2017

Single Cell Molecular Heterogeneity In Musculoskeletal Differentiation

Claire Marie Mcleod
University of Pennsylvania, cm.mcleod@gmail.com

Single Cell Molecular Heterogeneity In Musculoskeletal Differentiation

Abstract
Mesenchymal stem cells (MSCs) display substantial cell-to-cell variation that manifests across many aspects of cell phenotype and complicates the use of MSCs in regenerative applications. However, most conventional assays measure MSC properties in bulk and, as a consequence, mask this cell-to-cell variation. To better understand MSC heterogeneity and its underlying mechanisms, we quantitatively assessed MSC phenotype within the context of chondrogenesis, amongst clonal populations and single cells. Clonal MSCs differed in their contractility, ability to transmit extracellular strain the nucleus, capacity to form cartilage-like matrix, and transcriptomic signature. RNA FISH measurements of single cell gene expression found that both primary chondrocytes and chondrogenically-induced MSCs showed substantial mRNA expression heterogeneity. Surprisingly, variation in differentiation marker transcript levels only weakly associated with cartilage-like matrix production at the single cell level. This finding suggested that, although canonical markers have very clear functional roles in differentiation and matrix formation, their instantaneous mRNA abundance is only tenuously linked to the chondrogenic phenotype and matrix accumulation at the single cell level. One possible explanation for the apparent disconnect between gene and protein expression is that mRNA and protein exhibit different temporal dynamics. Using stochastic models of single cell behavior, we explored the impact of transcriptional stochasticity and temporal matrix dynamics on the perceived relationship between single cell mRNA and protein abundance. Simulations suggested that considering recent temporal fractions of protein (vs. total protein) increased the correlation between mRNA and protein abundance, and illustrated that mRNA stability was a crucial determinant of the timescale over which any such correlation persisted. Experimentally, non-canonical amino acid tagging was used to visualize and quantify temporal fractions of nascent extracellular matrix with high fidelity. The organization and temporal dynamics of the proteinaceous matrix depended on the biophysical features of the microenvironment, including the biomaterial scaffold and the niche constructed by cells themselves. Both chondrocytes and MSCs demonstrated marked cell-to-cell heterogeneity in nascent matrix production, consistent with model predictions. Ongoing work aims to combine these experimental measurements of nascent protein expression with instantaneous measures of mRNA abundance to better understand the mRNA-protein relationship, and to harness this understanding to improve regenerative therapies.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Bioengineering

First Advisor
Robert L. Mauck

Keywords
chondrogenesis, mesenchymal stem cells, single cell measurements

SINGLE CELL MOLECULAR HETEROGENEITY IN MUSCULOSKELETAL DIFFERENTIATION

Claire M. McLeod

A DISSERTATION

in

Bioengineering

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2017

Supervisor of Dissertation

_____

Robert L. Mauck
Mary Black Ralston Professor for Education and Research in Orthopaedic Surgery
Professor of Bioengineering

Graduate Group Chairperson

_____

Jason A. Burdick
Professor of Bioengineering

Dissertation Committee

Jason A. Burdick, Professor of Bioengineering
Maurizio Pacifici, Professor of Orthopaedic Surgery
Arjun Raj, Associate Professor of Bioengineering

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Rob Mauck. Working in Rob's lab has been a true privilege, and I am immensely grateful for the opportunity to learn from such a kind and brilliant mentor. His guidance has been invaluable, and I am tremendously appreciative of his support.

I am also grateful to my committee chair, Arjun Raj. Arjun's group has been my 'lab-away-from-lab', and I am incredibly fortunate to count him as a mentor. He has strongly influenced my perspectives on science, single cell heterogeneity and transcriptional stochasticity and without him, this would be a very different thesis.

I would also like to acknowledge the other members of my committee for their time and guidance. Jason Burdick and Maurizio Pacifici have both generously shared their expertise and provided insightful feedback.

This work has also been strongly influenced by other graduate students. In particular, it stems from investigations of clonal heterogeneity led by fellow graduate student Megan Farrell. Megan's measurements of clonal micromechanics and transcription (via RNAseq) are discussed in Chapters 3 and 4. This work also represents a close collaboration with Arjun's group, and I am especially grateful to fellow grad student Allison Cote. Our work together was a true partnership, and the studies described in Chapter 4 would not have been possible without such close collaboration. Our joint efforts to parse single cell transcriptional heterogeneity grew out of preliminary studies

performed by Megan, Margaret Dunagin and Patrick McClanahan, and I thank each of them for initiating the project. I would also like to acknowledge Tristan Driscoll and Suchin Heo, for kindly teaching me many new techniques and encouraging my growth as a scientist. Tristan, Suchin and Brian Cosgrove also contributed measurements of cellular traction force that are included in Chapter 3.

The members of the Mauck group and the broader McKay Lab community have been a constant source of friendship and scientific camaraderie. I am especially grateful to fellow grad students Brian Cosgrove and Bhavana Mohanraj, and postdocs Sun Peck and Sarah Gullbrand. I am also privileged to have work with and alongside many other excellent scientists, including Matt, Julia, John, Lachlan, DongHwa, Sylvia, Spence, Minwook, Claudia, Beth, Brendan, Bre, Justin, Sonia, Tonia, Eddie and Eric. Thank you to Liz for managing almost every aspect of the lab.

Finally, I would like to thank my partner, Todd. You're the best.

ABSTRACT

SINGLE CELL MOLECULAR HETEROGENEITY IN MUSCULOSKELETAL DIFFERENTIATION

Claire M. McLeod

Robert L. Mauck

Mesenchymal stem cells (MSCs) display substantial cell-to-cell variation that manifests across many aspects of cell phenotype and complicates the use of MSCs in regenerative applications. However, most conventional assays measure MSC properties in bulk and, as a consequence, mask this cell-to-cell variation. To better understand MSC heterogeneity and its underlying mechanisms, we quantitatively assessed MSC phenotype within the context of chondrogenesis, amongst clonal populations and single cells. Clonal MSCs differed in their contractility, ability to transmit extracellular strain the nucleus, capacity to form cartilage-like matrix, and transcriptomic signature. RNA FISH measurements of single cell gene expression found that both primary chondrocytes and chondrogenically-induced MSCs showed substantial mRNA expression heterogeneity. Surprisingly, variation in differentiation marker transcript levels only weakly associated with cartilage-like matrix production at the single cell level. This finding suggested that, although canonical markers have very clear functional roles in differentiation and matrix formation, their instantaneous mRNA abundance is only tenuously linked to the chondrogenic phenotype and matrix accumulation at the single cell level. One possible explanation for the apparent disconnect between gene and protein expression is that mRNA and protein exhibit different temporal dynamics. Using stochastic models of single cell behavior, we explored the impact of transcriptional stochasticity and temporal

matrix dynamics on the perceived relationship between single cell mRNA and protein abundance. Simulations suggested that considering recent temporal fractions of protein (vs. total protein) increased the correlation between mRNA and protein abundance, and illustrated that mRNA stability was a crucial determinant of the timescale over which any such correlation persisted. Experimentally, non-canonical amino acid tagging was used to visualize and quantify temporal fractions of nascent extracellular matrix with high fidelity. The organization and temporal dynamics of the proteinaceous matrix depended on the biophysical features of the microenvironment, including the biomaterial scaffold and the niche constructed by cells themselves. Both chondrocytes and MSCs demonstrated marked cell-to-cell heterogeneity in nascent matrix production, consistent with model predictions. Ongoing work aims to combine these experimental measurements of nascent protein expression with instantaneous measures of mRNA abundance to better understand the mRNA-protein relationship, and to harness this understanding to improve regenerative therapies.

TABLE OF CONTENTS

ACKNOWLEDGMENTS ..... **II**

ABSTRACT ..... **IV**

TABLE OF CONTENTS ..... **VI**

LIST OF TABLES ..... **XIV**

LIST OF ILLUSTRATIONS ..... **XV**

CHAPTER 1: INTRODUCTION **1**

CHAPTER 2: BACKGROUND **5**

Tissue engineering cartilage to recapitulate the extracellular matrix **5**

Articular cartilage in health and disease 5

The extracellular matrix of native articular cartilage 5

Tissue engineering cartilage to recapitulate the extracellular matrix 7

Mesenchymal stem cell heterogeneity **9**

The defining properties of mesenchymal stem cells 9

MSCs exhibit heterogeneity on multiple levels 9

Inter-clonal functional variation 11

Inter-clonal molecular variation 12

Biophysical variation amongst clones and individual cells 14

Intra-clonal heterogeneity and cell-to-cell variation 17

Potential origins & mechanisms of MSC heterogeneity 18

Measuring cell-to-cell molecular variation 22

Assaying gene expression in single cells 22

Assaying protein expression in single cells 25

Conclusions & Outlook 26

CHAPTER 3: MORPHOLOGIC AND MECHANOBIOLOGIC HETEROGENEITY IN CLONAL MESENCHYMAL STEM CELL POPULATIONS 28

Introduction 28

Results 30

Clonal variations in MSC proliferation and morphology 30

Clonal differences in nuclear morphology and heterochromatin content 30

Clonal variation in cytoskeletal-to-nuclear strain transmission 32

Cell-to-cell and clonal variation in contractility 34

Micromechanics of the cell and nascent pericellular matrix 37

Discussion 39

Methods 42

Cell isolation & culture 42

Fluorescent imaging and quantification 44

Micropost and traction force microscopy measures of contractility 44

Tensile stretch to assess nuclear deformation 45

Compression to assess matrix micromechanics 46

Statistics 47

CHAPTER 4: SINGLE-CELL DIFFERENCES IN INSTANTANEOUS MATRIX GENE EXPRESSION DO NOT PREDICT MATRIX DEPOSITION 48

Introduction 48

Results 49

Single cells express differentiation markers heterogeneously 49

RNA levels poorly predict single-cell functional potential 55

Transcriptomics does not identify better marker sets 60

Marker heterogeneity emerges rapidly after cell division 63

Marker genes do not identify a chondrocyte phenotype 65

Discussion 70

Methods 74

Cell Isolation and Expansion 74

Cell Encapsulation 75

Chondrogenic Pellet Culture and Biochemical Content 76

Live Cell Imaging and Tracking 76

RNA Fluorescence In-Situ Hybridization and Imaging 77

Quantification of Copy Number from RNA FISH Images 78

Quantification of Extracellular Matrix Deposition 78

RNA sequencing 80

Cell Volume and Area Measurements 80

Alcian Blue Staining 80

Statistical Comparisons 81

CHAPTER 5: HIGH FIDELITY VISUALIZATION OF CELL-TO-CELL VARIATION AND TEMPORAL DYNAMICS IN NASCENT EXTRACELLULAR MATRIX FORMATION 82

Introduction 82

Results 85

Methionine analogs enable the fluorescent labeling of extracellular matrix proteins. 85

Metabolic labeling tags pericellular collagenous network. 92

Biophysical features of the cellular microenvironment influence matrix organization 92

Nascent matrix intersperses with pre-existing matrix in developing (but not developed) microenvironments. 96

Patterns of matrix deposition identify phenotypic differences between cell types. 98

Discussion 105

Methods 109

Quantifying protein methionine content 109

Construct fabrication 109

Functional, noncanonical amino acid labeling 110

FUNCAT Imaging and Image Quantification 111

Alcian Blue Staining & Immunofluorescence 112

Matrix Digestion & Perturbation 112

Microcompression 113

CHAPTER 6: STOCHASTIC SIMULATIONS LINK GENE EXPRESSION HETEROGENEITY AND NASCENT PROTEIN PRODUCTION 114

Introduction 114

Experimental Results 116

Transcriptional dynamics differ between chondrogenic marker genes 116

A stochastic model of chondrocyte gene expression and matrix production 118

Establishing and parameterizing a stochastic model 118

Intrinsic heterogeneity emerges in single cell simulations using realistic model parameters 121

In silico correlation between temporal protein fraction and instantaneous gene expression 124

Correlations between nascent and total protein fractions: model and experiment 125

Discussion 125

Methods 129

Cell isolation & culture 129

RNA Fluorescence In-Situ Hybridization, Imaging, and Quantification 129

Fitting transcriptional parameters to experimental data 131

Computational simulations 131

CHAPTER 7: SUMMARY AND FUTURE DIRECTIONS 133

Summary 133

Future directions 136

Defining the role of observation window on mRNA-protein correlations 136

Leveraging heterogeneity to improve outcomes in cartilage tissue engineering 138

Modulating MSC heterogeneity 142

Conclusion 144

APPENDIX 1: RNA FISH PROBE SEQUENCES 145

APPENDIX 2: SUPPLEMENTARY STATISTICAL ANALYSIS RELATED TO CHAPTER 4 149

APPENDIX 3: METHIONINE CONTENT OF EXTRACELLULAR MATRIX PROTEINS 156

BIBLIOGRAPHY..**169**

LIST OF TABLES

Table 1: Single cell methods to assay mRNA and protein abundance. 22

Table 2: Mean and Coefficient of Variation Associated with Aggrecan RNA Count in Undifferentiated and Differentiated Cells 54

Table 3: Parameters used to model expression of chondrogenic markers......121

Supplementary Table 1: Additional statistical information related to GAPDH RNA abundance differentiating MSCs. 149

Supplementary Table 2: Correlation of marker gene RNA abundance in individual differentiating MSCs. 150

Supplementary Table 3: Coefficient of variation (CV) of single cell RNA count in small MSC colonies. 150

Supplementary Table 4: Additional statistical information related to aggrecan RNA abundance differentiating MSCs. 151

Supplementary Table 5: Additional statistical information related to divergence in aggrecan RNA abundance between sister cells. 152

Supplementary Table 6: Additional statistical information related to divergence in GAPDH RNA abundance between sister cells. 153

Supplementary Table 7: Additional statistical information related to RNA abundance in chondrocytes during passage-induced de-differentiation......154

Supplementary Table 8: Additional statistical information related to RNA abundance in chondrocytes during re-differentiation in agarose......155

LIST OF ILLUSTRATIONS

Figure 2-1: The cartilage extracellular matrix is primarily comprised of collagens and proteoglycans. 6

Figure 2-2: Bulk observations can mask heterogeneity. 8

Figure 2-3: Schematics of select methods to measure mRNA in single cells. 24

Figure 3-1: Clonal MSC populations differ morphologically and in their proliferation rates. 31

Figure 3-2: Passage-dependent clonal differences in nuclear morphology and chromatin organization. 33

Figure 3-3: Passage-dependent clonal differences in nuclear deformability. 34

Figure 3-4: Traction force exerted by cells differs between individual cells and amongst clonal populations. 36

Figure 3-5: Chondrogenically-induced clonal populations differ in their response to compression. 38

Figure 4-1: RNA FISH reveals heterogeneity in lineage marker expression in mesenchymal stem cells and chondrocytes. 52

Figure 4-2: Matrix production, GAPDH copy number, and viability of MSCs in 3D culture. 53

Figure 4-3: Marker gene expression is a poor predictor cartilage-like matrix production in individual MSCs. 58

Figure 4-4: Matrix staining intensity versus mRNA copy number and ROC curves for individual donors, markers, and marker ratios. 60

Figure 4-5: Genome-wide transcriptome profiling does not predict MSC functional potential. 61

Figure 4-6: Marker expression heterogeneity emerges shortly after cell division. 64

Figure 4-7: Heritability of marker copy number through cell division. 66

Figure 4-8: Chondrocyte de-differentiation and re-differentiation are not driven by altered absolute aggrecan expression. 68

Figure 4-9: Chondrocyte morphology with passage number. 69

Figure 5-1: HPG labeling enables high fidelity visualization of the extracellular matrix. 87

Figure 5-2: Large area scans confirm uniform labeling across gel. 88

Figure 5-3: Matrix labeling with azidohomoalanine (AHA). 89

Figure 5-4: Fluorescent tagging of HPG is highly specific. 89

Figure 5-5: Hyaluronidase digestion of fixed samples increases collagen II detection by immunofluorescence, but does not alter HPG labeling. 91

Figure 5-6: Enzymatic digestion differentially effects HPG-labeled matrix. 93

Figure 5-7: HPG labeling captures nascent matrix deposition. 94

Figure 5-8: Density of the cellular microenvironment impacts organization of nascent matrix. 95

Figure 5-9: Extent and organization of nascent matrix assembly depend on construct maturity and collagen cross-linking. 97

Figure 5-10: Apparent heterogeneity in matrix accumulation is cell type and timescale dependent. 101

Figure 5-11: Cell-to-cell heterogeneity in HPG-labeled nascent matrix formed by chondrocytes. 103

Figure 5-12: Exposure to BAPN increases nascent matrix synthesis. 103

Figure 5-13: Cell-to-cell heterogeneity in HPG-labeled nascent matrix formed by mesenchymal stem cells (MSCs). 105

Figure 6-1: Transcriptional differences in expression of chondrogenic marker genes. 117

Figure 6-2: Schematic of the reactions included in the stochastic simulations. ..118

Figure 6-3: Comparison of chondrocyte extracellular matrix production in the presence and absence of MMP inhibitors. 120

Figure 6-4: Cell-to-cell variation arises in simulations of single cell gene and protein expression. 122

Figure 6-5: In silico correlation between nascent protein and final mRNA count is time- and half-life dependent. 124

Figure 6-6: Variability of and correlation between nascent and total protein fractions. 126

Figure 7-1: Preliminary controls related to the combination of FUNCAT and RNA FISH. 138

Figure 7-2: Schematic of possible distributions underlying population-level improvement in differentiation potential ($\mu_1 \rightarrow \mu_2$). 143

Chapter 1: Introduction

Traumatic injury and degenerative processes damage articular cartilage and compromise normal joint function.[74,136] Cartilage has a limited intrinsic healing capacity, and treatment of the damaged joint surface often requires its replacement. Tissue-engineered cartilage replacements utilize cells and biomaterials to create and maintain constructs that mimic the structure and function of native tissue. [95,136] While chondrocytes, the cells native to cartilage, can be used to create robust engineered cartilage, their availability is limited. As a result, many approaches now rely on mesenchymal stem cells (MSCs), multipotent progenitor cells that are easily expanded and can be directed to undergo chondrogenesis.

While MSC-based approaches can yield engineered constructs that mimic native cartilage, the clinical use of MSC-based therapies is complicated by the extensive heterogeneity that typifies this cell source. Within a single population, MSCs undergo chondrogenesis to varying extents and over different timescales. This variability offers potential opportunities to identify and select the most strongly chondrogenic subpopulations and so improve the performance of MSC-based tissue engineered constructs. However, while there is growing appreciation of the inherent cell-to-cell variation amongst MSCs, the metrics most indicative of differentiation potential remain to be identified. Towards this end, this thesis investigates the heterogeneity of MSCs between single-cell derived clonal populations and amongst individual MSCs, with

particular focus on the ability of cells to produce a cartilage-like extracellular matrix. In parallel, MSC variability is contrasted with the relative homogeneity of chondrocytes.

The scientific context of this work is established in Chapter 2, which reviews the current literature surrounding MSC variability. This discussion recognizes three types of variation that exist between and within clonal MSC populations: i) variation in functional capacity to differentiate, ii) molecular content (e.g. gene and protein expression), and iii) cellular biophysical properties. This chapter also describes important features of cartilage's extracellular matrix, and provides an overview of techniques used to monitor gene and protein expression at the single cell level.

Chapter 3 identifies cellular mechanotransduction as a new dimension of clonal heterogeneity. Specifically, it details how clonal MSC populations differ in their cellular contractility, ability to transmit strain from the extracellular environment to the nucleus, and their ability to construct mechanically competent cartilage-like matrix. In Chapter 4, we also identify transcriptomic differences between clones, and relate these differences to their chondrogenic functional capacity.

Chapter 4 further extends with analysis with a primary focus on single cell differences in gene expression. Interestingly, there is high transcriptional variability within clonal MSC populations, and cell-to-cell variation persists throughout the course of differentiation. However, we show that this transcriptional variation is unable to explain differences in matrix production at the single cell level, and demonstrate that there is an apparent disconnect between the mRNA and protein abundance of traditional chondrogenic

markers. Intriguingly, despite their perceived homogeneity, chondrocytes also demonstrate cell-to-cell variation in gene expression. This lack of correlation between gene and protein expression is further reinforced by the finding that absolute expression of matrix genes increases with chondrocyte dedifferentiation, even as matrix protein synthesis decreases.

One potential explanation for this disconnect between gene and protein expression is that is that these molecules differ in their temporal dynamics. To facilitate an improved understanding of matrix synthesis dynamics, Chapter 5 presents a strategy to metabolically label matrix proteins as they are produced (via functional non-canonical amino acid tagging). This approach enables detailed visualization of matrix organization and can distinguish temporal matrix fractions. We develop this technique for 3D cultures of chondrocytes and MSCs and use it to examine how time, phenotype, and cellular microenvironment regulate matrix production at the single cell level. Cell-to-cell variability is apparent in these data as well, and intriguingly, nascent extracellular matrix appears to be more variable than total matrix protein content.

Building on this observation, Chapter 6 develops and applies a computational model to explore the impact of the observational window in our single cell measures on the perceived relationship between instantaneous mRNA expression, nascent protein expression, and total extracellular matrix protein deposition. To do so, the transcriptional dynamics of five matrix genes are experimentally quantified and used to parameterize stochastic simulations of single cell gene and protein expression. These simulations of mRNA and protein abundance suggest that the correlation between mRNA and protein

increases as the observational window is shortened. Furthermore, in the context of matrix genes and their associated proteins, mRNA half-life appears to be the key determinant of the mRNA-protein correlation in this 3D context.

Finally, Chapter 7 concludes this work with a discussion of the implications of MSC heterogeneity on cartilage tissue engineering efforts and outlines potential future work that may provide additional insight into the relationship between mRNA and protein abundance of extracellular matrix genes and methods to improve selection of superior cell populations for tissue engineering and repair applications.

Chapter 2: Background

Tissue engineering cartilage to recapitulate the extracellular matrix

Articular cartilage in health and disease

Articular cartilage covers the surfaces of articulating joints and is essential to healthy joint function. However, injury and disease often compromise the ability of cartilage to bear load and facilitate smooth joint movement. Unfortunately, the intrinsic healing capacity of cartilage is limited, and tissue lesions often require surgical intervention. However, these procedures have substantial limitations and there is considerable interest in the development of tissue engineered cartilage replacements. Towards this end, researchers aim to combine cells, biomaterial scaffolds, and suitable biochemical/biophysical cues to yield engineered constructs that mimic the structure and function of healthy native tissue, particularly the collagen and proteoglycan-rich extracellular matrix (ECM) that bears and distributes load in the healthy joint.

The extracellular matrix of native articular cartilage

Cartilage ECM is primarily comprised of collagens and proteoglycans, which respectively account for approximately 60% and 30% of the tissue's dry weight.[17] The collagens form a fibrillar, mesh-like network that provides tensile strength and mechanically constrains the swelling pressure engendered by the proteoglycans. These immobilized proteoglycans concentrate negative charges and attract soluble cations, creating high osmotic pressure and providing cartilage with its compressive resistance.

The collagen network primarily consists of three cartilage-specific collagens: collagen II, collagen IX, and collagen XI. Collagen II is a highly abundant, fibril-forming collagen that consists of three identical α1(II) chains and comprises 80-90% of the total collagen content.[59] Individual collagen II molecules organize head-to-tail to form a staggered polymer, and bundle into microfibrils and fibrils that also incorporate additional collagen IX and XI molecules (**Figure 2-1**).[59] Collagen IX binds to the surface of collagen II fibrils, and and consists of three distinct chains [α1(IX), α2(IX), and α3(IX)]. It is thought to bridge between fibrils and covalently anchor other matrix molecules to the collagen network.[59] In contrast, collagen XI integrates within collagen II fibrils and displays its N-terminal domain on the fibril surface.[59] The protruding N-terminal domains are thought to laterally cross-link fibrils and limit fibril diameter. Collagen XI molecules consist of two

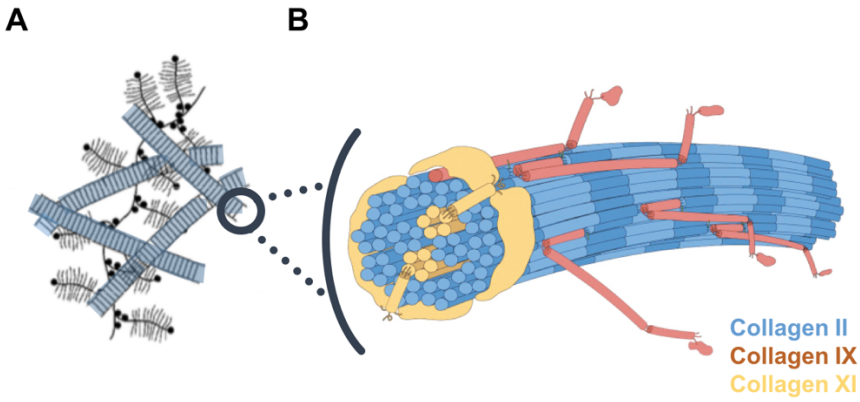

Figure 2-1: The cartilage extracellular matrix is primarily comprised of collagens and proteoglycans. A) Schematic depicting the primary molecular constituents of articular cartilage: aggrecan macromolecules and collagen fibrils. Reprinted from Trends in Cell Biology, Vol 5, Barbara M. Vertel, The Ins and Outs of Aggrecan, 458-464, 1995, with permission from Elsevier.[221] B) Structure of collagen II – collagen IX – collagen XI fibrils. Modified from Kadler et al. under a Creative Commons Attribution license (CC BY).[97]

unique alpha chains [α1(XI) and α2(XI)] and a third, post-translationally modified α1(II) chain.[59] In adult cartilage, an α1(V) chain may substitute for either the α1(XI)or α2(XI) chain.[59]

Other collagens found in cartilage include collagens VI and X. Collagen VI is a non-fibrillar molecule that is found immediately adjacent to chondrocytes and is thought to facilitate cell-matrix attachment.[26,131] Collagen X is found near chondrocytes in calcifying tissue, and is associated with cartilage hypertrophy.[131]

The proteoglycan fraction of cartilage predominantly consists of aggrecan molecules, which account for ~90% of proteoglycan content by mass.[17] Individual aggrecan molecules consist of a core protein synthesized from the *aggrecan* gene and many attached glycosaminoglycan chains (GAGs, e.g. chondroitin sulfate and heparin sulfate).[221] These protein-GAG complexes associate with link proteins and and hyaluronic acid, forming larger aggregates which become ensnared in the collagen network. Each sulfated GAG bears negative charges, and their collective aggregation yields the fixed negative charge density responsible for cartilage's compressive properties. Other proteoglycans include biglycan and decorin, which bind to collagens and other molecules, including chondrogenic growth factors. Additional minor proteoglycans further contribute to ECM structure and cellular function.[184]

Tissue engineering cartilage to recapitulate the extracellular matrix

ECMs are created and maintained by cells, and the choice of tissue engineering cell source strongly influences engineered cartilage quality. Chondrocytes, the cells resident

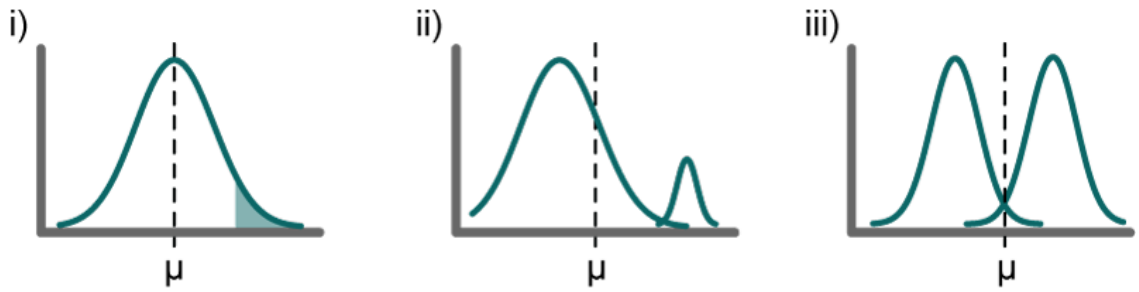

Figure 2-2: Bulk observations can mask heterogeneity, including i) 'tail' observations, ii) small subpopulations, and iii) bimodal behavior. Adapted from Altschuler et al.[3].

in native cartilage, excel at producing robust extracellular matrices *in vitro*, even under nutrient limiting conditions.[95,109] However, due to the difficulty of obtaining chondrocytes in sufficient number, chondrogenically induced stem cells are often used instead. Mesenchymal stem cells (MSCs) are readily obtained from adult tissue, expand well in culture, and can undergo chondrogenic differentiation. However, even with the most effective differentiation protocols, MSCs fail to fully match the performance of chondrocytes.[86,95] This performance gap likely exists in part due to marked variation in the ability of individual MSCs to undergo chondrogenesis: while some MSCs robustly undergo chondrogenesis, others fail to do so.[86] These underperforming, alternatively performing (e.g. osteogenic), or non-responsive subpopulations hinder the maturation of engineered tissues, but their poor performance is often masked by bulk assays that pool signal across entire cell populations (**Figure 2-2**). Recently, given the advent of single cell methods and a growing appreciation that ensemble measurements can mask important variation, new findings have begun to delineate MSC heterogeneity. In the following sections, we review the current understanding of heterogeneity between and

within MSC populations, and discuss how single cell techniques may be used to further parse this variability.

Mesenchymal stem cell heterogeneity

The defining properties of mesenchymal stem cells

As a cell type, MSCs are defined by three criteria. MSCs must 1) be plastic adherent; 2) express the surface markers CD105, CD73 and CD90, and lack expression of CD45, CD34, CD14 or CD11b, CD79 or CD19 and HLA-DR; and 3) be capable of differentiating into osteoblasts, adipocytes, and chondroblasts.[48] However, this operational definition does not precisely define a homogenous population of multipotent progenitors. Instead, it describes a heterogeneous group of cells that demonstrate variability between tissues of origin, between individual donors, amongst clonal subpopulations, and at the single cell level.

MSCs exhibit heterogeneity on multiple levels

While MSCs were first isolated from bone marrow,[153,167] they have since been identified in many connective tissues, including adipose tissue, the umbilical cord, and dental pulp.[206] In standard isolation techniques, adipose or bone marrow aspirates are progressively centrifuged, and filtered before being plated into culture. A small fraction of the cells (the presumed MSCs) will adhere to the tissue culture plastic, and proliferate. Both bone marrow- and adipose-derived MSCs are readily available,[58] yet they originate from stem cell niches that provide distinct biological, chemical, and mechanical cues. Tissue-dependent variation in differentiation capacity, surface markers, and

transcriptional and proteomic profiles has been widely studied, and the reader is referred to recent reviews for comparisons of MSCs across tissue sources.[102,143,206]

Even when derived from the same tissue of origin, MSCs demonstrate tremendous donor-to-donor variability. Intuitively, donor health may influence the availability and functional potential of MSCs.[113,224] Similarly, as donors age, MSC availability, self-renewal capacity, and differentiation potential have been reported to decline.[39,100,204] Surprisingly however, even MSCs isolated from young, healthy donors exhibit stark differences in proliferation rate, differentiation capacity, and ultimate clinical utility.[165] This functional variation extends to the molecular status of these cells.[148,149] For example, mass spectroscopy of MSCs isolated from six donors revealed that only 62% of all identified proteins were found in at least half of the donors, and only 13% of identified proteins were found in cells from each donor.[148] Such donor-to-donor variation has important clinical implications, and has motivated more detailed investigation of MSC variability.

Further study has revealed that donor- and tissue-dependent differences are superimposed upon cell-to-cell variation amongst MSCs within a single population. For example, multiple bone marrow aspirates simultaneously isolated from a single donor yield MSC cultures that proliferate at significantly different rates. [165] Even within a single isolate, cell-to-cell variation in MSC phenotype becomes evident during culture expansion and downstream use. This variation is commonly examined by comparing clonal subpopulations (groups of cells that are not only genetically identical, but also recently derived from a single parent cell).[87]

MSCs readily form clones, and their clonogenicity can be observed by sparsely plating an initial isolate and monitoring colony formation. Populations of cells from the same initial colony can be obtained by sub-culturing these colonies, or by seeding single cells into individual culture wells via limiting dilution or flow cytometry[200]. Inter-clonal heterogeneity (variation between clones) most obviously manifests as morphological and proliferative diversity. Clones differ in morphological phenotype, ranging from elongated spindle-like cells to large flattened cells and highly protrusive cells.[34,153] Similarly, individual clones proliferate both more quickly and more slowly than their corresponding polyclonal parent populations and differ in their self-renewal capacity, with select clones reaching early senescence.[185] Furthermore, inter-clonal variation is not limited to these phenotypic characteristics: it extends to include functional capacity, molecular signature, and the mechanical state of the cell.

Inter-clonal functional variation

The differentiation capacity of clonal MSCs was first studied to prove the existence of multipotent cells capable of committing towards adipogenic, osteogenic, and chondrogenic fates.[167] However, in parallel with the discovery of these tri-potential cells, other clonal populations were identified that had restricted differentiation potential and only differentiated towards a subset of the three canonical fates.[153,167,186] The relative frequency of clones with limited differentiation potential was donor dependent, with estimates suggesting that ~50% of clones are tri-potential, ~30% are bipotential (either osteo-chondro or osteo-adipo), and ~10% are unipotential osteoprogenitors.[186]

This functional heterogeneity is also apparent *in vivo*, and influences the utility of MSCs in multiple regenerative contexts. For example, clonal MSC populations implanted subcutaneously in mice demonstrate variable osteogenic capacity, with approximately half of clonal implants undergoing some degree of osteogenesis.[114] Similarly, clonal populations screened *in vitro* for above-average chondrogenic capacity result in repair of cartilage defects more robustly than unscreened populations.[93] Inter-clonal functional heterogeneity also extends to include commitment towards non-canonical fates. Stem cells derived from dental pulp demonstrate heterogeneous myogenic potential, and clones that are highly myogenic *in vitro* also engraft into muscle defects more efficaciously than the polyclonal parent populations.[232] Such functional variability potentially offers the opportunity to harness clonal identity and prospectively identify MSC subpopulations best suited to drive the functional restoration of diverse tissues. However, the challenge of performing *in vitro* screening of clonal functional capacity at a clinically-useful scale has motivated ongoing work to identify molecular or biophysical markers of MSC differentiation potential.

Inter-clonal molecular variation

Inter-clonal functional heterogeneity must derive from underlying molecular variation. While proteomic studies comparing individual clonal populations are challenging due to inherently limited cell number, comparisons between pooled fast- and slow-growing clones suggest broad trends. Fast-growing clones were more likely to be tri-potential than slow-growing clones,[140,186] and rapidly self-renewing MSCs engrafted into tissues more readily than slowly renewing MSCs.[121] Indeed, fast- and slow-growing MSCs

differed proteomically, with differential expression of proteins including intermediate filaments (e.g. lamin A/C), calcium-binding proteins (e.g. calmodulin), and glycolytic proteins (e.g. glyceraldehyde-3-phosphate dehydrogenase).[139] Furthermore, surface marker expression across clones has suggested that CD200 marks osteogenic subpopulations, while SSEA4 and CD140a are associated with adipogenic capacity.[168,183]

Clonal heterogeneity also extends to the transcriptome. Certainly, there is great divergence in transcriptional signature between high- and low-potential clones following exposure to differentiation conditions. There is also now evidence for clonal variation in basal gene expression in undifferentiated cells. Screens comparing stem cell gene expression between fast- and slow-growing clones have identified extensive differences in the expression of genes associated with the cell cycle and cellular division.[141,147] Fast-growing clones also expressed select growth factors (e.g. BMP2, FGF2, IGF1), lineage markers (e.g. aggrecan, alkaline phosphatase, collagen I, collagen II) and self-renewal markers (e.g. SOX2) more highly than slow-growing clones.[141] Conversely, other genes, including CD44, were more highly expressed in slow-growing clones.[141] Separately, direct comparisons of clonal transcriptomes indicated that clones with greater functional potential had enriched basal expression of genes implicated in skeletal and muscular development, including extracellular matrix components and MAP kinase signaling elements.[54,117,210] Notably, high baseline expression of calponin negatively correlated with clonal multipotency.[210] Within an osteogenic context, high potential clones also expressed extracellular matrix genes and genes regulated by osteogenic transcription

factors to a greater extent than poorly osteogenic clones.[117] Strikingly, basal expression of four genes, including decorin and lysyl oxidase-like 4 was more predictive of clonal osteogenesis than the expression of traditional osteoblastic markers including Runx2, collagen type I and osteopontin.[117] Such findings suggest that transcriptome-wide analysis of undifferentiated MSCs may be key to identifying prospective markers of stem cell fate; however, these predictors must also be validated through mechanistic studies identifying their role in maintenance of multi-potency and/or lineage specification.

Transcriptional activity is determined not only by the presence and activity of transcription factors, but also by the epigenetic status of the cell. DNA methylation, one type of epigenetic modification, is generally associated with a loss of gene expression and is crucial in stem cell differentiation. In undifferentiated adult stem cells, lineage-associated promoters are often hypomethylated.[13] Investigation of clonal MSC adipogenesis has found that while adipogenesis-associated promoters were hypomethylated in MSC clones, the specific pattern of methylation varied between clonal subpopulations. However, there was no clear relationship between a clone's methylation status and its gene expression pattern or ultimate adipogenic potential.[156,157]

Biophysical variation amongst clones and individual cells

The mechanical state of the cell has emerged as a potential biomarker indicative of cellular phenotype. Cellular mechanical properties reflect the underlying structure of the cell, including the cytoskeleton and nucleus. These structures change with differentiation, and also differ between committed cell types. Interestingly, increased nuclear deformability has been correlated with pluripotency. With differentiation,

chromatin condenses within the nucleus and the nuclear envelope is reinforced by increasing lamin A/C content.[161] Embryonic stem cells are 6-10 fold softer than their differentiated counterparts. [32,161] Mechanical differences of similar magnitude have been noted between individual, undifferentiated MSCs isolated and passaged together.[137] This mechanical variability may reflect the high degree of functional heterogeneity observed when these individual cells are tasked with a specific lineage transformation.

Mechanical differences also exist between clonal MSC populations. A study monitoring 32 MSC clones suggested that cellular mechanics can be prospectively used to predict differentiation capacity.[73] The functional potential of clones correlated with elastic and viscoelastic properties.[73] Clones with the highest adipogenic potential were were characterized by taller cells with lower elastic moduli.[73] Conversely, osteogenic capacity correlated with a higher elastic modulus, instantaneous modulus and relaxed modulus, while chondrogenic capacity correlated with elastic modulus and apparent viscosity.[73] Separately, efforts to biophysically sort MSCs in a high-throughput manner suggest that the cells of tri-potential MSC subpopulations are smaller, less stiff, and exhibit greater nuclear membrane fluctuations than cells with bi-potent (osteo-chondro) differentiation potential.[122]

Cell mechanics not only indicate cell phenotype, but also mediate the physical interaction between a cell and its environment. Many cell types, including MSCs, are able to sense and respond to mechanical cues. Biophysical stimuli including the elasticity of the microenvironment and exogenous forces have been widely examined as

determinants of stem cell fate. These cues elicit changes in biochemical signaling, gene expression, and ultimately cell phenotype and function.

A number of studies have begun to probe how individual MSCs response to biophysical cues. Dual adipogenic/osteogenic media causes polyclonal MSC populations to undergo mixed osteogenic and adipogenic differentiation; the relative balance between these two differentiated states is regulated by the physical stiffness of the cell microenvironment.[66,77,103] Cell response is most uniform in extremely soft or stiff environments, which favor adipogenesis and osteogenesis respectively. However, in environments of intermediate stiffness, commitment is variable: subpopulations of cells will differentiate towards each fate. Furthermore, in dynamic systems where an initially soft substrate can be stiffened, the ratio of adipogenic to osteogenic commitment is regulated by the timing of the soft-to-stiff transition.[77] Collectively, these findings begin to suggest that MSC subpopulations may have subtly different mechanical setpoints that govern mechanically-regulated fate commitment.

Any such setpoint likely relates to the tension sensed and contractility generated by an individual cell. On a population level, osteogenesis is associated with the ability of cells to spread and generate tension, while adipogenesis is promoted by conditions that restrict cell spreading and contractility. Interestingly, the traction force generated by an individual cell after short-term exposure to bipotential media serves as an indicator of its ultimate differentiation propensity: high contractility has been associated with osteogenic potential, while low contractility has been associated with adipogenic capacity.[66] Thus,

single cell or clonal heterogeneity in the ability of cells to generate traction may correspond to the functional variability observed.

Variation also extends to cellular activity along the pathways responsible for mechanotransduction. For example, calcium signaling is highly mechanosensitive, and regulates processes including differentiation and proliferation.[142] Investigations of baseline calcium signaling in MSCs have shown that some cells exhibit spontaneous calcium oscillations, while others do not.[101,104,208] The extent of variation is microenvironment dependent: the fraction of MSCs experiencing calcium oscillations increases with substrate stiffness, with 59-98% of cells oscillating on glass culture surfaces.[101,104,105,208] Furthermore, this variation extends to the cellular response to active mechanical stimulation. Subpopulations of undifferentiated MSCs encapsulated in hydrogels differentially respond to compression: calcium signaling was upregulated in one subset of cells, while it was downregulated in another.[23] It remains to be seen if similar variation extends to other mechanosensitve pathways, including MAPK and RhoA/ROCK signalling.

Intra-clonal heterogeneity and cell-to-cell variation

Studies investigating inter-clonal heterogeneity often imply that clonal subpopulations are relatively homogeneous. However, there is growing appreciation that even within a clone, cellular phenotype can be highly variable.[180] Cells within a clone can differ in their morphology and ability to differentiate. When intact colonies are exposed to adipogenic or osteogenic differentiation cues, differentiation initiates in the dense, inner portion of the colony.[233] Colony microdissection and subsequent analysis reveal spatial differences

in gene and protein expression. Cells located in the colony interior expressed extracellular matrix genes to a greater extent than cells located in the outer periphery, while 'outer' cells expressed higher levels of genes associated with cell proliferation and mitosis.[233] Building on this analysis, technological advances have enabled the interrogation of single cell gene expression amongst clonal and polyclonal populations (a major focus of this thesis). In parallel, single cell RNA sequencing of individual MSCs has shown that individual MSCs have variable basal expression of both early and late differentiation markers, and that markers of multiple lineages can be co-expressed in the same cell.[65] Furthermore, DNA methylation patterns are mosaic within individual clonal populations.[156,157] Thus, cell-to-cell variation exists at every level where inter-clonal heterogeneity has been noted.

Potential origins & mechanisms of MSC heterogeneity

It is tempting to speculate that MSC heterogeneity mirrors the diversity of environments present in the *in vivo* stem cell niche. [87] *In vivo*, MSCs reside in niches characterized by diverse cellular communities that present variable chemical and mechanical conditions. Indeed, microanatomical heterogeneity within the bone marrow niche has been shown to dictate cell-to-cell variation in osteolineage cells.[198] Upon isolation, MSCs from these heterogeneous environments mix together, and extant *in vivo* variation may persist into *in vitro* cultures. Indeed, there is mounting evidence that cultured cells retain 'memory' of their previous environments.[124,230] Furthermore, the mechanical properties of the stem cell microenvironment influence self-renewal capacity and regenerative potential.[69]

Perhaps then, the mechanisms responsible for in vitro cellular memory may also facilitate the maintenance of heterogeneity in primary cultures.

In addition to any heterogeneity derived from the *in vivo* niche, there is also apparent plasticity in MSC phenotype. In the study of intra-clonal spatial variation discussed above, the subculture of either 'inner' or 'outer' cells yielded new colonies with their own distinct inner and outer populations, suggesting that, in at least some dimensions, cellular variation is dynamic and reversible.[233] Potential dynamics of shifting phenotypic variability have been carefully studied in the context of hematopoietic stem cell (HSC) lineage commitment towards erythroid and myeloid fates. Clonal HSC populations heterogeneously express the surface protein Sca-1, a marker associated with the erythroid transcriptional signature. Subcultures of cells sorted for either the lowest or highest Sca-1 expression shift with time to reconstitute the original distribution of Sca-1 expression.[28] Similar behavior is observed in MSCs, which express Sca-1 heterogeneously between and within clonal populations.[78] At early passage, MSC fractions with either low or high Sca-1 expression were able to regenerate the distribution of Sca-1 expression in the parent population.[78] However, after extended passaging, sorted MSCs were less able to reconstitute the full range of Sca-1 expression.[78] Potential explanations of this behavior include spontaneous transcriptional fluctuations (of either transcriptome-wide programs[28] or individual regulators[166]) and epigenetic bistability[78].

Transcriptional fluctuations in the expression of individual genes might arise from the stochasticity inherent to many biological processes. While transcription at the population

level is often considered a process that proceeds at a constant, defined rate, transcription in individual cells is highly stochastic. Fundamentally, transcription requires the chemical interaction of RNA polymerases with an accessible promoter sequence and any requisite transcription factors. Thus, even if two stem cells were identical in every way, the transcriptional processes in each would be dictated by the random collisions of molecules within the nuclear milieu. The importance of such probabilistic interactions was elegantly shown in a now-classic experiment where two distinguishable yet near-identical genes were inserted into a cell.[53] Within individual cells, the expression of these two genes deviated, suggesting the existence of "intrinsic" random noise in gene expression. Notably, this intrinsic noise was superimposed upon any cell-to-cell variation controlled by "extrinsic" factors (e.g. epigenetic differences between cells, cell size, etc. – many of the types of variation described in the previous sections). Even so, intrinsic noise can give rise to substantial variation in copy number, and may drive cellular decision-making and phenotypic divergence.[6,175]

Population dynamics also likely contribute to the variation that emerges as the stem cells proliferate in culture. Upon cell division, a single, self-renewing cell splits into two daughter cells of approximately equal size. If the division is symmetrical, both daughter cells will possess the same self-renewal capacity: they will either both divide, or not divide. In contrast, asymmetric cell division will yield one self-renewing cell, and one that senesces in culture. Such dynamics allow an initially small fraction of cells to give rise to the majority of the population several days later. For example, one study reported that after 6 days of culture, 50% of progeny cells were derived from 9% of the initial

population.[226] Thus, much of the cell-to variation observed in polyclonal populations may actually be heterogeneity within a single dominant clone. Longer term tracking of MSC lineages over 12 passages confirms this notion of clonal dominance.[193] Interestingly, initially dominant clones were sometimes overtaken by other clonal subpopulations.[193]

This delayed dominance may be related to variation in the onset of senescence. With extended passage, MSCs suffer from decreased multilineage potential. [46,110,185,188] In parallel, the clonogeniticty of MSCs decreases and proliferation slows.[46,188] Functional capacity may also grow increasingly restricted with progressive culture. For example, the hierarchical lineage commitment hypothesis posits that through divisions, stem cells progressively lose the ability to commit to certain lineages. Alternatively, MSC functional heterogeneity may also be explained in part trans-differentiation, or transitions between partially restricted differentiation capacties.[164] The existence, structure, and governance of any such hierarchy or trans-differentiation processes remain to be elucidated.[164]

Notably, spontaneous genetic mutations are not thought to be the source of cell-to-cell variation amongst MSCs. Estimates of the mutation rate that would be required to give rise to the observed diversity are unfeasibly high: approximately one in three cells would need to experience a phenotype-altering mutation.[180] While such rates are possible, they are improbable, and thus genetic mutation is unlikely to be a dominant mechanism in the evolution of *in vitro* MSC heterogeneity. [180]

Measuring cell-to-cell molecular variation

Studies to discern the underpinnings of stem cell heterogeneity increasingly rely on methods to assay the molecular content of individual cells. To this end, adaptations of traditional methods and new techniques now allow one to assess gene and protein expression at the single cell level. Broadly, these approaches can be classified on the basis of their timing (continuous observation vs. fixed endpoint), their modality (imaging based vs. lysate based), and their ability to support multiplexed observations (high vs. low). In this section, we highlight select methods that may be of particular utility for assessing stem cell heterogeneity; a summary is provided in **Table 2-1**.

Table 2-1: Single cell methods to assay mRNA and protein abundance.

