptor. These interactions may be critical regulators of muscle homeostasis and should be further investigated.

This study presents a new set of candidate receptor expression signatures that may define human MuSC subpopulations (Fig. 18B) and provide human-specific receptor patterns (Fig. S9C). This approach is complimentary to receptor screening approaches, which have previously been useful to identify EGFR and CD82 as human MuSC receptor markers for flow cytometry (Alexander et al., 2016; Charville et al., 2015; Uezumi et al., 2016). The subpopulation-specific receptor genes identified here may allow for further comparison of molecular and functional human MuSC diversity across muscle groups (Garcia et al., 2018; Xu et al., 2015).

Our study has some limitations. First, the sample size is small, and donors are very diverse, thus limiting our ability to control for age and sex. We performed differential expression and gene set enrichment analyses within the MuSC1 and MuSC2 populations between the middle-age (43-69 yo) and aged (70-81 you) donors but found few age-cohort specific differences (data not shown). Nevertheless, our dataset still offers a new

transcriptomic cell reference atlas and computational data integration approaches as a resource to examine human muscle cell diversity in health, ageing and disease.

Future studies should aim at collecting muscle specimens in a more controlled manner, for example using a Bergström needle (Tarnopolsky et al., 2011; Sarver et al., 2017) from a unique anatomical site; though this would not be possible for some muscles presented in this study. These biopsies would allow for ageing and disease comparative analyses. Indeed, a recent report by Rubenstein et al. (2020) performed scRNA-seq on four human vastus lateralis muscle biopsies found that myofiber type composition and gene expression alterations based on donor age. Further, future work could also focus on collecting single-myonuclei from myofibers while discarding other non-myogenic cell types. This could illuminate the transcriptomic diversity of myofiber type, on differences that the local anatomy and tissue physiology may demonstrate, and to ultimately enrich our repertoire of know human muscle markers and understanding of its molecular regulators.

4.4 Supplementary material

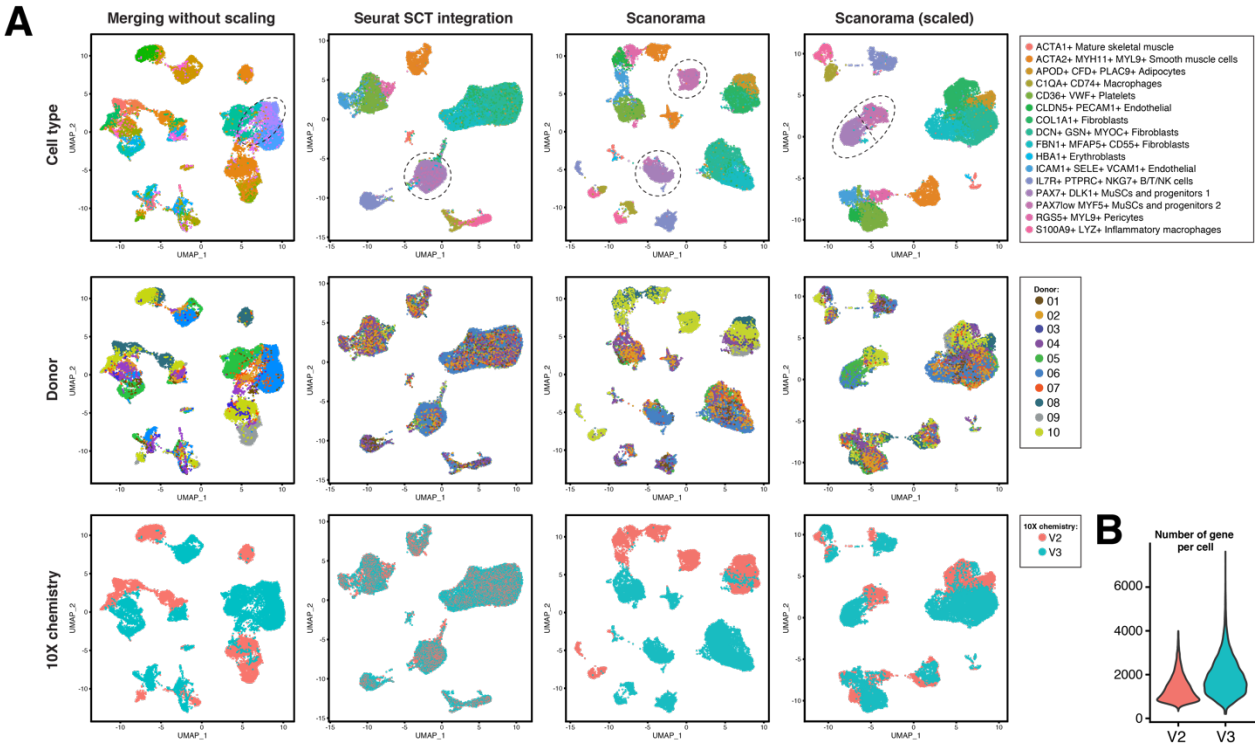


Figure S8 – Comparison of scRNA-seq integration and batch correction methods. We compared four scRNA-seq data integration methods to evaluate which most faithfully conserves donor, anatomical, and biological information while minimizes technical biases. **(A)** The n=10 donor datasets were first annotated independently using a nomenclature of 12 common cell type terms following unsupervised SNN clustering. Then we evaluated the integration method by UMAP and by coloring the data either by cell type, donor ID, or 10X library chemistry used. *First*, we integrated the data by merging the individually normalized gene expression matrices without any further correction. We saw strong technical biases that overwhelmed biological information as the different cell populations segregate by sample/donor and chemistry type. For instance, the two MuSC and progenitor subpopulations are grouped with fibroblasts and endothelial cells. *Second*, we tested the Seurat SCT integration method (Stuart et al., 2019b). This method first calculates a cross-correlation subspace from genes that are shared between datasets. We noticed that this method better "aligns" donor and chemistry type but at the expense of masking biological variability. For instance, we observed that the two MuSC and four stromal subpopulations (Fibroblast 1,2,3 and Adipocytes) were grouped together, hiding important biological heterogeneity. Although certainly useful to validate reproducibility in scRNA-seq experiments, the Seurat SCT integration approach overcorrected biological heterogeneity for heterogeneous samples. *Third*, we tested the Scanorama method (Hie et al., 2019), which relies on a computer vision algorithm that "stitches" datasets together even when the cell type composition between dataset is considerably different. We see that this method groups similar cell populations together while acknowledging donor differences. Yet, surprisingly, this method is also very sensitive at picking up differences in chemistry. To correct this chemistry effect, we scaled the Scanorama output by regressing out the chemistry and the number of genes detected per cell (significantly different between chemistry type) **(B)**. Using this integration method, we observed clear separation of the independently annotated cell populations. We present the resulting Scanorama-integrated dataset as a "consensus atlas" (Fig. 16B-C) of human muscle that describes donor-to-donor differences while grouping cells that are similar together and removing technical biases.

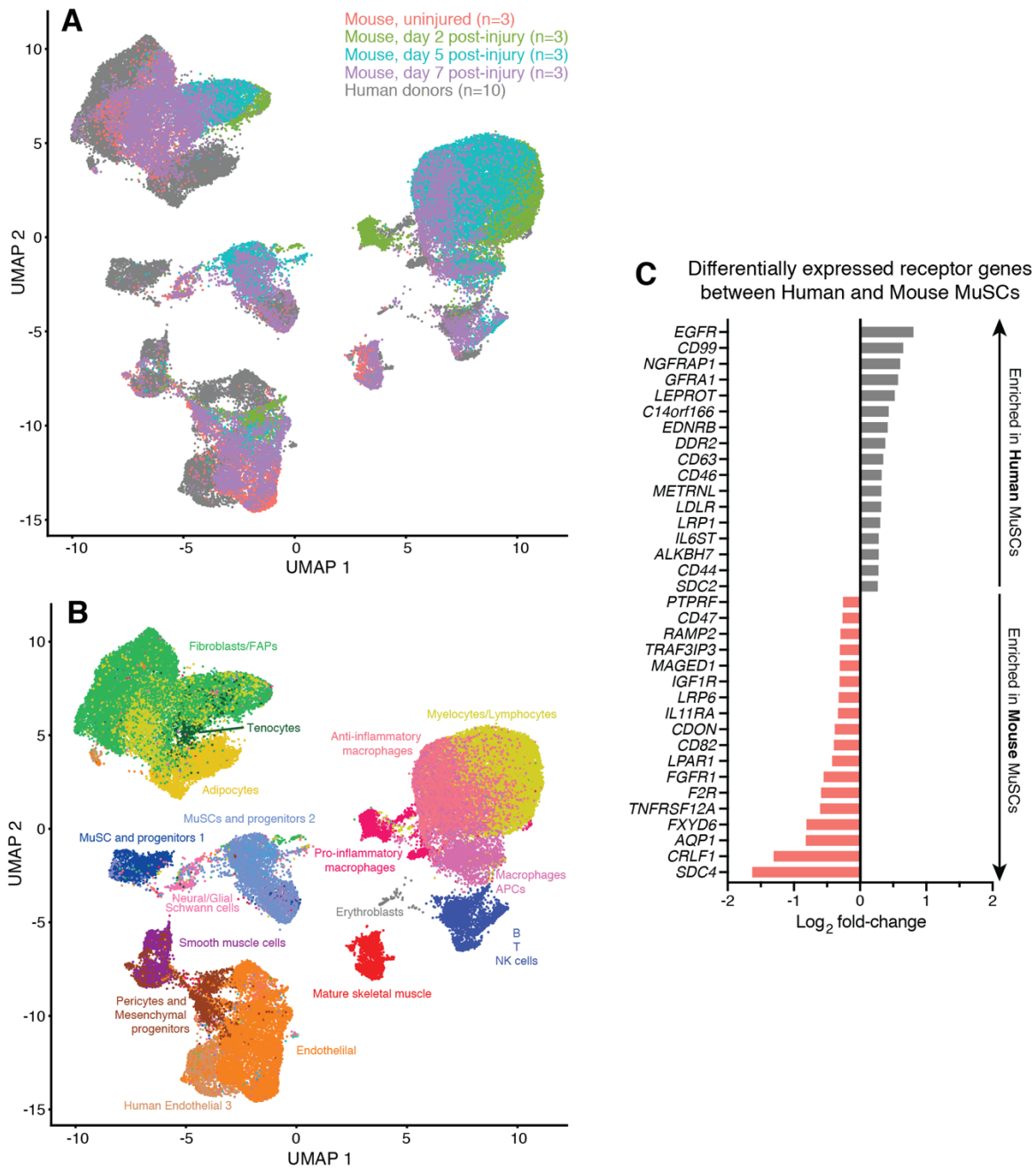


Figure S9 – Integration of human and mouse scRNA-seq data sets allows comparison of MuSC receptor gene expression across species. We generated an integrated scRNA-seq atlas including human sample datasets from Fig. 16 and an adult mouse muscle regeneration time-course from De Micheli et al. (2020). These datasets were integrated using first Scanorama and then Harmony for alignment across species. (**A**) Multi-species integrated atlas presented by UMAP plot a colored by sample type. (**B**) Multi-species integrated atlas presented by UMAP plot and annotated by cell-type clusters. (**C**) The human MuSC1 and MuSC2 clusters were grouped into a cumulative human MuSC cell population, which was compared to mouse MuSCs from the uninjured samples only. Receptor genes were analyzed between the mouse and human MuSC cells for differential expression. Differentially expressed genes with a FDR-corrected q-value < 0.05 are shown in (C).

4.5 Experimental section**

Human participation for muscle sample collection

All procedures were approved by the Institutional Review Board at Weill Cornell Medical College (WCMC IRB Protocol #1510016712) and were performed in accordance with relevant guidelines and regulations. All specimens were obtained at the NewYork-Presbyterian/Weill Cornell Medicine campus. All subjects provided written informed consent prior to participation. Samples were de-identified in accordance to IRB guidelines and only details concerning age, sex, and anatomic origin were included. Sample anatomic locations and donor details are provided in Fig. 16A.

Muscle digestion and single-cell sequencing library preparation

After collection from donors during surgery, the muscle samples were cleared from excessive fat and connective tissue and weighted. About 50-65 mg of tissue was then digested into a single-cell suspension following a previously reported protocol (Spinazzola et al., 2017). Briefly, the specimen was digested in 8 mg/mL Collagenase D (Roche) and 4.8 U/mL Dispase II (Roche) for 1 hr followed by manual dissociation, filtration, and red blood cell lysis. All single-cell suspensions were then frozen at -80^{o}C in 90% FBS, 10% DMSO and were re-filtered after thawing and prior to generating scRNA-seq libraries. The sequencing libraries were prepared using the Chromium Single Cell 3' reagent V2 or V3 kit (10X Genomics) in accordance with the manufacturer's protocol and diluted as to yield a recovery of ~6,000 single-cell transcriptomes with <5% doublet rate. The libraries were

**Complete information about reagents and resource is available in the published manuscript (De Micheli et al., 2020b) and in Appendix 3.

sequenced in multiplex (n=2 per sequencing run) on the NextSeq 500 sequencer (Illumina) to produce between 200 and 250 million reads per library.

Single-cell data analysis

Sequencing reads were processed with the Cell Ranger version 3.1 (10X Genomics) using the human reference transcriptome GRCh38. The downstream analysis was carried out with R 3.6.1 (2019-07-05). Quality control filtering, data clustering, visualization, and differential gene expression analysis was carried out using Seurat 3.1.0 R package. Each of the 10 datasets was first analyzed and annotated independently before integration with Scanorama (Hie et al., 2019). Filtering retained cells with >1000 unique molecular identifiers (UMIs), <20% UMIs mapped to mitochondrial genes, and genes expressed in at least 3 cells. Unsupervised shared nearest neighbor (SNN) clustering was performed with a resolution of 0.4 following which clusters were annotated with a common nomenclature of 12 cell type terms (Fig. S8). Differential expression analysis was achieved using either Seurat's "FindAllMarkers" (Fig. 16D) or "FindMarkers" (Fig. 17A) function using a Wilcoxon Rank Sum test and only considering genes with >log2(0.25) fold-change and expressed in at least 25% of cells in the cluster. P-values were corrected for false-discovery (FDR) and then reported as q-values. Integration of raw counts was achieved using the "scanorama.correct" function from Scanorama. The integrated values were finally scaled in Seurat regressing out the 10X chemistry type and the number of genes per cell. Visualization was done using uniform manifold approximation and projection (UMAP) (Becht et al., 2018). In Fig. S9, we integrated these human scRNA-

seq datasets with a cohort of adult mouse muscle scRNA-seq datasets collected 0-7 days post-notexin injury (De Micheli et al., 2020). For multi-species integration, scRNA-seq datasets were integrated using first Scanorama and then Harmony (Korsunsky et al., 2019) to align related cell populations across species. Mouse genes were converted to human orthologs using biomaRt Bioconductor R package (Durinck et al., 2009).

Pathway and gene set enrichment analysis

The list of differentially expressed genes between MuSC1 and MuSC2 (Fig. 17A) was used in Ingenuity Pathway Analysis (IPA) (QUIAGEN, 2019-08-30). Activated (canonical) pathways were calculated by "Core Analysis" setting a q-value cutoff of 0.05, which yielded 964 genes (366 down, 598 up). Top canonical pathways were chosen based of -log(p-value) and z-score values. Gene set enrichment analysis (GSEA, v.4.0.3) (Subramanian et al., 2005) was ran on the same gene list as IPA ranked by $\log_2$ fold-change and with default program settings. Gene sets database used: h.all.v7.0.symbols.gmt, c2.all.v7.0.symbols.gmt, c5.all.v7.0.symbols.gmt (Broad Institute). Gene sets enriched in phenotype were selected based on q-value and enrichment score (ES).

Ligand-receptor cell communication model

The model presented in Chapters 2 and 3 was used for this project. The model aims at scoring potential ligand-receptor interactions between MuSCs (receptor) and other cell types (ligand). We used the ligand-receptor interaction database from Ramilowski et al.

(2015). From the database, we considered 1915 ligand-receptor pairs (from 542 receptors and 518 ligands) to test for differential expression in our scRNA-seq dataset. To calculate the score for a given ligand-receptor pair, we multiply the average receptor expression in MuSCs by the average ligand expression per other cell type. We only considered receptors that are differentially expressed in either the MuSC1 or MuSC2 subpopulation when compared individually to all other cell types.

Availability of data

The human muscle scRNA-seq datasets supporting the conclusions of this project are available at the GEO repository under accession number GSE143704.

4.6 Chapter references

Addison, O., Marcus, R.L., LaStayo, P.C., and Ryan, A.S. (2014). Intermuscular Fat: A Review of the Consequences and Causes. International Journal of Endocrinology 2014, 1–11.

Alexander, M.S., Rozkalne, A., Colletta, A., Spinazzola, J.M., Johnson, S., Rahimov, F., Meng, H., Lawlor, M.W., Estrella, E., Kunkel, L.M., et al. (2016). CD82 Is a Marker for Prospective Isolation of Human Muscle Satellite Cells and Is Linked to Muscular Dystrophies. Cell Stem Cell 19, 800–807.

Andersen, D.C., Laborda, J., Baladron, V., Kassem, M., Sheikh, S.P., and Jensen, C.H. (2013). Dual role of delta-like 1 homolog (DLK1) in skeletal muscle development and adult muscle regeneration. Development 140, 3743.

Becht, E., McInnes, L., Healy, J., Dutertre, C.A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol.

Bentzinger, C.F., Wang, Y.X., Dumont, N.A., and Rudnicki, M.A. (2013). Cellular dynamics in the muscle satellite cell niche. EMBO Reports 14, 1062–1072.

Blau, H.M., Cosgrove, B.D., and Ho, A.T.V. (2015). The central role of muscle stem cells in regenerative failure with aging. Nature Medicine 21, 854.

Burzyn, D., Kuswanto, W., Kolodin, D., Shadrach, J.L., Cerletti, M., Jang, Y., Sefik, E., Tan, T.G., Wagers, A.J., Benoist, C., et al. (2013). A Special Population of Regulatory T Cells Potentiates Muscle Repair. Cell 155, 1282–1295.

Catalán, V., Gómez-Ambrosi, J., Rodríguez, A., Ramírez, B., Ortega, V.A., Hernández-Lizoain, J.L., Baixauli, J., Becerril, S., Rotellar, F., Valentí, V., et al. (2017). IL-32α-induced inflammation constitutes a link between obesity and colon cancer. Oncoimmunology 6, e1328338–e1328338.

Cereijo, R., Gavaldà-Navarro, A., Cairó, M., Quesada-López, T., Villarroya, J., Morón-Ros, S., Sánchez-Infantes, D., Peyrou, M., Iglesias, R., Mampel, T., et al. (2018). CXCL14, a Brown Adipokine that Mediates Brown-Fat-to-Macrophage Communication in Thermogenic Adaptation. Cell Metabolism 28, 750-763.e6.

Charrin, S., Latil, M., Soave, S., Polesskaya, A., Chrétien, F., Boucheix, C., and Rubinstein, E. (2013). Normal muscle regeneration requires tight control of muscle cell fusion by tetraspanins CD9 and CD81. Nature Communications 4, 1674.

Charville, G.W., Cheung, T.H., Yoo, B., Santos, P.J., Lee, G.K., Shrager, J.B., and Rando, T.A. (2015). Ex Vivo Expansion and In Vivo Self-Renewal of Human Muscle Stem Cells. Stem Cell Reports 5, 621–632.

Davegårdh, C., Broholm, C., Perfilyev, A., Henriksen, T., García-Calzón, S., Peijs, L., Hansen, N.S., Volkov, P., Kjøbsted, R., Wojtaszewski, J.F.P., et al. (2017). Abnormal epigenetic changes during differentiation of human skeletal muscle stem cells from obese subjects. BMC Medicine 15, 39.

De Micheli, A.J., Laurilliard, E.J., Heinke, C.L., Ravichandran, H., Fraczek, P., Soueid-Baumgarten, S., De Vlaminck, I., Elemento, O., and Cosgrove, B.D. (2020). Single-Cell Analysis of the Muscle Stem Cell Hierarchy Identifies Heterotypic Communication Signals Involved in Skeletal Muscle Regeneration. Cell Reports 30, 3583-3595.e5.

De Micheli, A.J., Spector, J.A., Elemento, O., and Cosgrove, B.D. (2020b). A reference single-cell transcriptomic atlas of human skeletal muscle tissue reveals bifurcated muscle stem cell populations. BioRxiv 2020.01.21.914713.

De Micheli, A.J., Swanson, J.B., Disser, N.P., Martinez, L.M., Walker, N.R., Oliver, D.J., Cosgrove, B.D., and Mendias, C.L. (2020c). Single-cell Transcriptomics Identify Extensive Heterogeneity in the Cellular Composition of Mouse Achilles Tendons. BioRxiv 801266.

Dell'Orso, S., Juan, A.H., Ko, K.-D., Naz, F., Gutierrez-Cruz, G., Feng, X., and Sartorelli, V. (2019). Single-cell analysis of adult skeletal muscle stem cells in homeostatic and regenerative conditions. Development dev.174177.

Durinck, S., Spellman, P., Birney, E., Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nature Protocols, 4, 1184–1191.

Enwere, E.K., Lacasse, E.C., Adam, N.J., and Korneluk, R.G. (2014). Role of the TWEAK-Fn14-cIAP1-NF-κB Signaling Axis in the Regulation of Myogenesis and Muscle Homeostasis. Front Immunol 5, 34–34.

Fernández-Hernando, C., Yu, J., Dávalos, A., Prendergast, J., and Sessa, W.C. (2010). Endothelial-Specific Overexpression of Caveolin-1 Accelerates Atherosclerosis in Apolipoprotein E-Deficient Mice. The American Journal of Pathology 177, 998–1003.

Garcia, S.M., Tamaki, S., Lee, S., Wong, A., Jose, A., Dreux, J., Kouklis, G., Sbitany, H., Seth, R., Knott, P.D., et al. (2018). High-Yield Purification, Preservation, and Serial Transplantation of Human Satellite Cells. Stem Cell Reports 10, 1160–1174.

Goncharov, N.V., Nadeev, A.D., Jenkins, R.O., and Avdonin, P.V. (2017). Markers and Biomarkers of Endothelium: When Something Is Rotten in the State. Oxidative Medicine and Cellular Longevity 2017, 9759735.

Hagan, M., Zhou, M., Ashraf, M., Kim, I.-M., Su, H., Weintraub, N.L., and Tang, Y. (2017). Long noncoding RNAs and their roles in skeletal muscle fate determination. Noncoding RNA Investig 1, 24.

Harmon, B.T., Orkunoglu-Suer, E.F., Adham, K., Larkin, J.S., Gordish-Dressman, H., Clarkson, P.M., Thompson, P.D., Angelopoulos, T.J., Gordon, P.M., Moyna, N.M., et al. (2010). CCL2 and CCR2 variants are associated with skeletal muscle strength and change in strength with resistance training. J Appl Physiol (1985) 109, 1779–1785.

Hie, B., Bryson, B., and Berger, B. (2019). Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. Nature Biotechnology 37, 685–691.

Järvinen, T.A., Järvinen, M., and Kalimo, H. (2014). Regeneration of injured skeletal muscle after the injury. Muscles Ligaments Tendons J 3, 337–345.

Karpus, O.N., Kiener, H.P., Niederreiter, B., Yilmaz-Elis, A.S., van der Kaa, J., Ramaglia, V., Arens, R., Smolen, J.S., Botto, M., Tak, P.P., et al. (2015). CD55 deposited on synovial collagen fibers protects from immune complex-mediated arthritis. Arthritis Research & Therapy 17, 6.

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nature Methods 16, 1289–1296.

Kuang, S., Chargé, S.B., Seale, P., Huh, M., and Rudnicki, M.A. (2006). Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. J Cell Biol 172, 103.

Larsson, L., Degens, H., Li, M., Salviati, L., Lee, Y. il, Thompson, W., Kirkland, J.L., and Sandri, M. (2018). Sarcopenia: Aging-Related Loss of Muscle Mass and Function. Physiological Reviews 99, 427–511.

Low, S., Barnes, J.L., Zammit, P.S., and Beauchamp, J.R. (2018). Delta-Like 4 Activates Notch 3 to Regulate Self-Renewal in Skeletal Muscle Stem Cells. STEM CELLS 36, 458–466.

Mittal, A., Kumar, A., Lach-Trifilieff, E., Wauters, S., Li, H., Makonchuk, D., Glass, D., and Kumar, A. (2010). The TWEAK-Fn14 system is a critical regulator of denervation-induced skeletal muscle atrophy in mice. The Journal of Cell Biology 188, 833–849.

Muffat, J., and Walker, D.W. (2010). Apolipoprotein D: An overview of its role in aging and age-related diseases. Cell Cycle 9, 269–273.

Mylona, E., Jones, K.A., Mills, S.T., and Pavlath, G.K. (2006). CD44 regulates myoblast migration and differentiation. Journal of Cellular Physiology 209, 314–321.

Pampeno, C., Derkatch, I.L., and Meruelo, D. (2014). Interaction of Human Laminin Receptor with Sup35, the [PSI+] Prion-Forming Protein from S. cerevisiae: A Yeast Model for Studies of LamR Interactions with Amyloidogenic Proteins. PLoS ONE 9, e86013.

