ELUCIDATION OF THE FATE AND FUNCTION OF MESENCHYMAL PROGENITORS IN TISSUE RENEWAL AND REGENERATION

by

Wilder Scott

BSc (Hons), The University of Saskatchewan, 2006

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Cellular and Physiological Sciences)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

March 2019

© Wilder Scott, 2019
The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

**Elucidation of the fate and function of mesenchymal progenitors in tissue renewal and regeneration.**

__________________________
Wilder Scott

__________________________
in partial fulfillment of the requirements for

Doctor of Philosophy

__________________________
the degree of

__________________________
in

Cell and developmental biology

---

**Examining Committee:**

T. Michael Underhill

Supervisor

Tim O’Connor

Supervisory Committee Member

Kelly McNagny

Supervisory Committee Member

Francis Lynn

University Examiner

Edward Conway

University Examiner

---

**Additional Supervisory Committee Members:**

Fabio Rossi

Supervisory Committee Member

Supervisory Committee Member
Abstract

Adult tissues contain multiple stem and progenitor cells that are critical for their renewal and regeneration. Tissue resident stem/progenitor populations include mesenchymal progenitors (MPs) whose function and fate are unclear. These studies have been hampered by the lack of suitable solitary markers of these cells which would enable both lineage tracing and functional analyses of MPs. Using a discovery-based approach, the gene Hypermethylated in cancer 1 (Hic1) has been identified as such a marker. Subsequently, two novel knock-in alleles of Hic1 were generated and characterized as part of this project and these analyses showed that Hic1 is restricted to quiescent MPs in muscle and other tissues. Single cell RNA-seq was employed to examine Hic1+ cells in muscle and this led to the identification of 3 predominant MP subpopulations with distinct function(s) and lineage potential. Further analysis in muscle injury models revealed that these cells exhibit diverse stage-specific activities, which coordinate multiple aspects of the regenerative process. Following regeneration, Hic1+ progeny contribute to several mesenchymal derivatives including Col22a1-expressing cells in the myotendinous junction.

In numerous tissues, MPs have been found to play a vital role in stem cell niches. Single cell-seq was employed to characterize MPs across tissues, and these analyses revealed extensive intra and inter-tissue heterogeneity. Within some tissues, unique populations could be identified that appeared based on gene expression to have niche-like properties. This was very evident in the populations characterized from bone marrow (BM), which contribute to the hematopoietic stem cell (HSC) niche. Deletion of Hic1 led to widespread MP hyperplasia, including in the BM. Interestingly, this led to a 2-fold increase in HSC number. Taken together, these results
suggest that the Hic1$^+$ MPs contribute to the HSC niche and that MP frequency regulates HSC niche capacity.

In summary, these findings identify $Hic1$ as a marker of MPs, and resulting genetic tools have been instrumental in defining MP subpopulations that exhibit transient and enduring roles in regeneration and non-cell autonomous activity in the HSC niche. This provides a solid foundation for understanding MP biology and their utility in cell-based and/or in situ modification to affect health and disease.
Lay Summary

Mesenchymal stem-like cells (MSCs) are thought to reside in most, if not all, tissues in mice and humans. Biomedical research is currently exploring MSCs for many stem cell therapy applications. This work serves to unify systemic murine MSCs using a genetic marker common across major organs that allows for their tracking, isolation and genetic manipulation. Herein we validate these novel tools and uncover the dynamic response of MSCs to muscle injury. We describe the myotenocyte, a novel MSC subpopulation with the specific responsibility of establishing connective tissue junctions between muscle and tendon. We also provide important insights into the function of MSCs in regeneration and regulating other stem cells. In aggregate, this work implicates MSCs as central players in tissue renewal and regeneration.
Preface

Animal studies were reviewed and approved by the University of British Columbia Committee on Animal Care under protocols A15-0191, A15-0202, A18-0045 and A17-0273.

All studies in this thesis were conceived and designed by R.W.S. and T.M.U. R.W.S. performed all experiments and bioinformatic data analysis. Manuscripts included in this thesis were written by R.W.S. and T.M.U.

A version of chapter 1 has been published. Scott, R.W., Underhill, T.M. (2016) Methods and Strategies for Lineage Tracing of Mesenchymal Progenitor Cells. Methods In Molecular Biology 1416:171-203. R.W.S. optimized all methodology and co-wrote the manuscript with TMU\(^1\).

A version of chapter 2 has been submitted and is currently under review. *Hic1* Marks and Regulates Quiescence of Mesenchymal Progenitors. R. Wilder Scott, Martin Arostegui, Heejung Choi, Ryan Vander Werff\(^1\), Petra Schreiner, Daniel Tausan, Fabio M. Rossi, and T. Michael Underhill.

A significant amount of the data in figures 3.30-3.32 was generated from experiments performed by Phuong Ha Nguyen (Dr. Fabio Rossi’s group) in collaboration with R. Wilder Scott.
# Table of Contents

Abstract .................................................................................................................................................. iii

Lay Summary .......................................................................................................................................... v

Preface ................................................................................................................................................... vi

Table of Contents ................................................................................................................................... vii

List of Tables ......................................................................................................................................... xi

List of Figures ......................................................................................................................................... xii

List of Abbreviations ............................................................................................................................. xv

Acknowledgements ............................................................................................................................... xviii

Dedication ............................................................................................................................................... xix

Chapter 1: Mesenchymal Progenitor cells in health and disease. ............................................................ 1

1.1.1 Mesenchymal cells - overview ...................................................................................................... 1

1.1.2 Are MSCs stem cells? .................................................................................................................... 2

1.1.3 MPs and MSCs in tissue regeneration ....................................................................................... 5

1.1.4 Cre recombinase and reporter expression ............................................................................... 6

1.1.5 Nestin and BM-MSCs .................................................................................................................. 9

1.1.6 LepR and BM-MSCs .................................................................................................................... 10

1.1.7 Pdgfra and Sca-1 – MP and “MSC” lineage markers .................................................................. 11

1.1.8 Gremlin 1 and BM-MSC/SSCs ............................................................................................... 12

1.1.9 Gli1 and tissue-resident MPs .................................................................................................... 14

1.1.10 Lineage markers of MP derivatives ......................................................................................... 15

1.2 MPs and the perivascular niche ...................................................................................................... 16
1.3 Overview of skeletal muscle regeneration ................................................................. 16
1.4 Skeletal muscle and the myotendinous junction ......................................................... 18
1.4.1 Fibrosis – excessive deposition of extracellular matrix ........................................... 19
1.5 The stem cell niche and MPs .................................................................................... 20
1.6 Quiescence – reversible growth arrest ......................................................................... 26
1.6.1 Hypermethylated in Cancer 1 (Hic1) ..................................................................... 28
1.7 Key technological advances ....................................................................................... 29
1.7.1 Genetic labelling, lineage tracing and cell purification ........................................... 29
1.7.2 Querying the transcriptome with next generation sequencing ............................... 31
1.8 Objectives of this work ............................................................................................. 32

Chapter 2: Methods ............................................................................................................ 34
2.1.1 Mice ...................................................................................................................... 34
2.1.2 Generation of anti-HIC1 antibody ......................................................................... 36
2.1.3 Fluorescence-activated cell sorting methodology .................................................. 36
2.1.4 Cell cycle analysis ............................................................................................... 40
2.1.5 Whole-mount LacZ staining ................................................................................ 41
2.1.6 Immunofluorescence and/or in situ X-gal staining ................................................. 41
2.1.7 Bone marrow harvest and flow cytometry .......................................................... 44
2.1.8 Hematopoietic progenitor growth assay in methylcellulose ................................. 44
2.1.9 Competitive bone marrow transplantation experiments ........................................ 45
2.1.10 Secondary bone marrow transplantation experiments ........................................ 45
2.1.11 Limiting dilution transplantation assay .............................................................. 45
2.1.12 In vitro BM-MP culturing experiments .............................................................. 46
2.1.13 Image acquisition and quantification .......................................................... 47
2.1.14 Gene expression profiling – popRNA-seq .................................................. 48
2.1.15 RNA-seq bioinformatic analyses ............................................................. 48
2.1.16 Single cell RNA-seq ................................................................................. 49
2.1.17 Statistical analyses .................................................................................... 50

Chapter 3: Hic1 Marks and Regulates Quiescence of Mesenchymal Progenitors .......55

3.1 Introduction ................................................................................................. 55
3.2 Results ........................................................................................................ 57
  3.2.1 HIC1 marks mesenchymal progenitors within skeletal muscle .............. 57
  3.2.2 Hic1 marks multiple MP populations in skeletal muscle .................... 59
  3.2.3 Hic1+ cells do not contribute to parenchymal or vascular lineages post injury .... 67
  3.2.4 Hic1+ cells coordinate multiple aspects of skeletal muscle regeneration ........ 71
  3.2.5 Hic1+ MP progeny contribute to multiple mesenchymal lineages in regenerated skeletal muscle ............................................................................................................................................................................. 82
  3.2.6 HIC1 regulates MP quiescence .............................................................. 87
  3.2.7 Systemic analysis of MPs – characterization of tdTomato+ cells across tissues .. 94
  3.2.8 ScRNA-seq provides novel insights into MP biology ............................ 96
  3.2.9 Hic1 identifies MPs in bone marrow including Cxcl12 abundant reticular marrow stromal cells. .............................................................................................................................. 104
  3.2.10 MP hyperplasia leads to expansion of the hematopoietic stem cell niche .... 106

Chapter 4: Conclusion ..........................................................................................111

4.1 Overview - Hic1 and MP biology ............................................................... 111
4.2 Case for *Hypermethylated in Cancer 1* representing a useful MP marker for identifying MPs in skeletal muscle........................................................................................................... 112

4.3 Fate mapping of the fibrogenic/adipogenic MP subset following skeletal muscle damage .......................................................................................................................................................................................... 113

4.4 *Hic1* identifies tenogenic, myotenogenic and mural/pericyte MP subsets ................. 115

4.5 The mouse adult mesenchymal progenitor atlas .......................................................................................................................... 116

4.6 Conditional deletion of *Hic1* in adult mice leads to MP hyperplasia ..................... 117

4.7 MP hyperplasia in the bone marrow enhances the “hematopoietic stem cell niche” . 119

4.8 Final remarks .................................................................................................................................................................................. 123

References.....................................................................................................................................................................................................126

Appendices........................................................................................................................................................................................................141

Appendix A - Example R code used to analyze MP datasets and generate figures for the muscle regeneration study........................................................................................................................................... 141

Appendix B - Example R code used to analyze MP datasets and generate figures for the mouse mesenchymal atlas........................................................................................................................................... 151
List of Tables

Table 1.1 Summary of recombinase-dependent reporter lines for lineage tracing studies............ 8
Table 2.1 Antibodies used in chapter 3 results................................................................. 51
List of Figures

Figure 1.1 The mesengenic process. ........................................................................................................ 3
Figure 3.1 | Identification of HicI+ cells in the TA muscle. ................................................................. 58
Figure 3.2 | Localization of Hic1 transcript and protein expression. ................................................. 60
Figure 3.3 | Localization of Hic1nLacZ+ reporter expression.............................................................. 61
Figure 3.4 | Perivascular location of Hic1 positive cells. ................................................................. 62
Figure 3.5 | Generation and characterization of a Hic1CreERT2 knock-in allele for lineage tracing of MPs.......................................................................................................................... 63
Figure 3.6 | Single cell transcriptomic analysis reveals four MP subpopulations at homeostasis in muscle. .................................................................................................................................................. 65
Figure 3.7 | tdTomato fluorescent reporter reveals extensive cytoplasmic ramifications of MPs, extended into their surrounding microenvironment................................................................................. 66
Figure 3.8 | Marked transformation of the quiescent MP phenotype following injury-induced activation........................................................................................................................................ 68
Figure 3.9 | Hic1+ cell dynamics following NTX-induced injury of the TA. ............................................ 69
Figure 3.10 | Hic1+ cells in skeletal muscle proliferate and are migratory after injury induced activation........................................................................................................................................... 70
Figure 3.11 | Hic1+ cells are found in the pancreas and exhibit similar behavior in response to caerulein-induced injury. ...................................................................................................................... 72
Figure 3.12 | Hic1+ cell dynamics following caerulein-induced injury of the pancreas. ....................... 73
Figure 3.13 | The activated MP phenotype displays stage-specific activities indicative of a highly-coordinated response to injury. ...................................................................................................... 75
Figure 3.14 | MPs increase cytokine production for 24-48 hours after skeletal muscle injury before entering the cell cycle. ................................................................. 76

Figure 3.15 | Proliferation peaks between day 2 and 4 in MPs.................................................. 77

Figure 3.16 | MPs participate in extracellular matrix remodeling. ........................................... 80

Figure 3.17 | MPs transiently express periostin before re-expression of basement membrane associated transcripts. ................................................................. 81

Figure 3.18 | Hic1-tracked MPs directly contribute to regeneration of the myotendinous junction, but not the myogenic nor the endothelial compartments. ........................................... 83

Figure 3.19 | TdTomato positive myotenocytes increase in numbers and are found to express the myotendinous junction specific FACET COL22a1, 14 days after muscle injury. .................. 85

Figure 3.20 | Hic1-reporter cell fate mapping demonstrates that these cells contribute to several mesenchymal cell lineages within regenerated skeletal muscle. ........................................... 86

Figure 3.21 | HIC1 regulates MP quiescence—deletion of Hic1 leads to an activated MP-like phenotype and an increase in MP number. ................................................................. 88

Figure 3.22 | Modified MP gene expression profile after deletion of Hic1. ..................................... 90

Figure 3.23 | Deletion of Hic1 in adult mice modifies the MP phenotype. ...................................... 91

Figure 3.24 | Deletion of Hic1 leads to mesenchymal hyperplasia and subsequent accumulation of adipocytes in the pancreas. ................................................................. 92

Figure 3.25 | Hic1 expression unifies a perivascular, stromal, reticular MP population organism wide as described in skeletal muscle. ................................................................. 95

Figure 3.26 | Single cell RNA-seq reveals inter and intra tissue heterogeneity and superpopulations shared among several organs. ................................................................. 97
Figure 3.27 | Single cell RNA-seq reveals inter and intra tissue heterogeneity and subpopulations shared among several organs. ................................................................. 99

Figure 3.28 | MPs across tissues express established mesenchymal markers. ...................... 100

Figure 3.29 | Unbiased shared nearest neighbor clustering classifies MPs from ten tissue libraries into 13 clusters. ................................................................. 103

Figure 3.30 | Hic1 identifies multiple subsets of bone marrow MPs, including Cxcl12 Abundant Reticular marrow stromal cells. ................................................................. 105

Figure 3.31 | MP numbers in the bone marrow increase as a result of Hic1 loss and induce a consequent increase in HSCs. ................................................................. 107

Figure 3.32 | Expansion of Hic1<sup>+</sup> MPs in bone marrow increases functional HSC frequency.  109
List of Abbreviations

ACK ammonium-chloride-potassium
ADSC adipose derived stem cell
AP adipogenic progenitors
APC allophycocyanin
BD Becton Dickinson
BM bone marrow
BM-MSC bone marrow derived mesenchymal stem/stromal cell
BSA bovine serum albumin
CAG synthetic cytomegalovirus enhancer beta actin promoter beta globin splice acceptor sequence
CFU colony forming unit
CFU-F colony forming unit fibroblast
Cre P1 bacteriophage derived Cre recombinase
CreERT2 tamoxifen inducible P1 bacteriophage derived Cre recombinase version 2
DAPI 4′,6-diamidino-2-phenylindole,
DNA deoxyribonucleic acid
EDTA ethylenediaminetetraacetic acid
EdU 5-ethynyl-2’-deoxyuridine
FACS fluorescence activated cell sorting
FAP fibro-adipo progenitor
FBS fetal bovine serum
FCM flow cytometry

FPKM Fragments Per Kilobase of transcript per Million mapped reads

FRT flippase recognition target

GFP green fluorescent protein

HEPES 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HSC hematopoietic stem cell

IF immunofluorescence

MP mesenchymal progenitor

MSC mesenchymal stem/stromal cell

MTJ myotendinous junction

NTX notexin snake venom phospholipase extract

PαS platelet derived growth factor and stem cell antigen surface marker combination

PBS phosphate buffered saline

PBS phosphate buffered saline with Tween 20

PCR polymerase chain reaction

PFA paraformaldehyde

PGK phosphoglycerate kinase

PI propidium Iodide

popRNA-seq population-based RNA sequencing

RNA ribonucleic acid

RNA-seq ribonucleic acid sequencing

RT room temperature

SC single cell
scRNA-seq single cell RNA sequencing
SNN shared nearest neighbor
SSC skeletal stem cell
SVF stromal vascular fraction
SVZ subventricular zone
TA tibialis anterior
TAM tamoxifen
tSNE t-distributed stochastic neighbor embedding
Acknowledgements

I am extremely grateful to Professor T. Michael Underhill for the opportunity to join his research program, an environment where the intellectual and financial resources available make student input the limiting variable. Also, for always remaining accessible for conversations well as providing students unlimited guidance and support through this challenging endeavor. It has truly been a pleasure. It is an understatement to express my eternal gratitude for the faculty of medicine graduate award I had the pleasure of receiving for several consecutive years.

I would also like to express my gratitude to the members of my supervisory committee, Professors Kelly McNagny, Tim O’Connor and Dr. Fabio Rossi for providing me additional positive role models for success inside and outside of the laboratory.

The expert training I received from many people at the BRC, especially Arthur Sampaio and Lin Yi, allowed me to develop the skills necessary to complete this work. Support from several undergraduate research assistants was extremely valuable especially HyunJi Kim who made many sacrifices toward my endeavors, but also Heejung Choi, Germain Ho and Amanda Yang. The many collaborative efforts with the now fellow student Martin Arostegui are particularly appreciated, without whom many experiments would have been limited to my two hands and minimal comic relief. Mentorship and friendship from since graduated fellow student Le Su was very much appreciated; his work ethic and dedication were an inspiration. Correspondence with the Sajita group at NYU was valuable as they were kind enough to reinstate the functional updated Seurat single cell analysis package upon my request. Finally, I am grateful to my dear mother Nancy Lynn Parsons, Karen Jocelyn Scott and Jessica Bryna Cait for unconditional moral support.
Dedication

To my late father, who inspired perpetual intellectual pursuit.
Chapter 1: Mesenchymal Progenitor cells in health and disease.

1.1.1 Mesenchymal cells - overview

Over 40 years ago, Friedenstein and colleagues described the isolation and characterization of a population of bone marrow (BM)-derived cells with colony forming activity\(^2-4\). The colonies derived from these cells appeared to resemble fibroblasts, a stromal cell responsible for extracellular matrix and collagen synthesis and remodeling\(^5,6\). Hence, they were termed colony-forming unit-fibroblast (CFU-F). Fibroblasts are derived from primitive mesenchyme, a loose connective tissue arising from the embryonic mesoderm germ layer\(^5,6\). Indeed, subsequent studies demonstrated that cells from a single CFU-F could variably contribute to various mesenchymal lineages, including but not limited to white adipocytes, chondrocytes and osteoblasts. Based on these properties, and others, BM CFU-Fs were later proposed to derive from mesenchymal stem cells (MSCs)\(^7,8\). The first MSC isolation from human BM was described in the late 1990’s by Pittenger \textit{et al.}, and these cells were found to have similar properties to that of previously described non-human MSCs\(^9\). As these cells were thought to represent endogenous stem cells with the ability to contribute to multiple lineages (even across germ layers), it was believed that they would represent an excellent source of stem cells to affect tissue regeneration. As a result, methods for their collection and expansion were quickly developed, and initial pre-clinical and clinical applications rationally targeted regeneration of skeletal tissues. MSC-based treatments have subsequently been expanded and they are now being “tested” in a myriad of conditions involving diverse tissues and organs. For the most part, MSC-based strategies have not delivered on their early promise, and it is
anticipated that a better understanding of MSC biology will enable development of more efficacious MSC therapeutics\textsuperscript{10}. In this regard, the MSC moniker has now been used to suggest that MSCs represent medicinal signaling cells, which can deliver a variety of factors and molecules to coordinate multiple facets of tissue regeneration, including immunomodulation, stem/progenitor cell function and activity, mechanical stabilization, innervation and revascularization\textsuperscript{11}.

1.1.2 Are MSCs stem cells?

The stem cell connotation brings with it a number of critical and well-defined criteria, which minimally include demonstration of extensive self-renewal activity and uni or multi-lineage contribution (Fig 1.1). Until recently these criteria, especially rigorous demonstration of self-renewal, were not satisfied for MSCs. Evidence for self-renewal of BM-MSCs comes from studies showing that \textit{in vitro} clonally-expanded MSCs, when transplanted into a recipient, are sufficient to recapitulate many aspects of bone, including osteoblasts, chondrocytes, marrow stromal cells and marrow adipocytes\textsuperscript{12,13}. In this regard, these MSC-derived bone spicules or organs are generated through endochondral ossification from donor MSCs and produce a BM microenvironment sufficient to support hematopoiesis of recipient hematopoietic stem cells (HSCs). To more properly reflect the location, function and activity of these cells, some groups have proposed that at least a subset of BM-derived “MSCs” be termed skeletal stem cells (SSCs), as this more accurately reflects their self-renewal and lineage potential\textsuperscript{14}. MSC-like cells in terms of clonogenicity and \textit{in vitro} multi-potency have been found in other tissues and organs; however, their relationship to BM-MSCs and SSCs has remained unclear\textsuperscript{15}. Furthermore, the
Figure 1.1 The mesengenic process.
A mesenchymal stem/progenitor cell is defined as a cell that can self-renew and differentiate into the mesenchymal derivatives shown.
self-renewal potential of these extra-medullary MSC-like cells has not been determined. Most recently, experts across disciplines in the field are proposing the retirement of the term MSC in favor of a new more rigorous classification system for this obviously heterogeneous cellular entity\textsuperscript{16,17}. Thus, pending consensus on the nature and relationship of these various mesenchymal cells, and consistent with their well-defined \textit{in vitro} properties (CFU-F activity and mesenchymal lineage potential), these cells will be referred herein to as mesenchymal progenitors (MPs). This population shares many properties, markers and activity with other described tissue-resident MPs that have been termed mesenchymal stromal cells, mesenchymal progenitor cells (MPCs) or fibro-adipo progenitors (FAPs)\textsuperscript{18,19}. Furthermore, these various described MP-like cell types do share a constellation of surface markers including Sca-1 and Cd140a (platelet-derived growth factor receptor \(\alpha\), PDGFRA), and are negative for blood (Ter119 and CD45) and endothelial (CD31) lineage markers (Lin\textsuperscript{−}). Within the BM, SSCs are located in the perivascular space in close association with sinusoids or arterioles\textsuperscript{10}. Similarly, extramedullary MPs are also often found in the perivascular space\textsuperscript{20}. To complicate matters further, within various adipose depots vessel-associated adipogenic progenitors also express several MP-associated markers (CD140a and Sca-1) and exhibit-overlapping lineage potential\textsuperscript{21}. The extent of overlap and heterogeneity within and between these various mesenchymal populations remains to be determined. The history and our current understanding of the field is described in detail in a relatively recent comprehensive review\textsuperscript{22}. 
1.1.3 MPs and MSCs in tissue regeneration

MSCs have been studied for the longest time in the context of bone. Transplant studies have provided an important framework for defining the nature and potential of these populations, and genetic lineage tracing methodology has now enabled us to identify, track, and study MPs in situ with minimal experimental intervention. Under steady state conditions these cell populations exist in a quiescent state and following exposure to the appropriate stimuli they take on an activated phenotype. In this regard, both BM-MSCs and SSCs expand following injury and their progeny directly contribute to the regenerated bone\textsuperscript{23,24}. Ablation of BM-MSC or SSCs severely compromises bone regeneration, however, it is assumed that this is a consequence of both direct (i.e. direct contribution to the skeleton), and indirect activities including potentially more generic regenerative functions\textsuperscript{23-25}. In many respects, extramedullary MPs exhibit similar activities and responses to injury. In skeletal muscle, a number of groups have shown that MPs participate in regeneration by providing trophic factor support for other muscle stem cells, such as satellite stem cells, and that their partial ablation impacts the timing of satellite cell differentiation and consequently myofiber regeneration\textsuperscript{18,19,26}. Similarly, in the mammalian heart and kidney (labeled with other markers), a tissue-resident MP population is quickly activated following injury\textsuperscript{27,28}. Regeneration is incomplete in these models, the progeny of these cells become myofibroblasts that in part underlie the excess connective tissue deposition associated with tissue repair and fibrosis.Interestingly, MP-like cells within the liver, hepatic stellate cells (LHSCs), also exhibit similar functional properties following injury and can also contribute to tissue fibrosis in this organ\textsuperscript{29}. In all of these situations, a quiescent MP-like cell is activated in response to injury and participates directly in the regenerative and/or reparative process.
Furthermore, BM-MSC/SSC and MPs share the expression of many markers and exhibit overlapping lineage potentials at least in vitro\textsuperscript{10,22,30}. Not surprisingly, however, BM-MSC/SSCs in comparison to MPs have unique properties, most notably the ability to generate an entire bone organ from clonally-expanded transplanted cells\textsuperscript{10}. In this regard, BM-MSC/SSC appear to exhibit “higher” intrinsic bone organ-forming activity, and this may in part reflect their biological function. Under pathological conditions such as in fibrodysplasia ossificans progressiva (FOP), MPs in other sites can also contribute to endochondral bone formation\textsuperscript{31}, but this may partly reflect enforced differentiation that is accompanied by expression of activated ALK2 receptor\textsuperscript{14}. Heterotopic bone formation is also observed in various diseases and following injury, and presumably tissue-resident MPs participate in this process, but this requires further study. In short, there are numerous functional parallels between BM-MSC/SSCs and extramedullary MPs, but also distinct differences and it is expected that strategies involving new and existing genetic tools will be crucial to furthering our understanding of these important stem/progenitor populations.