Method	Timing	Modality	Ability to multiplex
mRNA abundance			
RNA FISH	Endpoint	Image based	Low
Single cell qPCR	Endpoint	Lysate based	Intermediate
Single cell RNA sequencing	Endpoint	Lysate based	High
Molecular beacons	Continuous	Image based	Low
Spherical nucleic acids	Continuous	Image based	Low
Transgenic fluorescent reporters	Continuous	Image based	Low
Protein abundance			
Immunostaining	Endpoint	Image based	Low
Flow cytometry	Endpoint		Intermediate
Mass cytometry	Endpoint		Intermediate
Proximity assays	Endpoint	Lysate based	Low
Single cell western blot	Endpoint	Lysate based	Low
Single cell mass spec	Endpoint	Lysate based	High
Transgenic fluorescent reporters	Continuous	Image based	Low

Assaying gene expression in single cells

Single molecule RNA fluorescent *in situ* hybridization (FISH) is an imaging-based method that quantifies the absolute numbers of mRNA in fixed cells. Sets of

fluorescently labeled oligonucleotide probes tile along individual mRNA molecules in a sequence-specific manner, allowing mRNA molecules to be visualized as diffraction-limited spots (**Figure 2-3A**).[62,174] In standard RNA FISH, the number of genes simultaneously assayed is restricted by the availability of microscope filter sets (approximately 4 genes). However, recent strategies utilizing combinatorial and sequential barcoding have substantially increased the potential number of genes measured.[128,129]

Other endpoint methods for assaying single cell gene expression include single cell RT-qPCR, microarrays, and RNA-seq.[187,203] Each of these approaches measures the abundance of cDNA amplified from mRNA in the lysates of individual cells. The ability of single cell RNA-seq to report transcriptome-wide expression holds particular promise as the field develops and increased understanding of the many dimensions of cell-to-cell variation. However, these methods also require normalization (vs. absolute quantification), and cannot provide information about the spatial distribution of individual mRNA within an individual cell.

Alternative techniques allow gene expression to be monitored continuously in live cultures through imaging. Strategies such as molecular beacons and spherical nucleic acids (e.g. nanoflares) rely on oligonucleotides, fluorophores, and fluorescent quenchers to report gene expression (**Figure 2-3B,C**).[191,218] When the oligonucleotide probe binds to a target mRNA, the quencher separates from the fluorophore and permits fluorescent signal. These methods have been used to monitor gene expression in live MSCs

undergoing osteogenesis, and can be combined with FACS techniques to sort individual cells on the basis of their gene expression.[123,138]

Other efforts to monitor gene expression in real-time rely on transgenic methods. Expression of genes modified to include repetitive stem-loop motifs can be monitored using fluorescent bacteriophage proteins that bind to these sequences with high affinity (**Figure 2-3D**).[67,197] Alternatively, short-lived fluorescent reporter proteins have been considered as proxies for the expression of of genes the control of the same promoter.[209]

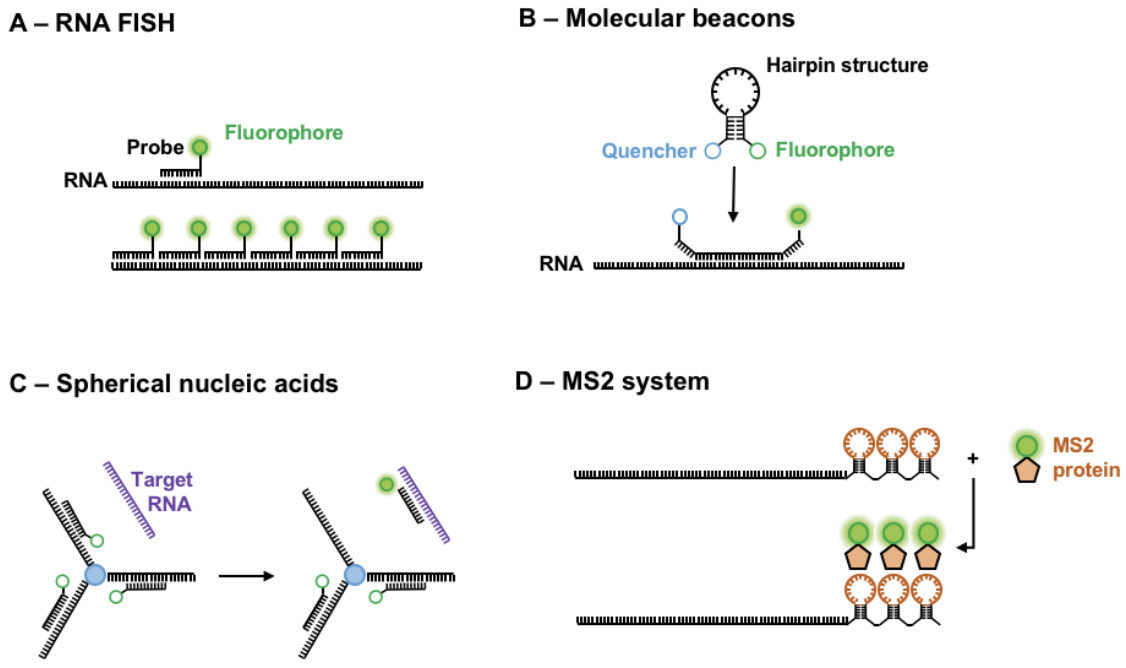

Figure 2-3: Schematics of select methods to measure mRNA in single cells. A) In RNA FISH, fluorescently labeled oligonucleotide probes tile along the target mRNA. B) Molecular beacons emit fluorescence upon binding to target mRNA. C) Spherical nucleic acids quench fluorescence until target mRNA binding occurs. D) In the MS2 system, fluorescently tagged proteins bind to motifs engineered into the mRNA sequence. Adapted by permission from Macmillan Publishers Ltd: Nature Cell Biology, Hoppe et al.,[85] copyright 2014.

Assaying protein expression in single cells

In addition to quantifying mRNA at the single cell level, it is also essential to map and measure the protein output from this message content on a cell-by-cell basis. Single cell measurements of protein expression are possible using a variety of techniques, many of which rely on antibody-based detection. The simplest of these is standard immunostaining, imaged at high magnification and quantified on a per-cell basis. Flow cytometry offers high throughput measurements of fluorescent antibody signal, and can be coupled with cell sorting and multiplexed to accommodate the measurement of 10-15 proteins.[202] Mass cytometry allows further multiplexing by leveraging mass spectroscopy to detect the levels of metallically-conjugated antibodies bound to individual cells.[202] Recently, imaging mass spectrometry has further extended this approach to enable the measurement of protein abundance in histological sections while preserving spatial information.[29]

An additional category of assays are those based on proximity, including proximity ligation (PLA) and proximity extension (PEA). In these approaches, pairs of antibodies conjugated to oligonucleotides are used to probe cell lysates.[75] When an antibody pair binds to the protein of interest, the two oligos are brought together and either ligated (PLA) or hybridized (PEA) to create a template for the synthesis of reporter DNA that is ultimately detected via qPCR or sequencing.[75] Interestingly, this method is compatible with lysate-based assays of single cell gene expression, and has recently been used to simultaneously examine the proteomic and transcriptomic state of single cells.[44]

Techniques for assaying protein expression in bulk lysates have recently been adopted to accommodate single cells. Western blots can be performed on individual cells that have settled into microwells in a polyacrymaide gel; the cells are lysed in their wells immediately prior to in-gel electrophoresis, blotting and detection.[88,98] There is also active development surrounding single cell mass spectrometry, which was recently used to quantify the abundance of thousands of proteins at the single cell level during embryonic stem cell differentiation.[18]

Conclusions & Outlook

Mesenchymal stem cells demonstrate many dimensions of heterogeneity: they differ between donors, as well as between and within clonal populations. Indeed, cell-to-cell variation seems inherent to the cell type, and we speculate that researchers will find heterogeneity wherever they look for it. The emergence of genome-wide single cell techniques holds the potential to identify new molecular targets that vary between cells and correspond to differences in functional potential. However, as our technological ability to interrogate biology at the single cell level grows, we will need to distinguish biological noise from variation that represents actionable signal. Our ability to discern such signal may be enhanced by the choice of measurement approach. For example, biological noise often has a temporal component (e.g. stochastic gene expression fluctuations). In these situations, endpoint measurements may reveal substantial variation, even if time-averaged behavior is similar between cells. In contrast, continuous measurements could be integrated or averaged over time to potentially smooth stochastic temporal fluctuations while retaining evidence of major cell-to-cell variation.

Regardless of measurement technique, the further study of MSC variation holds the potential to clarify the mechanisms and implications of cell-to-cell heterogeneity.

Chapter 3: Morphologic and mechanobiologic heterogeneity in clonal mesenchymal stem cell populations

Introduction

In vivo and *in vitro*, MSC fate commitment is strongly regulated by soluble chemical stimuli. Tissue engineering applications have leveraged this biochemical responsivity via the application of exogenous growth factors to direct MSC differentiation.[217] However, MSCs are notoriously heterogeneous and only a fraction of a given MSC population will respond to differentiation cues as intended.[164] Such functional heterogeneity complicates the efficient and widespread use of MSCs in clinical therapies.[86] Efforts to sort MSC populations via surface marker expression are limited by low yield as well as the lack of definitive markers that identify high-potency cells.[199] Recently, efforts to sort MSCs based on their biophysical properties (such as size and stiffness) have shown some promise in distinguishing MSC subpopulations predisposed towards specific lineages,[73,122] and suggest the existence of biophysically distinct subpopulations that may differ in their response to soluble differentiation cues.

Recent studies have also highlighted that the biophysical relationships between a cell and its microenvironment can influence cell fate.[89] These relationships depend not only upon the physical properties of the microenvironment, but also on the mechanical state of the cell. The cytoskeletal network provides a physical link to the extracellular environment and enables external physical cues to elicit biological responses.[89] For

example, mechanosensitive signaling cascades are initiated by force-induced changes to protein conformation[90], and extracellular force transmission to the nucleus modulates both intracellular signaling[50] and transcriptional activity[211]. Similarly, cytoskeletal contractility is requisite for force-induced changes in chromatin condensation in the nucleus.[82] Such condensation is a hallmark of gene silencing, and is typically associated with both differentiation and nuclear stiffening[161]. However, while mechanotransduction can influence cell fate, changes in differentiation status also modulate the mechanoresponsivity of the cell. Thus, the biophysical state of a cell is simultaneously a determinant and outcome of cellular differentiation.

Here, we build on the previous work linking biophysical properties of cells to differentiation potential, and evaluate mechanotransductive heterogeneity amongst clonal MSC subpopulations. First, we consider clone-to-clone variation in cellular morphology and their proliferative potential. Next, we examine how clonal variation persists in the context of expansion, with a particular focus on the nucleus (heterochromatic content and transmission of extracellular strain to the nucleus). Finally, we consider differences in how cells of differing clonal origins respond to the application of TGF to alter baseline contractile properties of the cells and how they construct a mechanically competent nascent matrix during early chondrogenesis in a 3D environment.

Results

Clonal variations in MSC proliferation and morphology

Clonal MSC populations were both morphologically and metabolically distinct from one another. Initial colonies differed in both the area occupied by the initial colony and cell density within that region (**Figure 3-1B**). Upon isolation, measurement of population doubling time showed wide variation between clones, ranging from 1.5 to 7 days (**Figure 3-1C**). In terms of morphology, some clones exhibited a classic spindle-like morphology, while others had multiple protrusions and lacked a clear polarization (**Figure 3-1D**). While cell area did not significantly differ between clones (**Figure 3-1E**), quantification of parameters such as cell aspect ratio, circularity, and solidity (**Figure 3-1F-H**) all supported the morphological differences we observed.

Clonal differences in nuclear morphology and heterochromatin content

In addition to these overall cell features, other intracellular metrics diverged between clones as well. For instance, DAPI staining showed wide variation in nuclear shape and morphology between clones. Moreover, for some clones, these nuclear features changed markedly with passage. For example, nuclear spread area increased dramatically in select clones (e.g. Clone C), but remained near constant with passage in others (e.g. Clone A, F; **Figure 3-2A,B**). Inside the nucleus, distinct morphological features could be observed as well. Staining for markers of heterochromatin showed wide variation amongst clones, and this too changed with passage number. For example, at passage 2, two of five clones (A and F) stained strongly for heterochromatin

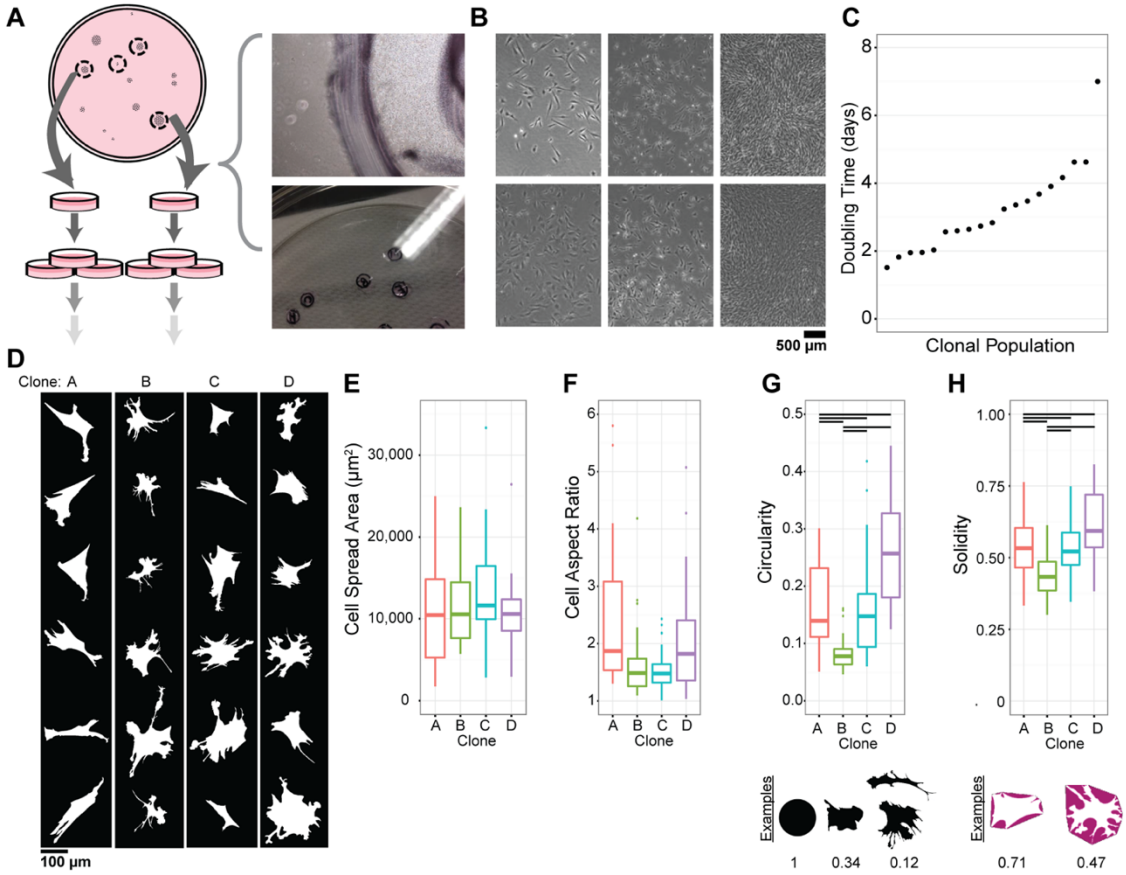

Figure 3-1: Clonal MSC populations differ morphologically and in their proliferation rates. (A) Schematic and photographs of the colony isolation process. Colonies formed on sparsely seeded plates, and were outlined before isolation. Top image shows microscopic view of an individual colony, lower images shows macroscopic view of a culture dish with multiple colonies identified. **(B)** Cell morphology and density in the initial colonies varied between clonal populations (each image taken from a different clonal population in the same plate). **(C)** Doubling time calculated between passages 1 and 2, for 19 clonal populations. **(D)** Cell silhouettes obtained from phalloidin-stained, passage 1 MSCs seeded on glass coverslips (scale: 100 µm). **(E - H)** Quantitative cell shape metrics of the four clones pictured in (D): cell spread area, cell aspect ratio, circularity and solidity (n = 30/condition; bars indicate $p < 0.05$ when compared via one-way ANOVA with Bonferroni post hoc test). For circularity and solidity, example silhouettes and corresponding quantification are included below the graphs.

(**Figure 3-2A,C**). By passage 4, heterochromatin staining intensity was attenuated across all clones and was more spatially uniform within individual nuclei, as indicated both visually (**Figure 3-2A**) and by the decreased variance of staining within individual cells (**Figure 3-2D**).

Clonal variation in cytoskeletal-to-nuclear strain transmission

Since mechanotransduction is influenced by the ability of a cell to transmit strain to the nucleus and the mechanics of the nucleus itself, inter-clonal differences in these nuclear metrics might also indicate the existence of mechanobiological heterogeneity in these cell populations. To interrogate clonal differences in cytoskeletal-nuclear strain transmission, we cultured clonal and polyclonal MSCs on nanofibrous scaffolds and monitored the deformation of the nucleus in response to applied strain. Nuclear deformability was calculated as the post-strain nuclear aspect ratio, normalized to the pre-strain nuclear aspect ratio, with a higher value indicating greater nuclear deformation (**Figure 3-3A**). At passage 2, nuclear deformation was similar between clones (**Figure 3-3B**, $p = 0.37$). By passage 4, however, clones differed in their response to applied strain (**Figure 3-3B-D**, $p = 0.02$). Two of five clonal populations remained highly deformable between passage 2 and passage 4 (Clones A and B, $p < 0.05$). In contrast, nuclear deformation in the remaining three clones was significantly attenuated by passage 4 ($p \geq 0.42$).

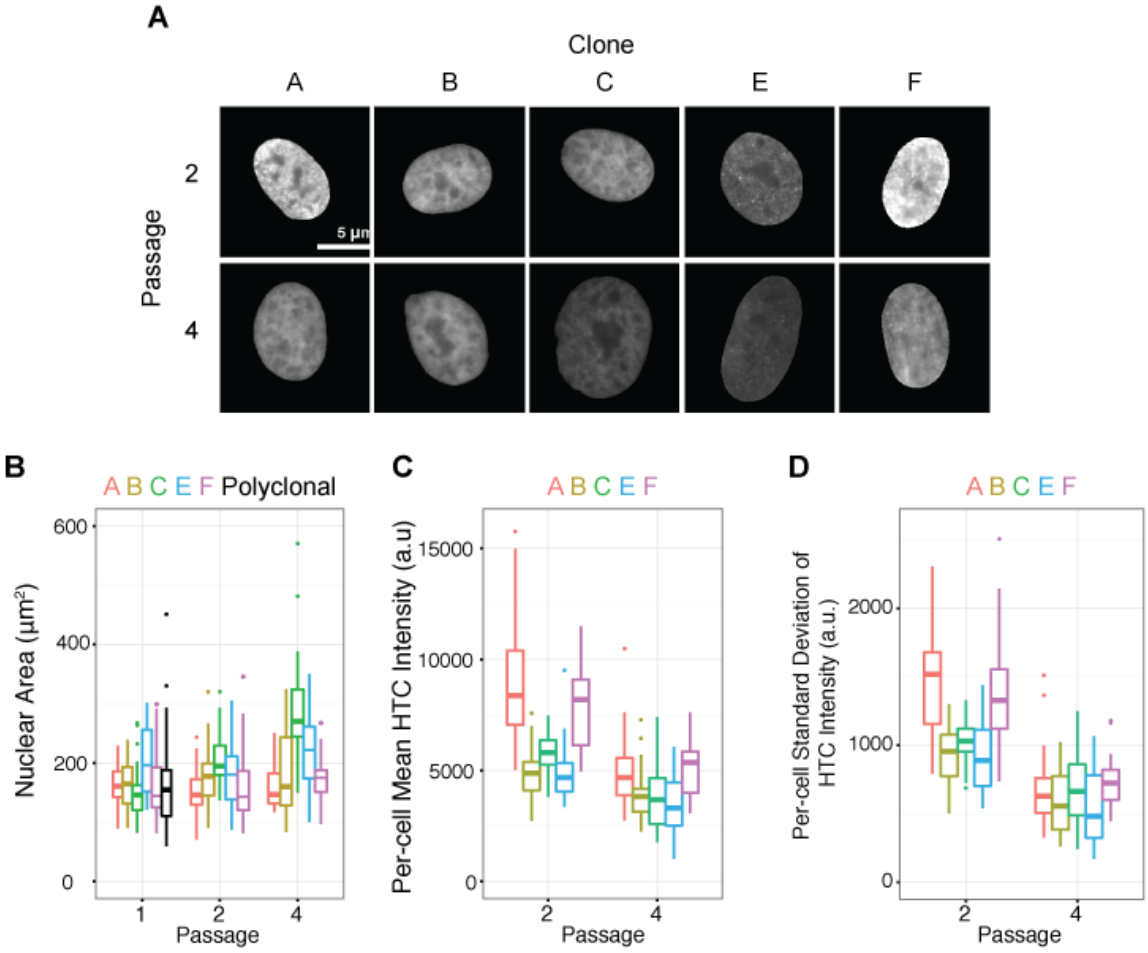

Figure 3-2: Passage-dependent clonal differences in nuclear morphology and chromatin organization. (A) Immunofluorescent staining for tri-methyl H3K27, an epigenetic marker of heterochromatin. **(B)** Projected nuclear area of polyclonal and clonal MSCs with passage. **(C-D)** Mean and standard deviation of tri-methyl H3K27 pixel intensity, calculated on a per-cell basis. (n = 21-60 cells/condition. Boxes denote the first and third quartiles. Whiskers extend to the most extreme observations within [1.5 * interquartile range] of the box hinges, and outlying observations are represented as individual points.)

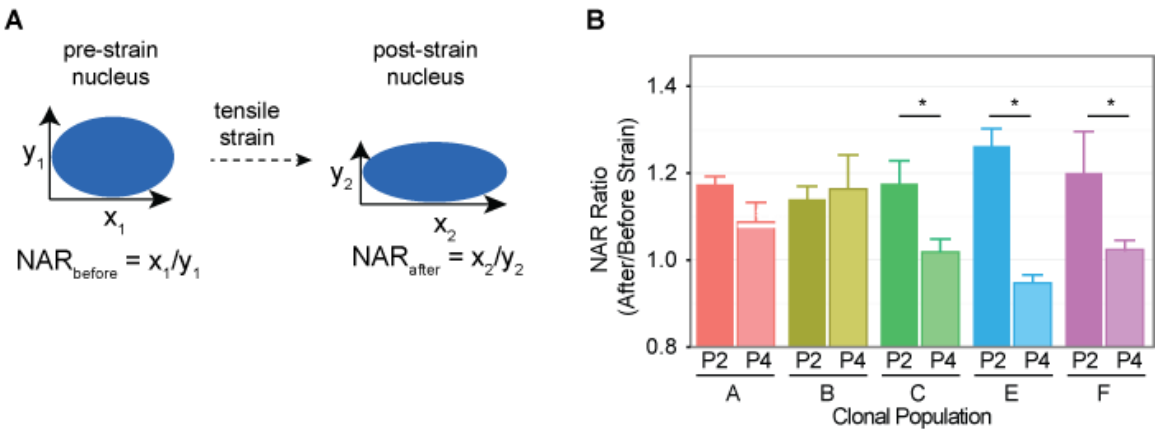

Figure 3-3: Passage-dependent clonal differences in nuclear deformability. (A) Schematic illustrating nuclear deformation and the calculation of nuclear aspect ratio (NAR). **(B)** Nuclear deformability in response to applied strain, measured as the ratio of NAR post-strain, to NAR pre-strain (n = 38-65 cells/condition, * indicates $p < 0.05$ P2 vs. P4 when compared via t-test with Bonferroni adjustment for multiple comparisons).

Cell-to-cell and clonal variation in contractility

Given their morphological differences at the cell and nuclear level, as well as their differences in strain transfer to the nucleus, we next hypothesized that clonal populations would differ in their baseline contractility levels. We have recently reported that baseline contractility in mesenchymal stem cells is required for nuclear deformation and ultimately regulation of chromatin condensation.

As a preliminary assessment, we first assessed MSC clonal heterogeneity in contractility using traction force microscopy. In this assay, the deformation of a hydrogel substrate is measured in the context of a cell before and after lysis in order to determine the cellular forces exerted on the substrate at a given point in time (**Figure 3-4A**). Consistent with our findings on glass, donor-matched clonal and polyclonal populations had similar spread areas on these hydrogels (polyclonal: 821 ± 121 μm^2, clone 1: 1226 ± 225 μm^2,

clone 2: 946 ± 96 μm^2; mean ± SEM, p = 0.17). Despite these comparable cell areas, however, clone 2 had an average traction stress that was 1.6x greater than that of clone 1 and the polyclonal population (**Figure 3-4B**; polyclonal: 274 ± 34 Pa, clone 1: 265 ± 37 Pa, clone 2: 451 ± 58 Pa, $p < 0.05$). This suggests that clonal subpopulations establish different levels of baseline contractility, and that these differences are maintained through at least two passages.

Given that our previous work had shown contractility was in part established through TGF/BMP basal signaling (in the absence of ligand),[81] we were curious as to whether the addition of these factors would increase traction levels to a maximum level that would be the same across clones. To test this, traction forces generated in a polyclonal MSC population were assessed using a higher throughput micropost array detector system (μPADS), with cells analyzed both before and with time after exposure to TGFβ (**Figure 3-4**C).[80] Prior to the addition of TGFβ ($t \leq 0$, **Figure 3-4C**), the baseline strain energy of individual cells varied between 55 and 540 mJ (168.5 ± 128.9, mean ± s.d., n = 15), consistent with our previous observations of marked heterogeneity across clonal populations using gel-based TFM measures. Upon addition of TGFβ, the average contractility of the population increased (**Figure 3-4C**, black line profile). Quite interestingly, we observed that the contractile response induced by TGFβ was heterogeneous at the single cell level. That is, some cells became more contractile, while others failed to respond at all (red and green profiles, respectively). This response (or lack thereof) did not seem depend on the initial contractile state of the cells. Taken together, these data suggest that individual MSCs and clonal populations differ in how

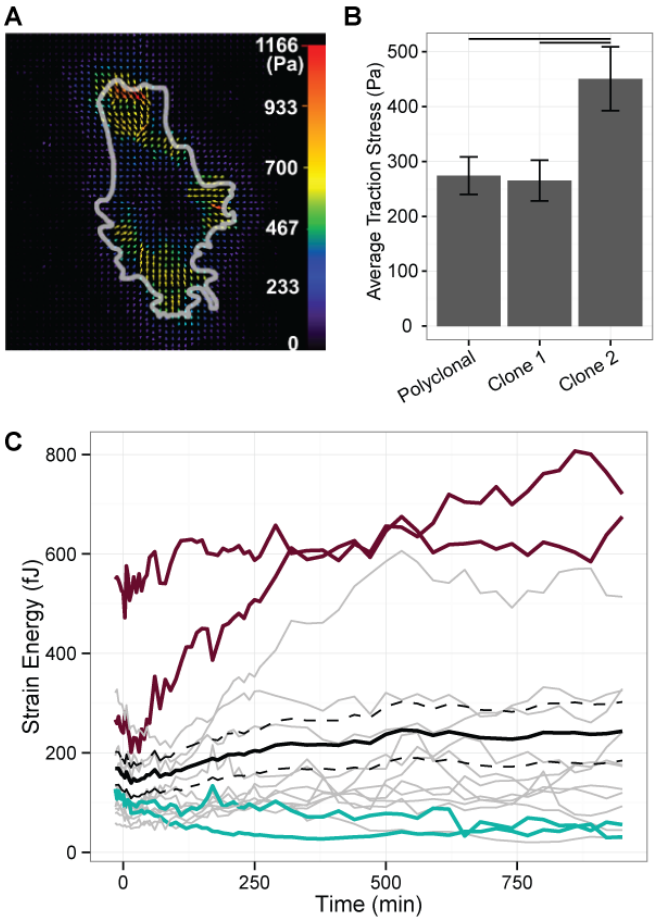

Figure 3-4: Traction force exerted by cells differs between individual cells and amongst clonal populations. (A) Traction force map of an individual MSC cultured on a polyacrylamide gel. **(B)** Mean traction stress per cell, for donor-matched polyclonal and clonal populations (n = 12-28 cells/condition, mean ± SEM, bars indicate $p < 0.05$ when compared via Mann-Whitney-U tests across groups). **(C)** Contractility of individual polyclonal MSCS, assessed on microposts before and after addition of TGFβ at t=0. Mean population response shown in black, with dashed lines indicating ± SEM. Individual cell traces are shown in grey, red, and teal.

they establish their baseline cellular traction stress generation as well as how their contractile state changes in response to agonists of contractility (including chondrogenic factors).

Micromechanics of the cell and nascent pericellular matrix

The combined effects of clonal heterogeneity are likely to culminate in clonal differences in functional capacity. Within a chondrogenic context, functionality is defined by the ability of cells to produce a robust extracellular matrix capable of resisting mechanical compression. Here, we evaluated this capacity by monitoring how cells encapsulated in agarose deform in response to 30% compression of the gel, and how this response changed with culture duration (**Figure 3-5A,B**). As a control we examined native chondrocytes, which demonstrated mean deformability ratios of 1.47 and 1.14, after 1 and 8 days of culture respectively. The effect of culture time was slightly attenuated in polyclonal MSCs, which demonstrated mean deformability ratios of 1.39 and 1.25 at days 1 and 8 respectively, as well as a donor-dependent effect (**Figure 3-5C**). Within the nine clonal populations examined, Day 8 deformability ratios reflected wide clonal variation and ranged from 1.17 to 1.40. Despite the stark differences in mean deformability, the variation around the mean was similar across populations. The evolution of pericellular mechanics between Day 1 and Day 8 also differed between clones (**Figure 3-5D**). Notably, select MSC clones and the donor-matched (Donor B) polyclonal populations did not show a significant change in deformability with time in culture ($p > 0.05$). In contrast, many individual clonal subpopulations showed significant decreases in this metric ($p < 0.05$), reflecting the progressive development of mechanically robust matrix. Corresponding Alcian blue staining showed that populations with lower deformation had greater pericellular accumulation of sulfated proteoglycans (**Figure 3-5E**).

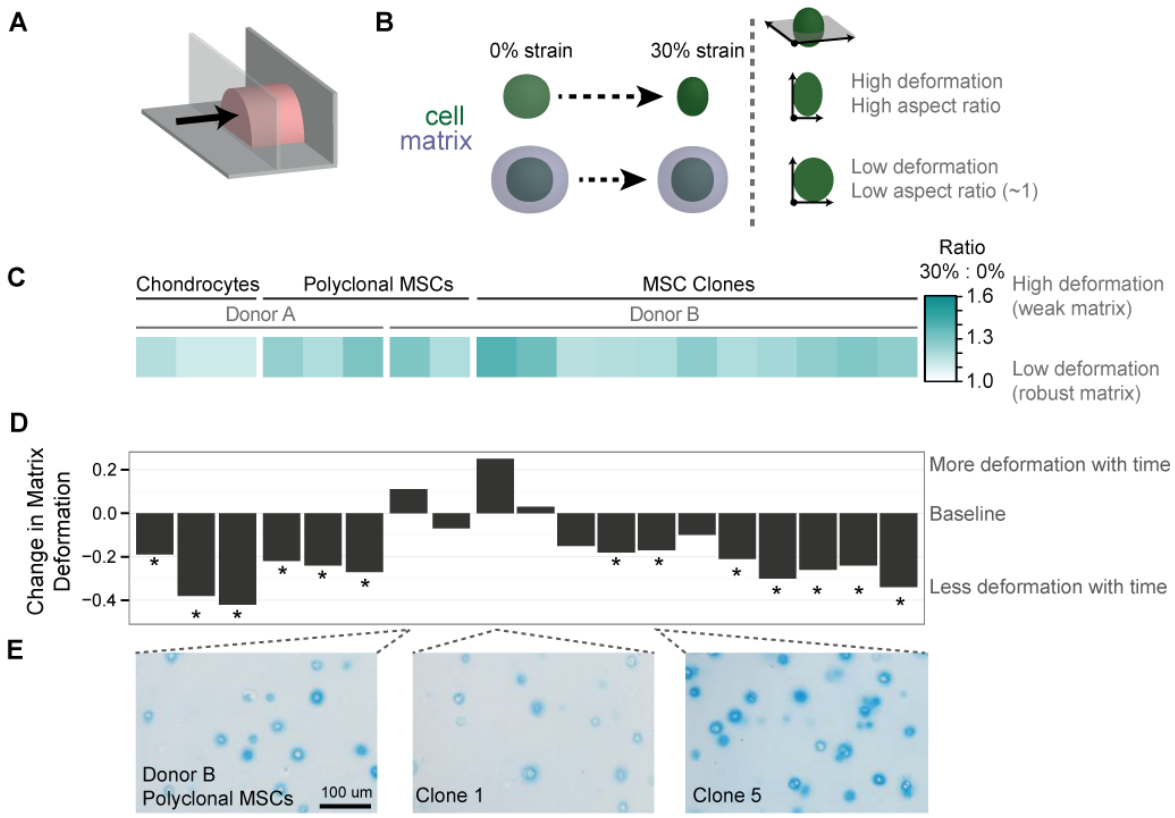

Figure 3-5: Chondrogenically-induced clonal populations differ in their response to compression. (A-B) Schematics illustrating microcompression to assess nascent extracellular matrix mechanics. (A) Illustration of gel orientation during compression. (B) Cell response to compression, measured by cell aspect ratio, serves as a proxy for the mechanics of the newly formed pericellular matrix. **(C)** Cellular deformability of donor-matched chondrocytes, polyclonal MSCs, and clonal MSCs after 8 days of culture, represented as post-compression cell aspect ratio normalized to pre-compression cell aspect ratio. (n = 3 fields of view/condition.) **(D)** Corresponding change in post-compression cell aspect ratio, between days 1 and 7. (same populations as C, * indicates $p<0.05$ day 1 vs day 8 when cellular deformability is compared via two-way ANOVA.) **(E)** Representative staining with Alcian blue, showing heterogeneous pericellular staining in polyclonal and clonal MSC populations.

Discussion

Polyclonal MSC populations isolated using standard techniques display heterogeneous chondrogenic potential.[167] Individual clonal populations differ in their ability to differentiate, their molecular fingerprint, and their mechanical properties.[43,122,164] To further delineate the variation present amongst clonal MSC populations, this work entailed a series of experiments targeted towards understanding the mechanobiologic facets of MSC heterogeneity. Specifically, our data show marked clone dependent variation in cell and nuclear morphology, nuclear chromatin condensation, nuclear deformability, contractility, response to contractile agonists, and establishment of a mechanical microenvironment.

We found that clonal MSC populations differed in their cellular morphology and ability to proliferate, corroborating previous reports.[153] Morphological diversity extended to the nucleus, with clonal populations also differing in their sub-nuclear organization. These differences in chromatin condensation support findings of epigenetic differences between clones.[156,157] Intriguingly, this heterogeneity attenuated with passage. Namely, staining to mark chromatin condensation revealed differences between clones at passage 2. However, with passage, clonal populations with high heterochromatin staining appeared to lose or reorganize H3K27me3 marks, and by passage 4, H3K27me3 patterns were similar between the five clones we examined. This finding builds on a previous report that H3K27 tri-methylation is stable with passage,[78] and newly identifies the clonal dependence of this behavior.

Measurements of nuclear deformation in response to applied strain exhibited similar passage and subpopulation dependence. Such attenuation of nuclear deformability could reflect either a change in nuclear mechanics or in the transmission of strain to the nuclear envelope. While we did not measure nuclear mechanics directly, we considered heterochromatin organization as a proxy for nuclear stiffness.[40] There was no consistent relationship between chromatin condensation status and nuclear deformability, suggesting that either changes in nuclear deformability are driven by altered strain transmission (i.e., nuclear connectivity), or that the mechanisms underlying decreased strain transmission are also clone dependent (e.g. select clones have altered nuclear mechanics, while others have altered strain transmission). Indeed, a challenging aspect of studying variation amongst clones is that a single outcome may have multiple potential causes, and different causes could drive phenotype in individual clonal populations.

The passage-dependent effects observed in nuclear morphology and deformability also highlighted that not all differences are preserved with expansion. This may reflect increasing intra-clonal diversity as the the cells in a clonal population reconstitute the heterogeneity of the parent polyclonal population (similar to how heterogeneity has been observed to reemerge in sorted polyclonal populations[28,78]). Alternatively, mechanisms such as mechanical memory may drive the apparent homogenization. For example, extended culture on stiff surfaces may override initial differences in clonal mechanobiological phenotype, and force clones towards to exhibit similar behavior with extended exposure to this modifier of their baseline mechanobiology.

Inter-clonal variation also extended to the clonal response to soluble factors. Cellular contractility differed between clones, potentially contributing to variation in nuclear pre-stress (and hence the observed differences in nuclear area and deformability). Furthermore, clonal populations exposed to TGFβ constructed extracellular microenvironments with variable mechanical properties. This finding provides a mechanical perspective on functional variation amongst clones. In cartilage, matrix quality is a key measure of differentiated phenotype: chondrocytes surround themselves with a pericellular matrix, which not only allows the tissue to bear compressive load but also mediates mechanical strain transfer from the tissue to the cell[76]. Previous studies have noted that within sparsely seeded agarose hydrogels similar to those used here, both chondrocytes[107] and polyclonal MSCs[222] develop pericellular matrices. In this study, we employed similar micromechanical techniques to identify clonal variation in pericellular matrix production and mechanics. In parallel, proteoglycan staining revealed intense staining localized in a compact manner around some clonal populations, but a lighter, more diffuse pericellular staining in others, indicating they had produced less matrix or ECM molecules of different molecular weights, sulfation levels (charges), and diffusivity. These observations further underscore the notion that clones differ in their ability to differentiate, and also establish that cells from different clones become situated in very different microenvironments with different mechanical properties. Any such micromechanical variation may feed back to the cells through mechanotransductive mechanisms and further drive divergence of phenotype.

Collectively, these findings add to the growing appreciation of clonal MSC heterogeneity. While it would be ideal to ultimately identify parameters that could be used to prospectively isolate the cells most capable of undergoing differentiation, any such effort may be complicated by the passage dependent changes described here. An alternative strategy may be the mitigation of heterogeneity – for example, can the combination of multiple orthogonal stimuli increase the percentage of MSCs exhibiting a desired response? By more fully understanding how clones respond to a variety of stimuli (e.g. active mechanical stimulation, mechanotransductive alteration, other growth factors), it may be possible to optimize the potential of polyclonal populations.

Methods

Cell isolation & culture

Cells were isolated from juvenile bovine tissue (Research 87, Boylston, MA). Chondrocytes were isolated from tibial plateau cartilage via a series of pronase and collagenase digestions. MSCs were isolated from femoral and tibial bone marrow. Marrow from matched tibia/femur pairs was plated over four dishes measuring 592 cm^2 in combined area, and cultured in basal media consisting of high glucose DMEM with 10% FBS and 1x antibiotic-antimycotic. MSC cultures were monitored for the formation of clearly demarcated cell colonies, which were isolated and sub-cultured 10-14 days after marrow isolation.

Colonies were considered as single-cell derived clonal populations, and were isolated via the trypsin drop method modified from Bartov *et al*. (**Figure 3-1**A).[8] Individual colonies were identified via bright field observation at 4x magnification, and a 7.5 mm diameter circle surrounding their location was marked on the outside of the culture plate using adhesive rings or permanent marker. The plate was then washed with phosphate buffered saline (PBS) and a cell scraper was used to remove cells outside of the identified colonies. Following PBS aspiration, a surgical spear was used to dry the plate surrounding each colony. A droplet of trypsin was added to each colony and held in place by surface tension. Colonies were incubated at 37°C for 2-5 min, gently agitated with a pipette, and transferred to either 6-well or 24-well culture plates, depending on colony size. In parallel, donor-matched polyclonal populations were obtained by trypsinizing cells from an entire plate using standard techniques. Following this initial passage (passage 1), both clonal and polyclonal populations were passaged when they reached ~80% confluence; cells were counted at each passage. At each replating, cells were reseeded at 5,000 cells/cm^2 in basal media for continued proliferation. Population growth rates were estimated from the cell counts taken at passages 2 and 3.

Subsequent experiments utilized chemically defined media, consisting high-glucose DMEM supplemented with 1 × antibiotic–antimycotic, 40 ng ml^{-1} dexamethasone, 50 µg ml^{-1} ascorbate 2-phosphate, 40 µg ml^{-1} L-proline, 100 µg ml^{-1} sodium pyruvate, 1.25 mg ml^{-1} bovine serum albumin, 5.35 µg ml^{-1} linoleic acid and 1 × insulin–transferrin–selenous acid premix (Corning CB-40350), either with or without 10 ng ml^{-1} TGFβ3 (R&D Systems, CM+ and CM- respectively).

Fluorescent imaging and quantification

At each passage, cells were seeded on collagen I-coated glass coverslips at a density of 3000 cells/cm^2 in CM+ and fixed with 4% paraformaldehyde after one day. To assess cell morphology, cells were stained with Alexa Fluor 488 phalloidin (Invitrogen, #A12379) and 4',6-diamidino-2-phenylindole (DAPI, ProLong Gold Reagent, ThermoFisher, #P36931). Five 20x fields of view were selected randomly and imaged using a fluorescent microscope (Nikon T30). Cells were manually segmented and morphology was quantified using ImageJ (n=30 cells/condition). Cell solidity was calculated as the ratio of cell spread area to convex hull area. Nuclear morphology and heterochromatin organization were assessed with DAPI and staining for histone H3 tri-methylation at K27 (1°: mouse monoclonal H3K27 trimethyl, Abcam ab6002; 2°: anti mouse IgG-TRITC (T2402, Sigma). Images focused around the nucleus mid-plane were taken using a fluorescent microscope (Zeiss Axioplan-2), and a custom Matlab script was used to segment individual nuclei using the DAPI signal and measure the corresponding intensity of H3K27 tri-methyl staining (n = 21 - 36 cells/condition).

Micropost and traction force microscopy measures of contractility

Traction force microscopy was performed as previously described.[50] Briefly, polyacrylamide hydrogels (5 kPa, v=0.45) containing 0.2 μm fluorescent microspheres (Invitrogen F8810) were prepared,[5] treated with 2 mg/mL sulfo-SANPAH (Pierce 22589), coated with fibronectin (20 μg/mL), and UV sterilized. Passage 1 MSCs were seeded onto the gels in basal media at a density of 3000 cells/cm^2 and allowed to attach for 20 hours. Phase contrast and fluorescent images of live cells and beads were acquired

before and after cell lysis with SDS buffer using a microscope equipped with live cell imaging chamber (DeltaVision Deconvolution). Image analysis was performed using the freely available plugin suite for ImageJ (n = 12- 25 cells/condition).[216]

The cellular contractile response to the addition of TGFβ was assessed using a custom micro-post array detector (mPAD).[66,231] MSCs were cultured on electrospun scaffolds for 7 days, re-seeded on the micropillars (spring constant: 18.19 nN/mm, post tips labeled wih DiI), and cultured for 1 day prior to observation. Micropillar deflection was then monitored using a fluorescence microscope equipped with a live-cell imaging chamber. Using measured post deflection, strain energy per cell was calculated as a function of time.[231]

Tensile stretch to assess nuclear deformation

At passage 2 and passage 4, polyclonal and clonal MSC populations were seeded onto aligned electrospun poly(ε-caprolactone) scaffolds, and cultured in CM- for one day. Next, live cells were stained with Hoechst (Sigma 33342) for 20 minutes at 37°C, and scaffolds were mounted into a custom stretching apparatus on a fluorescent microscope (Nikon T30).[50] Scaffolds were stetched from 0% to 15% grip-to-grip strain, in increments of 3%. Images of cell nuclei were collected at 0 and 15% strain (n = 20 - 43 cells/condition). Nuclei were automatically segmented using a custom MATLAB script and individual cells were tracked through each strain level. Nuclear aspect ratio (NAR, equal to the ratio of principle lengths) was calculated as the ratio of the long and short axes of each nucleus at each strain increment.

Compression to assess matrix micromechanics

Clonal and passage 2 MSCs were encapsulated in 2% agarose hydrogels. $3x10^6$ cells/mL were suspended in an agarose solution, and cast into cylindrical constructs measuring 4 mm in diameter and 2.25 mm in thickness. Hydrogels were cultured for 1 day in CM− or 8 days in CM+. On days 1 and 8, constructs were halved, and one half was stained with 4 μM calcein-AM in PBS for 30 min and tested using a custom microscope-mounted microcompression device.[107] The remaining half was fixed in 4% paraformaldehyde, paraffin embedded, sectioned to 8 μm, and stained with Alcian blue (pH 1.0) to histologically assess matrix accumulation.

Micromechanical testing was conducted using a custom unconfined compression testing device consisting of micrometer-controlled platens inside a glass-bottomed, PBS-filled well.[60,146] Constructs were placed in the PBS bath with the mid-sagittal plane downward and imaged at 0% and 30% compressive grip-to-grip stain using a inverted confocal microscope (Olympus Fluoview FV1000, 20X UPlanFL objective). Images were acquired through approximately 60 μm of the construct depth with a step interval of 2.34 μm per slice. Z-stack maximum intensity projections were binarized in MATLAB, and bounding boxes were computationally constructed around individual cells. Bounding box aspect ratio was calculated as the ratio of bounding box dimensions (Y/X). Parameters, including mean and standard deviation, were calculated for each image stack to measure the behavior of cells in each field of view (n=3 gels/condition). These parameters were then aggregated over each condition, and are presented as the condition mean ± standard deviation.

Statistics

Metrics describing cell morphology were compared between clones using one-way ANOVAs with Bonferroni post hoc tests. Average traction stress was compared between clones using Mann-Whitney U tests adjusted for multiple comparisons. Nuclear deformability was compared across clones and passage using a two-way ANOVA with Bonferroni post hoc tests for planned comparisons. Cellular deformability was compared across time (Day 1 vs Day 8) and clone using a two-way ANOVA with Bonferroni post hoc tests for planned comparisons.

Chapter 4: Single-cell differences in instantaneous matrix gene expression do not predict matrix deposition

Introduction

Despite the phenotypic heterogeneity in MSC populations, most studies that explore the molecular underpinnings of phenotype monitor differentiation via bulk assays of transcriptional state and protein synthesis averaged over an entire cell population. These ensemble measurements, by definition, mask population heterogeneity [3,155]. The advent of single cell methods allows for the measurement of cell-to-cell variation and the ability to quantify absolute gene expression in a single cell,[37,91,96] revealing, for example, marked transcriptional heterogeneity. Real-time fluorescent monitoring of changes in transcript levels in individual cells has also shown that individual MSCs differ in the timing and extent to which they upregulate an early osteogenic marker[138]. These findings underscore the limitations of coarse ensemble approaches and highlight the need for single-cell molecular profiling of these differentiation events. Although it is reasonable to speculate that the subpopulation of cells expressing high levels of marker genes would ultimately be the most chondrogenic, this hypothesis remains untested.

Given that individual MSCs are highly variable in their capacity to undergo chondrogenesis and accumulate cartilage-like matrix[86], we postulated that one could use single cell marker gene transcript levels as a means to enrich for MSC subpopulations most suited for therapeutic application. Here, we define this relationship by developing probe sets for RNA fluorescence in situ hybridization directed against transcripts of

markers of cartilage, bone, and fat, and use single-cell analysis to delineate the relationships between absolute transcript level and differentiated cell function. Specifically, we hypothesized that cells that robustly accumulate an aggrecan-rich, cartilage-like matrix would also express high levels of aggrecan mRNA, while at the same time suppressing markers of other lineages.

We find surprising levels of variability in the expression of aggrecan and other marker genes between individual MSCs both before and after differentiation. However, when we compare expression with functional capacity (defined by actual matrix deposition) on a single cell basis, we find a weak correlation between transcript abundance and protein expression. Transcriptome-wide analysis via RNA sequencing further suggests that neither an expanded set of marker genes, nor the principal components of global gene expression variation, correlate strongly with functional capacity. Indeed, even in fully differentiated chondrocytes derived from native tissue, absolute aggrecan mRNA expression is decoupled from cartilage-like matrix accumulation. Collectively, these findings suggest that sorting based solely on a small set of differentiation markers will not improve chondrogenic outcomes, and challenge the traditional notion that marker gene expression defines or is even strongly associated with phenotype.

Results

Single cells express differentiation markers heterogeneously

To quantify absolute gene expression of marker genes on a single cell basis during MSC differentiation and chondrocyte de-/re-differentiation, we paired classic cartilage tissue

engineering and cartilage biology experiments with single molecule RNA FISH[62,174]. Specifically, we monitored the simultaneous expression of aggrecan as a marker of chondrogenic differentiation, GAPDH as a reference gene, and osteopontin and lipoprotein lipase (LPL) as markers of alternate fates (osteogenesis and adipogenesis, respectively)[64,163,192]. For each gene, we designed fluorescently-labeled sequence-specific oligonucleotide probes to visualize individual mRNA molecules in intact fixed cells. Individual mRNA appeared as bright diffraction limited spots (**Figure 4-1a-b**) and subsequent spot counting yielded absolute copy number at the single cell level.