Pawlikowski, B., Vogler, T.O., Gadek, K., and Olwin, B.B. (2017). Regulation of skeletal muscle stem cells by fibroblast growth factors. Developmental Dynamics 246, 359–367.

Pedersen, L., Olsen, C.H., Pedersen, B.K., and Hojman, P. (2012). Muscle-derived expression of the chemokine CXCL1 attenuates diet-induced obesity and improves fatty acid oxidation in the muscle. American Journal of Physiology-Endocrinology and Metabolism 302, E831–E840.

Pisani, D.F., Clement, N., Loubat, A., Plaisant, M., Sacconi, S., Kurzenne, J.-Y., Desnuelle, C., Dani, C., and Dechesne, C.A. (2010). Hierarchization of Myogenic and Adipogenic Progenitors Within Human Skeletal Muscle. STEM CELLS 28, 2182–2194.

Pisconti, A., Bernet, J.D., and Olwin, B.B. (2012). Syndecans in skeletal muscle development, regeneration and homeostasis. Muscles Ligaments Tendons J 2, 1–9.

Ramilowski, J.A., Goldberg, T., Harshbarger, J., Kloppmann, E., Lizio, M., Satagopam, V.P., Itoh, M., Kawaji, H., Carninci, P., Rost, B., et al. (2015). A draft network of ligand–receptor-mediated multicellular signalling in human. Nature Communications 6, 7866.

Riddle, E.S., Bender, E.L., Thalacker-Mercer, A.E., (2018). Transcript profile distinguishes variability in human myogenic progenitor cell expansion capacity. Physiol Genomics 50, 817–827.

Rubenstein, A.B., Smith, G.R., Raue, U., Begue, G., Minchev, K., Ruf-Zamojski, F., Nair, V.D., Wang, X., Zhou, L., Zaslavsky, E., Trappe, T.A., Sealfon, S.C. (2020.) Single-cell transcriptional profiles of human skeletal muscle. Scientific Reports 10, 229.

Ryall, J.G., Dell'Orso, S., Derfoul, A., Juan, A., Zare, H., Feng, X., Clermont, D., Koulnis, M., Gutierrez-Cruz, G., Fulco, M., et al. (2015). The NAD+-Dependent SIRT1 Deacetylase Translates a Metabolic Switch into Regulatory Epigenetics in Skeletal Muscle Stem Cells. Cell Stem Cell 16, 171–183.

Sarver, D.C., Sugg, K.B., Disser, N.P., Enselman, E.R.S., Awan, T.M., and Mendias, C.L. (2017). Local cryotherapy minimally impacts the metabolome and transcriptome of human skeletal muscle. Scientific Reports 7.

Sato, S., Ogura, Y., and Kumar, A. (2014). TWEAK/Fn14 Signaling Axis Mediates Skeletal Muscle Atrophy and Metabolic Dysfunction. Frontiers in Immunology 5, 18.

Scimeca, M., Bonanno, E., Piccirilli, E., Baldi, J., Mauriello, A., Orlandi, A., Tancredi, V., Gasbarra, E., and Tarantino, U. (2015). Satellite Cells CD44 Positive Drive Muscle Regeneration in Osteoarthritis Patients. Stem Cells Int 2015, 469459–469459.

Sousa-Victor, P., Gutarra, S., García-Prat, L., Rodriguez-Ubreva, J., Ortet, L., Ruiz-Bonilla, V., Jardí, M., Ballestar, E., González, S., Serrano, A.L., et al. (2014). Geriatric muscle stem cells switch reversible quiescence into senescence. Nature 506, 316–321.

Spinazzola, J.M., and Gussoni, E. (2017). Isolation of Primary Human Skeletal Muscle Cells. Bio-Protocol 7, e2591.

Stuart, T., and Satija, R. (2019a). Integrative single-cell analysis. Nature Reviews Genetics 20, 257–272.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019b). Comprehensive Integration of Single-Cell Data. Cell 177, 1888-1902.e21.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 102, 15545.

Swindell, W.R., Johnston, A., Xing, X., Little, A., Robichaud, P., Voorhees, J.J., Fisher, G., and Gudjonsson, J.E. (2013). Robust shifts in S100a9 expression with aging: A novel mechanism for chronic inflammation. Scientific Reports 3, 1215.

Tajrishi, M.M., Zheng, T.S., Burkly, L.C., and Kumar, A. (2014). The TWEAK-Fn14 pathway: A potent regulator of skeletal muscle biology in health and disease. Cytokine & Growth Factor Reviews 25, 215–225.

Tarnopolsky, M.A., Pearce, E., Smith, K., and Lach, B. (2011). Suction-modified Bergström muscle biopsy technique: Experience with 13,500 procedures. Muscle & Nerve 43, 716–725.

Uezumi, A., Nakatani, M., Ikemoto-Uezumi, M., Yamamoto, N., Morita, M., Yamaguchi, A., Yamada, H., Kasai, T., Masuda, S., Narita, A., et al. (2016). Cell-Surface Protein Profiling Identifies Distinctive Markers of Progenitor Cells in Human Skeletal Muscle. Stem Cell Reports 7, 263–278.

Volonte, D., Liu, Y., and Galbiati, F. (2004). The modulation of caveolin-1 expression controls satellite cell activation during muscle repair. The FASEB Journal 19, 237–239.

Waddell, J.N., Zhang, P., Wen, Y., Gupta, S.K., Yevtodiyenko, A., Schmidt, J.V., Bidwell, C.A., Kumar, A., and Kuang, S. (2010). Dlk1 Is Necessary for Proper Skeletal Muscle Development and Regeneration. PLOS ONE 5, e15055.

van den Brink, S.C., Sage, F., Vértesy, Á., Spanjaard, B., Peterson-Maduro, J., Baron, C.S., Robin, C., and van Oudenaarden, A. (2017). Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. Nature Methods 14, 935–936.

Wang, Y.X., Feige, P., Brun, C.E., Hekmatnejad, B., Dumont, N.A., Renaud, J.-M., Faulkes, S., Guindon, D.E., and Rudnicki, M.A. (2019). EGFR-Aurka Signaling Rescues Polarity and Regeneration Defects in Dystrophin-Deficient Muscle Stem Cells by Increasing Asymmetric Divisions. Cell Stem Cell 24, 419–432.e6.

Watson, C., Whittaker, S., Smith, N., Vora, A.J., Dumonde, D.C., & Brown, K.A. (1996). IL-6 acts on endothelial cells to preferentially increase their adherence for lymphocytes. Clinical and experimental immunology, 105 1, 112-9.

Wu, Y., Tan, X., Liu, P., Yang, Y., Huang, Y., Liu, X., Meng, X., Yu, B., Wu, M., and Jin, H. (2019). ITGA6 and RPSA synergistically promote pancreatic cancer invasion and metastasis via PI3K and MAPK signaling pathways. Experimental Cell Research 379, 30–47.

Xu, X., Wilschut, K.J., Kouklis, G., Tian, H., Hesse, R., Garland, C., Sbitany, H., Hansen, S., Seth, R., Knott, P.D., Hoffman, W.Y., Pomerantz, J.H. (2015). Human Satellite Cell Transplantation and Regeneration from Diverse Skeletal Muscles. Stem Cell Reports 5, 419-434.

Zhang, L., Uezumi, A., Kaji, T., Tsujikawa, K., Andersen, D.C., Jensen, C.H., and Fukada, S. (2019). Expression and Functional Analyses of Dlk1 in Muscle Stem Cells and Mesenchymal Progenitors during Muscle Regeneration. International Journal of Molecular Sciences 20, 3269.

Chapter 5
Conclusions and future directions

5.1 Summary

In the past decade, the field of biology has witnessed considerable advancements in the manipulation, measurement, annotation, and analysis of single cells. Technologies such as scRNA-seq and CyTOF are poised to generate increasingly large datasets of gene and protein expression information. Combined with bioinformatic algorithms and models, the rapid accumulation of single-cell data is fueling tissue- and organ-level studies of complex molecular mechanisms and cellular processes. In a data-driven manner and often independent from prior hypotheses, single-cell technologies help us formulate and address new biological questions that promise to advance our understanding of multicellular systems such as skeletal muscle.

In this thesis, I presented two large transcriptomic datasets from 12 mouse and 10 human muscle tissues, totaling over 54,000 single-cell transcriptomes (Fig. S9A). The mouse dataset was generated from the tibialis anterior muscle under homeostasis and regenerative conditions. The human dataset was generated by combining distinct muscle biopsies from donors at different ages but biased towards elderly individuals. I also developed a new bioinformatic ligand-receptor interaction model to chart cell-signaling communication within the muscle tissue, which I applied to both the mouse and human single-cell projects. The code is freely available on GitHub (cf. Chapter 2.6.2). I adapted existing bioinformatic methods to curate and annotate transcriptomic atlases of muscle

tissue as a basis to model the myogenic differentiation lineage and discover new myogenic markers. The datasets and analysis framework provided a refined molecular definition of MuSCs from surface receptor expression and identified Syndecans as co-receptors to signaling pathways regulating myogenic progenitor fate. I also proposed new surface markers unique to mouse and human as well as cell-specific communication signals that influence myogenesis. In addition, the analysis framework and ligand-receptor model have also been applied to two other stem cell research projects (Achilles tendon, endometrial epithelium) not presented in this thesis (see: Fu et al., 2019; De Micheli et al., 2020c). The transcriptomic atlases (sequencing reads and annotated gene expression matrices) are published in the Gene Expression Omnibus (GEO) repository[††] for the scientific community to interrogate and re-analyze. In particular, the human dataset provides a valuable resource to explore the effect of ageing, related complications such as muscle atrophy (sarcopenia), and hereditary myopathies (Duchenne muscular dystrophy) on the MuSC transcriptome and how it may perturb tissue cell-signaling homeostasis. The ligand-receptor model can also help in identifying surface protein of dysregulated downstream pathways that may serve as drug targets for addressing such diseases. Compared to mouse MuSCs from the uninjured timepoint (Fig. 11), “homeostatic human MuSC” appear to be considerably more heterogenous and exist in two distinct subpopulations, but none express myogenic markers of differentiation such as *Myog* and *Myod1.* This heterogeneity can be explained by the diversity of human samples (and health) compared to those sourced from animal models. In contrast, in the

[††]GEO accessions numbers: GSE143435, GSE143437 (mouse study), GSE143704 (human study).

mouse study the greatest MuSC heterogeneity was observed at day 5 post-injury and was due to cell activation and differentiation. Though challenging experimentally, I would be informative to harvest “injured” human muscles from traumatic patients in order to compare the human and mouse myogenic lineage by trajectory inference analysis. Finally, ligand-receptor modeling also suggests that common mouse MuSC markers (Fig. S9C) might not define human MuSCs as accurately.

Overall, I hope that the datasets and analyses serve a resource to stimulate new discoveries in muscle biology and health.

Below is a summary of my accomplishments and findings for the mouse (Chapter 3) and human (Chapter 4) studies:

Chapter 3:

1. I developed a protocol to generate single-cell transcriptomes of mouse muscle-tissue cells, including MuSCs. I compared various technologies (for e.g. Drop-seq, 10X) and sample preparation strategies (for e.g. FACS, no-FACS) and selected the one that introduced the least biases.
2. I assembled, curated, and annotated a scRNA-seq reference atlas with over 34,000 single-cell from mouse muscles spanning different timepoints post-injury.
3. I examined the cell-type and gene expression dynamics of the regenerating muscle and revealed transient waves of amplification and diversification of multiple immune cell, FAPs, and myogenic cell types. I also presented the gene expression dynamics of FAPs markers post-injury.
4. I used single-cell trajectory inference analysis (Monocle) to build a model of myogenic differentiation. I demonstrated that the myogenic differentiation model can organize MuSC and progenitor cells into three distinct branches that represent “quiescent”, “cycling”, and “committed” cells, each with their unique set of markers. In particular, I discovered that *Id3* and *Btg2* (two transcription factors known to regulate neural progenitor differentiation and are direct target of *Pax7*) are enriched in the “quiescent” population, while the cycling population is enriched for cell-cycle genes such as *Ccnb2*.

5. I developed a new bioinformatic ligand-receptor interaction model which I used to reveal over 200 significant ligand-receptor pairs in the injured and un-injured muscle.
    a. My model allowed me to observe a diversification of Syndecan (Sdcs) surface receptor type expression in the MuSC and progenitor subpopulation post-injury (Sdc4 -> Sdc1,2,4).
    b. I validated Syndecan protein expression by CyTOF and showed transient induction of Sdcs in the “cycling” subpopulation, suggesting Sdcs might be key mediators of myogenic differentiation.
6. Finally, I used the results of the ligand-receptor model to motivate an *in vitro* screen of paracrine communication factors that influence MuSCs proliferation with dependency on Sdc receptors. Among ligands, I identify that PF4, THBS1, TGFβ1, and RSPO3, as well as ECM proteins fibronectin in Tenascin C, all interact with Sdcs.
    a. I (and colleagues) performed an *in vitro* MuSC ligand testing experiment with FACS-sorted MuSCs and showed that Sdc1 and Sdc4 were both necessary for FGF2/Fn-induced MuSC proliferation, while Sdc2 (but not Sdc4) was necessary for both TGFβ1-augmented and RSPO3-suppressed MuSC proliferation.
    b. These experiments proposed that Sdcs are involved in a notable fraction of MuSC cell-signaling communication channels and suggested that the temporal heterogeneity in Sdc expression observed at the transcript and

protein levels may provide a stage-specific target to both purify and regulate MuSC fate *in vitro*.

7. Altogether, my annotated temporal scRNA-seq atlas of muscle regeneration provides a reference resource to examine the role muscle-tissue cellular diversity and myogenic cell-signaling in aging, disease, and across species.

Chapter 4:

- I presented a curated and annotated scRNA-seq dataset of human muscle samples from 10 adult donors and diverse anatomical locations.
- I provided comparison of multiple scRNA-seq data-integration algorithms to best account for technical and patient variability. I showed that Scanorama delivered the most biologically consistent single-cell atlas.
- Using Scanorama-integrated data, I exposed the cellular diversity of human muscle tissue and novel markers. I revealed 16 distinct cell populations, including 3 subpopulations of endothelial cells, 3 subpopulations of fibroblasts, and 2 subpopulations of MuSC and progenitor cells. In particular, the 3 subpopulations of fibroblast respectively express *COL1A1*, *FBN1*, and *MFAP5*; markers that may help us better delineate new fibroblast subpopulations in muscle.
- I revealed that the human homeostatic muscle contains two distinct MuSC subpopulations, which I characterized by differential gene expression and pathway analyses. I considered the first subpopulation as “quiescent” as cells expressed known genes and new markers such as *PAX7*, *DLK1*, and *CHRDL2*. I considered the second subpopulation as “early-activated” as cells express markers of

inflammation including *IL32,* and *FN14*, a receptor involved in muscle wasting disease. I also identified two upregulated long non-coding RNAs previously undescribed in human myogenic cells. I hypothesize that the causes of this early activation are ageing and age-related muscle diseases.

- I used the ligand-receptor model (Chapter 3) and identified new surface markers and cell-interactions in muscle pathologies. In particular, I revealed that the "quiescent" MuSC subpopulation is enriched for many EGFR and NOTCH3 interactions, while the "early-activated" subpopulation is enriched for interactions involving the FN14 and RPSA receptors. I also proposed new surface markers for the two MuSC subpopulations, including *SDC2* and *CD44* respectively.
- My ligand-receptor model identified a broad set of surface markers that could refine the molecular definition of human MuSC subpopulations, as well as candidate cell-communication channels differentially involved in healthy and diseased muscles.
- I integrated our human datasets and with the mouse dataset presented in Chapter 3 to reveal specie-specific differentially expressed MuSC receptors.
- Altogether, my annotated scRNA-seq dataset offers a new muscle reference atlas to examine human muscle cell diversity in health, ageing and disease. The discovery of new MuSCs markers and downstream signaling pathways that regulate myogenesis may open new possibilities for treating myopathies.

5.2 Future directions

5.2.1 Characterizing transcriptomic biases of single-MuSCs during cell isolation

In this work, it has been observed that muscle tissue digestion and MuSC manipulation introduces biases in the cells' transcriptomic readout. These biases are especially noticeable for MuSCs that upon tissue dissociation loose key niche interactions that normally maintain their quiescent phenotype. In particular, it has been observed in both the mouse and human studies that MuSCs overexpress stress-response genes such as FOS, ERG1, and KLF6A. These genes have also been identified in other studies to be upregulated following collagenase digestion (O'Flanagan et al., 2019). Another recent study also demonstrated by bulk RNA-seq that quiescent MuSCs undergo major transcriptomic alterations during tissue digestion that erases important signature of quiescence (Machado et al., 2017). For these reasons, a rigorous characterization of these biases in MuSCs should be performed for scRNA-seq experiments.

These biases could be first addressed by optimizing the muscle digestion protocol to include fixation with paraformaldehyde (PFA) or methanol in a manner that is compatible with the 10X or future scRNA-seq technologies. Digestion enzyme concentrations and temperatures could also be adjusted to minimize cell stress. This would improve RNA quality and limit the non-physiological expression of stress-response genes and/or downregulation of stem-cell markers. Second, these biases could be addressed computationally, which would also provide an opportunity to "correct" existing datasets. Such corrections would require deriving a stress vector of up/down regulated genes for each cell type (since not all types of cells equally respond to tissue dissociation). This vector could be subtracted from the data or used in more advanced statistical

models. Machine learning algorithms could also help derive the weights of these vectors from highly heterogenous datasets. These efforts should reveal a more physiological signature of quiescence and improve the molecular definition of MuSCs.

5.2.2 Engineering artificial MuSC niches to evaluate ligand-receptor interactions

The ligands identified in Chapter 3 by ligand-receptor modeling were evaluated *in vitro* in standard culture plates. However, such two-dimensional culture systems do not recapitulate the complex architecture of the MuSC niche or muscle tissue during regeneration. Future ligand evaluation studies would deliver more physiologically relevant results if performed in culture systems that mimics the niche's three-dimensional and asymmetric organization of ligands (Fig. 3).

I invented an artificial MuSC niche using biphasic micro-hydrogel biomaterials assembled by droplet-microfluidic (c.f. Annex, Figs. S10-S12). This biphasic artificial niche recapitulates the bipolar (apical/basal) architecture of the MuSC niche. Each biomaterial phase of the artificial niche can be conjugated with desired ligands such as those identified by the ligand-receptor model (Fig. 15A-B) (for e.g. tenascin, fibronectin, Rspo3, Pf4). Chemokines and growth factors could either be conjugated to the hydrogel or supplemented in culture media. Ultimately, the ligand-receptor interactions identified from scRNA-seq experiments would allow for a more rational design and testing of artificial niches as physiological culture platforms for muscle tissue engineering applications.

5.2.3 Co-culture studies from ligand-receptor model results

The ligand-receptor model developed in this thesis also proposes potential cell-signaling communication channels between MuSCs and other cell types. These results should motivate co-culture studies and examine the presence of other cell types on MuSC fate *in vitro*. Co-culture strategies could help us better control and expand MuSCs *ex vivo*. For instance, it has been shown that T-cell conditioned media and four derived cytokines (IL-1α, IL-13, TNF-α, and IFN-γ) promote the long-term expansion of MuSCs in culture (Fu et al., 2015). In addition, the demonstrated dependency of MuSC survival on FAPs *in vivo* (Wosczyna et al., 2019) also suggest that there are crucial cell-signaling molecules between these two cell types that could enhance MuSC culture *in vitro*. For these reasons, my ligand-receptor model could be used to motivate co-culture studies and screens to identify signaling molecules that enhance MuSC function *in vitro.*

5.2.4 Investigate the effect of aging on the MuSC transcriptome

Ageing affects MuSCs in ways that are both intrinsic (cellular) and extrinsic (micro-environment) (Blau et al., 2015). Though ageing is not a research axis of this thesis, my two scRNA-seq datasets can provide the basis for exploring the consequence of ageing and related disease on the MuSC transcriptome. For example, MuSCs are known to accumulate intrinsic defects with age and certain diseases (Blau et al., 2015) and molecularly become more heterogeneous. Repeating the scRNA-seq injury timecourse experiment presented in Chapter 3 with older or dystrophic mice (for e.g. the *mdx* model) would allow us to identify new myogenic signatures of ageing and disease. In addition, the aged MuSC niche is also known to have an altered molecular content (Lukjanenko et

al., 2016). My receptor-ligand model could help identify new MuSC-niche interactions that are affected with ageing.

5.2.5 Single-nucleus RNA-seq to measure the heterogeneity within myofiber types

Myonuclei were annotated in this thesis as “mature skeletal muscle myofibers” given the expression of genes such as actin alpha 1 *(Acta1)* and myosin heave chain 1 (*Myh1)* (Figs. 9 and 16). Myosin proteins exist in various isoforms and are expressed by myonuclei in a fiber type-specific manner (for e.g. *Myh4* is expressed by fast-twitch type-2b fibers) (Schiaffino, 2018). However, in this thesis I was not able to specifically profile the myonuclei transcriptome as myofibers were discarded during sample preparation. The few myonuclei identified were either fragments of myofiber or myonuclei given their low percentage of mitochondrial RNA compared to other cell types.

Recently, Blackburn et al. (2019) proposed a method to measure the transcriptome of single myofibers using Smart-Seq and compared young and old mice muscles fibers (Blackburn et al., 2019). The study, however, does not consider differences in fiber type (such as slow- and fast-twitch) nor is designed to examine the heterogeneity of nuclei within a fiber. Because there is evidence that different type of myofibers may exhibit unique ageing characteristics (Verdijk et al., 2007; Kim et al., 2012) and that myonuclei positioning influences muscle function (Perillo and Folker, 2018; Roman and Gomes, 2018), single-nucleus RNA-sequencing (Habib et al., 2017) may help us investigate the gene expression heterogeneity of myonuclei. An important question to address could be whether myonuclei transcriptomic heterogeneity reflects myonuclei positioning, fiber type, and/or the neighboring physiology (for e.g. neuromuscular versus myotendinous

junctions, where myonuclei tend to aggregate). The analysis framework developed in this thesis could be used to address these questions.

5.2.6 Spatial transcriptomics: scRNA-seq and imaging CyTOF data

Single-cell RNA-seq does not capture the spatial positioning of cells within tissues. However, muscle regeneration displays important spatial events. For example, MuSC symmetric self-renewal divisions are longitudinal to myofibers while asymmetric divisions are oriented along the apical-basal axis of the niche. In both instances, specific protein such as Vangl2, Dystrophin, and Pard3 mark either daughter MuSC (Gros et al., 2009; Le Grand et al., 2009; Troy et al., 2012). Additionally, muscle repair also involves many spatially important juxtracrine cell-cell interactions. Therefore, there is a need add spatial positioning information to single-cell transcriptomic data.

Spatial transcriptomics promise to bridge spatial context and transcriptomic information. However, current methods such as Slide-seq (Rodriques et al., 2019) or RNA-seqFISH are either limited in resolution (~15 µm) or require complex protocols or optical equipment (Eng et al., 2019). In this thesis, I use mass cytometry (CyTOF) to validate scRNA-seq findings. CyTOF can also be used in an imaging modality to generate high-dimensional (>100 markers) images of tissue sections with 1 µm resolution. Thus, developing algorithms to integrate CyTOF images with scRNA-seq data could provide a more accessible path to spatial transcriptomics. Bioinformatic algorithms using non-linear distance metrics or probabilistic models will have to be developed to link a segmented cell in a CyTOF image to the “nearest” or “most probable” cell in a scRNA-seq dataset. Possible challenges could be the important difference in feature scale or type of technical

noise that is specific to each technology. While scRNA-seq can resolve over 15,000 genes, CyTOF can measure only a panel of about 100 protein markers. In addition, protein expression does not always correlate well to mRNA levels. A possible way to address these issues would be to use the CITE-seq technology (Stoeckius et al., 2017). CITE-seq can simultaneously measure the transcriptome and protein expression level information from a panel of DNA-tagged antibodies. By using the same antibody panel for CITE-seq and imaging CyTOF one would be able to associate a “CITE-seq cell” to a “segmented CyTOF cell” from similar protein expression profiles.

5.2.7 Large-scale integration of muscle single-cell data for a unified, multi-modal, and public resource

Collectively, the field is accumulating enormous amount of single-cell information from a wide variety of data modalities such as gene expression (scRNA-seq), chromatic accessibility (ATAC-seq), immunophenotyping (CITE-seq, CyTOF), and spatial transcriptomics. As new algorithms are continuously being developed to integrate single-cell datasets at scale, this enables, in a complementary manner, the analysis of datasets across diverse experimental settings, species, technologies, and laboratories. Akin to projects such as the Human Cell Atlas (Regev et al., 2017) or Tabula Muris (Schaum et al., 2018), a rigorous compilation and curation of past and new muscle single-cell datasets into a public and “interrogatable” resource will undeniably power new discoveries in the field of muscle biology and beyond.

5.3 Chapter references

Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat Methods *14*, 1083–1086.

Blackburn, D.M., Lazure, F., Corchado, A.H., Perkins, T.J., Najafabadi, H.S., and Soleimani, V.D. (2019). High-resolution genome-wide expression analysis of single myofibers using SMART-Seq. J. Biol. Chem. *294*, 20097–20108.