1.1.4 Cre recombinase and reporter expression

The aforementioned studies have relied greatly on genetic methods for the in situ labeling of MPs, coupled with the ability to track their progeny. Minimally this involves the use of a Cre (or other recombinase such a FRT) to induce the stable expression of a unique reporter gene\textsuperscript{32,33}. In this case, the expression of the reporter is absolutely dependent on Cre and following Cre mediated genome modification the reporter is expressed and continues to be expressed irrespective of the cellular context (i.e. stem, progenitor, differentiated cell, etc.). Typically, the
promoter driving the reporter is widely expressed in most, if not all, cell types. For these purposes, the Rosa26 promoter or the CAG promoter have proven popular and reasonably effective. A number of the more common reporter lines are highlighted in Table 1.134-42. The CAG promoter combines a CMV-derived enhancer with the chicken beta actin promoter, and provides high-level expression in a wide-variety of cell types. In the absence of Cre, the reporter of interest is not transcribed and following Cre-mediated deletion of an upstream cassette harboring repeated transcriptional stop signals or inversion of the region, reporter expression ensues. Importantly, the Cre-modified locus is inherited in all progeny, thereby enabling individual cells to be labeled and their progeny followed over time. The promoter driving the Cre determines which cell type(s) the Cre is expressed in. A large number of Cre lines have been generated, which can be used to label various MPs in various tissue compartments (see below). The addition of a modified estrogen ligand-binding domain (ER or ERT2) to the Cre, has provided temporal control of nuclear Cre activity32. In this manner, tamoxifen (TAM) treatment drives the translocation of the CreER/ERT2 from the cytoplasm into the nucleus thereby initiating the desired genomic modification, resulting in reporter gene transcription. For lineage tracing, there are advantages and disadvantages to constitutive and inducible Cres, however, the ability to temporally control Cre activity provides a major benefit in being able to carry out pulse-chase experiments at any age, where animals are briefly exposed to TAM and the fate of labeled cells can be assessed after days, months or even years32.
<table>
<thead>
<tr>
<th>Reporter Line</th>
<th>5' Component</th>
<th>3' Component</th>
<th>Promoter</th>
<th>Insertion Locus</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Availability*</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confetti</td>
<td>Multiple FPs</td>
<td>Multiple FPs</td>
<td>CAG</td>
<td>Rosa26</td>
<td>- multicolor, stochastic deletion to aid clonal analysis. - homozygous reporter mice can be used to create additional colors. - can be used for re-tracing of clones.</td>
<td>- requires strong Cre activity for good recombination. - some of the FPs are weaker and more difficult to detect. - requires cryosectioning to maintain FP activity.</td>
<td>Jax strain 017492</td>
<td>33</td>
</tr>
<tr>
<td>mTomato/mEGFP</td>
<td>mTomato</td>
<td>mEGFP</td>
<td>CAG</td>
<td>Rosa26</td>
<td>- membrane-localized FP enables visualization of cell morphology. - good for FCM and FACS, as &quot;cleaner&quot; populations can be identified. - useful for clonal analysis with low dose TAM. - FP expression appears to be well tolerated. - under certain conditions, membrane-bound FP may contribute to the ECM.</td>
<td>- requires moderate Cre activity for efficient recombination. - high levels of FP may be toxic in some cell types.</td>
<td>Jax strain 007676</td>
<td>34</td>
</tr>
<tr>
<td>nTomato/nEGFP</td>
<td>nTomato</td>
<td>nEGFP</td>
<td>CAG</td>
<td>Rosa26</td>
<td>- nuclear localized FP - good for FCM and FACS, as &quot;cleaner&quot; populations can be identified. - useful for clonal analysis with low dose TAM.</td>
<td>- typically requires moderate Cre activity for efficient recombination. - high levels of FP may be toxic in some cell types.</td>
<td>Jax strain 023035</td>
<td>35</td>
</tr>
<tr>
<td>LSL-tdTomato</td>
<td>stop cassette</td>
<td>dTomato</td>
<td>CAG</td>
<td>Rosa26</td>
<td>- strong FP expression. - very efficient deletion with even weak Cre activity.</td>
<td>- high levels of FP may be toxic in some cell types. - potential for greater leakiness due to efficient Cre-mediated recombination.</td>
<td>Jax strain 007914</td>
<td>36</td>
</tr>
<tr>
<td>LSL-LacZ</td>
<td>stop cassette</td>
<td>LacZ</td>
<td>Rosa26</td>
<td>Rosa26</td>
<td>- no autofluorescence issues. - stable signal, that is suitable for paraffin embedding.</td>
<td>- endogenous beta-galactosidase activity can be an issue. - detection in adult tissues with antibodies can be difficult. - requires specialized processing of tissue to maintain LacZ activity. - detection can be challenging depending on the tissue. - typically requires a decent Cre driver for good deletion.</td>
<td>Jax strain 003474</td>
<td>37</td>
</tr>
<tr>
<td>LSL-EYFP</td>
<td>stop cassette</td>
<td>EYFP</td>
<td>CAG</td>
<td>Rosa26</td>
<td>- efficient Cre-mediated recombination and reporter expression</td>
<td>- weaker FP, fluorescence overlaps with AF.</td>
<td>Jax strain 012567</td>
<td>38</td>
</tr>
<tr>
<td>LSL-ZsGreen1</td>
<td>stop cassette</td>
<td>ZsGreen1</td>
<td>CAG</td>
<td>Rosa26</td>
<td>- strong green FP, efficient Cre deletion with lower Cre expression.</td>
<td>- potential for greater leakiness due to efficient Cre-mediated recombination.</td>
<td>Jax strain 007906</td>
<td>39</td>
</tr>
<tr>
<td>FSF-LSL-tdTomato</td>
<td>2 stop</td>
<td>tdTomato</td>
<td>CAG</td>
<td>Rosa26</td>
<td>- reporter gene expression requires both Cre and Flp recombinase - allows labelling of cells based on two markers.</td>
<td>- requires mice lines that expression both recombinases.</td>
<td>Jax strain 021875</td>
<td>40</td>
</tr>
<tr>
<td>LSL-GFP-NLS-LacZ</td>
<td>stop</td>
<td>GFP-NLS-LacZ</td>
<td>Rosa26</td>
<td>Rosa26</td>
<td>- dual reporter with GFP and LacZ</td>
<td>- lower</td>
<td>Jax strain 008606</td>
<td>41</td>
</tr>
</tbody>
</table>

* Availability of mice is shown for the C57BL/6 background and is subject to change.

Table 1.1 Summary of recombinase-dependent reporter lines for lineage tracing studies
A wide variety of Cre lines have been generated, some of which have been successfully applied to the study of MPs and BM-MSC/SSC/MPS\textsuperscript{1,23,24,43-46}. Many of these were used to study BM-MSC/SSCs but may also have utility in investigating extramedullary MPs. At present, Cre lines that universally and specifically label MPs in all tissues and organs have not been identified. This may reflect the possibility that such markers do not exist, or we have yet to identify such a marker(s).

1.1.5 "Nestin and BM-MSCs"

Until recently our understanding of MP and SSC biology relied extensively on their isolation, culture and transplantation\textsuperscript{30,47}. In the past few years, several new genetic strategies have been used to identify both MPs and SSCs \textit{in situ}. Analysis of nestin-GFP transgenic mice identified a rare population of GFP\textsuperscript{+} cells within the BM\textsuperscript{13,48}. Further analysis of this population showed that they had appreciable CFU-F activity and \textit{in vitro} contributed to several mesenchymal lineages. More significantly, this population was found to reside in close proximity to blood vessels and HSCs, forming an important component of the HSC niche. Subsequent studies have demonstrated that the GFP\textsuperscript{+} cells could be separated into GFP\textsuperscript{hi} and GFP\textsuperscript{lo} with the former representing quiescent BM-MSCs within the HSC niche\textsuperscript{48}. The nestin-GFP transgenic lines have been useful for identifying BM-MSCs\textsuperscript{49}, but \textit{Nestin} itself as a marker has proven unreliable. Furthermore, lineage tracing with a \textit{Nestin}\textsuperscript{CreER} line has proven uninformative as it labels multiple populations within the BM including CD31\textsuperscript{+} endothelial cells\textsuperscript{23,24,50}. 
1.1.6 *LepR* and BM-MSCs

The Leptin receptor (*LepR*) has been recently identified as a useful marker for identifying BM-MSCs\(^{24,51}\). *LepR*-expressing cells within the BM were originally shown to overlap significantly with a group of BM cells defined by high levels of the cytokine CXCL12, these were termed CXCL12-abundant reticular cells or CARs\(^{51}\). Deletion of CXCL12 in *LepR*-expressing cells using a *Lepr\(^\text{Cre}\)* line demonstrated that CARs provided an essential source of several HSC maintenance factors including in addition to CXCL12, also scatter cell factor (SCF or Kit ligand). The *Lepr\(^\text{Cre}\)* line has recently been used to systematically study the nature of *LepR*\(^{+}\) cells within the BM, and these analyses demonstrated that this population overlaps extensively with what previously would be considered a BM-MSC\(^{24}\). Consistent with previous reports, isolated *LepR\(^{\text{Cre}}\)*-marked cells are positive for a number of MP/MSC-associated markers including Cd140a, Cd140b (Pdgfr\(\beta\)), CD105, CD51, but are low for Nes\(^{\text{GFP}}\). Interestingly, *LepR*\(^{+}\) cells exhibit limited contribution to pre-natal skeletal cells and post-natal skeletal growth, but afterwards contribute substantially to newly formed osteoblasts during skeletal turnover\(^{24}\). Furthermore, with aging, a large proportion of the adipocytes derive from a *LepR*\(^{+}\) cell. Following skeletal injury, the *LepR*\(^{+}\) cells contribute to both chondrocytic and osteogenic cells in the fracture callus. *In vitro* analyses demonstrated that *LepR*\(^{+}\)-derived cells represented the bulk of CFU-F activity within the BM and tested clones exhibited variable mesenchymal lineage potential. Analysis of the *Lepr\(^\text{Cre}\)* line has provided a valuable new marker for studying MSC biology, however, the non-inducible nature of the Cre has precluded more sophisticated lineage tracing approaches involving pulse-chase labeling and analysis. *LepR* was originally identified in the choroid plexus, which is rich in MP-like cells\(^{52}\) and was subsequently found to be
expressed in the hypothalamus and other brain regions, with limited expression in other tissues. However, it remains unclear if \textit{Lepr} represents a more generic marker of tissue-resident MP/MSCs. Furthermore, it remains to be determined if \textit{Lepr} expression in BM stromal cells defines the entire, or a subset, of the BM-MSC population and/or is associated with a specific BM-MSC function(s). These comments notwithstanding, analysis of the \textit{Lepr}^{Cre} mouse has provided important insights into the nature of BM-MSCs, their contribution to bone renewal and regeneration, and their \textit{in vivo} lineage potential.

\subsection{1.1.7 \textit{Pdgfra} and \textit{Sca-1} – MP and “MSC” lineage markers}

The stem cell antigen-1 (Sca-1) surface marker has been identified on a number of stem and progenitor cells\textsuperscript{53-55}. Mice deficient for Sca-1 display an osteopenic phenotype consistent with disruption of the MSC compartment in bone\textsuperscript{54}. The utility of Sca-1 as an MSC marker is problematic due to its widespread expression in other cell types including CD31\textsuperscript{+} endothelial cells. In recent studies, prospective isolation of mouse MSCs has been achieved using Sca-1 in combination with PDGFRA co-staining\textsuperscript{56,57}. \textit{Pdgfra} is expressed in mesenchymal cells\textsuperscript{58}. The PDGFRA-Sca-1 doubly positive population (P\alpha S) exhibited a much higher CFU-F forming activity than either singly-positive population\textsuperscript{57}. The P\alpha S fraction also exhibited many of the standard properties associated with MSCs, such as extensive \textit{in vitro} expansion, multi-lineage potential, and localization to the perivascular space \textit{in vivo}. \textit{Pdgfra} expression has also been used to identify MPs in several other tissues, including in skeletal muscle, lung, heart and fat depots\textsuperscript{19,27,59,60}. In the brain, \textit{Pdgfra} is also expressed in oligodendrocyte progenitors as well as in other sites both in the brain and body\textsuperscript{58,61}. Several groups have also demonstrated that \textit{Pdgfra}
is a useful marker for identifying both white and brown adipogenic progenitor cells (APs)\textsuperscript{59,62}. Studies carried out with $Pdgfra^{\text{Cre}}$ and $Pdgfra^{\text{CreERT2}}$ have shown that the majority of white adipocytes in both subcutaneous and epididymal fat can be labeled. The relationship of pre-adipocytes to MPs is still being resolved, however, it is likely that APs at least are a component of the adipose vascular stromal fraction, which has been shown to contain MPs with multi-potent lineage activity. The expression in these other cell populations limits the usefulness of $Pdgfra$ as a pan-MP/MSC marker, however, the degree of expression in some tissue compartments may be negligible thereby making the $Pdgfra^{\text{Cre}}$-based genetic tools useful for studying MPs with the noted caveats.

1.1.8 Gremlin 1 and BM-MSC/SSCs

A promising marker that has recently emerged for identifying BM-MSC/SSCs is $Gremlin1$ ($Grem1$)\textsuperscript{23,63}. GREMLIN1 acts as an extracellular bone morphogenetic protein (BMP) binding antagonist, with increased preference for BMP2 and BMP4\textsuperscript{64}. Genetic deletion of $Bmp2$ has limited impact on skeletal development, however, these mice exhibit increased fracture rates and a greatly reduced ability to initiate fracture repair\textsuperscript{65}. In earlier studies, $Grem1$ expression was associated with BM-MSCs\textsuperscript{63} and to further examine the utility of this marker for identifying BM-MSCs a mouse line harboring a $Grem1^{\text{CreERT}}$ transgenic BAC was generated\textsuperscript{23}. Consistent with other studies, $Grem1$-reporter positive cells exhibited variable expression of a number of MP/MSC markers including CD105 (endoglin)\textsuperscript{23}. In contrast to findings with the $Lepr^{\text{Cre}}$ mice, the $Grem1^{\text{CreERT}}$ mice enabled labeling of a pre-natal population with high skeletogenic potential. Similarly to the $Lepr^{\text{Cre}}$ mice, both lines identified a population that contributed substantially to
the fracture callus following bone injury and normal bone turnover. Interestingly, unlike labeling with the \textit{Lepr}^{\text{Cre}} mice, lineage tracing with the \textit{Grem1}^{\text{CreERT}} line yielded very few label positive BM adipocytes\textsuperscript{23}. Moreover, this line was also used to identify rare MP-like intestinal reticular cells associated with the vascular plexus within villi and also contributing to the periepithelial mesenchymal sheath within the crypts. The relationship between the \textit{Grem1}^{+} and \textit{LepR}^{+} BM populations needs additional study. The \textit{in vitro} properties of both populations overlap greatly, however, the lineage tracing studies have demonstrated several unique features not evident from the \textit{in vitro} studies. Together, analyses with these various Cre lines further supports the idea of a heterogeneous BM mesenchymal stem/progenitor cell population that participate in overlapping but not completely redundant BM functions. In this regard, the BM-MSC/SSC compartment can be minimally fractionated into at least two complementary populations, Grem1\textsuperscript{+} and LepR\textsuperscript{+}. The former population is associated with bone development, growth, renewal and regeneration (SSC)\textsuperscript{47}, while the later has a potentially more prominent role in the HSC niche, but also participates in post-natal bone regeneration and renewal. At this time, it remains unclear if the LepR\textsuperscript{+} cells represent a heterogeneous population, some of which inhabit the HSC niche. Another group identified mouse SSCs using the absence of lineage markers (Lin), Thy1, 6C3, CD105 and presence of AlphaV and CD200 surface markers and further fractionated the population and derivatives into eight distinct subsets using various combinations of the aforementioned markers\textsuperscript{66}. Most recently, a SSC has been isolated from human BM using the presence of the surface markers CD164, PDPN, CD73 and the absence of Lin and CD146. These populations were shown to be capable of generating cartilage and marrow containing ossicles when transplanted under the mouse kidney capsule\textsuperscript{67}. 
1.1.9 *Gli1* and tissue-resident MPs

The hedgehog (HH) signaling pathway plays an important role in multiple facets of stem/progenitor cell biology\(^\text{68}\). Gli1, a downstream constituent of the HH signaling pathway has recently been identified as a useful marker for identifying perivascular MPs\(^\text{44}\). *Gli1\(^{\text{CreERT2}}\)* mice were used to effectively label MPs in multiple tissues including, kidney, liver, lung, heart, skeletal muscle and BM\(^\text{44}\). Across these various tissues, labeled cells were found to co-express CD29, Sca-1, CD44 and CD105, with variable expression of PDGFRB, and were negative for CD45 and CD31. In the BM, as in other tissues, the Gli1\(^+\) cells were found in the perivascular space. In liver, lung and heart damage models, Gli1\(^+\) labelled cells acquired an activated phenotype following injury. In some of these models, Gli1-traced derivatives gave rise to a large number of aSMA\(^+\) myofibroblasts that had a substantive contribution to tissue fibrosis.

Numerous studies have shown an important role for Transforming growth factor β (TGFβ) in driving the emergence of aSMA\(^+\) myofibroblasts in fibrosis\(^\text{69}\). Consistent with these observations, Gli1\(^+\) MPs isolated from various tissues efficiently formed myofibroblasts following exposure to TGFβ1\(^\text{44}\). Together, these findings and others, indicate that MPs play an important role in tissue repair and fibrosis following injury and presumably in chronic disease. Similar to other MP markers, *Gli1* is also expressed in other cell types especially within the brain\(^\text{70,71}\), and this needs to be considered when using this line for lineage analysis of MPs.
1.1.10 Lineage markers of MP derivatives

MPs contribute to a number of mesenchymal lineages that are characterized by the expression of lineage-associated (but not always defining) transcription factors such as RUNX2, SOX9, SP7 (Osterix), peroxisome proliferation-activated receptor γ (PPARG), and Scleraxis (SCX). Runx2 and SP7 are expressed within the osteogenic lineage in addition to hypertrophic chondrocytes, whereas, PPARG and SCX are found within the adipogenic and tenogenic lineages, respectively. SOX9 is critically important in chondrogenesis, but also appears in a number of disparate cell types across other germ layers. Lineage tracing with SP7 has been useful for identifying MSC/SSCs with osteogenic potential within the BM, and in turn these studies have yielded important insights into MP behavior and contribution to bone lineages. The utility of these various markers for studying MP activity outside the BM remains to be defined. Prrxl is a pan-MP marker, and transgenic Cre lines harboring a portion of the Prrxl promoter have been used extensively for studying MP-like cells in limb skeletal development. Prrxl is expressed in adult MPs and thus provides a potentially useful tool for studying post-natal MP biology. Interestingly, Prrxl-labelled MPs contribute to subcutaneous fat, but have limited contribution to visceral white or brown fat. Thus, Prrxl may be useful for identifying and studying MPs within specific anatomical compartments. Lineage analysis has provided numerous important insights into MP biology in health and disease. As highlighted above, the different CREs exhibit advantages and disadvantages for the effective and specific labeling of MPs within various tissues and organs.
1.2 MPs and the perivascular niche

As the commonly used name Mesenchymal Stromal Cell implies, MPs are found within tissue stroma, or non-parenchymal compartment with at least endothelial cells, tissue resident macrophages/inflammatory cells, adipocytes and fibroblasts. MPs themselves are predominantly perivascular and possess intricate dendritic processes that span many endothelial cells. In recent studies, two groups have purified and described the lineage potential of MPs from skeletal muscle. In the paper by Joe et al., skeletal muscle resident MPs, aptly described based on their in vitro potential as Fibroadipogenic progenitors (FAPs), were enriched by the presence of cell surface markers CD34 and Sca-1 (Ly6a), and the absence of vascular endothelial marker CD31 and BM derived lineage markers including CD45. Concurrently Uzemi et al. reported isolation of a similar population using the cell surface marker platelet derived growth factor alpha (PDGFRα). This Lin− PDGFRα+ Sca-1+ (PαS) surface marker combination has been reportedly used to isolate an MP like population across a number of tissues, although this marker combination has proven unreliable for the isolation of similar populations from brain and adrenal gland. When purified and cultured these populations are capable of spontaneous and/or guided/induced differentiation into at least chondrocytes, adipocytes, osteoblasts and myofibroblasts.

1.3 Overview of skeletal muscle regeneration

Skeletal muscle exhibits potent regenerative capacity and as such serves as a useful model system to study mechanisms underlying tissue regeneration, and the fate and function of
stem and progenitor cells therein. Several skeletal muscle injury methods are currently in use to induce regeneration including physical injury by crush or repetitive micro abrasion, freezing, chemical agents such as BaCl$_2$ and glycerol, and myonecrotic agents such as cardiotoxin and notexin (NTX)$^{76-79}$. Reproducible skeletal muscle damage can be induced by NTX, a phospholipase extract from the venom of the Australian Tiger Snake (*Notechis Scutatus*). This enzyme cocktail is mostly specific to myofiber membrane phospholipids and by causing perforation of the sarcolemma, allows calcium influx and resultant hyper contraction leading to traumatic injury of the tissue. The tibialis anterior muscle (TA) is chosen for reproducibility as it is easily accessible and a small linear muscle bundle that is conducive to consistent distribution of damage and reliability of collection and processing. When properly administered, NTX can be considered a sterile injury that also results in a robust innate inflammatory response, including recruitment of at least blood borne neutrophils, eosinophils and monocytes to the site of injury. FAPs/MPs are known to proliferate and subsequently contribute adipose and structural connective tissue during regeneration in this model$^{18,26,80}$. The NTX model while leading to myonecrosis, has minimal effect on the health of other skeletal muscle resident cells including FAPs/MPs$^{18,81}$. Shortly after injury, a large percentage of this population (typically > 35%) were shown to enter the cell cycle$^{18}$. Likewise, skeletal muscle satellite stem cells also enter the cell cycle after injury and appear to undergo asymmetric division, thereby simultaneously maintaining the satellite stem cell pool and providing a source of myoblasts. Subsequently, in the latter stages of this regenerative process, these myoblasts align, and fuse to existing fibers or generate new fibers leading to the regeneration of multi-nucleated myofibres$^{82}$. Following expansion, MPs form a mesh-like network with overlapping processes that appear to impact multiple facets of the regenerative process, but this remains poorly understood. There is also
significant crosstalk amongst the various cell types supporting regeneration, but this is also largely unexplored. Components of the immune system play a significant role in regeneration, and much of the course of inflammatory events during the NTX induced injury model has been described\textsuperscript{81}. An infiltrate of at least neutrophils, eosinophils and monocytes inundates the damaged skeletal muscle within the first 24-48 hours\textsuperscript{83,84}. MPs have been shown to interact with the immune system through both the exchange of soluble factors and direct physical interaction with particularly innate inflammatory cells\textsuperscript{81,83}. Importantly, activated perivascular MPs are reported to become the fibrogenic myofibroblasts reflective of adaptive fibrosis during normal tissue regeneration, but are also retained and responsible for excessive connective tissue deposition in pathological fibrosis\textsuperscript{44}.

1.4 Skeletal muscle and the myotendinous junction

The myotendinous junction (MTJ) is a specific mechanical transduction system that exists at the interface between skeletal muscle and tendon fibers. The structure exists to effectively transmit force from the contractile machinery of the skeletal muscle sarcomere through the tendon to the bone to enable movement. The junction consists of a basket like structure that envelops the distal region of the myofiber and contains fingerlike projections that increase the contact surface between skeletal muscle and tendon\textsuperscript{85}. The fibril-associated collagen with interrupted triple helix COL22A1 has been shown to be specific to MTJs\textsuperscript{86,87}. In zebrafish, \textit{col22a1} knockouts present with muscular dystrophy due to MTJ disruption. Based on the same zebrafish work it was reported that detection of \textit{col22a1} during development by \textit{in situ} hybridization reveals diffuse localization to the somites with no signal at the myotome.
boundaries. The signal is reported to concentrate near the skeletal muscle cell extremities at later stages of development and is thought to be produced primarily by the myocytes\textsuperscript{85}. During zebrafish development myoseptal fibroblasts invade the myoseptal stroma and a subsequent dense collagen network is formed. Myoseptal fibroblasts express the tendon markers collagens 1 and 4, however, descriptions of the mouse MTJ remain limited\textsuperscript{88}. Following contractile muscle injury, the MTJ also sustains damage, however, at present inadequate information exists as to the cellular source of \textit{Col22a1} during regeneration of the MTJ. Finally, a recent study reveals age related changes in the MTJ, suggesting the observed lengthening of the structure could contribute to the skeletal muscle weakness associated with aging\textsuperscript{89}.

1.4.1 \textbf{Fibrosis – excessive deposition of extracellular matrix}

When these MPs adopt the fibrogenic myofibroblast phenotype and produce connective tissue and contractile machinery, it can be very beneficial to tissue regeneration. Complete regeneration and restoration of normal tissue function relies on the generation of a transitional or provisional extracellular matrix that provides essential scaffolding functions and wound stabilization\textsuperscript{26}. However, if this matrix is not subsequently remodeled and replaced by parenchymal cells, tissue function is lost. This occurs under conditions in which fibrogenic cells are not removed and/or there is continued aggravation within the regenerating tissue\textsuperscript{81}. When this phenomenon is allowed to progress, adaptive provisional matrix is not replaced by functional tissue and this is deemed pathological fibrosis. Consequently, tissue function is lost at the expense of tissue integrity. A prime example of this is seen in cirrhosis of the liver, where chronic repetitive injury/insult leads to liver failure due to replacement of hepatocytes by fibrotic
material. This is a pathological response and it could be said that MPs are the cell of origin. In studies from the Humphreys group, they have used a *Gli1*-based CreERT2 knock-in line in combination with Cre-dependent reporter and genetic ablation strategies to elegantly show that Gli1+ MPs are the source of myofibroblasts in several major organs. This pathological response is referred to as repair and often replaces regeneration in disease contexts. The switch from a regenerative to a more reparative phenotype is observed with aging where with a decrease in the capacity to regenerate, and renew, tissues become progressively more fibrotic. Tissue resident MPs are thought to exist in a quiescent state, described in section 1.5, and there are gaps in our understanding of their role in homeostasis and regeneration.

### 1.5 The stem cell niche and MPs

Another recurrent theme in MP biology is their participation in various stem and progenitor cell niches. The concept of a niche as a critical component of stem and progenitor cell biology has gained considerable traction over the last two decades and was based on observations in numerous systems that stem and progenitor cell behavior relies extensively on environmental cues. The niche is comprised of a collection of extrinsic signals within the local stem and progenitor cell microenvironment that shape their fate and function. Extracellular matrix structure and factors therein, auto and paracrine signals, cell contact and metabolic conditions all contribute to the specific nature of a niche. In many instances, niches are found in close association with blood vessels, a microenvironment that also contains perivascular stromal cells; the major focus of this section.
Niches come in a variety of shapes and sizes and are a critical component of tissues with tissue renewal and/or regenerative ability. In the liver, recent studies have shown that Axin2\(^+\) hepatocytes with progenitor cell-like properties reside in a niche adjacent to the central vein within the liver lobule. These cells appear to play a critical role in generating new hepatocytes as part of liver tissue renewal. The question remains, however, whether this exclusive potential is dependent on the cell type or the cell type is dependent on the microenvironment to retain stemness\(^92\). Other studies suggest self-renewing hepatocytes are distributed among all lobular zones and reconstitute the liver after injury\(^93\). In the brain the sub ventricular zone (SVZ) region is enriched for at least olfactory and perhaps other neural stem cells\(^94\). Definitive studies in the small and large intestine have identified Lgr5\(^+\) stem cells within the base of the crypts that are responsible for renewal of intestinal epithelia, and these populations are supported by a niche that contains in part Pdgfra\(^+\) and Gli1\(^+\) stromal cells\(^43,95-97\).

Several studies have described transient niches during development. Among others, the fetal liver harbors HSCs until the marrow is formed\(^98\). Regeneration is thought to recapitulate many aspects of development by re-initiation of developmental programs and, consistent with this, transient niches are thought to support tissue-resident stem/progenitor cells in tissue regeneration. For example, FAPs and skeletal muscle satellite stem cells both enter the cell cycle following injury with the former cell type providing critical trophic factor support\(^18\). The adult mammalian heart cannot regenerate, and this is likely a consequence of the absence of cardiac stem cells to support the generation of new cardiomyocytes\(^99\).

The prototypical stem cell compartment is the hematopoietic stem cell (HSC) and its study has shaped our definition of a stem cell. HSCs are endowed with the incredible ability to reconstitute the entire blood system through transplantation of a single cell in a sub lethally
irradiated mouse\textsuperscript{100}. An HSC can theoretically self-renew indefinitely, and this has been shown by serial transplantation of individual HSCs into recipients. The HSC cell cycle is regulated to meet demands of the blood system and there is heterogeneity within the HSC compartment at multiple levels\textsuperscript{100}. The main site of HSC maintenance and differentiation, where blood cells are formed, is the BM, which also represents one of the best-characterized niches to date. The complete nature of the HSC niche remains unclear though many advances have been made in characterizing key cellular and molecular components, but also the importance of niche location and its spatial organization needs to be considered.

Stromal cells are indispensable participants in numerous stem cell niches and in the BM, several signaling axes between BM-resident MPs (BM-MPs) and HSCs have been identified. A number of groups have recently established that BM-MPs play a vital role in the maintenance and regulation of hematopoiesis and serve as integral constituents of the HSC niche (reviewed in \textsuperscript{49,101}). Elegant studies from the Morrison group have since identified Lepr\textsuperscript{+} stromal cells as the major CFU-F forming population in BM. These cells also produce a number of factors, including CXCL12 and SCF (encoded by the Kitl gene) experimentally defined as essential for HSC maintenance and regulation. In this context, MPs provide trophic support that involves expression of various cytokines, molecules and factors. Conditional deletion of Cxcl12 within various proposed niche-associated cell types demonstrated that perivascular Lepr\textsuperscript{+} MPs are also essential for maintenance of HSC quiescence and re-populating activity\textsuperscript{51}. LeprCre mediated deletion of Cxcl12 reduces HSC retention in the BM without affecting HSC number. In contrast, Cxcl12 deletion by Prxr1Cre leads to severe HSC depletion. Further studies demonstrated that the niche factors likely act locally in a paracrine fashion since mice that are chimeric for the membrane bound SCF have normal HSCs only in close proximity to wild-type cells\textsuperscript{102}. 