To show that our measurements corresponded well with existing measurements of these systems, we first determined how absolute gene expression changed as MSCs underwent chondrogenic differentiation. To do so, we formed engineered constructs and used RNA FISH to quantify gene expression over three weeks in chemically defined media with or without TGFβ (chondrogenic induction media and control media respectively, **Figure 4-1c**). As expected, chondrogenic induction promoted proteoglycan synthesis and matrix accumulation (**Figure 4-1d**) and increased aggrecan copy number in comparison to control media (**Figure 4-1e**). Though there was considerable donor-to-donor variability in mean aggrecan levels and matrix deposition, the trends were similar between donors, with mean aggrecan copy number generally increasing over the first 7 days, before decreasing at later time points (**Figure 4-1e, Figure 4-2a, Appendix 2 Supplementary Table 1**). Mean GAPDH copy number increased with exposure to induction media (**Figure 4-2b**). Thus, in aggregate, this RNA FISH analysis aligned with

the canonical understanding of gene expression changes during chondrogenic differentiation[170].

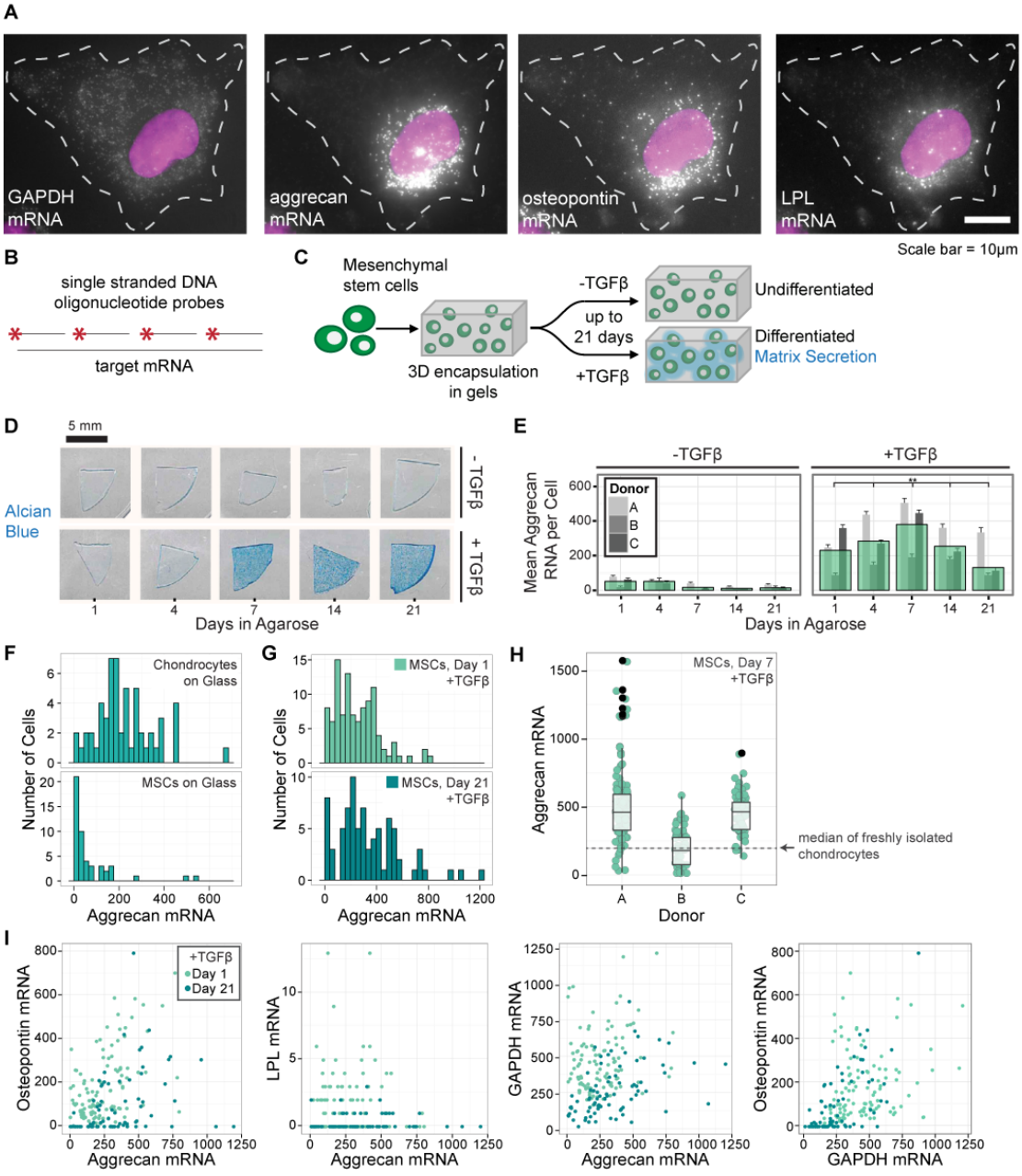

Figure 4-1 (previous page): RNA FISH reveals heterogeneity in lineage marker expression in mesenchymal stem cells and chondrocytes. a-b) Representative images (a) and schematic (b) of single molecule RNA FISH, in which fluorescently labelled DNA oligonucleotides enable quantification of absolute expression of multiple genes in the same cell. c) Chondrogenic induction scheme, involving cell encapsulation in 3D agarose constructs and exposure to TGFβ. d) Alcian blue staining for sulfated proteoglycans; Donor B shown. e) Mean aggrecan RNA counts in MSCs cultured in 3D for up to 21 days. Narrow bars represent the mean within an individual donor; overlaid bars represent the mean across donors. Error bars indicate standard error (n = 24-128 cells per donor and condition), ** indicates $p<0.01$ vs. –TGFβ conditions, and between +TGFβ time points. See Appendix 2 Supplementary Table 4 for all statistical comparisons. f-g) Distributions of single cell aggrecan expression for chondrocytes and MSCs plated on glass in basal media (f, n = 56 chondrocytes, 49 MSCs) and 3D encapsulated MSCs exposed to TGFβ for 1 and 21 days (g, n = 105 cells for day 1, 79 cells for day 21; Donor A shown.) h) Single cell aggrecan expression for each donor after 7 days of 3D culture with TGFβ relative to the median aggrecan expression in freshly isolated chondrocytes (dashed line; n = 103 cells for Donor A, 54 cells for Donor B, 65 cells for Donor C). i) Simultaneous expression of aggrecan, osteopontin, LPL, and GAPDH on day 1 and day 21; Donor A shown (n = 105 cells for day 1, 79 cells for day 21).

While these ensemble measures corresponded with previous findings, they did not provide information on cell-to-cell variability in expression of these lineage markers. Thus, we measured mRNA copy number on a cell-by-cell basis under baseline conditions and with differentiation. We assayed four conditions: naive MSCs in expansion culture, MSCs differentiating in engineered constructs after 1 and 21 days in induction media, and as a positive control, fully differentiated primary chondrocytes (**Figure 4-1f-g**). For each of these groups, single-cell analysis showed striking heterogeneity in expression, with aggrecan mRNA copy number per cell spanning three orders of magnitude (10^0 - 10^2). Consistent with the notion that stem cells exhibit greater variability than differentiated cells, naive MSCs showed the greatest heterogeneity in aggrecan expression (as measured by the coefficient of variation, **Table 4-1**), and the coefficient of variation decreased with exposure to induction media. However, the variability remained high even after long periods of time in differentiation culture (**Figure 4-1g**). Fully differentiated chondrocytes had the most homogeneous aggrecan expression of all the cell types and conditions we examined, though their mean aggrecan copy number was slightly lower than that of differentiated MSCs. These data

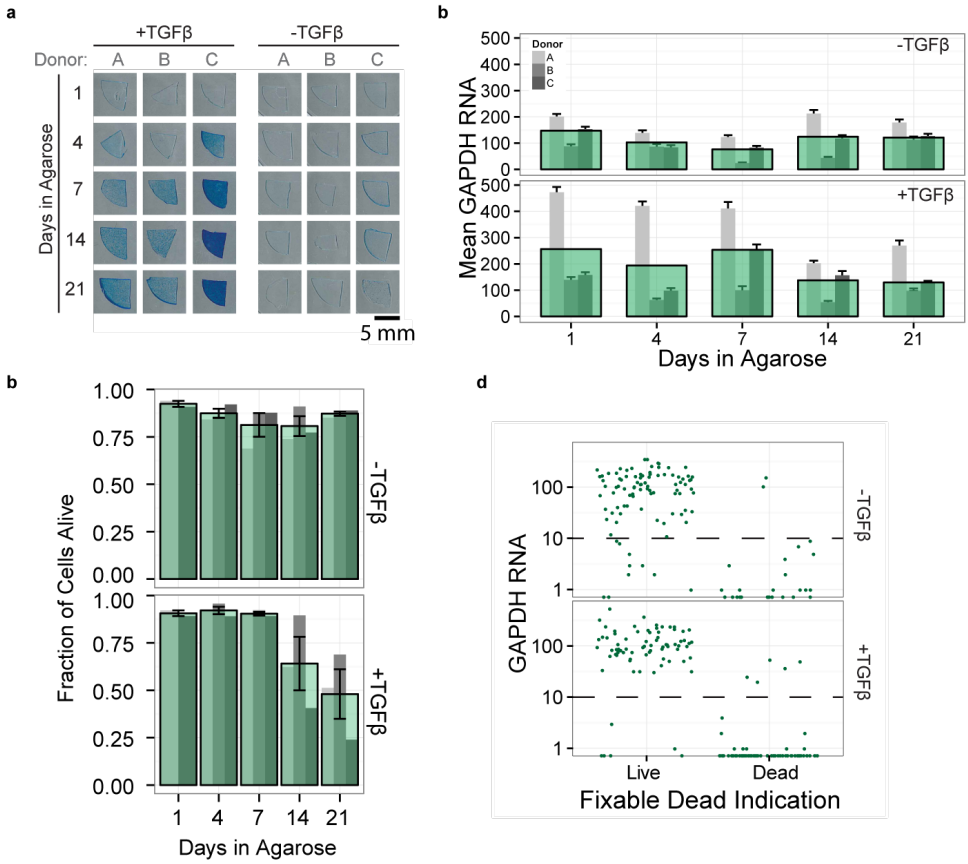

Figure 4-2: Matrix production, GAPDH copy number, and viability of MSCs in 3D culture. a) Alcian blue staining for sulfated proteoglycans for MSCs in 3D culture, for all donors and time points, with and without TGFβ induction of chondrogenesis. Donor B is also shown in Figure 1. Scale bar = 5 mm. b) Mean GAPDH RNA counts and c) cell viability over 21 days in 3D culture. Narrow bars represent the mean within an individual donor; overlaid bars represent the mean across donors. Error bars indicate standard error (n = 24-128 cells per donor and condition). RNA count means compared by t-tests with Satterthwaite approximation and simulated adjustment for multiple comparisons. See Appendix 2 Supplementary Table 1 for all statistical comparisons. d) Simultaneous RNA FISH and fixable dead staining established a threshold of GAPDH>10 to differentiate live cells from dead cells for further analysis. n = 85 cells for TGFβ-, 75 cells for TGFβ+.

show that MSCs exhibit substantial cell-to-cell expression heterogeneity and that, while chondrogenic culture promotes a chondrocyte-like gene expression pattern, copy

number remains highly variable between cells. Indeed, this variability within a population of differentiated MSCs overshadowed differences in mean expression between donors (3-4 orders of magnitude vs. a maximum ~2-fold difference, **Figure 4-1h**).

Table 4-1: Mean and Coefficient of Variation Associated with Aggrecan RNA Count in Undifferentiated and Differentiated Cells

	Mean Aggrecan	Aggrecan CV
Naive MSCs	69	1.60
Day 1 MSCs in gels	247	0.69
Day 21 MSCs in gels	334	0.72
Chondrocytes	225	0.57

This heterogeneity may either reflect different subpopulations that have adopted distinct fates or appear in cells that remain uncommitted. In the former scenario, if differentiated MSCs can express markers for only one fate at a time, then alternate lineage commitment should manifest as an anti-correlation between aggrecan and other lineage markers at the single cell level. To determine if this was the case, we performed RNA FISH for aggrecan, osteopontin and LPL in the same cells, with the latter two markers indicating an osteogenic and adipogenic lineage, respectively. Rather than identifying subpopulations that were distinctly chondrogenic or osteogenic, we instead observed a slight positive correlation between aggrecan and osteopontin (**Figure 4-1i**, Day 1 $\rho=0.49$, $p< 0.001$; Day 21 $\rho= 0.34$, $p<0.005$). Conversely, LPL expression was minimal, and did not correlate with either aggrecan or osteopontin expression (**Appendix 2 Supplementary Table 2**). These data suggested that heterogeneity in marker

expression after differentiation is not due to alternate lineage commitment, but rather highlights the fact that even differentiated MSCs can express high levels of markers for inappropriate lineages.

RNA levels poorly predict single-cell functional potential

Based on this tremendous cell-to-cell heterogeneity in chondrogenic gene expression, we next asked if aggrecan, Sox9, or collagen oligomeric matrix protein might serve as a means for separating robustly chondrogenic cells from the less chondrogenic ones in the initial heterogeneous population. For this to be possible, mRNA levels would need to correlate with chondrogenic capacity, indicated by the accumulation of a proteoglycan-rich extracellular matrix. To determine if such a connection existed, we seeded MSCs in 3D culture and induced chondrogenesis for 7 days, the point at which mean aggrecan expression peaked. We then performed immunofluorescent staining for aggrecan core protein (a central component of the cartilage-like extracellular matrix) simultaneously with RNA FISH using one of two probe sets: probes for markers of multiple fates (aggrecan, osteopontin, LPL, and GAPDH; Batch 1 samples) or probes for chondrogenic markers (Sox9, cartilage oligomeric matrix protein (COMP), and GAPDH; Batch 2 samples). We designated cells with evidence of extracellular staining for aggrecan core protein as ‘high-performing’ (comprising 12-62% of the population, depending on donor), and cells lacking extracellular staining as ‘low-performing’ (**Figure 4-3a-b**). Surprisingly, aggrecan expression did not strongly predict aggrecan core protein accumulation. Indeed, even within a single donor, the distribution of aggrecan mRNA abundance in high- and low- performing cells overlapped substantially (**Figure 4-3c**)**.** The mRNA/cell

distributions of other chondrogenic markers (COMP, Sox9), markers of alternative fates (osteopontin, LPL) and the housekeeping gene GAPDH (**Figure 4-4a**) also demonstrated similar overlap. While in aggregate the high-performing cells had a greater mean expression of aggrecan, COMP, and Sox9 and lower mean expression of osteopontin than low-performing cells, the magnitude of these differences was small and similar to the shift seen in GAPDH expression (**Figure 4-4a,** aggrecan: 1.35 fold increase, COMP: 1.14 fold increase, Sox9: 1.33 fold increase, GAPDH: 1.17 fold increase, osteopontin: -1.22 fold decrease). We also determined the expression ratios relative to commonly used normalization genes (i.e. aggrecan/GAPDH) or to genes indicating alternate lineage specification (i.e. aggrecan/osteopontin). These metrics also showed substantial overlap and small effect size **(Figure 4-4a**). Thus on this qualitative basis, neither absolute nor normalized single cell expression of marker genes was highly predictive of chondrogenic capacity at the single cell level.

To quantify the ability of transcript abundance to predict the extent of a cell's matrix accumulation, and thus sort high- from low-performing cells, we constructed receiver operating characteristic curves to determine the 'true positive' (sensitivity) and 'true negative' (specificity) rates associated with potential mRNA thresholds. We pooled data across donors assayed using the same probes (batches 1 and 2). Within each batch, we assessed the high/low classification performance of individual genes, gene expression ratios and linear combinations of gene expression levels **(Figure 4-3d-f, Figure 4-4b)**.

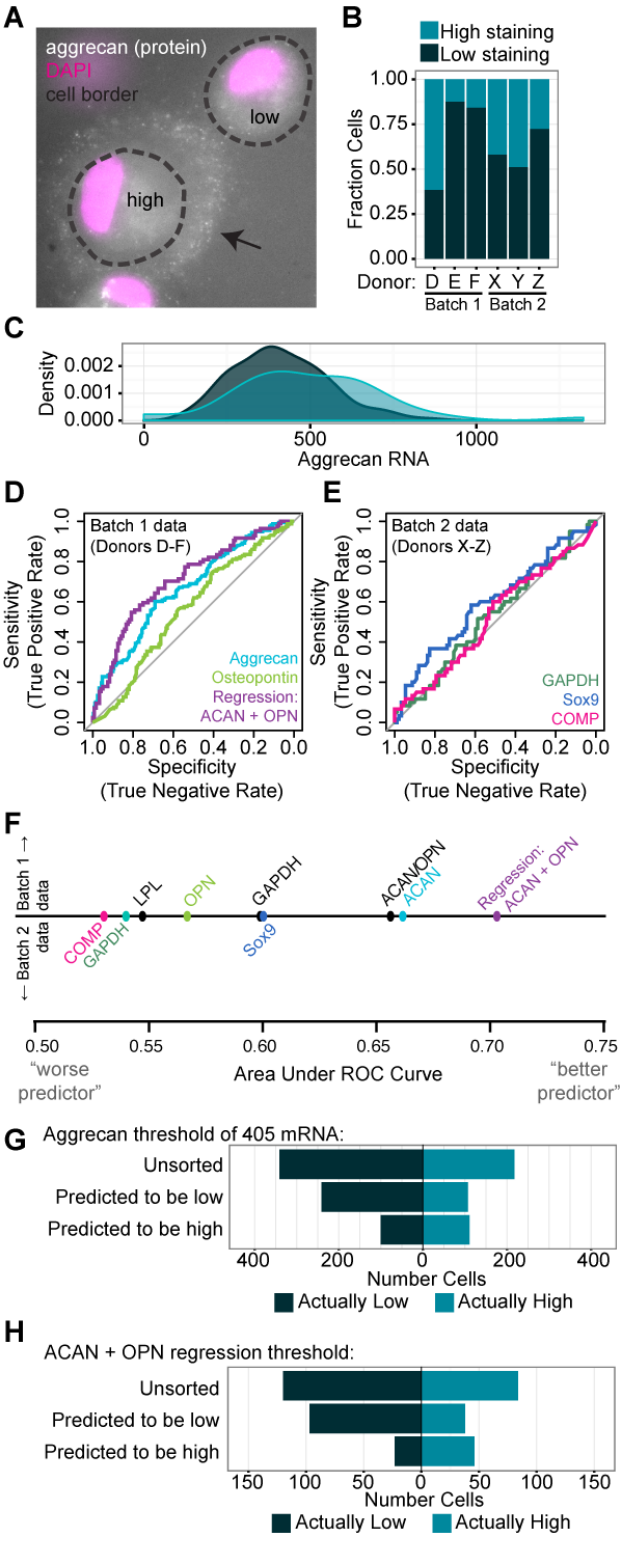
A
aggrecan (protein)
DAPI
cell border
low
high
B
High staining
Low staining
Fraction Cells
1.00
0.75
0.50
0.25
0.00
Donor: D E F X Y Z
Batch 1 Batch 2
C
Density
0.002
0.001
0.000
0
500
1000
Aggrecan RNA
D
Batch 1 data
(Donors D-F)
Sensitivity
(True Positive Rate)
Aggrecan
Osteopontin
Regression:
ACAN + OPN
Specificity
(True Negative Rate)
E
Batch 2 data
(Donors X-Z)
GAPDH
Sox9
COMP
F
Batch 1 data
Batch 2 data
LPL
OPN
GAPDH
ACAN+OPN
ACAN
Regression:
ACAN + OPN
COMP
GAPDH
Sox9
0.50
0.55
0.60
0.65
0.70
0.75
"worse predictor"
Area Under ROC Curve
"better predictor"
G
Aggrecan threshold of 405 mRNA:
Unsorted
Predicted to be low
Predicted to be high
400
200
0
200
400
Number Cells
Actually Low
Actually High
H
ACAN + OPN regression threshold:
Unsorted
Predicted to be low
Predicted to be high
150
100
50
0
50
100
150
Number Cells
Actually Low
Actually High

Figure 4-3 (previous page): Marker gene expression is a poor predictor cartilage-like matrix production in individual MSCs. a) Aggrecan core protein identified by immunostaining of MSCs showing high or low cartilage-like matrix formation after 7 days of 3D culture with TGFβ. b) Fraction of cells classified as high- or low-performing based on aggrecan protein staining (cells/donor: D: 78, E: 89, F: 51, X: 62, Y: 43, Z: 47). c) Distribution of aggrecan copy number in high- and low-performing MSC populations; probability density curve for Donor E shown (n = 153 cells). d-e) Receiver operating characteristic curves using individual gene expression and regression analysis on combinations of genes from batch 1 (d) and individual gene expression from batch 2 genes (e) to distinguish between high- and low-performing MSCs (cells/donor: D: 132, E: 153, F: 122, X: 57, Y: 42, Z: 47). f) Summary graph of area under the curve of receiver operating characteristic curves for individual gene expression, gene expression ratios, and regression analysis of combinations of gene. g-h) Simulated sorting of MSCs into anticipated high- and low-performing cells, using the optimized threshold of 405 aggrecan mRNA copies (g) and the optimized threshold from the ACAN+OPN regression (h).

While each metric discriminated between high- and low-performing cells better than random chance (represented by the diagonal line on the ROC plots, and an area under the curve = 0.5), the improvements in selection specificity were relatively small. Of the individual RNA types indicative of the chondrogenic lineage, aggrecan and Sox9 were best able to discriminate between high and low performing cells. For example, consider the optimized threshold of 405 aggrecan mRNA, which maximizes the Youden J statistic (sensitivity + specificity − 1). Conceptually, we can designate all cells with >405 aggrecan RNA as anticipated high performers, and others as anticipated low performers. For the donors studied, this unsorted population was 34% high- and 66% low-performing cells. Sorting based on this optimized aggrecan threshold misclassified 37% of all cells (i.e. percent of high cells predicted to be low, or low cells predicted to be high). 50% of high-performing cells were lost due to incorrect classification as ‘anticipated-low’ cells, and the fraction of high-performing cells in the ‘anticipated-high’ population was enriched only 35% over the unsorted population **(Figure 4-3g)**. A logistic regression model combining aggrecan and osteopontin expression improved on this performance only

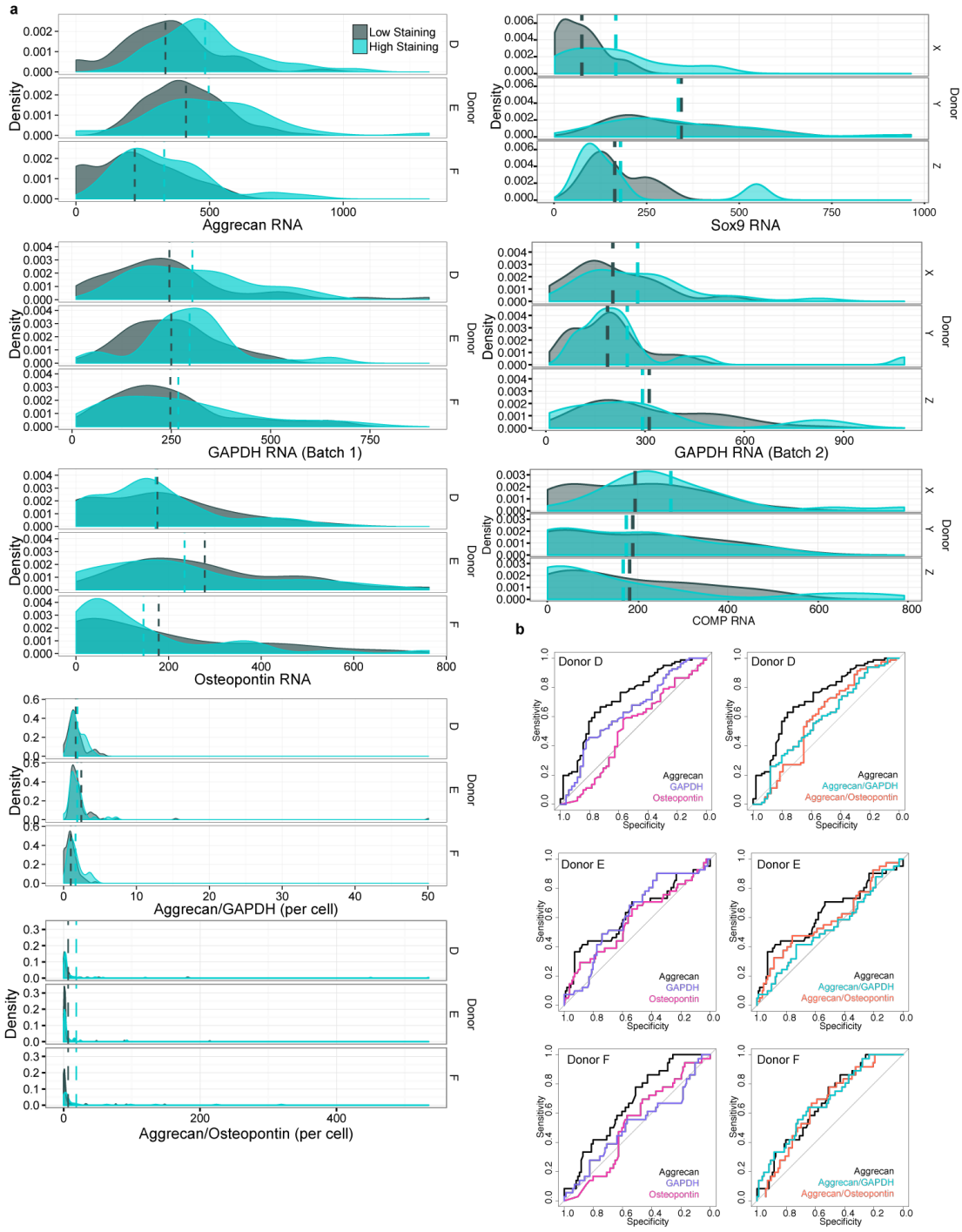
a
Low Staining
High Staining
Density
Donor
Aggrecan RNA
Sox9 RNA
GAPDH RNA (Batch 1)
GAPDH RNA (Batch 2)
COMP RNA
Osteopontin RNA
Aggrecan/GAPDH (per cell)
Aggrecan/Osteopontin (per cell)
b
Donor D
Donor E
Donor F
Sensitivity
Specificity
Aggrecan
GAPDH
Osteopontin
Aggrecan/GAPDH
Aggrecan/Osteopontin

Figure 4-4 (previous page): Matrix staining intensity versus mRNA copy number and ROC curves for individual donors, markers, and marker ratios. a) Distribution of aggrecan, GAPDH, osteopontin, aggrecan/GAPDH, and aggrecan/osteopontin gene expression within high- and low-performing MSC populations; separated by donor. Dashed lines represent the mean for each condition. For aggrecan/osteopontin, only cells that had at least one osteopontin (>95% of cells per donor) mRNA were included in the graph and mean statistics calculation. b) Receiver operating characteristic curves using individual gene expression and gene expression ratios to distinguish between high- and low-performing MSCs, separated by donor. Cells/donor: D: 132, E: 153, F: 122, X: 57, Y: 42, Z: 47.

slightly, where its optimized threshold yielded a 33% misclassification rate, and enriched the fraction of high cells by 37% (**Figure 4-3h**, vs 35% for aggrecan alone). Of the gene expression ratios, aggrecan/osteopontin was a better discriminator than aggrecan/GAPDH, though its selection performance did not surpass that of aggrecan alone. Sorting on a donor-by-donor basis was similarly ineffective (**Figure 4-4b**). Thus, sorting cells based on expression of aggrecan, other common differentiation markers, and linear combinations thereof would result in only marginal enrichment of the population, while substantially reducing available cell number.

Transcriptomics does not identify better marker sets

Based on the inability of aggrecan and other lineage specific markers to robustly predict matrix accumulation at the single cell level, we next utilized high throughput RNA sequencing to determine if other features of the transcriptome, and specifically factors present in the undifferentiated population, might prospectively identify MSCs with high differentiation potential. We expanded single cell-derived MSC colonies in monolayer, and collected a fraction of the cells for RNA sequencing and subsequent transcriptome analysis. The remaining fraction was expanded through an additional passage, formed into pellets, and cultured in the presence of TGFβ for 21 days to assay chondrogenic

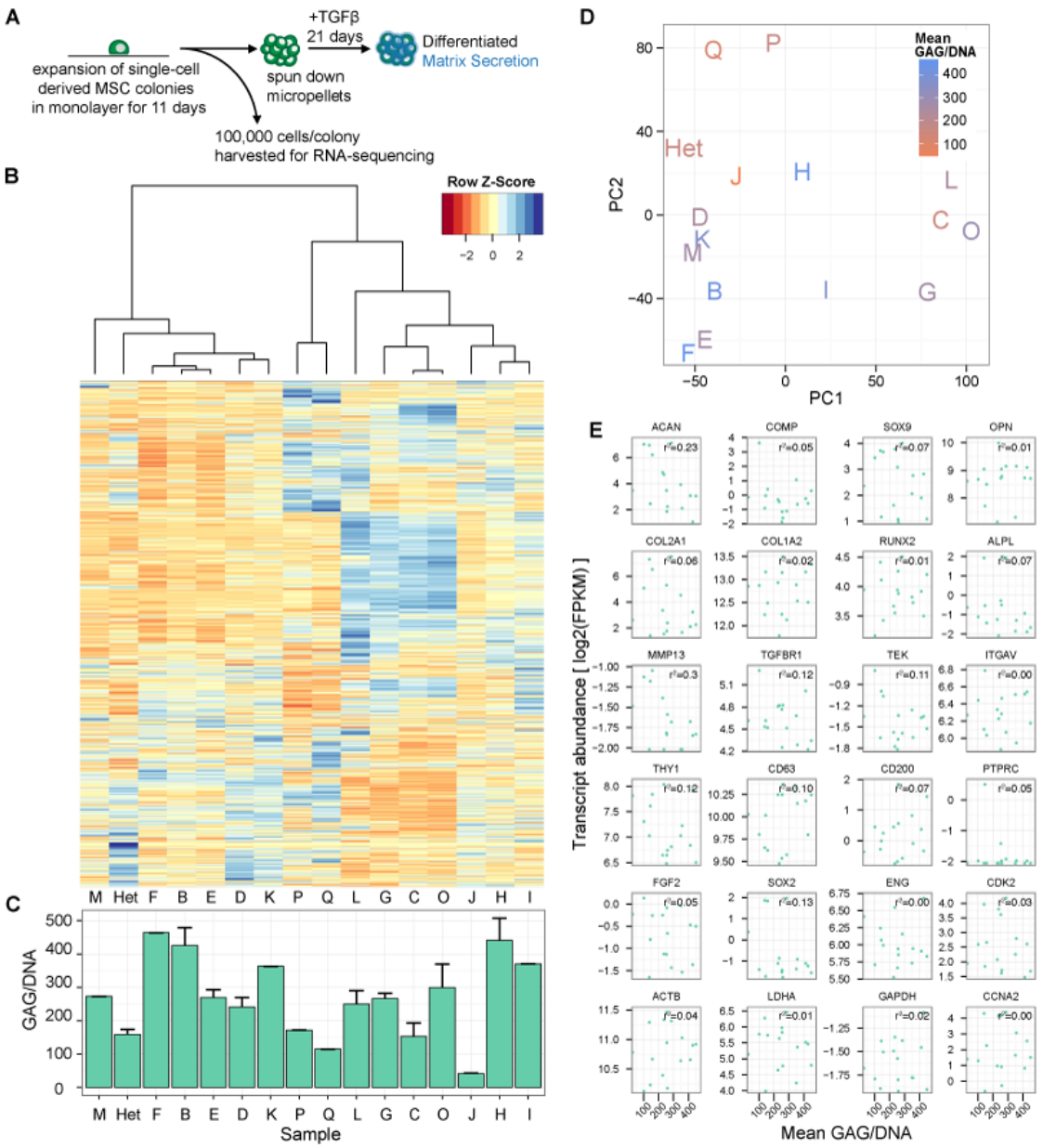

Figure 4-5: Genome-wide transcriptome profiling does not predict MSC functional potential. a) Schematic for RNA sequencing and testing of functional capacity of single cell derived clones. b) Unbiased clustering of clones (or heterogeneous population) based on fragments per kilobase of transcript per million reads (FPKM) of RNA sequencing results (subsetted for genes where at least one sample had FPKM > 1). c) Glycosaminoglycan deposition per DNA in micropellets derived from clonal or heterogeneous populations (from part b) cultured for 21 days in chondrogenic (TGFβ+) culture media. d) Principal component analysis of same RNA sequencing results as in part b, colored by GAG/DNA for each clone. e) log2 transformed FPKM of selected genes from RNA sequencing results as a function of GAG/DNA for each clone.

potential (**Figure 4-5a**). This evaluation of baseline MSC gene expression in clonal populations derived from single cells had the potential to identify markers that could be used to sort freshly isolated MSCs based on their gene expression signatures.

An initial comparison of differential expression between clones (**Figure 4-5b**), as compared to the deposition of extracellular matrix components of each clone (**Figure 4-5c**), revealed no striking patterns of gene expression that correlated with subsequent matrix deposition. We also used principal component analysis to determine if the variation between the gene expression of each clone could be used to predict functional capacity, but there was no relationship between clustering in either of the first two principal components and matrix deposition (**Figure 4-5d**).

Given that the full transcriptome lacked global predictive capacity, we next sought to broaden our conclusions from the FISH experiments by examining the sequencing data associated with individual genes. We selected a small subset of genes that corresponded to four categories of markers identified in previous studies: chondrogenic markers, stemness markers, cell-cycle associated genes, and housekeeping genes[27]. Consistent with our single-cell analysis results, none of these genes correlated strongly with functional potential on a clonal basis (**Figure 4-5e**). Even the most predictive genes, MMP13 and aggrecan, correlated only weakly (r^2= 0.3, p = < .05 and r^2=0.23, p = 0.062, respectively). Together, this transcriptomic analysis suggests that there is no expression signature at the RNA level that could pre-identify specific clones with high chondrogenic potential.

Marker heterogeneity emerges rapidly after cell division

Based on the inability of transcript levels to robustly predict matrix forming potential, we next asked whether it was propagated through cell division; that is, whether cells with a higher expression level would transmit this feature to their daughter cells. As an initial assay, we measured aggrecan copy number in every cell located within a series of small MSC colonies stimulated with TGFβ (where each individual colony likely arose from a single cell, **Figure 4-6b**). Results from this analysis showed that aggrecan copy number varied more within a single colony than it did between colonies (**Figure 4-6b**). This result suggests that with just a few cell divisions, aggrecan levels rapidly devolved to recapitulate the heterogeneity present in the bulk population. In contrast, GAPDH was less variable than aggrecan within each colony (lower coefficient of variation, **Appendix 2 Supplementary Table 3**), but showed greater differences in mean level between colonies (**Figure 4-6c**). Thus, not every gene demonstrated the high intra-colony variability observed in aggrecan expression, and some genes were differentially expressed between colonies. However, without live-cell time-lapse measurements of the cellular lineage, it was difficult to directly show that variability in aggrecan mRNA levels arose through randomization rather than heritable differences.

To overcome this limitation, we next continuously tracked MSCs as they migrated and divided in induction media by live cell microscopy for three days, and correlated terminal aggrecan expression between sister cells with respect to the time since their last division (**Figure 4-6d**). Shortly after division (<12 hours), sister cells had comparable aggrecan and GAPDH levels (**Figure 4-6e-f, Figure 4-7a-b**), suggesting symmetric partitioning of

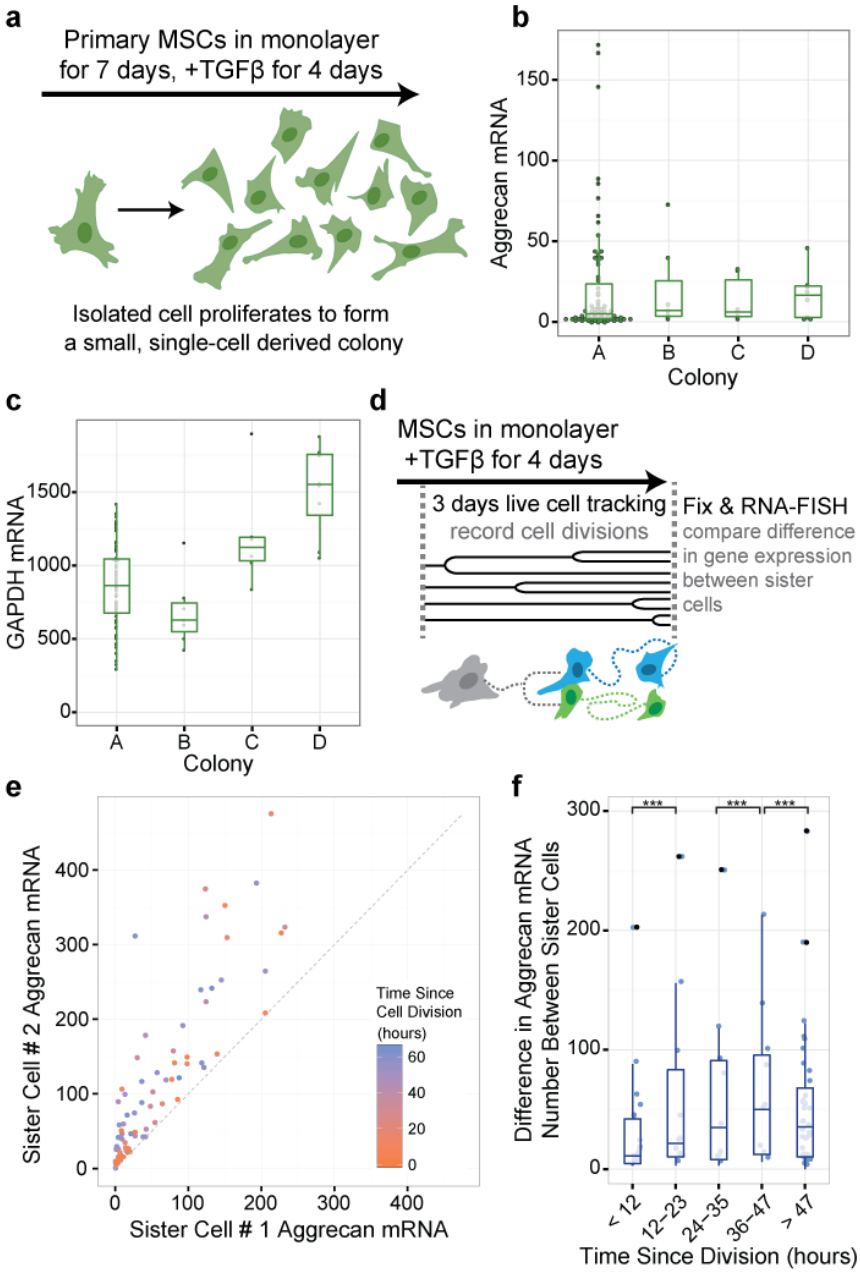

Figure 4-6: Marker expression heterogeneity emerges shortly after cell division.a-c) Gene expression in small MSC colonies. a) Colony formation scheme. b-c) Aggrecan and GAPDH expression in four colonies established from a single donor (n = 75 cells in colony A, 7 cells in colony B, 6 cells in colony C, 8 cells in colony D). d) Live-cell tracking scheme to identify sister cell pairs at various times post-cell division. e-f) Divergence in aggrecan gene expression between sister cells as a function of time since last division (n = 81 sister cell pairs, *** indicates $p<0.001$, see Appendix 2 Supplementary Table 5 for all comparisons).

RNA. However, after more than ~12 hours since division, sister cells showed increasingly divergent levels of aggrecan and GAPDH expression (**Figure 4-6e-f, Figure 4-7a-b, Appendix 2 Supplementary Table 6**). Within cell pairs, aggrecan and GAPDH divergence only weakly correlated, suggesting that the relative difference between sister cells was not globally regulated, underscoring the fact that aggrecan and GAPDH do not necessarily change together (**Figure 4-7c-d**). These findings may reflect a difference in cell function as a consequence of asymmetric cell division (i.e. sister cells have different target expression levels) or could simply identify how asynchronous dynamic fluctuations lead to temporal differences in expression level. In either case, these differences suggest that a sorted population of high aggrecan cells would not remain so for more than a couple days, and may explain why, at the single cell level, cells with high aggrecan RNA expression are not necessarily the cells with the greatest amount of matrix deposition.

Marker genes do not identify a chondrocyte phenotype

While aggrecan gene expression did not correlate with matrix deposition in MSCs, it is a canonical feature of the differentiated chondrocyte 'phenotype' and is widely considered to be a leading indicator of cartilage-specific extracellular matrix deposition (e.g. aggrecan core protein) [7,134,221]. It is also well accepted that, upon serial passaging and expansion in monolayer, chondrocyte matrix production decreases along with a multi-fold decrease in the aggrecan/GAPDH ratio (**Figure 4-8a**)[30,42,52,126,132]. This change in expression is associated with increases in cell size and proliferation rate [71,111,189]. To reconcile our finding of discordant aggrecan expression and matrix deposition in MSCs

with these classical experiments that define the chondrocyte 'phenotype', we performed RNA FISH on chondrocytes that were serially passaged in monolayer to induce "de-differentiation" and after they were "re-differentiated" in 3D culture (where one would expect a resumption of the cartilage phenotype)[12,15].

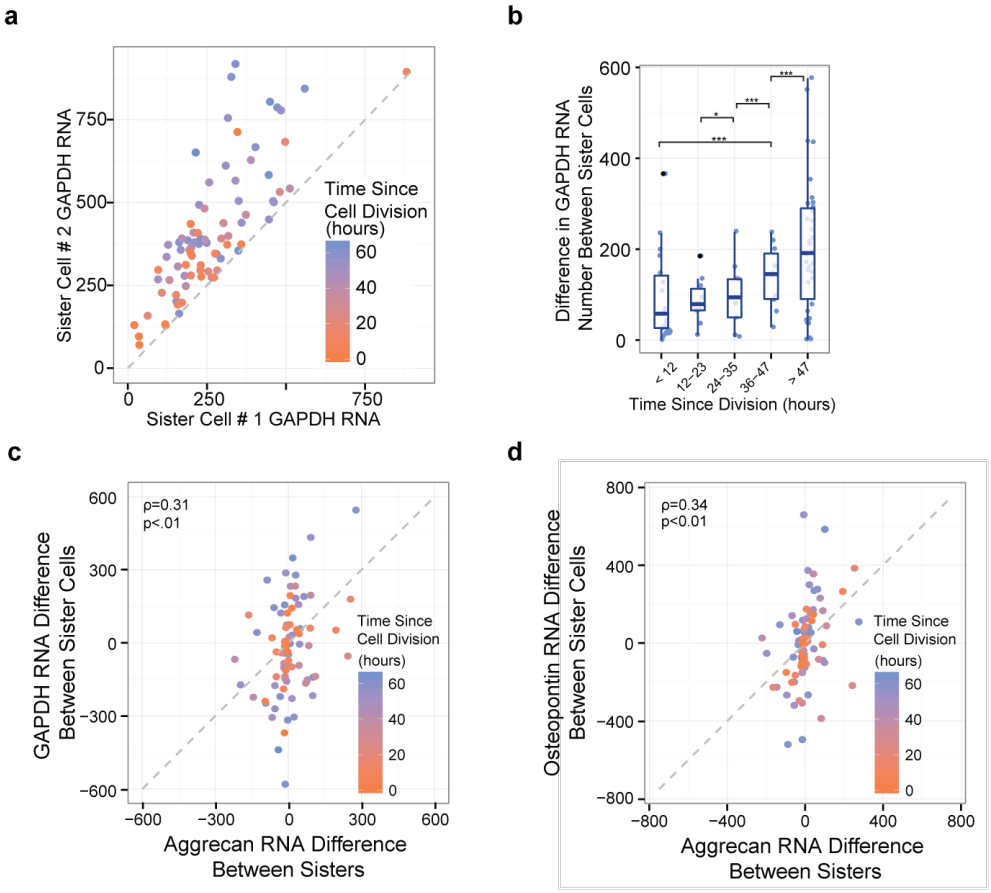

Figure 4-7: Heritability of marker copy number through cell division. a-b) Divergence in gene expression between sister cells as a function of time since their last division. c) Divergence in GAPDH vs aggrecan and d) osteopontin vs aggrecan between sister cells as a function of time since their last division. Box hinges denote the first and third quartiles. Whiskers extend from the hinges to the most extreme data points within (1.5 * interquartile range) of the hinges. Means compared by t-tests with Satterthwaite approximation and simulated adjustment for multiple comparisons, *** indicates p<0.001,* p<0.05. See Appendix 2 Supplementary Table 6 for all statistical comparisons. n = 81 sister cell pairs.

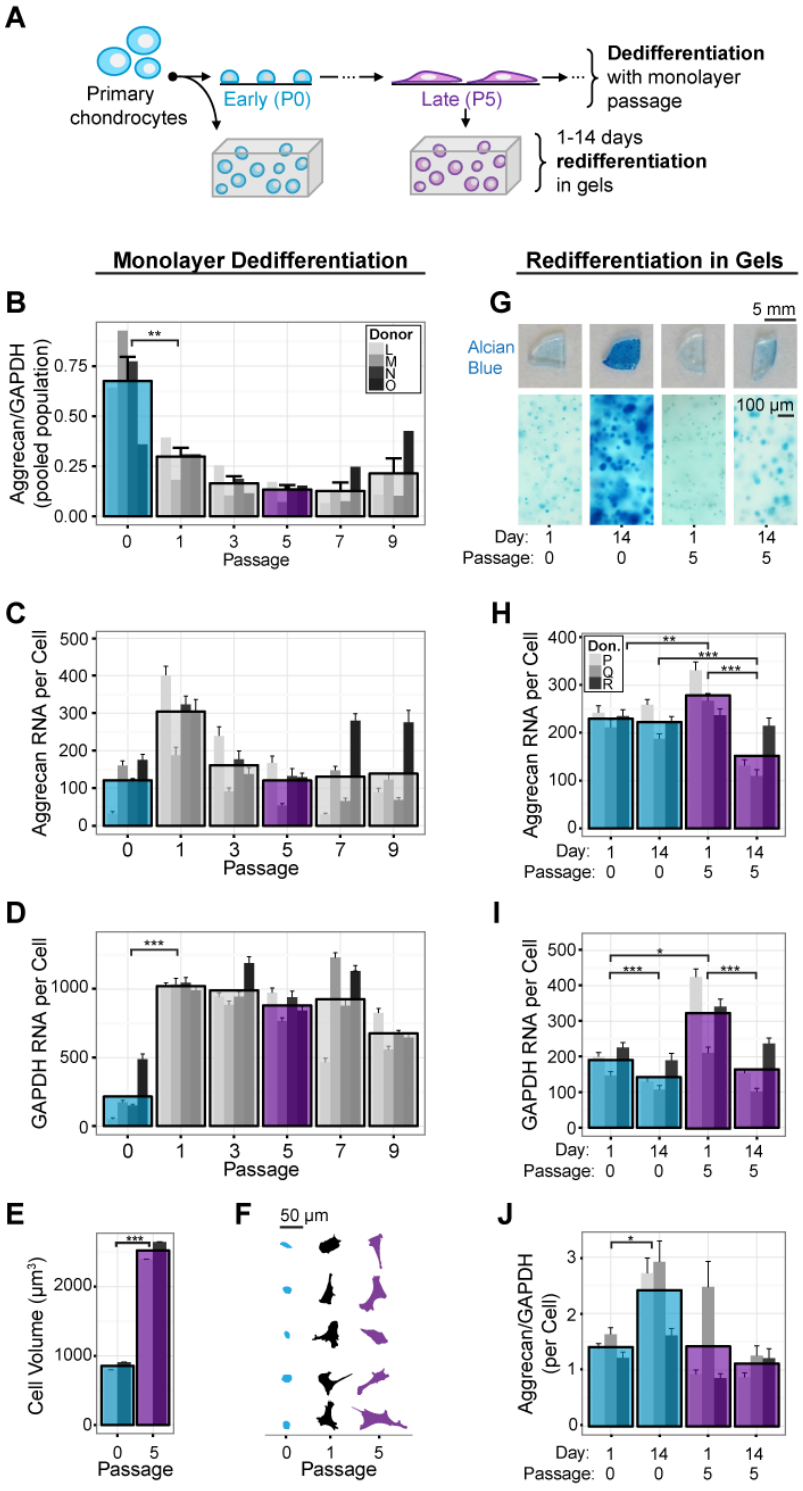
A
Primary chondrocytes
Early (P0)
Late (P5)
Dedifferentiation with monolayer passage
1-14 days redifferentiation in gels
Monolayer Dedifferentiation
Redifferentiation in Gels
B
Aggrecan/GAPDH (pooled population)
Donor
L
M
N
O
Passage
G
5 mm
Alcian Blue
100 µm
Day:
Passage:
C
Aggrecan RNA per Cell
H
Don.
P
Q
R
D
GAPDH RNA per Cell
I
E
Cell Volume (µm³)
F
50 µm
J
Aggrecan/GAPDH (per Cell)

Figure 4-8 (previous page): Chondrocyte de-differentiation and re-differentiation are not driven by altered absolute aggrecan expression. a) Chondrocyte de-differentiation and re-differentiation scheme. b-f) Analysis of chondrocytes de-differentiating with passage in monolayer culture. b) RNA FISH counts of aggrecan pooled over the population and normalized to GAPDH expression. c) Absolute aggrecan expression with passage number. d) Absolute GAPDH expression with passage number (n = 39-113 cells per donor per passage). e-f) Chondrocyte suspended cell volume (e) and morphology (f) with passage (n = 274-543 cells per donor per passage). g-j) Analyses of early passage (P0) and late passage (P5) chondrocytes re-differentiating in 3D culture. g) Alcian blue staining after 1 and 14 days of 3D culture. h) Absolute aggrecan expression. i) Absolute GAPDH expression. j) Single cell aggrecan expression normalized to GAPDH expression. (n = 46-65 cells per donor per condition). Narrow bars represent the mean within an individual donor; overlaid bars represent the mean across donors. Error bars indicate standard error; * indicates $p<0.05$, ** indicates $p<0.01$ and *** indicates $p<0.001$, see Appendix 2 Supplementary Table 7 and Supplementary Table 8 for all comparisons).