Blau, H.M., Cosgrove, B.D., and Ho, A.T.V. (2015). The central role of muscle stem cells in regenerative failure with aging. Nat Med *21*, 854–862

De Micheli, A.J., Swanson, J.B., Disser, N.P., Martinez, L.M., Walker, N.R., Oliver, D.J., Cosgrove, B.D., and Mendias, C.L. (2020c). Single-cell Transcriptomics Identify Extensive Heterogeneity in the Cellular Composition of Mouse Achilles Tendons. BioRxiv 801266.

Eng, C.-H.L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y., Yun, J., Cronin, C., Karp, C., Yuan, G.-C., et al. (2019). Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. Nature *568*, 235–239.

Fu, D.-J., De Micheli, A.J., Bidarimath, M., Ellenson, L.H., Cosgrove, B.D., Flesken-Nikitin, A., and Nikitin, A.Yu. (2019). PAX8 expressing epithelial cells are a cancer-prone source of clonal cyclical regeneration of endometrial epithelium. BioRxiv 853994.

Fu, X., Xiao, J., Wei, Y., Li, S., Liu, Y., Yin, J., Sun, K., Sun, H., Wang, H., Zhang, Z., et al. (2015). Combination of inflammation-related cytokines promotes long-term muscle stem cell expansion. Cell Res *25*, 655–673.

Gros, J., Serralbo, O., and Marcelle, C. (2009). WNT11 acts as a directional cue to organize the elongation of early muscle fibres. Nature *457*, 589–593.

Habib, N., Avraham-Davidi, I., Basu, A., Burks, T., Shekhar, K., Hofree, M., Choudhury, S.R., Aguet, F., Gelfand, E., Ardlie, K., et al. (2017). Massively parallel single-nucleus RNA-seq with DroNc-seq. Nature Methods *14*, 955–958.

Kim, J.-H., Torgerud, W.S., Mosser, K.H.H., Hirai, H., Watanabe, S., Asakura, A., and Thompson, L.V. (2012). Myosin light chain 3f attenuates age-induced decline in contractile velocity in MHC type II single muscle fibers: Age, velocity, and myosin light chains. Aging Cell *11*, 203–212.

Le Grand, F., Jones, A.E., Seale, V., Scimè, A., and Rudnicki, M.A. (2009). Wnt7a Activates the Planar Cell Polarity Pathway to Drive the Symmetric Expansion of Satellite Stem Cells. Cell Stem Cell *4*, 535–547.

Lukjanenko, L., Jung, M.J., Hegde, N., Perruisseau-Carrier, C., Migliavacca, E., Rozo, M., Karaz, S., Jacot, G., Schmidt, M., Li, L., et al. (2016). Loss of fibronectin from the aged stem cell niche affects the regenerative capacity of skeletal muscle in mice. Nature Medicine *22*, 897–905.

Machado, L., Esteves de Lima, J., Fabre, O., Proux, C., Legendre, R., Szegedi, A., Varet, H., Ingerslev, L.R., Barrès, R., Relaix, F., et al. (2017). In Situ Fixation Redefines Quiescence and Early Activation of Skeletal Muscle Stem Cells. Cell Reports *21*, 1982–1993.

O'Flanagan, C.H., Campbell, K.R., Zhang, A.W., Kabeer, F., Lim, J.L.P., Biele, J., Eirew, P., Lai, D., McPherson, A., Kong, E., et al. (2019). Dissociation of solid tumor tissues with cold active protease for single-cell RNA-seq minimizes conserved collagenase-associated stress responses. Genome Biology *20*, 210

Perillo, M., and Folker, E.S. (2018). Specialized Positioning of Myonuclei Near Cell-Cell Junctions. Front. Physiol. *9*, 1531.

Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., et al. (2017). The Human Cell Atlas. ELife *6*, e27041.

Rodriques, S.G., Stickels, R.R., Goeva, A., Martin, C.A., Murray, E., Vanderburg, C.R., Welch, J., Chen, L.M., Chen, F., and Macosko, E.Z. (2019). Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. 6.

Roman, W., and Gomes, E.R. (2018). Nuclear positioning in skeletal muscle. Seminars in Cell & Developmental Biology *82*, 51–56.

Schaum, N., Karkanias, J., Neff, N.F., May, A.P., Quake, S.R., Wyss-Coray, T., Darmanis, S., Batson, J., Botvinnik, O., Chen, M.B., et al. (2018). Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. Nature *562*, 367–372.

Schiaffino, S. (2018). Muscle fiber type diversity revealed by anti-myosin heavy chain antibodies. FEBS J *285*, 3688–3694.

Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P.K., Swerdlow, H., Satija, R., and Smibert, P. (2017). Simultaneous epitope and transcriptome measurement in single cells. Nature Methods *14*, 865–868.

Troy, A., Cadwallader, A.B., Fedorov, Y., Tyner, K., Tanaka, K.K., and Olwin, B.B. (2012). Coordination of Satellite Cell Activation and Self-Renewal by Par-Complex-Dependent Asymmetric Activation of p38α/β MAPK. Cell Stem Cell *11*, 541–553.

Verdijk, L.B., Koopman, R., Schaart, G., Meijer, K., Savelberg, H.H.C.M., and van Loon, L.J.C. (2007). Satellite cell content is specifically reduced in type II skeletal muscle

fibers in the elderly. American Journal of Physiology-Endocrinology and Metabolism *292*, E151–E157.

Wosczyna, M.N., Konishi, C.T., Perez Carbajal, E.E., Wang, T.T., Walsh, R.A., Gan, Q., Wagner, M.W., and Rando, T.A. (2019). Mesenchymal Stromal Cells Are Required for Regeneration and Homeostatic Maintenance of Skeletal Muscle. Cell Reports *27*, 2029-2035.e5.

Appendix 1

Artificial muscle stem cell niche project

The objective of this project was to engineer a cell-encapsulating poly(ethylene glycol) (PEG) hydrogel microbead system to control MuSC expansion in defined and scalable microenvironments with asymmetrically patterned biophysical and biomolecular features, which recapitulate the biphasic distribution and combinatorial complexity of the endogenous MuSC niche (i.e. artificial MuSC niche; Fig. S10). This approach recapitulates the biphasic distribution of features present in the MuSC niche through a serial encapsulation strategy in which two-phase PEG hydrogel beads are fabricated with microfluidics and with multiplexed molecular and mechanical features. The feature-set compositions that robustly promote stem-cell expansion can be evaluated in a combinatorial manner (Fig. S11, fluorophores bead barcoding).

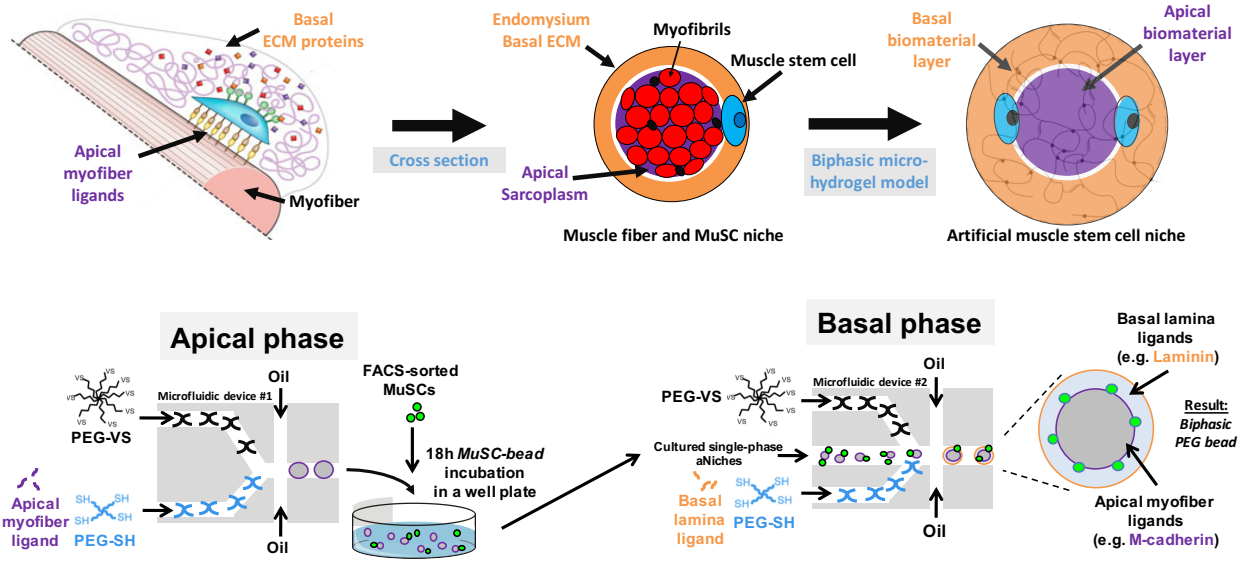


Figure S10 – Model for an artificial MuSC niche. *In vivo*, MuSCs reside at the interface between the apical myofiber plasma membrane and a basal lamina. MuSCs are therefore exposed to ligands from apical

myofibers and proteins from the ECM. I am proposing to recreate this bipolar geometry by a bi-phasic micro-hydrogel bead system where the inner and outer phases of the hydrogel respectively recapitulate the biology of the apical and basal sides. The first microfluidic device generates PEG micro-beads functionalized with apical myofiber ligands. Micro-beads are incubated with MuSCs until attachment of a single cell per micro-bead. Micro-beads with cells are then fed into a second microfluidic device that adds and additional layer of PEG functionalized with proteins from the basal lamina. Both microfluidic devices will be designed and fabricated to allow for mixing of PEG monomers by reacting 8-arm vinylsulfone (VS) and 4-arm thiol (SH) monomers, each at 10 kDa molecular weight. The result is a bi-phasic PEG micro-bead with spatially defined niches components that recapitulates the bipolar architecture of the native MuSC niche. Left niche schematic adapted from (Lutolf et al., 2009).

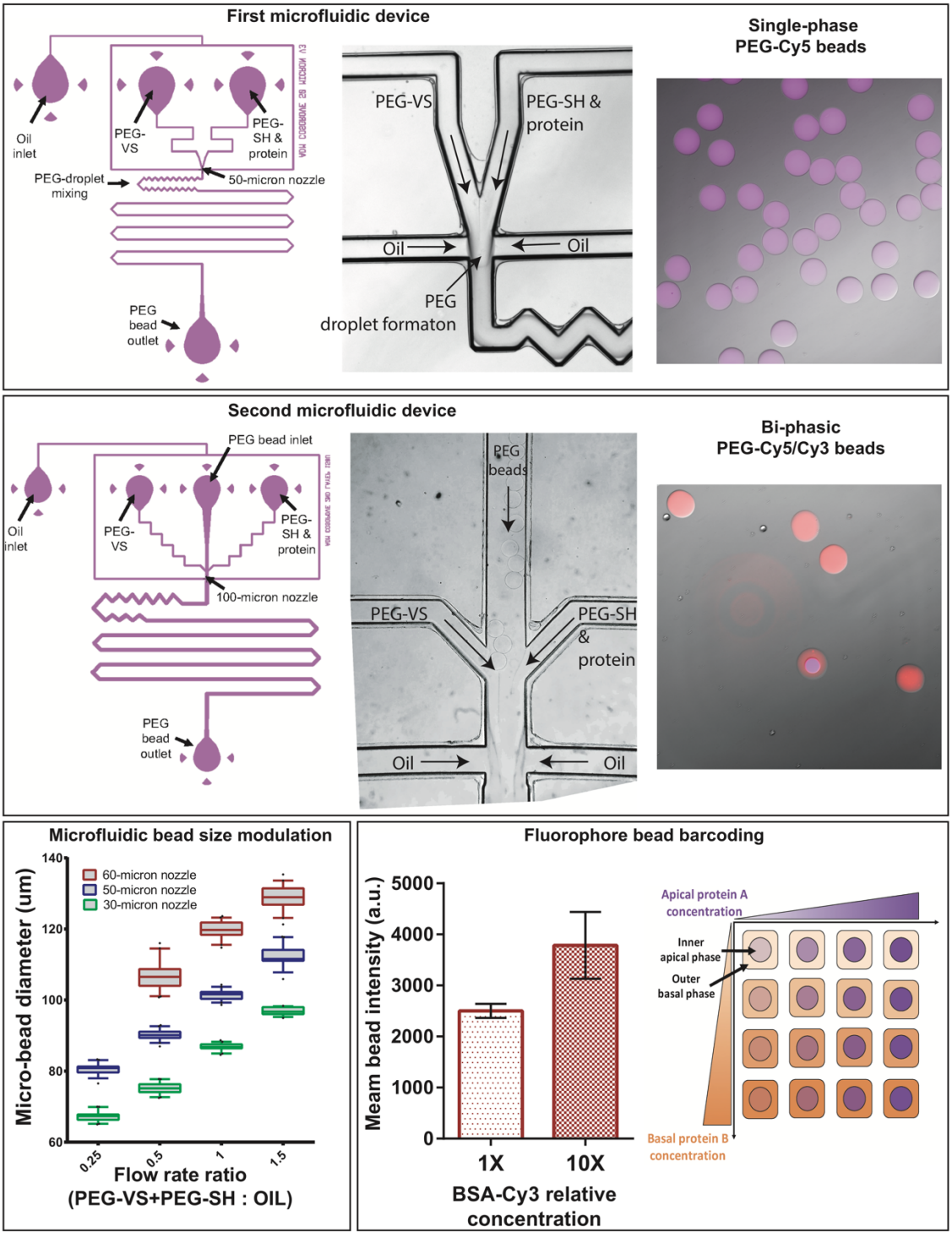


Figure S11 – Artificial MuSC niche platform characterization. Computer-assisted design of the first and second microfluidic devices and images of mono- and bi-phasic poly(ethylene) glycol (PEG) micro-beads with Cy3 and Cy5 fluorophores. Microfluidic bead size modulation using different device geometries and flow rate ratios. Fluorophore bead barcoding concept for combinatorial evaluation of bead composition.

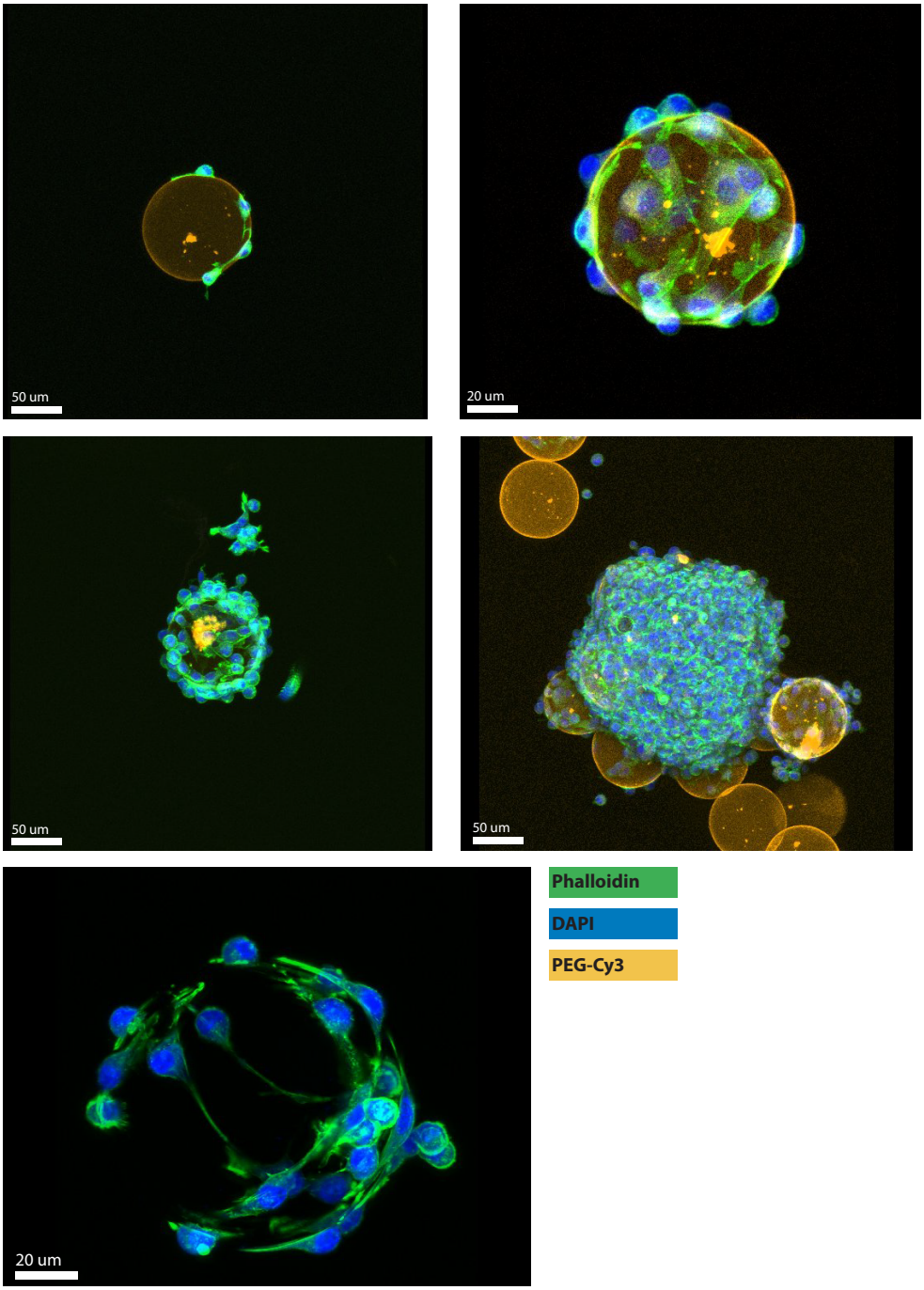


Figure S12 – Artificial MuSC niches with primary murine myoblasts. Confocal images of Cy3-labeled (to PEG hydrogel) and laminin-coated single-phase artificial niches micro-beads generated as described in Fig. S11 and cultured with primary murine myoblasts. Images were taken after 18 hours in culture. Cells are labeled with DAPI (nucleus) and Phalloidin (actin cytoskeleton).

Appendix 2

Biologically motivated biomaterial technologies for skeletal muscle cell-therapy‡‡

Abstract

Skeletal muscle is essential to locomotion, respiration, and metabolism, and can repair itself following minor damages, through the action of a resident muscle stem and progenitor cell population. After severe traumatic injuries, as well as in the context of degenerative disorders and aging, muscle repair is often insufficient but can be enhanced through a combination of cell- and biomaterials-based regenerative technologies. In this review, we summarize advances in biomaterials-based three- dimensional scaffolds designed with cellular, chemical, and physical components to aid muscle regeneration and enhance muscle recovery in various pathophysiological settings. We highlight a number of biologically motivated scaffold design parameters that affect choices between natural and synthetic polymers and how their properties impact myofiber alignment and maturation, vascularization, innervation, and immune- modulation. We also discuss how growth factor and adhesive peptide inclusions can enhance muscle stem and progenitor cell survival, self-renewal, and differentiation. Furthermore, we discuss injectable biomaterial scaffolds for the delivery of muscle cells, growth factors, and drugs into damaged muscles, and their therapeutic potential for treating muscular dystrophy, age-

‡‡This appendix is part of a review paper submitted to RSC's Biomaterial Science.

associated muscle atrophy, and traumatic volumetric muscle loss. These recent developments have poised the field to contribute a number of biomaterial technologies to enhance the treatment of muscle diseases.

A2.1. Introduction

Skeletal muscles are essential to many daily functions such as locomotion, respiration, and metabolism. Skeletal muscle tissue consists of aligned bundles of multinucleated striated myofibers that contract to generate force. Myofibers are surrounded by an extracellular matrix (ECM)-rich basal lamina, which provides myofiber anchorage and structural support, while coordinating their contractile forces. The ECM also provides a multitude of biochemical and biophysical signals that regulate muscle tissue homeostasis and repair. Further, muscle is extensively vascularized and innervated to receive the necessary nutrients and neural inputs that govern their function. The maintenance of this complex multi-cellular architecture is necessary to achieve muscle tissue contraction, and, to maintain function, its complete regeneration is required upon injury or damage.

Though muscle tissue is inherently regenerative, damage can exceed its regenerative capacity in the context of severe traumatic injuries and/or muscle-related disorders, and instead resolve in fibrosis, scaring, and diminished contractile function, impacting morbidity and mortality. Throughout life, healthy skeletal muscle has the ability to repair itself from minor injuries. These injuries are typically those induced by everyday activities, such as exercise-related strains, and rarely result in a major loss of muscle mass. However more severe damage, such as traumatic injury, infection, or surgery, can exceed the muscle's ability to regenerate itself. Similarly, muscle repair in the context of

aging-associated atrophy (sarcopenia) and muscle-wasting disorders such as Duchenne muscular dystrophy (DMD) is insufficient to maintain homeostasis (Blau et al., 2015).

These pathologies necessitate therapeutic interventions to aid in muscle repair and function. For example, reconstructive muscle surgery attempts to partially restore muscle contraction by transferring a healthy muscle from another part of the body but remain an invasive procedure with morbidity at the donor site (Gardetto et al., 2005). Furthermore, progressive muscle degeneration and defective regeneration can be ameliorated by delivering sufficient numbers of healthy muscle stem cells (MuSCs), through a variety of modes of delivery (Cosgrove et al., 2014; Hall et al., 2010; Price et al., 2007), or by enhancing the repair capacity of endogenous MuSCs (Brack et al., 2007; Chakkalakal et al., 2012; Shea et al., 2010) in mouse models.

The inability of muscle tissue to repair itself from trauma and degeneration has motivated the development of new therapeutic strategies leveraging on biomaterial technologies to enhance muscle cell-therapy. These biomaterials include engineered biocompatible scaffolds that attempt to reconstitute critical aspects of the muscle tissue environment. There are multiple modalities in which engineered biomaterials have been developed to enhance muscle cell and tissue engineering. First, given the rarity of MuSCs in muscle tissue and in biopsy samples (in which they represent ~2-5% of all mononuclear cells), there is considerable interest in using biomaterials to mimic in vitro the MuSCs niche microenvironment to expand MuSCs for subsequent cell- transplantation therapies (Cosgrove et al., 2009). For example, polymeric biomaterials based on poly(ethylene glycol) chemistry have been engineered as a more physiological cell culture substrate to

expand muscle stem and progenitor cells for transplantation (Cosgrove et al., 2014; Gilbert and Blau, 2011; Gilbert et al., 2010). Other versions of these biomaterial environments have attempted to directly generate functional muscle tissue in vitro with mature vascular beds and neuro-muscular junctions (NMJs) (Martin et al., 2015). Second, biomaterials have also been utilized as implantable or injectable scaffolds that provide temporary structural support to a muscle wound or defect, while also delivering myogenic progenitor cells, therapeutic proteins, and other factors to induce muscle regeneration in vivo (Bursac et al., 2015; Han et al., 2016). In this review, we summarize recent advances in biomaterial technologies, emphasizing on the design variables to address specific muscle diseases, for the therapeutic enhancement of muscle tissue regeneration.

A2.2. The biology of muscle regeneration

A2.2.1 Muscle stem cells, their microenvironment, and regenerative capabilities

Muscle stem cells (MuSCs) are critical to skeletal muscle regeneration throughout life (Cosgrove et al., 2009). MuSCs were first identified anatomically in 1961 as mononuclear cells residing at the periphery of muscle fibers, between the sarcolemma (myofiber membrane) and the basal lamina and were therefore given the name "*satellite cells*" (Mauro, 1961; Muir et al., 1965). This membrane bound compartment serves as an asymmetrically organized niche that governs MuSC function (Cosgrove et al., 2009; Kuang et al., 2008). In homeostasis, MuSCs are mostly quiescent and are characterized by the expression of the paired box protein 7 (Pax7) transcription factor (Grounds and McGeachie, 1987; Kuang et al., 2006; Snow, 1977). Upon muscle injury, they are alerted (Rodgers et al., 2014) into an activated state (Pax7+ Myf5+) (Jacobs et al., 1995) in

preparation for muscle repair. Following activation, they proliferate and some of their progeny (committed Pax7– Myf5+ MyoD+ myogenic progenitors called myoblasts) differentiate to generate fusion-competent Myogenin+ muscle cells (often called myocytes), while others self-renew to replenish the Pax7+ MuSC pool. Fused myofibers are characterized by expression of myosin heavy chain and other contractile machinery, are ensheathed by an extracellular basal lamina, and innervated by motor neurons at neuro-muscular junctions (NMJs) to provide contractile organization from the peripheral nervous system.

MuSCs have been identified as stem cells by lineage-tracing, cell-ablation, and transplantation experiments (Cosgrove et al., 2009). These studies have revealed that isolated or myofiber-associated Pax7+ MuSCs are also a potent source for treating muscle damage and age-related and degenerative disorders following transplantation. For example, Sacco et al. demonstrated that even a single MuSC isolated by FACS based on an α7-integrin+ CD34+ expression profile, is capable of extensive proliferation and myofiber repair when transplanted by direct intramuscular injection into the hindlimb muscles of pre-irradiated mice (Sacco et al., 2008). Similarly, MuSCs were used to improve the contractile function of muscle in the mdx dystrophic mouse and were also seen to repopulate the native MuSC niche (Cerletti et al., 2008). Transplantation of MuSCs in early adulthood can also ameliorate muscle atrophy in aging (Hall et al., 2010). The dysfunctional aged MuSC population can be rejuvenated in vitro through the combined effects of a pliant poly(ethylene glycol) hydrogel and p38 mitogen-activated protein kinase inhibition, and then transplanted to enhance muscle recovery in aged mice

(Cosgrove et al., 2014). These findings demonstrated that resident MuSCs are a promising cell source for treating muscle-related disorders and perform better than myoblast progenitors in long-term muscle regeneration evaluations (Montarras et al., 2005; Sacco et al., 2008).