22
Regulation of the release of Cxcl12 and HSC turnover has been tied to the nervous system in addition to a stem cell clock synchronized to the circadian rhythm\textsuperscript{103}.

Several \textit{in vivo} studies have identified essential factors required in support of the HSC niche, interestingly no known factor combination has been identified that is sufficient to sustain HSCs in culture. In this regard, it seems likely that other aspects of the niche are also required some of which are physical in nature, and more challenging to effectively model \textit{in vitro}. Some of these features likely include the three dimensional architecture of the niche, cell-cell interactions, oxygen tension and nutrient availability which require further exploration\textsuperscript{104}.

Not surprisingly, within the BM, HSC heterogeneity has been well defined, and it has been proposed that the periarteriolar niche supports greater quiescence compared to the perisinusoidal niche. At present, however, this remains inconclusive\textsuperscript{48}. Certainly, sinusoids are particularly vulnerable to irradiation, so periarteriolar HSCs may be important in regeneration of the blood system after irradiation\textsuperscript{105,106}. HSCs are also thought to require low oxygen tension and it follows then, that they are reported to reside in the hypoxic sinusoidal regions\textsuperscript{107,108}. Consistent with the observations that most HSCs within the BM reside in close proximity to blood vessels, especially sinusoids\textsuperscript{12,13}, endothelium is another source of CXCL12. However, endothelial cell-production has been reported to be 100-fold lower than CARs. Ablation studies of vascular cells may be confounded and less conclusive as blood flow to the niche may be compromised\textsuperscript{104}. Endothelial specific conditional deletion of SCF however causes HSC depletion without consequences to blood vessel integrity or stromal cell frequency\textsuperscript{109}. This study did not examine if there was potential compensatory up-regulation of these factors in associated stromal cells. The gold standard for analysis of HSC function and indirectly the non-cell autonomous capacity of the niche is evaluation of blood system reconstitution following
transplantation into non-lethally irradiated mice\textsuperscript{110}. Deletion of \textit{Scf} from both endothelial and Lepr\textsuperscript{+} cells results in compete loss of both HSC quiescence and serially transplantable HSCs\textsuperscript{111}.

Osteoblasts and especially bone lining cells have been also proposed to be key constituent of the HSC niche as these cells represent an additional source of CXCL12. However, osteoblasts express 1,000-fold lower levels of \textit{Cxcl12} than CAR cells and their ablation had a negligible effect on HSC frequency. Furthermore, they do not produce appreciable amounts of SCF and are not required for generation and/or maintenance of extramedullary niches\textsuperscript{104,109,112}.

Adipocytes in the marrow have also been observed to provide key factors that define the molecular components of the niche under certain conditions. A recent report demonstrated that a subset of Lepr\textsuperscript{+} stromal cells is biased to express an adipogenic fate. After irradiation, but not at homeostasis, adipocytes became a unique and necessary source of SCF for the regenerating hematopoietic system in the long bones, however, this mechanism does not appear to be operating in other skeletal elements such as the caudal vertebrae\textsuperscript{113}. Increased adiposity is a common feature in the aged BM\textsuperscript{113}. MPs are also responsible for transient adipocyte generation in other regenerative contexts where local energy stores are required\textsuperscript{18}. Perhaps the plasticity between these various mesenchymal cell types and their ability to support local adipogenesis enables them to serve as atypical transient contributors to the HSC niche in the marrow. To resolve these questions, further studies are required to address whether this phenomenon is unique to the BM or if adipocytes derived from other tissues are functionally equivalent if appropriately modified or conditioned.

Under pathological conditions, peripheral sites have also been reported to support and contain HSCs. In particular, along with the liver, the spleen can support extramedullary hematopoiesis and provides an ectopic HSC niche that is capable of expanding, on demand,
following severe blood loss or injury to, or disease of, the BM. In the absence of any challenge to the BM, HSCs in the spleen are rare, mainly found within the red pulp, but are reported to expand and undertake compensatory extramedullary hematopoiesis when necessary. There are numerous parallels between the HSC niches in the BM and spleen, and our understanding of these processes has been driven by the study of the BM. Interestingly, unlike the marrow, the spleen reticular perisinusoidal stromal cells providing niche factors are Tcf21⁺, likely reflective of their origin¹¹⁴.

Other studies have shown that HSCs are found in close association with quiescent nestinGFP⁺ MPs⁴⁸. In contrast, Greenbaum et al. showed that nestin–ve MPs within the BM were important for HSC maintenance¹¹². These roles assigned to seemingly contradictory populations may in part reflect our evolving understanding of the complexities of the HSC niche and/or the limitations of the various genetic tools used to define these entities. Furthermore, RNA-seq of the various stromal fractions within the BM do not support a prominent role for Nes-expressing cells in the HSC niche.

These comments notwithstanding, collectively these studies have provided unparalleled insights into the composition of the HSC niche and importance of MPs. The role of MPs in other stem/progenitor niches is beginning to emerge, and these studies will benefit greatly from the paradigms developed to interrogate MP function in the HSC niche. As has been observed in regeneration and renewal, niche capacity is also reduced with aging¹¹⁵. Thus, a better understanding of the MP function in these various contexts will likely aid development of therapeutics to modify MP fate and function for the treatment of a myriad of diseases.
1.6 Quiescence – reversible growth arrest

A general property of most if not all stem and progenitor cells is the ability to undergo variable lengths of growth arrest, which is commonly referred to as quiescence. Following a growth arrest signal, there are four well-defined possible outcomes\textsuperscript{116,117}. The first, differentiation, is a commitment to a functional cell type and expression of a genetic program reflective of specific phenotypes. There is debate as to the reversibility of this state as in some instances there is clear evidence that terminally differentiated functional cells such as hepatocytes can re-enter the cell cycle following the appropriate cues. In other cases, terminal differentiation is an irreversible transition to a post-mitotic state where under normal circumstances there is negligible ability to re-enter the cell cycle (i.e. adult cardiomyocytes, neurons, etc.). The second possible fate involves programmed cell death or apoptosis. A third fate, deemed senescence, serves important purposes but leads to aberrant cell function with typically cell and non-cell autonomous negative impacts\textsuperscript{118}. These three outcomes are defined in part by the growth arrested cells inability to re-enter the cell cycle. The fourth fate involves acquisition of a quiescent phenotype.

Cellular quiescence is considered to be the entry into, and maintenance of, a reversible state of withdrawal from the cell cycle, for variable periods of time from hours to years. In the adult organism, this ability is thought to be reserved for stem and progenitor cells that require carefully regulated temporal cell cycle control to maintain tissue homeostasis and support of renewal and regeneration while minimizing aberrant cell expansion and the potential consequences, such as oncogenesis. It was reported a decade ago that retention of the post-developmental ability to re-enter the cell cycle must not simply be a passive property but rather
the consequence of the deliberate induction of a genetic program\textsuperscript{116}. Whether or not this state of dormancy is simply the cessation of proliferation or a complete reduction of metabolic activity \textit{in vivo} is yet to be completely resolved and heterogeneity likely exists between cell types with respect to this property\textsuperscript{117}. Numerous quiescent stem and progenitor cells exist in adult tissues and can be activated to contribute to tissue homeostasis, such as the intestinal villus in which a conveyor belt-like system is used to support renewal of intestinal epithelial or the regeneration of skeletal muscle following damage\textsuperscript{84,119,120}. The cell responsible for myofiber regeneration, the skeletal muscle satellite stem cell, is proposed to regulate stem cell numbers and activation through the actions of various paracrine and autocrine signals and intrinsic epigenetic mechanisms\textsuperscript{121-125}. Studies of quiescence in hair follicle stem cells suggest that this state is maintained through the active repression of CDK4 by the transcription factor NFATc1, downstream of BMP signaling\textsuperscript{126}. These observations support a model involving active maintenance of quiescence by extrinsic factors.

It has been suggested that in response to focal damage that “distant” quiescent cells enter a primed G\textsubscript{0} state in which they exhibit faster activation kinetics if challenged. This modified state, termed G\textsubscript{alert}, thereby enables a poised quiescent stem and progenitor cell state that enables a quicker response if so required. This priming process appears to involve modulation of mTORC1 activity\textsuperscript{127}. A recent study of neural stem cells in \textit{Drosophila} demonstrates an arrest in the G2 phase, which is proposed to allow for an even more rapid cell cycle reentry\textsuperscript{119}. Collectively, these studies suggest that the quiescent state can be tailored to the cellular and ultimately tissue/host needs.

Additional features of cellular quiescence include reduced metabolic and transcriptional activities. However, this may be counteracted by the need to act quickly in response to the
appropriate stimuli with a concomitant increase in metabolism and transcription. There is considerable interest in discovering the underlying mechanisms regulating stem and progenitor cell quiescence. This information will support effective therapeutic strategies aimed at controlling cell growth, differentiation and organization as part of the tissue regeneration and engineering space.

1.6.1 *Hypermethylated in Cancer 1 (Hic1)*

A screen to identify regulators of MP quiescence was carried out by Collier *et al.* and this yielded a list of interesting candidate genes. Included in this list was a gene termed *Hypermethylated in Cancer 1* or *Hic1*. Interestingly, *Hic1* was also shown to be a mesenchyme restricted transcript in the developing murine embryo. *HIC1* was initially identified over 15 years ago and as the name implies, was discovered as part of the 17p13.3 chromosomal region found silenced by promoter hypermethylation in many human cancers. Subsequent analyses have reaffirmed the role *HIC1* represents as a tumor suppressor. *HIC1* is transcriptional repressor that contains 5 Kruppel-like zinc finger motifs DNA that coordinate DNA binding to the cognate sequence '5-[CG][CG]GGGCA[CA]CC-3'. The amino-terminus of the protein also contains a BTB/POZ domain (*Broad complex, Tramtrack and Bric à brac/Poxviruses and Zinc finger*) which facilitates homodimerization and transcriptional repression, along with an internal C-terminal binding protein (CtBP) region that also participates in gene repression. *Hic1* deletion in mice is embryonic lethal and prenatal animals exhibit widespread and diverse abnormalities including, growth retardation and severe defects in craniofacial, trunk (ventral body wall) and limb development. Mice heterozygous for *Hic1* develop normally, though
are reported to develop a susceptibility to age-related tumours with a latency of 90 weeks, and in these studies, isolated tumours exhibited increased hypermethylation of the wild-type Hic1 allele\textsuperscript{133}. In accordance with this, ectopic expression of HIC1 has been shown to induce growth arrest in established cancer cell lines\textsuperscript{129}. Consistent with a potential role in quiescence, other studies have linked HIC1 to the G1/S transition of the cell cycle including transcriptional regulation of cell-cycle genes encoding CCND1 and CDKN1A\textsuperscript{134-136}.

1.7 Key technological advances

Our understanding of stem and progenitor cell biology has accelerated over the last 1-2 decades and this has been greatly aided by sophisticated methods to query multiple aspects of the single cell, and the ability to study and manipulate these populations in their native environment(s). Numerous technologies have impacted these advances, and some of the salient strategies used in this thesis are highlighted here.

1.7.1 Genetic labelling, lineage tracing and cell purification

Lineage tracing methodologies have emerged over the last 15 years and have provided unparalleled insights into the fate and function of numerous stem and progenitor cells. One notable study published in 2007 by the Clevers group, identified the intestinal stem cell (ISC)-enriched gene Lgr5 and the subsequent development of a Lgr5-based CreERT2 knock-in line into the Lgr5 locus\textsuperscript{95}. This study was the first to provide a definitive marker of the ISC and the resulting genetic tools have been widely used to explore ISC biology in health and disease. As
discussed in previous sections, these approaches typically rely on two components, a robust solitary marker of the stem/progenitor population of interest coupled with a corresponding CreERT2 knock-in allele and for lineage tracing, a Cre-dependent reporter. This system enables one to carry out pulse-chase experiments, such that reporter gene expression can be induced in the population of interest and the fate of marked cells can be followed over time. In the case of the ISC, with time the reporter gene emerged in the transit amplifying progenitor compartment within the crypt, with label subsequently appearing in the differentiated progeny comprising the intestinal villus. The reporter was retained in the ISCs over time indicating that this population exhibited significant self-renewal capability, consistent with a bona fide stem cell. Furthermore, these genetic tools enable the characterization of these populations \textit{in vivo} using various microscopy-based methods including confocal microscopy\textsuperscript{60}. In particular, numerous groups have stressed the importance of confocal microscopy in combination with lineage reporter and immunodetection to confidently identify reporter/protein co-localization for more accurate fate determinations. These strategies have been augmented by both flow cytometry and fluorescence activated cell sorting (FACS) for the phenotypical characterization of populations of interest and their subsequent enrichment for further analysis. These enriched cell populations can be analyzed at the population or single cell level to interrogate the status of their genomic, epigenomic or transcriptomic profiles without confounding information or dilution of meaningful signal from other cell types. Lineage tracing, coupled with these additional methodologies, has become an indispensable tool in the kit of investigators working within the stem/progenitor cell fields.
1.7.2 Querying the transcriptome with next generation sequencing

Over the last few years transcriptional profiling using microarray-based approaches has been supplanted by next generation, also known as high throughput, sequencing (NGS) methods. With NGS becoming more accessible in recent years it has become routine to analyze samples at the whole transcriptome level. This allows for a comprehensive overview of gene expression as opposed to target based analysis such as qPCR. More importantly, these new methods provide a digital read-out of transcript abundance, thereby enabling comparison of transcript abundance across genes. Furthermore, NGS provides greater sensitivity with the ability to now interrogate the transcriptomes of single cells.

Recent advances in droplet based single cell sequencing technology have allowed for extremely high throughput profiling of thousands of single cells. The ability to label transcripts with unique cellular barcodes allows libraries containing transcripts from thousands of cells to be sequenced together, and subsequently demultiplexed and mapped back to their cell of origin. When isolating nucleic acids from a single cell it has generally been a requirement to amplify using PCR to obtain sufficient template for downstream library preparation. The advent of unique molecular barcodes provided a solution to the problem of PCR duplicates when amplifying cDNA from low input sources to generate enough material to successfully prepare a library. Library depth and overall transcript identification remains limited by the inability to efficiently capture all relevant transcripts in a cell and convert them to sequenceable templates. However, upcoming versions of nucleic acid isolation and library preparation chemistry have been designed to increase sensitivity. This technology combined with the computational analysis described below has provided a new perspective on cell heterogeneity.
The data output from the transcriptomes of thousands of cells is extensive to say the least and has recently created an increased demand for improved bioinformatic data management strategies\textsuperscript{143}. To meet this demand, dimension reduction and clustering algorithms have been created to guide analyses and aid generation of meaningful interpretations of these large and complex datasets\textsuperscript{144-147}. These tools have made it possible to profile thousands of cellular replicates and classify subpopulations. Furthermore, it is now possible to efficiently interrogate rare cell subpopulations or groups of cells that are distributed along the continuum of a biological or pathological process. This strategy, referred to as pseudotime, is particularly valuable to observe stem and progenitor cells as they make lineage decisions and undertake differentiation trajectories\textsuperscript{148-152}.

1.8 Objectives of this work

Currently the stem/progenitor source(s) of adult mesenchymal derivatives is being explored across different tissues. The amount of turnover in connective tissue lineages and the extent of contribution to pools of transient pro-regenerative, and persistent pathological, activated fibrogenic cells remain unclear. These analyses have been hampered by the absence of a robust universal marker for identifying and fate mapping of MPs within and across tissues. Herein I outline a series of studies using newly developed genetic tools to investigate MP biology. This is organized into 4 main concepts:

1- Initial experimental work in skeletal muscle focused on demonstrating the effectiveness of \textit{Hic1} as a solitary genetic marker of quiescent MPs. Novel \textit{Hic1}-based genetic reagents were generated, validated and used to explore MP behavior in
skeletal muscle homeostasis and regeneration. These tools in combination with transcriptional profiling have been used to generate a systemic adult MP atlas.

2- It has been observed in several tissues that when MPs exit quiescence, they become activated and enter the cell cycle yet very little is known about their fate or function following activation. Herein I examined the function of MPs in skeletal muscle regeneration and their roles in providing trophic support, modifying the regenerative microenvironment and as a source of both transient and enduring cell constituents in this process.

3- Importantly, mechanisms that regulate adult MP quiescence have yet to be explored. I addressed the function of Hic1 and its encoded product in regulating MP quiescence.

4- Stromal populations contribute trophic support in various tissues and have been shown to be vital residents of stem and progenitor cell niches. Deletion of Hic1 was found to lead to systemic MP hyperplasia, and the impact of this was tested on a well-characterized stem cell niche with known MP participation. For these purposes, the effect on the HSC niche was evaluated with regards to the non-cell autonomous activity of MPs.

In summary, these studies validated novel genetic tools to enable exploration of MP biology. These analyses revealed many new and exciting features of MPs including that this population shares a common mechanism regulating quiescence across tissues, exhibits inter- and intra-tissue heterogeneity likely reflective of tissue-specific functions and are central coordinators of tissue renewal and regeneration.
Chapter 2: Methods

2.1.1 Mice

Mice harboring modified \( Hic1 \) alleles, including \( Hic1^{\text{floxed}} \) (\( Hic1^f \)) and \( Hic1^{\text{CreERT2}} \), were generated by GenOway Inc. under contract using standard gene targeting methodology. For the \( Hic1^f \) mice a targeting vector was constructed that contained multiple components, including LoxP sites flanking the major coding exon 2 of \( Hic1 \), a PGK-neo cassette flanked by FRT sites distal to exon 2 and upstream of the 3’ LoxP site. An IRES-NLS-LacZ cassette distal to the 3’ LoxP site (downstream of exon 2) was added to enable efficient identification of \( Hic1 \) deleted cells along with \( Hic1 \)-expressing cells. This targeting construct was introduced into 129Sv/Pas ES cells and properly targeted ES cells were identified through a combination of PCR and southern analysis. Four targeted ES cell clones were used for subsequent injection into ~60 C57BL/6 blastocysts which led to 10 males with chimerism ranging from 80-100%. These male chimeras were bred with female Flippase-expressing mice for deletion of the PGK-neo cassette, and PGK-neo deleted mice were subsequently back-crossed to C57BL/6 mice for >10 generations. The \( Hic1^{\text{CreERT2}} \) mice were generated using similar methodology and involved introduction of an IRES-CreERT2 cassette into the 3’ untranslated region of the \( Hic1 \) gene. The CreERT2 gene was removed from plasmid pCAG-CreERT2 (Addgene #14797) and introduced into a targeting vector based on C57Bl/6 DNA sequences along with a FRT flanked PGKneo cassette. This vector was electroporated into C57Bl/6 ES cells and successfully targeted clones were injected into C57Bl/6 blastocysts. Male chimeras were subsequently bred to Flippase-expressing mice to delete the PGK-neo cassette. These mice were maintained on a C57BL/6
background. Other mouse lines used include: B6.Cg-Gt(ROSA)26Sortm14(CAG-tdTomato)Hze/J (Jax stock number 007914; herein referred to as Rosa^LSL-tdTomato), B6.Cg-Tg(UBC-Cre/ERT2)1Ejb/2J (Jax stock number 008085; herein referred to as UBC-CreERT2) and C57BL/6J (Jax stock number 000664); B6.129S4-Pdgfra^tm11(EGFP)Sor/J (Jax stock number 007669). All mice were maintained in (or as noted backcrossed to) a C57BL/6 background.

For lineage tracing experiments, Hic1^CreERT2 mice were interbred with Rosa^LSL-tdTomato mice (Ai14 line, JAX stock 007914) to generate Hic1^CreERT2/CreERT2 or +; Rosa26^LSL-tdTomato/+ mice. To induce CRE-ERT2 nuclear translocation, 8-9-week old mice were injected intraperitoneally with 100 mg/kg of Tamoxifen in 100 µL sunflower oil daily for 5 consecutive days. Skeletal muscle injury studies were initiated in these animals following a > 14 d washout period. For single cell RNA-seq experiments, mice were allowed a ten-day washout period to ensure clearance of tamoxifen before cell isolation.

For experiments involving conditional deletion of Hic1, Hic1^ff mice were interbred with UBC-CreERT2 mice to generate Wt or UBC-CreERT2; Hic1^ff mice. TAM was administered to mice > 8 weeks of age as described above.

Injury of the tibialis anterior muscles was induced by intramuscular injection of 0.2 µg of notexin (NTX) snake venom extract (Laxotan) in 20 µL PBS (10 µg/mL)^18. For acute pancreas injury, mice were IP injected with the cholecystokinin analog caerulein. Mice received 8 hourly injections of caerulein (at 50 µg/ml) within a single day and this was termed 0 d.

Animals were maintained and experimental protocols were conducted in accordance with approved and ethical treatment standards of the University of British Columbia.
2.1.2 **Generation of anti-HIC1 antibody**

A rabbit polyclonal antibody to HIC1 was generated based on the carboxyl 25 amino acids of HIC1 using ProSci Inc. (Poway, CA). This region exhibits 96% amino acid identity between mouse and human. A peptide (C25; AEVLSQGAHLAAGPDSRTIDRF) was conjugated to keyhole limpet hemocyanin and used to immunize 2 rabbits in complete Freund’s adjuvant. Collected serum was subsequently affinity purified on a C25 peptide-containing affinity column.

2.1.3 **Fluorescence-activated cell sorting methodology**

MPs were released from various muscles of the hind limb using a previously described protocol with modifications\textsuperscript{18}. Freshly dissected TA muscles were sequentially digested in a volume of 250 µL per TA containing 500 U/mL Collagenase II (Sigma C-6685-5G) for 30 minutes followed by a cocktail of 1.5 U/mL Collagenase D (Roche 11 088 882 001) and 2.4 U/mL Dispase II (Roche 04 942 078 001) for 60 minutes. Digested material was subsequently triturated by pipetting and passed through a 40 µm cell strainer. The resultant cell suspension was washed with FACS buffer (PBS containing 2 mM EDTA and 2% FBS) and centrifuged at 500xg for 5 minutes. To enrich for MPs from this whole skeletal muscle mononuclear suspension, cells were selected for the expression of various cell surface markers including CD45, CD31, Ter119 or Sca-1. The lin\textsuperscript{−} fraction was considered to be negative for CD45\textsuperscript{−}, Ter119\textsuperscript{−} and CD31\textsuperscript{−}. This was accomplished by incubation with a cocktail containing anti-Ter119-488 (Ablab 48-0031-01, 1:200), anti-CD45-PerCP (BD Pharmingen 557235, 1:400) and
anti-CD31-FITC (eBioscience 11-0311-85, 1:400). The Lin⁻ population was further sub-fractionated based on Ly6a/Sca-1 expression by the addition of anti-Ly-6A/E (Sca-1)-PE-Cy7 (eBioscience 25-5981-82, 1:3000) to the above antibody cocktail. This mixture containing cells and antibodies was subsequently incubated on ice in the dark for 30 minutes. Afterwards, 15 mLs FACS buffer was added to dilute the antibodies prior to centrifugation at 500xg for 5 minutes. The pellet was resuspended in FACS buffer containing Hoechst 33342 (Sigma B2261) to a final concentration of 4 µM. Stained cells were sorted using a BD Influx, and Hoescht and forward/side scatter parameters were used to identify viable single cells for all FACS enrichments. Sorted cells were collected into sort media (DMEM, 20% FBS, L-glutamine (2 mM), penicillin (100 units/ml) and streptomycin (100 µg/ml)) in cooled siliconized microcentrifuge tubes (Fisher Scientific; 02-681-320). For isolation of RNA, cells were microfuged at 500xg for 5 minutes and the isolated cell pellet was lysed in RNAzol (Sigma R4533).

To enrich for tdTomato-labelled Hic1⁺ MPs a similar FACS strategy was employed with the following antibody cocktail, anti-Ter119-647 (Ablab 67-0031-01, 1:200), anti-CD31-APC (BD Biosciences 551262, 1:400), anti-CD45-647 (Ablab 67-0047-01, 1:400). These antibodies were included to aid enrichment of tdTomato⁺ cells and reduce confounding cell contaminants. Furthermore, to attenuate the signal associated with tdTomato reporter expression in transgenic MPs and improve sorting selectivity, the PECy7 configuration was used to gate the tdTomato⁺ population.

To enrich for WT and Hic1 KO MPs the above-described approach was used with some modifications. The staining cocktail was altered to further enhance MP enrichment and included: anti-Ter119-647 (Ablab 67-0031-01, 1:200), anti-CD31-APC (BD Biosciences

To prepare single mononuclear suspensions from various other organs the following strategies were used. Skeletal muscles and hearts were processed as per skeletal muscle described above however for the heart preparation each ventricle was perfused with at least 5 mL PBS before dissection.

To dissociate lung tissue, blood was cleared by cardiac perfusion of each ventricle with at least 5 mL PBS before dissection. All lobes were introduced to enzyme cocktail 2 (1.5 U/mL Collagenase D [Roche 11 088 882 001] and 2.4 U/mL Dispase II [Roche 04 942 078 001]), minced and incubated at 37ºC on a bottle roller apparatus for one hour. The preparation was mixed by gentle vortex every fifteen minutes to counter the buoyancy of air-filled tissue fragments.

Similar to the lung preparation, to obtain the adipose stromal vascular fraction (SVF), epididymal white and interscapular brown adipose tissue deposits were dissected, minced and incubated in enzyme cocktail 2, rolling as above for one hour. Adipose preparations, however, were mixed every ten minutes to reduce phase separation of lipid dense tissue fragments and aqueous enzyme solution.

To dissociate the kidney, mice were transcardially perfused with cold PBS to remove blood. Kidneys were then dissected, and the capsule and adrenal gland removed before mincing and rolling incubation in enzyme cocktail 2 at 37ºC for 1 hour.
Since the pancreas is susceptible to rapid degradation, the whole organ was quickly yet very gently dissected and minced before 15-minute incubation in enzyme cocktail 2 rotating in an oven set to 37ºC.

To obtain mononuclear cell suspensions from brain and cerebellum, mice were transcardially perfused with PBS, whole brains were dissected, forebrain regions/cortex and cerebellum placed into separate 35mm dishes and bathed in enzyme cocktail 2 pre-warmed to 37ºC. With the dishes on a 37ºC heating block a 1 mL syringe with an 18G needle was used to inject enzyme cocktail 2 into multiple regions of the brain samples. After 5 minutes the samples were drawn up through the syringe and ejected back into the dish. This was repeated through a series of smaller needles until the tissue is relatively homogeneous and the material passes through a 23G or 25G syringe with ease. The resultant solution was then left to digest on the 37ºC block for 20 minutes before centrifugation through a 40% percoll gradient to remove myelin-associated material.

To dissociate liver, a procedure was developed for in situ digestions modeled after Maschmeyer et al. however using the retrograde perfusion technique as reported in Mederacke et al. and instead of subsequent density gradient centrifugation, single cell suspensions were FACS purified as described below\textsuperscript{153,154}.

To obtain a single cell suspension from BM that included the SVF, hindlimb longbones were dissected and placed into cold PBS on ice. Epiphyses were removed distal to the growth plate and marrow flushed with 4 mL per mouse of digest solution (500 µL stock collagenase I (20 mg per mL in HEPES) into a total of 4 mL sort media [20% FBS, L-glutamine (2 mM), penicillin (100 units/ml) and streptomycin (100 µg/ml)and 20 ng/mL FGF2 in DMEM]) into a 35 mm dish. Once all four bones were flushed and marrow collected the solution was transferred
into a 5 mL FACS tube and incubated at 37°C with gentle rocking for 30 minutes. Following the incubation, the cells were centrifuged at 500xg for 5 minutes, resuspended in FACS buffer to wash and resuspended in staining cocktail described below.