For dedifferentiation studies, we serially passaged chondrocytes nine times in monolayer with analysis at every other passage via RNA FISH. Consistent with classical findings [12,205], the normalized ratio of aggrecan to GAPDH expression level decreased with passage number (**Figure 4-8b**). However, and quite surprisingly, this change was not due to a decrease in absolute aggrecan copy number (**Figure 4-8c**). Rather, aggrecan copy number showed a small but significant increase from passage 0 (initial plating) to passage 1, before returning to passage 0 mean copy number at later passages. In contrast, there was a rapid increase in mean GAPDH copy number over the first passage (increasing ~4 fold) that remained at these elevated levels through additional passages (**Figure 4-8d**). Previous studies from our group have shown that global transcription (including expression of GAPDH and many other abundant "house-keeping" genes) correlates with and can be dictated by cell size [38]. We also found that chondrocyte spread cell area generally increased with passage number (**Figure 4-8f, Figure 4-9**) and that the mean volume of suspended cells increased by ~3-fold between primary isolation (passage 0) and passage 5 (**Figure 4-8e**). Taken together, these findings suggest that aggrecan expression does not decrease with chondrocyte de-

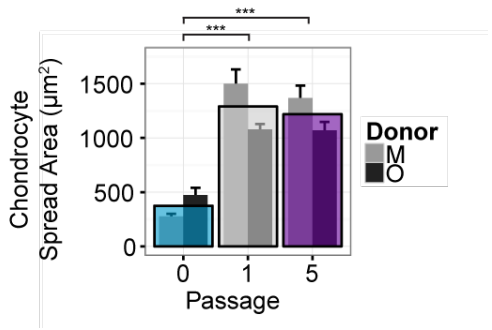

Figure 4-9: Chondrocyte morphology with passage number. Chondrocyte spread area with increasing passage number during de-differentiation (n = 25-27 cells per donor per condition). Narrow bars represent the mean within an individual donor; overlaid bars represent the mean across donors. Error bars indicate standard error. Compared via one-way ANOVA with Tukey post-hoc test, *** indicates p<0.01.

differentiation and does not correlate with chondrocyte functional potential at the population level. Instead, normalization to housekeeping genes obscures relatively minor changes in aggrecan gene expression that occurs during chondrocyte “de-differentiation”. These single cell data suggest that canonical markers of the chondrocyte phenotype do not accurately describe the molecular profile of de-differentiation.

To further explore how normalization may confound our interpretation of gene expression changes, we forced the re-differentiation of culture-expanded chondrocytes that had lost their ‘phenotype’. To do so, we encapsulated chondrocytes at early and late passage (passage 0 and 5, respectively) in 3D agarose hydrogels, and monitored matrix synthesis and gene expression over two weeks via Alcian blue staining and RNA FISH (**Figure 4-8a**). Consistent with classical studies [12,15], early passage chondrocytes produced matrix robustly upon encapsulation, while late passage (de-differentiated)

chondrocytes showed a significant attenuation in matrix deposition (**Figure 4-8g**). RNA FISH showed that after 1 day of agarose culture, late passage chondrocytes expressed more aggrecan and more GAPDH than early passage chondrocytes (**Figure 4-8h-i**). Over 14 days, mean aggrecan levels were maintained in early passage cells, but decreased in late passage cells. In keeping with our findings in monolayer, the aggrecan/GAPDH ratio was strongly influenced by changes in GAPDH (**Figure 4-8j**). These data further support the finding that absolute changes in aggrecan expression levels are not responsible for the loss of phenotype observed in serially passaged chondrocytes.

Discussion

In this work, quantitative single cell analysis of gene expression provided evidence that the abundance of mRNA markers is only weakly linked to the chondrogenic phenotype of cartilage and progenitor cells. Specifically, we found that both MSCs and chondrocytes exhibited rampant transcriptional heterogeneity. This observation was not altogether surprising for MSCs, given that a single MSC population is comprised of a heterogeneous pool of related but distinct clonal populations. However, the transcriptional heterogeneity within individual MSC colonies suggested that this overall population heterogeneity is not entirely due to the mixing of clonal populations of varying potency, but instead likely arose from random transcriptional processes. While such heterogeneity may confound the interpretation of ensemble measurements, if this variation reflected intrinsic differences in differentiation capacity or differentiated state, then it might be harnessed toward a productive end. That is, cell sorting based on this

variability could enable selection of 'superior' sub-populations for therapeutic applications. For example, the expression of 'stemness markers' such as SOX2[118], OCT4[106] and NANOG[116] can distinguish pluripotent cells from larger heterogeneous populations, and the expression of an early osteogenic marker enables enrichment of the stromal vascular fraction for osteogenic cells[138].

However, our data show that for naive MSCs, neither genome-wide transcriptional metrics nor the transcriptional abundance of MSC stemness and chondrogenic markers correlate with the ultimate functional capacity. Strikingly, the most predictive genes (aggrecan and MMP13) were negatively associated with chondrogenic capacity, potentially suggesting that high transcriptional promiscuity in naive MSCs reflects an inability to undergo robust lineage commitment. Furthermore, our single cell studies showed that while naïve MSCs and chondrocytes represent opposite ends of the differentiation spectrum, their absolute expression of canonical differentiation markers largely overlapped. When we monitored gene expression and cartilage-like matrix accumulation simultaneously on a cell-by-cell basis, marker expression taken at a single time point only weakly associated with cell output of extracellular matrix. Thus, we conclude that marker expression would only enable a slight enrichment of the population (~35% increase in high performing cells over the unsorted population) while drastically restricting available cell number for therapeutic application.

One possible explanation of the disconnect between an individual cell's transcript abundance and differentiated state is that, for many genes, transcription is a stochastic process comprised of long 'silent' periods punctuated by short transcriptional

bursts[33,72,176,177,209,236]. Bursting kinetics are strongly dependent on both the gene in question and the stimulus that is applied [41,159,177,209], along with position in cell cycle and cell volume [160,237]. For instance, stimulation (e.g. TGFβ) can induce a synchronized initial burst of target gene expression, but subsequent bursts are typically asynchronous [20,68,150,195]. Thus, two cells with fluctuating but equivalent gene expression over time may exhibit different copy number when sampled at a single time point. As recently reported [112,150], the rate of fluctuation (slow vs. fast) of a single gene manifests in the heterogeneity observed between and within small clonal clusters. Our findings of high intra-colony variability and sister cell divergence in MSCs suggest that marker copy number fluctuates rapidly over a short timescale. As a result, absolute marker gene expression is not strongly heritable in MSCs, and we speculate that cells sorted on the basis of such expression will undergo transcriptional shifts over time and with further population expansion. In other systems, such stochastic variation in gene expression not only marks but can also determine cell fate [31,51,133,175,178,207,225]. Here, it is surprising that for aggrecan, a gene whose product plays such a critical role in the extracellular matrix, such emergent heterogeneity in transcript abundance does not appear to reflect true variation in potency.

The disconnect between expression and functional capacity (matrix accumulation) may also reflect the time history of the system and the influence of other regulatory mechanisms. Aggrecan core protein undergoes co- and post-translational modifications, and may be subject to processing or secretory errors [130,221]. It may be that not every cell that produces core protein can appropriately modify the core and secrete it into the

extracellular space. Further, integration and retention of aggrecan core protein within the extracellular matrix relies on association with the hyaluronic acid and collagen network and other molecules [108,221], and even aggrecan that has been integrated into the established matrix may ultimately be degraded by aggrecanases produced locally [25]. Deficiencies in any of these steps could decouple even temporally-constant aggrecan mRNA expression from aggrecan core protein accumulation in the pericellular space. However, our transcriptome-wide data suggest that there is not a transcript level correlation between functional capacity and any of the genes involved in these processing steps.

Collectively, our findings in MSCs show that instantaneous aggrecan expression is only tenuously connected to matrix deposition. Moreover, differentiation of these cells fails to recapitulate the potential of native chondrocytes and does not prevent the expression of markers of alternate lineages even at the single cell level. Our finding that chondrocyte expression of aggrecan does not decrease with dedifferentiation also supports this weak connection, and raises questions as to the role of marker gene expression in defining phenotype. While aggrecan is one of the most conventional markers for the cartilage phenotype, its absolute expression did not correlate with cartilage-like matrix production and did not change as cells ‘de-differentiated’. If aggrecan expression does not change, other elements of the cell must be responsible for shifting cell fate and altering the transcriptional ‘focus’ of the cell. Here, our finding of major shifts in GAPDH with minor changes in aggrecan during de-differentiation suggest that de-differentiation may be better characterized as a shift in cell focus rather than a loss in specific programmatic

expression of marker genes. While it is not yet clear what cell-wide changes drive this process, future work utilizing transcriptomics may identify a more comprehensive set of markers that are predictive of differentiated cell function. Until phenotype and its basis in gene expression are more precisely defined, our results suggest that it may be ineffective to design therapies that seek to bolster phenotype by increasing expression of individual genes or regulating transcriptional control of individual promoter regions, even for those genes whose products are directly related to functional matrix assembly. Simply increasing the raw RNA signal available to the cell may be insufficient, and it may also be necessary to alter the transcriptional context in which this occurs. These findings challenge the traditional notion that marker gene expression defines or is even strongly associated with the chondrocyte phenotype, and identify new directions in progenitor cell biology to establish, enforce, and select subpopulations for therapeutic application.

Methods

Cell Isolation and Expansion

MSCs were isolated from the tibial and femoral bone marrow of juvenile bovine cows (3–6-months, Research 87, Boylston, MA) and expanded in a basal media consisting of high glucose DMEM with 10% FBS and 1x antibiotic-antimycotic. After the initial plating reached ~80% confluence, cells were passaged at a ratio 1:3 before use in experiments. For single cell derived colonies, bovine MSCs were isolated as described above and seeded sparsely onto glass coverslips. Individual colonies were allowed to expand for 3 days in basal media, followed by 4 days in chondrogenic induction media prior to fixation. All cells in each colony were manually located and imaged as described below.

Chondrocytes were isolated from articular cartilage from the trochlear groove of juvenile bovine knees. Cartilage was digested in basal media supplemented with type II collagenase (0.5 mg/mL, Sigma-Aldrich) for up to 18h. Isolated cells were filtered, washed, and plated in basal media. To improve cell yield for chondrocyte re-differentiation studies, cartilage was also digested in basal media with pronase (2.5 mg/mL, Calbiochem) for 1h prior to collagenase digestion. For all studies, chondrocytes were expanded in basal media and passaged 1:10 when plates reached ~80% confluence.

Cell Encapsulation

For 3D culture, MSCs (passage 2) or chondrocytes (passage 0 and passage 5) were encapsulated in 2% agarose micro-gels at a density of 2 million cells/mL. Molten 4% w/v agarose (type VII, Sigma, 44 °C) was mixed 1:1 with cells suspended in media and pipetted into small drops in a well plate. Round coverslips were placed on top of the molten drops to spread the mixture before the gel solidified, resulting in the formation of uniform micro-gels that were 10-12 mm in diameter (depending on coverslip diameter) and ~400 μm thick. Coverslips were removed from the micro-gels prior to culture. Micro-gels were supplied with fresh medium every three days and 24 hours before collection. MSC micro-gels were maintained in a chemically defined media consisting of high glucose DMEM supplemented with 1x antibiotic-antimycotic, 40 ng/mL dexamethasone, 50 μg/mL ascorbate 2-phosphate, 40 μg/mL L-proline, 100 μg/mL sodium pyruvate, 1.25 mg/mL bovine serum albumin, 5.35 μg/mL linoleic acid and 1× insulin–transferrin–selenous acid premix (Corning CB-40350), either with or without 10 ng/mL TGFβ3 (R&D

Systems)[144]. Chondrocyte micro-gels were cultured in basal media (high glucose DMEM + 10% FBS + 1x antibiotic-antimycotic) supplemented with 50 μg/mL ascorbate 2-phosphate. At defined time points, gels were fixed for 30 minutes in paraformaldehyde (PFA) and stored in 70% ethanol at 4°C.

Cell viability in gels was assessed using the LIVE/DEAD Cell Viability Assay Kit (Molecular Probes L-3224). A custom Matlab script quantified the number of live (calcein-AM positive) and dead (ethidium-homodiner-1 positive) cells in three 4X fields of view per micro-gel. To assess viability in conjunction with RNA FISH, a fixable, amine-binding green fluorescent dead cell stain (Molecular Probes L-23101) was employed. For fixable dead staining, micro-gels were washed with PBS, stained for 30 minutes in a 1:5000 dilution in PBS, washed with PBS again, and then fixed in PFA prior to RNA FISH analysis, as described below.

Chondrogenic Pellet Culture and Biochemical Content

Clonally derived passage 2 MSCs were formed into cell-rich pellets via centrifugation (200,000 cells/pellet) and cultured in chondrogenic induction media with TGFβ for 21 days[61]. Pellets were papain-digested and biochemically assayed for glycosaminoglycan (GAG) and DNA content using via the1,9-dimethylmethylene blue and Picogreen (Molecular Probes, Eugene, OR) assays, respectively[61].

Live Cell Imaging and Tracking

To investigate mRNA levels as a function of the time history of division, passage 2 MSCs were seeded into two-well LabTek chambered coverglass dishes (Fisher

Scientific) and cultured in chondrogenic induction media with TGFβ for 4 days. Seeded cells were supplied with fresh medium every 3 days and 24 hours before fixation. Over the last 3 days of culture, live cells were imaged using a Nikon Ti-E microscope with a custom environmental chamber. Transmitted light images were automatically acquired every 30 minutes over a period of 70 hours using a 10X air objective over a 289-image grid in each well of the two-well coverglass. Cell division was tracked manually using ImageJ, and matched to the corresponding RNA FISH quantification that followed.

RNA Fluorescence In-Situ Hybridization and Imaging

Single molecule RNA FISH was performed on samples[174]. Micro-gels and monolayer cells were fixed in PFA and permeabilized with 70% ethanol before in situ hybridization was performed using the specified pools of oligonucleotides. Monolayer and micro-gel samples were simultaneously co-stained with oligonucleotide probes for osteopontin labeled with Cy3, lipoprotein lipase labeled with Alexa 594, aggrecan labeled with Atto 647N, and GAPDH labeled with Atto700 (Stellaris oligonucleotides, Biosearch Technologies). See Appendix 3 for a complete list of sequences of oligonucleotide probes used in this study. Subsequently, samples were washed with 2x saline sodium citrate buffer (SSC) with 10% formamide (Ambion), and then 2x SSC supplemented with DAPI (Molecular Probes D3571) to stain the cell nuclei. Monolayer cells cultured in coverglass chambers were submerged in 2x SSC for imaging. Micro-gels were mounted in 2x SSC and compressed between a coverglass and slide for imaging. Cells in the micro-gel and small colonies were imaged using a Leica DMI600B automated widefield fluorescence microscope equipped with a 100x Plan Apo objective, a Pixis 1024BR

cooled charge-coupled device (CCD) camera, a Prior Lumen 220 light source, and filter sets specific for each fluorophore. Images in each fluorescence channel were taken as a series of optical z-sections (0.5-0.7 microns per section) spanning the vertical extent of each cell. To prevent differences in viability between conditions from confounding interpretation of single-cell gene expression, the fixable dead cell stain was used to establish a GAPDH copy number of > 10 mRNA as a threshold to identify live cells for inclusion in further analysis (**Figure 4-2d**). When this FISH analysis was applied to live-imaged cells, single plane scans were performed using a Nikon Ti-E microscope with a 63x Plan Apo objective.

Quantification of Copy Number from RNA FISH Images

Upon collecting images of RNA FISH samples, cell boundaries were manually identified and RNA spots were counted and localized using custom software written in MATLAB[174]. For spot counting in FISH images from live cell tracking, each cell was tracked through the acquired time series, and sister cells manually matched, with care taken to note the time since last division.

Quantification of Extracellular Matrix Deposition

Extracellular aggrecan protein content was quantified by immunostaining. Briefly, after the final wash stages of the FISH protocol, samples were incubated with primary antibody (Abcam ab3778, 1:50 in PBS) at 4^{o}C overnight, washed for 30 minutes in PBS, incubated with Alexa 488 secondary antibody (Invitrogen, 1:200 in PBS) at room temperature for 1 hour, washed with PBS for 30 minutes, and then mounted for imaging. For immunofluorescence images, a scorer blinded to the RNA FISH images examined

the DAPI, GFP (aggrecan core protein), and transmitted light images to classify cells with and without extracellular aggrecan core protein staining.

Receiver operating characteristic curves were constructed and analyzed using the pROC package in R[181]. Matrix deposition (high vs. low) was used as the binary outcome, and sensitivity and specificity were calculated for possible thresholds of RNA copy number, a linear combination of RNA counts, and RNA ratios. To construct the linear combination, the data sets corresponding to each batch (Batch 1: Donors D-F, assayed for aggrecan, osteopontin, LPL, GAPDH; Batch 2: Donors X-Z, assayed for COMP, Sox9, and GAPDH) were randomly split in half to create training and test data sets, to be used for model construction and evaluation respectively. Logistic regression was performed using glm in R, and non-significant terms were dropped. For batch 1, the final model was established as (Equation 1):

$$\ln\left(\frac{\widehat{\pi_i}}{1-\widehat{\pi_i}}\right) = \beta_{\mathrm{ACAN}}(aggrecan_i) - \beta_{\mathrm{OPN}}(osteopontin_i) - 1.52 \qquad (1)$$

where $\widehat{\pi_i}$, the estimated probability of the i-th cell having high matrix staining, was a function of the cell's aggrecan and osteopontin expression. Aggrecan associated positively (β_{ACAN} = 0.003, p<0.001), while osteopontin associated negatively (β_{OPN} = -0.001, p=0.05); other markers were not significantly associated with matrix deposition. For batch 2, the intercept was the only significant term and the model was not further analyzed. Having established this model on the training data set, its predictive performance was evaluated by constructing an ROC curve of the model applied to the test data set.

RNA sequencing

Poly-adenylated RNA from passage 1 clonal MSCs populations were isolated from monolayer culture in basal media. The Qiagen miRNeasy kit was used for RNA isolation, the NEBNext Poly(A) mRNA Magnetic Isolation Module was used for selection of poly-adenylated transcripts, and NEB Next Ultra Library Preparation Kit for Illumina was used for library preparation. Each sample was sequenced with 50 bp single-end reads on an Illumina HiSeq and to a depth of 15–25M reads. Reads were aligned to bosTau7 using STAR[47]. Reads per gene were quantified using HTSeq and a RefSeq bosTau7 from annotation release 103 [4]. FPKM (fragments per kilobase of transcript per million mapped reads) for each gene was calculated using R.

Cell Volume and Area Measurements

Chondrocyte area was measured in ImageJ by manually tracing images of phalloidin-stained cells sparsely plated onto glass coverslips. Chondrocyte suspended cell volumes were computed from the cell radii measured by an automatic cell counter (Nexcelom Cellometer) for chondrocytes in solution during passaging (immediately following trypsinization).

Alcian Blue Staining

Micro-gels were removed from 70% ethanol and equilibrated in 3% acetic acid for 30 minutes at room temperature. Gels were then transferred to Alcian blue solution (pH 1.0, Rowley Biochemical) for 30 minutes, washed three times in acid alcohol (1% hydrochloric acid in 70% ethanol) for 30 minutes, and then washed in PBS for 30

minutes before imaging. For macroscopic images, gels were photographed using a Ricoh photocopier and digital camera. For microscopic images, micro-gels were mounted in PBS and compressed between a coverglass and slide. Images were taken at 10x using a Nikon Eclipse Ni-E motorized upright microscope.

Statistical Comparisons

To compare mean single cell RNA counts, a generalized linear mixed model with a log-link function and by-donor random intercepts was constructed. For MSC RNA counts, media condition, and culture duration (with interaction term) were considered fixed effects, and an additional by-donor random slope effect was associated with media (-TGFβ vs + TGFβ). For MSC RNA divergence, time-since-division was considered a fixed effect. For chondrocyte RNA counts during dedifferentiation in monolayer, passage was considered a fixed effect and was also associated with a by-donor random slope. For chondrocyte RNA counts during re-differentiation, passage and culture duration (with interaction term) were considered fixed effects, and an additional by-donor random slope effect was associated with passage. In each model, estimated means were compared using Satterthwaite-based t-distributions with simulated adjustment for multiple comparisons (SAS Studio 3.3). Pooled chondrocyte aggrecan/GAPDH expression data were compared using a one-way ANOVA with Tukey's post-hoc test. Chondrocyte area and volume data were pooled across donors and compared using a one-way ANOVA with Tukey post-hoc tests. Sample size was chosen based on previous experience with these assays. Details of all statistical comparisons are provided in the supplementary tables included in Appendix 2.

Chapter 5: High fidelity visualization of cell-to-cell variation and temporal dynamics in nascent extracellular matrix formation

Introduction

In tissues throughout the body, the extracellular matrix (ECM) guides cell phenotype and imparts mechanical resilience over a lifetime of load-bearing use. These extracellular matrices are highly dynamic, and change in both structure and molecular composition as development progresses, and with aging and disease processes. In articular cartilage, the ECM transitions from a fibronectin-rich environment in early development, to one dominated by aggrecan and collagen II at tissue maturity[173]. Notably, in both developing and mature cartilage, matrix synthesis and turnover occur continuously, and are requisite for tissue homeostasis [152]. Unfortunately, this homeostasis is often disturbed by injury- and degeneration-induced damage to the cartilage matrix and its resident cells. Such damage fails to intrinsically heal, and has prompted the development of engineered cartilage replacements.

In the context of cartilage tissue engineering, chondrocytes and progenitor cells must not only create matrix, but also retain and assemble it in the pericellular space. The rates of ECM production, retention, and degradation define how rapidly an engineered construct can mature. Thus, the manner in which the matrices produced by individual cells interact and integrate with one another ultimately defines the functional properties of the tissue that forms [57,194]. Moreover, just as the *in vivo* ECM influences cell phenotype in native tissue, the structure and composition of the matrix in these *in vitro* constructs regulates

the extent and progression of chondrogenesis [21]. Thus, heightened understanding of matrix protein synthesis and remodeling is relevant to contexts spanning development, disease, and tissue engineering.

Towards the quantification matrix dynamics, ECM formation can be monitored via bulk biochemical measures across time and disease state. However, such ensemble approaches mask cell-to-cell variation and do not provide information regarding the spatial organization of the matrix. Alternatively, autoradiography with radiolabeled sulfate and proline can provide insight into the localization of proteoglycans and collagens around individual cells, and has demonstrated temporal changes in the rate and spatial distribution of secreted matrix [19,171,172]. However, this approach is inherently complicated by its use of radioisotopes. Moreover, the punctate pattern of autoradiographic grains offers limited information regarding the structure and organization of this nascent extracellular matrix.

To overcome these limitations, we introduce the use of a metabolic labeling approach, functional noncanonical amino acid tagging (FUNCAT), to enable high fidelity fluorescent observation of nascent extracellular matrix protein accumulation and assembly. Previously, FUNCAT has been used to visualize protein synthesis and intracellular trafficking in cell monolayers [10,45,215,234] , bacteria [135], larval zebrafish [83], and drosophila [55]. FUNCAT relies on “residue-specific” incorporation of non-canonical amino acids (ncAA) into proteins as they are synthesized [94]. While many ncAAs exist and collectively offer a diverse suite of functions, the ncAAs utilized in FUNCAT are restricted to those that contain bio-orthogonal functional groups that can be detected by highly selective

fluorescent tags following ncAA incorporation. Operationally, FUNCAT ncAA incorporation resembles pulse labeling: a canonical amino acid (cAA) is removed from the environment, and is replaced with a corresponding ncAA [10,94]. In the absence of the cAA, the endogenous translation machinery of the cell incorporates the ncAA into proteins during synthesis, yielding global incorporation of the ncAA across the nascent proteome [94]. This strategy contrasts with "site-specific" ncAA incorporation, which utilizes genetic manipulation to substitute ncAAs in targeted locations, and more advanced residue-specific strategies that rely on engineered biosynthetic machinery to incorporate ncAA [94].

In this study, we adapt the FUNCAT technique to enable the fluorescent visualization of extracellular matrix proteins in both native cartilage and in 3D engineered constructs. Our results demonstrate that the FUNCAT method enables high fidelity labeling of extracellular matrix proteins throughout the time course of matrix formation and homeostasis. We use this labeling approach to query cell-to-cell heterogeneity in matrix formation and to determine how the density of the microenvironment, crosslinking of nascent ECM proteins, and the pre-established ECM influence matrix protein distribution and assembly on a single cell basis using both primary chondrocytes and mesenchymal stem cells undergoing chondrogenic differentiation.

Results

Methionine analogs enable the fluorescent labeling of extracellular matrix proteins.

Current implementations of FUNCAT rely on the substitution of the cAA methionine, an amino acid that comprises between 0-6% of the residues in ECM proteins (**Figure 5-1a**, **Appendix 3**). For example, methionine constitutes 1.1% of fibronectin and 1.8% of collagen IV, an ECM protein found primarily in basal lamina. For articular cartilage, our tissue of interest, methionine represents ~1% of the amino acid content [16,99], with similar relative abundance amongst the major cartilage ECM constituents collagen II (1.08%) and aggrecan core protein (0.54%). In the present work, we separately considered labeling with two bio-orthogonal methionine analogs: homopropargylglycine (HPG) and azidohomoalanine (AHA)[45]. These analogs are structurally and functionally similar to native methionine, but include an alkyne and azide side chain respectively.

To assess the ability of HPG and AHA to identify nascent ECM proteins, we first cultured chondrocytes in thin (~400 μm thickness) 2% agarose gels for 9 days in the absence of native methionine and the continuous presence of either HPG or AHA. Following HPG or AHA culture and sample fixation, we identified incorporated HPG and AHA by labelling alkyne or azide residues with fluorescent azide or alkyne tags, respectively, via a copper-catalyzed azide-alkyne cycloaddition ("click") reaction. This staining procedure fluorescently and covalently labeled proteins synthesized in the presence of each ncAA, and was combined with additional fluorescent staining to identify the cell nuclei as well as the cell membrane (to distinguish intracellular proteins from extracellular matrix

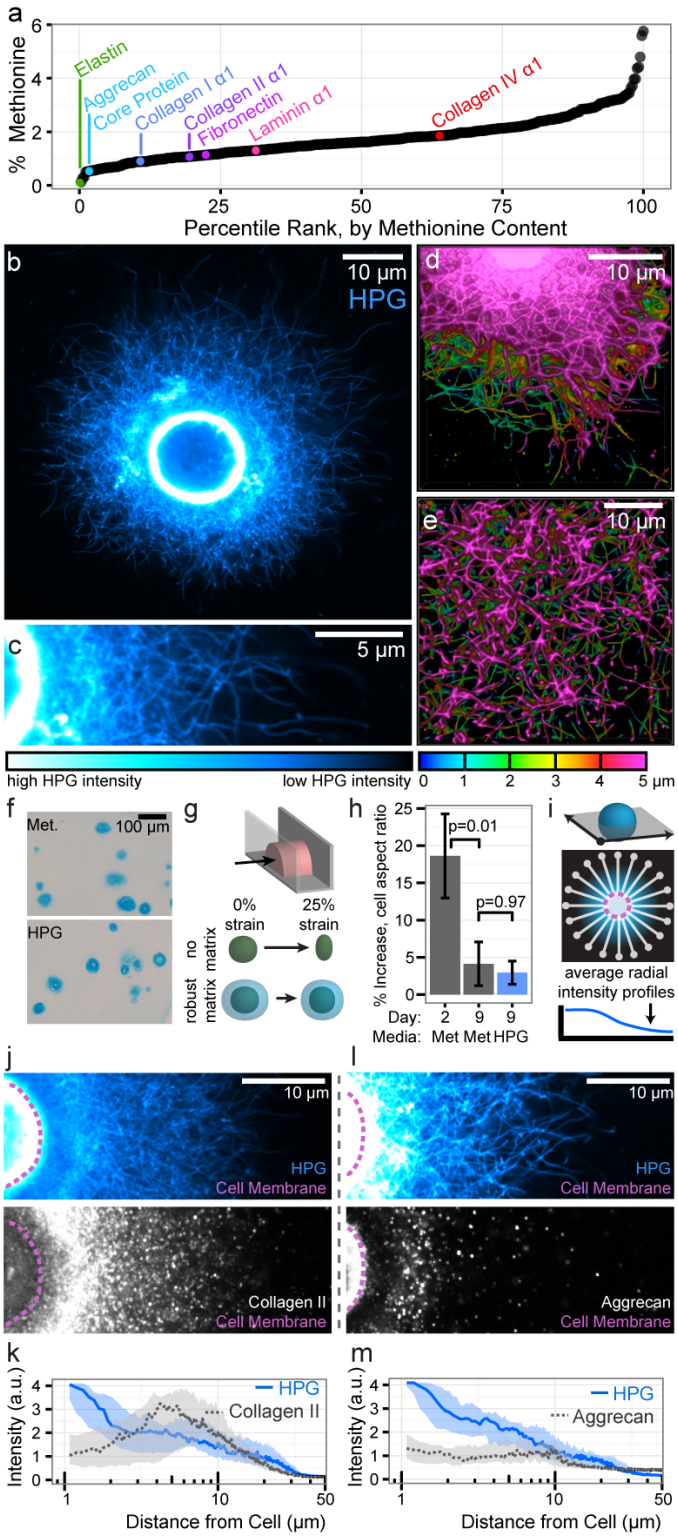
a
% Methionine
Elastin
Aggrecan
Core Protein
Collagen I α1
Collagen II α1
Fibronectin
Laminin α1
Collagen IV α1
Percentile Rank, by Methionine Content
b
10 μm
HPG
c
5 μm
d
10 μm
e
10 μm
high HPG intensity
low HPG intensity
0 1 2 3 4 5 μm
f
Met.
100 μm
HPG
g
0% strain
25% strain
no matrix
robust matrix
h
% Increase, cell aspect ratio
p=0.01
p=0.97
Day: 2 9 9
Media: Met Met HPG
i
average radial intensity profiles
j
10 μm
HPG
Cell Membrane
Collagen II
Cell Membrane
l
10 μm
HPG
Cell Membrane
Aggrecan
Cell Membrane
k
Intensity (a.u.)
HPG
Collagen II
Distance from Cell (μm)
m
Intensity (a.u.)
HPG
Aggrecan
Distance from Cell (μm)

Figure 5-1 (previous page): HPG labeling enables high fidelity visualization of the extracellular matrix. (a) Percentage methionine content of 361 extracellular matrix proteins, ordered by relative methionine abundance. (b, c) Confocal cross section of a chondrocyte cultured in agarose and continuously labeled with HPG for 9 days. (d, e) 3D reconstruction of 5 μm confocal stacks taken near the cell midplane (d) and below the cell (e) of day 9 chondrocytes. Color indicates vertical position in the stack. (f) Alcian blue staining of constructs cultured in control media with native methionine and labeling media containing HPG. (g) Schematic illustrating microcompression to assess extracellular matrix mechanics. (h) Percent increase in cellular aspect ratio following compression. Bars represent mean ± SEM (n=40 cells/group, compared via ANOVA with Tukey's post hoc test). i) Schematic illustrating radial profile quantification of extracellular HPG labeling. (j, k) Images and radial profile quantification of simultaneous collagen II immunostaining and HPG labeling. (l, m) Images and radial profile quantification of simultaneous aggrecan immunostaining and HPG labeling. For (k) and (m), lines represent median intensity profile, shaded areas represent 25th to 75th percentiles (n = 20 cells/group).

components). Stained gels were whole-mounted and imaged via confocal microscopy at high magnification (40-100X). Scans across the surface and through the depth suggested that labeling intensity was uniform throughout the gel (**Figure 5-2**). Individual cells exhibited extensive extracellular staining that had a clear fibrous structure (**Figure 5-1b-e**). Near the cell midplane (**Figure 5-1b-d**), densely packed HPG-tagged proteins extended outward from the cell body, forming a mesh-like structure. Labeled proteins often undulated through multiple z-planes, emphasizing the 3D nature of forming ECM. Below the cell (**Figure 5-1e**), protein was more loosely organized and fibrous structures were vertically oriented towards the cell body. The labeling patterns of HPG and AHA were similar (**Figure 5-3**), and so HPG was utilized for the remainder of the studies. Control samples cultured with native methionine instead of HPG showed minimal click reaction staining, confirming that the labeling is highly specific, with minimal off-target binding or labeling (**Figure 5-4**).

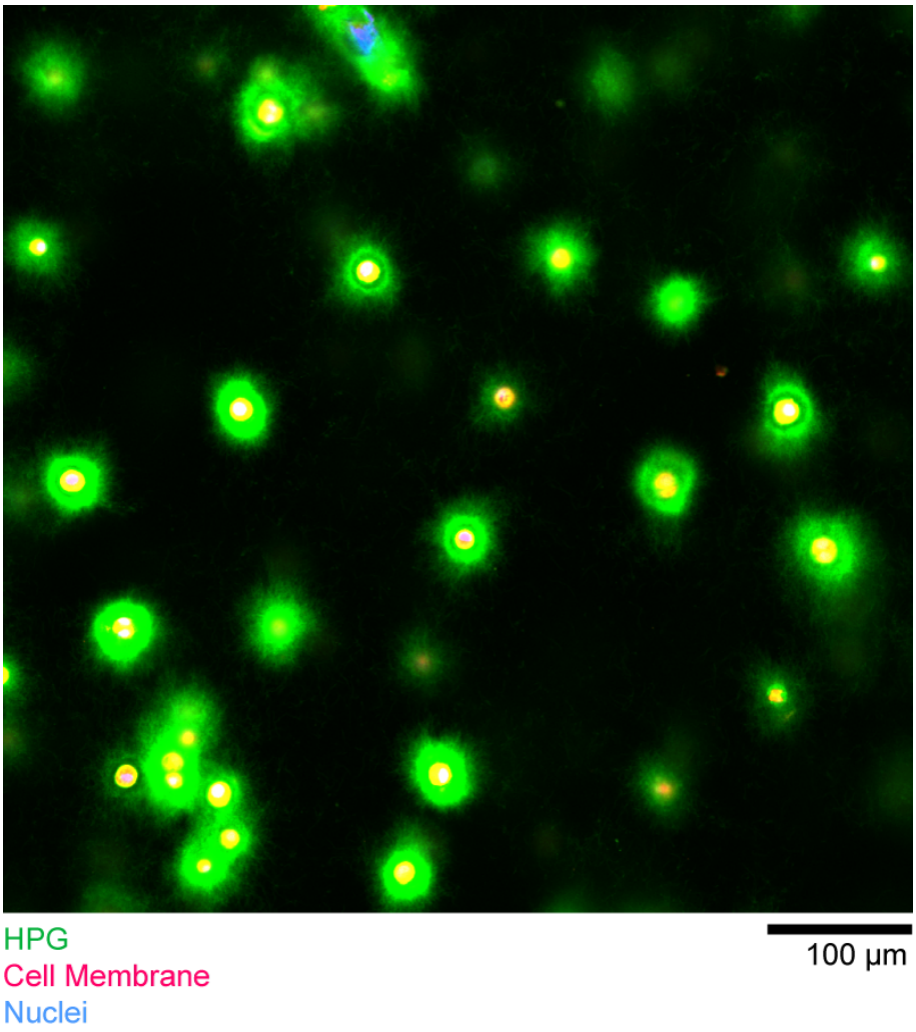

Figure 5-2: Large area scans confirm uniform labeling across gel. Chondrocytes in agarose, cultured in HPG labeling media for 9 days. Scan taken at 40X, ~ 50 µm into the gel.

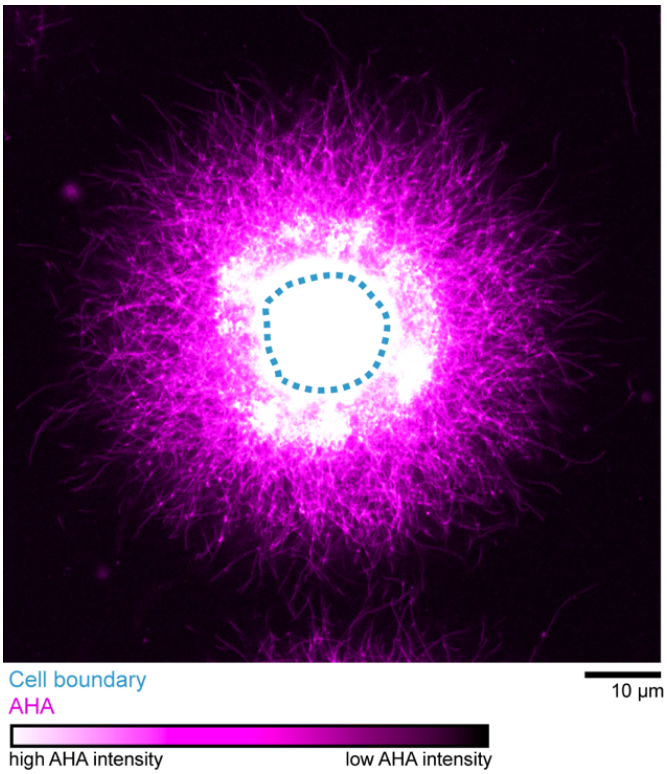

Figure 5-3: Matrix labeling with azidohomoalanine (AHA). Mid-plane confocal cross-section of a chondrocyte cultured in AHA labeling media for 7 days and tagged with Alexa Fluor 594-alkyne.

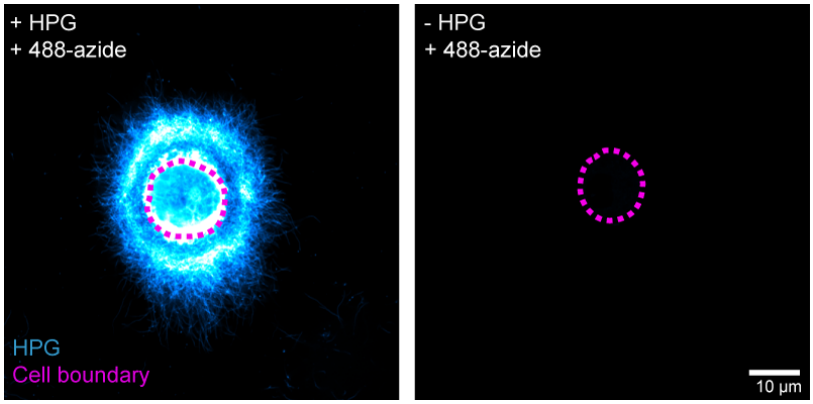

Figure 5-4: Fluorescent tagging of HPG is highly specific. Mid-plane confocal cross-sections of two chondrocytes, cultured in either labeling media (with HPG, left) or control methionine media (no HPG, right) for 9 days, and fluorescently tagged with 488-azide.

To confirm that the substitution of methionine with HPG did not impact matrix accumulation, we compared gels cultured for 9 days in either labeling media (with HPG) or control media (with native methionine). Matrix accumulation, marked by Alcian blue staining of histological sections, was similar between the labeling and control groups (**Figure 5-1f**). Furthermore, HPG incorporation did not alter the mechanical function of the formed matrix. We performed microcompression testing to monitor matrix stress shielding and infer the mechanical properties of the nascent ECM [107]. In this assay, cells with little or weak ECM deform readily with bulk compression of the gel, while cells with robust pericellular matrix accumulation deform much less (**Figure 5-1g**). At day 2, cells deformed readily in response to applied strain (18.6% ± 5.6%). By day 9, cell deformation was attenuated (4.1% ± 2.9%, $p = 0.01$ vs Day 2), and the extent of this attenuation was nearly identical between control and labeled groups (labeled: 2.9% ± 1.5% , $p = 0.97$ vs Day 9 Methionine, **Figure 5-1h**). Thus, the matrix produced in the presence of HPG was similar in both its distribution and mechanical function compared to matrix formed in standard culture conditions.

Because HPG should label all methionine-containing proteins, we next asked how the spatial pattern of HPG labeling compared with that of specific extracellular matrix proteins. For this and all subsequent studies, individual cells were identified via nuclear staining and imaged through the midplane of the cell body (**Figure 5-1i**). Simultaneous staining for HPG and either aggrecan core protein (**Figure 5-1j,k**) or collagen II (**Figure 5-1l,m**) emphasized the high degree of structural detail revealed by FUNCAT in comparison with traditional staining methods. Aggrecan core protein was restricted to

the pericellular space, an area where HPG labeling was often most intense (**Figure 5-1j,k**). In contrast, collagen II co-localized with the outer reaches of HPG labeling, both qualitatively (**Figure 5-1l,m**) and as quantified by average intensity profiles directed radially from the cell surface and extending into the gel (**Figure 5-1k,m**). Pre-treatment of fixed samples with hyaluronidase revealed additional collagen II staining closer to the cell, but did not influence the outer radius of staining, nor did it alter the co-localization between collagen II and HPG at the matrix boundary (**Figure 5-5**). Collectively, the tight co-localization of HPG signal with prevalent matrix proteins was consistent with the expectation that HPG would incorporate into, and label, the proteinaceous components of the extracellular matrix.

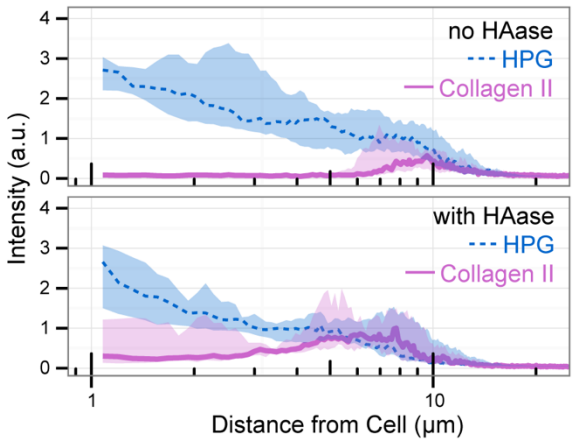

Figure 5-5: Hyaluronidase digestion of fixed samples increases collagen II detection by immunofluorescence, but does not alter HPG labeling. Radial intensity profiles for chondrocytes cultured in agarose and labeled with HPG for 9 days. Following fixation, samples were incubated in either PBS (control) or hyaluronidase before immunostaining and HPG tagging. Lines represent median intensity profile, shaded areas represent 25th to 75th percentiles (n = 9-11 cells/group).

Metabolic labeling tags pericellular collagenous network.

To better understand the identity and relative organization of the extracellular proteins labeled with HPG, we enzymatically digested chondrocyte-generated ECM prior to sample fixation on day 9. Alcian blue staining confirmed the efficacy of digestion with collagenase, hyaluronidase, and chondroitinase ABC (**Figure 5-6a**). Collagenase digestion of the fibrous collagen network completely removed HPG-labeled proteins accumulated near the cell, and yielded short fibrous protein fragments that were distributed throughout the gel (**Figure 5-6b,c**). In contrast, digestion with either hyaluronidase or chondroitinase ABC had only subtle effects on the pattern of HPG labeling (**Figure 5-6b,c**). Because hyaluronan and chondroitin sulfate are not proteins, they are not tagged directly by HPG labeling. Instead, hyaluronan digestion would be expected to disrupt pericellular proteoglycan aggregates, while chondroitinase ABC would be expected to remove the fixed negative charge density and cause a loss of charge-bound proteins in the pericellular space [229]. Thus, these results suggest that, at this time point, the majority of extracellular proteins labeled by HPG are collagens, or proteins that rely on collagens to be retained in the pericellular space.

Biophysical features of the cellular microenvironment influence matrix organization.

To assess the ability of HPG labeling to identify differences in matrix catabolism and organization, we next used continuous HPG exposure to monitor the time course of matrix accumulation by chondrocytes. With time, extracellular proteins progressively extended from the cell (**Figure 5-7**). Matrix distributed symmetrically around individual

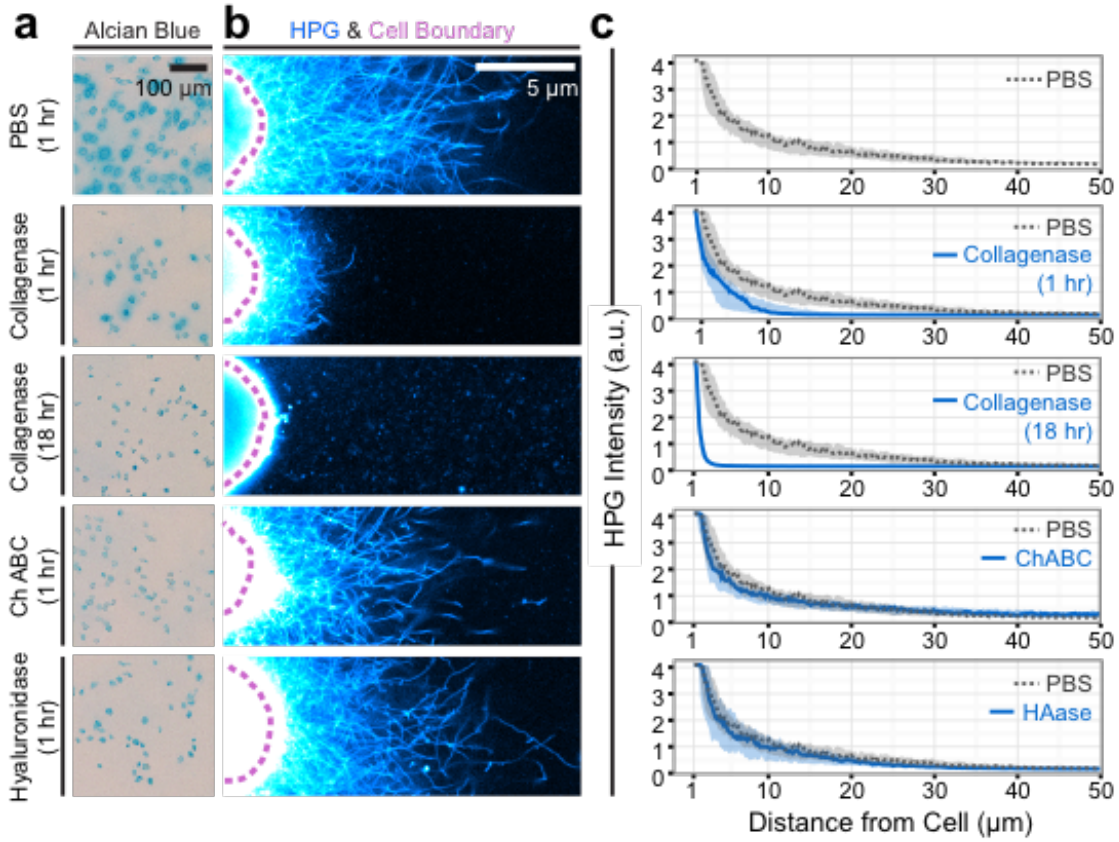

Figure 5-6: Enzymatic digestion differentially effects HPG-labeled matrix. (a) Alcian blue staining of chondrocyte/agarose constructs cultured and labeled with HPG for 9 days followed by enzymatic digestion prior to fixation. (b) Corresponding visualization of HPG-labeled matrix in digested samples. (c) Radial profile quantification of HPG intensity following digestion, compared to constructs incubated in PBS. Lines represent median intensity profile, shaded areas represent 25th to 75th percentiles (n = 20 cells/group).

chondrocytes (**Figure 5-11**), and overall labeling intensity increased with time in culture (**Figure 5-7**). By day 7, HPG labeling in chondrocytes often exhibited a low-intensity band adjacent to the cell surface, potentially reflecting the formation of a pericellular matrix rich in proteoglycans.