Commonly used surface antigens for MuSC prospective isolation by flow cytometry include α7-integrin (Sherwood et al., 2004), β1-integrin (Cerletti et al., 2008; Kuang et al., 2007), CD34 (Beauchamp et al., 2000), CXCR4 (Cerletti et al., 2008), syndecan-3/4 (Cornelison et al., 2001), VCAM1 (Liu et al., 2015a), M-cadherin (Beauchamp et al., 2000; Irintchev et al., 1994), neural cell adhesion molecule (NCAM) (Bosnakovski et al., 2008; Capkovic et al., 2008; Irintchev et al., 1994), c-met (Cornelison and Wold, 1997), and ABGC2 (Tanaka et al., 2009), all of which have been used in various combinations and represent largely overlapping cell populations, at least in terms of Pax7 expression analysis (Maesner et al., 2016). Further, other cell populations including pluripotent stem cell-derived myogenic progenitors, muscle-derived stem cells, mesoangioblasts, pericytes, and mesenchymal stromal (stem) cells have been demonstrated to have myogenic potential and putative stem-cell-like properties (Chal et al., 2015; Darabi et al., 2012; Darabi et al., 2011; Tedesco et al., 2010). Many of these myogenic cell populations, including MuSCs, have been more extensively evaluated in mouse models than from human samples and their molecular identity and methods for their isolation vary between species. For example, unlike mouse MuSCs, which can be isolated based on α7-integrin+ CD34+ expression, human MuSCs are not CD34+ (Peault et al., 2007). Instead, the human MuSCs reside within a β1-integrin+ CD56+ (NCAM+) CD34– cell population that

may be further refined by EGFR+ and/or CD82+ expression (Alexander et al., 2016; Castiglioni et al., 2014; Charville et al., 2015; Pisani et al., 2010; Uezumi et al., 2016; Zheng et al., 2007), but additional delineation of the unipotent adult MuSC population is required given the heterogeneous lineage potential of these cells.

Biomaterial design strategies have focused on controlling the fate and function of muscle stem and progenitor cells have been motivated by mimicking aspects of their in vivo regulatory niche (Bursac et al., 2015; Cosgrove et al., 2009; Gilbert and Blau, 2011). The MuSC niche is an instructive and dynamic microenvironment that is composed of a combination of supporting cells, biochemical (soluble and ECM factors), and biophysical (mechanical, topological) components. These cues direct the behavior of MuSC to maintain homeostasis and respond to injuries. Moreover, MuSCs are exposed to “apical” signals localized to the myofiber plasma membrane (through transmembrane expression and/or secretion) and “basal” signals associated with the basal lamina (ECM and sequestered proteins, as well as signals secreted from interstitial cells) (Cosgrove et al., 2009; Kuang et al., 2008; Yin et al., 2013). This asymmetric arrangement of niche signals is critical for MuSC polarity and contribute to self-renewal and fate decisions.

The basal lamina is a network of ECM proteins and composed primarily of collagen types IV and VI, laminin, fibronectin, as well as proteoglycans and glycoproteins. The basal lamina also serves as a reservoir for growth factors derived from systemic, myofiber sources, or MuSC themselves, which stimulate MuSC survival, activation, and proliferation. On the apical side, myofibers also secrete or present factors that influence the fate of MuSCs. For example, myofibers secrete SDF-1, a ligand to the MuSC receptor

CXCR4, which promotes MuSC migration (Ratajczak et al., 2003). Myofibers express Delta, M-cadherin, and myomaker on their plasma membrane to facilitate myofiber-MuSC interactions. M-cadherin and myomaker promote cell adhesion as well as the maturation and fusion of myogenic progenitors (Irintchev et al., 1994; Millay et al., 2013), whereas Delta promotes MuSC self-renewal and quiescence (Conboy et al., 2003). Many of these biochemical and physical factors have been incorporated into biomaterials in order to recapitulate the biology of the MuSC niche and provide design components to control the fate of MuSCs in culture and in vivo.

A2.2.2 Skeletal muscle regeneration

Muscle regeneration following injury involves the coordinated activation, self-renewal, and maturation of MuSCs, subsequently myoblasts, as mediated by both a dynamic and complex inflammatory response and local niche factors (Bentzinger et al., 2013a). After injury-induced myofiber damage and death, the rupture of the plasma membrane releases factors that activates the complement cascade, which attracts mononuclear cells such as neutrophils within two hours of injury (Tidball, 2017). Neutrophils invade the damaged area and remove necrotic tissue and cellular debris by phagocytosis through the production of myeloperoxidase and free radicals. Neutrophils secrete cytokines and growth factors such as TNF-α, IL-1β, IL-6, and FGF2, which play an important role in activating MuSCs and coordinating immune cell responses (Tidball, 2017). After neutrophil invasion, two populations of macrophages infiltrate the damaged tissue to begin the repair phase. Pro-inflammatory M1-biased macrophages continue the phagocytosis of damaged myofibers for another 1-2 days post injury. In addition, activated

M1-biased macrophages secrete a palette of pro-inflammatory cytokines (e.g., IL-1β, IL-4, IL-12, and IGF-1) that further stimulate, directly or indirectly, MuSC activation and proliferation. Subsequently, resident fibro/adipogenic progenitor cells (FAPs) are also activated and start remodeling the ECM and signaling to MuSCs (Heredia et al., 2013; Lemos et al., 2015).

Days later, the macrophage population gradually shift to a M2-biased anti-inflammatory phenotype which promotes myogenic differentiation. M2-biased macrophages encourage regeneration by secreting anti-inflammatory factors, such as IL-10, TGF-β, IGF-1, and VEGF, which commit myoblasts to differentiate and fuse into new myofibers. Secreted factors from M2-biased macrophages can promote the differentiation of myoblasts in culture (Saclier et al., 2013). Persistent and/or exacerbated inflammatory responses due to aging and muscular dystrophy have been associated with increased MuSC activation and premature myoblast death and differentiation, which is detrimental to the homeostatic maintenance and efficient regeneration of muscle tissue (Bouche et al., 2014; Palacios et al., 2010; Saclier et al., 2013; Tidball, 2017; Tierney et al., 2014). Therefore, immunomodulatory approaches have been proposed to protect muscle cells from the harsh inflammatory environment of muscle regeneration (San Emeterio et al., 2017). Furthermore, re-vascularization and re-innervation also occur during the M2 regenerative phase. Complete muscle repair depends on the successful restoration of the myotendinous and neuro-muscular junctions, as well as the vascular network (Turner and Badylak, 2013). In the case of severe muscle damage, such as in volumetric muscle loss,

excessive fibrosis will impair these outcomes and result in ischemia, muscle atrophy, and scar tissue formation resulting in deficiencies in muscle mass and strength.

The myofiber ECM plays a critical role in coordinating the muscle regeneration process. Aging and degenerative diseases are often accompanied by the accumulation of fibrotic tissue that alter the ECM composition and consequently affect muscle stem and progenitor cell-fate outcomes. Fibronectin, for example, an essential component of the MuSC niche, is lost with aging, leading to reduced MuSC maintenance and muscle regeneration (Lukjanenko et al., 2016). Moreover, reduction of fibronectin expression in transplanted young-adult MuSCs decreases their regenerative ability (Tierney et al., 2016), while fibronectin supplementation in aged muscles aids in MuSC rejuvenation and improves repair (Lukjanenko et al., 2016). Similarly, collagen VI, another critical ECM component of the MuSC niche, is extensively remodeled during the repair process and regulates muscle stiffness (Urciuolo et al., 2013). Mutations in collagen VI genes cause Bethlem myopathy (BM) and Ullrich congenital muscular dystrophy (UCMD) (Lampe and Bushby, 2005). Furthermore, ablation of collagen VI (COL6A1) impairs muscle regeneration by affecting MuSC self-renewal (Urciuolo et al., 2013) and knocking-down COL6A1 in transplanted MuSCs reduces their regenerative ability (Tierney et al., 2016). In all, modulating ECM components within engineered biomaterials provide an attractive approach to control cell fate outcomes for muscle therapeutics.

A2.2.3 Muscle disorders, trauma, and their clinical needs

Duchenne muscular dystrophy. Duchenne muscular dystrophy (DMD) is a genetic muscle-wasting disease that affects 1 in about 5000 boys and results in accelerated

myofiber damage and MuSC exhaustion (Guiraud et al., 2015; Rahimov and Kunkel, 2013). DMD is caused by various mutations in the Dystrophin gene, a large 427-kDa structural protein complex that anchors the cytoskeleton of myofibers to its surrounding ECM. Dystrophin mutations leads to a loss of membrane integrity and causes chronic muscle cell injury and death. MuSCs are therefore constantly enlisted to repair degenerating muscles and over time they exhaust. In patients with DMD, the MuSC-mediated repair is insufficient as MuSCs are progressively depleted while systemic muscle degeneration persists (Sacco et al., 2010). Multiple cell- and gene-therapy approaches have been evaluated though none offer a cure. Currently available treatments largely focus on cellular and genetic approaches to restore expression of the dystrophin protein in the skeletal muscle tissues of DMD patients, with the rationale that dystrophin replacement will ameliorate or cure the disease (Athanasopoulos et al., 2004; Blau, 2008; Bogdanovich et al., 2004; Goyenvalle et al., 2011; Rahimov and Kunkel, 2013). Myoblasts have been evaluated as cell therapy candidates in prior DMD clinical trials but were unable to contribute to the MuSC population in vivo and thus provide only transient repair benefit and dystrophin expression (Gussoni et al., 1997; Gussoni et al., 1992; Rinaldi and Perlingeiro, 2014; Sacco et al., 2008; Webster and Blau, 1990).

Recent advances towards cell-based DMD therapies have aimed to utilize MuSC or pluripotent stem cells-derived myogenic progenitors in combination with gene-correction strategies to provide long-term therapy to dystrophic muscles. Current challenges in these approaches include expanding patient-specific MuSCs, while conserving their Pax7+ stem-cell identity, and delivering them in therapeutically sufficient

numbers to multiple affected muscles. First, biomaterial technologies can help address these challenges by better recapitulating the composition of the niche to support more effective expansion of dystrophic MuSCs in vitro. Second, biomaterials may also aid clinical transplantation procedures, such through minimally invasive, syringe-injected biomaterial carriers to deliver "corrected" MuSCs or other myogenic precursors to affected dystrophic tissues. Third, biomaterial scaffolds could act as immunomodulators (Franz et al., 2011) to protect the injected cells from the aberrant immune responses within dystrophic muscles, while also providing the essential survival, migration, and differentiation signals.

Volumetric muscle loss. Traumatic injuries can also significantly diminish the contractile function of affected muscles. Volumetric muscle loss (VML), for instance, can result from traumatic muscle injuries or the surgical removal of tumors, in which critically large portions of muscle tissue are lost (Grogan et al., 2011). In VML, the muscle tissue and biological signals required for regeneration are damaged and/or lost such that muscle cannot repair itself autonomously. Reconstructive surgery is often required to heal muscles after VML, involving transplantation of muscle tissue from unaffected locations elsewhere in the patient's body. The success of this procedure is however limited, in part due to its invasiveness (morbidity at donor site), limited availability of healthy muscle in some critical cases, and the excessive generation of scar tissue. To address this issue, both in vitro and in vivo approaches have been explored. In vitro approaches consist in designing a biomaterial system that promotes the assembly and the maturation of contractile myofibers ready for transplantation. Biomaterials, however, may also hold

therapeutic potential as injectable, in situ polymerizable scaffolds that can provide temporary structural support to the wound, while modulating the inflammatory response and delivering the necessary biological cues for regeneration.

Aging-associated muscle atrophy. Muscle atrophy and dysfunction in the elderly is prevalent and affects more than 50% of the US population over 80 years of age (Doherty, 2003). Aging-associated muscle atrophy, also termed sarcopenia, severely impacts mortality, mobility and quality-of-life. Sarcopenia notably affects swallowing due pharyngeal muscle atrophy, vision due to ocular muscle atrophy, and urinary incontinence due to ureteral sphincter muscle damage and atrophy. Sarcopenia is exacerbated following injury and/or prolonged hospitalization. For example, muscle atrophy occurs in ~60% of aged patients after hip fracture (Di Monaco et al., 2011). There is an absence of efficacious drug or cell therapies to reduce muscle wasting in these diverse scenarios, despite more than a decade of clinical trials (Lin and Lue, 2012). MuSC dysfunction also contributions to the regenerative decline and atrophy in aging (Blau et al., 2015), and strategies to enhance and/or augment MuSC function in homeostasis and repair provide promise for ameliorating these pathologies (Cosgrove et al., 2014).

A2.3. Biomaterial systems for skeletal muscle regenerative engineering

Biomaterials derived from natural and synthetic sources have been investigated extensively in engineered microenvironments to control the behavior of cells and to assist in the repair of damaged tissues (Lutolf and Hubbell, 2005; Seliktar, 2012). Biomaterials can provide a three-dimensional scaffold designed to mimic the biochemical and

biophysical attributes of the native tissue ECM, and therefore provide physiological substrates for culturing cells in vitro and evaluate the influence of ECM mechanics and components on cell-fate decisions. In the context of skeletal muscle, biomaterials have been used to stimulate the maturation of myoblasts into assembled and aligned contractile myofiber bundles for in vitro modeling and for in vivo tissue engineering (Bursac et al., 2015; Han et al., 2016). Biomaterials have also been used to study regulators within the MuSC niche and to generate MuSCs for transplantation therapies for degenerative diseases (Cosgrove et al., 2014). Finally, biomaterials have been used as bulking agents for VML by providing a transient scaffold to promote muscle regeneration in situ (Qazi et al., 2015).

A2.3.1 Naturally-derived biomaterials for muscle cell culture and transplantation

Biomaterials can be derived from protein and/or glycoprotein components purified from natural extracellular matrices, produced synthetically through a mixture of polymeric and peptide sequences, or combine natural and synthetic components. Naturally-derived biomaterials are typically composed of ECM biopolymers such as collagen, laminin, fibronectin, hyaluronic acid, alginate, and chitosan. These biopolymers possess adhesion sequences that can promote cell attachment, enhance cell survival and migration, and stimulate differentiation-associated signal transduction pathways (reviewed in (Swinehart and Badylak, 2016)). Moreover, ECM biomaterials can have advantage features of biocompatibility, bioactivity (in terms of cell and/or protein adhesion), and degradability.

For example, collagen type I and IV hydrogel scaffolds contain many cell-binding epitopes recognized by integrin heterodimers and promote the differentiation and

maturation of C2C12 myoblasts in vitro (Smith et al., 2012). Laminin has often been combined to synthetic polymers to culture MuSCs. MuSCs and myoblasts adhere to laminin, a major component of their niche, via α7/β1-integrin heterodimers (Crawley et al., 1997). Alginate, though not an ECM component, is natural biomaterial that gelates in the presence of divalent ions. Alginate can be chemically modified to have different physical and biochemical properties. Alginate has been functionalized with cell adhesive peptides and chemically crosslinked to exhibit different stiffness and degradability rate. In the context of muscle regeneration, Liu et al. developed RGD-modified alginate microbeads to encapsulate human umbilical cord-derived mesenchymal stem cells for muscle tissue engineering (Liu et al., 2012). These alginate microbeads have a fast degradation rate (releasing cells within 8 days), which was shown to improve cell viability and differentiation in culture. Similarly, fibrin-based biomaterials have been used intensively as surgical bioadhesives. In a similar study, human umbilical cord- derived mesenchymal stem cells were encapsulated inside a fibrin gel containing fast- degradable microbeads to create macropores that enhance cell spreading, viability, and improve myogenic differentiation (Liu et al., 2013). Jana et al designed a chitosan scaffold with controlled topography in order to pre-align muscle cells in vitro (Jana et al., 2013). The microstructures of the scaffold, such as the pore size and orientation, were modulated by different chitosan concentration, freezing temperature and processing gradients, and facilitated for myofiber alignment from C2C12 cells in vitro.

Hyaluronic acid (HA), another ECM constituent, has been used as a clinical biomaterial for many wound-healing applications. HA biomaterials have been used to

deliver primary mouse myoblasts to damaged muscles (Wang et al., 2009). Rossi et al developed an in situ photo-polymerizable HA hydrogel for MuSC delivery applications (Rossi et al., 2011). After injection partially ablated mouse tibialis anterior muscles, this HA biomaterial scaffold, containing as few as 250 MuSCs, contributed to regeneration of muscle vasculature and innervation and aided in the formation of new MuSC niches. Matrigel, a mixture of basement membrane proteins containing laminin, collagen IV and other ECM proteins, and decellularized tissues have been used as biomaterials for in vitro culture and in vivo delivery of muscle progenitor cells. Decellularized tissues may better reconstitute both the complex proteomic content and architecture of the native ECM. Recently, Rao et al. engineered an injectable ECM hydrogel derived from decellularized skeletal muscle tissue to deliver myoblasts to ischemic muscles, resulting in improved cell survival and neovascularization (Rao et al., 2017).

A2.3.3 Synthetic biomaterials for muscle cell culture and transplantation

From a biomaterial design perspective, however, the intrinsic biological complexity of the natural polymers may make them difficult to fully characterize, thus restricting the ability to independently fine-tune their biomolecular or biophysical properties. For example, it is challenging to control the density of adhesive ligands, cytokine binding, and mechanical properties independently in a native collagen scaffold. Further, ECM biomaterials are challenging to purify and process and can be affected by lot-to-lot variability. In addition, when transplanted in vivo, these materials often elicit an immunological response, may be rejected, and can potentially be source of viral infection (Sobolev et al., 2009).

On the other hand, synthetic polymers can be used to generate biomaterial scaffolds with a greater control of biochemical and physical properties. These polymers are often biologically (Levit et al., 2013; Salimath and Garcia, 2016) inert and, unlike naturally derived biomaterials, require the addition of ECM proteins or peptides to support cells. Common synthetic polymers used for muscle tissue engineering are poly(glycolic acid) (PGA), poly(lactic-co-glycolic acid) (PLGA), poly(ethylene glycol) (PEG), poly(lactic acid) (PLLA), and poly(ethylene glycol) (PEG) (Grasman et al., 2015; Qazi et al., 2015). PEG is a frequently used as injectable biomaterial due to the diversity of biochemical functionalization strategies that have enabled controlled polymerization in situ and degradation rates. For example, PEG can be functionalized with vinyl sulfone or maleimides moieties that react almost exclusively by Michael-type addition with minimal effect on the surrounding cells and tissues (Lutolf and Hubbell, 2003; Salimath and Garcia, 2016). PEG can be functionalized with transglutaminase peptide sequences that are recognized by the coagulation factor XIII and allow for rapid polymerization after injection in vivo (Ehrbar et al., 2007). PEG can also be functionalized to permit cell adhesion, and laminin, fibrinogen, and RGD peptides sequences have been used to promote cell survival and myogenic differentiation for in vivo muscle therapies (Fuoco et al., 2012; Levit et al., 2013; Salimath and Garcia, 2016). Furthermore, PEG hydrogels can be modified to include matrix metalloproteinase-sensitive sequences to adjust its degradation properties and allow cells to migrate and remodel the biomaterial to promote tissue regeneration (Ehrbar et al., 2007). Together, these modifications provide PEG hydrogel biomaterials highly tunable feature sets.

A2.3.4 Physical properties of biomaterials for muscle tissue formation

The physical properties of a biomaterial scaffold, such as its stiffness (Young's modulus), porosity (pore size), and topography (microstructural features), are important design considerations for stem cell applications (Murphy et al., 2014). For muscle cells, the biomaterial stiffness impacts MuSC self-renewal behaviors and myoblast maturation (Cosgrove et al., 2009). Healthy adult muscle tissue has a bulk Young's modulus of ~10-12 kPa whereas geriatric and fibrotic tissue is stiffer (>18 kPa). Biomaterials with elasticities in similar ranges have been shown to either maintain MuSC culture in vitro using 12 kPa PEG hydrogels or to induce differentiation at higher stiffnesses (Cosgrove et al., 2014; Gilbert et al., 2010). The modulus can be modulated by varying monomer and/or cross-linking concentrations to induce mechanosensitive cellular responses. Intriguingly, more complicated strategies have been developed to dynamically change biomaterial moduli in situ by through photo-polymerization, photo- cross-linking, and/or photo-degradation (DeForest and Anseth, 2011; Tse and Engler, 2010). Other properties such as porosity and topography are similarly important design parameters. Pore size can affect nutrient, protein, and oxygen transport and regulate the formation of a new vasculature, which is essential for millimeter-thick tissue constructs. Topography generally refers to the biomaterial micro-architecture and can provide cues to stimulate cells to align and fuse into elongated myofibers (Ostrovidov et al., 2014).

Topologically-directed myoblast alignment promotes myogenic differentiation and myofiber maturation. For example, Huang et al. have shown that electrospun PLA/PLGA nanofibers can provide contact guidance to direct C2C12 into long myofibers with improved striation indicative of a mature myofiber phenotype (Huang et al., 2006). Page

et al. developed fibrin micro-threads that stimulated the longitudinal growth of muscle cells into fiber-like bundles and regenerated large defects with limited fibrosis when transplanted in vivo (Page et al., 2011).

A2.3.5 Biochemical functionalization of biomaterials to control muscle cell fates

Overcoming drug delivery and immune response challenges are critical for effective growth factors or small biomolecule therapies via intramuscular injection. One reason is the muscle regenerative environment that can limit the lifespan and bioactivity of the injected factors. In addition, such factors rapidly diffuse away from the site of injection and injury. The inability to sustain a desired local concentration over the muscle repair time course limits the potential of these therapeutic molecules. Biomaterials offer the opportunity to present drugs or biomolecules with controlled spatial and temporal presentation and improve bioavailability.

Biomaterials containing growth factors, cytokines, or other small drugs have been engineered to facilitate the survival of transplanted cells either by stimulating their regenerative functions or by modulating the in vivo healing environment. Insulin-like growth factor-1 (IGF-1) has been shown to be a central regulator of muscle repair (Musaro et al., 2001). Transgenic expression of IGF-1 in muscle was also demonstrated to accelerate functional muscle recovery by down-regulating pro-inflammatory cytokines and enhancing macrophage recruitment (Pelosi et al., 2007). The pro-angiogenic vascular endothelial growth factor (VEGF) is another important candidate that stimulate the formation of a neo-vasculature and improve in vivo regeneration. VEGF delivered to muscles has been shown restore perfusion in ischemic hindlimbs (Borselli et al., 2010).

Taken together, IGF-1 and VEGF was recently used to improve the survival of myoblasts using a shape-memory injectable scaffold (Wang et al., 2014). Wang et al. showed that by combining myoblasts with IGF-1 and VGEF the scaffold significantly reduced fibrosis, enhanced engraftment, improved contractile function, and enhanced vascularization in mouse tibialis anterior muscles following by myotoxin injury. In another study, Pumberger et al. created synthetic niches from porous alginate cyrogels that supported the adhesion of mesenchymal stromal cells and contained recombinant IGF-1 and VEGF (Pumberger et al., 2016). The authors showed that the paracrine factors are secreted by the mesenchymal stromal cells (MSCs) within functionalized alginate cyrogels following stimulation by IGF-1 and VEGF contribute to muscle repair and enhanced contractile.

Factors present in the native MuSC niche have also been utilized in biomaterial scaffolds to control the fate of MuSCs towards regenerative fates. For example, Notch activation through Delta ligands instruct MuSCs to asymmetrically self-renew both producing a cell committed to myogenic differentiation and one with conserved stem cell property (Cosgrove et al., 2009). Delta-IgGs have been used to expand a progenitor population called muscle-derived stem cells to enhance their engraftment in vivo (Parker et al., 2012). Furthermore, the other niche factor Wnt7a/Fzd7 has been shown to stimulate symmetric-self renewal and might be a potent candidate to stimulate MuSC expansion in vitro (Le Grand et al., 2009). Rudnicki and colleagues showed that overexpressing Wnt7a and fibronectin dramatically stimulated symmetric self-renewal of MuSCs cells in vitro (Bentzinger et al., 2013b) and in vivo (von Maltzahn et al., 2012).

Biomaterials have also been used to protect transplanted cells form the harsh inflammatory wound environment as well as immune modulators to enhance muscle tissue regeneration (Mao and Mooney, 2015). Rationally controlling the healing response by manipulating macrophage polarization, the adaptive immune response, or suppressing specific inflammatory pathways may stimulate repair and minimize the accumulation of scar tissue. In a mouse model for VML, Sadtler et al. screened for different types of tissue-derived ECM scaffolds to repair the defect (Sadtler et al., 2016). They identified that cardiac- and bone-derived scaffolds were best for tissue regeneration due to their ability to recruit immune cells, in particular CD4-positive T- helper 2 cells. The T-helper 2 response was described by an increase in IL-4 in the scaffold, an important cytokine for coordinating in muscle healing, by subsequent macrophage polarization, and functional tissue regeneration 6 weeks post injury.