Once enzymatic digestion was complete, all respective cell suspensions from above were filtered through a 40 μm cell strainer, washed with FACS buffer (PBS containing 2mM EDTA and 2% FBS), centrifuged at 500xg for five minutes, resuspended in antibody cocktail described above for isolation of tdTomato cells and incubated for thirty minutes on ice. Finally, cells were washed and centrifuged as above then resuspended in FACS buffer containing 4 μM Hoechst 33342 (Sigma B2261) and 1.5 μM Propidium Iodide (ThermoFisher P1304MP). To enrich for tdTomato-labelled Hic1+ MPs from BM, a similar FACS strategy was employed with the following additions to the antibody cocktail, anti-CD3-649 (KT3 Ablab AB11FB02MW258, 1:500), anti-CD79a-APC (BD Biolegend 133106, 1:500), anti-CD45RB/B220-647 (RA-6B2 Ablab 67-0030-05, 1:500).

2.1.4 Cell cycle analysis

To enable detection of proliferating cells, standard methodology involving analysis of the incorporation of the nucleoside triphosphate analogue 5-ethyl-2’-deoxyuridine (EdU) was used. For these purposes, animals were injected intraperitoneally with 0.5 mg of EdU (ThermoFisher E10415) in 250 µL PBS daily during the 3-d pulse period between injury and collection at 4DPI. EdU+ cells were subsequently detected by flow cytometry and in histological samples using the Click-iT Plus EdU Pacific Blue Flow Cytometry Assay kit (ThermoFisher...
C10636) and the Click-iT Plus EdU Alexa Fluor 647 Imaging kit (ThermoFisher C10640), respectively, both as per manufacturer instructions.

2.1.5 Whole-mount LacZ staining

LacZ activity in some tissue samples was detected using established whole-mount X-gal staining \(^1\). Briefly, to enhance penetration of the fixative in whole tissues, dissected tissues were cut into \(\sim 4 \text{ mm}^3\) pieces and fixed for 18-24 hours at 4°C in LacZ fixative (100 mM MgCl\(_2\), 0.2 % glutaraldehyde, 5 mM ethylene glycol tetra-acetic acid in PBS). Samples were subsequently washed 3 times for 30 minutes each in PBS then permeabilized in a solution containing 2 mM MgCl\(_2\), 0.01% deoxycholate, 0.02% NP40, 5 mM potassium ferricyanide, and 5 mM potassium ferrocyanide overnight at 37°C. This was followed by an overnight incubation in LacZ staining solution (2 mM MgCl\(_2\), 0.01 % deoxycholate, 0.02% NP40, 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide and 1mg/mL 5-bromo-4-chloro-3-indolyl-\(\beta\)-D-galactopyranoside) at 37°C. Samples were post fixed in 4% PFA overnight at 4°C and visualized and/or paraffin embedded for histological analysis. Six \(\mu\)m sections were prepared and counterstained with nuclear fast red for analysis.

2.1.6 Immunofluorescence and/or in situ X-gal staining

For in situ X-gal staining on histological sections, tissues were prepared in the following manner. Mice were terminally anesthetized by intraperitoneal injection of Avertin (400 mg/kg)
and fixed by intracardial perfusion of cold PBS followed by cold LacZ fixative as described above, before dissection and immersion of tissue samples in LacZ fixative for 3 hours on ice.

To enable detection of native tdTomato expression in processed tissue samples, mice were terminally anesthetized by intraperitoneal injection of Avertin and fixed by intracardial perfusion of 10 mM EDTA in PBS followed by 4% paraformaldehyde. Tissues were immersed in 4% paraformaldehyde for 48 hours at 4°C.

For cryosectioning, with the exception of bones which first underwent a 72-hour decalcification in 14% EDTA in PBS with the solution replaced every 24 hours, the above collected and fixed samples were washed with PBS then incubated through a cryoprotective series of sucrose solutions of increasing concentration from 10-50% (for ≥3 hours each before embedding into OCT compound (Tissue Tek 4583). Tissues were immersed into OCT in disposable plastic cryomolds (Polysciences 18646A) and frozen in an isopentane bath cooled by liquid nitrogen. Cryosections were cut (Leica CM3050S) at a thickness of 5-30 µm and mounted onto Superfrost Plus slides (VWR 48311-703). This material was used for either LacZ and/or IF staining.

For detection of LacZ on sections, in situ LacZ staining with X-gal was carried out. Slides were thawed at RT, washed 3x 10 minutes in PBS and then incubated overnight at 37°C in a humidified chamber with the aforementioned X-gal staining solution. Slides were counterstained with nuclear fast red and mounted with Aqua Polymount. For additional IF staining of LacZ stained samples, freshly stained slides were subsequently washed with PBS for 3x 5 minutes and IF staining was carried out as described below.

For IF staining, slides were thawed at room temperature, washed 3x 10 minutes in PBS and incubated for 1 hour in PBS containing 10 mg/mL sodium borohydride (Sigma 213462) to
quench autofluorescence. Following this treatment, slides were washed with PBS and incubated in blocking solution containing 2.5% BSA (Sigma A7030) and 2.5% Goat serum (Gemini 100-190) for 90 minutes at room temperature prior to incubation in primary antibody (listed below) overnight at 4°C. AlexaFluor conjugated secondary antibodies were typically diluted 1:500 and applied to the slides for 45 minutes. After each antibody incubation, 3x 5-minute PBS washes were performed and sections were counterstained with DAPI (600 nM) and mounted with Aqua Polymount (Polysciences 18606). The following modifications were used for anti-Tenomodulin staining: 0.1% Tween-20 (OmniPur 9480) in PBS (PBST) was used for all washing steps. PBST with 5% donkey serum (Millipore S30) and 2.5% BSA was used as the blocking solution and the antibodies were prepared in the same for incubation.

For detection of HIC1 with anti-HIC1 (C25; 1:2000 dilution) a modified tyramide signal amplification staining strategy was used. Cryosectioned material was stained using TSA Kit #16 (ThermoFisher T20926) as per manufacturer’s instructions from step 2.4 to step 4.4 with the following modifications. Either anti-Ly6a/Sca-1 (Abcam51317, 1:200) or anti-CD31 (BD 550274, 1:50) was included in the antibody cocktail in step 3.3 and Alexa Fluor 555 goat anti-rat IgG (ThermoFisher A21434, 1:500) in step 3.6. Furthermore step 4.2 was performed overnight at 4°C. To highlight the basement membrane using an antibody of the same species without cross reactivity, separate consecutive antibody incubations were performed instead. Between steps 4.3 and 4.4 the sections were incubated with rabbit anti-laminin (Ab11575) for 90 minutes, followed by 3x 5 minutes PBS washes. Laminin staining was subsequently detected without amplification by incubation with goat anti-rabbit 488 (ThermoFisher A11034) for 45 minutes, and the slides were counterstained with DAPI (ThermoFisher D3571) and mounted with Aqua Polymount.
2.1.7 Bone marrow harvest and flow cytometry

Mice were sacrificed as described above and femurs and tibias were collected for BM harvest by flushing with ice-cold FACs buffer (PBS containing 2 mM EDTA and 2% FBS). Red blood cells were lysed with ACK lysing buffer (Thermo Fisher Scientific), and preparations were washed and filtered through a 70 µm mesh filter. Filtered cell suspensions were incubated with primary antibodies against indicated cell surface markers for 30 minutes at 4 °C (in FACS buffer, at \( \sim 3 \times 10^7 \) cells/ml). The antibodies used in flow cytometry and the dilutions are listed in Table 2.1.

SLAM marker positive populations were enumerated using was an LSRII (Becton Dickenson) flow cytometer and data were collected using FacsDIVA software. Cell sorting was performed on a FACS Influx (Becton Dickenson) or FACS Aria (Becton Dickenson) sorter. Sorting gates were strictly defined based on fluorescence minus one staining controls. Flow cytometry data analysis was performed using FlowJo 10.4.1 (Treestar) software.

2.1.8 Hematopoietic progenitor growth assay in methylcellulose

Whole BM was harvested as described above and plated at \( 2 \times 10^4 \) cells per 2 ml culture. Methylcellulose media (M3434; STEMCELL Technologies) was used to support colony formation and CFUs were scored at day 8 after plating.
2.1.9 Competitive bone marrow transplantation experiments

Adult seven-week wildtype C57Bl6 Ly5.1 recipients were irradiated by two exposures to 5.25 Gy from Cs137, 6 hours apart in a Gammacell 220 instrument. BM cells from age and sex-matched mice were mixed 1:1 with Ly5.2 GFP+ BM control cells. For transplantation, 4 x 10^6 cells were delivered via intravenous tail vein injection into CD45.1 hosts 6 hours after lethal irradiation. After 16 weeks, we determined the percentage contribution of GFP negative (Hic1^f/f or UBC-Cre;Hic1^f/f) versus positive (Ly5.2 GFP) cells in donor-derived (CD45.2+) granulocytes (Gr1^+), using the indicated FACS antibodies on a BD LSRII flow cytometer.

2.1.10 Secondary bone marrow transplantation experiments

Whole BM was extracted from primary recipients six months after transplant and cryopreserved. Recipient mice were lethally irradiated as described above, 24 hours prior to secondary transplant. Cryopreserved BM from the primary recipients was thawed and reconstituted such that 3 x 10^6 cells were transplanted via intravenous injection into Ly5.1 secondary recipient mice. BM was isolated four months after transplant for analysis as described for the primary transplant.

2.1.11 Limiting dilution transplantation assay

Whole BM cell suspensions from Ly5.2 Hic1^f/f and UBC-Cre;Hic1^f/f donors were prepared in different dilutions and each was transplanted into lethally irradiated Ly5.1 mice.
(together with 500K Sca-1− Ly5.1 helper BM cells per transplantation). Blood and BM was collected from each recipient six months after the transplant and stained with conjugated primary antibodies to the indicated surface markers and analyzed using a BD LSRII flow cytometer to detect the appearance of Gr1+ Ly5.2 donor cells. A threshold of 0.5% test cells was set to define a positive result.

### 2.1.12 In vitro BM-MP culturing experiments

Cells were sorted as described above from BM suspensions using indicated gating into high (4.5g/L) glucose Dulbecco’s modified eagle medium (DMEM, Thermo Fisher), supplemented with 10% (v/v) fetal bovine serum (FBS), 10 ng/ml bFGF (Peprotech), 1% (v/v) Penicillin Streptomycin (ThermoFisher) and 2 mM L-glutamine (Thermo Fisher). Media was changed every 7 days.

For clonal expansion experiments, FACS sorted single tdTomato+/CD31−/CD45− cells were seeded in 96 well plates on a feeder layer of gamma irradiated mouse-derived C3H10T1/2 cells. The cells were cultured in high (4.5 g/L) glucose DMEM (Thermo Fisher), supplemented with 20% v/v fetal bovine serum (FBS), 10 ng/ml bFGF (Peprotech), 2 mM L-glutamine (Thermo Fisher) and 1% (v/v) Penicillin Streptomycin (ThermoFisher) for 14 days. After 14 days single clones and extent of colony expansion was quantified from images collected by epi-fluorescence imaging using a Zoe Fluorescent Cell Imaging system (Bio-Rad).
2.1.13 Image acquisition and quantification

Microscopy images were acquired using two different platforms including a widefield epi-fluorescence Olympus BX63 compound microscope and a Nikon Eclipse Ti inverted microscope equipped with a C2Si confocal system.

For enumeration and analysis of the distribution of tdTomato<sup>+</sup> cells in time course experiments, a total of 15 images were analyzed per time point (5 images/mouse; 3 mice/time point). tdTomato<sup>+</sup> cell bodies defined by DAPI nuclear stain were enumerated and apposition to vessel assigned per cell. The distribution of a low percentage of cells (1-2%) was considered ambiguous and they were not included in the enumeration. Enumeration and distribution of X-gal positive nuclei in samples from *Hic1<sup>nLacZ/</sup>*</sup> mice was carried out in a similar manner.

The enumeration of nLacZ-expressing cells in various tissues was performed using 3 different methods that were consequence of density of DAPI<sup>+</sup> nuclei. The specific method used for each sample is described in the figure legends. For the majority of the tissue samples, the number of nLacZ<sup>+</sup> nuclei was normalized to the total number of DAPI<sup>+</sup> nuclei within the counted region. For these purposes, a minimum of 3 regions were randomly selected from images of each sample (N=4-5) for analysis. BM counts were extrapolated from analysis of an area of 2,400 µm<sup>2</sup> in each image. For the kidney, spleen and skin samples, all counts were normalized to the analyzed area on the slides and not DAPI staining.

To evaluate CXCL5 expression in “activated” MPs, the percentage of doubly positive tdTomato<sup>+</sup> CXCL5<sup>+</sup> cells in injured TA muscles was counted. Cells counts were carried out on two regions of 1 x 1 mm cropped from stitched images of one whole skeletal muscle section for
all samples (for R samples damaged areas were identified and selected by increased mononuclear infiltrate)

2.1.14 Gene expression profiling – popRNA-seq

Total RNA was isolated using RNazol (Sigma R4533) as per manufacturer’s instructions with the following modifications: StepV.2.-linear polyacrylamide (GenElute, Sigma 56575) was added as a carrier to the RNA solution prior to the addition of cold isopropanol. Total RNA was precipitated overnight at -20°C. Precipitated RNA was centrifuged at 21,000xg for 30 minutes at 4°C. SUPERase-IN RNase inhibitor (ThermoFisher AM2696) was added at 1:20 to the resuspended RNA solution. Sample integrity was tested on an Agilent Bioanalyzer 2100 RNA 6000 Nano chip (5067-1511). RNA samples with an RNA Integrity Number > 8 were used to prepare libraries following the standard protocol for the TruSeq Stranded mRNA library kit (Illumina) on the Illumina Neoprep automated nanofluidic library prep instrument. Paired end sequencing was performed on the Illumina NextSeq 500 using the High Output Kit (Illumina).

2.1.15 RNA-seq bioinformatic analyses

Illumina sequencing output generated bcl file was de-multiplexed by bcl2fastq2. De-multiplexed read sequences were then aligned to the Mouse Genome mm10 reference sequence using TopHat splice junction mapper with Bowtie 2

(http://ccb.jhu.edu/software/tophat/index.shtml) or STAR

expression were estimated using Cufflinks (http://cole-trapnell-lab.github.io/cufflinks/). For some datasets, enriched pathways or processes were identified using Ingenuity Pathway Analysis (IPA). For these purposes, IPA was performed on gene lists that had been filtered for a p-value <0.05 using Cufflinks.

2.1.16 Single cell RNA-seq

Single cell suspensions with the exception of marrow were generated as described for pop-RNA-seq with the following downstream modifications. Target cells were enriched by FACS sorting into 0.22 µm vacuum filtered SC collection media (DMEM containing 5% FBS) with 1.5 µM Propidium Iodide. Viable target cells were subsequently further purified, and debris reduced by sorting a second time and collected into 10 µL SC collection media. Cells were counted and quality control was determined by hemocytometer. If > 98% of visible objects were verified to be single tdTomato⁺ cells the suspension was input into a Chromium Controller (10x Genomics), captured and libraryed with the Chromium single cell 3’ reagent kit v2 (10x Genomics). cDNA libraries were sequenced on a Nextseq 500 (Illumina) to a minimum depth of 50,000 reads per cell. A transgenic reference genome was generated by the concatenation of the sequences for tdTomato to the mm10 reference genome and subsequent use of the cellranger mkref pipeline. Demultiplexing, alignment to the modified mm10 reference genome, principal component analysis, clustering, non-linear reduction (tSNE) and differential expression was performed using the cellranger count pipeline (10x Genomics). Aggregated library datasets were created using the cellranger aggr pipeline. Graphical output was generated using the cellrangerRkit R package (10x Genomics) and the output from the geneBCmatrix
(cellrangerRkit) was formatted using an in house developed R function to generate violin plots and plotted using the grammar of graphics 2 R package. A parallel set of principal component and downstream analysis was performed using the Seurat 2.3.4 R package (Satija lab, NYU).

Cell ranger count or aggr output expression matrices were used for further analysis by applying the backSPIN algorithm for python setting d=5 for 32 total possible clusters and f=2000 to return the 2,000 most significant genes used to generate the clusters.

The same output matrices were input to the Seurat single cell analysis pipeline for R (Satija lab, NYU) where principal component analysis, shared nearest neighbor clustering, non-linear dimensional reduction (tSNE and UMAP) and differential gene expression analysis was performed using standard parameter settings.

Graphical output from the backSPIN analysis in figure 3.26b, c was generated using the Rtsne, Pheatmap and grammar of graphics 2 R packages. The set interaction diagram was generated using the UpSet application provided by the visR.0.9.33 graphical user interface for R. All other graphical output from the single cell analysis was generated using the Seurat visualization tools built into the R package.

2.1.17 Statistical analyses

All data are represented as mean ± standard deviation and the sample number is indicated in the figure legends. For pairwise comparisons, unpaired t-tests were used to calculate P values. For comparison of > 2 means, one-way ANOVA with Boneferroni’s post ad-hoc tests was employed. Both statistical methods were carried using Prism (Graphpad Software). For figures 3.30-3.32, statistical analysis was performed using unpaired Student’s t-test for analysis between
2 groups followed by Tukey’s post hoc test. Sample size and/or replicate number for each experiment are indicated in the figure legends. Results with p values of < 0.05 were considered statistically significant. Stem cell frequencies (or stem cell initiating frequencies and significance) were calculated using ELDA: Extreme Limiting Dilution Analysis online software\textsuperscript{155}.

Table 2.1 Antibodies used in chapter 3 results

<table>
<thead>
<tr>
<th>Flow cytometry</th>
<th>vendor</th>
<th>Catalogue number</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>For Enrichment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anti-Ly-6A/E (Sca-1)-PE-Cy7</td>
<td>eBioscience</td>
<td>25-5981-82</td>
<td>1:3000</td>
</tr>
<tr>
<td>anti-Ter119-488</td>
<td>Ablab</td>
<td>48-0031-01</td>
<td>1:200</td>
</tr>
<tr>
<td>anti-CD31-PerCP</td>
<td>eBioscience</td>
<td>46-0311-82</td>
<td>1:400</td>
</tr>
<tr>
<td>anti-CD45-PerCP</td>
<td>BD Pharmingen</td>
<td>557235</td>
<td>1:400</td>
</tr>
<tr>
<td>anti-CD31-FITC</td>
<td>eBioscience</td>
<td>11-0311-85</td>
<td>1:400</td>
</tr>
<tr>
<td>anti-CD45.2-FITC</td>
<td>eBioscience</td>
<td>11-0454-85</td>
<td>1:400</td>
</tr>
<tr>
<td><strong>For Timecourse</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anti-Ter119-647</td>
<td>Ablab</td>
<td>67-0031-01</td>
<td>1:200</td>
</tr>
<tr>
<td>anti-CD31-APC</td>
<td>BD Biosciences</td>
<td>551262</td>
<td>1:400</td>
</tr>
<tr>
<td>anti-CD117-APC</td>
<td>eBioscience</td>
<td>17-1172-82</td>
<td>1:500</td>
</tr>
<tr>
<td>anti-CD11b-647</td>
<td>Ablab</td>
<td>67-0055-01</td>
<td>1:500</td>
</tr>
<tr>
<td>anti-F4/80-647</td>
<td>Ablab</td>
<td>67-0035-05</td>
<td>1:500</td>
</tr>
<tr>
<td>anti-CD45-647</td>
<td>Ablab</td>
<td>67-0047-01</td>
<td>1:400</td>
</tr>
<tr>
<td>Marrow analysis</td>
<td>Company</td>
<td>Catalog Number</td>
<td>Dilution</td>
</tr>
<tr>
<td>------------------------------------------------------</td>
<td>--------------</td>
<td>----------------------</td>
<td>----------</td>
</tr>
<tr>
<td>anti-Ly6G-Gr1 efluor450</td>
<td>eBioscience</td>
<td>48-5931-82</td>
<td>1:200</td>
</tr>
<tr>
<td>anti-CD45.1-PE</td>
<td>eBioscience</td>
<td>12-0453-83</td>
<td>1:800</td>
</tr>
<tr>
<td>anti-CD45.2-APC</td>
<td>eBioscience</td>
<td>17-0454-82</td>
<td>1:400</td>
</tr>
<tr>
<td>anti-Ly6A/E (Sca-1)-PeCy7</td>
<td>Biolegend</td>
<td>108114</td>
<td>1:200</td>
</tr>
<tr>
<td>anti-CD117 (Ckit) PerCP/Cy5.5</td>
<td>Biolegend</td>
<td>105824</td>
<td>1:200</td>
</tr>
<tr>
<td>anti-CD150 (SLAM)-PE</td>
<td>Biolegend</td>
<td>115903</td>
<td>1:400</td>
</tr>
<tr>
<td>anti-CD48 eFluor450 (PB)</td>
<td>eBioscience</td>
<td>48-0481-80</td>
<td>1:400</td>
</tr>
<tr>
<td>FITC lineage cocktail</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anti-Cd11c</td>
<td>Ablab</td>
<td>N418</td>
<td>1:400</td>
</tr>
<tr>
<td>anti-Ter119</td>
<td>Ablab</td>
<td></td>
<td>1:400</td>
</tr>
<tr>
<td>anti-Cd3e</td>
<td>Ablab</td>
<td>2C11</td>
<td>1:400</td>
</tr>
<tr>
<td>anti-Cd4</td>
<td>Ablab</td>
<td>GK1.5</td>
<td>1:400</td>
</tr>
<tr>
<td>anti-Gr1 (Ly6G)</td>
<td>Ablab</td>
<td>RB6-8C5</td>
<td>1:400</td>
</tr>
<tr>
<td>anti-NK1.1</td>
<td>Ablab</td>
<td>PK136</td>
<td>1:400</td>
</tr>
<tr>
<td>anti-Cd19</td>
<td>Ablab</td>
<td>1D3</td>
<td>1:400</td>
</tr>
<tr>
<td>anti-B220</td>
<td>Ablab</td>
<td>RA-6B2</td>
<td>1:400</td>
</tr>
<tr>
<td>Antibody Name</td>
<td>Vendor</td>
<td>Catalog Number</td>
<td>Dilution</td>
</tr>
<tr>
<td>---------------</td>
<td>--------</td>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>anti-Cd3</td>
<td>Ablab</td>
<td>KT3</td>
<td>1:400</td>
</tr>
<tr>
<td>anti-Cd11b</td>
<td>Ablab</td>
<td>M1/70</td>
<td>1:400</td>
</tr>
<tr>
<td><strong>Histo IF Primary</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>anti-Laminin</td>
<td>Abcam</td>
<td>11575</td>
<td>1:100</td>
</tr>
<tr>
<td>anti-CD31</td>
<td>BD</td>
<td>550274</td>
<td>1:50</td>
</tr>
<tr>
<td>anti-Ly6a (Sea-1)</td>
<td>Abcam</td>
<td>51317</td>
<td>1:200</td>
</tr>
<tr>
<td>anti-CXCL5</td>
<td>Abcam</td>
<td>198505</td>
<td>1:100</td>
</tr>
<tr>
<td>anti-Perilipin</td>
<td>Cell Signaling</td>
<td>9349S</td>
<td>1:200</td>
</tr>
<tr>
<td>anti-Periostin</td>
<td>Abcam</td>
<td>14041</td>
<td>1:100</td>
</tr>
<tr>
<td>anti-Tagln</td>
<td>Abcam</td>
<td>14106</td>
<td>1:100</td>
</tr>
<tr>
<td>anti-αSMA</td>
<td>Abcam</td>
<td>5694</td>
<td>1:200</td>
</tr>
<tr>
<td>anti-CD146</td>
<td>Abcam</td>
<td>75769</td>
<td>1:250</td>
</tr>
<tr>
<td>anti-Collagen 1a1</td>
<td>Abcam</td>
<td>34710</td>
<td>1:100</td>
</tr>
<tr>
<td>anti-Collagen 22a1</td>
<td>Ronen Schweitzer</td>
<td>generously provided</td>
<td>1:150</td>
</tr>
<tr>
<td>anti-Tenomodulin (-Santa Cruz</td>
<td>49324</td>
<td>1:50</td>
<td></td>
</tr>
<tr>
<td>anti-CD146/MCAM</td>
<td>Abcam</td>
<td>75769</td>
<td>1:250</td>
</tr>
<tr>
<td><strong>Histo IF Secondary</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alexa Fluor 488 goat anti-mouse IgG (H+L)</td>
<td>ThermoFisher</td>
<td>A11029</td>
<td>1:500</td>
</tr>
<tr>
<td>Antibody</td>
<td>Manufacturer</td>
<td>Catalog Number</td>
<td>Dilution</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------------</td>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>Alexa Fluor 555 goat anti-rat IgG (H(^\text{+})L)</td>
<td>ThermoFisher</td>
<td>A21434</td>
<td>1:500</td>
</tr>
<tr>
<td>Alexa Fluor 647 goat anti-rat IgG (H(^\text{+})L)</td>
<td>ThermoFisher</td>
<td>A21247</td>
<td>1:500</td>
</tr>
<tr>
<td>Alexa Fluor 488 donkey anti-rat IgG (H(^\text{+})L)</td>
<td>ThermoFisher</td>
<td>A11055</td>
<td>1:500</td>
</tr>
<tr>
<td>Alexa Fluor 488 goat anti-rabbit IgG (H(^\text{+})L)</td>
<td>ThermoFisher</td>
<td>A11034</td>
<td>1:500</td>
</tr>
<tr>
<td>Alexa Fluor 594 goat anti-rabbit IgG (H(^\text{+})L)</td>
<td>ThermoFisher</td>
<td>A11037</td>
<td>1:500</td>
</tr>
<tr>
<td>Alexa Fluor 647 goat anti-rabbit IgG (H(^\text{+})L)</td>
<td>ThermoFisher</td>
<td>A21245</td>
<td>1:500</td>
</tr>
<tr>
<td>Alexa Fluor 405 donkey anti goat IgG</td>
<td>Abcam</td>
<td>175665</td>
<td>1:250</td>
</tr>
</tbody>
</table>
Chapter 3: *Hic1* Marks and Regulates Quiescence of Mesenchymal Progenitors

3.1 Introduction

Mesenchymal progenitors (MPs) are thought to modulate tissue maintenance and regeneration by providing trophic support to tissue specific stem cells\(^{13,51,156-158}\), a notion that has supported their use in cell therapy approaches. A clear example of such role stems from studies in adult skeletal muscle, whose efficient regeneration requires the coordinated action of distinct tissue-resident stem and progenitor cell populations\(^{159}\). Figuring prominently in this scenario are Pax7\(^+\) satellite cells, which are essential for post-natal muscle formation and regeneration following damage\(^{159-161}\). In addition to satellite cells, a population of mesenchymal fibro/adipogenic progenitors (FAPs) was identified within the skeletal muscle interstitium that also play an important role in regeneration\(^{18,19,159}\). Kardon and colleagues used Tcf7l2-based transgenic tools to identify and conditionally ablate a muscle resident MP population, showing that even a small decrease (~45%) in these cells resulted in dysregulated satellite differentiation and impacted skeletal muscle regeneration\(^{159}\). In addition to playing a positive role in regeneration, these progenitors are a main source of the pathogenic extracellular matrix-expressing cells underlying skeletal muscle fibrosis, which is often observed in aging, chronic disease or in conditions in which regeneration is impaired\(^{81,162}\).