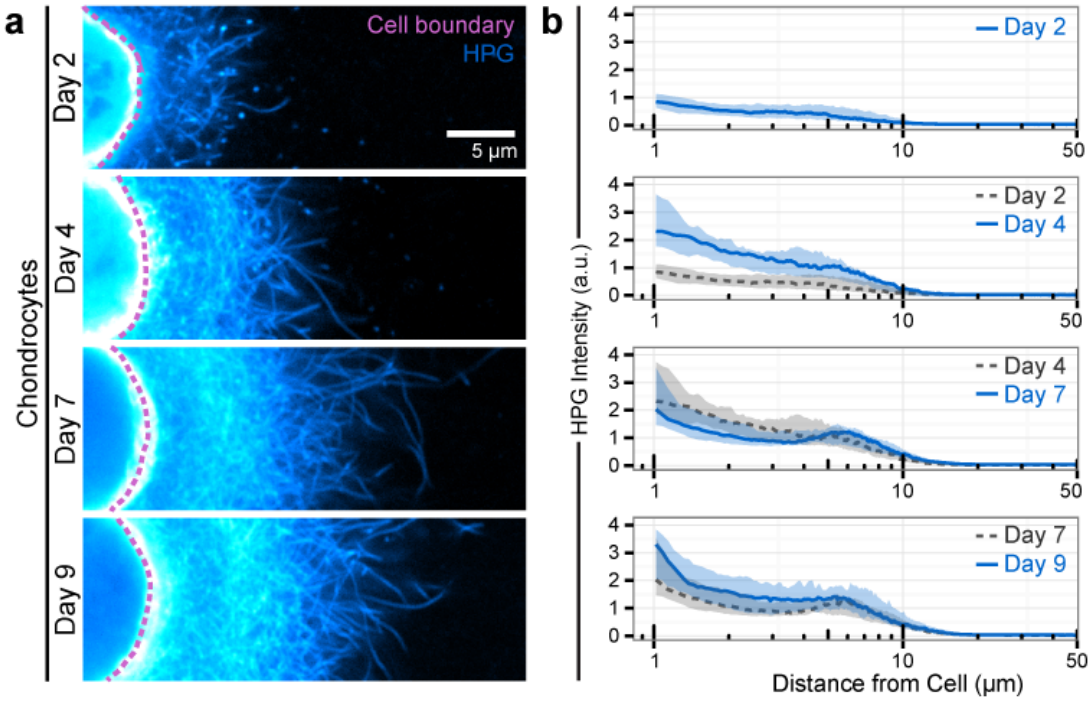

Figure 5-7: HPG labeling captures nascent matrix deposition. Images (a) and radial profile quantification (b) of chondrocytes cultured in 2% agarose and continuously labeled with HPG for up to 9 days. Colormap and scale bar are consistent across all images. Lines represent median intensity profile, shaded areas represent 25th to 75th percentiles (n = 20 cells/group).

The structure of this accumulating matrix was modulated by the cellular microenvironment, including the choice of the biomaterial scaffold and its physical properties. We have previously demonstrated that increasing hydrogel density promotes matrix formation, but prevents matrix distribution throughout the material 32. To examine this behavior at the single cell level, we cultured chondrocytes in 2% agarose gels (the standard condition), as well as 1% and 3% agarose gels, for 9 days. Extracellular proteins distributed readily throughout the 1% gel, resulting in a disperse protein network that extended between neighboring cells (**Figure 5-8**). Conversely, in denser 2% and 3%

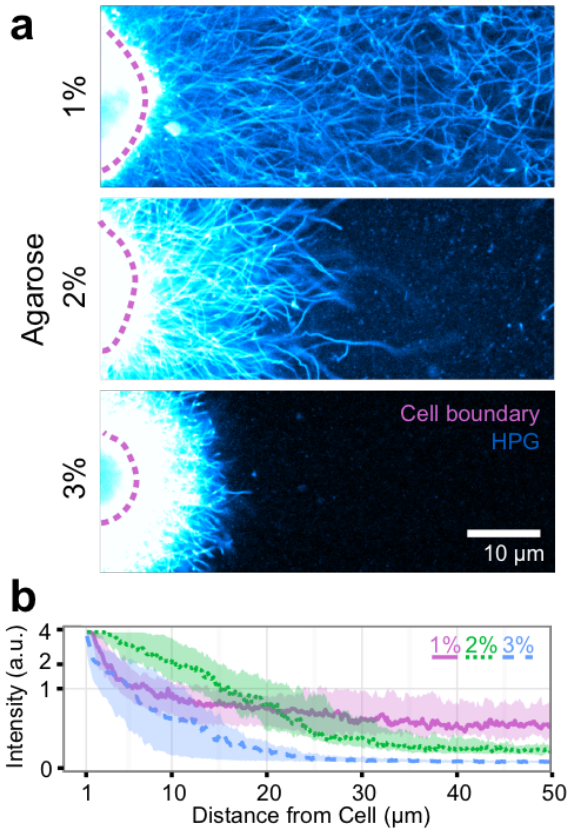

Figure 5-8: Density of the cellular microenvironment impacts organization of nascent matrix. (a) Images of HPG-labeled chondrocytes cultured in agarose of varying densities. (b) Radial profile quantification of HPG intensity in 1%, 2%, and 3% agarose hydrogels.

gels, individual cells retained discrete extracellular protein; accumulated protein was most compact in 3% agarose (**Figure 5-8**). Thus, as the microenvironment became increasingly dense and less permissive, matrix proteins were spatially restricted to their point of origin.

Nascent matrix intersperses with pre-existing matrix in developing (but not developed) microenvironments.

Next, we asked how proteinaceous matrix forms and is organized within the context of an existing extracellular matrix. That is, do matrix proteins accumulate in concentric rings (like growing trees), or does new matrix interdigitate with the pre-existing structure? To answer this question, gels were cultured in control media for 2, 7, 19, and 40 days and switched to HPG labeling media for the two final days of culture. Only proteins synthesized during the final two days would be labeled with HPG, while pre-existing protein would be unlabeled. Results from this study showed that, between 2 and 4 days, nascently produced fibrous proteins accumulated in the pericellular space (**Figure 5-9**). Fibrous structures also formed between days 7 and 9, and were distributed within the pre-existing matrix (**Figure 5-9a**). Between 19-21 days and 40-42 days, however, nascent proteins were primarily restricted to the immediate pericellular space and lacked a fibrous structure (**Figure 5-9b**). This likely reflects an increasingly dense microenvironment and a potential shift in the molecular composition of the nascent matrix. Cartilage explants cultured for 3 days in labeling media showed little evidence of nascent proteins in the extracellular space (**Figure 5-9c**), suggesting that matrix turnover in mature tissue occurs on a timescale longer than the labeling period. Collectively, these results are consistent with the notion that the permissivity of the existing matrix dictates nascent matrix elaboration and organization. At early time points in culture, cells form matrix rapidly and incorporate this newly formed material within the existing matrix. As the existing matrix matures, increased matrix density restricts nascent proteins to the more immediate pericellular space.

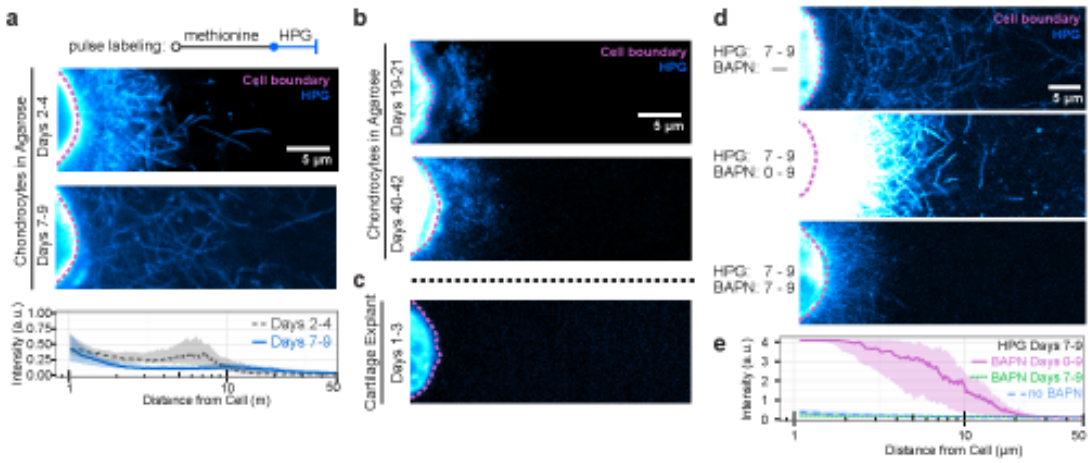

Figure 5-9: Extent and organization of nascent matrix assembly depend on construct maturity and collagen cross-linking. (a,b) Images (a) and radial profile quantification (b) of chondrocytes cultured in agarose for 4 and 9 days that were pulse-labeled with HPG for the final 2 days of culture. (c) Chondrocytes cultured in agarose for 21 and 42 days, and pulse labeled for the final 2 days of culture. (d) Pulse-labeled chondrocyte in a cartilage explant, cultured for 3 days. (e) Pulse-labeled (final 2 days of culture) chondrocytes cultured in the presence or absence of BAPN, a collagen cross-linking enzyme inhibitor. BAPN was administered either continuously or over the final 2 days of culture. (f) Radial intensity profiles of day 7-9 pulse labeled chondrocytes in the presence and absence of BAPN. Colormaps are consistent within subfigures. Lines represent median intensity profile, shaded areas represent 25th to 75th percentiles (n = 20 cells/group).

Because matrix formation and accumulation are highly regulated processes, we also asked how perturbation of normal matrix assembly might alter nascent matrix protein synthesis and organization. To do so, we used β-aminopropionitrile (BAPN) to inhibit the collagen crosslinking enzyme, lysyl oxidase. Here, continuous exposure to BAPN for 9 days dramatically altered matrix organization, but did so in a non-intuitive manner. In the context of BAPN, continuous labeling revealed dense and intensely labeled proteins adjacent to the cell (**Figure 5-12**). The extent of extracellular protein deposition was sharply truncated at 7.9 ± 0.5 μm from the cell, suggesting that matrix proteins were unable to extend into the hydrogel as they had in control samples. Pulse labeling with

HPG from 7-9 days showed that continuous BAPN treatment (from day 0) influenced not only the spatial organization of matrix proteins, but also their synthesis dynamics. In pulse-labeled control groups, matrix protein labeling was of low intensity and interspersed within the pre-existing matrix (**Figure 5-9d,e; Figure 5-12**). Strikingly, BAPN treatment increased the amount of nascent extracellular proteins produced, and restricted these proteins to the immediate pericellular space. Intriguingly, these results suggest that collagen crosslinking is necessary for the cell to project collagen outward into the microenvironment, and for cells to achieve matrix homeostasis.

Patterns of matrix deposition identify phenotypic differences between cell types.

Having shown that chemical manipulation of matrix crosslinking fundamentally altered the distribution of extracellular matrix proteins, we next wondered if FUNCAT labeling could also detect phenotypic differences between cell types that would be reflected in their nascent matrix production and assembly. For example, using standard histological techniques and biochemical assays, we have previously shown that chondrocytes and mesenchymal stem cells (MSCs) differentially produce and organize their ECM [56]. To examine this phenomenon with higher fidelity and with single cell resolution, we next cultured MSCs and chondrocytes in the presence of HPG. Continuous labeling between days 0-9 showed that matrix protein accumulation by MSCs lagged behind that of chondrocytes, and differed in its spatial organization. Notably, there was high MSC-to-MSC variation in the labeled extracellular protein present (**Figure 5-10**a,c, **Figure 5-13a,b**), consistent with the heterogeneous nature of this cell population [36,164,180]. This

heterogeneity manifested in two ways: in both the intensity of the labeled matrix proteins, and the pattern of matrix protein distribution. Select MSCs produced labeled fibrous proteins by day 9, but these molecules were often found far away from the cell, suggesting that MSCs may be less able than chondrocytes to sequester ECM within the pericellular space, or may have a higher level of matrix degradation and turnover. MSCs also sometimes distributed extracellular proteins asymmetrically around the cell body (45% symmetric, 42% asymmetric, 13% without matrix at Day 9; **Figure 5-13a**). In contrast, the organization of continuously labeled matrix protein was very similar between chondrocytes, and was symmetric around individual cells (**Figure 5-11**). Consistent with continuous labeling, MSCs that were pulse labeled between days 7-9 showed high cell-to-cell variability in matrix accumulation patterns (**Figure 5-10b,d**, **Figure 5-13b**). Intriguingly, nascent matrix proteins in pulse labeled chondrocytes were also markedly heterogeneous (**Figure 5-10b,d, Figure 5-11b**) during this time point. These data indicate that, as the timescale considered shortens, apparent heterogeneity in matrix protein production and organization increases. Taken together, these observations suggest that continuous and pulsed HPG labeling can capture metabolic and phenotypic differences that may be of biological significance in the assembly of the cellular microenvironment at the single cell level.

Discussion

In this work, noncanonical amino acid tagging revealed cell-, time- and microenvironment-dependent patterns in extracellular matrix assembly and organization. Specifically, we considered the role of these factors in a chondrogenic context, with a

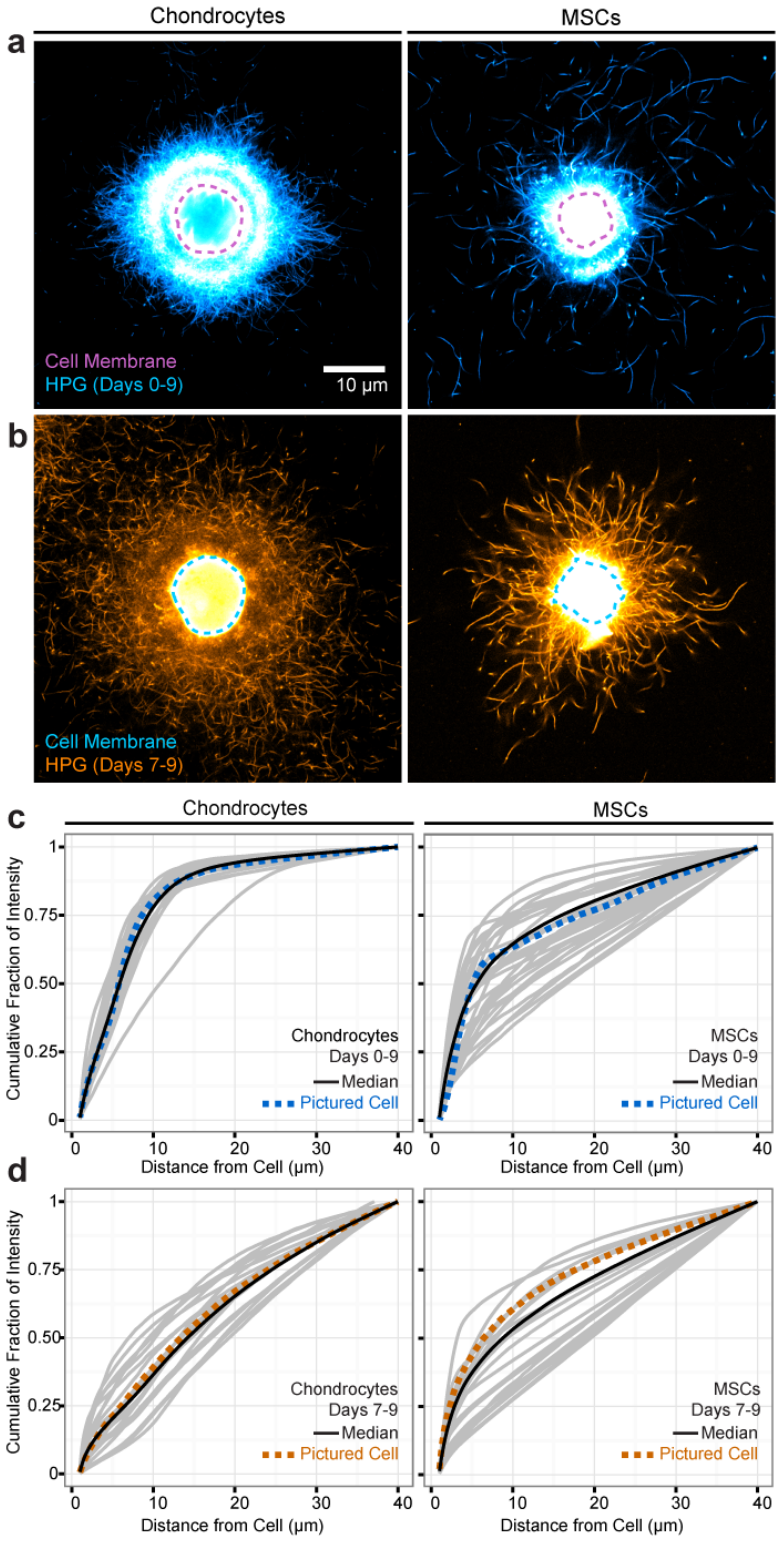
Chondrocytes
MSCs
a
Cell Membrane
HPG (Days 0-9)
10 µm
b
Cell Membrane
HPG (Days 7-9)
c
Chondrocytes
MSCs
Cumulative Fraction of Intensity
Chondrocytes
Days 0-9
Median
Pictured Cell
MSCs
Days 0-9
Median
Pictured Cell
Distance from Cell (µm)
d
Chondrocytes
Days 7-9
Median
Pictured Cell
MSCs
Days 7-9
Median
Pictured Cell
Distance from Cell (µm)

Figure 5-10 (previous page): Apparent heterogeneity in matrix accumulation is cell type and timescale dependent. (a,b) Representative chondrocytes and MSCs labeled with HPG continuously (a; blue) or between days 7-9 (b; orange). Pulse labeled samples were imaged with higher gain settings than continuously labelled samples, given that the shorter labeling period results in much lower overall intensity (see Fig S7 for direct comparison). (c,d) Quantified matrix distribution of individual cells, represented as the cumulative fraction of total intensity versus distance from the cell. Median is shown in black; representative cells from (a) and (b) are shown in dashed blue and orange lines. Other individual cells are represented in grey (n = 20 cells/group).

focus on how native chondrocytes and differentiating MSCs establish and respond to the cellular niche. Visualization of detailed matrix structures highlighted the differing nature of the proteinaceous ECM produced by these two cell types. Chondrocytes produced a well-organized, mechanically-robust matrix proteins that formed a physical barrier and ultimately restricted molecular mobility within the extracellular space. Interestingly, pulse labeling revealed persistent heterogeneity in the active remodeling of this network, which summated to a consistent profile of matrix protein distribution. In contrast, MSCs produced matrix proteins that were loosely organized, poorly retained and often asymmetrically distributed around individual cells. Similarly, on an individual cell basis, MSCs' nascent matrix proteins remained heterogeneous in both amount and distribution. Such differences in matrix protein organization likely impact not only the bulk mechanics that develop over time within the construct, but also, and perhaps more importantly, the biochemical and biophysical environment perceived by individual cells at a given time point. ECM organization modulates microscale transport properties, including the mobility and availability of soluble molecules such as growth factors [227]. Furthermore, matrix organization and connectivity influence how forces are transduced to the cell [79].

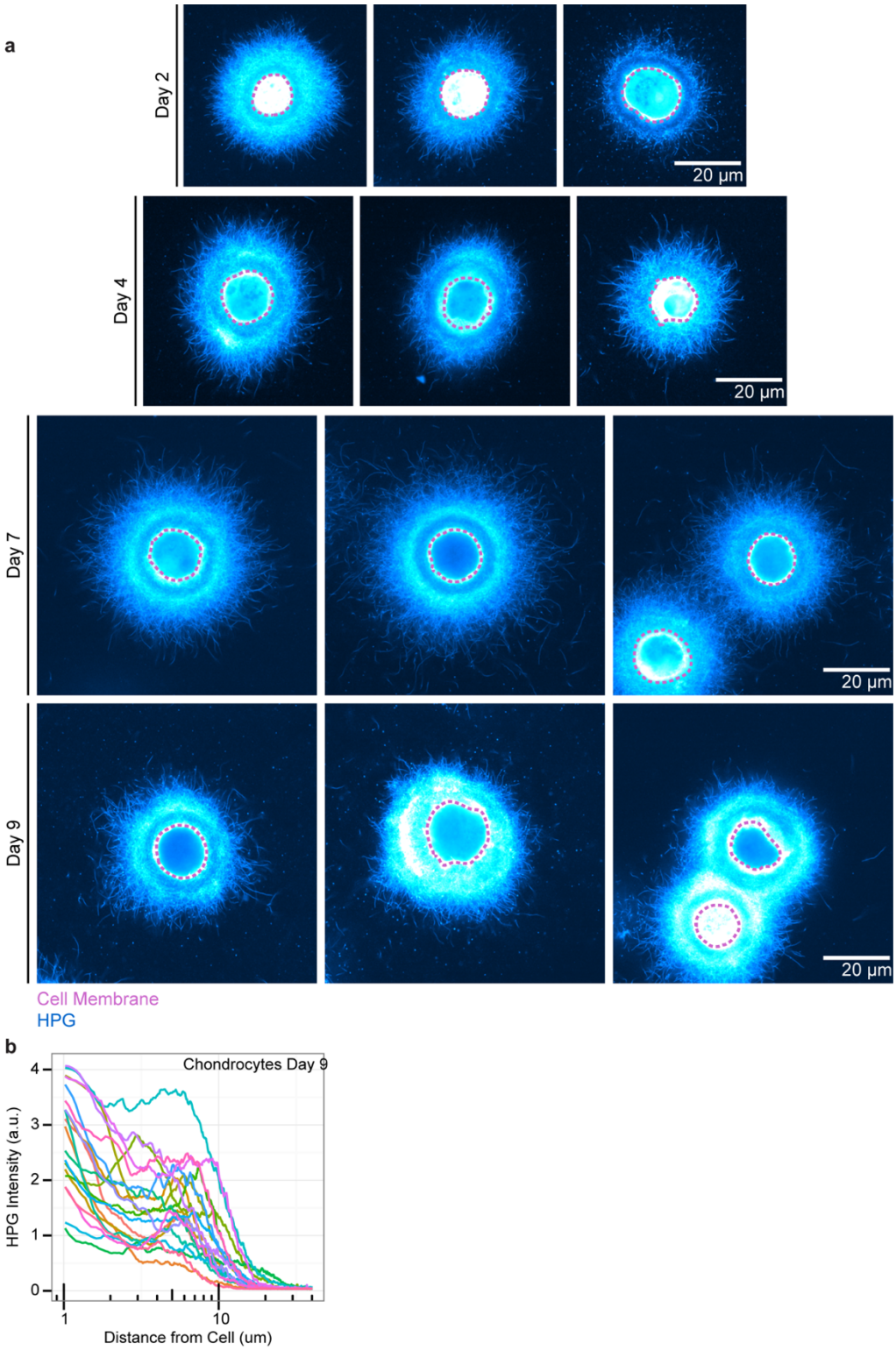
a
Day 2
Day 4
Day 7
Day 9
20 µm
Cell Membrane
HPG
b
Chondrocytes Day 9
HPG Intensity (a.u.)
0
1
2
3
4
1
10
Distance from Cell (um)

Figure 5-11 (previous page): Cell-to-cell heterogeneity in HPG-labeled nascent matrix formed by chondrocytes. (a) A panel of representative chondrocytes cultured in agarose and labeled with HPG for up to 9 days. Colormap and scale bar are consistent across all images. **(b)** Radial intensity profiles for 20 individual chondrocytes cultured and labeled for 9 days.

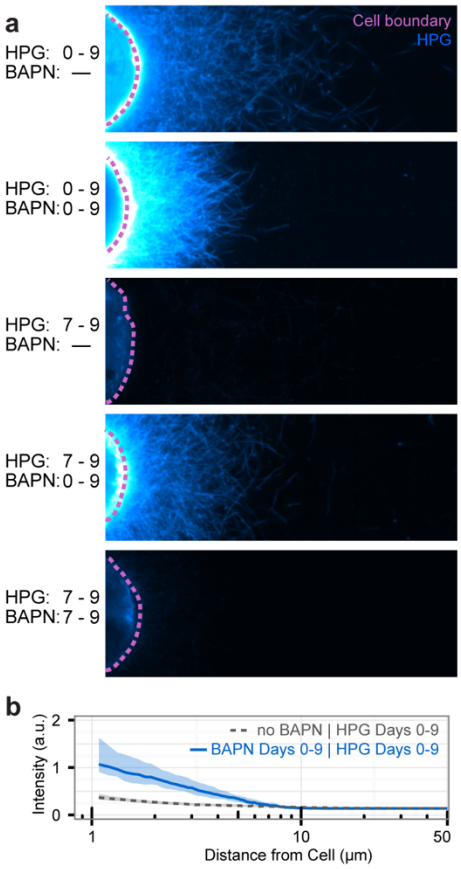

Figure 5-12: Exposure to BAPN increases nascent matrix synthesis. (a) Continuous and pulse labeling (final 2 days) of chondrocytes cultured in the presence and absence of BAPN, a collagen cross-linking enzyme inhibitor. Pulse labeled cells are the same cells shown in Figure 5, visualized here using the same imaging settings as the continuously labeled groups. **(b)** Radial intensity profiles of chondrocytes continuously labeled with HPG for 9 days in the presence and absence of BAPN. Lines represent median intensity profile, shaded areas represent 25th to 75th percentiles (n = 20 cells/group).

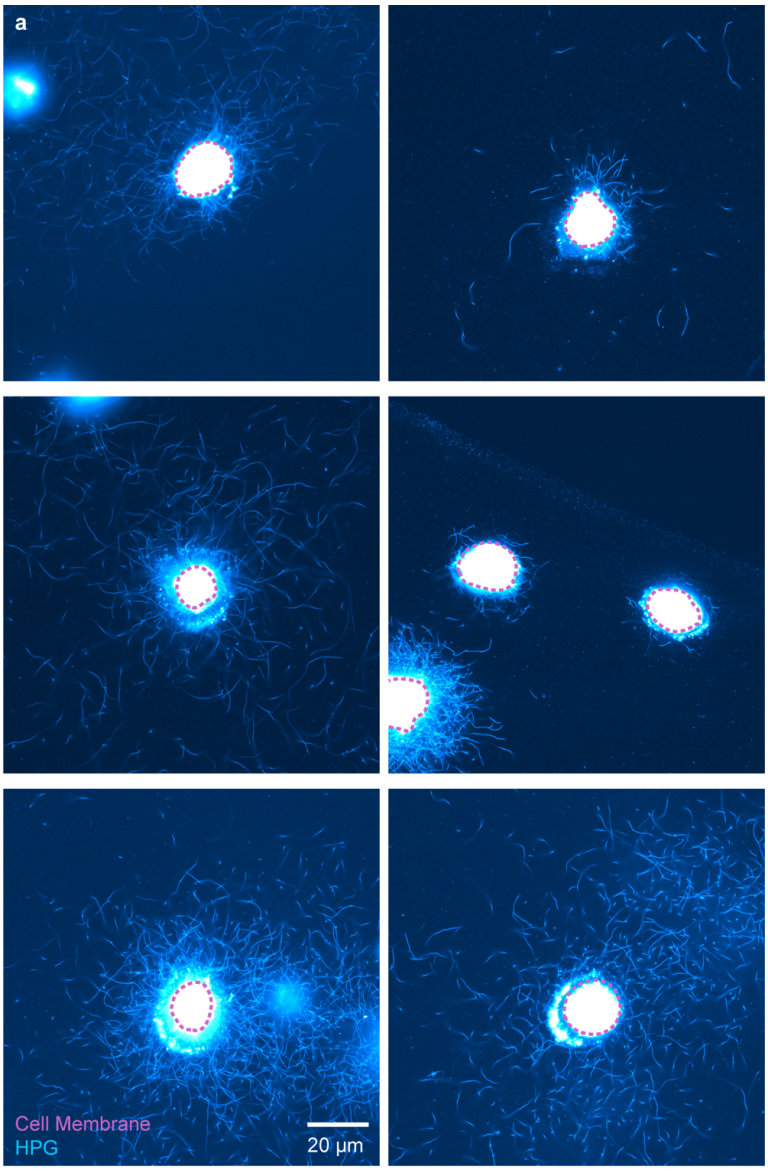
a
Cell Membrane
HPG
20 µm
b
MSCs Day 9
HPG Intensity (a.u.)
0
1
2
3
4
1
10
Distance from Cell (µm)

Figure 5-13 (previous page): Cell-to-cell heterogeneity in HPG-labeled nascent matrix formed by mesenchymal stem cells (MSCs). (a) A panel of representative MSCs cultured in agarose and labeled with HPG continuously for 9 days. **(b)** Radial intensity profiles for 20 individual MSCs cultured and labeled for 9 days.

Thus, the collective organization of matrix proteins is both a key product and determinant of cell phenotype, and is uniquely visualized by FUNCAT labeling.

Both construct maturity and biomaterial density regulated the distribution of nascent matrix proteins in the extracellular space. As the permissivity of the microenvironment decreased, nascent proteins were retained more closely to the cell. That is, the more permissive the microenvironment, the more readily newly-formed matrix could disperse throughout the construct. This finding may explain, in part, why the intentional removal of proteoglycans from developing cartilage constructs hastens subsequent construct growth [158]. Perhaps by decreasing the density of the pericellular microenvironment, mild digestion creates space for increased collagen deposition. Similarly, microenvironmental permissivity may play a role in matrix biosynthesis following injury and during degeneration. In the degenerative condition of osteoarthritis, the density of the extracellular matrix is decreases, weakening its mechanical integrity, particularly in the pericellular space [228]. At the same time, osteoarthritic cartilage has higher matrix synthesis activity than healthy tissue [127]. Thus, not only will osteoarthritic cells produce more nascent matrix, but this matrix will have greater mobility in less dense, degenerating tissue (consistent with observations of expanded pericellular matrix in degenerative cartilage [228]).

The utility of FUNCAT labeling also extends to the examination of how chemical and biological perturbations influence matrix assembly and organization. Previously, BAPN treatment of cartilage explants and alginate hydrogels indicated a reduction in the stable incorporation of nascent collagen into the tissue ECM or gel microenvironment [11,145]. Here, our results build on these existing bulk analyses and show that this defect in crosslinking manifests organizationally, with a buildup of matrix proteins in the immediate pericellular space, and a decrease in protein content further removed from the cell. It is possible that, upon BAPN washout, this pool of unincorporated collagen becomes available for rapid cross-linking and network assembly, potentially explaining why BAPN pre-treatment can improve integration strength between two pieces of cartilage [145]. Strikingly, fibrous matrix proteins organized perpendicularly to the cell membrane in both control and treated conditions. In tendon cells, the secretion of aligned matrix is a force-dependent process, requiring cellular contractility and a competent cytoskeleton [24]. Our finding that BAPN localized matrix proteins, but inhibited their wider distribution, suggests that collagen crosslinking is required to establish a functional framework against which newly formed matrix molecules are 'pushed' progressively further from the cell.

Modulation of cartilage cross-linking is only one of many possible matrix assembly perturbations. A similar labeling strategy could be used to finely assess the role of specific matrix components; for example, how does knockdown or knockout of individual proteoglycans or structural matrix proteins influence the organization and timing of matrix assembly? Notably, a FUNCAT-based approach would also be able to identify

contexts in which turnover is altered, but where total matrix protein content remains unchanged – a situation that traditional staining procedures would be unable to detect. This investigation of the biosynthetic response to injury, disease, or altered genetic program could be performed either *in vitro* (as we assessed tissue engineered construct formation here) or *in vivo*. Previous studies have examined protein synthesis in drosophila and zebrafish systems and mice [22,55,83]; scaling FUNCAT to mammalian systems with *in vivo* models of cartilage degeneration (e.g. meniscus destabilization) or joint development could lend insight into how matrix proteins form and reorganize during these processes. Recent studies indicate that HPG and AHA can incorporate into developing murine embryos [22], supporting the potential scalability of FUNCAT analysis to such model systems.

An important limitation of the FUNCAT procedure is that labeling is restricted to methionine-containing proteins. The vast majority of ECM proteins contain methionine: across a set of 361 standard ECM proteins, 85% contain >1% methionine, and 30% contain >2% methionine (Fig 1a and Table SX). Comparatively, the proteins most commonly found in the cartilage-like matrix (i.e. collagen II and aggrecan) are methionine-poor. Given that we were able to visualize, with high fidelity, ECM predominantly composed of these two molecules, FUNCAT should also be amenable to the examination of the ECM in other tissues, including those rich in fibronectin (1.1% Met), laminins (0.9-2.4% Met), and other collagens (0.6-3.1% Met). Indeed, proteomic analysis has indicated successful *in vivo* incorporation of AHA and HPG across tissues including heart, lung, brain, muscle, and kidney[22]. A notable exception is the protein

elastin, which contains only a single methionine residue; such low methionine content suggests that AHA- or HPG-based FUNCAT would be unable to fully detail matrix protein dynamics in elastin-rich tissues (e.g. arteries)[2]. Similarly, FUNCAT is also unable to directly capture the dynamics of non-proteinaceous ECM components, including glycosaminoglycans. Such components may be amenable to assessment through similar metabolic labeling strategies that incorporate azide-modified sugars to monitor nascent glycan synthesis [9,119,120,235].

A second limitation of the FUNCAT procedure described herein is its reliance on copper to catalyze the azide-alkyne reaction. Copper-catalyzed reactions are well-suited to fixed samples, such as those examined here, but have more limited utility in live-cell imaging applications, where metal toxicity impairs cell viability [92]. A cell-compatible alternative, copper-free click chemistry, relies on strain-promoted cycloaddition between a biomolecule containing azide (e.g. AHA) and cyclooctynes conjugated to fluorophores or affinity tags. Direct comparisons of copper- and copper-free click chemistries in proteomic applications suggest that the copper-free chemistry has decreased sensitivity and increased non-specific labeling in comparison to copper-catalyzed reactions [14,125]. Alternatively, the modification of copper-catalyzed click reactions to include ligand acceleration[14,84] or copper chelation[220] has recently improved cytocompatability to levels comparable with the copper-free chemistry, albeit with modified or additional reagents. Thus, the choice of specific cycloaddition chemistry will require one to establish an application-specific balance between desired biocompatibility, sensitivity, and reaction complexity.

In conclusion, our findings highlight that the spatial organization of extracellular proteins is a sensitive readout modulated by diverse stimuli. While our work focuses on chondrogenesis, the metabolic labeling approach used here offers the ability to query extracellular matrix formation by different cell types, in many different contexts, including tissue degeneration, wound healing, fibrosis, and other pathologic changes in tissue phenotype. Most importantly, given the fidelity and localization of this labeling, these alterations in normal tissue homeostasis can be examined at the single cell level, where such global changes in tissue structure and function first originate. Future applications of this technique will elucidate the spatiotemporal aspects of proteinaceous matrix assembly and remodeling, and contribute to our understanding of how dynamic cell-matrix interactions regulate tissue formation, degeneration, and repair.

Methods

Quantifying protein methionine content

Using the UniProtKB/Swiss-Prot database[219], human proteins of interest were identified using the gene ontology terms ‘extracellular matrix structural constituents’ (GO:0005201) and ‘proteinaceous extracellular matrix’ (GO:0005578). Methionine content was calculated by counting the number of methionine residues relative to the number of total amino acids in each protein sequence.

Construct fabrication

MSCs and chondrocytes were isolated from juvenile bovine knees (Research 87, Boylston MA) and passaged once before encapsulation in 3D gels. For all conditions,

cells were encapsulated at a density of 1.5 - 2x10^6 cells/mL. To obtain 1%, 2% and 3% (w/v) agarose gels, molten 4% agarose was combined with a warm cell suspension and solidified into thin sheets (~400 μm thick) underneath coverslips [36]. 4mm x 5mm μ-gels were cut from the sheets and cultured individually in 48 well plates.

Functional, noncanonical amino acid labeling

Samples were cultured in a chemically defined media: high glucose DMEM without glutamine, methionine, or cystine (Life Technologies 21013024) supplemented with either 0.1 mM L-methionine (Sigma M5308; control media) or 0.1 mM L-homopropargylglycine (Molecular Probes C10186; labeling media), 10 ng/ ml TGFβ-3, 0.1 μM dexamethasone, 4 mM L-glutamine, 0.201 mM cystine (Sigma C7602), 100 μg/mL sodium pyruvate, 1.25 mg/mL bovine serum albumin, 0.1% ITS premix, 50 μg/mL ascorbate 2-phosphate, 40 μg/mL proline, and 1% penicillin-streptomycin-amphotericin. Medium was replenished every 2-3 days, and finally fixed in 4% phosphate-buffered paraformaldehyde for 15 minutes before storage in PBS at 4°C. Cartilage plugs (4 mm diameter) were cultured in labeling media for 3 days, and fixed in 4% PFA overnight.

HPG incorporated into fixed samples was covalently tagged with Alexa Fluor 488. Samples collected up to day 9 were stained and imaged as intact gels. To enhance imaging, cartilage and gels collected at 21 and 42 days were cryosectioned to yield 40 μm thick cross-sections. Both intact gels and cryosections were first stained with a 1:1000 dilution of a plasma membrane stain (Molecular Probes C10046) in PBS for 30 minutes at room temperature. Next, samples were rinsed twice with PBS, and incubated

in a click reaction labeling solution (prepared from Molecular Probes C10428 according to product instructions, and including Alexa Fluor 488 azide) for 40 minutes at room temperature. Samples were washed in reaction rinse buffer (Molecular Probes C10428, 5 minutes at room temperature), and then once with PBS. Nuclei were labeled with Hoechst in PBS for 15 minutes at room temperature, and samples were washed twice with PBS before imaging. The same labeling procedure was followed for AHA samples, with Alexa Fluor 594 alkyne (Molecular Probes C10102 and A10275).

FUNCAT Imaging and Image Quantification

Samples were mounted in PBS and imaged using a Nikon A1 confocal microscope. To acquire large fields of view (**Figure 5-2**), confocal slices were captured at 40X. To visualize staining near individual cells, cells within 100 μm of the gel surface were identified via nuclear staining, and 100X confocal sections were taken through the cell midplane. To quantify the staining associated with each cell, 20 radial intensity profiles emanating from the cell center were mapped, truncated to include only the extracellular domain (demarcated by membrane staining), and averaged over each cell. Any encroaching matrix from nearby cells was manually excluded from this quantification. For easier visualization, post-processing was used to enhance the contrast; all images in a subfigure figure were imaged and post-processed identically. Quantified HPG intensity values shown in graphs were not transformed and are comparable between Figure 5-7 and Figure 5-9A. Sample size was selected based on a bootstrap resampling of a dataset of n=40 cells; samples sizes of n=5, 10, 20, 30 and 40 cells were simulated over 100 bootstrap replicates. Aggregate metrics were near-identical between n=20, 30, and

40 cells, suggesting that a sample size of n=20 cells per condition balanced sufficient statistical coverage with experimental efficiency.

To determine matrix radius, background signal levels were determined for each profile. Radius was considered as the first distance where intensity dipped below 110% of the background signal level. Profiles were averaged to determine the matrix radius of each cell; values reported are the mean ± standard error of 20 cells.

Alcian Blue Staining & Immunofluorescence

Histological sections or intact μ-gels were stained with Alcian blue pH 1.0 (Rowley Biochemical) to identify sulfated proteoglycans [36]. Immunofluorescence staining was performed following the cell membrane, HPG and nuclear staining described above. μ-gels were stained for aggrecan (1°: Abcam ab3778, 1:50 in PBS overnight at 4°C; 2°: ThermoFisher AlexaFluor 546 goat anti-mouse, 1:200 in 5% BSA for 1 hr at room temperature). Additional μ-gels were stained for collagen II (1°: DSHB ii-ii6b3, 10 μg/mL in PBS overnight at 4°C; 2°: ThermoFisher AlexaFluor 546 goat anti-mouse, 1:200 in 5% BSA for 1 hr at room temperature). Collagen II staining was performed with and without hyaluronidase digestion (300 μg/ml, 2.5 hours at room temperature).

Matrix Digestion & Perturbation

Following 9 days of continuous culture in labeling media, agarose μ-gels were digested with a panel of enzymes. Highly purified collagenase (Worthington CLSPA) was suspended at 300 U/mL in PBS. Chondrotinase ABC (Sigma C2905) was suspended at 0.4U/mL in 50 mM Tris, 60 mM sodium acetate, 0.02% BSA, pH 8.0. Hyaluronidase

(Type IV-S from bovine testes, Sigma H3884) was suspended at 300 μg/mL in PBS. Digestion was performed at 37°C for either 1 or 18 hours. To assess the impact of a decrease in collagen crosslinking, β-aminopropionitrile (BAPN, Sigma A3134) was added to labeling media at 300 μM [151]. BAPN was applied for varying time periods as indicated in the results and figure legends.

Microcompression

Agarose gels seeded with $2x10^6$ chondrocytes/mL were cast between two glass plates and cut to yield cylindrical samples measuring 4 mm in diameter and 2 mm thick. Constructs were cultured in either control media (native methionine) or labeling media (HPG) for up to 9 days. Unfixed samples were halved through the mid-sagittal plane and stained with calcein AM and Hoechst to mark cell bodies and nuclei. Using a confocal-mounted device [60], samples were compressed to 30% strain in increments of 5%. Before each compressive step, 3D stacks depicting a region of interest were imaged. Following each compressive step, the sample was allowed to equilibrate for 5 minutes. Individual cells depicted in stack maximum Z projections were computationally segmented and manually tracked through multiple strain levels. For identified cells, the cell aspect ratio was calculated as the ratio of the long axis to the short axis. Data are presented as the ratio of strained to unstrained cell aspect ratio, calculated on a single cell basis.

Chapter 6: Stochastic simulations link gene expression heterogeneity and nascent protein production

Introduction

Ensemble measurements of genome-wide mRNA and protein abundance suggest that mRNA and protein expression generally correlate, with gene-to-gene differences in mRNA levels explaining approximately 40-80% of the variation in protein content.[38,223] However, amongst individual cells and for a single gene, mRNA abundance only weakly corresponds with protein quantity. For example, in eukaryotes, a screen of >120 genes compared single cell mRNA abundance with the signal of corresponding fluorescent reporter proteins; this analysis showed that, at the single cell level, mRNA and protein were uncorrelated.[212] Discordance between mRNA and protein has also been demonstrated in mammalian cells[1], but single cell mRNA-protein relationships may also be cell-type and gene dependent.[162] Within the specific context of chondrogenesis, we have previously noted that, at the single cell level, the transcriptional abundance of extracellular matrix genes does not correlate with the amount of matrix protein accumulated by individual cells undergoing differentiation.[36]

A possible explanation for the apparent disconnect between mRNA and protein abundance is that these molecules often differ in their temporal dynamics. For example, extracellular matrix proteins tend to accumulate with time, but gene expression is a stochastic process that occurs in rapid bursts separated by periods of transcriptional silence. Generally, the half-lives of corresponding mRNAs and proteins are

uncorrelated[190], and this mismatch is particularly exacerbated for matrix components. For instance, the half-life of many matrix proteins can range from months to years, while typical mRNA half-lives range from 1 hour to 2 days (median = 9 hours).[190,201] While it may be possible for mRNA and protein molecules with similar longevities to correlate, or for quantities of a stable mRNA and unstable protein to correspond, it is difficult to imagine how an unstable mRNA and highly stable protein could do so. Indeed, previous studies exploring the influence of protein half-life on the mRNA-to-protein relationship suggest that when a protein does not rapidly turn over, the correlation between mRNA and protein quantity weakens.[177]

To explicate this discordance between message and output, this study employs a computational model to examine the interplay between temporal dynamics to message and output in the mRNA-protein relationship. We experimentally quantify the transcriptional dynamics of five chondrogenic matrix components, and use this data to parameterize stochastic simulations of gene and protein expression in single cells. Using these simulations, we temporally sample the molecular abundance of mRNA and protein in individual 'cells' and explore how sampling window influences the observed correlation between mRNA and protein expression of chondrogenic matrix genes. Specifically, we consider the relationship between instantaneous mRNA expression, nascent protein expression, and total protein deposition over the course of early tissue development in engineered constructs.

Experimental Results

Transcriptional dynamics differ between chondrogenic marker genes

To establish a biologically relevant parameter space, we experimentally quantified single cell transcription of five chondrogenic marker genes: ACAN, COL6a1, COL9a1, COL11a1, and HSPG2. The corresponding proteins are integral components of the chondrocyte extracellular matrix: aggrecan, collagen VI and perlecan are found in the pericellular matrix, while collagens IX and XI integrate into forming fibrous network. As expected, each gene was expressed in chondrocytes cultured in agarose for 7 days in the presence of TGFβ (**Figure 6-1A**). ACAN, COL6 and COL11 had high median abundance (340-560 mRNA), while COL9 and HSPG2 had low median abundance (70-110 mRNA). For each gene, there was substantial cell-to-cell variation in mRNA abundance, with copy number ranging over multiple orders of magnitude (10^1 – 10^3).

To more fully understand how mRNA levels may change within a single cell over time, we measured two additional parameters: the mRNA half-life and the fraction of cells with active transcription sites. To measure RNA half-life, we compared mRNA abundance before and after transcriptional inhibition. mRNA half-life varied widely amongst these genes, ranging from approximately 3 hours (COL6, COL11) to >24 hours (COL9). ACAN and HSPG2 showed an intermediate mRNA stability, with half-lives on the order of 6-9 hours (**Figure 6-1B**). We note that mRNA half-life is cell-type and context-dependent,[49,182] and that while our measurement of ACAN half-life was similar to other reports, the half-lives of the collagens and HSPG2 differed from values measured in monolayer cultures.[196,213]

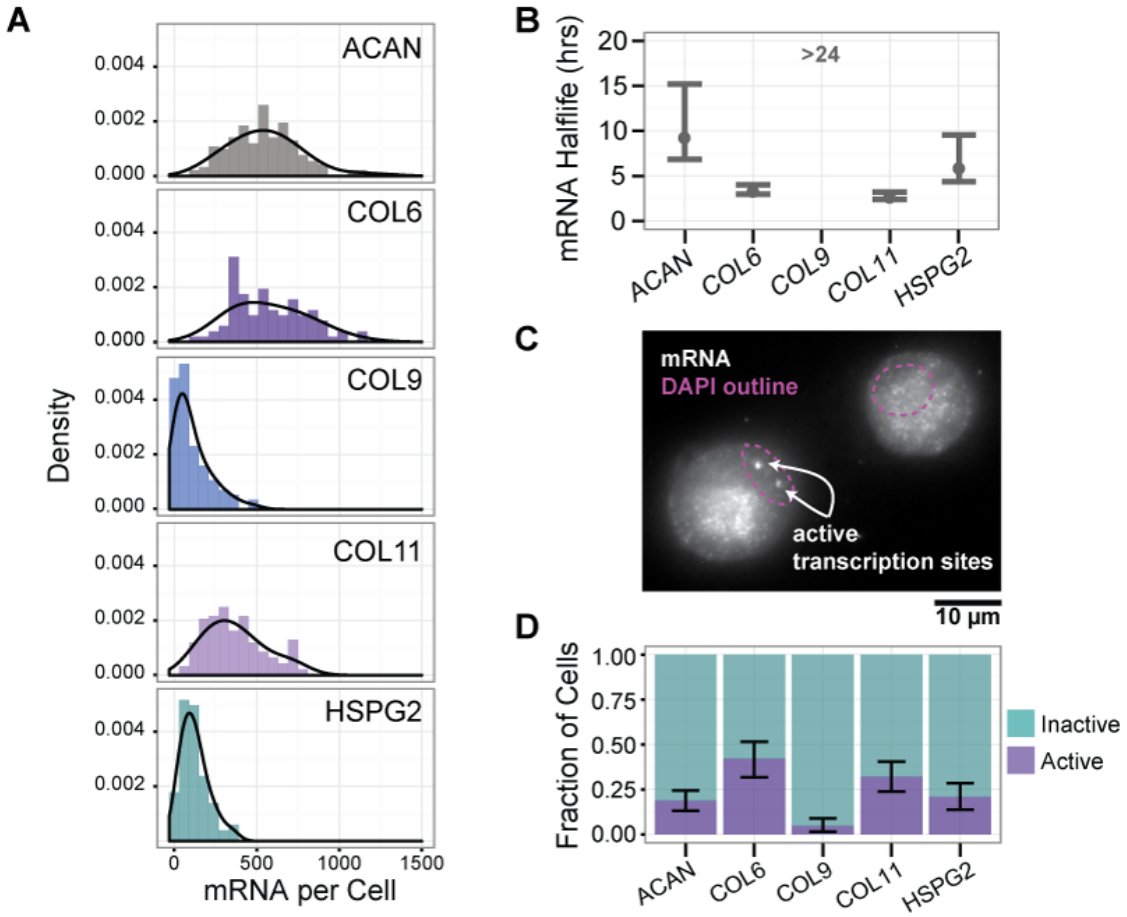

Figure 6-1: Transcriptional differences in expression of chondrogenic marker genes. A) Distribution of mRNA copy number amongst individual chondrocytes cultured in agarose for 7 days (n = 85-154 cells/group). B) mRNA halflife of chondrogenic marker genes (n = 84-234 / group, mean ± bootstrapped 95% confidence interval). C) Maximum projection of a COL6 RNA FISH image (white: COL6, magenta: DAPI). The white foci in the nucleus of the lower left cell indicate that it was actively transcribing COL6. D) Fraction of cells that were actively transcribing chondrogenic marker genes (n = 104 – 160 cells / group, measured fraction ± bootstrapped 95% confidence interval).