A2.3.6 Muscle stem and progenitor cells for muscle tissue engineering

Many different cell populations – such as myoblasts, MuSCs, mesenchymal stem cells, pericytes, and mesoangioblasts – can participate in the regeneration of muscle and have been investigated for therapy. Biomaterial constructs have been used to deliver muscle stem and progenitor cells and the guidance for their regenerative contributions, by both directing their differentiation and promoting their retention into a reservoir stem cell population. Cells types not directly participating in myogenesis have also been incorporated into biomaterials as paracrine regulators of myofiber regeneration.

The delivery of cells to a damaged muscle is however challenging for a couple of reasons. For instance, due to the harsh inflammatory environment injected cells face

when transplanted, bolus injection has often lead to poor survival and engraftment. Additionally, cells delivered by bolus injection are only locally distributed and cannot effectively regenerate large damages such as in VML. Various types of biomaterial scaffold have therefore been developed with the objective of protecting the cells and providing them with signals to survive, engraft, and efficiently participate in the repair process.

The cell source is critically important for therapeutic success. Allogenic cell sources may very likely to be an immunological concern and be rejected by the patient. Though autologous sources are preferred, it is often difficult to obtain from a minimally-invasive biopsy large quantities of MuSCs for culture and transplantation given their rarity in vivo and difficulty to expand ex vivo. For instance, culturing MuSCs on rigid substrates, such as a plastic dish, can significantly affect their regenerative potential (Gilbert et al., 2010). Culture techniques have therefore shifted towards more physiological niche-like biomaterials that enhance MuSCs survival, conserve their identity, and stimulate their expansive self-renewal fate (Cosgrove et al., 2009). For some muscle diseases such as DMD, transplanting functional and genetically corrected MuSCs may be the most efficacious strategy for any long-term benefit if cell-delivery challenges can be overcome. For DMD patients, transplanting dystrophin-corrected MuSCs will allow the native muscle tissue to repair itself using a healthy source of stem cells. A possible strategy would be to transplant “artificial niches” as biomaterials encapsulating MuSCs with the necessary cues to drive the repopulation of the diseased tissue with a new permanent and healthy pool of mutation-free MuSCs.

More abundant but less well-defined autologous cell types suited for engineering muscle repair would include myoblasts and muscle progenitors derived from embryonic stem cells (ECSs) or induced pluripotent stem cells (iPSCs) sources. For example, Darabi et al. showed that expressing PAX7 in ECSs and iPCs induced myogenic differentiation, which upon transplantation into the dystrophic mdx mouse resulted in the repopulation of the MuSC niche and restored muscle function (Darabi et al., 2012). This study highlights the potential of human ECSs/iPSCs for muscle therapies, especially for DMD where large quantities of cells need to be manufactured and corrected in vitro. New advances in myogenic differentiation from iPS sources provide promising approaches to meet these goals (Chal et al., 2015).

Other adult cell types, such as mesenchymal stem cells (MSCs), can also be used to indirectly stimulate muscle regeneration through paracrine secretion (Ferrari et al., 1998). Though they do not necessary contribute to the formation of new myofibers, MSCs secrete factors that contribute to muscle repair (Sassoli et al., 2012). MSCs can be incorporated into injectable biomaterials in order to leverage their paracrine effect (Mao et al., 2017).

Mesoangioblasts and pericytes have also been used for muscle repair, though their in vivo contribution is more limited compared to stem cells. Fuoco et al. developed an injectable PEG-fibrinogen hydrogel-mesoangioblasts formulation that photo-polymerizes in situ (Fuoco et al., 2012). They demonstrated that the hydrogel promoted myofiber differentiation in vitro and that upon intramuscular injection the mesoangioblasts had enhanced engraftment and decrease apoptosis. The hydrogel was tested in a tibialis

anterior freeze injury model in a dystrophic (aSGKO/SCID) mice where histological evaluation displayed enhanced integration of the mesoangioblasts into the regenerative myofibers compared to controls. Pericytes have also been used as a myogenic cell source. The same group used a similar PEG-fibrinogen hydrogel in order to rejuvenate old pericytes (Fuoco et al., 2012). They showed in vitro that the hydrogel could restore the diminished myogenic and vasculorgenic capability of aged pericytes, hence allowing these cells to then produce vascularized muscle in vivo indistinguishable from that using younger cells.

A2.3.6 Addressing vascularization and innervation challenges

Vascularization. Skeletal muscle cells demand oxygen for metabolic and contractile activities. Trauma can result in the destruction of vasculature, hence severely limiting the supply of blood to the injury site. Revascularization is critical for engineering thick muscle tissue in vitro and therefore biomaterials need to possess an architecture conductive to blood vessel formation. In addition, biomaterials that contain also proangiogenic factors and/or endothelial cells, which has been shown to quickly restore perfusion and limit ischemia. Borselli et al. developed an injectable alginate scaffold that locally delivered VEGF and IGF1 to an ischemic muscle (Borselli et al., 2010). They demonstrated that the synergy between these two growth factors promoted angiogenesis, protecting the muscle tissue from hypoxia, and promoted satellite cell activation and myogenesis within 3 weeks of injury.

Innervation. Myofiber contraction is elicited through the neuro-muscular junction (NMJ) through acetylcholine receptors on the muscle cell membrane. Aging-associated muscle atrophy and weakness is associated with dysregulation of the NMJ leading to a loss of motor neurons (Jang and Van Remmen, 2011). MuSCs and myofiber cells help regulate the NMJ and changes in their function in aging contribute to this loss (Liu et al., 2015b). In the absence of full re-innervation, tissue engineered muscle constructs do not maintain mature contractile myofibers long-term. Various biomaterials have been suggested to address the challenge of re-innervation. For example, Wang et al. showed a fibrin hydrogel combined with agrin and laminin enhanced the clustering of acetylcholine receptors on muscle fibers, a first step towards the formation of a new NMJ (Wang et al., 2013). VEGF has also been identified to promote neuronal regeneration by stimulating directly neurons and glial cells (Wang et al., 2013). Shvartsman et al. developed an alginate scaffold to deliver VEGF in a localized a sustained fashion in muscle ischemic injuries (Shvartsman et al., 2014). They showed that VEGF increased the expression of nerve growth factors in ischemic muscle, which supported both the restoration of blood supply and reinnervation through axonal regeneration.

A2.3.7 Recent advances in millimeter-scale myofiber-mimicking scaffolds

A number of recent reports indicate that an advantageous strategy for muscle tissue engineering and muscle cell delivery therapy involves generating millimeter-scale long- aspect ratio scaffolds that resemble mature myofibers in their architecture. Mozetic et al. recently described an additive manufacturing strategy to create muscle cell scaffolds with precise control of cell shape and scaffold architecture (Mozetic et al., 2017). The

authors used a thermosensitive pluronic/alginate hydrogel to bioprint 3D myoblast scaffolds and showed they enhanced 3D alignment along the axis prescribed within the scaffold. Quarta et al. used a collagen-based extrusion method to generate ECM-rich myofiber mimics for delivering muscle stem cells to injured mice and showed that these scaffolds could enhance control of quiescence and activation of MuSCs in vitro and in vivo (Quarta et al., 2016). Further they demonstrated that similar scaffolds could enhance the co-delivery of MuSCs and other supporting (non-myogenic) muscle cell populations for in vivo delivery (Quarta et al., 2017). In a related strategy, Constantini et al. developed a 3D hydrogel printing approach to generate unidirectionally patterned PEG-fibrinogen-alginate scaffolds, which enhanced myoblast alignment and fusion in vitro and in vivo (Costantini et al., 2017).

A2.4. Concluding Remarks

Biologically-inspired biomaterials technologies offer great promise for aiding a diverse array of muscle stem and progenitor cell-based therapies. Numerous design principles have been established using both synthetic and natural polymer systems and the field is poised to engineer hybrid technologies to optimize muscle regenerative therapies. Future challenges still to be addressed include the translation of these approaches to the recently revealed human MuSC population (Alexander et al., 2016; Xu et al., 2015) and scaling their implementation for clinical applications and addressing regulatory development milestones.

5. References

Alexander, M.S., Rozkalne, A., Colletta, A., Spinazzola, J.M., Johnson, S., Rahimov, F., Meng, H., Lawlor, M.W., Estrella, E., Kunkel, L.M., et al. (2016). CD82 Is a Marker for Prospective Isolation of Human Muscle Satellite Cells and Is Linked to Muscular Dystrophies. Cell Stem Cell 19, 800-807.

Athanasopoulos, T., Graham, I.R., Foster, H., and Dickson, G. (2004). Recombinant adeno-associated viral (rAAV) vectors as therapeutic tools for Duchenne muscular dystrophy (DMD). Gene Ther 11 Suppl 1, S109-121.

Beauchamp, J.R., Heslop, L., Yu, D.S., Tajbakhsh, S., Kelly, R.G., Wernig, A., Buckingham, M.E., Partridge, T.A., and Zammit, P.S. (2000). Expression of CD34 and Myf5 defines the majority of quiescent adult skeletal muscle satellite cells. J Cell Biol 151, 1221-1234.

Bentzinger, C.F., Wang, Y.X., Dumont, N.A., and Rudnicki, M.A. (2013a). Cellular dynamics in the muscle satellite cell niche. EMBO reports 14, 1062-1072.

Bentzinger, C.F., Wang, Y.X., von Maltzahn, J., Soleimani, V.D., Yin, H., and Rudnicki, M.A. (2013b). Fibronectin Regulates Wnt7a Signaling and Satellite Cell Expansion. Cell stem cell 12, 75-87.

Blau, H.M. (2008). Cell therapies for muscular dystrophy. N Engl J Med 359, 1403-1405.

Blau, H.M., Cosgrove, B.D., and Ho, A.T. (2015). The central role of muscle stem cells in regenerative failure with aging. Nat Med 21, 854-862.

Bogdanovich, S., Perkins, K.J., Krag, T.O., and Khurana, T.S. (2004). Therapeutics for Duchenne muscular dystrophy: current approaches and future directions. J Mol Med 82, 102-115.

Borselli, C., Storrie, H., Benesch-Lee, F., Shvartsman, D., Cezar, C., Lichtman, J.W., Vandenburgh, H.H., and Mooney, D.J. (2010). Functional muscle regeneration with combined delivery of angiogenesis and myogenesis factors. Proc Natl Acad Sci U S A 107, 3287-3292.

Bosnakovski, D., Xu, Z., Li, W., Thet, S., Cleaver, O., Perlingeiro, R.C., and Kyba, M. (2008). Prospective isolation of skeletal muscle stem cells with a Pax7 reporter. Stem cells (Dayton, Ohio) 26, 3194-3204.

Bouche, M., Munoz-Canoves, P., Rossi, F., and Coletti, D. (2014). Inflammation in muscle repair, aging, and myopathies. BioMed research international 2014, 821950.

Brack, A.S., Conboy, M.J., Roy, S., Lee, M., Kuo, C.J., Keller, C., and Rando, T.A. (2007). Increased Wnt signaling during aging alters muscle stem cell fate and increases fibrosis. Science 317, 807-810.

Bursac, N., Juhas, M., and Rando, T.A. (2015). Synergizing Engineering and Biology to Treat and Model Skeletal Muscle Injury and Disease. Annual review of biomedical engineering 17, 217-242.

Capkovic, K.L., Stevenson, S., Johnson, M.C., Thelen, J.J., and Cornelison, D.D. (2008). Neural cell adhesion molecule (NCAM) marks adult myogenic cells committed to differentiation. Exp Cell Res 314, 1553-1565.

Castiglioni, A., Hettmer, S., Lynes, M.D., Rao, T.N., Tchessalova, D., Sinha, I., Lee, B.T., Tseng, Y.H., and Wagers, A.J. (2014). Isolation of Progenitors that Exhibit Myogenic/Osteogenic Bipotency In Vitro by Fluorescence-Activated Cell Sorting from Human Fetal Muscle. Stem cell reports 2, 92-106.

Cerletti, M., Jurga, S., Witczak, C.A., Hirshman, M.F., Shadrach, J.L., Goodyear, L.J., and Wagers, A.J. (2008). Highly efficient, functional engraftment of skeletal muscle stem cells in dystrophic muscles. Cell 134, 37-47.

Chakkalakal, J.V., Jones, K.M., Basson, M.A., and Brack, A.S. (2012). The aged niche disrupts muscle stem cell quiescence. Nature 490, 355-360.

Chal, J., Oginuma, M., Al Tanoury, Z., Gobert, B., Sumara, O., Hick, A., Bousson, F., Zidouni, Y., Mursch, C., Moncuquet, P., et al. (2015). Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. Nat Biotechnol 33, 962-969.

Charville, G.W., Cheung, T.H., Yoo, B., Santos, P.J., Lee, G.K., Shrager, J.B., and Rando, T.A. (2015). Ex Vivo Expansion and In Vivo Self-Renewal of Human Muscle Stem Cells. Stem cell reports 5, 621-632.

Conboy, I.M., Conboy, M.J., Smythe, G.M., and Rando, T.A. (2003). Notch-mediated restoration of regenerative potential to aged muscle. Science 302, 1575-1577.

Cornelison, D.D., Filla, M.S., Stanley, H.M., Rapraeger, A.C., and Olwin, B.B. (2001). Syndecan-3 and syndecan-4 specifically mark skeletal muscle satellite cells and are implicated in satellite cell maintenance and muscle regeneration. Dev Biol 239, 79-94.

Cornelison, D.D., and Wold, B.J. (1997). Single-cell analysis of regulatory gene expression in quiescent and activated mouse skeletal muscle satellite cells. Dev Biol 191, 270-283.

Cosgrove, B.D., Gilbert, P.M., Porpiglia, E., Mourkioti, F., Lee, S.P., Corbel, S.Y., Llewellyn, M.E., Delp, S.L., and Blau, H.M. (2014). Rejuvenation of the muscle stem cell population restores strength to injured aged muscles. Nat Med 20, 255-264.

Cosgrove, B.D., Sacco, A., Gilbert, P.M., and Blau, H.M. (2009). A home away from home: challenges and opportunities in engineering in vitro muscle satellite cell niches. Differentiation 78, 185-194.

Costantini, M., Testa, S., Mozetic, P., Barbetta, A., Fuoco, C., Fornetti, E., Tamiro, F., Bernardini, S., Jaroszewicz, J., Swieszkowski, W., et al. (2017). Microfluidic-enhanced 3D bioprinting of aligned myoblast-laden hydrogels leads to functionally organized myofibers in vitro and in vivo. Biomaterials 131, 98-110.

Crawley, S., Farrell, E.M., Wang, W., Gu, M., Huang, H.Y., Huynh, V., Hodges, B.L., Cooper, D.N., and Kaufman, S.J. (1997). The alpha7beta1 integrin mediates adhesion and migration of skeletal myoblasts on laminin. Exp Cell Res 235, 274-286.

Darabi, R., Arpke, R.W., Irion, S., Dimos, J.T., Grskovic, M., Kyba, M., and Perlingeiro, R.C. (2012). Human ES- and iPS-derived myogenic progenitors restore DYSTROPHIN and improve contractility upon transplantation in dystrophic mice. Cell Stem Cell 10, 610-619.

Darabi, R., Santos, F.N., Filareto, A., Pan, W., Koene, R., Rudnicki, M.A., Kyba, M., and Perlingeiro, R.C. (2011). Assessment of the myogenic stem cell compartment following transplantation of Pax3/Pax7-induced embryonic stem cell-derived progenitors. Stem cells (Dayton, Ohio) 29, 777-790.

DeForest, C.A., and Anseth, K.S. (2011). Cytocompatible click-based hydrogels with dynamically tunable properties through orthogonal photoconjugation and photocleavage reactions. Nat Chem 3, 925-931.

Di Monaco, M., Vallero, F., Di Monaco, R., and Tappero, R. (2011). Prevalence of sarcopenia and its association with osteoporosis in 313 older women following a hip fracture. Arch Gerontol Geriatr 52, 71-74.

Doherty, T.J. (2003). Invited review: Aging and sarcopenia. J Appl Physiol (1985) 95, 1717-1727.

Ehrbar, M., Rizzi, S.C., Hlushchuk, R., Djonov, V., Zisch, A.H., Hubbell, J.A., Weber, F.E., and Lutolf, M.P. (2007). Enzymatic formation of modular cell-instructive fibrin analogs for tissue engineering. Biomaterials 28, 3856-3866.

Ferrari, G., Cusella-De Angelis, G., Coletta, M., Paolucci, E., Stornaiuolo, A., Cossu, G., and Mavilio, F. (1998). Muscle regeneration by bone marrow-derived myogenic progenitors. Science 279, 1528-1530.

Franz, S., Rammelt, S., Scharnweber, D., and Simon, J.C. (2011). Immune responses to implants - a review of the implications for the design of immunomodulatory biomaterials. Biomaterials 32, 6692-6709.

Fuoco, C., Salvatori, M.L., Biondo, A., Shapira-Schweitzer, K., Santoleri, S., Antonini, S., Bernardini, S., Tedesco, F.S., Cannata, S., Seliktar, D., et al. (2012). Injectable polyethylene glycol-fibrinogen hydrogel adjuvant improves survival and differentiation of transplanted mesoangioblasts in acute and chronic skeletal-muscle degeneration. Skelet Muscle 2, 24.

Gardetto, A., Raschner, C., Schoeller, T., Pavelka, M.L., and Wechselberger, G. (2005). Rectus femoris muscle flap donor-site morbidity. Br J Plast Surg 58, 175-182.

Gilbert, P.M., and Blau, H.M. (2011). Engineering a stem cell house into a home. Stem Cell Res Ther 2, 3.

Gilbert, P.M., Havenstrite, K.L., Magnusson, K.E., Sacco, A., Leonardi, N.A., Kraft, P., Nguyen, N.K., Thrun, S., Lutolf, M.P., and Blau, H.M. (2010). Substrate elasticity regulates skeletal muscle stem cell self-renewal in culture. Science 329, 1078-1081.

Goyenvalle, A., Seto, J.T., Davies, K.E., and Chamberlain, J. (2011). Therapeutic approaches to muscular dystrophy. Hum Mol Genet 20, R69-78.

Grasman, J.M., Zayas, M.J., Page, R.L., and Pins, G.D. (2015). Biomimetic scaffolds for regeneration of volumetric muscle loss in skeletal muscle injuries. Acta Biomater 25, 2-15.

Grogan, B.F., Hsu, J.R., and Skeletal Trauma Research, C. (2011). Volumetric muscle loss. J Am Acad Orthop Surg 19 Suppl 1, S35-37.

Grounds, M.D., and McGeachie, J.K. (1987). A model of myogenesis in vivo, derived from detailed autoradiographic studies of regenerating skeletal muscle, challenges the concept of quantal mitosis. Cell Tissue Res 250, 563-569.

Guiraud, S., Aartsma-Rus, A., Vieira, N.M., Davies, K.E., van Ommen, G.J., and Kunkel, L.M. (2015). The Pathogenesis and Therapy of Muscular Dystrophies. Annu Rev Genomics Hum Genet 16, 281-308.

Gussoni, E., Blau, H.M., and Kunkel, L.M. (1997). The fate of individual myoblasts after transplantation into muscles of DMD patients. Nat Med 3, 970-977.

Gussoni, E., Pavlath, G.K., Lanctot, A.M., Sharma, K.R., Miller, R.G., Steinman, L., and Blau, H.M. (1992). Normal dystrophin transcripts detected in Duchenne muscular dystrophy patients after myoblast transplantation. Nature 356, 435-438.

Hall, J.K., Banks, G.B., Chamberlain, J.S., and Olwin, B.B. (2010). Prevention of muscle aging by myofiber-associated satellite cell transplantation. Sci Transl Med 2, 57ra83.
Han, W.M., Jang, Y.C., and Garcia, A.J. (2016). Engineered matrices for skeletal muscle satellite cell engraftment and function. Matrix Biol.

Heredia, J.E., Mukundan, L., Chen, F.M., Mueller, A.A., Deo, R.C., Locksley, R.M., Rando, T.A., and Chawla, A. (2013). Type 2 innate signals stimulate fibro/adipogenic progenitors to facilitate muscle regeneration. Cell 153, 376-388.

Huang, N.F., Patel, S., Thakar, R.G., Wu, J., Hsiao, B.S., Chu, B., Lee, R.J., and Li, S. (2006). Myotube assembly on nanofibrous and micropatterned polymers. Nano Lett 6, 537-542.

Irintchev, A., Zeschnigk, M., Starzinski-Powitz, A., and Wernig, A. (1994). Expression pattern of M-cadherin in normal, denervated, and regenerating mouse muscles. Dev Dyn 199, 326-337.

Jacobs, S.C., Wokke, J.H., Bar, P.R., and Bootsma, A.L. (1995). Satellite cell activation after muscle damage in young and adult rats. Anat Rec 242, 329-336.

Jana, S., Cooper, A., and Zhang, M. (2013). Chitosan scaffolds with unidirectional microtubular pores for large skeletal myotube generation. Adv Healthc Mater 2, 557-561.

Jang, Y.C., and Van Remmen, H. (2011). Age-associated alterations of the neuromuscular junction. Exp Gerontol 46, 193-198.

Kuang, S., Charge, S.B., Seale, P., Huh, M., and Rudnicki, M.A. (2006). Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. J Cell Biol 172, 103-113.

Kuang, S., Gillespie, M.A., and Rudnicki, M.A. (2008). Niche regulation of muscle satellite cell self-renewal and differentiation. Cell Stem Cell 2, 22-31.

Kuang, S., Kuroda, K., Le Grand, F., and Rudnicki, M.A. (2007). Asymmetric self-renewal and commitment of satellite stem cells in muscle. Cell 129, 999-1010.

Lampe, A.K., and Bushby, K.M. (2005). Collagen VI related muscle disorders. J Med Genet 42, 673-685.

Le Grand, F., Jones, A.E., Seale, V., Scime, A., and Rudnicki, M.A. (2009). Wnt7a activates the planar cell polarity pathway to drive the symmetric expansion of satellite stem cells. Cell Stem Cell 4, 535-547.

Lemos, D.R., Babaeijandaghi, F., Low, M., Chang, C.K., Lee, S.T., Fiore, D., Zhang, R.H., Natarajan, A., Nedospasov, S.A., and Rossi, F.M. (2015). Nilotinib reduces muscle fibrosis in chronic muscle injury by promoting TNF-mediated apoptosis of fibro/adipogenic progenitors. Nat Med.

Levit, R.D., Landazuri, N., Phelps, E.A., Brown, M.E., Garcia, A.J., Davis, M.E., Joseph, G., Long, R., Safley, S.A., Suever, J.D., et al. (2013). Cellular encapsulation enhances cardiac repair. J Am Heart Assoc 2, e000367.

Lin, C.S., and Lue, T.F. (2012). Stem cell therapy for stress urinary incontinence: a critical review. Stem Cells Dev 21, 834-843.

Liu, J., Xu, H.H., Zhou, H., Weir, M.D., Chen, Q., and Trotman, C.A. (2013). Human umbilical cord stem cell encapsulation in novel macroporous and injectable fibrin for muscle tissue engineering. Acta Biomater 9, 4688-4697.

Liu, J., Zhou, H., Weir, M.D., Xu, H.H., Chen, Q., and Trotman, C.A. (2012). Fast-degradable microbeads encapsulating human umbilical cord stem cells in alginate for muscle tissue engineering. Tissue Eng Part A 18, 2303-2314.

Liu, L., Cheung, T.H., Charville, G.W., and Rando, T.A. (2015a). Isolation of skeletal muscle stem cells by fluorescence-activated cell sorting. Nat Protoc 10, 1612-1624.

Liu, W., Wei-LaPierre, L., Klose, A., Dirksen, R.T., and Chakkalakal, J.V. (2015b). Inducible depletion of adult skeletal muscle stem cells impairs the regeneration of neuromuscular junctions. Elife 4.

Lukjanenko, L., Jung, M.J., Hegde, N., Perruisseau-Carrier, C., Migliavacca, E., Rozo, M., Karaz, S., Jacot, G., Schmidt, M., Li, L., et al. (2016). Loss of fibronectin from the aged stem cell niche affects the regenerative capacity of skeletal muscle in mice. Nat Med 22, 897-905.

Lutolf, M.P., and Hubbell, J.A. (2003). Synthesis and physicochemical characterization of end-linked poly(ethylene glycol)-co-peptide hydrogels formed by Michael-type addition. Biomacromolecules 4, 713-722.

Lutolf, M.P., and Hubbell, J.A. (2005). Synthetic biomaterials as instructive extracellular microenvironments for morphogenesis in tissue engineering. Nat Biotechnol 23, 47-55.

Maesner, C.C., Almada, A.E., and Wagers, A.J. (2016). Established cell surface markers efficiently isolate highly overlapping populations of skeletal muscle satellite cells by fluorescence-activated cell sorting. Skelet Muscle 6, 35.

Mao, A.S., and Mooney, D.J. (2015). Regenerative medicine: Current therapies and future directions. Proc Natl Acad Sci U S A 112, 14452-14459.

Mao, A.S., Shin, J.W., Utech, S., Wang, H., Uzun, O., Li, W., Cooper, M., Hu, Y., Zhang, L., Weitz, D.A., et al. (2017). Deterministic encapsulation of single cells in thin tunable microgels for niche modelling and therapeutic delivery. Nat Mater 16, 236-243.

Martin, N.R., Passey, S.L., Player, D.J., Mudera, V., Baar, K., Greensmith, L., and Lewis, M.P. (2015). Neuromuscular Junction Formation in Tissue-Engineered Skeletal Muscle Augments Contractile Function and Improves Cytoskeletal Organization. Tissue Eng Part A 21, 2595-2604.