Despite the clear evidence of important roles for stromal progenitors in health and disease, our understanding of their contribution to regenerating tissues is limited by the lack of reliable markers to define them. Mesenchymal populations within various tissues have been identified
using a variety of markers including Nestin, Gli1, Grem1, Lepr, PDGFRA, SCA-1 (Ly6a) and CD34, and combinations thereof\textsuperscript{13,18,19,23,24,44}. Cre lines based on these markers have provided fundamental insights into their lineage contributions \textit{in vivo}\textsuperscript{1,34}. Mesenchymal cells with potential stem cell activity, or mesenchymal stem cells (MSCs), were originally identified in BM\textsuperscript{14,22}. Within this tissue, Lepr expression has been used to identify mesenchymal cells that contribute to the adipogenic and the osteogenic lineages\textsuperscript{24}, and like LepR, Grem1 expression can also be used to identify bone mesenchymal cells with endogenous osteogenic lineage potential\textsuperscript{23}. Gli1 appears to mark an MSC-like population across multiple tissues including bone, and Gli1\textsuperscript{+} cells contribute to myofibroblasts in various fibrotic models, including those of the heart, lung, kidney, liver and BM\textsuperscript{44,46}. While these markers allowed an initial glance at the \textit{in vivo} role of MPs, it is unclear whether they label specialized, possibly lineage committed subsets of cells, and a marker capable of reliably identifying immature progenitors across multiple tissues is currently lacking. As a consequence, the extent of MP heterogeneity and the molecular mechanisms modulating MP fate and function are poorly understood. Here, we identify the transcriptional repressor Hic1 as a marker of multiple lineages of MPs both within skeletal muscle and across tissues. Novel transgenic tools based on this marker allow the lineage tracing and molecular characterization of these cells and their progeny in skeletal muscle, highlighting their role as a signaling hub in the response to damage, and revealing the extent of their contributions to mature lineages in homeostasis and regeneration. In addition, conditional deletion of Hic1 points to a role for this factor in controlling the quiescence, and therefore the abundance of tissue resident MPs in homeostasis, revealing a first insight into the mechanisms underlying their regulation. Interestingly, the observed increase of MP hyperplasia following
Hic1 loss extends into the BM and this is associated with an approximate doubling of HSCs, which may reflect an enhanced capacity within the HSC niche.

3.2 Results

3.2.1 HIC1 marks mesenchymal progenitors within skeletal muscle

To identify MP-specific markers we fractionated whole skeletal muscle into multiple populations (Fig. 3.1a and b) and focused on the analysis of the Lin⁻ (CD31⁻, CD45⁻, Ter119⁻) Sca-1⁺ population which we previously showed was enriched for MPs. RNA-seq analysis was applied to these fractions (entire population or popRNA-seq) to identify markers enriched in the Lin⁻ Sca-1⁺ fraction (Fig. 3.1c). As expected, we noticed a clear enrichment (Fig. 3.2a) of several known MP-related markers including Pdgfra, Thy1, Gli1 and Lepr. Interestingly, pericyte markers including Rgs5, Pdgfrb, Mcam, Cspg4, Kcnj8 were substantially enriched in the Lin⁻ Sca-1⁻ fraction as were markers reflective of the tenogenic lineage (Scx, Mlx, Tnmd, Kera) (Fig. 3.2b). This fraction also contained a strong satellite signature associated with expression of Pax7 (Fig. 3.1c and 2.2c). Within the Lin⁻ Sca-1⁺ fraction we noted a substantial enrichment in the transcript for Hypermethylated in cancer 1 (Hic1) (Fig. 3.1c and 2.2d). Hic1 encodes for a transcriptional repressor with potential tumor suppressor activity, that has been shown to directly regulate genes involved in the cell cycle. It was originally identified based on observations that the HIC1 locus was hypermethylated and transcriptionally silenced in diverse human tumors. Characterization of Hic1 in the developing mouse embryo showed that
Figure 3.1 | Identification of Hic1$^+$ cells in the TA muscle.
a) flow cytometry plots of sorted cell populations used to generate the RNA-seq samples shown in 1b. b, schematic overview of strategy used to purify MPs from TA muscle into fractions indicated by FACS gates in a. Numbers in parentheses indicate percent of the total mononuclear fraction from 3 independent isolations. c, heat map from RNA-seq analysis of the various fractions indicated in a. Select genes associated with various cell types within muscle are shown in the right panel, with Hic1 showing enrichment in the Lin$^-$ Sca-1$^+$ fraction.
Hic1 is predominantly restricted to mesenchyme within various tissues and organs. This was further corroborated using a citrine knock-in to the Hic1 locus.

To validate the observed enrichment of Hic1 in the FAP fraction of the RNA-seq data, immunofluorescence (IF) was carried out on adult tibialis anterior (TA) muscle. HIC1 was detected within the nuclei of blood vessel-associated Sca-1+ cells in the interstitial space of the muscle. No positive cells were found associated with the myofibers as defined by laminin staining of the muscle basal lamina (Fig. 3.2e). Analysis of a novel mouse strain containing a nuclear LacZ knock-in Hic1nLacZ/+ confirmed the presence of Hic1+ cells within the interstitial space (Fig. 3.2e, 3.3a and b) in association with capillaries (Fig. 3.4a). Within arterioles and arteries, Hic1+ cells were observed in the adventitia (Fig. 3.4b). Infrequent, vessel-associated Hic1+ cells were also detected in the periphery of tendons (Fig. 3.3a bottom panels) and around the myotendinous junction. In summary, Hic1/HIC1 is expressed in skeletal muscle resident fibro/adipogenic progenitors.

3.2.2 Hic1 marks multiple MP populations in skeletal muscle

To more effectively mark, isolate and enable fate mapping of Hic1+ cells in skeletal muscle regeneration, a CreERT2 knock-in line was generated and bred to a Rosa26 conditional tdTomato, Rosa26LSL-tdTom, reporter line (Fig. 3.5a and b). Flow cytometry (FCM) analysis of tdTomato+ cells immediately after TAM treatment, showed that 94.8 ± 0.7% of this population was Lin−, Sca-1+ FAPs (Fig. 3.5c). In addition, the transcriptome of tdTomato+ cells highly correlated (R2=.93) with that of the Lin− Sca-1+ fraction from Fig. 3.1c (Fig. 3.1c, 3.5b, 3.5d), suggesting extensive overlap between the two subsets. Indeed, labelled cells were enriched for
Figure 3.2 | Localization of Hic1 transcript and protein expression.

a-c, expression of various genes indicative of different skeletal muscle cell types, represented as Fragment Per Kilobase of transcript per Million reads (FPKM) and derived from RNA-seq analysis of FACS-enriched populations from muscle (N=3-4). d, Hic1, Ly6a and Mcam transcript abundance in the different fractions from the RNA-seq analyses (N=3-4) data reflect the mean ± s.d. from at least 3 biological replicates. e, anti-HIC1 staining of 8-week-old adult TA muscle. HIC1+ cells (white arrowheads) reside outside the Laminin+ myofiber basal lamina, are Sca-1+ and located in close proximity to CD31+ blood vessels. MF, myofiber.
Figure 3.3 | Localization of $Hic1^{nlacZ/+}$ reporter expression.

a, the distribution of Hic1$^+$ cells was evaluated in whole-mount (upper left panel) and *in situ* X-gal stained sections of TA muscles and the myotendinous junction (bottom series) from $Hic1^{nlacZ/+}$ mice. The muscle fibers have been stained with phalloidin (PLN). b, strategy used to generate a conditional $Hic1$ allele which incorporates a cassette containing an IRES element upstream of nuclear LacZ. Deletion of the major coding exon of $Hic1$ leads to expression of nuclear LacZ. PCR demonstrating deletion of exon 2 of $Hic1$. TN, tendon.
Figure 3.4 | Perivascular location of *Hic1* positive cells.
a, *Hic1* cells were enumerated from stained samples and found to localize in juxtaposition to blood vessels, data reflect the mean ± s.d. from at least 3 biological replicates. b, identification of *Hic1* cells within the adventitia of a small arteriole.
Figure 3.5 | Generation and characterization of a Hic1CreERT2 knock-in allele for lineage tracing of MPs.

a, to maintain wild-type HIC1 expression, an IRES-CreERT2 cassette was inserted into the 3’ untranslated region of the Hic1 gene. The FLpPed PGK-neo cassette was deleted through interbreeding to a germline FLP-expressing line, and all experiments utilized the PDK-neo deleted line. Unlike Hic1LacZ/LacZ mice, Hic1CreERT2/CreERT2 mice are viable, healthy and fertile. b, overview of the Hic1CreERT2 knock-in allele and associated lineage-tracing strategy. c, flow cytometry analysis of enzymatically dissociated TA muscles from Hic1CreERT2; Rosa26LSL-tdTomato mice post-TAM shows that tdTomato+ cells are negative for CD31 and CD45, and predominantly express Sca-1. d, pair-wise comparisons in log2 FPKM scatter plots from RNA-seq analyses between tdTomato-enriched cells versus the fractions shown in Fig 3.1b, c, e, individual gene expression plots from RNA-seq analyses of tdTomato+ and other marker-enriched muscle mononuclear fractions. Hic1 and Ly6a transcript abundance in the various fractions. f, expression of various genes within the tdTomato-enriched fraction in comparison to other indicated fractions, represented as FPKM and derived from RNA-seq analysis of FACS-enriched populations from muscle (N=3-4). e,f, data reflect the mean ± s.d. from at least 3 biological replicates.
the expression of MP and stem/progenitor cell markers such as Cd34, Gli1, Pdgfra, Thy1, but lacked markers associated with myogenic cell fate such as Pax7 (Fig. 3.5e and f).

While FAPs comprise the vast majority of cells labelled with this approach, FCM data indicates that about 5% of tdTomato+ cells are negative for Sca-1, implying additional cell types also express Hic1. To further explore the heterogeneity within Hic1 expressing cells, 2,173 Lin−; tdTomato+ cells were profiled using single cell (sc) RNA-seq. Cluster analysis revealed three subsets, a major one expressing FAP markers as expected, and two smaller subsets characterized by absent/low Ly6a expression representing ~11% of the total (Fig. 3.6a). These 2 populations were also low for FAP markers Cd34, Ly6a and Pdgfra, but were enriched for markers reflective of tenogenic cells (Scx, Mκ, Tnmd, Kera) and pericytes (Rgs5, Kcnj8, Mcam, Pdgfrb) (Fig. 3.6a-c). Taken together, these findings indicate that Hic1 is expressed in at least 3 phenotypically distinct mesenchymal cell populations within skeletal muscle.

In situ analysis of TdTomato+ cells revealed a basement membrane shared with CD31+ capillaries (Fig. 3.7a and b), and further enabled the visualization of a surprisingly intricate network of dendritic processes that circumscribed host and neighboring vessels. Given their high frequency and co-expression of Pdgfra, this population represents FAPs (Fig. 3.7c). Under homeostatic conditions these cells appeared quiescent, as shown by the negligible levels (0.17 ± 0.06%) of EdU incorporation and of transcripts associated with proliferation and cytokinesis (Fig. 3.14e). However, when TdTomato+ sorted cells were placed in culture, they exhibited robust colony forming activity with a frequency of 1 in 17 ± 5.7, a frequency congruent with that reported by other groups for skeletal muscle and LepR+ cells in the BM. Collectively, these findings demonstrate that Hic1 is expressed by multiple quiescent mesenchymal populations.
Figure 3.6 | Single cell transcriptomic analysis reveals four MP subpopulations at homeostasis in muscle.
a, t-SNE plot of scRNA-seq data from enriched tdTomato⁺ MPs one-week post-TAM induction. b, heat-map of scRNA-seq data showing enriched genes in the 4 different clusters. FAP-fibro/adipo progenitors, TP tenocyte progenitors, PP-pericyte progenitors. c, violin plots showing expression of select genes from the scRNA-seq data.
Figure 3.7 | tdTomato fluorescent reporter reveals extensive cytoplasmic ramifications of MPs, extended into their surrounding microenvironment.

a, representative transverse and longitudinal sections of the TA muscle from *Hic1*<sup>CreERT2</sup>-reporter mice. Numerous tdTomato<sup>+</sup> cells can be observed within the muscle interstitium. Individual tdTomato<sup>+</sup> cells are shown in the lower panel with their extensive dendritic-like processes and close association with CD31<sup>+</sup> endothelial cells. 

b, high magnification view of a Laminin-stained tdTomato<sup>+</sup> cell on a capillary.

c, representative images of interstitial muscle and a tendon-containing region from the TA of a *Hic1*<sup>CreERT2/CreERT2</sup>; *Rosa26<sup>LSL-tdTom</sup> Pdgfra<sup>H2bEGFP</sup> line.
3.2.3  **Hic1**

**cells do not contribute to parenchymal or vascular lineages post injury**

The availability of a reporter system labeling multiple mesenchymal lineages allows the morphological and molecular characterization of their response to damage. To this end, reporter expression was visualized at different time points following skeletal muscle damage (Fig. 3.8a). In whole-mount preparations of the TA muscle, 3 d post-injury (p.i.), a substantial increase in reporter gene expression was observed using either conditional tdTomato or LacZ (*Rosa26\(^{LSL}\)LacZ\(^{\Lambda}\)*) reporters (Fig. 3.8b). Collection at various time points post-injury revealed dynamic changes within the traced population (Fig. 3.9). An increase in cell numbers (1.9 fold) compared to uninjured controls was observed starting two days after damage. This correlated with an ~ 9-fold increase in labelled cells that were no longer associated with vessels (Fig. 3.10a, b and c). TdTomato\(^{+}\) cell number peaked on 3-4 d with a total increase of 2.7-fold over baseline. At this stage tdTomato\(^{+}\) cells exhibit a markedly modified morphology, forming an extensive interlaced network of dendritic processes that circumscribed the regenerating myofibers. By 7 d post injury (p.i), tdTomato\(^{+}\) cell number begins to recede, and by 14 d p.i., tdTomato cell number and distribution have returned to baseline (Fig. 3.9, 3.10b). This is in alignment with previous reports showing that FAPs undergo extensive apoptosis from 4 - 7 d post-injury, returning to pre-injury levels by 9 d\(^{81}\). At 28 d post-injury tdTomato\(^{+}\) cell distribution is indistinguishable from un-injured controls (Fig. 3.9, 3.10b). At this stage, tdTomato cells are again vessel-associated, and resumed their characteristic quiescent morphology within the interstitium of the regenerated tissue (Fig. 3.9). As shown in Figure 3.9, no appreciable tdTomato\(^{+}\) myofibers nor CD31\(^{+}\) endothelial cells were detected, indicating that the mesenchymal lineages traced in the
Figure 3.8 | Marked transformation of the quiescent MP phenotype following injury-induced activation.

a, experimental plan for lineage tracing of MPs using the $\text{Hic1}^{\text{CreERT2}},$ reporter mice and NTX induced muscle injury. Following TAM-administration a wash-out period of at least 2 weeks was used prior to sample collection. b, two different Cre-dependent reporter lines were used to follow MP activation following NTX injection into the TA muscle. Whole mount stained and imaged TA muscles are shown in non-damaged and 3 d post damage.
Figure 3.9 | *Hic1<sup>+</sup>* cell dynamics following NTX-induced injury of the TA.
Analysis of TA muscle sections at various time points after NTX-induced injury shows numerous changes in tdTomato<sup>+</sup> MP number, morphology and distribution.
**Figure 3.10 | Hic1+ cells in skeletal muscle proliferate and are migratory after injury induced activation.**

a, numerous tdTomato+ cells are apparent within the injured muscle interstitium; at D4 these are not vessel associated. b, quantification of tdTomato+ MP number at various time points post-injury. Following injury, tdTomato cell numbers return to baseline by D14. c, enumeration of the distribution of tdTomato+ cells indicates that ~1/3 of the cells exhibit a non-vessel association and this returns to baseline by D14. b, c, **P<0.01, ***P<0.001, statistical significance (one-way ANOVA with Bonferroni’s post ad-hoc tests) relative to D0 baseline. ns, not significant.
Hic1\textsuperscript{CreERT2}; tdTomato conditional reporter mice do not contribute to these two lineages. This is in contrast to previous literature, suggesting vessel-associated MPs retain myogenic potential\textsuperscript{20}. However, our findings are also consistent with a recent report on the origin of new endothelial cells in the heart following injury, where labelled MPs were reported not to appreciably contribute to the endothelial cell pool\textsuperscript{167}. MPs have been reported in essentially every tissue analyzed\textsuperscript{20,44}. A preliminary analysis of multiple tissues revealed vessel associated, tdTomato\textsuperscript{+} cells were present in liver, pancreas, lung, kidney, central nervous system etc., suggesting that they may play similar roles across the organism. To ascertain whether the response to damage of Hic1-expressing cells was comparable across these different organs, MP activation following injury was also assessed in an unrelated tissue, the pancreas. In the pancreas, conditional Hic1\textsuperscript{CreERT2}; tdTomato cells can be visualized in close association with CD31\textsuperscript{+} vessels (Fig. 3.11b and d). Following caerulein-induced pancreatitis robust MP activation is observed (Fig. 3.11a-c and 3.12) which is also coupled with MP expansion and dissociation from vessels (Fig. 3.11b-d). Similar to that observed in skeletal muscle, following repair, MPs returned to a pre-injury frequency and distribution, with negligible contribution to the parenchyma or CD31\textsuperscript{+} vasculature. In aggregate, our findings show that the Hic1 expressing cells are highly dynamic following damage. To further explore the full extent of their activities and fates, purified MPs from TA muscle were subjected to transcriptomic analysis.

3.2.4 Hic1\textsuperscript{+} cells coordinate multiple aspects of skeletal muscle regeneration

To better understand the function of Hic1\textsuperscript{+} MPs during skeletal muscle regeneration, RNA-seq and scRNA-seq was used to characterize the nature of their transcriptome 1, 2, 3, 4, 5,
Figure 3.11 | Hic1⁺ cells are found in the pancreas and exhibit similar behavior in response to caerulein-induced injury.

a, schematic overview of pancreas injury induction. b, detection of tdTomato⁺ cells along with IF staining of shown proteins in the acinar (left) and islets (right, demarcated by dashed line) of 9-week old pancreas. TdTomato⁺ cells are associated with blood vessels (white arrowheads). c and d, quantification of the number and distribution of tdTomato⁺ cells following injury. c and d, a minimum of 5 sections from 3 independent mice were used to enumerate tdTomato⁺ number and distribution.
Figure 3.12 | *Hic1* cell dynamics following caerulein-induced injury of the pancreas.
Visualization of tdTomato$^+$ cells in combination with IF detection of indicated proteins at various times following caerulein-induced injury. Dashed lines delineate the undamaged pancreatic islets.
7, and 10 days post-injury. Bioinformatic and ingenuity pathway analyses (IPA) revealed a
dynamic temporal succession of transcriptional programs indicative of multiple diverse functions
such as cytokine production, cell proliferation, trophic factor production, extracellular matrix
production and remodeling, and basement membrane synthesis (Fig. 3.13a). Within the first 24
hours p.i. numerous cytokine genes were upregulated, including those associated with neutrophil
recruitment (Cxcl1 and 5), adhesion, recruitment, expansion and maturation of monocytes
(Cxcl2, Cxcl14, Csf2 and Ccl7). Interestingly, cytokine production was transient, and peaked at
24 hours p.i. and quickly diminished thereafter (Fig. 3.13a-d and 3.14a). At this time point, 29 ±
3.0% of the tdTomato+ cells were positive for the neutrophil chemotactic cytokine CXCL5 (Fig.
3.14b and c). ScRNA-seq analysis of 4,820 cells revealed that this transcript was abundant in the
Ly6a+, Cd34+, Pdgfra+ sub-clusters, and nearly absent in the pericytic and tenogenic clusters
(Fig. 3.13d).

On the second day following injury, numerous genes associated with cell cycle entry
(Cdkn1a), G1-S phase transition (Ccnd1, d2 and d3, Cdk4) and proliferation (Mki67) were
coordinately upregulated (Fig. 3.14d). These transcripts peaked on 3 and 4 d, consistent with
increased cell numbers and EdU incorporation (Fig. 3.14e), as well as with the peak of DNA
synthesis previously reported for FAPs by Joe et al.18. On day 2 p.i., cells expressing
proliferation-related genes clustered separately and represented ~ 15% of the profiled population
(377 cells out of 2,556 cells) (Fig. 3.15a and b). Transcript signatures representative of all the
three Hic1-expressing lineages identified in undamaged tissue were detected within this Mki67-
expressing cluster, indicating that they all participated in cell proliferation (Fig. 3.15c).
Transcripts associated with the cell cycle were greatly reduced in all clusters by 4 d, with
negligible detection at 14 d.
Figure 3.13 | The activated MP phenotype displays stage-specific activities indicative of a highly-coordinated response to injury.

a, RNA-seq analysis of tdTomato+ enriched MPs at various time points post-injury. The profiles of select genes representative of the identified cellular activities are shown. Cxcl5, cytokines; Mki67, cell cycle; Postn, provisional matrix; Lamc1, basement membrane. b, scRNA-seq t-SNE clustering of cells from D0, D1, D2, and D4. c, heat-map of scRNA-seq hierarchically clustered genes that highlight various MP functions in regeneration. The vertical colored lines reflect the programs indicated in a. T, tenogenic cluster; P, pericytic cluster. d, t-SNE plots of scRNA-seq data for Pdgfra and Cxcl5. The pericytic and tenogenic lineages are indicated by magenta and blue dashed lines, respectively.
Figure 3.14 | MPs increase cytokine production for 24-48 hours after skeletal muscle injury before entering the cell cycle.

a, t-SNE plot of scRNA-seq samples collected at the indicated times p.i., along with tSNE heatmaps of individual cytokine genes. b, anti-CXCL5 staining in control and NTX-injured TA muscles, with CXCL5, tdTomato^+ doubly positive cells present the post-injury population. c, quantification of CXCL5^+, tdTomato^+ cells post-injury. d, numerous genes associated with cell cycle entry, mitosis and cytokinesis are elevated on D1-3 post-injury. e, IF of EdU^+, tdTomato^+ cells at 0 and 4 d post-injury. FCM analysis of doubly tdTomato^+ an EdU^+ cells post-injury, along with quantification.
Figure 3.15 | Proliferation peaks between day 2 and 4 in MPs.

a, flow cytometric analysis of EdU incorporation in Hic1-reporter cells in samples from notexin-induced TA muscle preparations. CD31 and CD45-FITC conjugated antibodies were used to identify endothelial and BM-derived cells, respectively and PE filter/channel was used for detection of tdTomato\(^+\) cells. EdU Click-it plus Pacific blue azide kit was used for detection of EdU\(^+\) cells. b, t-SNE plots of clusters expressing Mki67 and different times post-injury. c, tSNE plot and tSNE heatmaps of genes associated with the tenogenic and pericytic lineages, Scx and Rgs5, respectively. Genes associated with the cell cycle (Birc5 and Ccnd1) are enriched in these populations p.i.
Shortly following the beginning of MP expansion, transcripts related to the production of extracellular matrix (ECM) became notable. Numerous transcripts reflective of the production of provisional ECM, known to be transiently deposited during successful skeletal muscle regeneration\textsuperscript{168}, were present, with \textit{Periostin} (\textit{Postn}) being highly abundant (Fig. 3.16a) both at the transcript and protein levels (Fig. 3.16c). At the single cell level, the message for \textit{Postn} was most evident in a population of PDGFR\(\alpha^+\) cells exhibiting a myofibroblast-like signature (\textit{Acta2}, \textit{Col1a1}, \textit{1a2}, etc.) (Fig. 3.16b and c). Interestingly, this population was also enriched for expression of \textit{Adam12}, a gene associated with fibrogenic cells in skin and skeletal muscle\textsuperscript{80}. However, \textit{Postn} was detected within all three Hic1-labelled lineages to varying extents, with pericytes having the least amount. ECM production was in decline by 7 d day p.i., a time at which numerous proteases associated with matrix remodeling were induced, such that by 10 days p.i. much of the provisional matrix had been removed and regenerated myofibers began to emerge. By 14 days p.i., extracellular POSTN was no longer detectable (Fig. 3.17a). Thus, Hic1\(^+\) cells transiently acquire a myofibroblast-like phenotype and secrete an ECM-rich scaffold, providing a favorable microenvironment to promote myofiber regeneration.

Interestingly, several genes (\textit{Col8a2}, \textit{Col14a1}, \textit{Col15a1}, \textit{Fbln1} and 5, \textit{Hspg2}, \textit{Lama2}, \textit{Lama4}, \textit{Lamc1}, \textit{Lamb2}, \textit{Nid2}) associated with the basal lamina were expressed with clearly distinct kinetics relative to ECM genes (Fig. 3.17b). These transcripts were abundant in quiescent MPs, quickly declined following activation (Fig. 3.17b), and slowly increased during ECM production, returning to baseline levels only as late as 14 d post-injury. Thus, changes in expression of basal-lamina-associated genes in Hic1\(^+\) FAPs corresponded with the disruption and restoration of the laminin sheath that envelopes myofibers (Fig. 3.17b).
In summary, the combination of lineage tracing, popRNA-seq and scRNA-seq enables efficient deconvolution of the heterogeneity of MPs and the precise identification of not only the role of these cells in regeneration, but also which subset carries out a given function. In doing so, we reveal how MPs, through the temporally-ordered secretion of cytokines and ECM components, modulate the regenerative environment.
Figure 3.16 | MPs participate in extracellular matrix remodeling.
a, multiple genes associated with provisional ECM production and turnover are coordinately expressed starting at D1 through to D14. b, distribution of Postn transcripts in t-SNE derived clusters of single cells. c, tSNE heatmaps demonstrating that activated MPs exhibit dynamic expression of numerous genes associated with a myofibroblast-like phenotype. Refer to a, for sample ID.
Figure 3.17 | MPs transiently express periostin before re-expression of basement membrane associated transcripts.

a, IF staining of 4 (ctrl - contralateral TA), 4 and 14 d post-injury TA muscles from Hic1; TdTomato reporter mice with anti-POSTN. b, various constituents of the basement membrane are produced during the later stages of the regenerative process as determined by RNA-seq analysis of sorted Hic1CreERT2; tdTomato cells. tSNE heatmaps derived from scRNA-seq analyses demonstrate a similar pattern of gene expression. Refer to a, for sample ID.
3.2.5 Hic1+ MP progeny contribute to multiple mesenchymal lineages in regenerated skeletal muscle

Next, we tested whether Hic1+ MPs influence regeneration solely through trophic factors, or whether they also produce differentiated progeny. In successfully regenerated skeletal muscle, overall MP frequency and distribution was indistinguishable from the un-injured control (Fig. 3.8-10). However, FCM analysis 14 days post-injury revealed a substantive increase in the Sca-1- subset of tdTomato+ cells (25.1 ± 1.53% versus 5.2 ± 0.4% in uninjured muscle) (Fig. 3.18a). Bulk RNA-seq analyses at this time point also revealed an increase in the expression of genes associated with pericytes (Rgs5, Mcam, Notch3, Kcnj8) as well as with tenocytes (Tnmd, Mkx, Kera, Scx, etc., Fig. 3.18b and c). Interestingly Col22a1, a gene required for robust myotendinous junctions\(^{85,87}\), was also substantially increased. Congruent with these transcriptomic signatures, following damage tdTomato+ cells were found embedded in a COL22a1-rich ECM at the distal ends of myofibers, which they enveloped with processes that also inserted into the tendon (Fig. 3.19a). In contrast, participation of tdTomato+ cells in these basket-like anatomical structures was not observed in undamaged tissue, in which the traced cells associated with myotendinous junctions were restricted to CD31+ blood vessels, with no appreciable contribution to the junction itself or to tenocytes. ScRNA-seq profiling confirmed the expansion of the tenogenic tdTomato+ Ly6a-, Cd34-, Pdgfra- subpopulation, and revealed a significant overlap between expression of tenogenic genes and Col22a1. In aggregate, our data strongly suggests that myotendinous junction cells are derived from Hic1-expressing progenitors and represent modified tenocytes (Fig. 3.19b and c).
Figure 3.18 | *Hic1*-tracked MPs directly contribute to regeneration of the myotendinous junction, but not the myogenic nor the endothelial compartments.

a, analysis of Sca-1 expression in tdTomato$^+$ cells post injury shows a significant increase in Sca-1- tdTomato$^+$ cells. b, RNA-seq analysis of tdTomato$^+$ enriched fraction from D14 TA muscle post-injury shows enrichment of transcripts associated with several mesenchymal lineages including tendon (*Scx, Mkx, Tnmd, Kera*) and the myotendinous junction (*Col22a1*). c, hierarchical clustering of pericyte-associated genes at various times p.i.
In addition to the tenogenic cluster, another prominent Hic1-expressing subpopulation in undamaged tissue was defined as pericyte-like based on the expression of Rgs5, Acta2, Mcam, Notch3, Kcnj8 and Tagln. Fourteen days p.i. this subpopulation increased modestly, from 5.5% to 7.3% of the total tdTomato+ scRNA-seq profiled cells (Fig. 3.20a and b). Consistent with this, ACTA2+, TAGLN+ or MCAM+ pericytes positive for tdTomato were evident at homeostasis and their frequency did not change appreciably post-injury. Importantly, Hic1-expressing pericytes or their progeny did not contribute to myogenesis.