To establish transcriptional activity, we considered a cell as being in an actively transcribing state if bright FISH foci were visible within the nucleus. Such spots represent the co-localization of many nascent mRNA at a transcription locus (**Figure**

6-1C). Sporadic transcription site activity suggests that each of the five genes was expressed in bursts. The relative fraction of cells bursting was highly gene dependent (**Figure 6-1**D, p = 1.024e-10, n = 104 – 160 cells/group). Notably, while select gene pairs exhibited similar mRNA distributions, mRNA half-lives, or bursting fractions, the combination of these three metrics yielded a distinct transcriptional signature for each gene considered.

A stochastic model of chondrocyte gene expression and matrix production

Establishing and parameterizing a stochastic model

To model the stochasticity inherent to gene and protein expression in chondocytes, we considered a random telegraph model of transcription and translation.[70,177,214] The model describes transcription from two copies of a single gene (**Figure 6-2**). Each gene copy independently switches between an inactive and active state (respectively denoted I and A). Transcription from the active gene yields individual mRNA molecules (M). mRNAs are translated to create proteins (P), and are also allowed to decay over time.

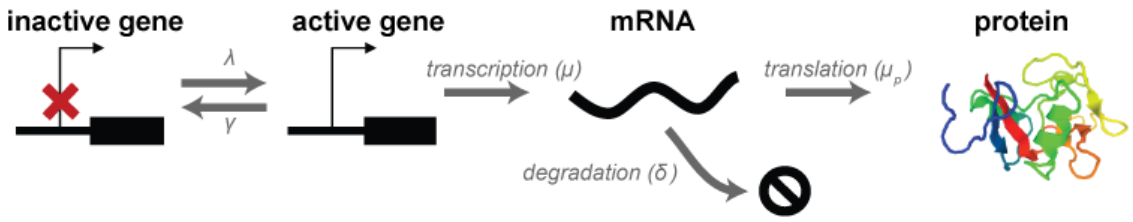

Figure 6-2: Schematic of the reactions included in the stochastic simulations.

These chemical processes are described by the equations:

$$I \xrightarrow{\lambda} A$$

$$A \xrightarrow{\gamma} I$$

$$A \xrightarrow{\mu} A + M \qquad \text{(Eq. 6-1)}$$

$$M \xrightarrow{\mu_p} M + P$$

$$M \xrightarrow{\delta} \emptyset$$

where λ is the rate of gene activation, γ is the rate of gene inactivation, μ is the transcription rate from an active gene, μ_p is the rate of protein translaation, and δ is rate of mRNA decay. This model incorporates the assumption that all translated protein is retained by the cell indefinitely; this assumption is consistent with experimental data showing that broad-spectrum inhibition of matrix degradation during the first nine days of culture had minimal impact on matrix quantity and organization (**Figure 6-3**), as well as measurements of matrix turnover in native cartilage tissue. That is not to say that formed matrix is not lost to the medium (some proteins do not assemble into the network but instead diffuse out of the construct), but rather indicates that this process is not governed by protein degradation per se, and is instead determined by the ability of the formed protein to couple into the network.

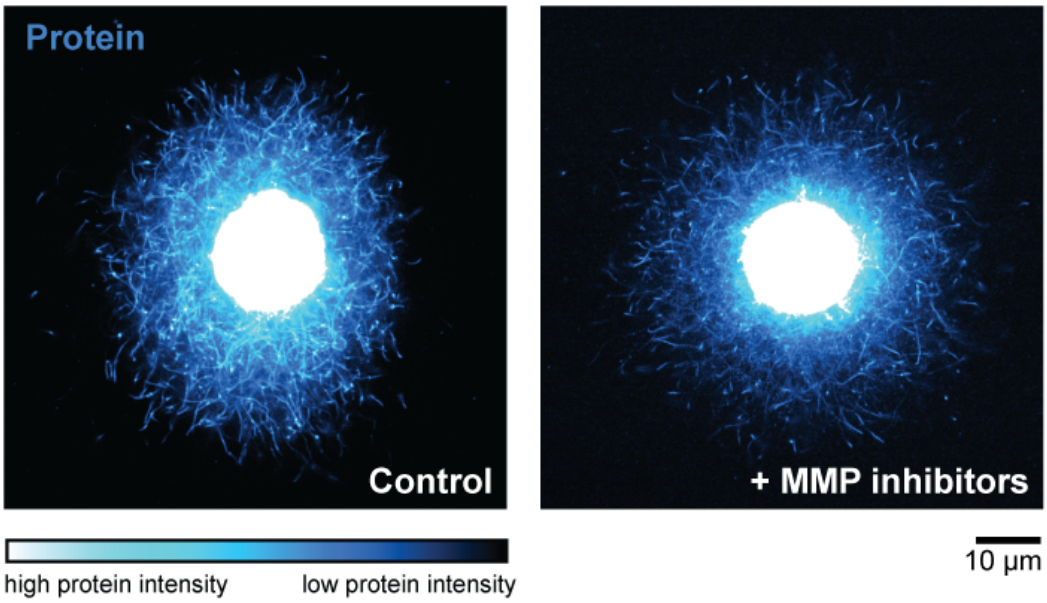

Figure 6-3: Comparison of chondrocyte extracellular matrix production in the presence and absence of MMP inhibitors. Constructs were cultured for 9 days in media containing the methionine analog HPG, which was fluorescently tagged to label proteins following sample fixation.

Having established the model, we next estimated the transcriptional parameters λ, γ, and μ by fitting experimentally measured mRNA distributions with the steady-state solution to the master equation using the maximum likelihood method.[177] The translation rate μ_p was coarsely estimated from the number of aggrecan molecules capable of packing into a the pericellular matrix of a chondrocyte; our estimate fell within the range of translation rates observed in mammalian cells.[190] Details regarding parameter estimation are discussed in the methods section, and the identified parameters are summarized in **Table 6-1**. Notably, after accounting for the fraction of cells that were actively transcribing, transcription rates (μ) were consistent with measured transcription rates in populations of mammalian cells (10^0 – 10^2 mRNA/cell/hour).[190] Furthermore,

even though the fraction of cells with active transcription sites was not used as a model input, the steady state bursting ratios λ/(λ + γ) output from the model were consistent with the experimental measurements of transcriptional activity (**Table 6-1**, **Figure 6-1D**). Collectively, these model outputs suggested that the fitting procedure for data from chondrocytes in 3D culture faithfully recapitulated biological behavior in mammalian cells.

Table 6-1: Parameters used to model expression of chondrogenic markers.

	λ (s^{-1})	γ (s^{-1})	δ (s^{-1})	μ (s^{-1})	$\frac{\lambda}{\lambda+\gamma}$
ACAN	1.10×10^{-4}	6.16×10^{-4}	2.04×10^{-5}	3.71×10^{-2}	.15
COL6	2.65×10^{-4}	6.14×10^{-4}	5.66×10^{-5}	5.41×10^{-2}	.30
COL9	7.87×10^{-6}	1.66×10^{-4}	8.02×10^{-6}	9.58×10^{-3}	.045
COL11	2.06×10^{-4}	5.02×10^{-4}	7.13×10^{-5}	4.51×10^{-2}	.29
HSPG2	6.09×10^{-5}	3.25×10^{-4}	3.21×10^{-5}	1.26×10^{-2}	.15

Intrinsic heterogeneity emerges in single cell simulations using realistic model parameters

Having obtained biologically-informed parameter sets, we next used a Gillespie Monte-Carlo method to simulate the dynamics of many individual cells.[70] At the simulation endpoint (t = 216 hours = 9 days), the stochastic model reproduced the mRNA distributions observed experimentally (**Figure 6-4B**), and also showed cell-to-cell variation in final protein content (**Figure 6-4C**). Over time, individual cells randomized away from the initial condition of zero mRNA, and the simulated population achieved a steady state mRNA distribution after approximately 2 days. Individual cell traces demonstrated large fluctuations with time (e.g. transcriptional burstiness, **Figure 6-4C**). Protein content increased near-linearly with time, with variation in the slopes of individual

cell traces (**Figure 6-4D**). Consistent with our previous experimental results,[36] instantaneous measures of mRNA abundance correlated poorly with simultaneous protein quantities. This lack of correlation held across the multiple genes considered in the simulations (**Figure 6-4E**).

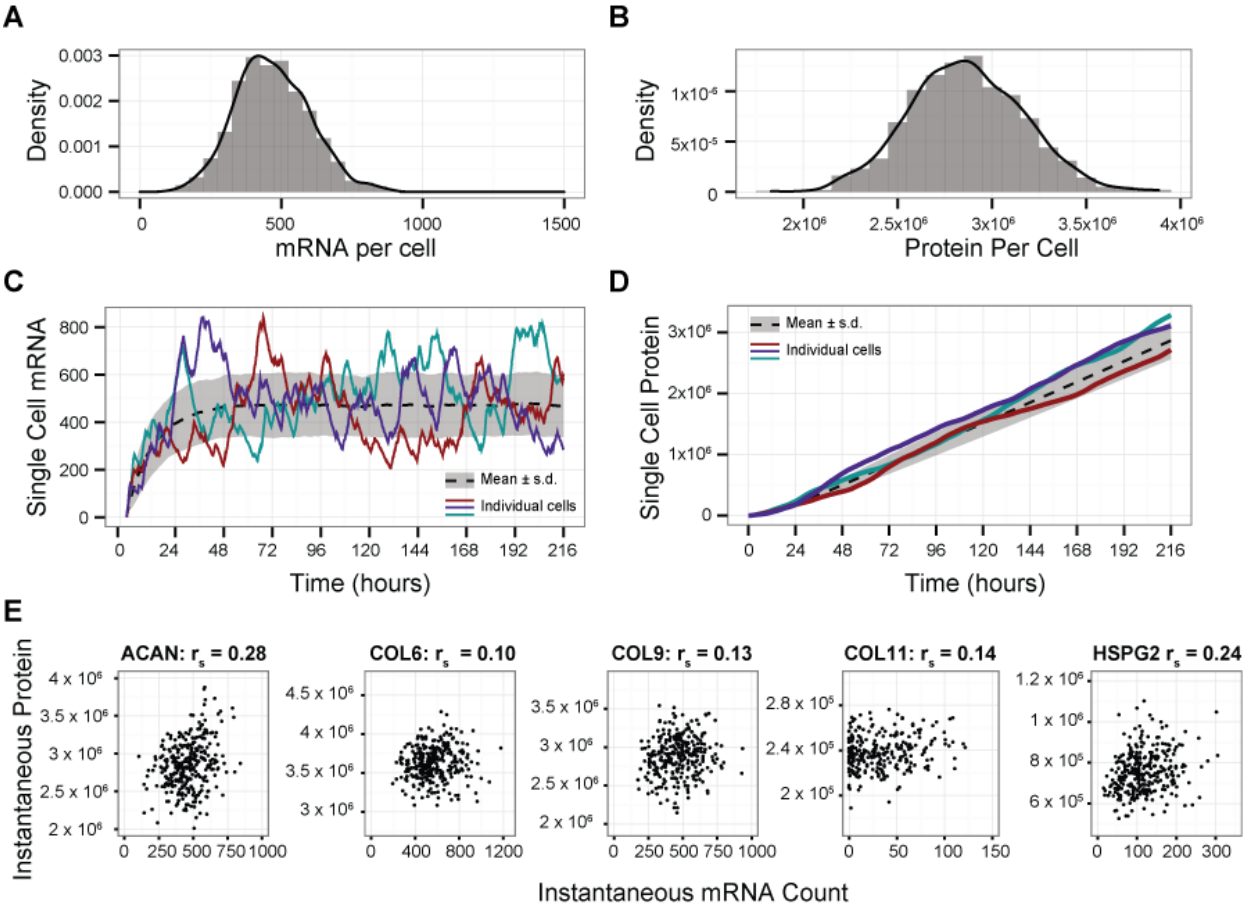

Figure 6-4: Cell-to-cell variation arises in simulations of single cell gene and protein expression. A) Distribution of ACAN mRNA at the final simulation time point, t = 216 hours (9 days). Grey bars indicate histogram of observations; line indicates estimate of the associated continuous distribution. B) Distribution of aggrecan protein at the final simulation time point, t = 216 hours. C) Simulated single cell ACAN mRNA over time. D) Simulated single cell ACAN protein over time. The same example cells are shown in C and D, dotted line indicates mean response and grey region indicates standard deviation. E) Scatterplots and associated Spearman correlations relating instantaneous mRNA and protein counts at t = 216 hours for ACAN, COL6, COL9, COL11 and HSPG2. For clarity, plots represent 300 points randomly sampled from 1000 simulations/condition.

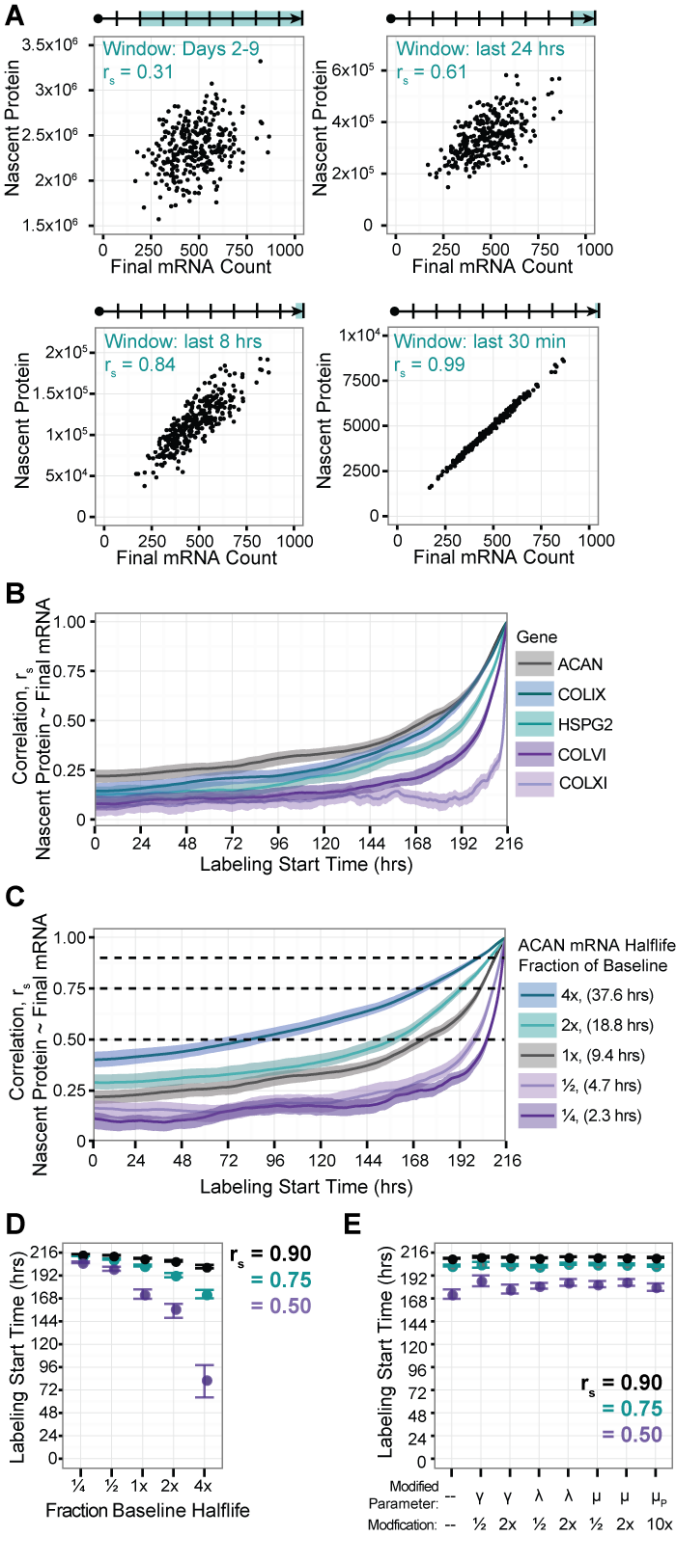
A
Window: Days 2-9
r_s = 0.31
Window: last 24 hrs
r_s = 0.61
Window: last 8 hrs
r_s = 0.84
Window: last 30 min
r_s = 0.99
Nascent Protein
Final mRNA Count
B
Correlation, r_s
Nascent Protein ~ Final mRNA
Labeling Start Time (hrs)
Gene
ACAN
COLIX
HSPG2
COLVI
COLXI
C
ACAN mRNA Halflife
Fraction of Baseline
4x, (37.6 hrs)
2x, (18.8 hrs)
1x, (9.4 hrs)
½, (4.7 hrs)
¼, (2.3 hrs)
D
Labeling Start Time (hrs)
Fraction Baseline Halflife
r_s = 0.90
= 0.75
= 0.50
E
Modified Parameter:
Modfication:

Figure 6-5 (previous page): In silico correlation between nascent protein and final mRNA count is time- and half-life dependent. A) Scatterplots of temporally labeled nascent protein and final instantaneous mRNA counts. (n= 1000 simulated cells, with 300 randomly selected cells plotted.) B) Simulated correlation between nascent protein and final mRNA for chondrogenic marker genes, as a function of the labeling start time. C) Simulated correlation between nascent protein and final mRNA for ACAN, as a function of time and the mRNA half life parameter. D-E) Labeling start times corresponding to r_s = 0.90, 0.75 and 0.50 for altered values of mRNA half-life (D) and additional model parameters (E). For B-E, n= 1000 simulations/ condition; values shown are mean ± bootstrapped 95% confidence intervals.

In silico correlation between temporal protein fraction and instantaneous gene expression

Next, we considered how the correlation of final mRNA expression with *nascent* protein expression might differ from the correlation between final mRNA expression and total protein content. Nascent protein fractions were identified as those produced in specific temporal windows, similar to how proteins might be tagged in a pulse labeling experiment. For long sampling windows, mRNA and protein were only weakly correlated. However, as the sampling window shorted, the correlation between instantaneous mRNA and nascent protein increased (**Figure 6-5A,B**). The relationship between window duration and correlation was gene-dependent; ACAN and COL9 demonstrated the most persistent relationships, while COLXI mRNA and protein quickly lost correlation (**Figure 6-5B**). mRNA half-life strongly influenced correlation persistence when all other parameters were held constant (**Figure 6-5C,D**). Increasing or decreasing mRNA half-life by four-fold altered the time at which mRNA and protein exhibited a correlation of 0.50 by >1 day (**Figure 6-5D**). Independent perturbation of other model parameters had minimal effect, suggesting that mRNA half-life was the dominant determinant of mRNA-protein correlation in this system (**Figure 6-5E**).

Correlations between nascent and total protein fractions: model and experiment

These simulations also allowed us to consider how apparent heterogeneity in protein expression may change with time. Using metabolic labeling of nascent protein, we previously noted that nascent protein fractions exhibited greater cell-to-cell variability than measurements of total protein per cell.[146] This observation was replicated by the simulated populations, which further demonstrated the strong dependence of single-cell protein variation on the temporal window considered (**Figure 6-6A**). *In vitro*, nascent protein expression moderately correlated with total protein expressed by the same cell. Notably, the nascent-total protein correlation observed *in* vitro was greater than that predicted by the simulations (**Figure 6-6B,C**). This difference is likely attributable to the role of extrinsic noise sources, which were not incorporated into the model described here.

Discussion

In this study, quantitative single cell analysis was combined with stochastic simulations to illustrate the conditions under which instantaneous mRNA levels correlate with nascent protein fractions more strongly than with total protein. In simulations, this mRNA-nascent protein correlation was markedly gene-dependent and highly sensitive to changes in the mRNA half-life. Intuitively, long or short mRNA half-lives may smooth or exacerbate the impact of stochastic gene (in)activation, and resultant transcriptional burstiness. This finding complements previous work demonstrating that for a given

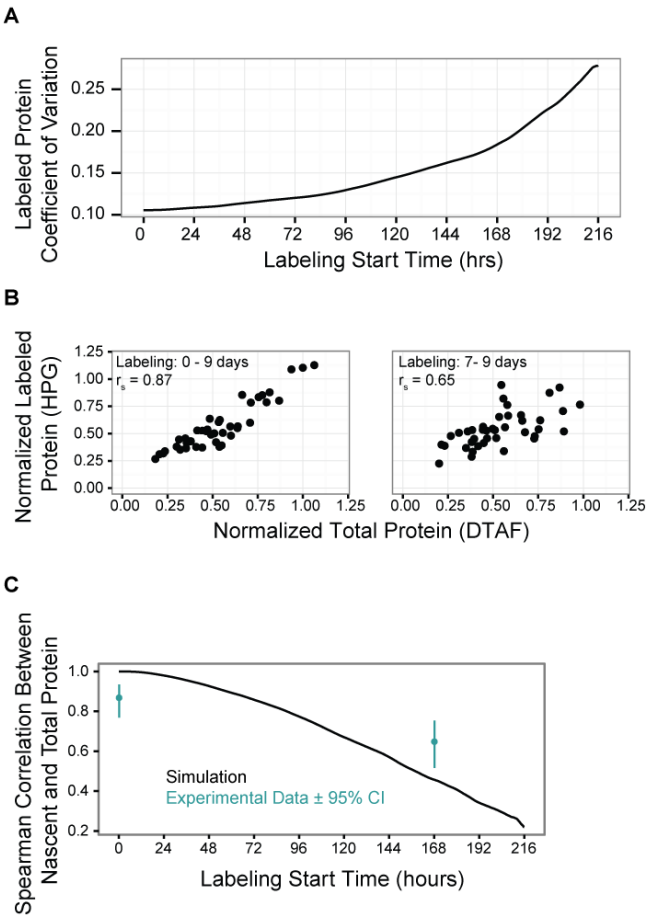

Figure 6-6: Variability of and correlation between nascent and total protein fractions. A) Coefficient of variation describing labeled protein levels, as a function of labeling start time. B) Experimental measurements of nascent protein metabolically labeled with HPG and total protein stained with DTAF. (n = 40-42 cells/condition.) C) Correlation between nascent and total protein measurements in single cells. Experimental data is shown in green and simulation data is shown in black.

mRNA half-life, protein half-life is a strong determinant of the mRNA-protein correlation.[177]

Across the transcriptome, mRNA half-lives range over multiple orders of magnitude (<1 hour to years). mRNA stability is influenced by multiple determinants, including the

mRNA's inherent nucleotide sequence and a host of external factors. Collectively, these regulatory mechanisms allow the stability of a given mRNA to vary across different environments and conditions. For example, one genetically encoded determinant of stability is the 'optimality' of a mRNA's codons (essentially how efficiently the mRNA can be translated).[169,182] As tRNA abundance changes, so too may codon optimality, and thus mRNA stability.[169] Dynamic external regulators of mRNA stability include cytokines and mRNA binding proteins.[182] Two particularly relevant to chondrogenesis are TGFβ and IL-1. TGFβ treatment upregulates the expression of RNA binding proteins that enhance mRNA stability, while IL-1 stabilizes mRNA associated with some inflammatory genes.[182] Indeed, for many genes, including those encoding key components of the extracellular matrix, mRNA stability is reduced in osteoarthritis.[213] Thus, while we identified aggrecan and collagen IX as having the most temporally persistent mRNA – protein relationships, this finding may be highly context dependent. In other situations, the perceived mRNA-to-protein correlation may differ.

While we found weak correlations between mRNA and protein levels for all but the shortest of labeling windows, we acknowledge the importance of the dynamic range of molecular count in these calculations. For example, if a similar assay were performed on a mixture of two cell populations which differed substantially in their average expression (e.g. one highly expressing, and one non-expressing), the correlation measurement would be dominated by the difference between the two subpopulations instead of differences between individual cells. In contrast, isolated consideration of either subpopulation would likely suggest a weak correlation (similar to the conditions

described here). This notion is consistent with the well established experimental strategy of overexpressing a gene to bolster protein expression.

One limitation of the model presented here is that it considered only intrinsic variation, and assumed that each gene-copy, and indeed each cell, were subject to identical regulation. Even for isogenic populations of relatively homogenous cells (e.g. chondrocytes), extrinsic noise arising from factors including an individual cell's specific lineage, microenvironment, and cell cycle stage also regulates mRNA abundance. Extrinsic noise is likely of even greater importance for cell types with significant cell-to-cell variation in phenotype (e.g. MSCs). For example, the epigenetic and metabolic differences observed amongst MSC populations represent an additional layer of noise not accounted for by this model. Existing frameworks to identify and model the relative contributions of intrinsic and extrinsic noise utilize either instantaneous measurements of dual fluorescent reporter expression, or time-series measurements of a single reporter.[53,63] Such experiments in naïve, differentiating, and committed cell types might help to elucidate the relative importance of extrinsic noise in the initiation and maintenance of chondrogenesis. Multilayered models might even shed light on how intrinsic noise helps to establish sources of extrinsic noise that emerge within an initially identical population, potentially leading to the predisposition of individual cells towards specific phenotypic states.

In summary, this model provides support to the notion of a temporal influence on the correlation between gene and protein expression for matrix molecules. Experimental evidence of such a relationship might be demonstrated by studies combining single cell

quantitation of mRNA abundance (e.g. RNA FISH) with measurement of nascent protein fractions (e.g. FUNCAT). Progress towards these experiments, and other future directions, will be discussed in the next chapter.

Methods

Cell isolation & culture

Passage 1 chondrocytes isolated from the bovine calve knee joints (Research 87) were encapsulated in thin 2% agarose microgels (μgels) and cultured in chemically defined media supplemented with TGFβ for up to 7 days as previously described in Refs. [36,146].

MMP-mediated matrix degradation was inhibited through the addition of an MMP inhibitor (Sigma GM6001, 10 μM) for the duration of culture.

Nascent protein was labeled during synthesis through the incorporation of a non-canonical methionine analog (azidohomoalanine or homopropargylglycine), and was tagged following fixation via a click reaction.[146] Total protein was detected by staining samples with 200 μM DTAF for 1 hour at room temperature. Proteins were imaged using a Nikon confocal microscope at 100x magnification.

RNA Fluorescence In-Situ Hybridization, Imaging, and Quantification

To measure single cell gene expression, single molecule RNA FISH was performed as previously described.[36,179] Briefly, day 7 μgels were fixed in 4% PFA and stored in 70% ethanol prior to *in situ* hybridization. Samples were stained with oligonucleotide probes targeting aggrecan, collagen IX, collagen XI, and perlecan (Stellaris oligonucleotides,

Biosearch Technologies) and DAPI. Probes were conjugated with either Alexa 594 or Atto 647N, and their sequences are listed in Appendix 1. For imaging, gels were compressed and mounted in 2x saline sodium citrate buffer. Image stacks encompassing the full cell height were collected using a Leica DMI600B equipped with a 100x Plan Apo objective. To quantify single cell gene expression, cells were manually segmented and individual mRNA spots were counted with a custom MATLAB script. The transcriptional state (active vs. inactive) of individual cells was assessed by manually examining images for bright probe foci located within the nucleus, indicative of multiple mRNA that have not yet diffused away from an active transcriptional site.[177] Relative fractions of actively transcribing cells were compared via chi-squared test for equality of proportions.

RNA decay was measured by inhibiting transcription in living agarose-encapsulated chondrocytes. Samples (Day 7 μgels) were fixed after either 0 or 4 hours of treatment with 1 μg/ml actinomycin D (Sigma A1410). RNA FISH was used to quantify the expression of ACAN, COL6a1, COL9a1, COL11a1, and HSPG2 at both timepoints; n > 84 cells were measured per condition. Decay was assumed to be exponential, and RNA half life ($\hat{t}_{1/2}$) was estimated as:

$$\hat{t}_{1/2} = \frac{t}{\log_2(\widehat{N}_0 - \widehat{N}_4)}$$

where $\widehat{N}_0$ and $\widehat{N}_4$ are the mean mRNA copy number measured after 0 and 4 hours respectively.

For both mRNA half-life and transcriptional state measurements, 95% confidence intervals were constructed through bootstrapping (500 bootstrap samples/condition).

Fitting transcriptional parameters to experimental data

The transcriptional parameters λ, γ, and μ by were obtained by fitting experimentally measured mRNA distributions with the steady-state solution to the master equation using the maximum likelihood method as described in the supplement of Ref. [177]. The translation rate μ_p was estimated by coarsely calculating the number of aggrecan molecules capable of packing into a chondrocyte's pericellular matrix. To do so, aggrecan molecules were considered as cylinders 400 nm long and 80 nm in diameter,[154] and were estimated to occupy the pericellular matrix volume with 50% packing efficiency. The pericellular volume was considered as a spherical shell with an inner radius of 15 μm and an outer radius of 22.5 μm (dimensions corresponding to chondrocytes cultured in agarose for 9 days).[146] Under these conditions, approximately 3 million aggrecan molecules could pack into the periceullar volume. The accumulation of these molecules over 9 days approximately corresponded to μ_p = 0.5 min^{-1}. Changing this estimate by an order of magnitude had minimal impact on the correlation-labeling window relationship, and thus this estimate was deemed sufficient (**Figure 6-5E**).

Computational simulations

To model stochastic gene expression, we used the Gillespie algorithm to govern the evolution of mRNA and protein counts over time.[70] At each simulation step, the algorithm randomly selected which of the five possible reactions (Eq. 6-1) would occur next, taking

into account weights based on the reactions' relative propensities (the reactant concentration multiplied by the reaction rate constant). Next, the algorithm identified the time until the next reaction would occur by selecting randomly from an exponential probability distribution function. Finally, the algorithm executed the chosen reaction by changing the number of molecules in the system and advancing the simulation time. For each condition described, gene and protein expression of 1,000 cells was simulated over 9 days of simulation time using custom C code.[177] Simulation data was recorded at 10 minute intervals for subsequent analysis.

Chapter 7: Summary and Future Directions

Summary

Mesenchymal stem cells display substantial cell-to-cell variation. Their heterogeneity manifests across many aspects of cell phenotype, including differentiation capacity, molecular signature, and cellular biophysics. Such pervasive variability complicates the use of MSCs in regenerative applications and potentially limits their therapeutic efficacy.[86] As a result, strategies to leverage or otherwise mitigate this variability could ultimately improve the clinical performance of MSC-based therapies. However, most conventional assays measure MSC properties in bulk and, as a consequence, mask this cell-to-cell variation. To better understand the nature of MSC heterogeneity and its underlying mechanisms, we have quantitatively assessed MSC phenotype at the clonal and single cell level. We conducted this investigation within the context of chondrogenesis, and have contrasted MSC heterogeneity against the single cell behavior of chondrocytes, the relatively homogenous cells found in native cartilage.

In Chapter 3, we observed that MSC heterogeneity extends to include aspects of cellular mechanotransduction. Clonal MSC populations differed morphologically and also varied in their contractility, ability to transmit extracellular strain the nucleus, and capacity to form mechanically competent cartilage-like matrix. Clones also differed in their transcriptional signature, but variability within a clone was similar to the variation observed between clones (Chapter 4).

In Chapter 4, we used single molecule RNA FISH to query the mRNA expression levels of conventional differentiation markers (aggrecan, cartilage oligomeric matrix protein, Sox9, osteopontin, and lipoprotein lipase) at the single cell level, and found that both primary chondrocytes and chondrogenically-induced MSCs showed substantial mRNA expression heterogeneity. Even small MSC colonies and sister cell pairs showed high cell-to-cell variation in transcript abundance, suggesting that marker mRNA expression was not heritable through cell division. Surprisingly, this variation in marker transcript levels only weakly associated with cartilage-like matrix production at the single cell level. Furthermore, transcriptome-wide analysis of single cell derived clones suggested that other markers, either alone or in linear combination, did not correlate with functional potential. This RNA-protein disconnect was also apparent in fully differentiated cells: as primary chondrocytes dedifferentiated with monolayer expansion, mRNA expression of the cartilage marker aggrecan did not correlate with the extent of cartilage-like matrix accumulation. Together, these quantitative analyses suggested that efforts to sort chondrogenically “superior” MSC subpopulations based on these markers would only marginally enrich the progenitor population. Our results also suggested that, although canonical markers have very clear functional roles in differentiation and matrix formation, their instantaneous mRNA abundance is only tenuously linked to the chondrogenic phenotype and matrix accumulation at the single cell level.

One possible explanation for the apparent disconnect between gene and protein expression is that mRNA and protein exhibit different temporal dynamics. Chiefly, in comparison to mRNA, extracellular matrix proteins are stable over much longer time

frames and tend to accumulate in developing tissue engineered constructs. As a result, the assessment of mRNA and protein abundance over time might provide insight into what governs their correlation. However, quantifying the spatiotemporal organization of matrix proteins is challenging using standard techniques. In Chapter 5 we addressed these challenges by using noncanonical amino acid tagging to fluorescently label extracellular matrix synthesized in the presence of bio-orthogonal methionine analogs. This strategy labeled matrix proteins with high resolution, without compromising their distribution or mechanical function. We demonstrated that the organization and temporal dynamics of the proteinaceous matrix depended on the biophysical features of the microenvironment, including the biomaterial scaffold and the niche constructed by cells themselves. Pulse labeling experiments revealed that, in immature constructs, nascent matrix was highly fibrous and interdigitated with pre-existing matrix, while in more developed constructs, nascent matrix lacked fibrous organization and was retained in the immediate pericellular matrix. Inhibition of collagen crosslinking increased matrix synthesis, but compromised matrix organization. Finally, these data demonstrated marked cell-to-cell heterogeneity in nascent matrix production amongst both chondrocytes and mesenchymal stem cells undergoing chondrogenesis. Collectively, these results introduced fluorescent noncanonical amino acid tagging as a strategy to investigate spatiotemporal matrix organization, and demonstrated its ability to identify differences in phenotype, microenvironment, and matrix assembly at the single cell level.

We then united these notions of transcriptional stochasticity and temporal matrix dynamics in Chapter 6, where we assessed the impact of mRNA kinetics (production

and degradation) on the correlation between mRNA and protein, and defined its dependence on the observational window utilized. To do so, we computationally simulated stochastic gene and protein expression at the single cell level for five matrix genes: aggrecan, collagen IV, collagen IX, collagen XI, and perlecan. For each gene, simulation parameters reflecting transcriptional dynamics were identified through experimental measurement of single cell gene expression. These experimentally derived parameters yielded simulations that demonstrated both transcriptional bursting and a lack of correlation between instantaneous mRNA and total protein levels – results that were consistent with our experimental observations in Chapter 4. Intriguingly, shortening the window of protein considered (analogous to how nascent protein labeling would tag a temporal fraction of the proteinaceous matrix) increased the correlation between mRNA and protein abundance. This model further illustrated how mRNA stability was a crucial determinant of the timescale over which any such correlation persists.

Future directions

Defining the role of observation window on mRNA-protein correlations

A logical extension of the work presented here would be to experimentally test the hypotheses put forth by the model described in Chapter 6 first, that timing and observational window strongly influence the correlation between mRNA and protein at the single gene/single cell level, and second, that mRNA abundance will correlate strongly with nascent matrix deposition, but only weakly with total matrix deposition. One strategy to experimentally test this hypothesis is to combine FUNCAT and RNA FISH,

thus enabling the simultaneous measurement of nascent protein and instantaneous mRNA abundance. Towards this end, we have obtained preliminary data suggesting that the RNA FISH and FUNCAT techniques are compatible. The combination of FUNCAT and RNA FISH techniques minimally impacted the signal of each. Our early data show that mRNA abundance was unaltered by exposure to the FUNCAT reaction mixture in samples cultured with native methionine (**Figure 7-1A**). FUNCAT intensity was similarly unaffected (**Figure 7-1B**). Furthermore, while all of the FUNCAT imaging presented in Chapter 5 was performed on a confocal microscope, labeled nascent matrix can also be imaged on fluorescent microscopes optimized for FISH imaging (**Figure 7-1C-E**). Indeed, cell-by-cell comparisons of detected signal demonstrate strong correlation between the two microscope types, with the exception of situations with high out-of-plane fluorescence due to vertically stacked cells (**Figure 7-1E**). Thus, we are poised to experimentally query the relationship between instantaneous mRNA expression, and nascent and total protein expression. These studies are ongoing.

An alternative strategy to test these hypotheses would be to combine continuous measurements of RNA abundance with assays for either nascent or total protein levels. For example, temporal averages of mRNA over time in individual cells could be reported through the use of molecular beacons or genetically engineered reporter proteins. Following a period of live cell imaging, these time-averaged mRNA measurements could be compared with nascent protein labeled during the same window, or final total protein levels.

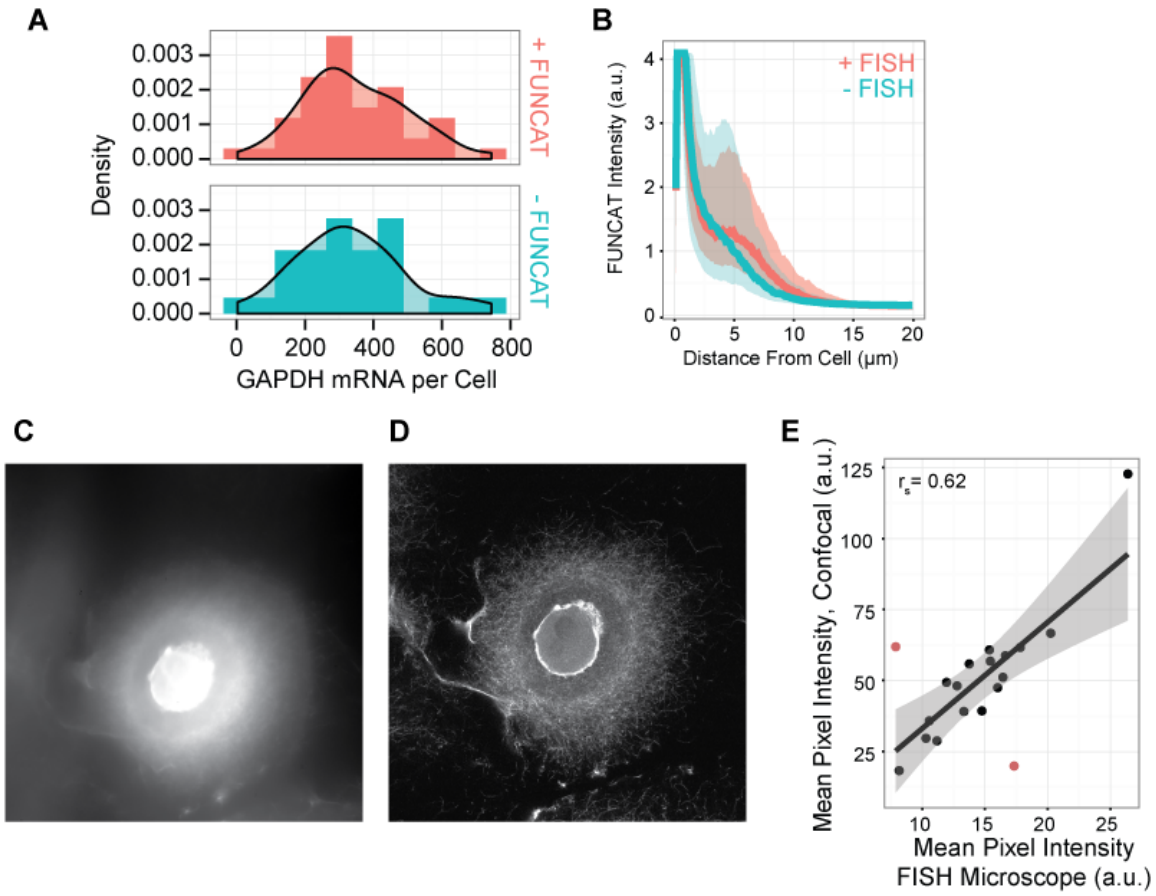

Figure 7-1: Preliminary controls related to the combination of FUNCAT and RNA FISH. A) Histograms and smoothed PDFs of single cell GAPDH mRNA abundance, with and without the FUNCAT reaction. B) FUNCAT intensity with and without RNA FISH hybridization. Lines represent median intensity ± 25th and 75th percentiles. C-D) Images of the same FUNCAT-labeled cell, taken on an inverted fluorescence microscope optimized for FISH (C) and a confocal microscope (D). E) Comparison of mean FUNCAT pixel intensity detected using an inverted fluorescence microscope and a confocal microscope. The same cells were imaged using each microscope; the pink points indicated locations where there was substantial out-of-plane fluorescent intensity. These points were included in the calculation of linear fit (black line ± grey confidence interval) and the correlation. n = 19 cells.

Leveraging heterogeneity to improve outcomes in cartilage tissue engineering

A second possible avenue of future study is the application of our understanding of heterogeneity to develop more robust tissue engineered cartilage replacements. The

path towards this goal is less well defined, but may include implementing cell sorting prior to construct fabrication, or using transcriptome-wide screens to identify novel predictive markers of chondrogenesis whose mRNA fluctuations are less stochastic in nature.

We initially hypothesized that highly chondrogenic MSCs could be prospectively identified based on their aggrecan mRNA levels. However, perhaps the most striking finding of this work is that, at the single cell level, instantaneous aggrecan mRNA abundance was a poor predictor of the amount of aggrecan protein content accumulated during chondrogenic culture. Computational simulations of how sorting might perform suggested that selecting cells based on instantaneous aggrecan expression would only slightly enrich the fraction of highly chondrogenic cells. At the same time, the number of cells would be greatly reduced due to the exclusion of cells with high chondrogenic capacity.

However, this is not to say that efforts to sort MSCs should be abandoned. While our findings are consistent with a growing body of evidence that finds discordance between mRNA and protein abundance at the single cell level, they contrast with the success of recent efforts using live-cell mRNA measurements to identify the most osteogenic MSCs. Both molecular beacons (targeting alkaline phosphatase) and SmartFlares (targeting Sox9 and Runx2) have been used to isolate highly osteogenic MSC subpopulations.[123,138] This raises the question: why does mRNA-based sorting appear to be effective for osteogenic progenitors? These mRNA do not appear to have extraordinary stability (alkaline phosphotase: $t_{1/2}$ = 21 hrs, Sox9: 1.7 hrs, Runx2: 6.1 hrs,

in embryonic stem cells), so their performance is not attributable to mRNA dynamics.[196] One possibility is that the markers used to query osteogenic potential are simply better suited to this type of application than the chondrogenic markers we considered. For example, the osteogenic markers may exhibit greater differential expression between high- and low-potential cells, and perhaps these differences would be detectable despite temporal fluctuations in mRNA abundance.

If this were the case, one way to improve cell selection for chondrogenesis would be to identify more predictive markers. In Chapter 4, we noted that the poor performance of our *in silico* sort did not seem to be due to our choice of marker: other chondrogenic markers, markers of alternative fates, and combinations thereof were similarly unable to predict single cell matrix accumulation, either alone or in combination. Additionally, we performed paired RNA sequencing and proteoglycan assays, and found that at the ensemble level, aggrecan was one of the best predictors of extracellular matrix accumulation. However, one limitation of this RNA sequencing study is that it compared gene expression in naïve MSCs with their ultimate matrix deposition. The osteogenic selection studies have identified the need for an osteogenic priming period prior to cell sorting. This priming period appears necessary to elevate gene expression in the responsive cells. It is possible that a similar transcriptomic screen following priming exposure to chondrogenic stimuli could identify non-canonical markers of chondrogenesis.

The disparity between the osteogeneic sorting experiments and our projections of chondrogenic sorting may also be due to measurement technique. Here, we performed

RNA FISH in intact, fixed cells and obtained measurements of the exact mRNA copy number. The live cell techniques instead yield more approximate measurements and may smooth temporal fluctuations in mRNA abundance. For example, perhaps oligo-mRNA binding and any re-conjugation between SmartFlare's fluorophore and quencher are slightly delayed following changes in mRNA abundance. Indeed, molecular beacons targeting the adipogenesis marker *PPARG* take approximately 25 minutes to fully detect control cDNA in a microplate assay.[115] Depending on the specific kinetics of an mRNA and its corresponding probe, a temporal lag could allow fluorescent intensity to reflect some quasi-average of mRNA abundance over time. If so, it is possible that a temporally-averaged measurement would be more predictive of cell propensity than truly instantaneous measures. Thus, despite our RNA-FISH findings, it may be worthwhile to assess chondrogenic potential in experimentally sorted populations.

Even given an mRNA target and measurement technique capable of identifying high potential cells, there are additional challenges associated with the practical implementation of a cell sorting strategy. If cells must be primed to assess their chondrogenic potential, it would be necessary to culture MSC monolayers in pro-chondrogenic media. However, traditional chondrogenic media induces high cellular contractility and often causes MSCs to detach from the culture surface after ~ 3-5 days – a duration shorter than the priming windows used for osteogenic sorting.[123,138] Thus, to avoid cell detachment, it would be necessary to determine if shorter priming durations would be sufficient to elevate target mRNA expression, or if priming could occur without the upregulation of contractility. Additionally, the use of molecular beacons requires

cellular permeabilization, and while SmartFlares can be endocytosed, there are concerns about the mechanisms and kinetics of their uptake.[35]

If mRNA expression is not the ideal metric to predict chondrogeneis, it is possible that other assays may prove more effective. We observed correlation between nascent and total protein at the single cell level, suggesting that there is indeed temporally persistent cell-to-cell heterogeneity in matrix production. Intuitively then, something must determine a cell's chondrogenic potential. As discussed in Chapter 2, there are many other dimensions of MSC heterogeneity, and they may hold the key to successful chondrogenic enrichment. For example, biophysical sorting of MSCs could be smoothly integrated into isolation and expansion procedures[122], and it may be possible to identify a biophysical chondrogenic signature.

Modulating MSC heterogeneity

In Chapter 5, our data hinted at the notion that matrix heterogeneity may be influenced by the encapsulating biomaterial construct. At high concentrations of agarose (4% w/v), the accumulation of extracellular matrix proteins was restricted to a tight band around the cells. Conversely, in low concentration gels (2% w/v), the matrix diffused throughout the construct, and effectively homogenized the environment. At an intermediate concentration (3%), the matrix remained associated with individual cells, but extended sufficiently from the cells to exhibit substantial cell-to-cell variation. While it is not yet clear if gel density modulated cell-to-cell variation in matrix quantity (vs. modulating variation in spatial distribution), these data raise the question of if the choice of biomaterial platform governs MSC heterogeneity.

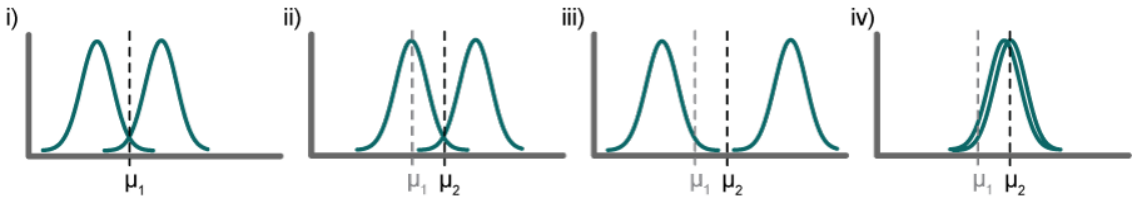

Figure 7-2: Schematic of possible distributions underlying population-level improvement in differentiation potential ($\mu_1 \rightarrow \mu_2$). i) Example distribution of baseline performance. The choice of bimodal behavior is for illustrative purposes only, and may not reflect the true distribution of chondrogenic differentiation capacity. ii – iv) Possible distributions exhibiting improved mean behavior. ii) Uniform improvement. iii) Improvement amongst high performing cells only. iv) Improvement amongst low performing cells only.

Numerous biomaterial systems have been shown to enhance MSC chondrogenesis. Is this accomplished by improving the performance of the entire MSC population (i.e. shifting the entire distribution, **Figure 7-2i,ii**)? Alternatively, do sub-populations differentially respond to these microenvironmental stimuli? For example, the most chondrogenic cells could become even more chondrogenic, while the least chondrogenic cells remain unaffected (**Figure 7-2iii**). Alternatively, an environmental stimulus could improve the performance of the least chondrogenic cells, while having minimal effect on the most chondrogenic subpopulations (**Figure 7-2iv**).

It would be of particular interest to examine how biologically active biomaterials regulate such cell-to-cell variability (e.g. hyaluronic acid ± N-cadherin moieties). Additional stimuli of interest include hypertrophic conditions, co-culture between MSCs and chondrocytes, and the use of active mechanical loading to induce chondrogenesis. It is tempting to speculate that if two pro-chondrogenic stimuli shift the distribution in different ways, they might combine synergistically. Additionally, improved understanding of the biological

pathways underlying MSC heterogeneity may reveal opportunities to modify poorly performing cells and enhance their chondrogenic capacity, or to identify them early in the process, and remove them from the population.