Mauro, A. (1961). Satellite cell of skeletal muscle fibers. J Biophys Biochem Cytol 9, 493-495.

Millay, D.P., O'Rourke, J.R., Sutherland, L.B., Bezprozvannaya, S., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2013). Myomaker is a membrane activator of myoblast fusion and muscle formation. Nature 499, 301-305.

Montarras, D., Morgan, J., Collins, C., Relaix, F., Zaffran, S., Cumano, A., Partridge, T., and Buckingham, M. (2005). Direct isolation of satellite cells for skeletal muscle regeneration. Science 309, 2064-2067.

Mozetic, P., Giannitelli, S.M., Gori, M., Trombetta, M., and Rainer, A. (2017). Engineering muscle cell alignment through 3D bioprinting. Journal of Biomedical Materials Research Part A, n/a-n/a.

Muir, A.R., Kanji, A.H., and Allbrook, D. (1965). The structure of the satellite cells in skeletal muscle. J Anat 99, 435-444.

Murphy, W.L., McDevitt, T.C., and Engler, A.J. (2014). Materials as stem cell regulators. Nat Mater 13, 547-557.

Musaro, A., McCullagh, K., Paul, A., Houghton, L., Dobrowolny, G., Molinaro, M., Barton, E.R., Sweeney, H.L., and Rosenthal, N. (2001). Localized Igf-1 transgene expression sustains hypertrophy and regeneration in senescent skeletal muscle. Nature genetics 27, 195-200.

Ostrovidov, S., Hosseini, V., Ahadian, S., Fujie, T., Parthiban, S.P., Ramalingam, M., Bae, H., Kaji, H., and Khademhosseini, A. (2014). Skeletal muscle tissue engineering:

methods to form skeletal myotubes and their applications. Tissue Eng Part B Rev 20, 403-436.

Page, R.L., Malcuit, C., Vilner, L., Vojtic, I., Shaw, S., Hedblom, E., Hu, J., Pins, G.D., Rolle, M.W., and Dominko, T. (2011). Restoration of skeletal muscle defects with adult human cells delivered on fibrin microthreads. Tissue Eng Part A 17, 2629-2640.

Palacios, D., Mozzetta, C., Consalvi, S., Caretti, G., Saccone, V., Proserpio, V., Marquez, V.E., Valente, S., Mai, A., Forcales, S.V., et al. (2010). TNF/p38alpha/polycomb signaling to Pax7 locus in satellite cells links inflammation to the epigenetic control of muscle regeneration. Cell Stem Cell 7, 455-469.

Parker, M.H., Loretz, C., Tyler, A.E., Duddy, W.J., Hall, J.K., Olwin, B.B., Bernstein, I.D., Storb, R., and Tapscott, S.J. (2012). Activation of Notch signaling during ex vivo expansion maintains donor muscle cell engraftment. Stem cells (Dayton, Ohio) 30, 2212-2220.

Peault, B., Rudnicki, M., Torrente, Y., Cossu, G., Tremblay, J.P., Partridge, T., Gussoni, E., Kunkel, L.M., and Huard, J. (2007). Stem and progenitor cells in skeletal muscle development, maintenance, and therapy. Mol Ther 15, 867-877.

Pelosi, L., Giacinti, C., Nardis, C., Borsellino, G., Rizzuto, E., Nicoletti, C., Wannenes, F., Battistini, L., Rosenthal, N., Molinaro, M., et al. (2007). Local expression of IGF-1 accelerates muscle regeneration by rapidly modulating inflammatory cytokines and chemokines. FASEB J 21, 1393-1402.

Pisani, D.F., Dechesne, C.A., Sacconi, S., Delplace, S., Belmonte, N., Cochet, O., Clement, N., Wdziekonski, B., Villageois, A.P., Butori, C., et al. (2010). Isolation of a highly myogenic CD34-negative subset of human skeletal muscle cells free of adipogenic potential. Stem cells (Dayton, Ohio) 28, 753-764.

Price, F.D., Kuroda, K., and Rudnicki, M.A. (2007). Stem cell based therapies to treat muscular dystrophy. Biochim Biophys Acta 1772, 272-283.

Pumberger, M., Qazi, T.H., Ehrentraut, M.C., Textor, M., Kueper, J., Stoltenburg-Didinger, G., Winkler, T., von Roth, P., Reinke, S., Borselli, C., et al. (2016). Synthetic niche to modulate regenerative potential of MSCs and enhance skeletal muscle regeneration. Biomaterials 99, 95-108.

Qazi, T.H., Mooney, D.J., Pumberger, M., Geissler, S., and Duda, G.N. (2015). Biomaterials based strategies for skeletal muscle tissue engineering: existing technologies and future trends. Biomaterials 53, 502-521.

Quarta, M., Brett, J.O., DiMarco, R., De Morree, A., Boutet, S.C., Chacon, R., Gibbons, M.C., Garcia, V.A., Su, J., Shrager, J.B., et al. (2016). An artificial niche preserves the quiescence of muscle stem cells and enhances their therapeutic efficacy. Nat Biotechnol 34, 752-759.

Quarta, M., Cromie, M.J., Chacon, R., Blonigan, J., Garcia, V., Akimenko, I., Hamer, M., Paine, P., Stok, M., Shrager, J.B., et al. (2017). Bioengineered constructs combined with exercise enhance stem cell-mediated treatment of volumetric muscle loss. Nature communications 8.

Rahimov, F., and Kunkel, L.M. (2013). The cell biology of disease: cellular and molecular mechanisms underlying muscular dystrophy. J Cell Biol 201, 499-510.

Rao, N., Agmon, G., Tierney, M.T., Ungerleider, J.L., Braden, R.L., Sacco, A., and Christman, K.L. (2017). Engineering an Injectable Muscle-Specific Microenvironment for Improved Cell Delivery Using a Nanofibrous Extracellular Matrix Hydrogel. ACS Nano 11, 3851-3859.

Ratajczak, M.Z., Majka, M., Kucia, M., Drukala, J., Pietrzkowski, Z., Peiper, S., and Janowska-Wieczorek, A. (2003). 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 (Dayton, Ohio) 21, 363-371.

Riddle, E.S., Bender, E.L., Thalacker-Mercer, A.E., (2018). Transcript profile distinguishes variability in human myogenic progenitor cell expansion capacity. Physiol Genomics 50, 817–827.

Rinaldi, F., and Perlingeiro, R.C. (2014). Stem cells for skeletal muscle regeneration: therapeutic potential and roadblocks. Transl Res 163, 409-417.

Rodgers, J.T., King, K.Y., Brett, J.O., Cromie, M.J., Charville, G.W., Maguire, K.K., Brunson, C., Mastey, N., Liu, L., Tsai, C.R., et al. (2014). mTORC1 controls the adaptive transition of quiescent stem cells from G0 to G(Alert). Nature 510, 393-396.

Rossi, C.A., Flaibani, M., Blaauw, B., Pozzobon, M., Figallo, E., Reggiani, C., Vitiello, L., Elvassore, N., and De Coppi, P. (2011). In vivo tissue engineering of functional skeletal muscle by freshly isolated satellite cells embedded in a photopolymerizable hydrogel. FASEB J 25, 2296-2304.

Sacco, A., Doyonnas, R., Kraft, P., Vitorovic, S., and Blau, H.M. (2008). Self-renewal and expansion of single transplanted muscle stem cells. Nature 456, 502-506.

Sacco, A., Mourkioti, F., Tran, R., Choi, J., Llewellyn, M., Kraft, P., Shkreli, M., Delp, S., Pomerantz, J.H., Artandi, S.E., et al. (2010). Short telomeres and stem cell exhaustion model Duchenne muscular dystrophy in mdx/mTR mice. Cell 143, 1059-1071.

Saclier, M., Cuvellier, S., Magnan, M., Mounier, R., and Chazaud, B. (2013). Monocyte/macrophage interactions with myogenic precursor cells during skeletal muscle regeneration. The FEBS journal 280, 4118-4130.

Sadtler, K., Estrellas, K., Allen, B.W., Wolf, M.T., Fan, H., Tam, A.J., Patel, C.H., Luber, B.S., Wang, H., Wagner, K.R., et al. (2016). Developing a pro-regenerative biomaterial scaffold microenvironment requires T helper 2 cells. Science 352, 366-370.

Salimath, A.S., and Garcia, A.J. (2016). Biofunctional hydrogels for skeletal muscle constructs. J Tissue Eng Regen Med 10, 967-976.

San Emeterio, C.L., Olingy, C.E., Chu, Y., and Botchwey, E.A. (2017). Selective recruitment of non-classical monocytes promotes skeletal muscle repair. Biomaterials 117, 32-43.

Sassoli, C., Zecchi-Orlandini, S., and Formigli, L. (2012). Trophic actions of bone marrow-derived mesenchymal stromal cells for muscle repair/regeneration. Cells 1, 832-850.

Seliktar, D. (2012). Designing cell-compatible hydrogels for biomedical applications. Science 336, 1124-1128.

Shea, K.L., Xiang, W., LaPorta, V.S., Licht, J.D., Keller, C., Basson, M.A., and Brack, A.S. (2010). Sprouty1 regulates reversible quiescence of a self-renewing adult muscle stem cell pool during regeneration. Cell Stem Cell 6, 117-129.

Sherwood, R.I., Christensen, J.L., Conboy, I.M., Conboy, M.J., Rando, T.A., Weissman, I.L., and Wagers, A.J. (2004). Isolation of adult mouse myogenic progenitors: functional heterogeneity of cells within and engrafting skeletal muscle. Cell 119, 543-554.

Shvartsman, D., Storrie-White, H., Lee, K., Kearney, C., Brudno, Y., Ho, N., Cezar, C., McCann, C., Anderson, E., Koullias, J., et al. (2014). Sustained delivery of VEGF maintains innervation and promotes reperfusion in ischemic skeletal muscles via NGF/GDNF signaling. Mol Ther 22, 1243-1253.

Smith, A.S., Passey, S., Greensmith, L., Mudera, V., and Lewis, M.P. (2012). Characterization and optimization of a simple, repeatable system for the long term in vitro culture of aligned myotubes in 3D. Journal of cellular biochemistry 113, 1044-1053.

Snow, M.H. (1977). Myogenic cell formation in regenerating rat skeletal muscle injured by mincing. I. A fine structural study. Anat Rec 188, 181-199.

Sobolev, O., Stern, P., Lacy-Hulbert, A., and Hynes, R.O. (2009). Natural killer cells require selectins for suppression of subcutaneous tumors. Cancer Res 69, 2531-2539.

Swinehart, I.T., and Badylak, S.F. (2016). Extracellular matrix bioscaffolds in tissue remodeling and morphogenesis. Dev Dyn 245, 351-360.

Tanaka, K.K., Hall, J.K., Troy, A.A., Cornelison, D.D., Majka, S.M., and Olwin, B.B. (2009). Syndecan-4-expressing muscle progenitor cells in the SP engraft as satellite cells during muscle regeneration. Cell Stem Cell 4, 217-225.

Tedesco, F.S., Dellavalle, A., Diaz-Manera, J., Messina, G., and Cossu, G. (2010). Repairing skeletal muscle: regenerative potential of skeletal muscle stem cells. J Clin Invest 120, 11-19.

Tidball, J.G. (2017). Regulation of muscle growth and regeneration by the immune system. Nat Rev Immunol 17, 165-178.

Tierney, M.T., Aydogdu, T., Sala, D., Malecova, B., Gatto, S., Puri, P.L., Latella, L., and Sacco, A. (2014). STAT3 signaling controls satellite cell expansion and skeletal muscle repair. Nat Med 20, 1182-1186.

Tierney, M.T., Gromova, A., Sesillo, F.B., Sala, D., Spenle, C., Orend, G., and Sacco, A. (2016). Autonomous Extracellular Matrix Remodeling Controls a Progressive Adaptation in Muscle Stem Cell Regenerative Capacity during Development. Cell reports 14, 1940-1952.

Tse, J.R., and Engler, A.J. (2010). Preparation of hydrogel substrates with tunable mechanical properties. Current protocols in cell biology / editorial board, Juan S Bonifacino [et al] Chapter 10, Unit 10 16.

Turner, N.J., and Badylak, S.F. (2013). Biologic scaffolds for musculotendinous tissue repair. Eur Cell Mater 25, 130-143.

Uezumi, A., Nakatani, M., Ikemoto-Uezumi, M., Yamamoto, N., Morita, M., Yamaguchi, A., Yamada, H., Kasai, T., Masuda, S., Narita, A., et al. (2016). Cell-Surface Protein Profiling Identifies Distinctive Markers of Progenitor Cells in Human Skeletal Muscle. Stem cell reports 7, 263-278.

Urciuolo, A., Quarta, M., Morbidoni, V., Gattazzo, F., Molon, S., Grumati, P., Montemurro, F., Tedesco, F.S., Blaauw, B., Cossu, G., et al. (2013). Collagen VI

regulates satellite cell self-renewal and muscle regeneration. Nature communications 4, 1964.

von Maltzahn, J., Renaud, J.M., Parise, G., and Rudnicki, M.A. (2012). Wnt7a treatment ameliorates muscular dystrophy. Proc Natl Acad Sci U S A 109, 20614-20619.

Wang, L., Cao, L., Shansky, J., Wang, Z., Mooney, D., and Vandenburgh, H. (2014). Minimally invasive approach to the repair of injured skeletal muscle with a shape-memory scaffold. Mol Ther 22, 1441-1449.

Wang, L., Shansky, J., and Vandenburgh, H. (2013). Induced formation and maturation of acetylcholine receptor clusters in a defined 3D bio-artificial muscle. Mol Neurobiol 48, 397-403.

Wang, W., Fan, M., Zhang, L., Liu, S.H., Sun, L., and Wang, C.Y. (2009). Compatibility of hyaluronic acid hydrogel and skeletal muscle myoblasts. Biomed Mater 4, 025011.

Webster, C., and Blau, H.M. (1990). Accelerated age-related decline in replicative life-span of Duchenne muscular dystrophy myoblasts: implications for cell and gene therapy. Somat Cell Mol Genet 16, 557-565.

Xu, X., Wilschut, K.J., Kouklis, G., Tian, H., Hesse, R., Garland, C., Sbitany, H., Hansen, S., Seth, R., Knott, P.D., et al. (2015). Human Satellite Cell Transplantation and Regeneration from Diverse Skeletal Muscles. Stem cell reports 5, 419-434.

Yin, H., Price, F., and Rudnicki, M.A. (2013). Satellite cells and the muscle stem cell niche. Physiol Rev 93, 23-67.

Zheng, B., Cao, B., Crisan, M., Sun, B., Li, G., Logar, A., Yap, S., Pollett, J.B., Drowley, L., Cassino, T., et al. (2007). Prospective identification of myogenic endothelial cells in human skeletal muscle. Nat Biotechnol 25, 1025-1034.

Appendix 3

Material and resources

S3.1 Related to Chapter 3

Antibody (clone)	Vendor	Metal tag
CD8a (53-6.7)	Fluidigm	153Eu
CD9 (KMC8)	Fluidigm	158Gd
CD11b (M1/70)	Fluidigm	143Nd
CD3e (145-2C11)	Fluidigm	152Sm
CD4 (RM4-5)	Fluidigm	145Nd
CD45.2 (104)	Fluidigm	147Sm
CD206/MMR (C068C2)	Fluidigm	169Ti
CD80 (16-10A1)	Fluidigm	171Yb
CD31/PECAM-1 (390)	Fluidigm	165Ho
CD140a (APA5)	Fluidigm	148Nd
CD140b (APB5)	Fluidigm	151Eu
CyclinB1 (GNS-1)	Fluidigm	164Dy
Cell-ID™ IdU	Fluidigm	
CD15 (mc-480)	Lederer Lab, Harvard Medical School	159Tb
CD29 (HMB1-1)	Lederer Lab, Harvard Medical School	175Lu
CD34 (MEC14.7)	Lederer Lab, Harvard Medical School	166Er
CD106/VCAM-1 (429)	Lederer Lab, Harvard Medical School	155Gd
Integrin alpha 7	R&D Systems	161Dy
Pax7	R&D Systems	156Gd
Myogenin	R&D Systems	167Er
Myosin Heavy Chain	R&D Systems	172Yb
VE-Cadherin	R&D Systems	145Nd
Laminin alpha 1	R&D Systems	160Gd
Sca-1/Ly6	R&D Systems	170Er
C1q R1/CD93	R&D Systems	149Sm
CX3CR1	R&D Systems	154Sm
Syndecan-2/CD362	R&D Systems	162Dy
MyoD (5.8A)	Novus Biologicals	150Nd
CD82	Thermo Fisher Scientific	144Nd
COL1A1 (3G3)	Santa Cruz Biotechnology	141Pr
Syndecan-3 (G-2)	Santa Cruz Biotechnology	174Yd

Scx (N-term)	Abgent	163Dy
Fibromodulin (3E9D10)	Proteintech	142Nd
CD138/Sdc1 (281-2)	BD	176Yb
Syndecan-4 (KY/8.2)	BD	172Yd

Table S1 – CyTOF reagent panel.

Protein	Final concentration	Vendor and catalogue #
Laminin (matrix protein)	3.125 $\mu g/cm^2$	Thermo Fisher Scientific 23017015
Fibronectin (matrix protein)	5 $\mu g/cm^2$	Sigma-Aldrich FC010-10MG
Tenascin C (matrix protein)	2 $\mu g/cm^2$	R&D Systems 3358-TC-050
CXCL12/SDF-1 (paracrine protein)	100 ng/mL	R&D Systems 460-SD-010/CF
CCL5/RANTES (paracrine protein)	50 ng/mL	R&D Systems 478-MR-025/CF
THBS1 (paracrine protein)	5 μg/mL	R&D Systems 7859-TH-050
PF4/CXCL4 (paracrine protein)	2 μg/mL	R&D Systems 595-P4-025
TGFβ1 (paracrine protein)	25 ng/mL	R&D Systems 7666-MB-005/CF
MDK (paracrine protein)	150 ng/mL	R&D Systems 9760-MD-050
RSPO3 (paracrine protein)	100 ng/mL	R&D Systems 4120-RS-025/CF
FGF2 (paracrine protein)	1.25 ng/mL	Thermo Fisher Scientific PHG0264
Syndecan-1/CD138 (nAb)	9 μg/mL	R&D Systems MAB2966
Syndecan-2/CD362 (nAb)	9 μg/mL	R&D Systems AF6585
Syndecan-4 (nAb)	9 μg/mL	BD Biosciences 550350

Table S2 – Proteins and neutralizing antibodies (nAbs) and their concentration used for *in vitro* MuSC paracrine ligand testing.

S3.2 Related to Chapter 4

Reagents		
Dispase II (neutral protease, grade II)	Sigma-Aldrich	04942078001
Collagenase D, from Clostridium histolyticum	Sigma-Aldrich	11088866001

Commercial kits		
Chromium Single Cell 3' Library & Gel Bead Kit *v2*	10X Genomics	CG00052 (protocol)
Chromium Single Cell 3' Library & Gel Bead Kit *v3*	10X Genomics	CG000183 (protocol)

Deposited data		
Human ligand-receptor database	Ramilowski et al., 2015	https://www.ncbi.nlm.nih.gov/pubmed/26198319
Human scRNAseq dataset	This paper	GSE143704

Software packages and algorithms		
Cell Ranger 3.1.0 (July 24, 2019)	10X Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest
Seurat 3.1.0	Stuart et al., 2019b	https://github.com/satijalab/seurat
Scanorama (online version as of 2019-11-19)	Hie et al., 2019	https://github.com/brianhie/scanorama
Harmony (online version as of 2019-11-19)	Korsunsky et al., 2019	https://github.com/immunogenomics/harmony
biomaRt 2.43.1 (online version as of 2020-01-08)	Durinck et al., 2009	https://bioconductor.org/packages/release/bioc/html/biomaRt.html
Gene Set Enrichment Analysis (4.0.3)	Subramanian et al., 2005	http://software.broadinstitute.org/gsea/index.jsp
Ingenuity Pathway Analysis (IPA, 2019-08-30)	QUIAGEN	https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/

Table S3 – Reagents and resources used for the project presented in Chapter 4.

Bibliography

Abou-Khalil, R., Mounier, R., and Chazaud, B. (2010). Regulation of myogenic stem cell behaviour by vessel cells: The “ménage à trois” of satellite cells, periendothelial cells and endothelial cells. Cell Cycle *9*, 892–896.

Addison, O., Marcus, R.L., LaStayo, P.C., and Ryan, A.S. (2014). Intermuscular Fat: A Review of the Consequences and Causes. International Journal of Endocrinology 2014, 1–11.

Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat Methods *14*, 1083–1086.

Almada, A.E., and Wagers, A.J. (2016). Molecular circuitry of stem cell fate in skeletal muscle regeneration, ageing and disease. Nature Reviews Molecular Cell Biology 17, 267–279.

Alexander, M.S., Rozkalne, A., Colletta, A., Spinazzola, J.M., Johnson, S., Rahimov, F., Meng, H., Lawlor, M.W., Estrella, E., Kunkel, L.M., et al. (2016). CD82 Is a Marker for Prospective Isolation of Human Muscle Satellite Cells and Is Linked to Muscular Dystrophies. Cell Stem Cell *19*, 800–807.

Alles, J., Karaiskos, N., Praktiknjo, S.D., Grosswendt, S., Wahle, P., Ruffault, P.-L., Ayoub, S., Schreyer, L., Boltengagen, A., Birchmeier, C., et al. (2017). Cell fixation and preservation for droplet-based single-cell transcriptomics. BMC Biol *15*, 44.

Andersen, D.C., Laborda, J., Baladron, V., Kassem, M., Sheikh, S.P., and Jensen, C.H. (2013). Dual role of delta-like 1 homolog (DLK1) in skeletal muscle development and adult muscle regeneration. Development 140, 3743.

Arnold, L., Henry, A., Poron, F., Baba-Amer, Y., van Rooijen, N., Plonquet, A., Gherardi, R.K., and Chazaud, B. (2007). Inflammatory monocytes recruited after skeletal muscle injury switch into antiinflammatory macrophages to support myogenesis. The Journal of Experimental Medicine *204*, 1057–1069.

Becht, E., McInnes, L., Healy, J., Dutertre, C.A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol.

Baitsch, D., Bock, H.H., Engel, T., Telgmann, R., Muller-Tidow, C., Varga, G., Bot, M., Herz, J., Robenek, H., von Eckardstein, A., et al. (2011). Apolipoprotein E induces

antiinflammatory phenotype in macrophages. Arterioscler Thromb Vasc Biol 31, 1160-1168.

Beauchamp, J.R., Heslop, L., Yu, D.S.W., Tajbakhsh, S., Kelly, R.G., Wernig, A., Buckingham, M.E., Partridge, T.A., and Zammit, P.S. (2000). Expression of Cd34 and Myf5 Defines the Majority of Quiescent Adult Skeletal Muscle Satellite Cells. The Journal of Cell Biology *151*, 1221–1234.

Beaudart, C., Rizzoli, R., Bruyère, O., Reginster, J.-Y., and Biver, E. (2014). Sarcopenia: burden and challenges for public health. Arch Public Health 72, 45.

Becht, E., McInnes, L., Healy, J., Dutertre, C.A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol.

Bentzinger, C.F., Wang, Y.X., Dumont, N.A., and Rudnicki, M.A. (2013). Cellular dynamics in the muscle satellite cell niche. EMBO Reports 14, 1062–1072.

Bentzinger, C.F., Wang, Y.X., von Maltzahn, J., Soleimani, V.D., Yin, H., and Rudnicki, M.A. (2013a). Fibronectin Regulates Wnt7a Signaling and Satellite Cell Expansion. Cell Stem Cell *12*, 75–87.

Bentzinger, C.F., Wang, Y.X., Dumont, N.A., and Rudnicki, M.A. (2013b). Cellular dynamics in the muscle satellite cell niche. EMBO Reports *14*, 1062–1072.

Biressi, S., and Rando, T.A. (2010). Heterogeneity in the muscle satellite cell population. Semin Cell Dev Biol 21, 845-854.

Bjornson, C.R.R., Cheung, T.H., Liu, L., Tripathi, P.V., Steeper, K.M., and Rando, T.A. (2012). Notch Signaling Is Necessary to Maintain Quiescence in Adult Muscle Stem Cells. STEM CELLS *30*, 232–242.

Blackburn, D.M., Lazure, F., Corchado, A.H., Perkins, T.J., Najafabadi, H.S., and Soleimani, V.D. (2019). High-resolution genome-wide expression analysis of single myofibers using SMART-Seq. J. Biol. Chem. *294*, 20097–20108.

Blanco-Bose, W.E., Yao, C.-C., Kramer, R.H., and Blau, H.M. (2001). Purification of Mouse Primary Myoblasts Based on α7 Integrin Expression. Experimental Cell Research *265*, 212–220.

Blau, H.M., Cosgrove, B.D., and Ho, A.T. (2015). The central role of muscle stem cells in regenerative failure with aging. Nat Med 21, 854-862.

Brack, A.S., Conboy, M.J., Roy, S., Lee, M., Kuo, C.J., Keller, C., and Rando, T.A. (2007). Increased Wnt signaling during aging alters muscle stem cell fate and increases fibrosis. Science *317*, 807–810.

van den Brink, S.C., Sage, F., Vértesy, Á., Spanjaard, B., Peterson-Maduro, J., Baron, C.S., Robin, C., and van Oudenaarden, A. (2017). Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. Nature Methods *14*, 935–936.