Within regenerated skeletal muscle, tdTomato+ cells contributed with variable extents to other mesenchymal cell types. Rare interstitial tdTomato+ perilipin+ adipocytes were observed; consistent with the rarity of adipocyte generation during successful regeneration in mice, they were relatively infrequent, and we did not detect an adipogenic cluster in the scRNA-seq analyses (Fig. 3.20c and d). Furthermore, after damage numerous non-vessel associated tdTomato+ cells were observed within the skeletal muscle endo, ecto and perimysium, suggesting that the progeny of Hic1-expressing cells also contributes to the capsular components of the regenerated skeletal muscle (Fig. 3.20e).

In summary, Hic1-expressing cells contribute to multiple stromal elements following skeletal muscle damage, including pericytes, tenocytes, adipocytes and as yet, undescribed specialized myotendinous junction cells.
Figure 3.19 | TdTomato positive myotenocytes increase in numbers and are found to express the myotendinous junction specific FACET COL22a1, 14 days after muscle injury.

a, t-SNE plots of Hic1;tdTomato, tendon-expressed genes from 2,173 and 3,527 scRNA-seq profiled cells from D0 and D14, respectively. b, violin plots of select genes from scRNA-seq analyses at D14 post-injury, the cluster ID corresponds to that in Fig. 3.6a-c. c, anti-COL22a1 staining of uninjured and D14 post-injury TA muscles shows a marked increase in tdtomato⁺ cells embedded in a COL22a1 rich matrix in the regenerated muscle (white arrowheads). TN, tendon.
Figure 3.20 | Hic1-reporter cell fate mapping demonstrates that these cells contribute to several mesenchymal cell lineages within regenerated skeletal muscle.

a, tSNE heat maps of pericyte-associated genes at 0 d and 14 d p.i. b, IF staining of 14 d p.i. TAs with anti-MCAM, anti-ACTA2 and anti-TAGLN. c, hierarchical clustering of adipocyte-associated genes at various times p.i. d, identification of tdTomato+ interstitial adipocytes at 14 d p.i. by anti-perilipin staining of Hic1CreERT2/CreERT2; tdTomato derived TA samples. e, visualization of tdTomato+ cells within the fibrous tissue that envelopes the muscle fibers at 14 d p.i.
3.2.6 HIC1 regulates MP quiescence

In adult mice under homeostatic conditions there is limited cellular turnover, and both muscle satellite stem cells and FAPs/MPs are in a quiescent state. On the second day following injury, as described above and in our previous work, these two populations enter the cell cycle\(^1\), and quickly expand for the next 3-4 days\(^{159}\). We used a mouse strain expressing LacZ from the Hic1 locus to follow Hic1 expression during this dynamic process. Under homeostasis, X-gal\(^+\) cells can be readily observed in the interstitium of Hic1\(^{nLacZ/+}\) skeletal muscle. Interestingly, despite the significant expansion of this population, following injury few X-gal\(^+\) nuclei were evident in the damaged region (Fig. 3.21a). However, ten days p.i., the frequency and distribution of X-gal\(^+\) cells returned to pre-injury levels. Thus, Hic1 expression was readily detectable in quiescent MPs, declined in activated MPs, and returned to normal as these cells re-entered quiescence (Fig. 3.21a), suggesting a possible role in regulating this process. Consistent with this, Coller et al.\(^{116}\) identified Hic1 as a quiescence program-associated gene in human lung fibroblasts and heterotopic expression of HIC1 has been shown to induce growth arrest\(^{129,135}\). To test its role in maintaining quiescence in vivo, Hic1 was ubiquitously deleted in 9-week old mice (UBC-CreERT2; Hic1\(^{f/f}\) or \(^{+/+}\)). In the Hic1\(^{ff}\) strain, CRE activity results in the replacement of the Hic1 coding sequence with that of nuclear LacZ, allowing the identification of the knock-out cells still transcribing the locus. Multiple tissues were collected at various times post-TAM injection and X-gal stained to identify Hic1\(^+\) cells (Fig. 3.21b-d). In 7 of the 11 tissues examined, an increase in X-gal\(^+\) cells relative to untreated controls was apparent 10 days following TAM administration. This increase extended to all tissues examined to 250 days later,
Figure 3.21 | HIC1 regulates MP quiescence—deletion of Hic1 leads to an activated MP-like phenotype and an increase in MP number.

a, chronological series of X-gal stained muscle sections following notexin-induced TA-muscle injury in Hic1^nLacZ/+ mice. At 3 and 4 d.p.i. there are fewer X-gal cells apparent, but by 10-14, X-gal+ cells emerge in a distribution and frequency similar to the un-injured 0 d control sample.
b, deletion of Hic1 leads to significant short and a stable long-term increase in nLacZ+ MPs in multiple tissues. Mean ± SD are shown and derived from 3 sections/tissue from the indicated mice. c, X-gal stained tissues from UBC-CreERT2; Hic1^f/f and UBC-CreERT2; Hic1^fl mice at 250 d post-TAM administration. All tissues were collected and processed for in situ X-gal staining of cryosectioned samples. b and c, mean ± SD are shown, in a N is indicated in the legend, and an N=8 for c. Significance, in b, *P<0.05, **P<0.01, ***P<0.001, statistical significance (one-way ANOVA with Bonferroni’s post ad-hoc tests) relative to corresponding 9-week Hic1^nLacZ/+ baseline. ns, not significant. c, statistical significance (*** P<0.001) was determined by unpaired t-test.
but its extent was similar to that detected at the earlier time point. At this time, the efficiency of 
Hic1 deletion was determined by RNA-seq to be 92%, 86%, 91% and 97% (N ≥ 6) in skeletal 
muscle, heart, liver and lung, respectively. Thus, deletion of Hic1 leads to a rapid increase in 
MPs which is maintained over time, leading to a higher number of Hic1-expressing cells in most 
tissues at steady state. The increase in MP number is evident by flow and histological lacZ 
staining (Fig. 3.21b-c and 3.23a-c). To uncover the transcriptomic changes that occur in MPs 
due to Hic1 loss, MPs were sorted from skeletal muscle and profiled by RNA sequencing. 
Ingenuity Pathway analysis of RNA-seq data at 10 d post-TAM showed an enrichment for 
programs associated with cell survival, proliferation, migration and invasion in Hic1 deleted cells 
versus control (Fig. 3.22a and b). These programs are also induced, albeit to a higher extent, 
during damage-induced MP activation, indicating that deletion of Hic1 leads to a partially 
activated MP phenotype (Fig. 3.23d). Taken together, these findings indicate that HIC1 directly 
regulates quiescence and through this mechanism, the number of MPs present in the tissue at 
homeostasis.

Hic1 expression marks the FAP subpopulation, and these cells efficiently generate both 
myofibroblasts and adipocytes18. In the Hic1 cKOs, ectopic fat was identified in multiple 
tissues, with the pancreas exhibiting the most extreme phenotype (Fig. 3.24a). Interestingly, 
while the number of tissue resident MPs plateaued shortly after deletion of Hic1, the number of 
parenchymal adipocytes continued to accumulate with age (Fig2.24b). Congruent with Hic1 
being expressed in MPs and not their differentiated progeny, the ectopic adipocytes are negative 
for nLacZ (Fig. 3.24a and b). Thus activated, Hic1-deleted MPs generate low numbers of 
differentiated progeny with time, which then accumulates in tissues. This is consistent with the 
notion that the plateau reached by MP numbers following Hic1 deletion is not due to a cessation
Figure 3.22 | Modified MP gene expression profile after deletion of Hic1.

a, IPA interrogation of transcriptomic data (N=4) from sorted Hic1-deleted MPs from 8 week old adult mice 10 d post-TAM shows up-regulation of programs associated with proliferation, migration and invasion, and reduction in indices associated with cell death. b, expression of select genes in sorted MPs from the entire hind limb musculature of Hic1<sup>f/f</sup> and UBC-CreERT2; Hic1<sup>f/f</sup> mice 10 d post-TAM. Gene expression is presented as FPKM (N=4).
Figure 3.23 | Deletion of Hic1 in adult mice modifies the MP phenotype.
a, representative FCM plots of the sorting strategy used to enrich for wild-type and Hic1-deleted MPs.  
b, flow cytometry analysis of Lin- Sca-1+ MPs in TA muscle from Hic1+/f and Hic1-deleted mice 10 d post-TAM injection.  
c, histological analysis of TA muscle shows an increase in X-gal stained cells.  
d, heatmap of differentially expressed genes from RNA-seq analysis of sorted MPs from 8 week old TA muscle of Hic1+/f mice and UBC-CreERT2; Hic1+/f mice—all mice received TAM.  
Differentially expressed genes were identified using cufflinks with an N=4.
Figure 3.24 | Deletion of Hic1 leads to mesenchymal hyperplasia and subsequent accumulation of adipocytes in the pancreas.

a, Anti-perilipin staining of pancreas from mice with the indicated genotype at 5 months post-TAM (mice were treated with TAM at 8 weeks of age). b, detection of perilipin adipocytes (white arrows) in pancreas from Hic1-deleted and non-deleted mice at 9 weeks post-deletion (upper panel) and 250 d post-deletion (lower panel). Bottom panel sections were in situ stained for LacZ, then processed for immunodetection of perilipin.
of proliferation, but rather to other mechanisms likely including the continuous production of differentiated progeny.

Quiescence is a key defining property of many stem and progenitor cells, but not of their post-mitotic derivatives, and is typically defined by the ability of a growth-arrested cell to enter the cell cycle. Quiescent MPs can be activated to enter the cell cycle by a number of signals, including but not limited to inflammation and damage, and expand several-fold to produce a microenvironment and progeny to support regeneration and restoration of tissue function. Following regeneration, quiescent MPs "re-appear" in preparation for the next tissue renewal or regenerative event. In this regard, quiescence is vital to MP function; however, with the possible exception of the hedgehog pathway, the mechanisms regulating this fundamental property are poorly defined. The expression pattern of Hic1 in many tissues overlaps with that of the downstream HH effector Gli1, however, Gli1 is also found in brain neural stem cells\textsuperscript{169}, whereas Hic1 expression in the brain is restricted to MPs\textsuperscript{170}. In the lung, Gli1 has been shown to regulate quiescence of lung mesenchyme, and deletion of Gli1 leads to an inappropriate expansion of this compartment\textsuperscript{117}. Similarly, conditional deletion of Hic1 leads to an immediate expansion of MPs across tissues, which suggests that HIC1 participates in the active maintenance of MP quiescence. The robust expression of Hic1 in tissue-resident MPs in homeostasis and its re-emergence in MPs post-injury reflects its role in the quiescent state and the need of MPs to enter the cell cycle following activation. Together, these findings indicate that Hic1 exhibits non-redundant functions and serves as a master regulator of MP quiescence.
3.2.7 Systemic analysis of MPs – characterization of tdTomato$^+$ cells across tissues

In previous sections, *Hic1* was identified as a marker of MPs in skeletal muscle and a related-labelled population in the pancreas. To characterize *Hic1* expression in other tissues and organs, a similar strategy as described above involving the *Hic1$^{nLacZ}$* and *Hic1$^{CreERT2}$*; Rosa26$^{LSL}$tdTomato lines were employed. The later tool has the added advantage of enabling visualization of the complex MP morphology *in situ* as well as isolation of MPs by fluorescence activated cell sorting (FACS).

To characterize the expression pattern of *Hic1* in tissues at homeostasis, histological sections of ten organs from 9-week *Hic1$^{nLacZ/+}$* mice were processed and stained *in situ* with X-Gal. Histological sections were also co-stained with the basement membrane protein laminin and endothelial marker CD31 to highlight MP distribution in skeletal muscle, heart, lung, pancreas, brown adipose tissue, white adipose tissue, cerebral cortex, liver and kidney (Fig 3.25 right). Visualization of β-galactosidase reporter activity revealed a pattern remarkably similar to that observed in skeletal muscle. X-gal positive nuclei were observed in immediate proximity to abluminal basement membrane surfaces of blood vessels throughout all organs and tissues examined.

IF in several tissues was performed to examine the distribution and morphology of tdTomato$^+$ cells in the *Hic1*; reporter mice. Visualization of the tdTomato reporter cells at the histological level showed a similar distribution to the LacZ reporter mice while also revealing the intricate cellular ramifications reminiscent of skeletal muscle resident MPs (Fig. 3.25). These analyses revealed several characteristics of the tdTomato$^+$ population including the marked differences in abundance of these cells within tissues. Consistent with the nLacZ enumeration
Figure 3.25 | \textit{Hic1} expression unifies a perivascular, stromal, reticular MP population organism wide as described in skeletal muscle.

Left, \textit{Hic1}^{C\text{reERT2}^\text{R26^\text{LSL TOMato}}} mice sampled 10 days after tamoxifen administration at 8 weeks to highlight the morphology and distribution of reporter positive cells at homeostasis across representative cryosections of skeletal muscle, heart, lung, pancreas, BAT, cerebellum, WAT, liver and kidney. Right, \textit{Hic1}^{\text{LacZ}^+} reporter identifies a perivascular stromal cell in all tissues examined. The ten tissues shown in the left panel with alternate representative regions from kidney and brain are stained for LacZ and structural markers indicated.
carried out in Fig. 3.21b, both heart and lung contain an abundance of tdTomato$^+$ cells. In all cases, the tdTomato$^+$ cells were confined to the tissue interstitium and appeared non-parenchymal in nature. Congruent with the findings in skeletal muscle and pancreas, tdTomato$^+$ cells resided in the immediate proximity to CD31$^+$ blood vessels, indicating that these cells occupy the perivascular space in all tissues examined.

To better define the nature of tdTomato$^+$ populations within tissues and organs shown in Fig. 3.25, single cell RNA-seq was used to examine their transcriptomes and to query their relatedness to the populations characterized from skeletal muscle. *Hic1; tdTom* mice were allowed to complete the post-natal adolescent growth period and reach young adulthood (eight weeks) before activation of the tdTomato reporter by TAM treatment. To effectively isolate MPs from these various tissues, a template dissociation protocol was developed then specific optimizations for each tissue were incorporated. Furthermore, MPs typically comprise < 10% of the total cell number in any tissue, therefore a FACS-based enrichment strategy was used to enable increased sensitivity for selective interrogation of this cellular compartment (Fig 3.26a). Consistent with the histological findings, flow cytometry analyses of tdTomato$^+$ cells revealed highly variable frequencies within tissues (Fig 3.26a). Using these newly developed methods, tdTomato$^+$ cells were isolated from ten dissociated tissues and scRNA-seq profiles were successfully generated from 28,757 cells using 10x methodology.

3.2.8 ScRNA-seq provides novel insights into MP biology

To interrogate the generated scRNA-seq datasets a variety of bioinformatic tools were employed. Initially, the individual tissue-specific MP datasets were combined to generate an
Figure 3.26 | Single cell RNA-seq reveals inter and intra tissue heterogeneity and superpopulations shared among several organs.

a, Flow cytometric measurement of the percentage of tdTomato positive cells in all tissues sorted for single cell capture (N=1). b, tSNE dimensional reduction plot of MPs colored by supercluster. c, Heatmap of an unbiased hierarchical biclustering of single cell RNA-seq expression data from ten tissue sources using the backSPIN algorithm. The first cellular bifurcation in to superclusters is indicated by red and black dotted lines. d, Number of cells per supercluster expressing pericyte marker Rgs5 and FAP marker Pdgfra. e, tSNE plots of each tissue colored by supercluster markers Rgs5 (red) and Pdgfra (blue) or both (purple) indicating the relative proportions of pericyte like compared to FAP like MPs across profiled tissues. Cells expressing neither marker are colored white. Plot border is colored by tissue identity from a.
aggregate expression matrix. The backSPIN top-down hierarchical biclustering algorithm was used to divide cells into groups irrespective of their source. The two groups formed by the first bifurcation, similar to skeletal muscle, were enriched for \( Pdgfra \) and \( Rgs5 \) (Fig 3.26 b-d). Analysis at the individual tissue level revealed a similar trend, in that two predominant populations existed at varying ratios that can be marked as either \( Rgs5 \) or \( Pdgfra \) positive. In limited situations, there was overlap in the expression of these two markers (discussed below). As the bifurcations progressed it became evident that MP populations from lung, kidney and marrow exhibited unique profiles (Fig. 3.27, clusters 0, 14 and 21) as indicated by the homogeneous bar color. Similarly, multiple liver MP subpopulations also appeared to be specialized (Fig. 3.27, clusters 11 and 12 and 23). Conversely, many clusters were formed from several tissue libraries representing MP subsets that were conserved across multiple tissues (Fig. 3.27, clusters 4, 5, 6, 13, 20 and 22). Bar plots of commonly used markers for MSCs, perivascular mural cells, FAPs and pericytes indicated that tdTomato\(^+\) MPs express many of these established markers to varying extents (Fig. 3.28). Mesenchymal markers such as \( Pdgfra \), \( Cd248 \), \( Cxcl12 \), \( Ly6a \), and \( Pth1r \) are broadly expressed in this population, whereas genes indicative of other germ layers such as epithelium or endoderm were absent, as were genes specific to parenchymal cell function\(^{18,43,44,46,54,109,171-174}\). Furthermore, \( Gli1 \) has previously been used to identify and lineage trace adult pleural and visceral organ-associated mesenchymal cells, and it is broadly expressed within the Hic1\(^+\) population albeit at a low level\(^ {44,46}\). Collectively, these findings suggest that \( Hic1 \) identifies common and unique stromal populations that exhibit appreciable gene signature overlap with that reported for various adult mesenchymal populations.
Figure 3.27 | Single cell RNA-seq reveals inter and intra tissue heterogeneity and subpopulations shared among several organs.
Bar plot of individual clusters generated using the backSPIN algorithm colored to identify tissue source library.
Gene indicative of FAPs/MSCs (Pdgfra, Cd34, Ly6a, Gli1, Cxcl12), hepatic stellate cells (Lrat, Pthr1, Cd248, Gfap, Ngfr), myofibroblasts and mural cells (Acta2), pericytes (Pdgfrb, Rgs5, Kcnj8, Mcam, Abcc9), or liver sinusoidal endothelial cells (Lyve1) are shown. Each vertical bar represents the absolute gene expression (reads) within a single cell.
As noted above, the Hic1 MP can be sub-divided into two mostly mutually exclusive populations marked by Rgs5 and Pdgfra expression (Fig 3.26e). Interestingly, in the liver both markers are co-expressed in a population representative of liver hepatic stellate cells (LHSC). This is based on expression of LHSC-associated genes such as Lecithin Retinol Acyltransferase (Lrat), Des, Gfap, Pdgfra and Cd248 (Fig. 3.28). This provides an explanation for the fraction of cells from cluster 0 (red) that express Pdgfra in Figure 3.26d. There remained a second unidentified population on the liver library however, that was negative for both markers and segregated as a separate population within the atlas tSNE. Unexpectedly, after further analysis of cluster markers this unique liver population was identified as liver sinusoidal endothelial cells (LSECs) as they expressed the vascular markers Adgrl4, Clec4g, Mrc1, Lyve1, Sox18, Cldn5, Podxl and Pecam1. LSECs are typically defined by the co-expression of the lymphatic gene Lyve1 along with endothelial markers. Interestingly, similar to Hic1 this population also exhibits low-level expression of Gli1 (Fig. 3.28). Analogous to that reported in Mederacke et al., the two subtypes of tdTomato+ MP in the liver can be distinguished from one another by the presence (LHSCs), or absence (LSECs), of blue 4-500 nm fluorescence indicative of retinol droplets and this is reflected in the differential expression of Lrat in these populations (Fig. 3.28). These analyses reveal an example of the intra- and inter-tissue tdTomato+ cell heterogeneity observed, in that liver tdTomato+ cells cluster independently from each other and those from other tissues.

To further explore MP population heterogeneity, subsequent analyses included the use of a distinct approach that employed the shared nearest neighbor (SNN) clustering method provided as part of the Seurat single cell seq analysis tool pipeline. Principle component analysis was run to identify important variation and a tSNE projection was again calculated to visualize high
dimensional space on a two-dimensional plot. At this stage, we colored the projection by expression levels of Cxcl12 (which is expressed by all tdTomato\textsuperscript{+} cells), Pdgfra and Rgs5 to identify the FAP and pericyte superclusters, respectively, and the expression of both to identify the LHSCs. The absence of both markers and the presence of Sox18 expression were used to visualize the LSEC population (Fig. 3.29a). As expected, each marker identified distinct regions within the tSNE landscape. To investigate how MP populations are related across tissues and gain insights as to the identity of the clusters, a tSNE plot colored by cluster was generated (Fig. 3.29b). This plot demonstrates that although the superclusters group similar cell types across tissues (Fig 3.29c FAP and PP), some tissues retain sufficient variation to segregate from the mixed clusters. This perspective suggests that most systemic FAPs identified by Pdgfra expression are similar, however, lung, brain and kidney each contain unique FAP-like populations (Fig. 3.29c lungFAP, kidFAP and brnFAP). Similarly, pericyte progenitors from most tissues cluster together (Fig 3.29c PP), while brain and kidney encompass pericyte progenitor like populations with unique molecular profiles (Fig 3.29c brnPP, kidPP, juxPP). Interestingly, in this context tenogenic and myotenogenic progenitors identified and described in detail above (section Fig. 3.19) co-cluster with systemic FAPs (Fig 3.29c MTP). These results are congruent with the findings shown by the backSPIN analysis and further support the observed inter and intra-tissue heterogeneity of characterized MPs and this likely in part reflects tissue-specific function(s). In addition to the observed distinct LSEC signature cluster, further analysis revealed a unique subpopulation within the Pdgfra\textsuperscript{+} supercluster that mapped to the marrow library. Surprisingly, this population was observed to co-express high levels of Lepr and Grem1, which have been used to independently identify mesenchymal stromal cell and skeletal
Figure 3.29 | Unbiased shared nearest neighbor clustering classifies MPs from ten tissue libraries into 13 clusters.
a, tSNE heatmap projection of MPs from all tissues colored by super cluster markers Pdgfra, co-expression of Pdgfra and Rgs5, Rgs5 and Sox18. b, Cluster map tSNE dimension reduction plot colored by cluster identity. c, Outlines in two-dimensional space of assigned clusters and their classification/category. d, Identification of a unique population from BM that expresses BM mesenchymal stromal markers.
stem cells in the BM, respectively (Fig. 3.29d)\textsuperscript{23,24}. The unexpected co-expression of these two markers in this population prompted further interrogation.

### 3.2.9 \textit{Hic1} identifies MPs in bone marrow including \textit{Cxcl12} abundant reticular marrow stromal cells.

In adult BM \textit{Hic1};\textit{tdTomato} labels a rare (< 0.1%) stromal cell population uniformly distributed throughout the tissue (Fig 3.30a). To assess colony forming unit potential of the \textit{tdTomato}\textsuperscript{+} cells, a colony forming unit-fibroblast assay (CFU-F) was performed. Under these conditions, the rare \textit{tdTomato} cells formed adherent colonies (Fig 3.30b) which represented the majority of the total colony forming activity (Fig 3.30c).

Similar to the other tissues examined, scRNA-seq analysis revealed that the BM MPs clustered into several subtypes (Fig 3.30d) including a \textit{Pdgfra} and \textit{Ly6a} (P\textalpha S) expressing FAP like population and a \textit{Rgs5}-expressing pericyte like population (Fig 3.30e). In addition, a population indicative of osteogenic lineage that expressed \textit{Alp}, \textit{Sp7} and high levels of \textit{Coll1a1} was detected and likely represents a rare osteogenic progenitor population (Fig 3.30f). The majority of \textit{tdTomato}\textsuperscript{+} MPs isolated from BM however, make up a \textit{Pdgfra}\textsuperscript{+}, \textit{Ly6a}\textsuperscript{−}, \textit{Cspg4}\textsuperscript{−}, \textit{Nes}\textsuperscript{−}, \textit{Ptn}\textsuperscript{+}, \textit{Notch2}\textsuperscript{+}, \textit{Grem1}\textsuperscript{hi}, \textit{Lepr}\textsuperscript{hi} and \textit{Kitl}\textsuperscript{hi} population that expresses high levels of \textit{Cxl12} transcript in comparison to other MPs (Fig 3.30f). This population is consistent with the \textit{Cxl12} abundant reticular (CAR) cells shown to be responsible for the majority of trophic support in the HSC niche\textsuperscript{24}.  

104
Figure 3.30 | Hic1 identifies multiple subsets of bone marrow MPs, including Cxcl12 Abundant Reticular marrow stromal cells.

a. Fluorescence micrograph of BM (DAPI intense area) from homozygous Hic1^CreERT2^;R26^LSL-TdTomato^ animals demonstrating the rare distribution of tdTomato^+ cells. P - periosteum, CB - Cortical bone, M - marrow. b, A representative adherent tdTomato positive colony from a culture of whole BM. c, Quantification of tdTomato positive colonies relative to the total number of adherent colonies. d, tSNE dimension reduction plot colored by cluster of sorted BM MPs generated from single cell RNA-seq expression data. e, The same tSNE from d colored by expression of the supercluster markers Pdgfra (blue) and Rgs5 (red). F, tSNE heatmaps of gene markers associated with the various mesenchymal cell types in the BM.
3.2.10 MP hyperplasia leads to expansion of the hematopoietic stem cell niche

As discussed above, conditional deletion of Hic1 in adult mice leads to MP hyperplasia in various tissues including the BM (Fig 3.21), and to determine if these “additional” MPs retained colony-forming activity, CFU-F assays were conducted. Cultures established from the BM of Hic1\(^{ff}\) animals formed an average of 43.7 ± 0.7 colonies whereas hemizygous UBC-CreERT2; Hic1\(^{ff}\) cultures yielded 90.9 ± 0.1 colonies per million whole BM cells seeded (Fig 3.31a). To evaluate the impact of increased BM-MPs on the hematopoietic compartment, various aspects of HSC activity were assessed. Using similar conditions as described above, conditional deletion of Hic1 was achieved by treating eight-week control Hic1\(^{ff}\) and UBC-CreERT2; Hic1\(^{ff}\) animals with TAM and BM was collected for analysis 1-year post-treatment. To confirm the increase in MPs in the BM at this time, LacZ reporter signal was also quantified. The X-gal stained area detected above background in control Hic1\(^{ff}\) marrow was confirmed to be increased 13 ± 1 fold compared to UBC-CreERT2; Hic1\(^{ff}\) marrow (Fig 3.31b). The increase in LacZ signal is congruent with the observed increase in the colony forming activity of the UBC-CreERT2\(^+\) BM.