Buas, M.F., and Kadesch, T. (2010). Regulation of skeletal myogenesis by Notch. Experimental Cell Research *316*, 3028–3033.

Burzyn, D., Kuswanto, W., Kolodin, D., Shadrach, J.L., Cerletti, M., Jang, Y., Sefik, E., Tan, T.G., Wagers, A.J., Benoist, C., et al. (2013). A Special Population of Regulatory T Cells Potentiates Muscle Repair. Cell *155*, 1282–1295.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nature Biotechnology *36*, 411–420.

Capkovic, K., Stevenson, S., Johnson, M., Thelen, J., and Cornelison, D. (2008). Neural cell adhesion molecule (NCAM) marks adult myogenic cells committed to differentiation. Experimental Cell Research *314*, 1553–1565.

Catalán, V., Gómez-Ambrosi, J., Rodríguez, A., Ramírez, B., Ortega, V.A., Hernández-Lizoain, J.L., Baixauli, J., Becerril, S., Rotellar, F., Valentí, V., et al. (2017). IL-32α-induced inflammation constitutes a link between obesity and colon cancer. Oncoimmunology 6, e1328338–e1328338.

Cereijo, R., Gavaldà-Navarro, A., Cairó, M., Quesada-López, T., Villarroya, J., Morón-Ros, S., Sánchez-Infantes, D., Peyrou, M., Iglesias, R., Mampel, T., et al. (2018). CXCL14, a Brown Adipokine that Mediates Brown-Fat-to-Macrophage Communication in Thermogenic Adaptation. Cell Metabolism 28, 750-763.e6.

Cerletti, M., Jurga, S., Witczak, C.A., Hirshman, M.F., Shadrach, J.L., Goodyear, L.J., and Wagers, A.J. (2008). Highly Efficient, Functional Engraftment of Skeletal Muscle Stem Cells in Dystrophic Muscles. Cell *134*, 37–47.

Chakkalakal, J.V., Christensen, J., Xiang, W., Tierney, M.T., Boscolo, F.S., Sacco, A., and Brack, A.S. (2014). Early forming label-retaining muscle stem cells require p27kip1 for maintenance of the primitive state. Development 141, 1649-1659.

Charrin, S., Latil, M., Soave, S., Polesskaya, A., Chrétien, F., Boucheix, C., and Rubinstein, E. (2013). Normal muscle regeneration requires tight control of muscle cell fusion by tetraspanins CD9 and CD81. Nature Communications 4, 1674.

Charville, G.W., Cheung, T.H., Yoo, B., Santos, P.J., Lee, G.K., Shrager, J.B., and Rando, T.A. (2015). Ex Vivo Expansion and In Vivo Self-Renewal of Human Muscle Stem Cells. Stem Cell Reports *5*, 621–632.

Chazaud, B., Sonnet, C., Lafuste, P., Bassez, G., Rimaniol, A.-C., Poron, F., Authier, F.-J., Dreyfus, P.A., and Gherardi, R.K. (2003). Satellite cells attract monocytes and use macrophages as a support to escape apoptosis and enhance muscle growth. The Journal of Cell Biology *163*, 1133–1143.

Chazaud, B. (2016). Inflammation during skeletal muscle regeneration and tissue remodeling: application to exercise-induced muscle damage management. Immunol Cell Biol 94, 140-145.

Chen, H., Lau, M.C., Wong, M.T., Newell, E.W., Poidinger, M., and Chen, J. (2016). Cytofkit: A Bioconductor Package for an Integrated Mass Cytometry Data Analysis Pipeline. PLoS Comput Biol 12, e1005112.

Cho, D.S., and Doles, J.D. (2017). Single cell transcriptome analysis of muscle satellite cells reveals widespread transcriptional heterogeneity. Gene *636*, 54–63.

Choo, H.-J., Canner, J.P., Vest, K.E., Thompson, Z., and Pavlath, G.K. (2017). A tale of two niches: differential functions for VCAM-1 in satellite cells under basal and injured conditions. American Journal of Physiology-Cell Physiology *313*, C392–C404.

Christov, C., Chrétien, F., Abou-Khalil, R., Bassez, G., Vallet, G., Authier, F.-J., Bassaglia, Y., Shinin, V., Tajbakhsh, S., Chazaud, B., et al. (2007). Muscle Satellite Cells and Endothelial Cells: Close Neighbors and Privileged Partners. MBoC *18*, 1397–1409.

Cohen, M., Giladi, A., Gorki, A.-D., Solodkin, D.G., Zada, M., Hladik, A., Miklosi, A., Salame, T.-M., Halpern, K.B., David, E., et al. (2018). Lung Single-Cell Signaling Interaction Map Reveals Basophil Role in Macrophage Imprinting. Cell *175*, 1031-1044.e18.

Collins, C.A., Olsen, I., Zammit, P.S., Heslop, L., Petrie, A., Partridge, T.A., and Morgan, J.E. (2005). Stem Cell Function, Self-Renewal, and Behavioral Heterogeneity of Cells from the Adult Muscle Satellite Cell Niche. Cell *122*, 289–301.

Conboy, I.M., and Rando, T.A. (2002). The Regulation of Notch Signaling Controls Satellite Cell Activation and Cell Fate Determination in Postnatal Myogenesis. Developmental Cell 13.

Conboy, I.M., Conboy, M.J., Smythe, G.M., and Rando, T.A. (2003). Notch-Mediated Restoration of Regenerative Potential to Aged Muscle. Science *302*, 1575.

Cornelison, D.D.W., and Wold, B.J. (1997). Single-Cell Analysis of Regulatory Gene Expression in Quiescent and Activated Mouse Skeletal Muscle Satellite Cells. Developmental Biology *191*, 270–283.

Cornelison, D.D.W., Olwin, B.B., Rudnicki, M.A., and Wold, B.J. (2000). MyoD–/– Satellite Cells in Single-Fiber Culture Are Differentiation Defective and MRF4 Deficient. Developmental Biology *224*, 122–137.

Cornelison, D.D.W., Filla, M.S., Stanley, H.M., Rapraeger, A.C., and Olwin, B.B. (2001). Syndecan-3 and Syndecan-4 Specifically Mark Skeletal Muscle Satellite Cells and Are Implicated in Satellite Cell Maintenance and Muscle Regeneration. Developmental Biology *239*, 79–94.

Cornelison, D.D., Wilcox-Adelman, S.A., Goetinck, P.F., Rauvala, H., Rapraeger, A.C., and Olwin, B.B. (2004). Essential and separable roles for Syndecan-3 and Syndecan-4 in skeletal muscle development and regeneration. Genes Dev 18, 2231-2236.

Cosgrove, B.D., Sacco, A., Gilbert, P.M., and Blau, H.M. (2009). A home away from home: Challenges and opportunities in engineering in vitro muscle satellite cell niches. Differentiation *78*, 185–194.

Cosgrove, B.D., Gilbert, P.M., Porpiglia, E., Mourkioti, F., Lee, S.P., Corbel, S.Y., Llewellyn, M.E., Delp, S.L., and Blau, H.M. (2014). Rejuvenation of the muscle stem cell population restores strength to injured aged muscles. Nat Med 20, 255-264.

Davegårdh, C., Broholm, C., Perfilyev, A., Henriksen, T., García-Calzón, S., Peijs, L., Hansen, N.S., Volkov, P., Kjøbsted, R., Wojtaszewski, J.F.P., et al. (2017). Abnormal epigenetic changes during differentiation of human skeletal muscle stem cells from obese subjects. BMC Medicine 15, 39.

De Micheli, A.J., Laurilliard, E.J., Heinke, C.L., Ravichandran, H., Fraczek, P., Soueid-Baumgarten, S., De Vlaminck, I., Elemento, O., and Cosgrove, B.D. (2020). Single-Cell Analysis of the Muscle Stem Cell Hierarchy Identifies Heterotypic Communication Signals Involved in Skeletal Muscle Regeneration. Cell Reports 30, 3583-3595.e5.

De Micheli, A.J., Spector, J.A., Elemento, O., and Cosgrove, B.D. (2020b). A reference single-cell transcriptomic atlas of human skeletal muscle tissue reveals bifurcated muscle stem cell populations. BioRxiv 2020.01.21.914713.

De Micheli, A.J., Swanson, J.B., Disser, N.P., Martinez, L.M., Walker, N.R., Oliver, D.J., Cosgrove, B.D., and Mendias, C.L. (2020c). Single-cell Transcriptomics Identify Extensive Heterogeneity in the Cellular Composition of Mouse Achilles Tendons. BioRxiv 801266.

Dell'Orso, S., Juan, A.H., Ko, K.-D., Naz, F., Perovanovic, J., Gutierrez-Cruz, G., Feng, X., and Sartorelli, V. (2019). Single cell analysis of adult mouse skeletal muscle stem cells in homeostatic and regenerative conditions. Development *146*, dev174177.

Docheva, D., Hunziker, E.B., Fassler, R., and Brandau, O. (2005). Tenomodulin is necessary for tenocyte proliferation and tendon maturation. Mol Cell Biol 25, 699-705.

Duan, C., Ren, H., and Gao, S. (2010). Insulin-like growth factors (IGFs), IGF receptors, and IGF-binding proteins: Roles in skeletal muscle growth and differentiation. General and Comparative Endocrinology *167*, 344–351.

Dumont, N.A., Wang, Y.X., and Rudnicki, M.A. (2015). Intrinsic and extrinsic mechanisms regulating satellite cell function. Development *142*, 1572–1581.

Durinck, S., Spellman, P., Birney, E., Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nature Protocols, 4, 1184–1191.
Efremova, M., Vento-Tormo, M., Teichmann, S.A., and Vento-Tormo, R. (2019). CellPhoneDB v2.0: Inferring cell-cell communication from combined expression of multi-subunit receptor-ligand complexes. BioRxiv 680926.

Eng, C.-H.L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y., Yun, J., Cronin, C., Karp, C., Yuan, G.-C., et al. (2019). Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. Nature *568*, 235–239.

Enwere, E.K., Holbrook, J., Lejmi-Mrad, R., Vineham, J., Timusk, K., Sivaraj, B., Isaac, M., Uehling, D., Al-awar, R., LaCasse, E., et al. (2012). TWEAK and cIAP1 Regulate Myoblast Fusion Through the Noncanonical NF-κB Signaling Pathway. Sci. Signal. *5*, ra75.

Enwere, E.K., LaCasse, E.C., Adam, N.J., and Korneluk, R.G. (2014). Role of the TWEAK-Fn14-cIAP1-NF-κB Signaling Axis in the Regulation of Myogenesis and Muscle Homeostasis. Front. Immunol. *5*.

Evano, B., and Tajbakhsh, S. (2018). Skeletal muscle stem cells in comfort and stress. Npj Regenerative Medicine *3*, 24.

Farioli-Vecchioli, S., Saraulli, D., Costanzi, M., Leonardi, L., Cina, I., Micheli, L., Nutini, M., Longone, P., Oh, S.P., Cestari, V., et al. (2009). Impaired terminal differentiation of hippocampal granule neurons and defective contextual memory in PC3/Tis21 knockout mice. PloS one 4, e8339.

Farrell, J.A., Wang, Y., Riesenfeld, S.J., Shekhar, K., Regev, A., and Schier, A.F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. Science *360*, eaar3131.

Fernández-Hernando, C., Yu, J., Dávalos, A., Prendergast, J., and Sessa, W.C. (2010). Endothelial-Specific Overexpression of Caveolin-1 Accelerates Atherosclerosis in Apolipoprotein E-Deficient Mice. The American Journal of Pathology 177, 998–1003.

Fielding, R.A., Manfredi, T.J., Ding, W., Fiatarone, M.A., Evans, W.J., and Cannon, J.G. (1993). Acute phase response in exercise. III. Neutrophil and IL-1 beta accumulation in skeletal muscle. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology *265*, R166–R172.

Fletcher, R.B., Das, D., Gadye, L., Street, K.N., Baudhuin, A., Wagner, A., Cole, M.B., Flores, Q., Choi, Y.G., Yosef, N., et al. (2017). Deconstructing Olfactory Stem Cell Trajectories at Single-Cell Resolution. Cell Stem Cell 20, 817-830.e8.

Fu, D.-J., De Micheli, A.J., Bidarimath, M., Ellenson, L.H., Cosgrove, B.D., Flesken-Nikitin, A., and Nikitin, A.Yu. (2019). PAX8 expressing epithelial cells are a cancer-prone source of clonal cyclical regeneration of endometrial epithelium. BioRxiv 853994.

Fu, X., Xiao, J., Wei, Y., Li, S., Liu, Y., Yin, J., Sun, K., Sun, H., Wang, H., Zhang, Z., et al. (2015). Combination of inflammation-related cytokines promotes long-term muscle stem cell expansion. Cell Res *25*, 655–673.

Garcia, S.M., Tamaki, S., Lee, S., Wong, A., Jose, A., Dreux, J., Kouklis, G., Sbitany, H., Seth, R., Knott, P.D., et al. (2018). High-Yield Purification, Preservation, and Serial Transplantation of Human Satellite Cells. Stem Cell Reports 10, 1160–1174.

Gaylinn, B.D., Eddinger, T.J., Martino, P.A., Monical, P.L., Hunt, D.F., and Murphy, R.A. (1989). Expression of nonmuscle myosin heavy and light chains in smooth muscle. Am J Physiol 257, C997-1004.

Giordani, L., He, G.J., Negroni, E., Sakai, H., Law, J.Y.C., Siu, M.M., Wan, R., Corneau, A., Tajbakhsh, S., Cheung, T.H., et al. (2019). High-Dimensional Single-Cell Cartography Reveals Novel Skeletal Muscle-Resident Cell Populations. Molecular Cell *74*, 609-621.e6.

Girardi, F., Taleb, A., Giordani, L., Cadot, B., Datye, A., Ebrahimi, M., Gamage, D.G., Millay, D.P., Gilbert, P.M., and Grand, F.L. (2019). TGFβ signaling curbs cell fusion and muscle regeneration. BioRxiv 557009.

Goel, A.J., Rieder, M.-K., Arnold, H.-H., Radice, G.L., and Krauss, R.S. (2017). Niche Cadherins Control the Quiescence-to-Activation Transition in Muscle Stem Cells. Cell Rep *21*, 2236–2250.

Goncharov, N.V., Nadeev, A.D., Jenkins, R.O., and Avdonin, P.V. (2017). Markers and Biomarkers of Endothelium: When Something Is Rotten in the State. Oxidative Medicine and Cellular Longevity 2017, 9759735.

Gros, J., Serralbo, O., and Marcelle, C. (2009). WNT11 acts as a directional cue to organize the elongation of early muscle fibres. Nature *457*, 589–593.

Haber, A.L., Biton, M., Rogel, N., Herbst, R.H., Shekhar, K., Smillie, C., Burgin, G., Delorey, T.M., Howitt, M.R., Katz, Y., et al. (2017). A single-cell survey of the small intestinal epithelium. Nature *551*, 333–339.

Habib, N., Avraham-Davidi, I., Basu, A., Burks, T., Shekhar, K., Hofree, M., Choudhury, S.R., Aguet, F., Gelfand, E., Ardlie, K., et al. (2017). Massively parallel single-nucleus RNA-seq with DroNc-seq. Nature Methods *14*, 955–958.

Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol *20*, 296.

Hagan, M., Zhou, M., Ashraf, M., Kim, I.-M., Su, H., Weintraub, N.L., and Tang, Y. (2017). Long noncoding RNAs and their roles in skeletal muscle fate determination. Noncoding RNA Investig 1, 24.

Harmon, B.T., Orkunoglu-Suer, E.F., Adham, K., Larkin, J.S., Gordish-Dressman, H., Clarkson, P.M., Thompson, P.D., Angelopoulos, T.J., Gordon, P.M., Moyna, N.M., et al. (2010). CCL2 and CCR2 variants are associated with skeletal muscle strength and change in strength with resistance training. J Appl Physiol (1985) 109, 1779–1785.

Heredia, J.E., Mukundan, L., Chen, F.M., Mueller, A.A., Deo, R.C., Locksley, R.M., Rando, T.A., and Chawla, A. (2013). Type 2 Innate Signals Stimulate Fibro/Adipogenic Progenitors to Facilitate Muscle Regeneration. Cell *153*, 376–388.

Hie, B., Bryson, B., and Berger, B. (2019). Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. Nature Biotechnology *37*, 685–691.

Ho, M.M., Manughian-Peter, A., Spivia, W.R., Taylor, A., and Fraser, D.A. (2016). Macrophage molecular signaling and inflammatory responses during ingestion of atherogenic lipoproteins are modulated by complement protein C1q. Atherosclerosis 253, 38-46.

Hwang, B., Lee, J.H., and Bang, D. (2018). Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp Mol Med 50, 96.

Ilicic, T., Kim, J.K., Kolodziejczyk, A.A., Bagger, F.O., McCarthy, D.J., Marioni, J.C., and Teichmann, S.A. (2016). Classification of low quality cells from single-cell RNA-seq data. Genome Biology *17*, 29.

Irintchev, A., Zeschnigk, M., Starzinski-Powitz, A., and Wernig, A. (1994). Expression pattern of M-cadherin in normal, denervated, and regenerating mouse muscles. Dev. Dyn. *199*, 326–337.

Janssen, E., Dzeja, P.P., Oerlemans, F., Simonetti, A.W., Heerschap, A., de Haan, A., Rush, P.S., Terjung, R.R., Wieringa, B., and Terzic, A. (2000). Adenylate kinase 1 gene

deletion disrupts muscle energetic economy despite metabolic rearrangement. EMBO J 19, 6371-6381.

Järvinen, T.A., Järvinen, M., and Kalimo, H. (2014). Regeneration of injured skeletal muscle after the injury. Muscles Ligaments Tendons J 3, 337–345.

Joe, A.W.B., Yi, L., Natarajan, A., Le Grand, F., So, L., Wang, J., Rudnicki, M.A., and Rossi, F.M.V. (2010). Muscle injury activates resident fibro/adipogenic progenitors that facilitate myogenesis. Nature Cell Biology *12*, 153–163.

Karpus, O.N., Kiener, H.P., Niederreiter, B., Yilmaz-Elis, A.S., van der Kaa, J., Ramaglia, V., Arens, R., Smolen, J.S., Botto, M., Tak, P.P., et al. (2015). CD55 deposited on synovial collagen fibers protects from immune complex-mediated arthritis. Arthritis Research & Therapy 17, 6.

Kester, L., and van Oudenaarden, A. (2018). Single-Cell Transcriptomics Meets Lineage Tracing. Cell Stem Cell *23*, 166–179.

Kim, J.-H., Torgerud, W.S., Mosser, K.H.H., Hirai, H., Watanabe, S., Asakura, A., and Thompson, L.V. (2012). Myosin light chain 3f attenuates age-induced decline in contractile velocity in MHC type II single muscle fibers: Age, velocity, and myosin light chains. Aging Cell *11*, 203–212.

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nature Methods *16*, 1289–1296.

Kotecha, N., Krutzik, P.O., and Irish, J.M. (2010). Web-based analysis and publication of flow cytometry experiments. Curr Protoc Cytom Chapter 10, Unit10 17.

Kuang, S., Chargé, S.B., Seale, P., Huh, M., and Rudnicki, M.A. (2006). Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. The Journal of Cell Biology *172*, 103–113.

Kuang, S., Kuroda, K., Le Grand, F., and Rudnicki, M.A. (2007). Asymmetric Self-Renewal and Commitment of Satellite Stem Cells in Muscle. Cell *129*, 999–1010.

Kuang, S., Gillespie, M.A., and Rudnicki, M.A. (2008). Niche Regulation of Muscle Satellite Cell Self-Renewal and Differentiation. Cell Stem Cell *2*, 22–31.

Kumar, D., Shadrach, J.L., Wagers, A.J., and Lassar, A.B. (2009). Id3 is a direct transcriptional target of Pax7 in quiescent satellite cells. Mol Biol Cell 20, 3170-3177.

Kumar, M.P., Du, J., Lagoudas, G., Jiao, Y., Sawyer, A., Drummond, D.C., Lauffenburger, D.A., and Raue, A. (2018). Analysis of Single-Cell RNA-Seq Identifies

Cell-Cell Communication Associated with Tumor Characteristics. Cell Reports *25*, 1458-1468.e4.

Kuschel, R., Deininger, M.H., Meyermann, R., Bornemann, A., Yablonka-Reuveni, Z., and Schluesener, H.J. (2000). Allograft inflammatory factor-1 is expressed by macrophages in injured skeletal muscle and abrogates proliferation and differentiation of satellite cells. J Neuropathol Exp Neurol 59, 323-332.

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494–498.

Larsson, L., Degens, H., Li, M., Salviati, L., Lee, Y. il, Thompson, W., Kirkland, J.L., and Sandri, M. (2018). Sarcopenia: Aging-Related Loss of Muscle Mass and Function. Physiological Reviews 99, 427–511.

Le Grand, F., Jones, A.E., Seale, V., Scimè, A., and Rudnicki, M.A. (2009). Wnt7a Activates the Planar Cell Polarity Pathway to Drive the Symmetric Expansion of Satellite Stem Cells. Cell Stem Cell *4*, 535–547.

Liu, L., Cheung, T.H., Charville, G.W., and Rando, T.A. (2015). Isolation of skeletal muscle stem cells by fluorescence-activated cell sorting. Nat Protoc *10*, 1612–1624.

Liu, N., Garry, G.A., Li, S., Bezprozvannaya, S., Sanchez-Ortiz, E., Chen, B., Shelton, J.M., Jaichander, P., Bassel-Duby, R., and Olson, E.N. (2017). A Twist2-dependent progenitor cell contributes to adult skeletal muscle. Nature Cell Biology *19*, 202–213.

Liu, Z., Jin, Y.Q., Chen, L., Wang, Y., Yang, X., Cheng, J., Wu, W., Qi, Z., and Shen, Z. (2015). Specific marker expression and cell state of Schwann cells during culture in vitro. PloS one 10, e0123278.

Londhe, P., and Davie, J.K. (2013). Interferon-γ Resets Muscle Cell Fate by Stimulating the Sequential Recruitment of JARID2 and PRC2 to Promoters to Repress Myogenesis. Sci. Signal. *6*, ra107.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

Low, S., Barnes, J.L., Zammit, P.S., and Beauchamp, J.R. (2018). Delta-Like 4 Activates Notch 3 to Regulate Self-Renewal in Skeletal Muscle Stem Cells: Regulation of Self-Renewal in Skeletal Muscle Stem Cells. Stem Cells *36*, 458–466.

Lukjanenko, L., Jung, M.J., Hegde, N., Perruisseau-Carrier, C., Migliavacca, E., Rozo, M., Karaz, S., Jacot, G., Schmidt, M., Li, L., et al. (2016). Loss of fibronectin from the aged stem cell niche affects the regenerative capacity of skeletal muscle in mice. Nature Medicine *22*, 897–905.

Lutolf, M.P., Gilbert, P.M., and Blau, H.M. (2009). Designing materials to direct stem-cell fate. Nature 462, 433–441.
Lyons, G.E., Ontell, M., Cox, R., Sassoon, D., and Buckingham, M. (1990). The expression of myosin genes in developing skeletal muscle in the mouse embryo. J Cell Biol 111, 1465-1476.

Machado, L., Esteves de Lima, J., Fabre, O., Proux, C., Legendre, R., Szegedi, A., Varet, H., Ingerslev, L.R., Barrès, R., Relaix, F., et al. (2017). In Situ Fixation Redefines Quiescence and Early Activation of Skeletal Muscle Stem Cells. Cell Reports *21*, 1982–1993.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202–1214.

Maesner, C.C., Almada, A.E., and Wagers, A.J. (2016). Established cell surface markers efficiently isolate highly overlapping populations of skeletal muscle satellite cells by fluorescence-activated cell sorting. Skeletal Muscle *6*, 35.

Mashinchian, O., Pisconti, A., Le Moal, E., and Bentzinger, C.F. (2018). The Muscle Stem Cell Niche in Health and Disease. In Current Topics in Developmental Biology, (Elsevier), pp. 23–65.

Mauro, A. (1961). SATELLITE CELL OF SKELETAL MUSCLE FIBERS. The Journal of Biophysical and Biochemical Cytology *9*, 493–495.

Mittal, A., Kumar, A., Lach-Trifilieff, E., Wauters, S., Li, H., Makonchuk, D., Glass, D., and Kumar, A. (2010). The TWEAK-Fn14 system is a critical regulator of denervation-induced skeletal muscle atrophy in mice. The Journal of Cell Biology 188, 833–849.

Montarras, D., Morgan, J., Collins, C., Relaix, F., Zaffran, S., Cumano, A., Partridge, T., and Buckingham, M. (2005). Direct Isolation of Satellite Cells for Skeletal Muscle Regeneration. Science *309*, 2064.

Motohashi, N., and Asakura, A. (2014). Muscle satellite cell heterogeneity and self-renewal. Frontiers in cell and developmental biology 2.

Mourikis, P., and Tajbakhsh, S. (2014). Distinct contextual roles for Notch signalling in skeletal muscle stem cells. BMC Dev Biol *14*, 2.