To determine the impact of additional BM-MPs on the HSC compartment we analyzed wildtype and mutant whole BM for HSC content by flow cytometry using the SLAM marker (CD150\(^+\) CD244\(^-\) CD48\(^-\)) combination. Long term HSCs were found at a frequency of 0.0138 ± 0.0012% in marrow isolated from control Hic1\(^{ff}\) animals while the frequency was increased by 1.75-fold to 0.0241 ± 0.0012% in BM-MP enriched from the UBC-CreERT2; Hic1\(^{ff}\) marrow (Fig. 3.31c). In accordance with this observation, when whole BM was cultured in a methocellulose assay, a significant increase in CFU was observed in UBC-CreERT2; Hic1\(^{ff}\) compared to Hic1\(^{ff}\) control marrow (Fig 3.31d). The immunophenotypic evidence supported an
Figure 3.31 | MP numbers in the bone marrow increase as a result of Hic1 loss and induce a consequent increase in HSCs.

a, Colony Forming Unit – Fibroblast assay comparing adherent colony forming potential of Hic1 CKO vs control BM harvested 1-year after TAM-induced conditional deletion in 8-week-old mice. Representative wells are shown of cresyl violet stained colonies used for enumeration.

b, Quantification of LacZ reporter signal detected in BM from wildtype, heterozygous and Hic1 null mice.

c, Phenotypic analysis of LT-HSC percentage by flow cytometry using the SLAM surface marker combination.

d, Quantification of the in vitro colony forming potential in the HSC compartment.
increase in HSC numbers, however, the properties of these additional putative HSCs required further interrogation. Since the potential of HSCs cannot be measured reliably by immunophenotypic analyses alone, functional analyses were employed to determine the potential of the additional HSCs generated in the Hic1-deleted samples. To accomplish this, a competitive repopulation assay was carried out using cryopreserved BM from Hic1f/f control and UBC-CreERT2; Hic1f/f animals. These marrow samples were mixed at 50% ± 2% with marrow from animals ubiquitously-expressing eGFP and transplanted into lethally irradiated Ly5.1 recipient mice. Marrow was harvested from recipients six months after transplant and the Gr1+ fraction was analyzed by flow cytometry to determine the ratio of GFP- competing test cells to GFP+ donor cells. In the Gr1+ fraction from recipients transplanted with Hic1f/f/GFP- control test cells, a 51:49 ± 5.96 ratio was observed (Fig 3.32a). In contrast, the ratio of UBC-CreERT2; Hic1f/f/GFP- test cells compared to GFP+ competing cells was ~81:19 ± 3.5 (Fig 3.32a). These results indicate that additional BM-MPs are associated with increased repopulating activity in the BM compartment.

To resolve if the increase in BM-MPs impacts HSC niche capacity and/or intrinsic HSC repopulating activity, serial transplantation assays were employed. This approach will also further support that LT-HSCs are responsible for the repopulation activity observed in the lethally irradiated primary transplant recipients. For these purposes, marrow was isolated from primary recipients six months following transplantation and serially transplanted to secondary lethally irradiated Ly5.1 recipients in a non-competitive assay. After four months the Gr1+ fraction was analyzed by flow cytometry. The resultant ratios remained similar to that observed for the primary recipients at 50:50 ± 1.59 and 85:15 ± 2.94 for control Hic1f/f and UBC-CreERT2; Hic1f/f GFP- test cells, respectively, compared to GFP+ cells (Fig 3.32 b). These
Figure 3.32 | Expansion of Hic1\(^+\) MPs in bone marrow increases functional HSC frequency.

a, Comparison of repopulation potential by competitive transplant assay of BM from UBC-CreERT2;Hic1\(^{fl}\) vs Hic1\(^{fl}\) control mice 6 months after transplant.  
b, Results of a subsequent secondary transplant assay 4 months after the secondary transplant.  
c, BM was harvested from mice 12 months after conditional deletion of Hic1 at 8 weeks and a limiting dilution assay was performed to determine the increase in functional HSCs.
results indicate that the competitive advantage of the *Hic1*-deleted marrow is abrogated upon exposure to the new secondary recipient niche microenvironment. This further suggests that the observed increased potential of the HSC population in the *Hic1*-deleted BM is likely a consequence of increased niche capacity and not any property(s) intrinsic to the HSCs.

If the increase in BM-MPs leads to increased HSC niche capacity, this should be reflected in an increase in functional HSCs. Competitive transplant assays measure HSC activity at the population level and aid determination of overall fitness. To assess HSC potential at the individual cell level and to more accurately enumerate functional HSCs in the BM-MP-expanded animals, a limiting dilution assay was performed. The frequency of reconstituting cells in *Hic1*+/− marrow was estimated to be 1/19,430 while in marrow derived from *UBC-CreERT2; Hic1*+/− a significantly higher (p=0.0361) frequency of 1/9,251 was observed (Fig 3.32c). These findings indicate that the expansion of Hic1+ BM-MPs culminates in an increase in HSCs with functional reconstitution potential, likely reflecting a modified niche environment involving enhanced physical and/or trophic support.
Chapter 4: Conclusion

4.1 Overview - Hic1 and MP biology

MPs are emerging as central players in many fundamental biological programs. In all tissues and organs, MPs are a universal component of the stroma and reside within the perivascular space of small and large blood vessels. Reports of MPs regulating microenvironment composition and function in various tissues in homeostasis and during regeneration or repair are rapidly accumulating\textsuperscript{24,43,92,114,180,181}. This is congruent with current attempts to target and use MPs to treat a broad spectrum of diseases. Indeed, hundreds of clinical trials are underway attempting to harness the therapeutic potential of MPs. To date, limited efficacy has been demonstrated and this is likely due to the limited understanding of MP biology. As the number of reports of stromal cells participating in health and disease increases it is becoming clear that a single potentially universal marker for MPs would enable studies aimed at elucidating MP fate and function. Herein, we have identified \textit{Hic1} as a genetic marker of murine MPs across all organs examined and begun an in-depth characterization of their transcriptomes across multiple organs at homeostasis. These analyses have also been extended to better understand the dynamic behavior of MPs following injury-induced activation in a skeletal muscle regeneration model. We further demonstrated that \textit{Hic1}-based transgenic tools serve as useful reagents for the identification, isolation and manipulation of MPs revealing yet to be recognized fates and functions.
4.2 Case for Hypermethylated in Cancer 1 representing a useful MP marker for identifying MPs in skeletal muscle

Since no solitary universal marker was available to identify and isolate adult MPs, we sought to use a whole transcriptome analysis of sorted cell populations from skeletal muscle to identify potential genes that were enriched in the MP fraction versus other fractions. To optimize conditions for efficient enrichment of MPs, a reiterative strategy was initially employed involving FACS followed by RNA-seq. Once optimized, in-depth RNA seq was performed and the dataset was interrogated for genes specifically enriched in the MP fraction with an emphasis on transcription factors, signaling molecules and potential surface markers. Interrogation of these datasets revealed that Hic1 co-clustered tightly with Gli1, an existing marker associated with the mesenchymal stromal/progenitor cells in multiple tissues. In many tissues, resident MPs have been found within the perivascular space. Consistent with this, anti-HIC1 staining of skeletal muscle sections showed HIC1 was expressed in a non-endothelial cell type that was located in the perivascular compartment. Further evidence to support a perivascular location of Hic1 cells in skeletal muscle derived from analyses of two novel transgenic lines including Hic1\textsuperscript{nLacZ} and Hic1; reporter mice where 96.9% ± 0.8% and 96.2% ± 0.5% of nLacZ\textsuperscript{+} and tdTomato\textsuperscript{+} cells, respectively, were found on blood vessels. Collectively, these findings show that Hic1 identifies a blood vessel-associated cell type consistent with the proposed location of adult mesenchymal stromal cells\textsuperscript{20,182}. Furthermore, PDGFRA and Sca-1 expression marks skeletal muscle FAPs, and FAPs are Hic1; reporter gene positive by flow cytometry. Consistent with the Joe et al. report, isolated Hic1; reporter cells exhibited similar colony forming activity when cultured \textit{in vitro}\textsuperscript{18}. The most compelling data in support of an MP identity for Hic1\textsuperscript{+} cells
is derived from the popRNA-seq and scRNA-seq datasets. In both cases, transcriptomes of Hic1+ reporter positive cells exhibit an expression signature highly correlated with previous cell types identified in the literature as mesenchymal stromal cells, mesenchymal stem cells, medicinal signaling cells, skeletal stem cells and connective tissue stem cells, etc. Interestingly, reciprocal analysis of the complete Hic1+ MP transcriptomic datasets revealed a small subset of cells that were Ly6a and Pdgfra negative and this was confirmed by surface staining of reporter positive cells. These new and unexpected subsets include progenitors with a pericyte and tenogenic/myotenogenic expression signature. Furthermore, after injury, similar to the FAP population, these novel sub-populations retained their tenogenic/pericytic markers, but up-regulated genes associated with cell cycle entry and progression. These activities are also indicative of the MP phenotype. Consistent with progenitor activity, Hic1+ cells exhibit a hierarchical organization, expand following activation, and contribute to either transient or enduring mesenchymal lineages.

Collectively, these findings using a multi-pronged approach reinforce the concept that Hic1 represents a new and formidable marker for identifying skeletal muscle resident MPs.

4.3 Fate mapping of the fibrogenic/adipogenic MP subset following skeletal muscle damage

FAPs have previously been characterized in at least skeletal muscle, skin, adipose tissue (ADSC), cortical bone and marrow (MSC), and have been characterized as being PDGFRA+ and Sca-1+. Hic1CreERT2 allowed us to fate map FAPs for the first time and isolate single cells by FACS. This further demonstrated that Hic1 expression was highly enriched in the fibro/adipo
subset of MPs. At this stage FAPs could be divided into two main clusters based on their ECM secretome likely indicating there is some specialization between tissue-resident FAPs associated with different structures in the muscle. Lineage tracing combined with pop and scRNA-seq after NTX-induced skeletal muscle injury revealed a robust and distinct series of temporally choreographed contributions to the regenerative program. This was associated with multiple distinct dynamic programs involved in various activities, some of which included immune cell recruitment, and matrix production and remodeling. This implicates MPs as immunomodulatory pro-regenerative participants in skeletal muscle regeneration, as has been previously suggested in other contexts, strengthening their reputation as therapeutic entities or targets. All of these activities are required on a temporary basis and this is consistent with the observation that Hic1; reporter positive cells for the most part return to pre-injury levels. Consistent with this, previous studies have shown that the expanded FAP population appears mostly transient in nature and they are cleared through immune cell-mediated apoptosis.81

After MP activation and associated cell cycle entry, it remains to be determined whether all MPs exhibit similar fate potentials and specifically how and when quiescent MPs re-appear. This could involve symmetrical or asymmetric division events that ultimately lead to reconstitution of the MP quiescence compartment following activation. Intuitively the underlying mechanisms would also serve to maintain DNA fidelity by limiting the rounds of division in cells destined to remain quiescent and provide high fidelity progenitors as required. It is established that terminally differentiated hepatocytes can enter the cell cycle to regenerate liver parenchyma. Recently however it has been demonstrated that certain hepatocytes have a higher propensity for proliferation based on their proximity to proposed niches and related detection of downstream WNT signaling and TERT expression92,93. Further experimentation
needs to be conducted to understand the hierarchy of tissue resident MPs and their progeny, and the mechanisms driving reacquisition of a quiescent phenotype.

4.4 *Hic1* identifies tenogenic, myotenogenic and mural/pericyte MP subsets

The transcriptomic analyses at the population and single cell levels also revealed the existence of two other MP subsets not captured by the PαS marker combination, tenocyte/myotenocyte and pericyte progenitors. This provided an explanation for the detection of *Hic1* in the Sca-1 negative fraction since unlike FAPs, these other two main populations did not express *Ly6a* and have thus far remained somewhat elusive. The former is a novel population of cells that express the tenogenic transcription factor *Scleraxis* (*Scx*) and are observed to increase in numbers after skeletal muscle regeneration, contributing to the remodeled tendon. Unexpectedly, within this tenogenic compartment a specialized tenocyte subset was also identified that expresses the unique *Col22a1* transcript, indicative of the MTJ. Our understanding of adult myotendinous junction turnover at homeostasis and during regeneration is limited. These findings provide a foundation and a strategy for careful interrogation of the myotenocyte life cycle. In this regard, it will be critical to further analyze cell fate decisions in these populations during self-renewal and differentiation. The ability to control these myotenocyte progenitors could serve to enhance engraftment during connective tissue transplant or other regenerative approaches. The latter, a pericyte progenitor subpopulation, observed to enter the cell cycle, was also identified however it remains unclear to the extent of contribution to fibrogenic cells and this is being explored in more detail in a separate concurrent study.
Collectively, these data reveal the existence of novel MP subsets in skeletal muscle with unappreciated lineage potential that share in common *Hic1* expression.

### 4.5 The mouse adult mesenchymal progenitor atlas

To further explore the utility of *Hic1* as a universal MP marker, scRNA-seq analyses were carried out on enriched MPs from a variety of tissues. With the exception of LSECs (discussed below), tdTomato$^+$ cells from the various analyzed tissues expressed a signature analogous to that observed in skeletal muscle and overlapped with expression profiles reported for mesenchymal cells in the literature. A somewhat unintended consequence of the characterization of MPs from many murine organs was the ability to generate a substantial foundation toward a comprehensive atlas of MPs. This aggregated dataset allows for interrogation of systemic MPs as a whole providing a resource for comparisons between tissue resident MPs and systemic subpopulations of MPs. Specific markers for the various subsets should be further explored and validated in more detail as an extension of these findings. In combination with the recently published whole mouse atlas, our MP atlas can be used to define markers and pathways exclusive to, and absent from, MPs$^{183}$. In general, it is expected that a variety of useful MP-related markers will be identified to enable studies directed at better understanding tissue level MP specialization.

*Hic1/HIC1* has been reported by other groups to be expressed in other non-MP cell types, including inner granular and Purkinje neurons in the cerebellum$^{184,185}$, basal epithelial cells in the mammary gland$^{186}$, a limited number of immune cell types (discussed below) and in various cell lines$^{187,188}$. From our analysis of both *Hic1* knock-in alleles along with IF staining, we have not
detected appreciable \textit{Hic1} expression in these compartments. In the mammary gland, \textit{Hic1} is
detected to the interstitial stromal and in the brain, \textit{Hic1} identifies perivascular stromal cell types
and is excluded from all neurons and glia. We have not observed \textit{Hic1} expression in any
epithelial cell type nor endoderm-derived cell types. Interestingly, for the most part \textit{Hic1} also
appears to be excluded from endothelial cells, with the one exception being the LSECs. The
LSECs are unlike most other endothelial cells in that they are devoid of a basement membrane
and fenestrated in nature. Due to the absence of a basement membrane, the HSCs are more
closely associated with the endothelium. Thus, the presence of a prominent tdTomato$^+$ signal in
LSECs was thought to be an artifact and potentially the result of material transfer between HSCs
and LSECs. Consistent with this possibility, both cell types exhibit strong endocytic activities.
However, scRNA-seq analyses revealed comparable \textit{Hic1} expression in the LSECs compared to
other tissues. Furthermore, the presence of other typically mesenchymally-restricted transcripts
such as \textit{Gli1, Vim, Pth1r, Vcam1,} and \textit{Lepr} in the LSEC population suggests that these cells may
have an alternate origin and/or phenotype compared to other endothelial cells. The relationship
between LSECs and HSCs needs further investigation, in addition to the role of \textit{Hic1} in these
populations. Future studies would also determine if \textit{Hic1} serves as a marker of mature and/or
LSEC progenitors and these tools could be used to explore LSEC biology.

\textbf{4.6 \ Conditional deletion of \textit{Hic1} in adult mice leads to MP hyperplasia}

Herein \textit{Hic1} is described as a marker of quiescent tissue resident MPs and this was
evident when adult animals were analyzed after conditional deletion of \textit{Hic1}. It has been
reported that \textit{Hic1} is a tumour suppressor and MP hyperplasia was observed in all tissues
examined to varying degrees. The fact that despite loss of Hic1, “null” cells do not continue to proliferate indefinitely *in vivo* suggests there are other mechanisms operating to regulate growth arrest. Additionally, this data implies a mechanism(s) controlling MP cell number and reestablishment of “homeostasis” in knockout animals. More detailed analysis should be undertaken to yield insights into the MP phenotype, and this may help identify compensatory mechanisms that are invoked to regulate cell number and additional cell cycle entry. In the absence of Hic1, the MP phenotype appears to be only marginally impacted, as there is a limited effect on skeletal muscle regeneration in this background in comparison to WT controls.

Unfortunately, Hic1 is not expressed once MPs differentiate or transition to a senescent state and therefore these Hic1+ progeny can no longer be effectively traced with the nLacZ knock-in allele. Preliminary analyses have revealed the presence of ectopic adipocytes in numerous tissues following Hic1 deletion and this may in part reflect dysregulated MPs. However, whether these adipocytes represent Hic1+ progeny remains to be conclusively demonstrated by lineage tracing.

Consistent with the role of Hic1 in quiescence, it was observed that after skeletal muscle injury induced activation of MPs, Hic1 expression was transiently decreased as measured by RNA-seq and LacZ reporter expression. These observations suggest that Hic1 expression is reestablished as “select” MPs are fated to return to a quiescence program or Hic1 re-expression determines which cells will become quiescent.

It remains unclear whether a quiescent state is achieved in Hic1 null cells after an activation cycle. Further studies should be carried out to determine whether or not Hic1 null MPs are able to properly re-enter a quiescent state thereby retaining the capacity to undergo additional cycles of activation and quiescence. Other collaborative studies described in the next section indicate an increase in hematopoietic niche capacity in the BM as a result of MP
hyperplasia. It is likely that this pathway can be exploited to prime MPs directly and also to potentially modify niche activity to support stem cell mobilization, such as observed in the HSC niche. However, the short and long-term consequences of such changes in MP status remains to be determined, especially in the context of stem/progenitor cell exhaustion.

4.7 MP hyperplasia in the bone marrow enhances the “hematopoietic stem cell niche”

Like in all other tissues examined, the Hic1LacZ$^+$ MP population in the BM undergoes an expansion after conditional deletion of Hic1 in adult mice. BM stromal cells classically identified as MSCs have recently been distinguished from the skeletal stem cells (SSC) using a variety of markers and cell fate potential$^{10,14}$. In accordance with these discoveries, a more detailed transcriptomic and histological analysis of MPs genetically labeled by Hic1 in the BM reveal intra-tissue heterogeneity some of which parallels that observed in other tissues. When categorized by single cell gene expression profile, MP clusters include FAP and PP populations. In addition, the marrow contains somewhat unique subpopulations. One population exhibited an osteogenic profile, and likely represents a rare tdTomato$^+$ subset observed on the endosteal surfaces. The second is a subtype that aligns with the current accepted description of a BM stromal cell or Cxcl12$^{\text{high}}$ reticular cell (CAR). The abundant expression of Cxcl12 or stromal cell-derived factor 1 (SDF1) is the main source of the required hematopoietic stem cell (HSC) maintenance signal, without which HSCs are lost. The MP subpopulation we have identified expresses high levels of Cxcl12 as well as Kitl, another key HSC factor, and Lepr, the definitive marker of marrow stromal cells that contribute to the HSC niche. TdTomato positive MPs contain the majority of the colony-forming unit - fibroblast (CFU-F) activity in the BM.
Collectively, these observations indicate that BM CAR cells are identified by Hic1. Consistent with this premise, following conditional deletion of Hic1, the increase of LacZ+ MPs in the BM exhibit an increase in colony forming potential. The increase in CFU-F potential, using the established assay of MSC potency, suggests an increase in bona fide marrow stromal cells and this includes CAR MPs.

As mentioned in an earlier section, the potential ramifications of generating an increased number of MSCs/MPs are extensive. As these cells are found within stem cell niches and it is thought that the number of niches within the BM greatly influences HSC abundance. The increase in Hic1+ cells in the BM led us to directly test if this generated additional HSC niches that could support functional HSCs. To test this, HSC frequency and repopulation capacity was assayed. First, a phenotypic analysis of SLAM markers by flow cytometry indicated a two-fold increase in LT-HSCs. To further evaluate this, a competitive repopulation assay was performed which revealed that the Gr1+ BM fraction derived from recipient animals four months after transplant with WT test marrow contained a similar ratio of competing cells. In contrast, recipients of Hic1−/− BM were comprised of over eighty percent derivatives of the test population. This suggests a competitive advantage is conferred to the BM of Hic1−/− animals at least with respect to HSC regenerative capacity. To rule out an effect on the stem cells themselves and provide evidence to support the increase of LT-HSCs, these data were further validated by performing a secondary competitive transplant assay. No difference in the ratio of test marrow to competing marrow was observed indicating the increase in regenerative capacity did not change over time nor was a consequence of exposure to the recipient marrow microenvironment. Collectively, this also suggested but was not directly tested, that the recipients were able to provide adequate niche support for engraftment of the additional HSCs from both transplants.
To further address if the increase in HSCs was intrinsic to this population, the frequency of HSCs was quantified using a functional assay. BM was isolated from wild type or Hic1 null animals and transplanted to wildtype recipients and blood system reconstitution was evaluated using a limiting dilution assay. Under these conditions, similar to what was observed with the increase in the marrow MPs, a two-fold increase in HSC numbers was detected. Taken together, these results suggest that the increase in niche stromal cells in the Hic1−/− BM confers an opportunity to indirectly expand the HSC population. In this regard, the expansion of MPs may lead to non-cell autonomous effects on the BM HSC niche, most likely by providing more physical cellular contact and/or trophic support of HSCs. As a result, the steady state fraction of HSCs increases proportionally. Further extrapolation of these findings suggests that CAR cells alone control niche capacity and HSC numbers. However, this does not preclude the possibility that a similar increase in other niche cell types would not also seed an increase in the number of niches. The possibility also remains that a Lepr low/negative marrow MP subpopulation is responsible for these observations. However, that seems unlikely based on previous genetic ablation experiments using Lepr-Cre demonstrating the requirement of factors produced by Lepr+ cells and endothelial cells, the latter of which has not been observed to express Hic1 in the BM51.

Analysis of the BM from UBC-CreERT2;Hic1−/− animals robustly shows that the observed increase in stromal MPs leads to increased numbers of HSCs though the exact mechanism remains to be determined. It is possible that the quiescence/activation axis in BM MPs is modified after Hic1 KO and the full extent of the participation of activated MPs in the niche is unknown. Therefore, it is not entirely clear whether the proposed increased niche
capacity is a result of more niches or a result of a more activated MP phenotype creating niche microenvironments conducive to increased HSC support.

Ongoing studies will determine whether MPs in the spleen will be comparable to MPs in the marrow and include the proposed Lepr⁺ spleen perisinusoidal niche cells. At the time of writing, single cell profiling reveals a Tcf21⁺, Lepr⁺ CAR cell among Hic1CreERT2;tdTomato⁺ MPs sorted from spleen. The role of this population in extramedullary splenic hematopoiesis remains to be determined.

The widespread distribution of MPs within tissue and organ stroma, and the proposed stromal support role of BM MSCs in the hematopoietic stem cell niche intuitively suggests a role for MPs in various other stem/progenitor cell niches. Within these compartments, it is expected that different stem cell niches will involve the production of overlapping but also different trophic factors. This is reflected in the data where for example, a unique Rspo3 positive subpopulation was identified in brain MPs. This population expresses genes previously associated with the choroid plexus epithelium and thus may play a role in supporting ependymal cells in this region. Future experiments will test these and other individual hypotheses using the MP atlas to guide functional evaluation and contribution. Conditional loss of function of selected systemic and tissue specific factors have the potential to reveal the non-cell autonomous contribution of MPs to the renewal of vascular and parenchymal organ components, especially those that are known to undergo extensive remodeling. Furthermore, these genetic mouse lines will also serve as excellent models to explore MP-dependent contributions in various regenerative and disease models.

Work by Burrows et al identified a subpopulation of T lymphocytes specific to the small intestine that express Hic1. They observed that following Hic1 loss these cells were no
longer detected in the small intestine. This population has also been reported to be depleted after mice are fed a vitamin A deficient diet. We are further evaluating the potential expression of \textit{Hic1} in these relatively rare immune cell subsets. Vitamin A plays a fundamental role in regulating gastrointestinal immune cell homing and activity\textsuperscript{191-195}, and thus it is possible that \textit{Hic1}-related genetic programs may be operating in these populations to regulate their quiescence which may be reflected as memory in these lymphocytes. The loss of the type three innate lymphoid cells (ILC3) population in the small intestine following conditional deletion of \textit{Hic1} may also reflect that the gut homing ability of these specialized lymphocytes may be disrupted\textsuperscript{190}. As part of the HSC niche experiments, the cell autonomous action of \textit{Hic1} in hematopoiesis was tested by conditionally deleting \textit{Hic1} using VavCre\textsuperscript{196}. No competitive advantage was observed when BM from these animals was tested in competitive transplant assays against wildtype marrow. The expression and function of \textit{Hic1} in select immune cell subtypes requires further investigation but does not detract from the overarching role for \textit{Hic1} in quiescent MPs as described herein.

\section*{4.8 Final remarks}

Several key contributions emerge from this thesis that impact our understanding of MP biology in biomedical research and regenerative medicine. First, the identification and validation of a systemic marker for murine and possibly human mesenchymal stem/progenitor cells. This allows for in-depth analysis of murine MP biology as a systemic entity and provides proof-of-concept that these key tools can provide insights to previously untenable questions. Moreover, use of lineage tracing and transcriptional profiling of MPs after injury-induced activation,
revealed the dynamic role for MPs in coordinating multiple facets of the regenerative program. These analyses also identified unique sub-populations that have yielded fundamental insights into regeneration of structures, such as the MTJ. In aggregate, this body of work supports the potential use of Hic1-based genetic tools to explore the fate and function of MPs in a multitude of physiological and pathological programs.

Since Hic1 is a direct retinoic acid target gene, this implicates the retinoid signaling pathway operating in part through Hic1 as key to the regulation of systemic MP quiescence. This likely exists as part of a larger group of RA regulated programs involving additional RA-regulated genes that participate in regulation of the MP quiescence state, subsequent activation and return to quiescence\textsuperscript{197}. This linkage further reinforces the potentially important role for RA in regulating tissue homeostasis, renewal and regeneration and this may in part by accomplished by direct regulation of the MP phenotype.

The collection of single cell RNA-seq libraries that were generated from MPs across tissues provides an excellent resource to enable MP comparisons within and between tissues. It will also be interesting to use this approach to extend our analysis of the MP activated transcriptome under different injury conditions in distinct tissues. Herein, injury of skeletal muscle and pancreas leads to a clear activation phenotype associated with MP expansion. However, it remains to be determined to what extent MP activities are shared and whether there are distinct tissue-specific invariant programs between these and other compartments. Taken together, these studies will further our understanding of MP biology in tissue renewal and regeneration. Furthermore, this will provide the foundation to initiate studies aimed at genetically ablating or manipulating MP behavior to better understand the fate and function of MPs in health and disease. In this regard, it is expected that a better understanding of MP
biology will provide the framework for the rational use of MPs in various therapeutic strategies that may involve delivery of MPs or their \textit{in situ} modification (i.e. pharmaceutical or biological agents). There is also a strong link with MP/MSC dysfunction and aging, especially in the progeroid syndromes and thus it will be interesting to test whether rejuvenation or replacement of MPs could impact aspects of chronological aging. Ultimately, one could contemplate the treatment of the systemic diseases associated with aging by targeting the MP system to delay degenerative processes and/or extend life or health span.

Taken together this work uses the murine homologue to a putative human tumour suppressor, the repressive transcription factor \textit{Hic1} to unify a systemic population of MPs and provide a foundation for gaining insights into diverse aspects of MP biology and the potential therapeutic manipulation thereof. Furthermore, the transgenic tools described herein provide an opportunity to investigate the fate and function of this population in embryonic development, health and disease.
References


Mederacke, I., Hsu, C. C., Troeger, J. S., Huebener, P., Mu, X., Dapito, D. H., Pradere, J. P. & Schwabe, R. F. Fate tracing reveals hepatic stellate cells as dominant contributors to


44 Kramann, R., Schneider, R. K., DiRocco, D. P., Machado, F., Fleig, S., Bondzie, P. A., Henderson, J. M., Ebert, B. L. & Humphreys, B. D. Perivascular Gli1+ progenitors are


Greicius, G., Kabiri, Z., Sigmundsson, K., Liang, C., Bunte, R., Singh, M. K. & Virshup, D. M. PDGFRAlpha(+) pericryptal stromal cells are the critical source of Wnts and


Wu, H., Uchimura, K., Donnelly, E. L., Kirita, Y., Morris, S. A. & Humphreys, B. D. Comparative Analysis and Refinement of Human PSC-Derived Kidney Organoid


Appendices

Appendix A - Example R code used to analyze MP datasets and generate figures for the muscle regeneration study.

```
library(cellrangerRkit)
library(RSvgDevice)

gbm_pheatmap(log_gene_bc_matrix(gbmd0), prioritized_genes_Km4, cells_to_plot,
            n_genes=10, colour=S11colors_paper, limits=c(-3,3))

cclegenelist <- c("Ptp4a1","Txnip","Cnd3","Cdkn1a","Cdkn1b","Cenh","Cnd1","Cdk2","Cnd2","Mki67","Cdkn3","Ccne1","Ccne2","Cdkn2b","Cdkn2a","Cdkn1c")

load("Desktop/ws20170725_NTXd124ANALYSIS_VARIABLE_BACKUP/analysis_resultsd1.Rdata")
load("Desktop/ws20170725_NTXd124ANALYSIS_VARIABLE_BACKUP/analysis_resultsd2.Rdata")
load("Desktop/ws20170725_NTXd124ANALYSIS_VARIABLE_BACKUP/analysis_resultsd4.Rdata")

#### LOG 10
visualize_gene_markers(gbm_logd0, genesTSNEHM_paper[1],
                       tsne_projd0[c("TSNE.1","TSNE.2")], limits=c(0,0.5), marker_size = 2)
visualize_gene_markers(gbm_logd1, genesTSNEHM_paper[1],
                       tsne_projd1[c("TSNE.1","TSNE.2")], limits=c(0,2), marker_size = 2)
visualize_gene_markers(gbm_logd2, genesTSNEHM_paper[1],
                       tsne_projd2[c("TSNE.1","TSNE.2")], limits=c(0,3), marker_size = 2)
visualize_gene_markers(gbm_logd4, genesTSNEHM_paper[1],
                       tsne_projd4[c("TSNE.1","TSNE.2")], limits=c(0,3), marker_size = 2)

#### LOG 2
visualize_gene_markers(gbm_log2d0, genesTSNEHM_paper[1],
                       tsne_projd0[c("TSNE.1","TSNE.2")], limits=c(0,2.25), marker_size = 2)
dev.copy(pdf, "crTSNE_HM_Pdgfra_d0_log2exp.pdf")
dev.off()
visualize_gene_markers(gbm_log2d1, genesTSNEHM_paper[1],
                       tsne_projd1[c("TSNE.1","TSNE.2")], limits=c(0,2.25), marker_size = 2)
dev.copy(pdf, "crTSNE_HM_Pdgfra_d1_log2exp.pdf")
dev.off()
visualize_gene_markers(gbm_log2d2, genesTSNEHM_paper[1],
                       tsne_projd2[c("TSNE.1","TSNE.2")], limits=c(0,2.25), marker_size = 2)
dev.copy(pdf, "crTSNE_HM_Pdgfra_d2_log2exp.pdf")
dev.off()
```

141
str(QQQQ_14)
wsvlndropcols <- c("Cluster")
QQQQ1_14 <- subset(QQQQ_14, Cluster == 1)
str(QQQQ1_14)
QQQQ1a_14 <- QQQQ1_14[, !(names(QQQQ1_14) %in% wsvlndropcols)]
str(QQQQ1a_14)
QQQQ1b_14 <- melt(QQQQ1a_14)
QQQQ2_14 <- subset(QQQQ_14, Cluster == 2)
str(QQQQ2_14)
QQQQ2a_14 <- QQQQ2_14[, !(names(QQQQ2_14) %in% wsvlndropcols)]
str(QQQQ2a_14)
QQQQ2b_14 <- melt(QQQQ2a_14)
QQQQ3_14 <- subset(QQQQ_14, Cluster == 3)
str(QQQQ3_14)
QQQQ3a_14 <- QQQQ3_14[, !(names(QQQQ3_14) %in% wsvlndropcols)]
str(QQQQ3a_14)
QQQQ3b_14 <- melt(QQQQ3a_14)
QQQQ4_14 <- subset(QQQQ_14, Cluster == 4)
str(QQQQ4_14)
QQQQ4a_14 <- QQQQ4_14[, !(names(QQQQ4_14) %in% wsvlndropcols)]
str(QQQQ4a_14)
QQQQ4b_14 <- melt(QQQQ4a_14)
tst50 <- ggplot(QQQQ1b_14, aes(x=variable, y=value)) + geom_violin(trim=F, fill = km4col1, colour="black", scale = "count", adjust = 3) + scale_y_log10() + theme(axis.text=element_text(size=5))
tst51 <- ggplot(QQQQ2b_14, aes(x=variable, y=value)) + geom_violin(trim=F, fill = km4col2, colour="black", scale = "count", adjust = 3) + scale_y_log10() + theme(axis.text=element_text(size=5))
tst52 <- ggplot(QQQQ3b_14, aes(x=variable, y=value)) + geom_violin(trim=F, fill = km4col3, colour="black", scale = "count", adjust = 3) + scale_y_log10() + theme(axis.text=element_text(size=5))
tst53 <- ggplot(QQQQ4b_14, aes(x=variable, y=value)) + geom_violin(trim=F, fill = km4col4, colour="black", scale = "count", adjust = 3) + scale_y_log10() + theme(axis.text=element_text(size=5))
tst50 + ylim(-1000, 50000)
tst51 + ylim(-1000, 50000)
tst52 + ylim(-1000, 50000)
tst53 + ylim(-1000, 50000)
tst50 + geom_jitter(shape=1, position=position_jitter(0.2), size = 0.01)
tst51 + geom_jitter(shape=1, position=position_jitter(0.2), size = 0.01)
tst52 + geom_jitter(shape=1, position=position_jitter(0.2), size = 0.01)
tst53 + geom_jitter(shape=1, position=position_jitter(0.2), size = 0.01)
tst54 <- ggplot(QQQQ1b_14, aes(x=variable, y=value)) + geom_violin(trim=T, fill = km4col1, colour="black", scale = "count", adjust = 5) + scale_y_log10() + theme(axis.text=element_text(size=5))

143
tst55 <- ggplot(QQQQ2b_14, aes(x=variable, y=value)) + geom_violin(trim=T, fill = km4col2, colour="black", scale = "count", adjust = 5) + scale_y_log10() + theme(axis.text = element_text(size=5))
tst56 <- ggplot(QQQQ3b_14, aes(x=variable, y=value)) + geom_violin(trim=T, fill = km4col3, colour="black", scale = "count", adjust = 5) + scale_y_log10() + theme(axis.text = element_text(size=5))
tst57 <- ggplot(QQQQ4b_14, aes(x=variable, y=value)) + geom_violin(trim=T, fill = km4col4, colour="black", scale = "count", adjust = 5) + scale_y_log10() + theme(axis.text = element_text(size=5))
tst54 + geom_jitter(shape=1, position=position_jitter(0.2), size = 0.01)
tst55 + geom_jitter(shape=1, position=position_jitter(0.2), size = 0.01)
tst56 + geom_jitter(shape=1, position=position_jitter(0.2), size = 0.01)
tst57 + geom_jitter(shape=1, position=position_jitter(0.2), size = 0.01)

### DAY 14

tst54
tst55
tst56
tst57

tst54 + geom_jitter(shape=1, position=position_jitter(0.2), size = 0.01)
tst55 + geom_jitter(shape=1, position=position_jitter(0.2), size = 0.01)
tst56 + geom_jitter(shape=1, position=position_jitter(0.2), size = 0.01)
tst57 + geom_jitter(shape=1, position=position_jitter(0.2), size = 0.01)

### DAY 0 VIOLINS

tst60 <- ggplot(QQQQ1b, aes(x=variable, y=value)) + geom_violin(trim=T, fill = km4col1, colour="black", scale = "count", adjust = 5) + scale_y_log10() + theme(axis.text = element_text(size=5))
tst61 <- ggplot(QQQQ2b, aes(x=variable, y=value)) + geom_violin(trim=T, fill = km4col2, colour="black", scale = "count", adjust = 5) + scale_y_log10() + theme(axis.text = element_text(size=5))

dev.copy(pdf, "crSC_violin_D14_Cluster_1.pdf")
dev.off()
dev.copy(pdf, "crSC_violin_D14_Cluster_2.pdf")
dev.off()
dev.copy(pdf, "crSC_violin_D14_Cluster_3.pdf")
dev.off()
dev.copy(pdf, "crSC_violin_D14_Cluster_4.pdf")
dev.off()
dev.copy(pdf, "crSC_violin_D14_Cluster_1_with_small_dots.pdf")
dev.off()
dev.copy(pdf, "crSC_violin_D14_Cluster_2_with_small_dots.pdf")
dev.off()
dev.copy(pdf, "crSC_violin_D14_Cluster_3_with_small_dots.pdf")
dev.off()
dev.copy(pdf, "crSC_violin_D14_Cluster_4_with_small_dots.pdf")
dev.off()

******************************************************************************

### DAY 0 VIOLINS

tst60 <- ggplot(QQQQ1b, aes(x=variable, y=value)) + geom_violin(trim=T, fill = km4col1, colour="black", scale = "count", adjust = 5) + scale_y_log10() + theme(axis.text = element_text(size=5))
tst61 <- ggplot(QQQQ2b, aes(x=variable, y=value)) + geom_violin(trim=T, fill = km4col2, colour="black", scale = "count", adjust = 5) + scale_y_log10() + theme(axis.text = element_text(size=5))
```r
tst62 <- ggplot(QQQQ3b, aes(x=variable, y=value)) + geom_violin(trim=T, fill=km4col3, colour="black", scale="count", adjust=5) + scale_y_log10() + theme(axis.text=element_text(size=5))
tst63 <- ggplot(QQQQ4b, aes(x=variable, y=value)) + geom_violin(trim=T, fill=km4col4, colour="black", scale="count", adjust=5) + scale_y_log10() + theme(axis.text=element_text(size=5))
tst60
tst61
tst62
tst63
tst60 + geom_jitter(shape=1, position=position_jitter(0.2), size=0.01)
tst61 + geom_jitter(shape=1, position=position_jitter(0.2), size=0.01)
tst62 + geom_jitter(shape=1, position=position_jitter(0.2), size=0.01)
tst63 + geom_jitter(shape=1, position=position_jitter(0.2), size=0.01)
dev.copy(pdf, "crSC_violin_D0_Cluster_1.pdf")
dev.off()
dev.copy(pdf, "crSC_violin_D0_Cluster_2.pdf")
dev.off()
dev.copy(pdf, "crSC_violin_D0_Cluster_3.pdf")
dev.off()
dev.copy(pdf, "crSC_violin_D0_Cluster_4.pdf")
dev.off()
dev.copy(pdf, "crSC_violin_D0_Cluster_1_with_small_dots.pdf")
dev.off()
dev.copy(pdf, "crSC_violin_D0_Cluster_2_with_small_dots.pdf")
dev.off()
dev.copy(pdf, "crSC_violin_D0_Cluster_3_with_small_dots.pdf")
dev.off()
dev.copy(pdf, "crSC_violin_D0_Cluster_4_with_small_dots.pdf")
dev.off()
```

### CYTOKINE HEATMAP


gbm_pheatmap(log_gene_bc_matrix(gbmntx), cytokinegenelist, cells_to_plot_ntx_libid_fixed, n_genes=15, colour=S11colors_04710, limits=c(-1.5,1.5))
dev.copy(pdf, "crSC_cytokine_heatmap_limits1pt5.pdf")
dev.copy(pdf, "crSC_cytokine_heatmap_limits1pt5_rowclustered.pdf")
dev.off()

```r
```
tst58 <- ggplot(QQQQb_14, aes(x=variable, y=value)) + geom_violin(trim=F, fill = km4col4, colour="black", scale = "count", adjust = 0.1) + scale_y_log10() + theme(axis.text=element_text(size=5))

***************************************************************************

### DAY 0124 kmeans 4 cluster heatmap

scp -r
wscott@brclogin1.brc.ubc.ca:/brcwork/underhill_lab/wscott/170515_NS500668_0199_AHV73C BGX2/AGGSC8dmg/outs/analysis/kmeans Desktop/outs/analysis/
scp -r
wscott@brclogin1.brc.ubc.ca:/brcwork/underhill_lab/wscott/170515_NS500668_0199_AHV73C BGX2/AGGSC8dmg/outs/analysis/diffexp Desktop/outs/analysis/
cpp_npx_dsktp <- "Desktop"
ar <- load_cellranger_analysis_results(cpp_npx_dsktp)
cluster_result_km4_npx <- ar$clustering$kmeans_kmeans_4_clusters$Cluster
prioritized_genes_npx_km4_npx <- prioritize_top_genes(gbmnpx, cluster_result_km4_npx, "sseq", min_mean=0.5)
cells_to_plot_npx_km4_npx <- order_cell_by_clusters(gbmnpx, cluster_result_km4_npx)

### BY graph cluster

cell_result_graphtest <- ar$clustering$graphclust$clusters
prioritized_genes_npx <- prioritize_top_genes(gbmnpx, cluster_result_graphtest, "sseq", min_mean=0.5)
cells_to_plot_npx <- order_cell_by_clusters(gbmnpx, cluster_result_graphtest)

***************************************************************************

cluster_result_km4_d0 <- analysis_resultsd0$kmeans$kmeans_4_clusters$Cluster
prioritized_genes_npx_km4_d0 <- prioritize_top_genes(gbmd0, cluster_result_km4_d0, "sseq", min_mean=0.5)
cells_to_plot_npx_km4_d0 <- order_cell_by_clusters(gbmd0, cluster_result_km4_d0)

gbm_pheatmap(log_gene_bc_matrix(gbmnpx), prioritized_genes_npx_km4_npx, cells_to_plot_npx_km4_npx, n_genes=20, colour=S11colors_04710, limits=c(-1.5,1.5))
devo.copy(pdf, "crSC_Heatmap_npx_km4.pdf")
devo.off()
gbm_pheatmap(log_gene_bc_matrix(gbmd2), prioritized_genes_ntx_km4_d2, cells_to_plot_ntx_km4_d2, n_genes=20, colour=S11colors_04710, limits=c(-1.5,1.5))
cluster_result_km4_d4 <- analysis_resultsd4$kmeans$kmeans_4_clusters$Cluster
prioritized_genes_ntx_km4_d4 <- prioritize_top_genes(gbmd4, cluster_result_km4_d4, "sseq", min_mean=0.5)
cells_to_plot_ntx_km4_d4 <- order_cell_by_clusters(gbmd4, cluster_result_km4_d4)
gbm_pheatmap(log_gene_bc_matrix(gbmd4), prioritized_genes_ntx_km4_d4, cells_to_plot_ntx_km4_d4, n_genes=20, colour=S11colors_04710, limits=c(-1.5,1.5))
gbm_pheatmap(log_gene_bc_matrix(gbmd1), cytokinegenelist, cells_to_plot_ntx_km4_d1, n_genes=20, colour=S11colors_04710, limits=c(-1.5,1.5))
gbm_pheatmap(log_gene_bc_matrix(gbmd2), cytokinegenelist, cells_to_plot_ntx_km4_d2, n_genes=20, colour=S11colors_04710, limits=c(-1.5,1.5))
gbm_pheatmap(log_gene_bc_matrix(gbmd4), cytokinegenelist, cells_to_plot_ntx_km4_d4, n_genes=20, colour=S11colors_04710, limits=c(-1.5,1.5))
gbm_pheatmap(log_gene_bc_matrix(gbmd1), cellcyclegenelist, cells_to_plot_ntx_km4_d1, n_genes=20, colour=S11colors_04710, limits=c(-1.5,1.5))
gbm_pheatmap(log_gene_bc_matrix(gbmd2), cellcyclegenelist, cells_to_plot_ntx_km4_d2, n_genes=20, colour=S11colors_04710, limits=c(-1.5,1.5))
gbm_pheatmap(log_gene_bc_matrix(gbmd4), cellcyclegenelist, cells_to_plot_ntx_km4_d4, n_genes=20, colour=S11colors_04710, limits=c(-1.5,1.5))
*********************************************************************
genesd14 <- c("Mkx","Kera","Fmod","Col12a1","Six2","Tnmd","Sex","Col11a1","Col22a1","Chad","Chodl"
)tsne_projd14 <- analysis_results_d14$tsne
gbm_log2d14 <- log_gene_bc_matrix(gbm_bcnormd14,base=2)
max(exprs(gbmd14[which(fData(gbm)$symbol == 'Mkx'),]))visualize_gene_markers(gbm_log2d14, genesd14[1], tsne_projd14[c("TSNE.1","TSNE.2")], limits=c(0,2), marker_size = 2)
dev.copy(pdf, "crTSNE_HM_Mkx_d14_log2exp.pdf")dev.off()
*********************************************************************
### FIGURE 5 TENOGENIC GENE COMPARISON tSNE Heatmaps
genesfig5teno <- c("Col22a1","Mkx","Sex","Tnmd","Hic1","tdTomato")max(exprs(gbmd0[which(fData(gbm)$symbol == 'Col22a1'),]))max(exprs(gbmd14[which(fData(gbm)$symbol == 'Col22a1'),]))visualize_gene_markers(gbm_log2d0,genesfig5teno[1],tsne_projd0[c("TSNE.1","TSNE.2")],limits=c(0,4), marker_size = 2)
dev.copy(pdf, "crTSNE_HM_figure5_Col22a1_D0_log2exp.pdf")dev.off()
dev.copy(pdf, "crTSNE_HM_figure5_Mkx_D0_log2exp.pdf")dev.off()
### REMAKE tSNEs AS SVG
visualize_gene_markers(gbm_log2d0,genesfig5teno[1],tsne_projd0[c("TSNE.1","TSNE.2")],limits=c(0,4), marker_size = 2)
**visualize_gene_markers(gbm_log2d14,genesfig5teno[6],tsne_projd14[c("TSNE.1","TSNE.2")],limits=c(0,5.5), marker_size = 2)**

ggsave(filename="crTSNE_HM_figure5_Col22a1_D0_log2exp.svg", device = devSVG, width = 8 , height = 8)

**visualize_gene_markers2(gbm_log2d0,genesfig5teno[1],tsne_projd0[c("TSNE.1","TSNE.2")],limits=c(0, 4), marker_size = 2)**

ggsave(filename="crTSNE_HM_figure5_Col22a1_D0_log2exp.svg", device = devSVG, width = 8 , height = 8)

**genesSUPP4 <- c("Pdgfrb","Mcam","Kcnj8","Tagln","Rgs5","Acta2","Cspg4","Notch3")**

### gbmd0

```r
print(max(exprs(gbmd0[which(fData(gbmd0)$symbol == 'Pdgfrb'),])))
```

### gbmd14

```r
print(max(exprs(gbmd14[which(fData(gbmd14)$symbol == 'Notch3'),])))
```

### exponent list for log2 (-1) to get limits

```r
8
7.5
7.25
8
4.5
1
6
```
visualize_gene_markers2(gbm_log2ntx2, genesSUPP3tsne[[1]],
  tsne_projntx2[, c("TSNE.1", "TSNE.2")], limits = c(0, 1.5), marker_size = 2)
ggsave(filename = "crTSNE_HM_Adam12_ntx2_log2exp.svg", device = devSVG, width = 8, height = 8)

******************************************************************************
genes_SUPP3tsne <-

{print(max(exprs(gbmntx2[which(fData(gbmntx2)$symbol == 'Adam12'),])))}

[1] 11  ###   2.5
tsneprojntx2 <- analysis_results_ntx2$tsne
use_genesntx2 <- get_nonzero_genes(gbmntx2)

gbm_bcnormntx2 <- normalize_barcode_sums_to_median(gbmntx2[use_genesntx2,])


gbm_log2ntx2 <- log_gene_bc_matrix(gbm_bcnormntx2, base = 2)

print(dim(gbm_log2ntx2))

visualize_gene_markers2(gbm_log2ntx2, genes_SUPP3tsne[[1]],
  tsne_projntx2[, c("TSNE.1", "TSNE.2")], limits = c(0, 1.5), marker_size = 2)
ggsave(filename = "crTSNE_HM_Adam12_ntx2_log2exp.svg", device = devSVG, width = 8, height = 8)

******************************************************************************
genes_ext_basmem <-
c("Hspg2", "Col16a1", "Col8a1", "Fbln7", "Col18a1", "Lama4", "Col14a1", "Nid2", "Lamb1", "Col15a1", "Dag1", "Fbln2", "Lamb2", "Fbln1", "Col4a1", "Col4a2", "Lamc1", "Nid1", "Lama2", "Fbln5")

genes_ext_cytokines <-

### this gives max raw count values

{print("basement membrane")
print(max(exprs(gbmntx2[which(fData(gbmntx2)$symbol == 'Hspg2'),)]))
print("cytokines")
print(max(exprs(gbmntx2[which(fData(gbmntx2)$symbol == 'Ccl19'),)]))
}

### this will exclude non zero genes to generate mean values from log2

{print("basement membrane")
print("cytokines")
}

### this gives max log2 values

{print("basement membrane")
print("cytokines")
}
### basement membrane

```r
genes_ext_basmem <- c("basement membrane")

visualize_gene_markers2(gbm_log2ntx2, genes_ext_basmem[1],
tsne_projntx2[cbind("TSNE.1","TSNE.2")], limits=c(0,4), marker_size = 1)

ggsave(filename="crTSNE_HM_Hspg2_ntx2_log2exp.svg", device = devSVG, width =8 , height = 8)
```

### cytokines

```r
genes_ext_cytokines <-
c("Ccl19","Cxcl3","Csf2","Cxcl14","Cxcl5","Cxcl1","Cxc12","Ccl7","Ccl8","Csf1","Ccl2","Cxl12","Il33","Il34","Tnfsf9")

visualize_gene_markers2(gbm_log2ntx2, genes_ext_cytokines[1],
tsne_projntx2[cbind("TSNE.1","TSNE.2")], limits=c(0,0.25), marker_size = 1)

ggsave(filename="crTSNE_HM_Ccl19_ntx2_log2exp.svg", device = devSVG, width =8 , height = 8)
```

**********************************************************

#### Library ID Tsne for D 0 1 2 4 (BY R. WILDER SCOTT)

```r
fourcol <- c("#F564E3","#DE8C00","#F8766D","#619CFF")

gbmntxlibid <- gbmntx@phenoData@data$barcode
{
    rm(gbmntxlibid)  ### clears variables
    rm(wlnlibid)
    rm(wi)
    rm(wj)
    rm(wk)
    rm(wl)

    gbmntxlibid <- gbmntx@phenoData@data$barcode          #### loads the barcode list in the proper order
    wlnlibid <- length(gbmntx@phenoData@data$barcode)   #### detects and stores barcode list length

    wl = matrix()                                           #### declares matrix to store library ids
    for (wi in gbmntxlibid[1:wlnlibid]){
        wj <- as.character(wi)                 #### makes factor into character string
        wk <- substr(wj, start=18, stop = 18)  #### Take the last character in the string (barcode is 16 plus - and LibID number is 18)
        wk <- as.integer(wk)                                   #### defines value as integer for matrix input
        wl <- c(wl,wk)                                         #### input value to end of vector
    }
    wl <- wl[-1]                         #### removes first vector position since empty from frist loop

    str(wl)
}
```
Barcode <- gbmntx@phenoData@data$barcode
Cluster <- w1
gbmntxlibiddf <- data.frame(Barcode, Cluster)
visualize_clusters(gbmntxlibiddf$Cluster,tsneProj0124[c("TSNE.1","TSNE.2")],colour=fourcol[c(4,1,2,3)], marker_size=1)
ggsave(filename="ws20171219_gbmntxLibIDtsne.svg", device = devSVG, width = 8, height = 8)

*****************************************************************************
FeaturePlot(object = gbmSntx, features.plot = "Inmt", do.hover = TRUE, data.hover = c("ident", "PC1", "nGene"))
DotPlot(object = gbmSntx, genes.plot = features.plot, plot.legend = TRUE)

Appendix B - Example R code used to analyze MP datasets and generate figures for the
mouse mesenchymal atlas.

library(Seurat)
library(dplyr)
library(RSvgDevice)
qmpsDATA <- Read10X(data.dir="/brcwork/underhill_lab/wscott/ws20180518QMPs/outs/filtered_gene_bc_matrices_mex/mm10_plusCreTomSSM2GFP")
dense.size <- object.size(x = as.matrix(x = qmpsDATA))
dense.size
sparse.size <- object.size(x = qmpsDATA)
sparse.size
dense.size/sparse.size
qmps <- CreateSeuratObject(raw.data = qmpsDATA, min.cells = 3, min.genes = 200, project = "qmps201805")
mito.genes <- grep(pattern = "^MT-", x = rownames(x = qmps@data), value = TRUE)
percent.mito <- Matrix::colSums(qmps@raw.data[mito.genes, ])/Matrix::colSums(qmps@raw.data)
qmps <- AddMetaData(object = qmps, metadata = percent.mito, col.name = "percent.mito")
VlnPlot(object = qmps, features.plot = c("nGene", "nUMI", "percent.mito"), nCol = 3)
ggsave(filename="~/ws20180517_Seurat_qmps_vln_metaData.svg", nCol = 3)
ggsave(filename="~/ws20180517_Seurat_qmps_percentmito.svg", nCol = 3)
ggsave(filename="~/ws20180517_Seurat_qmps_dotplot_metaData.svg", nCol = 3)
qmps <- FilterCells(object = qmps, subset.names = c("nGene", "percent.mito"), low.thresholds = c(200, -Inf), high.thresholds = c(5000, 0.05))
qmps <- NormalizeData(object = qmps, normalization.method = "LogNormalize", scale.factor = 10000)
qmps <- FindVariableGenes(object = qmps, mean.function = ExpMean, dispersion.function = LogVMR, x.low.cutoff = 0.0125, x.high.cutoff = 3, y.cutoff = 0.5)
length(x = qmps@var.genes)
qmps <- ScaleData(object = qmps, vars.to.regress = c("nUMI", "percent.mito"))
qmps <- RunPCA(object = qmps, pc.genes = qmps@var.genes, do.print = TRUE, pcs.print = 1:5, genes.print = 5)  ### default is 20 perhaps test running 100 since elbow looks like it is at 20
PrintPCA(object = qmps, pcs.print = 1:5, genes.print = 5, use.full = FALSE)
VizPCA(object = qmps, pcs.use = 1:4)
ggsave(filename="/ws20180517_Seurat_qmps_4PC.svg", device = devSVG, width = 8, height = 8)
VizPCA(object = qmps, pcs.use = 1:15)
ggsave(filename="/ws20180517_Seurat_qmps_15PC.svg", device = devSVG, width = 8, height = 8)
PCAPlot(object = qmps, dim.1 = 1, dim.2 = 2)
ggsave(filename="/ws20180517_Seurat_qmps_PCAplot.svg", device = devSVG, width = 8, height = 8)
qmps <- ProjectPCA(object = qmps, do.print = FALSE)
PCHeatmap(object = qmps, pc.use = 1, cells.use = 500, do.balanced = TRUE, label.columns = FALSE)
PCHeatmap(object = qmps, pc.use = 1:18, cells.use = 500, do.balanced = TRUE, label.columns = FALSE, use.full = FALSE)
ggsave(filename="/ws20180517_Seurat_qmps_PCAheatmap.svg", device = devSVG, width = 8, height = 8)
qmps <- JackStraw(object = qmps, num.replicate = 100)
JackStrawPlot(object = qmps, PCs = 1:19)
ggsave(filename="/ws20180517_Seurat_qmps_jackstraw.svg", device = devSVG, width = 8, height = 8)
PCElbowPlot(object = qmps)
ggsave(filename="/ws20180517_Seurat_qmps_elbowplot.svg", device = devSVG, width = 8, height = 8)
save.image(file = "~/qmps.Rdata")
save(qmps, file = "~/qmps.Robj")
rm(list=ls())
q()