Muffat, J., and Walker, D.W. (2010). Apolipoprotein D: An overview of its role in aging and age-related diseases. Cell Cycle 9, 269–273.

Mukund, K., and Subramaniam, S. (2019). Skeletal muscle: A review of molecular structure and function, in health and disease. WIREs Syst Biol Med.

Mylona, E., Jones, K.A., Mills, S.T., and Pavlath, G.K. (2006). CD44 regulates myoblast migration and differentiation. Journal of Cellular Physiology 209, 314–321.

Nabeshima, Y., Hanaoka, K., Hayasaka, M., Esuml, E., Li, S., Nonaka, I., and Nabeshima, Y. (1993). Myogenin gene disruption results in perinatal lethality because of severe muscle defect. Nature *364*, 532–535.

Novak, M.L., and Koh, T.J. (2013). Macrophage phenotypes during tissue repair. J Leukoc Biol 93, 875-881.

O'Flanagan, C.H., Campbell, K.R., Zhang, A.W., Kabeer, F., Lim, J.L.P., Biele, J., Eirew, P., Lai, D., McPherson, A., Kong, E., et al. (2019). Dissociation of solid tumor tissues with cold active protease for single-cell RNA-seq minimizes conserved collagenase-associated stress responses. Genome Biology *20*, 210.

Palacios, D., Mozzetta, C., Consalvi, S., Caretti, G., Saccone, V., Proserpio, V., Marquez, V.E., Valente, S., Mai, A., Forcales, S.V., et al. (2010). TNF/p38α/Polycomb Signaling to Pax7 Locus in Satellite Cells Links Inflammation to the Epigenetic Control of Muscle Regeneration. Cell Stem Cell *7*, 455–469.

Pampeno, C., Derkatch, I.L., and Meruelo, D. (2014). Interaction of Human Laminin Receptor with Sup35, the [PSI+] Prion-Forming Protein from S. cerevisiae: A Yeast Model for Studies of LamR Interactions with Amyloidogenic Proteins. PLoS ONE 9, e86013.

Pawlikowski, B., Vogler, T.O., Gadek, K., and Olwin, B.B. (2017). Regulation of skeletal muscle stem cells by fibroblast growth factors: Regulation of Skeletal Muscle SCs by FGFs. Dev. Dyn. *246*, 359–367.

Péault, B., Rudnicki, M., Torrente, Y., Cossu, G., Tremblay, J.P., Partridge, T., Gussoni, E., Kunkel, L.M., and Huard, J. (2007). Stem and Progenitor Cells in Skeletal Muscle Development, Maintenance, and Therapy. Molecular Therapy *15*, 867–877.

Pedersen, L., Olsen, C.H., Pedersen, B.K., and Hojman, P. (2012). Muscle-derived expression of the chemokine CXCL1 attenuates diet-induced obesity and improves fatty acid oxidation in the muscle. American Journal of Physiology-Endocrinology and Metabolism 302, E831–E840.

Perdas, E., Stawski, R., Kaczka, K., Nowak, D., and Zubrzycka, M. (2019). Altered levels of circulating nuclear and mitochondrial DNA in patients with Papillary Thyroid Cancer. Sci Rep *9*, 14438.

Perillo, M., and Folker, E.S. (2018). Specialized Positioning of Myonuclei Near Cell-Cell Junctions. Front. Physiol. *9*, 1531.

Picelli, S., Bjorklund, A.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Meth *10*, 1096–1098.

Pinho, S., and Frenette, P.S. (2019). Haematopoietic stem cell activity and interactions with the niche. Nat Rev Mol Cell Biol 20, 303–320.

Pisani, D.F., Clement, N., Loubat, A., Plaisant, M., Sacconi, S., Kurzenne, J.-Y., Desnuelle, C., Dani, C., and Dechesne, C.A. (2010). Hierarchization of Myogenic and Adipogenic Progenitors Within Human Skeletal Muscle. STEM CELLS *28*, 2182–2194.

Pisconti, A., Cornelison, D.D.W., Olguín, H.C., Antwine, T.L., and Olwin, B.B. (2010). Syndecan-3 and Notch cooperate in regulating adult myogenesis. The Journal of Cell Biology *190*, 427–441.

Pisconti, A., Bernet, J.D., and Olwin, B.B. (2012). Syndecans in skeletal muscle development, regeneration and homeostasis. Muscles Ligaments Tendons J *2*, 1–9.

Pisconti, A., Banks, G.B., Babaeijandaghi, F., Betta, N.D., Rossi, F.M.V., Chamberlain, J.S., and Olwin, B.B. (2016). Loss of niche-satellite cell interactions in syndecan-3 null mice alters muscle progenitor cell homeostasis improving muscle regeneration. Skeletal Muscle *6*, 34.

Porpiglia, E., Samusik, N., Ho, A.T., Cosgrove, B.D., Mai, T., Davis, K.L., Jager, A., Nolan, G.P., Bendall, S.C., Fantl, W.J., et al. (2017). High-resolution myogenic lineage mapping by single-cell mass cytometry. Nat Cell Biol 19, 558-567.

Pretheeban, T., Lemos, D.R., Paylor, B., Zhang, R.-H., and Rossi, F.M. (2012). Role of stem/progenitor cells in reparative disorders. Fibrogenesis Tissue Repair *5*, 20.

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nature Methods *14*, 979–982.

Ramilowski, J.A., Goldberg, T., Harshbarger, J., Kloppmann, E., Lizio, M., Satagopam, V.P., Itoh, M., Kawaji, H., Carninci, P., Rost, B., et al. (2015). A draft network of ligand–receptor-mediated multicellular signalling in human. Nature Communications *6*, 7866.

Ratajczak, M.Z. (2003). 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.

Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., et al. (2017). The Human Cell Atlas. ELife 6, e27041.

Reilly, G.C., and Engler, A.J. (2010). Intrinsic extracellular matrix properties regulate stem cell differentiation. Journal of Biomechanics *43*, 55–62.

Ren, H., Yin, P., and Duan, C. (2008). IGFBP-5 regulates muscle cell differentiation by binding to IGF-II and switching on the IGF-II auto-regulation loop. J Cell Biol 182, 979-991.

Rhoads, R.P., Johnson, R.M., Rathbone, C.R., Liu, X., Temm-Grove, C., Sheehan, S.M., Hoying, J.B., and Allen, R.E. (2009). Satellite cell-mediated angiogenesis in vitro coincides with a functional hypoxia-inducible factor pathway. American Journal of Physiology-Cell Physiology *296*, C1321–C1328.

Riddle, E.S., Bender, E.L., Thalacker-Mercer, A.E., (2018). Transcript profile distinguishes variability in human myogenic progenitor cell expansion capacity. Physiol Genomics 50, 817–827.

Rocheteau, P., Gayraud-Morel, B., Siegl-Cachedenier, I., Blasco, M.A., and Tajbakhsh, S. (2012). A subpopulation of adult skeletal muscle stem cells retains all template DNA strands after cell division. Cell 148, 112-125.

Rodgers, J.T., King, K.Y., Brett, J.O., Cromie, M.J., Charville, G.W., Maguire, K.K., Brunson, C., Mastey, N., Liu, L., Tsai, C.-R., et al. (2014). mTORC1 controls the adaptive transition of quiescent stem cells from G0 to GAlert. Nature *510*, 393–396.

Rodriques, S.G., Stickels, R.R., Goeva, A., Martin, C.A., Murray, E., Vanderburg, C.R., Welch, J., Chen, L.M., Chen, F., and Macosko, E.Z. (2019). Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. 6.

Roman, W., and Gomes, E.R. (2018). Nuclear positioning in skeletal muscle. Seminars in Cell & Developmental Biology *82*, 51–56.

Rood, J.E., Stuart, T., Ghazanfar, S., Biancalani, T., Fisher, E., Butler, A., Hupalowska, A., Gaffney, L., Mauck, W., Eraslan, G., et al. (2019). Toward a Common Coordinate Framework for the Human Body. Cell 179, 1455–1467.

Rosenberg, A.B., Roco, C.M., Muscat, R.A., Kuchina, A., Sample, P., Yao, Z., Graybuck, L.T., Peeler, D.J., Mukherjee, S., Chen, W., et al. (2018). Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. Science *360*, 176.

Rubenstein, A.B., Smith, G.R., Raue, U., Begue, G., Minchev, K., Ruf-Zamojski, F., Nair, V.D., Wang, X., Zhou, L., Zaslavsky, E., Trappe, T.A., Sealfon, S.C. (2020.) Single-cell transcriptional profiles of human skeletal muscle. Scientific Reports 10, 229.

Ryall, J.G., Schertzer, J.D., and Lynch, G.S. (2008). Cellular and molecular mechanisms underlying age-related skeletal muscle wasting and weakness. Biogerontology 9, 213–228.

Ryall, J.G., Dell'Orso, S., Derfoul, A., Juan, A., Zare, H., Feng, X., Clermont, D., Koulnis, M., Gutierrez-Cruz, G., Fulco, M., et al. (2015). The NAD+-Dependent SIRT1 Deacetylase Translates a Metabolic Switch into Regulatory Epigenetics in Skeletal Muscle Stem Cells. Cell Stem Cell *16*, 171–183.

Sacco, A., Doyonnas, R., Kraft, P., Vitorovic, S., and Blau, H.M. (2008). Self-renewal and expansion of single transplanted muscle stem cells. Nature *456*, 502–506.

Sanes, J.R. (2003). The Basement Membrane/Basal Lamina of Skeletal Muscle. Journal of Biological Chemistry *278*, 12601–12604.

Santos, A.J.M., Lo, Y.-H., Mah, A.T., and Kuo, C.J. (2018). The Intestinal Stem Cell Niche: Homeostasis and Adaptations. Trends in Cell Biology 28, 1062–1078.

Sarver, D.C., Sugg, K.B., Disser, N.P., Enselman, E.R.S., Awan, T.M., and Mendias, C.L. (2017). Local cryotherapy minimally impacts the metabolome and transcriptome of human skeletal muscle. Scientific Reports 7.

Sato, S., Ogura, Y., and Kumar, A. (2014). TWEAK/Fn14 Signaling Axis Mediates Skeletal Muscle Atrophy and Metabolic Dysfunction. Frontiers in Immunology 5, 18.

Scapini, P., Lapinet-Vera, J.A., Gasperini, S., Calzetti, F., Bazzoni, F., and Cassatella, M.A. (2000). The neutrophil as a cellular source of chemokines. Immunological Reviews *177*, 195–203.

Schaum, N., Karkanias, J., Neff, N.F., May, A.P., Quake, S.R., Wyss-Coray, T., Darmanis, S., Batson, J., Botvinnik, O., Chen, M.B., et al. (2018). Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. Nature *562*, 367–372.

Schiaffino, S. (2018). Muscle fiber type diversity revealed by anti-myosin heavy chain antibodies. FEBS J *285*, 3688–3694.

Schweitzer, R., Chyung, J.H., Murtaugh, L.C., Brent, A.E., Rosen, V., Olson, E.N., Lassar, A., and Tabin, C.J. (2001). Analysis of the tendon cell fate using Scleraxis, a specific marker for tendons and ligaments. Development 128, 3855-3866.

Scimeca, M., Bonanno, E., Piccirilli, E., Baldi, J., Mauriello, A., Orlandi, A., Tancredi, V., Gasbarra, E., and Tarantino, U. (2015). Satellite Cells CD44 Positive Drive Muscle Regeneration in Osteoarthritis Patients. Stem Cells Int 2015, 469459–469459.

Seale, P., Sabourin, L.A., Girgis-Gabardo, A., Mansouri, A., Gruss, P., and Rudnicki, M.A. (2000). Pax7 Is Required for the Specification of Myogenic Satellite Cells. Cell *102*, 777–786.

Shafiee, G., Keshtkar, A., Soltani, A., Ahadi, Z., Larijani, B., and Heshmat, R. (2017). Prevalence of sarcopenia in the world: a systematic review and meta- analysis of general population studies. J Diabetes Metab Disord 16, 21.

Sherwood, R.I., Christensen, J.L., Conboy, I.M., Conboy, M.J., Rando, T.A., Weissman, I.L., and Wagers, A.J. (2014). Isolation of Adult Mouse Myogenic Progenitors: Functional Heterogeneity of Cells within and Engrafting Skeletal Muscle. 12.

Shireman, P.K., Contreras-Shannon, V., Ochoa, O., Karia, B.P., Michalek, J.E., and McManus, L.M. (2007). MCP-1 deficiency causes altered inflammation with impaired skeletal muscle regeneration. Journal of Leukocyte Biology *81*, 775–785.

Skelly, D.A., Squiers, G.T., McLellan, M.A., Bolisetty, M.T., Robson, P., Rosenthal, N.A., and Pinto, A.R. (2018). Single-Cell Transcriptional Profiling Reveals Cellular Diversity and Intercommunication in the Mouse Heart. Cell Reports *22*, 600–610.

Snow, M.H. (1977). Myogenic cell formation in regenerating rat skeletal muscle injured by mincing I. A fine structural study. The Anatomical Record *188*, 181–199.

Sousa-Victor, P., Gutarra, S., Garcia-Prat, L., Rodriguez-Ubreva, J., Ortet, L., Ruiz-Bonilla, V., Jardi, M., Ballestar, E., Gonzalez, S., Serrano, A.L., et al. (2014). Geriatric muscle stem cells switch reversible quiescence into senescence. Nature 506, 316-321.
Stuart, T., and Satija, R. (2019). Integrative single-cell analysis. Nat Rev Genet 20, 257-272.

Spinazzola, J.M., and Gussoni, E. (2017). Isolation of Primary Human Skeletal Muscle Cells. Bio Protoc *7*, e2591.

Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P.K., Swerdlow, H., Satija, R., and Smibert, P. (2017). Simultaneous epitope and transcriptome measurement in single cells. Nature Methods *14*, 865–868.

Stuart, T., and Satija, R. (2019a). Integrative single-cell analysis. Nature Reviews Genetics 20, 257–272.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., III, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019b). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888-1902.e21.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set

enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 102, 15545.

Svensson, V., Vento-Tormo, R., and Teichmann, S.A. (2018). Exponential scaling of single-cell RNA-seq in the past decade. Nature Protocols *13*, 599–604.

Swindell, W.R., Johnston, A., Xing, X., Little, A., Robichaud, P., Voorhees, J.J., Fisher, G., and Gudjonsson, J.E. (2013). Robust shifts in S100a9 expression with aging: A novel mechanism for chronic inflammation. Scientific Reports 3, 1215.

Tajrishi, M.M., Zheng, T.S., Burkly, L.C., and Kumar, A. (2014). The TWEAK-Fn14 pathway: A potent regulator of skeletal muscle biology in health and disease. Cytokine & Growth Factor Reviews 25, 215–225.

Tanaka, K.K., Hall, J.K., Troy, A.A., Cornelison, D.D.W., Majka, S.M., and Olwin, B.B. (2009). Syndecan-4-Expressing Muscle Progenitor Cells in the SP Engraft as Satellite Cells during Muscle Regeneration. Cell Stem Cell *4*, 217–225.

Tarnopolsky, M.A., Pearce, E., Smith, K., and Lach, B. (2011). Suction-modified Bergström muscle biopsy technique: Experience with 13,500 procedures. Muscle & Nerve 43, 716–725.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nature Methods *6*, 377–382.

Tedesco, F.S., and Cossu, G. (2012). Stem cell therapies for muscle disorders: Current Opinion in Neurology *25*, 597–603.

Tidball, J.G. (2017). Regulation of muscle growth and regeneration by the immune system. Nature Reviews Immunology *17*, 165–178.

Tierney, M.T., and Sacco, A. (2016). Satellite Cell Heterogeneity in Skeletal Muscle Homeostasis. Trends Cell Biol 26, 434-444.

Tierney, M.T., Stec, M.J., Rulands, S., Simons, B.D., and Sacco, A. (2018). Muscle Stem Cells Exhibit Distinct Clonal Dynamics in Response to Tissue Repair and Homeostatic Aging. Cell Stem Cell 22, 119-127 e113.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nature Biotechnology *32*, 381–386.

Troy, A., Cadwallader, A.B., Fedorov, Y., Tyner, K., Tanaka, K.K., and Olwin, B.B. (2012). Coordination of Satellite Cell Activation and Self-Renewal by Par-Complex-Dependent Asymmetric Activation of p38α/β MAPK. Cell Stem Cell *11*, 541–553.

Uezumi, A., Fukada, S., Yamamoto, N., Takeda, S., and Tsuchida, K. (2010). Mesenchymal progenitors distinct from satellite cells contribute to ectopic fat cell formation in skeletal muscle. Nature Cell Biology *12*, 143–152.

Uezumi, A., Ito, T., Morikawa, D., Shimizu, N., Yoneda, T., Segawa, M., Yamaguchi, M., Ogawa, R., Matev, M.M., Miyagoe-Suzuki, Y., et al. (2011). Fibrosis and adipogenesis originate from a common mesenchymal progenitor in skeletal muscle. Journal of Cell Science *124*, 3654–3664.

Uezumi, A., Nakatani, M., Ikemoto-Uezumi, M., Yamamoto, N., Morita, M., Yamaguchi, A., Yamada, H., Kasai, T., Masuda, S., Narita, A., et al. (2016). Cell-Surface Protein Profiling Identifies Distinctive Markers of Progenitor Cells in Human Skeletal Muscle. Stem Cell Reports *7*, 263–278.

Urciuolo, A., Quarta, M., Morbidoni, V., Gattazzo, F., Molon, S., Grumati, P., Montemurro, F., Tedesco, F.S., Blaauw, B., Cossu, G., et al. (2013). Collagen VI regulates satellite cell self-renewal and muscle regeneration. Nature Communications *4*, 1964.

Valdés, J.A., Flores, S., Fuentes, E.N., Osorio-Fuentealba, C., Jaimovich, E., and Molina, A. (2013). IGF-1 induces IP3-dependent calcium signal involved in the regulation of myostatin gene expression mediated by NFAT during myoblast differentiation. Journal of Cellular Physiology *228*, 1452–1463.

van Velthoven, C.T.J., de Morree, A., Egner, I.M., Brett, J.O., and Rando, T.A. (2017). Transcriptional Profiling of Quiescent Muscle Stem Cells In Vivo. Cell reports 21, 1994-2004.

Vento-Tormo, R., Efremova, M., Botting, R.A., Turco, M.Y., Vento-Tormo, M., Meyer, K.B., Park, J.-E., Stephenson, E., Polański, K., Goncalves, A., et al. (2018). Single-cell reconstruction of the early maternal–fetal interface in humans. Nature *563*, 347–353.

Verdijk, L.B., Koopman, R., Schaart, G., Meijer, K., Savelberg, H.H.C.M., and van Loon, L.J.C. (2007). Satellite cell content is specifically reduced in type II skeletal muscle fibers in the elderly. American Journal of Physiology-Endocrinology and Metabolism *292*, E151–E157.

Volonte, D., Liu, Y., and Galbiati, F. (2004). The modulation of caveolin-1 expression controls satellite cell activation during muscle repair. The FASEB Journal 19, 237–239.

Waddell, J.N., Zhang, P., Wen, Y., Gupta, S.K., Yevtodiyenko, A., Schmidt, J.V., Bidwell, C.A., Kumar, A., and Kuang, S. (2010). Dlk1 Is Necessary for Proper Skeletal Muscle Development and Regeneration. PLOS ONE 5, e15055.

Wagner, A., Regev, A., and Yosef, N. (2016). Revealing the vectors of cellular identity with single-cell genomics. Nat Biotechnol 34, 1145-1160.

Wang, Y.J., Schug, J., Lin, J., Wang, Z., Kossenkov, A., the HPAP Consortium, and Kaestner, K.H. (2019). Comparative analysis of commercially available single-cell RNA sequencing platforms for their performance in complex human tissues (Genomics).

Wang, Y.X., Feige, P., Brun, C.E., Hekmatnejad, B., Dumont, N.A., Renaud, J.-M., Faulkes, S., Guindon, D.E., and Rudnicki, M.A. (2019). EGFR-Aurka Signaling Rescues Polarity and Regeneration Defects in Dystrophin-Deficient Muscle Stem Cells by Increasing Asymmetric Divisions. Cell Stem Cell *24*, 419-432.e6.

Wang, Y.X., and Rudnicki, M.A. (2011). Satellite cells, the engines of muscle repair. Nat Rev Mol Cell Biol 13, 127-133.

Watson, C., Whittaker, S., Smith, N., Vora, A.J., Dumonde, D.C., & Brown, K.A. (1996). IL-6 acts on endothelial cells to preferentially increase their adherence for lymphocytes. Clinical and experimental immunology, 105 1, 112-9.

Wen, Y., Bi, P., Liu, W., Asakura, A., Keller, C., and Kuang, S. (2012). Constitutive Notch Activation Upregulates Pax7 and Promotes the Self-Renewal of Skeletal Muscle Satellite Cells. Molecular and Cellular Biology *32*, 2300–2311.

Wosczyna, M.N., and Rando, T.A. (2018). A Muscle Stem Cell Support Group: Coordinated Cellular Responses in Muscle Regeneration. Dev Cell 46, 135-143.

Wosczyna, M.N., Konishi, C.T., Perez Carbajal, E.E., Wang, T.T., Walsh, R.A., Gan, Q., Wagner, M.W., and Rando, T.A. (2019). Mesenchymal Stromal Cells Are Required for Regeneration and Homeostatic Maintenance of Skeletal Muscle. Cell Reports *27*, 2029-2035.e5.

Wu, Y., Tan, X., Liu, P., Yang, Y., Huang, Y., Liu, X., Meng, X., Yu, B., Wu, M., and Jin, H. (2019). ITGA6 and RPSA synergistically promote pancreatic cancer invasion and metastasis via PI3K and MAPK signaling pathways. Experimental Cell Research 379, 30–47.

Xin, Y., Kim, J., Ni, M., Wei, Y., Okamoto, H., Lee, J., Adler, C., Cavino, K., Murphy, A.J., Yancopoulos, G.D., et al. (2016). Use of the Fluidigm C1 platform for RNA sequencing of single mouse pancreatic islet cells. Proc Natl Acad Sci USA *113*, 3293.

Xu, X., Wilschut, K.J., Kouklis, G., Tian, H., Hesse, R., Garland, C., Sbitany, H., Hansen, S., Seth, R., Knott, P.D., Hoffman, W.Y., Pomerantz, J.H. (2015). Human Satellite Cell Transplantation and Regeneration from Diverse Skeletal Muscles. Stem Cell Reports 5, 419-434.

Yablonka-Reuveni, Z., Danoviz, M.E., Phelps, M., and Stuelsatz, P. (2015). Myogenic-specific ablation of Fgfr1 impairs FGF2-mediated proliferation of satellite cells at the myofiber niche but does not abolish the capacity for muscle regeneration. Front. Aging Neurosci. *7*.

Yao, C.C., Ziober, B.L., Sutherland, A.E., Mendrick, D.L., and Kramer, R.H. (1996). Laminins promote the locomotion of skeletal myoblasts via the alpha 7 integrin receptor. J. Cell Sci. *109*, 3139.

Yin, H., Price, F., and Rudnicki, M.A. (2013). Satellite Cells and the Muscle Stem Cell Niche. Physiological Reviews *93*, 23–67.

Yuniati, L., Scheijen, B., van der Meer, L.T., and van Leeuwen, F.N. (2019). Tumor suppressors BTG1 and BTG2: Beyond growth control. J Cell Physiol 234, 5379-5389.

Zakrzewski, W., Dobrzyński, M., Szymonowicz, M., and Rybak, Z. (2019). Stem cells: past, present, and future. Stem Cell Res Ther 10, 68.

Zhang, L., Uezumi, A., Kaji, T., Tsujikawa, K., Andersen, D.C., Jensen, C.H., and Fukada, S. (2019). Expression and Functional Analyses of Dlk1 in Muscle Stem Cells and Mesenchymal Progenitors during Muscle Regeneration. International Journal of Molecular Sciences 20, 3269.

Zammit, P.S. (2017). Function of the myogenic regulatory factors Myf5, MyoD, Myogenin and MRF4 in skeletal muscle, satellite cells and regenerative myogenesis. Seminars in Cell & Developmental Biology *72*, 19–32.

Zammit, P.S., Relaix, F., Nagata, Y., Ruiz, A.P., Collins, C.A., Partridge, T.A., and Beauchamp, J.R. (2006). Pax7 and myogenic progression in skeletal muscle satellite cells. J. Cell Sci. *119*, 1824.

Zordan, P., Rigamonti, E., Freudenberg, K., Conti, V., Azzoni, E., Rovere-Querini, P., and Brunelli, S. (2014). Macrophages commit postnatal endothelium-derived progenitors to angiogenesis and restrict endothelial to mesenchymal transition during muscle regeneration. Cell Death Dis 5, e1031.

Zhou, J.X., Taramelli, R., Pedrini, E., Knijnenburg, T., and Huang, S. (2017). Extracting Intercellular Signaling Network of Cancer Tissues using Ligand-Receptor Expression Patterns from Whole-tumor and Single-cell Transcriptomes. Scientific Reports *7*, 8815.

Ziegenhain, C., Vieth, B., Parekh, S., Reinius, B., Guillaumet-Adkins, A., Smets, M., Leonhardt, H., Heyn, H., Hellmann, I., and Enard, W. (2017). Comparative Analysis of Single-Cell RNA Sequencing Methods. Molecular Cell *65*, 631-643.e4.