Conclusion

To conclude, tissue engineering with adult mesenchymal stem cells is a promising avenue of research in regenerative medicine. However, not all MSCs have the same functional potential, and individual MSCs differ in their responses to the same differentiation cues. Here, we have begun to understand MSC heterogeneity as it applies to cartilage tissue engineering, with the ultimate goal of developing strategies to identify MSC subpopulations best suited for cartilage tissue engineering. These studies create a foundation for future tissue engineering work wherein heterogeneity is minimized to "synchronize" MSCs or prospectively isolate only those MSCs most likely to robustly undergo chondrogenesis, potentially improving therapeutic outcomes.

Appendix 1: RNA FISH probe sequences

Oligonucleotide probe sequences used to quantify mRNA abundance.

Target	Aggrecan	Collagen VI alpha 1	Collagen IX alpha 1
	ACAN	COL6A1	COL9A1
# probes:	32 probes	32 probes	32 probes
Probe 1	gtccttgtctccatagcaac	cagcacaaagaacaggtcca	acgcttgacagtcactgatg
Probe 2	acacgtcataggtttcgttg	tggtgaaggacttgaccttg	aatcatcttgtccaatcctg
Probe 3	cttggaaggtgaacttctcc	tgtcgttcaggttgtcaatg	ctggatgctgctttatctat
Probe 4	tttctgggatgtccacaaag	acttcgtcgctatagtgcag	ctgtaaggcagttgatccaa
Probe 5	cccttacttcagggacaaac	ctttccgaagtacttgacgg	ggttggaatcctgaagtcta
Probe 6	actgatgtcctctactccag	ccacaaccaagtacttgttc	cttccagtcatccgaaaagt
Probe 7	gcagagatttctggaccttc	acggagaagactttgatgcc	ctgaggaatcctgaatctgc
Probe 8	actgatgtcctctactccag	atggtctggctgatgacttc	gtttggccattaatcttcac
Probe 9	aaggtcctctactccagaag	tgatcatgtcggtgatggtg	aaaggctgcagtttggagac
Probe 10	agaaggaagtccactaaggt	cacacttgctccacattatt	ggaatcgaacaaggagggca
Probe 11	agtctgctgagatcctctac	tccttggataccatcgaatc	cgccaatcatgatcttatgc
Probe 12	tctactccagaagcagagac	gcaatgtcattattgtcctc	cgaaaagagtcgcactgctc
Probe 13	actgatgtcctctactccag	tccaaaatctcgcattcgtc	taaggactcaatcctgttgc
Probe 14	caccagaaggaagaacactg	actcacagcaagagcacatt	tttgacctcttggctttata
Probe 15	actgatgtcctctactccag	gcacgaagaggatgtcgatg	tcagcatccattgaagttca
Probe 16	actgatgtcctctactccag	gcgatctcgaagttctgtag	accaggaattcctggagaac
Probe 17	cctccagaaggaagtctact	ggtcaatgaccttgacgatg	aatcctatatctccgagttc
Probe 18	agaaggaagtccactaaggt	ctcaaacttgaccagctcat	atgcaaacctgcttgatgtg
Probe 19	gaaggtattctgccaaggtc	atctggttgtggctgtactg	catctgagcaaagtgctctt
Probe 20	gtccactgaaatcaggaagg	caatgcggttgttcggagtg	aacctgatggtcagatcgac
Probe 21	caaagagtccagaggtttcc	caagccaaagacatccttga	ttgctgaatagcaccattgt
Probe 22	ccactaatgtcaggaaaccc	caggagatgacgttgagctg	caaaaggaggggtgttgtct
Probe 23	gtcacgccagatatttctcc	gttttccttgaccaatgaga	agaacacgcgctcactacaa
Probe 24	gatttctggttgtccagagg	gcagtgatgttcttcaggaa	aaatccttgcagagctggaa
Probe 25	caattgctccagaagacagt	ggacatttcttgtctatgca	gggctgcttcttaactaatc
Probe 26	aagttccagaatgcctgaag	caggaggatggtgatgtcgg	agcaccaattctcttgacag
Probe 27	ccaaagatcccaaatggtct	gcgtttggtgatgtcaaagt	agcagaaaggcaagctgacg
Probe 28	accactggattcaaaaagct	gtgtagttctgcaggaactg	gggtttggatacagtacagt
Probe 29	tgggactgatgacacttcta	gagttgccatccgagaagag	tggacgagcccaaaatacga
Probe 30	cagatcagcttcatggaagg	cgaagaccacgtcatactca	acttcttttcattgaaccca
Probe 31	ggattcgatttctagacgcg	taactgggcacgcggaacag	acacaggcctacagaaatgc
Probe 32	attgatctcgtatcggtcct	gacaccgtctggtagaagac	ggagcagaaagggcttttta

Target	Collagen XI alpha 1	Cartilage Oligomeric Matrix Protein	GAPDH
	COL11A1	COMP	GAPDH
# probes:	32 probes	32 probes	20 probes
Probe 1	aaatcgcgattgcagaccta	gtaatctccttgacctgctg	agaattaaaagcagccctgg
Probe 2	cgcactcacaactgtgaact	catcaccgtgtttttcagga	tcattgatggcgacgatgtc
Probe 3	gttgttatggtgaaatccca	gttggctgtctcggtacaag	agaccatgtagtgaaggtca
Probe 4	gttgtttttgatattccctc	ctggtattgatgcagcgaac	ttgactgtgccgttgaactt
Probe 5	gtgtggtcttcatacaggaa	cgtacaaacctgcttgttgg	ctttccattgatgacgagct
Probe 6	ttccattggtgtcaacgatt	cggtctcacactcgttaatg	gatctcgctcctggaagatg
Probe 7	cccctcaaaaacttcttcat	attgacgcatacagagttgg	tctccatggtagtgaagacc
Probe 8	tggactgtaatgctcacagt	gaaagccgtctaagtctgtg	agagatgatgaccctcttgg
Probe 9	tgatccagatacactcctag	cagttgtcctttctgcactg	ttgttatacttctcgtggtt
Probe 10	ttcttcaactggatttggct	gcagttgtccttctcattga	gaggcattgctgacaatctt
Probe 11	cataatcctctccagttaga	cttgtcgccatccgtatttc	ccaaagtggtcatggatgac
Probe 12	attagtgggacttgttggtt	cattctggtcgtcgttcttc	aagcagggatgatattctgg
Probe 13	atctgtttctgctggaacac	ctttctgatctgagttgggc	ttgagctcagggatgacctt
Probe 14	ttcaataagcatgccaggtt	ttgtcacaggcatctcctac	gtcagatccacaacagacac
Probe 15	actaacatggtaccaggagg	gtctgcgttgctcttctgag	catacttggcaggtttctcc
Probe 16	gaagtccatcaaatcctcga	gtcttggtcgctgtcacaa	gtcctcagtgtagcctagaa
Probe 17	gaatccaactggaccttgag	cagttgtccttcgagtcctg	tgagtgtcgctgttgaagtc
Probe 18	accaactggaccttgaagac	ccatcattgtcgtcatcgtc	aatgagcttgacaaagtggt
Probe 19	accaggatctccaggaaaac	ctagacggcagttgtcccta	ttgctgtagccgaattcatt
Probe 20	ggaaccaaatatttcctcca	accttgtccgcatcaaagtc	tactccttggaggccatgtg
Probe 21	acttcatatgctcgatgtct	ggacacacatcgatcttgtc	
Probe 22	gttgcagatctttgcaagtt	acccagttagggtctatctg	
Probe 23	tcctttggccatgatgaaat	atctccataccctggttgag	
Probe 24	gaaattttgccgagcagagg	gtcgctgttcatcgtctgta	
Probe 25	tgatcccaagaaccgaagtg	tggaacgtgccttcgaaatc	
Probe 26	tgtcctaatcttagccaaga	ggtagccaaagatgaaacca	
Probe 27	aaatgggttggtggtaccaa	catgaccacatagaagctgg	
Probe 28	acatccttgatggatgggaa	cagtacgtctgctccatctg	
Probe 29	ggaattcagagacacgtgca	taggatgtcttgtccttcca	
Probe 30	ccaaaagtcttccaggtttt	cctcgtagaatcgcactcta	
Probe 31	gaagagttgctcggtatgtg	catggtcgtatccaggatca	
Probe 32	ttctcatacacgttttcctg	atgatgttctcctgggagaa	

Target	Perlecan	Lipoprotein Lipase	Osteopontin
	HSPG2	LPL	OPN
# probes:	32 probes	32 probes	24 probes
Probe 1	gaaatttaccagtgctcgga	ggagtcctgagagcaaattt	cccaagaggcagaagcaaat
Probe 2	ctcacaatcgagctcatcag	caggaatgaggtggcaagtg	gtttaactggaagggcggag
Probe 3	agtctccagttgatgatagg	tgctgtggttgaagtgacag	tacagcatctgggtatttgt
Probe 4	cacaggacaagccggaatag	caactctcatacattcctgt	aagtctgcttctgagatggg
Probe 5	tctgtgtccagaaagcaagt	gacactggataatgctgctg	gacacagaattctgtggtgc
Probe 6	ccaatggcatggaggagaat	ataaacttggccacatcctg	gggtattttgtttgttgtca
Probe 7	atcttcacgagaccagtaga	agttaaattcatccgccatc	ggctttcattggacttactt
Probe 8	atgatggagattccgacagg	atcccaagagatgcacattg	gtcatctagatcgtctgttt
Probe 9	gtacattgcgaatggtcagg	ttcttattggtcagacttcc	cgtcggagtcattagagttg
Probe 10	ccatggcgaacatattggag	catactcgaagttaggtcca	aatggtcagggtcatcagtg
Probe 11	cacacagaggtaaaagccgg	caggagaaaggcgacttgga	tcagaatggtgagactcgtc
Probe 12	aaggacgaccacttgaatcc	ccctggtgaatgtgtgtaaa	gcgattgttggaatatcagt
Probe 13	gtggaggacagagacgatga	ttctggattccgatacttcg	ccgtagggataaacggagtg
Probe 14	gaagatgtacaagcgggagc	cgttagggtaaatgtccaca	tcttgacttcagtccgtaag
Probe 15	ttggtgactgtgactgtgat	caatgttacatcctggttgg	ttagatcggcggaacttctt
Probe 16	acagagtctacaggctcaat	ctctgcaatcacacggagag	tgtggcatctggactctgaa
Probe 17	tgtgaggaggatgactcgat	taggccttacttggattttc	ctgactcgtcttcttaggtg
Probe 18	tactgacatggcacacgtac	ggcagagacctttctcaaag	tggcgtgagttctttggaaa
Probe 19	ctggatggtgacgatgagtg	tgaccttgttgatctcgtag	gaatgcttgttcttatcctt
Probe 20	tgggaagacgacgactcgat	tacatcttgctgcttctttt	tttcctgactctcaatcaga
Probe 21	cgggtcaatggagatcacag	gcatctgagaacgagtcttc	aattcttggctgagtttgga
Probe 22	catggacgaggcacttgaag	cctggttggtgtatgtatta	atgatctaggtctagcttgt
Probe 23	tgaggttcacaacactctgg	cagtgccatacagagagatc	tttcaggtgtttgtcttctt
Probe 24	tctgacggtttagctgatga	aaaagggatgttctcgctct	attgacctcagaagaggcac
Probe 25	tctacaatcacttccactcg	ttgtggaaacttcaggcagg	
Probe 26	gatgagtacggatggcagag	aatatccacctccgtgtaaa	
Probe 27	catgcaagggaagacagctg	ggaccagctgaagtaggaat	
Probe 28	aaggagatgaagtcgggctg	cacctttttttgagtctctc	
Probe 29	tgaggtcgtggaagacgatc	gataagacattttctcccgg	
Probe 30	acacagacactggaactcgt	tatcacaggtgactttcctt	
Probe 31	ggaagccgttgtcatagaag	acttgtcatggcatttcaca	
Probe 32	ttaatggggtcttcggagac	cagccagactttctattcag	

Target	Sox9
	SOX9
# probes:	7 probes
Probe 1	aaggggtccaggagattcat
Probe 2	ctgctcgtcggtcatcttca
Probe 3	cttggggaacgtgttctcct
Probe 4	gacggcctcgcggatgcaca
Probe 5	cgtagcccttgagcacctgg
Probe 6	ggcatgggcaccagcgtcca
Probe 7	Ttgttcttgctcgagccgtt

Appendix 2: Supplementary statistical analysis related to Chapter 4

Supplementary Table 1	
Comparison	**Adjusted P-value**
-TGFB vs +TGFB, Day 01	0.3572
-TGFB vs +TGFB, Day 04	0.0827
-TGFB vs +TGFB, Day 07	0.0361
-TGFB vs +TGFB, Day 14	0.9479
-TGFB vs +TGFB, Day 21	0.9980
+TGFB, Day 1 vs Day 4	0.0022
+TGFB, Day 1 vs Day 7	0.4637
+TGFB, Day 1 vs Day 14	<.0001
+TGFB, Day 1 vs Day 21	0.0002
+TGFB, Day 4 vs Day 7	0.0019
+TGFB, Day 4 vs Day 14	0.0002
+TGFB, Day 4 vs Day 21	0.0009
+TGFB, Day 7 vs Day 14	<.0001
+TGFB, Day 7 vs Day 21	0.0002
+TGFB, Day 14 vs Day 21	0.0016

Supplementary Table 1: Additional statistical information related to GAPDH RNA abundance differentiating MSCs. Companion table to Supplementary **Figure 4-1**b. P-values indicated for all planned contrasts between –TGFβ and +TGFβ groups over 21 days of culture. GAPDH RNA count means compared by t-tests with Satterthwaite approximation and simulated adjustment for multiple comparisons.

Supplementary Table 2			
Timepoint/Gene Pair	r^2	Pearson's correlation coefficient	p-value
Day 1 aggrecan - osteopontin	0.24	0.49	1.0e-07
Day 21 aggrecan - osteopontin	.11	.34	.0019
Day 1 aggrecan-LPL	7.5e-5	.0087	.92
Day 21 aggrecan-LPL	.0022	.047	.67
Day 1 osteopontin-LPL	1.9e-5	.0044	.96
Day 21 osteopontin-LPL	.04	.21	.059

Supplementary Table 2: Correlation of marker gene RNA abundance in individual differentiating MSCs. r^2 and Pearson's correlation coefficient between abundance of aggrecan, osteopontin, and LPL in MSCs cultured with TGFβ in agarose for either 1 or 21 days. Test for the significance of the Pearson correlation coefficient was conducted with 103 degrees of freedom for Day 1 comparisons, and and 77 degrees of freedom for Day 21 comparisons. n= 105 cells at Day 1 and 79 cells at Day 21.

Supplementary Table 3				
	Colony A	Colony B	Colony C	Colony D
Aggrecan CV	1.71	1.34	1.07	0.91
GAPDH CV	0.32	0.35	0.31	0.20

Supplementary Table 3: Coefficient of variation (CV) of single cell RNA count in small MSC colonies. GAPDH RNA abundance was less variable in small colonies than aggrecan RNA abundance. (n = 75 cells in colony A, 7 cells in colony B, 6 cells in colony C, 8 cells in colony D).

Supplementary Table 4	
Comparison	**Adjusted P-value**
-TGFB vs +TGFB, Day 01	0.0067
-TGFB vs +TGFB, Day 04	0.0043
-TGFB vs +TGFB, Day 07	0.0020
-TGFB vs +TGFB, Day 14	0.0017
-TGFB vs +TGFB, Day 21	0.0031
+TGFB, Day 1 vs Day 4	0.0006
+TGFB, Day 1 vs Day 7	0.0002
+TGFB, Day 1 vs Day 14	0.0047
+TGFB, Day 1 vs Day 21	0.0002
+TGFB, Day 4 vs Day 7	0.0004
+TGFB, Day 4 vs Day 14	0.0021
+TGFB, Day 4 vs Day 21	0.0002
+TGFB, Day 7 vs Day 14	0.0002
+TGFB, Day 7 vs Day 21	<.0001
+TGFB, Day 14 vs Day 21	0.0002

Supplementary Table 4: Additional statistical information related to aggrecan RNA abundance differentiating MSCs. Companion table to **Figure 4-1**e. P-values indicated for all planned contrasts between –TGFβ and +TGFβ groups over 21 days of culture. Aggrecan RNA count means compared by t-tests with Satterthwaite approximation and simulated adjustment for multiple comparisons.

Supplementary Table 5		
Comparison		**Adjusted P-value**
< 12 hrs	**[12-23 hrs]**	<.0001
< 12 hrs	**[24-35 hrs]**	<.0001
< 12 hrs	**[36-47 hrs]**	<.0001
< 12 hrs	**> 47 hrs**	<.0001
[12-23 hrs]	**[24-35 hrs]**	0.8759
[12-23 hrs]	**[36-47 hrs]**	0.7564
[12-23 hrs]	**> 47 hrs**	0.9997
[24-35 hrs]	**[36-47 hrs]**	<.0001
[24-35 hrs]	**> 47 hrs**	<.0001
[36-47 hrs]	**> 47 hrs**	<.0001

Supplementary Table 5: Additional statistical information related to divergence in aggrecan RNA abundance between sister cells. Companion table to Figure 4-6f. Means compared by t-tests with Satterthwaite approximation and simulated adjustment for multiple comparisons, n = 81 sister cell pairs.

Supplementary Table 6		
Comparison		Adjusted P-value
< 12 hrs	**[12-23 hrs]**	0.1649
< 12 hrs	**[24-35 hrs]**	0.6435
< 12 hrs	**[36-47 hrs]**	<.0001
< 12 hrs	**> 47 hrs**	<.0001
[12-23 hrs]	**[24-35 hrs]**	0.0173
[12-23 hrs]	**[36-47 hrs]**	<.0001
[12-23 hrs]	**> 47 hrs**	<.0001
[24-35 hrs]	**[36-47 hrs]**	<.0001
[24-35 hrs]	**> 47 hrs**	<.0001
[36-47 hrs]	**> 47 hrs**	<.0001

Supplementary Table 6: Additional statistical information related to divergence in GAPDH RNA abundance between sister cells. Companion table to Figure 4-7b. Means compared by t-tests with Satterthwaite approximation and simulated adjustment for multiple comparisons, n = 81 sister cell pairs.

Supplementary Table7				
Passage Comparison		**Aggrecan/GAPDH (Pooled Population) Adjusted P-value**	**Aggrecan per Cell Adjusted P-value**	**GAPDH per Cell Adjusted P-value**
0	1	0.0079	0.2108	0.0006
0	3	0.0004	0.9342	0.0006
0	5	0.0002	0.9999	0.0008
0	7	0.0002	1.0000	0.0008
0	9	0.0011	0.9984	0.0025
1	3	0.7036	0.6621	1.0000
1	5	0.5040	0.2891	0.9945
1	7	0.4593	0.1658	0.9929
1	9	0.9377	0.3623	0.7026
3	5	0.9993	0.9783	0.9981
3	7	0.9982	0.8850	0.9972
3	9	0.9944	0.9934	0.7784
5	7	1.0000	0.9988	1.0000
5	9	0.9510	0.9999	0.9343
7	9	0.9309	0.9941	0.9431

Supplementary Table 7: Additional statistical information related to RNA abundance in chondrocytes during passage-induced de-differentiation. Companion table to Figure 4-8b-d. P-values indicated for all planned contrasts between chondrocytes at different passages (passages 0-9). Pooled aggrecan/GAPDH levels compared via one-way ANOVA with Tukey's post-hoc test. Single cell RNA count means compared by t-tests with Satterthwaite approximation and simulated adjustment for multiple comparisons.

Supplementary Table 8			
Comparison	**Aggrecan per Cell Adjusted P-value**	**GAPDH per Cell Adjusted P-value**	**Aggrecan/GAPDH per Cell Adjusted P-value**
Day 1, P0 vs P5	0.0016	0.0406	0.9983
P0, Day 1 vs 14	0.0897	0.0006	0.0176
Day 14, P0 vs P5	0.0002	0.6714	0.0588
P5, Day 1 vs 14	<.0001	<.0001	0.1982

Supplementary Table 8: Additional statistical information related to RNA abundance in chondrocytes during re-differentiation in agarose. Companion table to Figure 4-8h-j. P-values indicated for all planned contrasts between chondrocytes at different passages (passages 0 and 5) and culture time points (days 1 and 14). Pooled aggrecan/GAPDH levels compared via one-way ANOVA with Tukey's post-hoc test. Single cell RNA count means compared by t-tests with Satterthwaite approximation and simulated adjustment for multiple comparisons.

Appendix 3: Methionine Content of Extracellular Matrix Proteins

Methionine content of extracellular matrix structural constituents (GO:0005201) & proteinaceous extracellular matrix (GO:0005578).

Retrieved from UniProtKB/Swiss-Prot October2016, and filtered for human proteins.

UniProt ID	Protein Name	Gene	# Met	# Amino Acids	% Met
P08253	72 kDa type IV collagenase	MMP2	12	660	1.82%
Q9UHI8	A disintegrin and metalloproteinase with thrombospondin motifs 1	ADAMTS1	15	967	1.55%
Q9H324	A disintegrin and metalloproteinase with thrombospondin motifs 10	ADAMTS10	12	1103	1.09%
P58397	A disintegrin and metalloproteinase with thrombospondin motifs 12	ADAMTS12	30	1594	1.88%
Q76LX8	A disintegrin and metalloproteinase with thrombospondin motifs 13	ADAMTS13	24	1427	1.68%
Q8WXS8	A disintegrin and metalloproteinase with thrombospondin motifs 14	ADAMTS14	17	1223	1.39%
Q8TE58	A disintegrin and metalloproteinase with thrombospondin motifs 15	ADAMTS15	11	950	1.16%
Q8TE57	A disintegrin and metalloproteinase with thrombospondin motifs 16	ADAMTS16	26	1224	2.12%
Q8TE56	A disintegrin and metalloproteinase with thrombospondin motifs 17	ADAMTS17	16	1095	1.46%
Q8TE60	A disintegrin and metalloproteinase with thrombospondin motifs 18	ADAMTS18	15	1221	1.23%
Q8TE59	A disintegrin and metalloproteinase with thrombospondin motifs 19	ADAMTS19	28	1207	2.32%
O95450	A disintegrin and metalloproteinase with thrombospondin motifs 2	ADAMTS2	26	1211	2.15%
P59510	A disintegrin and metalloproteinase with thrombospondin motifs 20	ADAMTS20	35	1910	1.83%
O15072	A disintegrin and metalloproteinase with thrombospondin motifs 3	ADAMTS3	27	1205	2.24%
O75173	A disintegrin and metalloproteinase with thrombospondin motifs 4	ADAMTS4	14	837	1.67%
Q9UNA0	A disintegrin and metalloproteinase with thrombospondin motifs 5	ADAMTS5	10	930	1.08%
Q9UKP5	A disintegrin and metalloproteinase with thrombospondin motifs 6	ADAMTS6	17	1117	1.52%
Q9UKP4	A disintegrin and metalloproteinase with thrombospondin motifs 7	ADAMTS7	15	1686	0.89%

Q9UP79	A disintegrin and metalloproteinase with thrombospondin motifs 8	ADAMTS8	11	889	1.24%
Q9P2N4	A disintegrin and metalloproteinase with thrombospondin motifs 9	ADAMTS9	26	1935	1.34%
P22303	Acetylcholinesterase	ACHE	9	614	1.47%
Q9Y215	Acetylcholinesterase collagenic tail peptide	COLQ	14	455	3.08%
Q8N6G6	ADAMTS-like protein 1	ADAMTSL1	17	1762	0.96%
Q86TH1	ADAMTS-like protein 2	ADAMTSL2	14	951	1.47%
P82987	ADAMTS-like protein 3	ADAMTSL3	30	1691	1.77%
Q6UY14	ADAMTS-like protein 4	ADAMTSL4	7	1074	0.65%
Q6ZMM2	ADAMTS-like protein 5	ADAMTSL5	4	481	0.83%
P16112	Aggrecan core protein	ACAN	13	2415	0.54%
O00468	Agrin	AGRN	13	2067	0.63%
P05186	Alkaline phosphatase, tissue-nonspecific isozyme	ALPL	15	524	2.86%
P01009	Alpha-1-antitrypsin	SERPINA1	10	418	2.39%
P37840	Alpha-synuclein	SNCA	4	140	2.86%
O75443	Alpha-tectorin	TECTA	21	2155	0.97%
Q9NP70	Ameloblastin	AMBN	25	447	5.59%
Q99217	Amelogenin, X isoform	AMELX	11	191	5.76%
Q99218	Amelogenin, Y isoform	AMELY	9	206	4.37%
Q6UX39	Amelotin	AMTN	3	209	1.44%
P51693	Amyloid-like protein 1	APLP1	16	650	2.46%
P03950	Angiogenin	ANG	3	147	2.04%
Q9BY76	Angiopoietin-related protein 4	ANGPTL4	7	406	1.72%
P07355	Annexin A2	ANXA2	9	339	2.65%
P23352	Anosmin-1	ANOS1	5	680	0.74%
Q9BXN1	Asporin	ASPN	12	380	3.16%
P98160	Basement membrane-specific heparan sulfate proteoglycan core protein	HSPG2	43	4391	0.98%
Q96PL2	Beta-tectorin	TECTB	5	329	1.52%
P21810	Biglycan	BGN	8	368	2.17%
P13497	Bone morphogenetic protein 1	BMP1	10	986	1.01%
P12644	Bone morphogenetic protein 4	BMP4	8	408	1.96%
Q96GW7	Brevican core protein	BCAN	8	911	0.88%

Q2TAL6	Brorin	VWC2	5	325	1.54%
P27797	Calreticulin	CALR	5	417	1.20%
Q8N4T0	Carboxypeptidase A6	CPA6	12	437	2.75%
Q66K79	Carboxypeptidase Z	CPZ	16	652	2.45%
Q9NQ79	Cartilage acidic protein 1	CRTAC1	13	661	1.97%
O75339	Cartilage intermediate layer protein 1	CILP	26	1184	2.20%
Q8IUL8	Cartilage intermediate layer protein 2	CILP2	14	1156	1.21%
P21941	Cartilage matrix protein	MATN1	8	496	1.61%
P49747	Cartilage oligomeric matrix protein	COMP	8	757	1.06%
O75718	Cartilage-associated protein	CRTAP	9	401	2.24%
P48509	CD151 antigen	CD151	5	253	1.98%
P36222	Chitinase-3-like protein 1	CHI3L1	7	383	1.83%
O15335	Chondroadherin	CHAD	3	359	0.84%
Q6NUI6	Chondroadherin-like protein	CHADL	2	762	0.26%
O43405	Cochlin	COCH	10	550	1.82%
Q76M96	Coiled-coil domain-containing protein 80	CCDC80	29	950	3.05%
P02452	Collagen alpha-1(I) chain	COL1A1	13	1464	0.89%
P08123	Collagen alpha-2(I) chain	COL1A2	10	1366	0.73%
P02458	Collagen alpha-1(II) chain	COL2A1	16	1487	1.08%
P02461	Collagen alpha-1(III) chain	COL3A1	17	1466	1.16%
P02462	Collagen alpha-1(IV) chain	COL4A1	31	1669	1.86%
P08572	Collagen alpha-2(IV) chain	COL4A2	25	1712	1.46%
Q01955	Collagen alpha-3(IV) chain	COL4A3	29	1670	1.74%
P53420	Collagen alpha-4(IV) chain	COL4A4	24	1690	1.42%
P29400	Collagen alpha-5(IV) chain	COL4A5	26	1685	1.54%
Q14031	Collagen alpha-6(IV) chain	COL4A6	25	1691	1.48%
P20908	Collagen alpha-1(V) chain	COL5A1	23	1838	1.25%
P05997	Collagen alpha-2(V) chain	COL5A2	24	1499	1.60%
P25940	Collagen alpha-3(V) chain	COL5A3	12	1745	0.69%
P12109	Collagen alpha-1(VI) chain	COL6A1	11	1028	1.07%
P12110	Collagen alpha-2(VI) chain	COL6A2	13	1019	1.28%
P12111	Collagen alpha-3(VI) chain	COL6A3	37	3177	1.16%
A8TX70	Collagen alpha-5(VI) chain	COL6A5	55	2615	2.10%

A6NMZ7	Collagen alpha-6(VI) chain	COL6A6	47	2263	2.08%
Q02388	Collagen alpha-1(VII) chain	COL7A1	18	2944	0.61%
P27658	Collagen alpha-1(VIII) chain	COL8A1	23	744	3.09%
P25067	Collagen alpha-2(VIII) chain	COL8A2	9	703	1.28%
P20849	Collagen alpha-1(IX) chain	COL9A1	12	921	1.30%
Q14055	Collagen alpha-2(IX) chain	COL9A2	11	689	1.60%
Q14050	Collagen alpha-3(IX) chain	COL9A3	6	684	0.88%
Q03692	Collagen alpha-1(X) chain	COL10A1	11	680	1.62%
P12107	Collagen alpha-1(XI) chain	COL11A1	24	1806	1.33%
P13942	Collagen alpha-2(XI) chain	COL11A2	16	1736	0.92%
Q99715	Collagen alpha-1(XII) chain	COL12A1	45	3063	1.47%
Q05707	Collagen alpha-1(XIV) chain	COL14A1	33	1796	1.84%
P39059	Collagen alpha-1(XV) chain	COL15A1	30	1388	2.16%
Q07092	Collagen alpha-1(XVI) chain	COL16A1	26	1604	1.62%
Q9UMD9	Collagen alpha-1(XVII) chain	COL17A1	33	1497	2.20%
P39060	Collagen alpha-1(XVIII) chain	COL18A1	15	1754	0.86%
Q14993	Collagen alpha-1(XIX) chain	COL19A1	18	1142	1.58%
Q96P44	Collagen alpha-1(XXI) chain	COL21A1	12	957	1.25%
Q8NFW1	Collagen alpha-1(XXII) chain	COL22A1	20	1626	1.23%
Q17RW2	Collagen alpha-1(XXIV) chain	COL24A1	24	1714	1.40%
Q96A83	Collagen alpha-1(XXVI) chain	COL26A1	6	441	1.36%
Q8IZC6	Collagen alpha-1(XXVII) chain	COL27A1	38	1860	2.04%
Q6UXH8	Collagen and calcium-binding EGF domain-containing protein 1	CCBE1	8	406	1.97%
Q96CG8	Collagen triple helix repeat-containing protein 1	CTHRC1	4	243	1.65%
P45452	Collagenase 3	MMP13	12	471	2.55%
P29279	Connective tissue growth factor	CTGF	13	349	3.72%
P54108	Cysteine-rich secretory protein 3	CRISP3	5	245	2.04%
Q9H0B8	Cysteine-rich secretory protein LCCL domain-containing 2	CRISPLD2	12	497	2.41%
P07585	Decorin	DCN	5	359	1.39%
Q13316	Dentin matrix acidic phosphoprotein 1	DMP1	6	513	1.17%
Q9NZW4	Dentin sialophosphoprotein	DSPP	4	1301	0.31%
Q07507	Dermatopontin	DPT	8	201	3.98%

Q12959	Disks large homolog 1	DLG1	13	904	1.44%
Q03001	Dystonin	DST	158	7570	2.09%
Q14118	Dystroglycan	DAG1	19	895	2.12%
Q9Y5L3	Ectonucleoside triphosphate diphosphohydrolase 2	ENTPD2	9	495	1.82%
Q12805	EGF-containing fibulin-like extracellular matrix protein 1	EFEMP1	11	493	2.23%
O95967	EGF-containing fibulin-like extracellular matrix protein 2	EFEMP2	8	443	1.81%
P19957	Elafin	PI3	3	117	2.56%
P15502	Elastin	ELN	1	786	0.13%
Q96A84	EMI domain-containing protein 1	EMID1	12	441	2.72%
Q9Y6C2	EMILIN-1	EMILIN1	7	1016	0.69%
Q9BXX0	EMILIN-2	EMILIN2	18	1053	1.71%
Q9NT22	EMILIN-3	EMILIN3	6	766	0.78%
Q9NRM1	Enamelin	ENAM	21	1142	1.84%
Q9HCU0	Endosialin	CD248	8	757	1.06%
Q8IUX8	Epidermal growth factor-like protein 6	EGFL6	9	553	1.63%
Q99645	Epiphycan	EPYC	5	322	1.55%
Q96RT1	Erbin	ERBIN	31	1412	2.20%
P43003	Excitatory amino acid transporter 1	SLC1A3	26	542	4.80%
Q16610	Extracellular matrix protein 1	ECM1	5	540	0.93%
O94769	Extracellular matrix protein 2	ECM2	13	699	1.86%
Q86XX4	Extracellular matrix protein FRAS1	FRAS1	61	4008	1.52%
P35555	Fibrillin-1	FBN1	52	2871	1.81%
P35556	Fibrillin-2	FBN2	59	2912	2.03%
Q75N90	Fibrillin-3	FBN3	38	2809	1.35%
P05230	Fibroblast growth factor 1	FGF1	2	155	1.29%
P31371	Fibroblast growth factor 9	FGF9	2	208	0.96%
Q06828	Fibromodulin	FMOD	7	376	1.86%
P02751	Fibronectin	FN1	27	2386	1.13%
P23142	Fibulin-1	FBLN1	8	703	1.14%
P98095	Fibulin-2	FBLN2	12	1184	1.01%
Q9UBX5	Fibulin-5	FBLN5	11	448	2.46%
Q53RD9	Fibulin-7	FBLN7	7	439	1.59%

Q5H8C1	FRAS1-related extracellular matrix protein 1	FREM1	39	2179	1.79%
Q5SZK8	FRAS1-related extracellular matrix protein 2	FREM2	66	3169	2.08%
P0C091	FRAS1-related extracellular matrix protein 3	FREM3	34	2139	1.59%
P09382	Galectin-1	LGALS1	2	135	1.48%
Q08380	Galectin-3-binding protein	LGALS3BP	9	585	1.54%
Q6ZMI3	Gliomedin	GLDN	15	551	2.72%
Q3B7J2	Glucose-fructose oxidoreductase domain-containing protein 2	GFOD2	17	385	4.42%
P78417	Glutathione S-transferase omega-1	GSTO1	7	241	2.90%
P35052	Glypican-1	GPC1	11	558	1.97%
Q8N158	Glypican-2	GPC2	10	579	1.73%
P51654	Glypican-3	GPC3	23	580	3.97%
O75487	Glypican-4	GPC4	18	556	3.24%
P78333	Glypican-5	GPC5	18	572	3.15%
Q9Y625	Glypican-6	GPC6	19	555	3.42%
Q92896	Golgi apparatus protein 1	GLG1	38	1179	3.22%
Q96RW7	Hemicentin-1	HMCN1	73	5635	1.30%
Q8NDA2	Hemicentin-2	HMCN2	42	5059	0.83%
Q9Y251	Heparanase	HPSE	9	543	1.66%
P10915	Hyaluronan and proteoglycan link protein 1	HAPLN1	2	354	0.56%
Q9GZV7	Hyaluronan and proteoglycan link protein 2	HAPLN2	3	340	0.88%
Q96S86	Hyaluronan and proteoglycan link protein 3	HAPLN3	2	360	0.56%
Q86UW8	Hyaluronan and proteoglycan link protein 4	HAPLN4	2	402	0.50%
Q8WWQ2	Inactive heparanase-2	HPSE2	10	592	1.69%
P05019	Insulin-like growth factor I	IGF1	5	195	2.56%
P23229	Integrin alpha-6	ITGA6	15	1130	1.33%
P05556	Integrin beta-1	ITGB1	15	798	1.88%
Q01638	Interleukin-1 receptor-like 1	IL1RL1	12	556	2.16%
Q17R60	Interphotoreceptor matrix proteoglycan 1	IMPG1	18	797	2.26%

Q9BZV3	Interphotoreceptor matrix proteoglycan 2	IMPG2	23	1241	1.85%
P03956	Interstitial collagenase	MMP1	10	469	2.13%
Q96I82	Kazal-type serine protease inhibitor domain-containing protein 1	KAZALD1	3	304	0.99%
O60938	Keratocan	KERA	10	352	2.84%
O00515	Ladinin-1	LAD1	6	517	1.16%
P25391	Laminin subunit alpha-1	LAMA1	40	3075	1.30%
P24043	Laminin subunit alpha-2	LAMA2	46	3122	1.47%
Q16787	Laminin subunit alpha-3	LAMA3	44	3333	1.32%
Q16363	Laminin subunit alpha-4	LAMA4	32	1823	1.76%
O15230	Laminin subunit alpha-5	LAMA5	41	3695	1.11%
P07942	Laminin subunit beta-1	LAMB1	32	1786	1.79%
P55268	Laminin subunit beta-2	LAMB2	16	1798	0.89%
Q13751	Laminin subunit beta-3	LAMB3	28	1172	2.39%
A4D0S4	Laminin subunit beta-4	LAMB4	18	1761	1.02%
P11047	Laminin subunit gamma-1	LAMC1	22	1609	1.37%
Q13753	Laminin subunit gamma-2	LAMC2	22	1193	1.84%
Q9Y6N6	Laminin subunit gamma-3	LAMC3	14	1575	0.89%
Q14766	Latent-transforming growth factor beta-binding protein 1	LTBP1	19	1721	1.10%
Q14767	Latent-transforming growth factor beta-binding protein 2	LTBP2	16	1821	0.88%
Q8N2S1	Latent-transforming growth factor beta-binding protein 4	LTBP4	10	1624	0.62%
Q9NZU1	Leucine-rich repeat transmembrane protein FLRT1	FLRT1	12	646	1.86%
O43155	Leucine-rich repeat transmembrane protein FLRT2	FLRT2	13	660	1.97%
Q9NZU0	Leucine-rich repeat transmembrane protein FLRT3	FLRT3	12	649	1.85%
O75829	Leukocyte cell-derived chemotaxin 1	LECT1	8	334	2.40%
P16150	Leukosialin	SPN	10	400	2.50%
P51884	Lumican	LUM	4	338	1.18%
Q08397	Lysyl oxidase homolog 1	LOXL1	3	574	0.52%
Q9Y4K0	Lysyl oxidase homolog 2	LOXL2	20	774	2.58%
P39900	Macrophage metalloelastase	MMP12	12	470	2.55%

Q7Z304	MAM domain-containing protein 2	MAMDC2	15	686	2.19%
O00339	Matrilin-2	MATN2	19	956	1.99%
O15232	Matrilin-3	MATN3	6	486	1.23%
P09237	Matrilysin	MMP7	9	267	3.37%
Q9NQ76	Matrix extracellular phosphoglycoprotein	ME	8	525	1.52%
P08493	Matrix Gla protein	MGP	3	103	2.91%
P51512	Matrix metalloproteinase-16	MMP16	15	607	2.47%
Q9ULZ9	Matrix metalloproteinase-17	MMP17	10	603	1.66%
Q99542	Matrix metalloproteinase-19	MMP19	10	508	1.97%
O60882	Matrix metalloproteinase-20	MMP20	14	483	2.90%
O75900	Matrix metalloproteinase-23	MMP23A	5	390	1.28%
Q9Y5R2	Matrix metalloproteinase-24	MMP24	10	645	1.55%
Q9NPA2	Matrix metalloproteinase-25	MMP25	9	562	1.60%
Q9NRE1	Matrix metalloproteinase-26	MMP26	5	261	1.92%
Q9H239	Matrix metalloproteinase-28	MMP28	5	520	0.96%
P14780	Matrix metalloproteinase-9	MMP9	9	707	1.27%
P01033	Metalloproteinase inhibitor 1	TIMP1	4	207	1.93%
P16035	Metalloproteinase inhibitor 2	TIMP2	6	220	2.73%
P35625	Metalloproteinase inhibitor 3	TIMP3	6	211	2.84%
Q99727	Metalloproteinase inhibitor 4	TIMP4	4	224	1.79%
P55083	Microfibril-associated glycoprotein 4	MFAP4	4	255	1.57%
P55081	Microfibrillar-associated protein 1	MFAP1	14	439	3.19%
P55001	Microfibrillar-associated protein 2	MFAP2	2	183	1.09%
Q13361	Microfibrillar-associated protein 5	MFAP5	4	173	2.31%
P20774	Mimecan	OGN	2	298	0.67%
Q02505	Mucin-3A	MUC3A	74	3323	2.23%
Q99102	Mucin-4	MUC4	34	2169	1.57%
P98088	Mucin-5AC	MUC5AC	45	5654	0.80%
Q6W4X9	Mucin-6	MUC6	31	2439	1.27%
Q9H8L6	Multimerin-2	MMRN2	15	949	1.58%
Q9H1U4	Multiple epidermal growth factor-like domains protein 9	MEGF9	4	602	0.66%
Q99972	Myocilin	MYOC	8	504	1.59%
Q2UY09	NA	COL28A1	14	1125	1.24%

Q15063	NA	POSTN	17	836	2.03%
Q9BQB4	NA	SOST	2	213	0.94%
Q6UXI9	Nephronectin	NPNT	8	565	1.42%
O95631	Netrin-1	NTN1	10	604	1.66%
O00634	Netrin-3	NTN3	4	580	0.69%
Q9HB63	Netrin-4	NTN4	11	628	1.75%
O00533	Neural cell adhesion molecule L1-like protein	CHL1	15	1208	1.24%
O14594	Neurocan core protein	NCAN	24	1321	1.82%
Q8IVL1	Neuron navigator 2	NAV2	53	2488	2.13%
P22894	Neutrophil collagenase	MMP8	8	467	1.71%
P14543	Nidogen-1	NID1	8	1247	0.64%
Q14112	Nidogen-2	NID2	8	1375	0.58%
Q9GZU5	Nyctalopin	NYX	4	481	0.83%
A1E959	Odontogenic ameloblast-associated protein	ODAM	8	279	2.87%
Q68BL7	Olfactomedin-like protein 2A	OLFML2A	11	652	1.69%
Q9UBM4	Opticin	OPTC	5	332	1.51%
Q99983	Osteomodulin	OMD	12	421	2.85%
Q7RTW8	Otoancorin	OTOA	25	1153	2.17%
O95428	Papilin	PAPLN	9	1278	0.70%
O14936	Peripheral plasma membrane protein CASK	CASK	23	926	2.48%
Q92626	Peroxidasin homolog	PXDN	24	1479	1.62%
P80108	Phosphatidylinositol-glycan-specific phospholipase D	GPLD1	18	840	2.14%
Q8TCT1	Phosphoethanolamine/phosphocholine phosphatase	PHOSPHO1	7	267	2.62%
P36955	Pigment epithelium-derived factor	SERPINF1	6	418	1.44%
Q63HQ2	Pikachurin	EGFLAM	21	1017	2.06%
P21246	Pleiotrophin	PTN	3	168	1.79%
Q7Z5L7	Podocan	PODN	4	613	0.65%
Q6PEZ8	Podocan-like protein 1	PODNL1	5	512	0.98%
Q5K4E3	Polyserase-2	PRSS36	9	855	1.05%
P51888	Prolargin	PRELP	6	382	1.57%
Q32P28	Prolyl 3-hydroxylase 1	P3H1	16	736	2.17%

Q8IVL5	Prolyl 3-hydroxylase 2	P3H2	14	708	1.98%
O00622	Protein CYR61	CYR61	6	381	1.57%
Q14129	Protein DGCR6	DGCR6	4	220	1.82%
Q8TB73	Protein NDNF	NDNF	6	568	1.06%
P48745	Protein NOV homolog	NOV	6	357	1.68%
Q96JX3	Protein SERAC1	SERAC1	16	654	2.45%
Q9GZT5	Protein Wnt-10a	WNT10A	8	417	1.92%
O00744	Protein Wnt-10b	WNT10B	6	389	1.54%
O96014	Protein Wnt-11	WNT11	11	354	3.11%
Q9UBV4	Protein Wnt-16	WNT16	12	365	3.29%
P09544	Protein Wnt-2	WNT2	11	360	3.06%
Q93097	Protein Wnt-2b	WNT2B	6	391	1.53%
P56704	Protein Wnt-3a	WNT3A	6	352	1.70%
P56705	Protein Wnt-4	WNT4	6	351	1.71%
P41221	Protein Wnt-5a	WNT5A	12	380	3.16%
Q9H1J7	Protein Wnt-5b	WNT5B	9	359	2.51%
Q9Y6F9	Protein Wnt-6	WNT6	5	365	1.37%
O00755	Protein Wnt-7a	WNT7A	8	349	2.29%
P56706	Protein Wnt-7b	WNT7B	6	349	1.72%
Q9H1J5	Protein Wnt-8a	WNT8A	8	351	2.28%
Q93098	Protein Wnt-8b	WNT8B	5	351	1.42%
O14904	Protein Wnt-9a	WNT9A	4	365	1.10%
O14905	Protein Wnt-9b	WNT9B	4	357	1.12%
P28300	Protein-lysine 6-oxidase	LOX	5	417	1.20%
P04628	Proto-oncogene Wnt-1	WNT1	5	370	1.35%
P56703	Proto-oncogene Wnt-3	WNT3	5	355	1.41%
Q8IWL2	Pulmonary surfactant-associated protein A1	SFTPA1	6	248	2.42%
Q8IWL1	Pulmonary surfactant-associated protein A2	SFTPA2	5	248	2.02%
P35247	Pulmonary surfactant-associated protein D	SFTPD	8	375	2.13%
A6NMY6	Putative annexin A2-like protein	ANXA2P2	9	339	2.65%
Q16473	Putative tenascin-XA	TNXA	4	311	1.29%
P23471	Receptor-type tyrosine-protein phosphatase zeta	PTPRZ1	42	2315	1.81%

P78509	Reelin	RELN	70	3460	2.02%
Q8NC24	RELT-like protein 2	RELL2	9	303	2.97%
Q6XPR3	Repetin	RPTN	4	784	0.51%
P10745	Retinol-binding protein 3	RBP3	30	1247	2.41%
Q8N474	Secreted frizzled-related protein 1	SFRP1	10	314	3.18%
P02768	Serum albumin	ALB	7	609	1.15%
O75093	Slit homolog 1 protein	SLIT1	20	1534	1.30%
O94813	Slit homolog 2 protein	SLIT2	16	1529	1.05%
O75094	Slit homolog 3 protein	SLIT3	16	1523	1.05%
Q8IVN8	Somatomedin-B and thrombospondin type-1 domain-containing protein	SBSPON	5	264	1.89%
Q15465	Sonic hedgehog protein	SHH	5	462	1.08%
P09486	SPARC	SPARC	5	303	1.65%
Q14515	SPARC-like protein 1	SPARCL1	12	664	1.81%
Q9H4F8	SPARC-related modular calcium-binding protein 1	SMOC1	7	434	1.61%
Q9H3U7	SPARC-related modular calcium-binding protein 2	SMOC2	3	446	0.67%
Q9HCB6	Spondin-1	SPON1	27	807	3.35%
Q9BUD6	Spondin-2	SPON2	3	331	0.91%
P08254	Stromelysin-1	MMP3	7	477	1.47%
P09238	Stromelysin-2	MMP10	10	476	2.10%
P24347	Stromelysin-3	MMP11	5	488	1.02%
Q9UQE7	Structural maintenance of chromosomes protein 3	SMC3	35	1217	2.88%
P01730	T-cell surface glycoprotein CD4	CD4	7	458	1.53%
Q7Z7G0	Target of Nesh-SH3	ABI3BP	6	1075	0.56%
P24821	Tenascin	TNC	23	2201	1.04%
Q9UQP3	Tenascin-N	TNN	21	1299	1.62%
Q92752	Tenascin-R	TNR	22	1358	1.62%
P22105	Tenascin-X	TNXB	55	4242	1.30%
Q08629	Testican-1	SPOCK1	5	439	1.14%
Q92563	Testican-2	SPOCK2	6	424	1.42%
Q9BQ16	Testican-3	SPOCK3	6	436	1.38%
Q6ZMP0	Thrombospondin type-1 domain-containing protein 4	THSD4	19	1018	1.87%

P35442	Thrombospondin-2	THBS2	12	1172	1.02%
P35443	Thrombospondin-4	THBS4	9	961	0.94%
P48307	Tissue factor pathway inhibitor 2	TFPI2	4	235	1.70%
Q03167	Transforming growth factor beta receptor type 3	TGFBR3	20	851	2.35%
P01137	Transforming growth factor beta-1	TGFB1	7	390	1.79%
Q15582	Transforming growth factor-beta-induced protein ig-h3	TGFBI	14	683	2.05%
Q07654	Trefoil factor 3	TFF3	2	80	2.50%
P0DKB5	Trophoblast glycoprotein-like	TPBGL	4	382	1.05%
Q9UJW2	Tubulointerstitial nephritis antigen	TINAG	10	476	2.10%
Q9GZM7	Tubulointerstitial nephritis antigen-like	TINAGL1	15	467	3.21%
Q9UBB9	Tuftelin-interacting protein 11	TFIP11	26	837	3.11%
O00300	Tumor necrosis factor receptor superfamily member 11B	TNFRSF11B	5	401	1.25%
Q6UXA7	Uncharacterized protein C6orf15	C6orf15	5	325	1.54%
Q8WVF2	Unique cartilage matrix-associated protein	UCMA	5	138	3.62%
O75445	Usherin	USH2A	78	5202	1.50%
P15692	Vascular endothelial growth factor A	VEGFA	8	232	3.45%
P13611	Versican core protein	VCAN	55	3396	1.62%
Q53GQ0	Very-long-chain 3-oxoacyl-CoA reductase	HSD17B12	10	312	3.21%
Q6UXI7	Vitrin	VIT	9	678	1.33%
P04004	Vitronectin	VTN	7	478	1.46%
P04275	von Willebrand factor	VWF	56	2813	1.99%
Q6PCB0	von Willebrand factor A domain-containing protein 1	VWA1	5	445	1.12%
Q5GFL6	von Willebrand factor A domain-containing protein 2	VWA2	12	755	1.59%
O95388	WNT1-inducible-signaling pathway protein 1	WISP1	5	367	1.36%
O76076	WNT1-inducible-signaling pathway protein 2	WISP2	2	250	0.80%
O95389	WNT1-inducible-signaling pathway protein 3	WISP3	6	354	1.69%
P60852	Zona pellucida sperm-binding protein 1	ZP1	8	638	1.25%
Q05996	Zona pellucida sperm-binding protein 2	ZP2	22	745	2.95%

P21754	Zona pellucida sperm-binding protein 3	ZP3	8	424	1.89%
Q12836	Zona pellucida sperm-binding protein 4	ZP4	6	540	1.11%
O60844	Zymogen granule membrane protein 16	ZG16	1	167	0.60%

BIBLIOGRAPHY

1. Albayrak C, Jordi CA, Zechner C, et al. Digital Quantification of Proteins and mRNA in Single Mammalian Cells. *Mol Cell*. 2016;61(6):914-924. doi:10.1016/j.molcel.2016.02.030.

2. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Molecular Biology of the Cell. 2002.

3. Altschuler SJ, Wu LF. Cellular heterogeneity: do differences make a difference? *Cell*. 2010;141(4):559-563. doi:10.1016/j.cell.2010.04.033.

4. Anders S, Pyl PT, Huber W. HTSeq - A Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2014;31(2):166-169. doi:10.1093/bioinformatics/btu638.

5. Aratyn-Schaus Y, Oakes PW, Stricker J, Winter SP, Gardel ML. Preparation of Complaint Matrices for Quantifying Cellular Contraction. *J Vis Exp*. 2010;(46):e2173-e2173. doi:10.3791/2173.

6. Balázsi G, van Oudenaarden A, Collins JJ. Cellular decision making and biological noise: from microbes to mammals. *Cell*. 2011;144(6):910-925. doi:10.1016/j.cell.2011.01.030.

7. Barry F, Boynton RE, Liu B, Murphy JM. Chondrogenic Differentiation of Mesenchymal Stem Cells from Bone Marrow: Differentiation-Dependent Gene Expression of Matrix Components. *Exp Cell Res*. 2001;268(2):189-200. doi:10.1006/excr.2001.5278.

8. Bartov E, Jerdan JA, Glaser BM. A simple technique for isolating pure cell populations from mixed primary cultures. *J Tissue Cult Methods*. 1988;11(4):181-183. doi:10.1007/BF01407311.

9. Beahm BJ, Dehnert KW, Derr NL, et al. A visualizable chain-terminating inhibitor of glycosaminoglycan biosynthesis in developing zebrafish. *Angew Chem Int Ed Engl*. 2014;53(13):3347-3352. doi:10.1002/anie.201310569.

10. Beatty KE, Liu JC, Xie F, et al. Fluorescence visualization of newly synthesized proteins in mammalian cells. *Angew Chem Int Ed Engl*. 2006;45(44):7364-7367. doi:10.1002/anie.200602114.

11. Beekman B, Verzijl N, Bank RA, von der Mark K, TeKoppele JM. Synthesis of collagen by bovine chondrocytes cultured in alginate; posttranslational modifications and cell-matrix interaction. *Exp Cell Res*. 1997;237(1):135-141.

12. Benya PD, Shaffer JD. Dedifferentiated chondrocytes reexpress the differentiated collagen phenotype when cultured in agarose gels. *Cell*. 1982;30(1):215-224.

13. Berdasco M, Esteller M. DNA methylation in stem cell renewal and multipotency. *Stem Cell Res Ther*. 2011;2(5):42. doi:10.1186/scrt83.

14. Besanceney-Webler C, Jiang H, Zheng T, et al. Increasing the efficacy of bioorthogonal click reactions for bioconjugation: a comparative study. *Angew Chem Int Ed Engl*. 2011;50(35):8051-8056. doi:10.1002/anie.201101817.

15. Bonaventure J, Kadhom N, Cohen-Solal L, et al. Reexpression of Cartilage-Specific Genes by Dedifferentiated Human Articular Chondrocytes Cultured in Alginate Beads. *Exp Cell Res*. 1994;212(1):97-104. doi:10.1006/excr.1994.1123.

16. Bondarenko LB, Kovalenko VM. Pyrazinamide Effects on Cartilage Type II Collagen Amino Acid Composition. *Int J Pept*. 2012;2012:781785. doi:10.1155/2012/781785.

17. Buckwalter J, Mankin H, Grodzinsky A. Articular Cartilage and Osteoarthritis. *AAOS Instr Course Lecutures*. 2005;54:465-480.

18. Budnik B, Levy E, Slavov N. Mass-spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *bioRxiv*. 2017:102681. doi:10.1101/102681.

19. Buschmann MD, Gluzband YA, Grodzinsky AJ, Kimura JH, Hunziker EB, Hunziker B. Chondrocytes in Agarose Culture Synthesize a Mechanically Functional Extracellular Matrix. *J Orthop Res*. 1992;10(6):745-758. doi:10.1002/jor.1100100602.

20. Cai L, Dalal CK, Elowitz MB. Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature*. 2008;455(7212):485-490. doi:10.1038/nature07292.

21. Cai R, Nakamoto T, Kawazoe N, Chen G. Influence of stepwise chondrogenesis-mimicking 3D extracellular matrix on chondrogenic differentiation of mesenchymal stem cells. *Biomaterials*. 2015;52:199-207. doi:10.1016/j.biomaterials.2015.02.033.

22. Calve S, Witten AJ, Ocken AR, Kinzer-Ursem TL. Incorporation of non-canonical amino acids into the developing murine proteome. *Sci Rep*. 2016;6:32377. doi:10.1038/srep32377.

23. Campbell J, Bader D, Lee D. Mechanical loading modulates intracellular calcium signaling in human mesenchymal stem cells. *J Appl Biomater …*. 2008.

24. Canty EG, Starborg T, Lu Y, et al. Actin filaments are required for fibripositor-mediated collagen fibril alignment in tendon. *J Biol Chem*. 2006;281(50):38592-38598. doi:10.1074/jbc.M607581200.

25. Caterson B, Flannery CR, Hughes CE, Little CB. Mechanisms of proteoglycan metabolism that lead to cartilage destruction in the pathogenesis of arthritis. *Drugs Today (Barc)*. 35(4-5):397-402.

26. Cescon M, Gattazzo F, Chen P, Bonaldo P. Collagen VI at a glance. *J Cell Sci*. 2015;128(19):3525-3531. doi:10.1242/jcs.169748.

27. Chan CKF, Seo EY, Chen JY, et al. Identification and Specification of the Mouse Skeletal Stem Cell. *Cell*. 2015;160(1-2):285-298. doi:10.1016/j.cell.2014.12.002.

28. Chang HH, Hemberg M, Barahona M, Ingber DE, Huang S. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature*. 2008;453(7194):544-547. doi:10.1038/nature06965.

29. Chang Q, Ornatsky OI, Siddiqui I, Loboda A, Baranov VI, Hedley DW. Imaging Mass Cytometry. *Cytom Part A*. 2017;91(2):160-169. doi:10.1002/cyto.a.23053.

30. Cheng T, Maddox NC, Wong AW, Rahnama R, Kuo AC. Comparison of gene expression patterns in articular cartilage and dedifferentiated articular chondrocytes. *J Orthop Res*. 2012;30(2):234-245. doi:10.1002/jor.21503.

31. Choi PJ, Cai L, Frieda K, Xie XS. A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science*. 2008;322(5900):442-446. doi:10.1126/science.1161427.

32. Chowdhury F, Na S, Li D, et al. Material properties of the cell dictate stress-induced spreading and differentiation in embryonic stem cells. *Nat Mater*. 2010;9(1):82-88. doi:10.1038/nmat2563.

33. Chubb JR, Trcek T, Shenoy SM, Singer RH. Transcriptional pulsing of a developmental gene. *Curr Biol*. 2006;16(10):1018-1025. doi:10.1016/j.cub.2006.03.092.

34. Colter DC, Sekiya I, Prockop DJ. Identification of a subpopulation of rapidly self-renewing and multipotential adult stem cells in colonies of human marrow stromal cells. *Proc Natl Acad Sci U S A*. 2001;98(14):7841-7845. doi:10.1073/pnas.141221698.

35. Comenge J, Held M, Levy R, Carolan G, Mason D. The Spherical Nucleic Acids mRNA Detection Paradox. *Sci Res*. April 2016. doi:10.14293/S2199-1006.1.SOR-CHEM.AZ1MJU.v2.

36. Cote AJ, McLeod CM, Farrell MJ, et al. Single-cell differences in matrix gene expression do not predict matrix deposition. *Nat Commun*. 2016;7:10865. doi:10.1038/ncomms10865.

37. Crosetto N, Bienko M, van Oudenaarden A. Spatially resolved transcriptomics and beyond. *Nat Rev Genet*. 2014;16(1):57-66. doi:10.1038/nrg3832.

38. Csárdi G, Franks A, Choi DS, Airoldi EM, Drummond DA. Accounting for Experimental Noise Reveals That mRNA Levels, Amplified by Post-Transcriptional Processes, Largely Determine Steady-State Protein Levels in Yeast. Snyder M, ed. *PLOS Genet*. 2015;11(5):e1005206. doi:10.1371/journal.pgen.1005206.

39. D'Ippolito G, Schiller PC, Ricordi C, Roos BA, Howard GA. Age-Related Osteogenic Potential of Mesenchymal Stromal Stem Cells from Human Vertebral Bone Marrow. *J Bone Miner Res*. 1999;14(7):1115-1122. doi:10.1359/jbmr.1999.14.7.1115.

40. Dahl KN, Ribeiro AJS, Lammerding J. Nuclear shape, mechanics, and mechanotransduction. *Circ Res*. 2008;102(11):1307-1318. doi:10.1161/CIRCRESAHA.108.173989.

41. Dar RD, Razooky BS, Singh A, et al. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proc Natl Acad Sci*. 2012;109(43):17454-17459. doi:10.1073/pnas.1213530109.

42. Darling EM, Athanasiou K a. Rapid phenotypic changes in passaged articular chondrocyte subpopulations. *J Orthop Res*. 2005;23(2):425-432. doi:10.1016/j.orthres.2004.08.008.

43. Darling, Cruz G. Cellular Mechanical Biomarkers for Lineage-Specific , Stem Cell Enrichment. *ORS Abstr*. (1):54673.

44. Darmanis S, Gallant CJ, Marinescu VD, et al. Simultaneous Multiplexed Measurement of RNA and Proteins in Single Cells. *Cell Rep*. 2016;14(2):380-389. doi:10.1016/j.celrep.2015.12.021.

45. Dieterich DC, Hodas JJL, Gouzer G, et al. In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. *Nat Neurosci*. 2010;13(7):897-905. doi:10.1038/nn.2580.

46. Digirolamo C, Stokes D, Colter D, Phinney D, Class R, Prockop D. Propagation and senescence of human marrow stromal cells in culture : a simple colony-forming assay identifies samples with the greatest potential to propagate and differentiate. *Br J Haematol*. 1999;(107):275-281.

47. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21. doi:10.1093/bioinformatics/bts635.

48. Dominici M, Le Blanc K, Mueller I, et al. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy*. 2006;8(4):315-317. doi:10.1080/14653240600855905.

49. Dozin B, Quarto R, Rossi F, Cancedda R. Stabilization of the Messenger RNA Follows Transcriptional Activation of Type-II Collagen Gene in Differentiating Chicken Chondrocyte. *J Biol Chem*. 1990;265(13):7216-7220.

50. Driscoll TP, Cosgrove BD, Heo S-J, Shurden ZE, Mauck RL. Cytoskeletal to Nuclear Strain Transfer Regulates YAP Signaling in Mesenchymal Stem Cells. *Biophys J*. 2015;108(12):2783-2793. doi:10.1016/j.bpj.2015.05.010.

51. Eldar A, Elowitz MB. Functional roles for noise in genetic circuits. *Nature*. 2010;467(7312):167-173. doi:10.1038/nature09326.

52. Elima K, Vuorio E. Expression of mRNAs for collagens and other matrix components in dedifferentiating and redifferentiating human chondrocytes in culture. *FEBS Lett*. 1989;258(2):195-198.

53. Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. *Science*. 2002;297(5584):1183-1186. doi:10.1126/science.1070919.

54. Elsafadi M, Manikandan M, Atteya M, et al. Characterization of Cellular and Molecular Heterogeneity of Bone Marrow Stromal Cells. *Stem Cells Int*. 2016;2016:1-18. doi:10.1155/2016/9378081.

55. Erdmann I, Marter K, Kobler O, et al. Cell-selective labelling of proteomes in Drosophila melanogaster. *Nat Commun*. 2015;6:7521. doi:10.1038/ncomms8521.

56. Erickson IE, Huang AH, Chung C, Li RT, Burdick JA, Mauck RL. Differential maturation and structure-function relationships in mesenchymal stem cell- and chondrocyte-seeded hydrogels. *Tissue Eng Part A*. 2009;15(5):1041-1052. doi:10.1089/ten.tea.2008.0099.

57. Erickson IE, Huang AH, Sengupta S, Kestle S, Burdick JA, Mauck RL. Macromer density influences mesenchymal stem cell chondrogenesis and maturation in photocrosslinked hyaluronic acid hydrogels. *Osteoarthritis Cartilage*. 2009;17(12):1639-1648. doi:10.1016/j.joca.2009.07.003.

58. Estes BT, Diekman BO, Gimble JM, Guilak F. Isolation of adipose-derived stem cells and their induction to a chondrogenic phenotype. *Nat Protoc*.

2010;5(7):1294-1311. doi:10.1038/nprot.2010.81.

59. Eyre D. Collagen of articular cartilage. *Arthritis Res*. 2002;4(1):30-35. doi:10.1186/ar380.

60. Farrell MJ, Comeau ES, Mauck RL. Mesenchymal stem cells produce functional cartilage matrix in three-dimensional culture in regions of optimal nutrient supply. *Eur Cell Mater*. 2012;23:425-440. doi:vol023a33 [pii].

61. Farrell MJ, Shin JI, Smith LJ, Mauck RL. Functional consequences of glucose and oxygen deprivation on engineered mesenchymal stem cell-based cartilage constructs. *Osteoarthr Cartil*. 2015;23(1):134-142. doi:10.1016/j.joca.2014.09.012.

62. Femino AM, Fay FS, Fogarty K, Singer RH. Visualization of single RNA transcripts in situ. *Science*. 1998;280(5363):585-590.

63. Finkenstädt B, Woodcock DJ, Komorowski M, et al. Quantifying intrinsic and extrinsic noise in gene transcription using the linear noise approximation: an application to single cell data. *Ann Appl Stat*. 2013;7(4):1960-1982. doi:10.1214/13-AOAS669.

64. Frank O, Heim M, Jakob M, et al. Real-time quantitative RT-PCR analysis of human bone marrow stromal cells during osteogenic differentiation in vitro. *J Cell Biochem*. 2002;85(4):737-746. doi:10.1002/jcb.10174.

65. Freeman BT, Jung JP, Ogle BM. Single-Cell RNA-Seq of Bone Marrow-Derived Mesenchymal Stem Cells Reveals Unique Profiles of Lineage Priming. Covas DT, ed. *PLoS One*. 2015;10(9):e0136199. doi:10.1371/journal.pone.0136199.

66. Fu J, Wang Y-K, Yang MT, et al. Mechanical regulation of cell function with geometrically modulated elastomeric substrates. *Nat Methods*. 2010;7(9):733-736. doi:10.1038/nmeth.1487.

67. Fusco D, Accornero N, Lavoie B, et al. Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr Biol*. 2003;13(2):161-167. doi:10.1016/S0960-9822(02)01436-7.

68. Gandhi SJ, Zenklusen D, Lionnet T, Singer RH. Transcription of functionally related constitutive genes is not coordinated. *Nat Struct Mol Biol*. 2011;18(1):27-34. doi:10.1038/nsmb.1934.

69. Gilbert PM, Havenstrite KL, Magnusson KEG, et al. Substrate elasticity regulates skeletal muscle stem cell self-renewal in culture. *Science*. 2010;329(5995):1078-1081. doi:10.1126/science.1191035.

70. Gillesple DT. Exact stochastic simulation of coupled chemical reactions. *J Phys Chem*. 1977;81(25):2340-2361. doi:10.1021/j100540a008.

71. Glowacki J, Trepman E, Folkman J. Cell shape and phenotypic expression in chondrocytes. *Proc Soc Exp Biol Med*. 1983;172(1):93-98.

72. Golding I, Paulsson J, Zawilski SM, Cox EC. Real-time kinetics of gene activity in individual bacteria. *Cell*. 2005;123(6):1025-1036. doi:10.1016/j.cell.2005.09.031.

73. González-Cruz RD, Fonseca VC, Darling EM. Cellular mechanical properties reflect the differentiation potential of adipose-derived mesenchymal stem cells. *Proc Natl Acad Sci U S A*. 2012;109(24):E1523-9. doi:10.1073/pnas.1120349109.

74. Grässel S, Lorenz J. Tissue-engineering strategies to repair chondral and osteochondral tissue in osteoarthritis: use of mesenchymal stem cells. *Curr Rheumatol Rep*. 2014;16(10):452. doi:10.1007/s11926-014-0452-5.

75. Greenwood C, Ruff D, Kirvell S, Johnson G, Dhillon HS, Bustin SA. Proximity assays for sensitive quantification of proteins. *Biomol Detect Quantif*. 2015;4:10-16. doi:10.1016/j.bdq.2015.04.002.

76. Guilak F, Alexopoulos LG, Upton ML, et al. The Pericellular Matrix as a Transducer of Biomechanical and Biochemical Signals in Articular Cartilage. *Ann N Y Acad Sci*. 2006;1068(1):498-512. doi:10.1196/annals.1346.011.

77. Guvendiren M, Burdick JA. Stiffening hydrogels to probe short- and long-term cellular responses to dynamic mechanics. *Nat Commun*. 2012;3:792. doi:10.1038/ncomms1792.

78. Hamidouche Z, Rother K, Przybilla J, et al. Bistable Epigenetic States Explain Age-Dependent Decline in Mesenchymal Stem Cell Heterogeneity. *Stem Cells*. November 2016. doi:10.1002/stem.2514.

79. Han WM, Heo S-J, Driscoll TP, et al. Microstructural heterogeneity directs micromechanics and mechanobiology in native and engineered fibrocartilage. *Nat Mater*. 2016;15(4):477-484. doi:10.1038/nmat4520.

80. Heo S-J, Driscoll TP, Thorpe SD, et al. Differentiation alters stem cell nuclear architecture, mechanics, and mechano-sensitivity. *Elife*. 2016;5:283-293. doi:10.7554/eLife.18207.

81. Heo S-J, Han WM, Szczesny SE, et al. Mechanically Induced Chromatin Condensation Requires Cellular Contractility in Mesenchymal Stem Cells. *Biophys J*. 2016;111(4):864-874. doi:10.1016/j.bpj.2016.07.006.

82. Heo S-J, Thorpe SD, Driscoll TP, et al. Biophysical Regulation of Chromatin Architecture Instills a Mechanical Memory in Mesenchymal Stem Cells. *Sci Rep*. 2015;5:16895. doi:10.1038/srep16895.

83. Hinz FI, Dieterich DC, Tirrell DA, Schuman EM. Noncanonical Amino Acid Labeling in Vivo to Visualize and Affinity Purify Newly Synthesized Proteins in Larval Zebrafish. *ACS Chem Neurosci*. 2012;3(1):40-49. doi:10.1021/cn2000876.

84. Hong V, Steinmetz NF, Manchester M, Finn MG. Labeling live cells by copper-catalyzed alkyne--azide click chemistry. *Bioconjug Chem*. 2010;21(10):1912-1916. doi:10.1021/bc100272z.

85. Hoppe PS, Coutu DL, Schroeder T. Single-cell technologies sharpen up mammalian stem cell research. *Nat Cell Biol*. 2014;16(10):919-927. doi:10.1038/ncb3042.

86. Huang AH, Farrell MJ, Mauck RL. Mechanics and mechanobiology of mesenchymal stem cell-based engineered cartilage. *J Biomech*. 2010;43(1):128-136. doi:10.1016/j.jbiomech.2009.09.018.

87. Huang S. Non-genetic heterogeneity of cells in development: more than just noise. *Development*. 2009;136(23):3853-3862. doi:10.1242/dev.035139.

88. Hughes AJ, Spelke DP, Xu Z, Kang C-C, Schaffer D V, Herr AE. Single-cell western blotting. *Nat Methods*. 2014;11(7):749-755. doi:10.1038/nmeth.2992.

89. Humphrey JD, Dufresne ER, Schwartz MA. Mechanotransduction and extracellular matrix homeostasis. *Nat Rev Mol Cell Biol*. 2014;15(12):802-812. doi:10.1038/nrm3896.

90. Humphrey JD, Dufresne ER, Schwartz MA. Mechanotransduction and extracellular matrix homeostasis. *Nat Rev Mol Cell Biol*. 2014;15(12):802-812. doi:10.1038/nrm3896.

91. Itzkovitz S, van Oudenaarden A. Validating transcripts with probes and imaging technology. *Nat Methods*. 2011;8(4 Suppl):S12-9. doi:10.1038/nmeth.1573.

92. Jewett JC, Bertozzi CR. Cu-free click cycloaddition reactions in chemical biology. *Chem Soc Rev*. 2010;39(4):1272-1279.

93. Jiang L, Ma A, Song L, et al. Cartilage regeneration by selected chondrogenic clonal mesenchymal stem cells in the collagenase- induced monkey osteoarthritis model. 2013. doi:10.1002/term.

94. Johnson JA, Lu YY, Van Deventer JA, Tirrell DA. Residue-specific incorporation

of non-canonical amino acids into proteins: recent developments and applications. *Curr Opin Chem Biol*. 2010;14(6):774-780. doi:10.1016/j.cbpa.2010.09.013.

95. Johnstone B, Alini M, Cucchiarini M, et al. Tissue engineering for articular cartilage repair--the state of the art. *Eur Cell Mater*. 2013;25:248-267.

96. Junker JP, van Oudenaarden A. Every cell is special: genome-wide studies add a new dimension to single-cell biology. *Cell*. 2014;157(1):8-11. doi:10.1016/j.cell.2014.02.010.

97. Kadler KE, Hill A, Canty-Laird EG. Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators. *Curr Opin Cell Biol*. 2008;20(5):495-501. doi:10.1016/j.ceb.2008.06.008.

98. Kang C-C, Yamauchi KA, Vlassakis J, Sinkala E, Duncombe TA, Herr AE. Single cell–resolution western blotting. *Nat Protoc*. 2016;11(8):1508-1530. doi:10.1038/nprot.2016.089.

99. Kaplan S, Shimizu C. Growth hormone effects on cartilage amino acids composition. *Am J Dis Child*. 1963;105:576-579.

100. Katsara O, Mahaira LG, Iliopoulou EG, et al. Effects of donor age, gender, and in vitro cellular aging on the phenotypic, functional, and molecular characteristics of mouse bone marrow-derived mesenchymal stem cells. *Stem Cells Dev*. 2011;20(9):1549-1561. doi:10.1089/scd.2010.0280.

101. Kawano S, Shoji S, Ichinose S, Yamagata K, Tagami M, Hiraoka M. Characterization of Ca(2+) signaling pathways in human mesenchymal stem cells. *Cell Calcium*. 2002;32(4):165-174.

102. Kern S, Eichler H, Stoeve J, Klüter H, Bieback K. Comparative Analysis of Mesenchymal Stem Cells from Bone Marrow, Umbilical Cord Blood, or Adipose Tissue. *Stem Cells*. 2006;24(5):1294-1301. doi:10.1634/stemcells.2005-0342.

103. Khetan S, Guvendiren M, Legant WR, Cohen DM, Chen CS, Burdick J a. Degradation-mediated cellular traction directs stem cell fate in covalently crosslinked three-dimensional hydrogels. *Nat Mater*. 2013;12(5):458-465. doi:10.1038/nmat3586.

104. Kim T-J, Seong J, Ouyang M, et al. Substrate rigidity regulates Ca2+ oscillation via RhoA pathway in stem cells. *J Cell Physiol*. 2009;218(2):285-293. doi:10.1002/jcp.21598.

105. Kim T-J, Sun J, Lu S, Qi Y-X, Wang Y. Prolonged Mechanical Stretch Initiates Intracellular Calcium Oscillations in Human Mesenchymal Stem Cells. Leipzig ND,

ed. *PLoS One*. 2014;9(10):e109378. doi:10.1371/journal.pone.0109378.

106. King FW, Liszewski W, Ritner C, Bernstein HS. High-Throughput Tracking of Pluripotent Human Embryonic Stem Cells with Dual Fluorescence Resonance Energy Transfer Molecular Beacons. *Stem Cells Dev*. 2011;20(3):475-484. doi:10.1089/scd.2010.0219.

107. Knight MM, Lee DA, Bader DL. The influence of elaborated pericellular matrix on the deformation of isolated articular chondrocytes cultured in agarose. *Biochim Biophys Acta - Mol Cell Res*. 1998;1405(1-3):67-77. doi:10.1016/S0167-4889(98)00102-5.

108. Knudson CB. Hyaluronan receptor-directed assembly of chondrocyte pericellular matrix. *J Cell Biol*. 1993;120(3):825-834. doi:10.1083/jcb.120.3.825.

109. Kock L, van Donkelaar CC, Ito K. Tissue engineering of functional articular cartilage: the current status. *Cell Tissue Res*. 2012;347(3):613-627. doi:10.1007/s00441-011-1243-1.

110. Kretlow JD, Jin Y-Q, Liu W, et al. Donor age and cell passage affects differentiation potential of murine bone marrow-derived stem cells. *BMC Cell Biol*. 2008;9:60. doi:10.1186/1471-2121-9-60.

111. Kumar D, Lassar AB. The Transcriptional Activity of Sox9 in Chondrocytes Is Regulated by RhoA Signaling and Actin Polymerization. *Mol Cell Biol*. 2009;29(15):4262-4273. doi:10.1128/MCB.01779-08.

112. Kumar RM, Cahan P, Shalek AK, et al. Deconstructing transcriptional heterogeneity in pluripotent stem cells. *Nature*. 2014;516(7529):56-61. doi:10.1038/nature13920.

113. Kuznetsov S a, Mankani MH, Bianco P, Robey PG. Enumeration of the colony-forming units-fibroblast from mouse and human bone marrow in normal and pathological conditions. *Stem Cell Res*. 2009;2(1):83-94. doi:10.1016/j.scr.2008.07.007.

114. Kuznetsov SA, Krebsbach PH, Satomura K, et al. Single-colony derived strains of human marrow stromal fibroblasts form bone after transplantation in vivo. *J Bone Miner Res*. 1997;12(9):1335-1347. doi:10.1359/jbmr.1997.12.9.1335.

115. Labriola NR, Darling EM. Temporal heterogeneity in single-cell gene expression and mechanical properties during adipogenic differentiation. *J Biomech*. 2015;48(6):1058-1066. doi:10.1016/j.jbiomech.2015.01.033.

116. Lahm H, Doppler S, Dreßen M, et al. Live Fluorescent RNA-Based Detection of

Pluripotency Gene Expression in Embryonic and Induced Pluripotent Stem Cells of Different Species. *Stem Cells*. 2015;33(2):392-402. doi:10.1002/stem.1872.

117. Larsen KH, Frederiksen CM, Burns JS, Abdallah BM, Kassem M. Identifying A Molecular Phenotype for Bone Marrow Stromal Cells With *In Vivo* Bone Forming Capacity. *J Bone Miner Res*. 2009;25(4):091012153414059-54. doi:10.1359/jbmr.091018.

118. Larsson HM, Lee ST, Roccio M, et al. Sorting Live Stem Cells Based on Sox2 mRNA Expression. Najbauer J, ed. *PLoS One*. 2012;7(11):e49874. doi:10.1371/journal.pone.0049874.

119. Laughlin ST, Baskin JM, Amacher SL, Bertozzi CR. In vivo imaging of membrane-associated glycans in developing zebrafish. *Science*. 2008;320(5876):664-667. doi:10.1126/science.1155106.

120. Laughlin ST, Bertozzi CR. Metabolic labeling of glycans with azido sugars and subsequent glycan-profiling and visualization via Staudinger ligation. *Nat Protoc*. 2007;2(11):2930-2944. doi:10.1038/nprot.2007.422.

121. Lee RH, Hsu SC, Munoz J, et al. A subset of human rapidly self-renewing marrow stromal cells preferentially engraft in mice. *Blood*. 2006;107(5):2153-2161. doi:10.1182/blood-2005-07-2701.

122. Lee WC, Shi H, Poon Z, et al. Multivariate biophysical markers predictive of mesenchymal stromal cell multipotency. *Proc Natl Acad Sci*. 2014;111(42):E4409-E4418. doi:10.1073/pnas.1402306111.

123. Li B, Menzel U, Loebel C, Schmal H, Alini M, Stoddart MJ. Monitoring live human mesenchymal stromal cell differentiation and subsequent selection using fluorescent RNA-based probes. *Sci Rep*. 2016;6(1):26014. doi:10.1038/srep26014.

124. Li CX, Talele NP, Boo S, et al. MicroRNA-21 preserves the fibrotic mechanical memory of mesenchymal stem cells. *Nat Mater*. October 2016. doi:10.1038/nmat4780.

125. Li S, Zhu H, Wang J, et al. Comparative analysis of Cu (I)-catalyzed alkyne-azide cycloaddition (CuAAC) and strain-promoted alkyne-azide cycloaddition (SPAAC) in O-GlcNAc proteomics. *Electrophoresis*. 2016;37(11):1431-1436. doi:10.1002/elps.201500491.

126. Lin Z, Fitzgerald JB, Xu J, et al. Gene expression profiles of human chondrocytes during passaged monolayer cultivation. *J Orthop Res*. 2008;26(9):1230-1237. doi:10.1002/jor.20523.

127. Lorenzo P, Bayliss MT, Heinegård D. Altered patterns and synthesis of extracellular matrix macromolecules in early osteoarthritis. *Matrix Biol*. 2004;23(6):381-391. doi:10.1016/j.matbio.2004.07.007.

128. Lubeck E, Cai L. Single-cell systems biology by super-resolution imaging and combinatorial labeling. *Nat Methods*. 2012;9(7):743-748. doi:10.1038/nmeth.2069.

129. Lubeck E, Coskun AF, Zhiyentayev T, Ahmad M, Cai L. Single-cell in situ RNA profiling by sequential hybridization. *Nat Methods*. 2014;11(4):360-361. doi:10.1038/nmeth.2892.

130. Luo W, Guo C, Zheng J, et al. Aggrecan from start to finish. *J Bone Miner Metab*. 2000;18(2):51-56. doi:10.1007/s007740050011.

131. Luo Y, Sinkeviciute D, He Y, et al. The minor collagens in articular cartilage. *Protein Cell*. February 2017. doi:10.1007/s13238-017-0377-7.

132. Ma B, Leijten JCH, Wu L, et al. Gene expression profiling of dedifferentiated human articular chondrocytes in monolayer culture. *Osteoarthr Cartil*. 2013;21(4):599-603. doi:10.1016/j.joca.2013.01.014.

133. Maamar H, Raj A, Dubnau D. Noise in Gene Expression Determines Cell Fate in Bacillus subtilis. *Science (80-)*. 2007;317(5837):526-529. doi:10.1126/science.1140818.

134. Mackay AM, Beck SC, Murphy JM, Barry FP, Chichester CO, Pittenger MF. Chondrogenic differentiation of cultured human mesenchymal stem cells from marrow. *Tissue Eng*. 1998;4(4):415-428. doi:10.1089/ten.1998.4.415.

135. Mahdavi A, Szychowski J, Ngo JT, et al. Identification of secreted bacterial proteins by noncanonical amino acid tagging. *Proc Natl Acad Sci U S A*. 2014;111(1):433-438. doi:10.1073/pnas.1301740111.

136. Makris EA, Gomoll AH, Malizos KN, Hu JC, Athanasiou KA. Repair and tissue engineering techniques for articular cartilage. *Nat Rev Rheumatol*. 2015;11(1):21-34. doi:10.1038/nrrheum.2014.157.

137. Maloney JM, Nikova D, Lautenschläger F, et al. Mesenchymal stem cell mechanics from the attached to the suspended state. *Biophys J*. 2010;99(8):2479-2487. doi:10.1016/j.bpj.2010.08.052.

138. Marble HD, Sutermaster BA, Kanthilal M, Fonseca VC, Darling EM. Gene expression-based enrichment of live cells from adipose tissue produces subpopulations with improved osteogenic potential. *Stem Cell Res Ther*. 2014;5(5):12. doi:10.1186/scrt502.

139. Mareddy S, Broadbent J, Crawford R, Xiao Y. Proteomic profiling of distinct clonal populations of bone marrow mesenchymal stem cells. *J Cell Biochem*. 2009;106(5):776-786. doi:10.1002/jcb.22088.

140. Mareddy S, Crawford R, Brooke G, Xiao Y. Clonal isolation and characterization of bone marrow stromal cells from patients with osteoarthritis. *Tissue Eng*. 2007;13(4):819-829. doi:10.1089/ten.2006.0180.

141. Mareddy S, Dhaliwal N, Crawford R, Xiao Y. Stem cell-related gene expression in clonal populations of mesenchymal stromal cells from bone marrow. *Tissue Eng Part A*. 2010;16(2):749-758. doi:10.1089/ten.TEA.2009.0307.

142. Matta C, Zakany R. Calcium signalling in chondrogenesis: implications for cartilage repair. *Front Biosci (Schol Ed)*. 2013;5:305-324.

143. Mattar P, Bieback K. Comparing the Immunomodulatory Properties of Bone Marrow, Adipose Tissue, and Birth-Associated Tissue Mesenchymal Stromal Cells. *Front Immunol*. 2015;6:560. doi:10.3389/fimmu.2015.00560.

144. Mauck RL, Yuan X, Tuan RS. Chondrogenic differentiation and functional maturation of bovine mesenchymal stem cells in long-term agarose culture. *Osteoarthr Cartil*. 2006;14(2):179-189. doi:10.1016/j.joca.2005.09.002.

145. McGowan KB, Sah RL. Treatment of cartilage with beta-aminopropionitrile accelerates subsequent collagen maturation and modulates integrative repair. *J Orthop Res*. 2005;23(3):594-601. doi:10.1016/j.orthres.2004.02.015.

146. McLeod CM, Mauck RL. High fidelity visualization of cell-to-cell variation and temporal dynamics in nascent extracellular matrix formation. *Sci Rep*. 2016;6:38852. doi:10.1038/srep38852.

147. Menicanin D, Bartold PM, Zannettino ACW, Gronthos S. Identification of a Common Gene Expression Signature Associated with Immature Clonal Mesenchymal Cell Populations Derived from Bone Marrow and Dental Tissues. *Stem Cells Dev*. 2010;19(10):1501-1510. doi:10.1089/scd.2009.0492.

148. Mindaye ST, Ra M, Lo Surdo JL, Bauer SR, Alterman M. Global proteomic signature of undifferentiated human bone marrow stromal cells: Evidence for donor-to-donor proteome heterogeneity. *Stem Cell Res*. 2013;11(2):793-805. doi:10.1016/j.scr.2013.05.006.

149. Mindaye ST, Lo Surdo J, Bauer SR, Alterman MA. The proteomic dataset for bone marrow derived human mesenchymal stromal cells: Effect of in vitro passaging. *Data Br*. 2015;5:864-870. doi:10.1016/j.dib.2015.10.020.

150. Molina N, Suter DM, Cannavo R, Zoller B, Gotic I, Naef F. Stimulus-induced modulation of transcriptional bursting in a single mammalian gene. *Proc Natl Acad Sci U S A*. 2013;110(51):20563-20568. doi:10.1073/pnas.1312310110.

151. Moskalewski S, Langeveld CH, Scherft JP. Influence of beta-aminoproprionitrile (BAPN) on cell growth and elastic fiber formation in cultures of auricular chondrocytes. *Experientia*. 1983;39(10):1147-1148. doi:10.1007/BF01943153.

152. Mueller MB, Tuan RS. Anabolic/Catabolic balance in pathogenesis of osteoarthritis: identifying molecular targets. *PM R*. 2011;3(6 Suppl 1):S3-11. doi:10.1016/j.pmrj.2011.05.009.

153. Muraglia A, Cancedda R, Quarto R. Clonal mesenchymal progenitors from human bone marrow differentiate in vitro according to a hierarchical model. *J Cell Sci*. 2000;113 (Pt 7:1161-1166.

154. Ng L, Grodzinsky AJ, Patwari P, Sandy J, Plaas A, Ortiz C. Individual cartilage aggrecan macromolecules and their constituent glycosaminoglycans visualized via atomic force microscopy. *J Struct Biol*. 2003;143(3):242-257. doi:10.1016/j.jsb.2003.08.006.

155. Nimmo R, May GE, Enver T. Primed and ready: understanding lineage commitment through single cell analysis. *Trends Cell Biol*. 2015:1-9. doi:10.1016/j.tcb.2015.04.004.

156. Noer A, Boquest AC, Collas P. Dynamics of adipogenic promoter DNA methylation during clonal culture of human adipose stem cells to senescence. *BMC Cell Biol*. 2007;8:18. doi:10.1186/1471-2121-8-18.

157. Noer A, Sørensen AL, Boquest AC, Collas P. Stable CpG Hypomethylation of Adipogenic Promoters in Freshly Isolated , Cultured , and Differentiated Mesenchymal Stem Cells from Adipose Tissue □. 2006;17(August):3543-3556. doi:10.1091/mbc.E06.

158. O'Connell GD, Nims RJ, Green J, Cigan AD, Ateshian GA, Hung CT. Time and dose-dependent effects of chondroitinase ABC on growth of engineered cartilage. *Eur Cell Mater*. 2014;27:312-320.

159. Octavio LM, Gedeon K, Maheshri N. Epigenetic and conventional regulation is distributed among activators of FLO11 allowing tuning of population-level heterogeneity in its expression. Akhtar A, ed. *PLoS Genet*. 2009;5(10):e1000673. doi:10.1371/journal.pgen.1000673.

160. Padovan-Merhar O, Nair GP, Biaesch AG, et al. Single Mammalian Cells Compensate for Differences in Cellular Volume and DNA Copy Number through

Independent Global Transcriptional Mechanisms. *Mol Cell*. 2015;58(2):339-352. doi:10.1016/j.molcel.2015.03.005.

161. Pajerowski J, Dahl K, Zhong F. Physical plasticity of the nucleus in stem cell differentiation. *Proc Natl Acad Sci*. 2007;1.

162. Park S, Lee JY, Hong S, et al. Dual transcript and protein quantification in a massive single cell array. *Lab Chip*. 2016;16(19):3682-3688. doi:10.1039/C6LC00762G.

163. Pelttari K, Steck E, Richter W. The use of mesenchymal stem cells for chondrogenesis. *Injury*. 2008;39 Suppl 1:S58-65. doi:10.1016/j.injury.2008.01.038.

164. Pevsner-Fischer M, Levin S, Zipori D. The origins of mesenchymal stromal cell heterogeneity. *Stem Cell Rev*. 2011;7(3):560-568. doi:10.1007/s12015-011-9229-7.

165. Phinney DG, Kopen G, Righter W, Webster S, Tremain N, Prockop DJ. Donor variation in the growth properties and osteogenic potential of human marrow stromal cells. *J Cell Biochem*. 1999;75(3):424-436.

166. Pina C, Fugazza C, Tipping AJ, et al. Inferring rules of lineage commitment in haematopoiesis. *Nat Cell Biol*. 2012;14(3):287-294. doi:10.1038/ncb2442.

167. Pittenger MF. Multilineage Potential of Adult Human Mesenchymal Stem Cells. *Science (80-)*. 1999;284(5411):143-147. doi:10.1126/science.284.5411.143.

168. Pontikoglou C, Langonné A, Ba MA, et al. CD200 expression in human cultured bone marrow mesenchymal stem cells is induced by pro-osteogenic and pro-inflammatory cues. *J Cell Mol Med*. 2016;20(4):655-665. doi:10.1111/jcmm.12752.

169. Presnyak V, Alhusaini N, Chen Y-H, et al. Codon Optimality Is a Major Determinant of mRNA Stability. *Cell*. 2015;160(6):1111-1124. doi:10.1016/j.cell.2015.02.029.

170. Puetzer JL, Petitte JN, Loboa EG. Comparative review of growth factors for induction of three-dimensional in vitro chondrogenesis in human mesenchymal stem cells isolated from bone marrow and adipose tissue. *Tissue Eng Part B Rev*. 2010;16(4):435-444. doi:10.1089/ten.TEB.2009.0705.

171. Quinn TM, Grodzinsky AJ, Hunziker EB, Sandy JD. Effects of injurious compression on matrix turnover around individual cells in calf articular cartilage explants. *J Orthop Res*. 1998;16(4):490-499. doi:10.1002/jor.1100160415.

172. Quinn TM, Schmid P, Hunziker EB, Grodzinsky AJ. Proteoglycan deposition around chondrocytes in agarose culture: construction of a physical and biological interface for mechanotransduction in cartilage. *Biorheology*. 2002;39(1-2):27-37.

173. Quintana L, zur Nieden NI, Semino CE. Morphogenetic and regulatory mechanisms during developmental chondrogenesis: new paradigms for cartilage tissue engineering. *Tissue Eng Part B Rev*. 2009;15(1):29-41. doi:10.1089/ten.teb.2008.0329.

174. Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods*. 2008;5(10):877-879. doi:10.1038/nmeth.1253.

175. Raj A, van Oudenaarden A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*. 2008;135(2):216-226. doi:10.1016/j.cell.2008.09.050.

176. Raj A, van Oudenaarden A. Single-molecule approaches to stochastic gene expression. *Annu Rev Biophys*. 2009;38:255-270. doi:10.1146/annurev.biophys.37.032807.125928.

177. Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol*. 2006;4(10):e309. doi:10.1371/journal.pbio.0040309.

178. Raj A, Rifkin SA, Andersen E, van Oudenaarden A. Variability in gene expression underlies incomplete penetrance. *Nature*. 2010;463(7283):913-918. doi:10.1038/nature08781.

179. Raj A, Tyagi S. Detection of individual endogenous RNA transcripts in situ using multiple singly labeled probes. *Methods Enzymol*. 2010;472:365-386. doi:10.1016/S0076-6879(10)72004-8.

180. Rennerfeldt DA, Van Vliet KJ. Concise Review: When Colonies Are Not Clones: Evidence and Implications of Intracolony Heterogeneity in Mesenchymal Stem Cells. *Stem Cells*. 2016;34(5):1135-1141. doi:10.1002/stem.2296.

181. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12(1):77. doi:10.1186/1471-2105-12-77.

182. Ross J. mRNA stability in mammalian cells. *Microbiol Rev*. 1995;59(3):423-450.

183. Rostovskaya M, Anastassiadis K. Differential Expression of Surface Markers in Mouse Bone Marrow Mesenchymal Stromal Cell Subpopulations with Distinct

Lineage Commitment. Shi X-M, ed. *PLoS One*. 2012;7(12):e51221. doi:10.1371/journal.pone.0051221.

184. Roughley PJ. Articular cartilage and changes in arthritis: noncollagenous proteins and proteoglycans in the extracellular matrix of cartilage. *Arthritis Res*. 2001;3(6):342-347. doi:10.1186/ar326.

185. Russell KC, Lacey MR, Gilliam JK, Tucker HA, Phinney DG, O'Connor KC. Clonal analysis of the proliferation potential of human bone marrow mesenchymal stem cells as a function of potency. *Biotechnol Bioeng*. 2011;108(11):2716-2726. doi:10.1002/bit.23193.

186. Russell KC, Phinney DG, Lacey MR, Barrilleaux BL, Meyertholen KE, O'Connor KC. In vitro high-capacity assay to quantify the clonal heterogeneity in trilineage potential of mesenchymal stem cells reveals a complex hierarchy of lineage commitment. *Stem Cells*. 2010;28(4):788-798. doi:10.1002/stem.312.

187. Saliba A-E, Westermann AJ, Gorski SA, Vogel J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res*. 2014;42(14):8845-8860. doi:10.1093/nar/gku555.

188. Schellenberg A, Lin Q, Schüler H, et al. Replicative senescence of mesenchymal stem cells causes DNA-methylation changes which correlate with repressive histone marks. *Aging (Albany NY)*. 2011;3(9):873-888.

189. Schiltz JR, Mayne R, Holtzer H. The Synthesis of Collagen and Glycosaminoglycans by Dedifferentiated Chondroblasts in Culture. *Differentiation*. 1973;1(2):97-108. doi:10.1111/j.1432-0436.1973.tb00106.x.

190. Schwanhäusser B, Busse D, Li N, et al. Global quantification of mammalian gene expression control. *Nature*. 2011;473(7347):337-342. doi:10.1038/nature10098.

191. Seferos D, Giljohann D, Hill H, Prigodich A, Mirkin C. Nano-Flares: Probes for Transfection and mRNA Detection in Living Cells. 2007. doi:10.1021/JA0776529.

192. Sekiya I, Larson BL, Vuoristo JT, Cui J-G, Prockop DJ. Adipogenic differentiation of human adult stem cells from bone marrow stroma (MSCs). *J Bone Miner Res*. 2004;19(2):256-264. doi:10.1359/JBMR.0301220.

193. Selich A, Daudert J, Hass R, et al. Massive Clonal Selection and Transiently Contributing Clones During Expansion of Mesenchymal Stem Cell Cultures Revealed by Lentiviral RGB-Barcode Technology. *Stem Cells Transl Med*. 2016;5(5):591-601. doi:10.5966/sctm.2015-0176.

194. Sengers BG, Van Donkelaar CC, Oomens CWJ, Baaijens FPT. The local matrix

distribution and the functional development of tissue engineered cartilage, a finite element study. *Ann Biomed Eng*. 2004;32(12):1718-1727.

195. Shah K, Tyagi S. Barriers to transmission of transcriptional noise in a c-fos c-jun pathway. *Mol Syst Biol*. 2014;9(1):687-687. doi:10.1038/msb.2013.45.

196. Sharova L V, Sharov AA, Nedorezov T, Piao Y, Shaik N, Ko MSH. Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. *DNA Res*. 2009;16(1):45-58. doi:10.1093/dnares/dsn030.

197. Shav-Tal Y, Darzacq X, Shenoy SM, et al. Dynamics of single mRNPs in nuclei of living cells. *Science*. 2004;304(5678):1797-1800. doi:10.1126/science.1099754.

198. Silberstein L, Goncalves KA, Kharchenko PV, et al. Proximity-Based Differential Single-Cell Analysis of the Niche to Identify Stem/Progenitor Cell Regulators. *Cell Stem Cell*. 2016;19(4):530-543. doi:10.1016/j.stem.2016.07.004.

199. Sivasubramaniyan K, Lehnen D, Ghazanfari R, et al. Phenotypic and functional heterogeneity of human bone marrow- and amnion-derived MSC subsets. *Ann N Y Acad Sci*. 2012;1266(1):94-106. doi:10.1111/j.1749-6632.2012.06551.x.

200. Smith JR, Pochampally R, Perry A, Hsu S-C, Prockop DJ. Isolation of a Highly Clonogenic and Multipotential Subfraction of Adult Stem Cells from Bone Marrow Stroma. *Stem Cells*. 2004;22(5):823-831. doi:10.1634/stemcells.22-5-823.

201. Sophia Fox AJ, Bedi A, Rodeo SA. The basic science of articular cartilage: structure, composition, and function. *Sports Health*. 2009;1(6):461-468. doi:10.1177/1941738109350438.

202. Spitzer MH, Nolan GP, Goltsev Y, et al. Mass Cytometry: Single Cells, Many Features. *Cell*. 2016;165(4):780-791. doi:10.1016/j.cell.2016.04.019.

203. Ståhlberg A, Bengtsson M. Single-cell gene expression profiling using reverse transcription quantitative real-time PCR. *Methods*. 2010;50:282-288. doi:10.1016/j.ymeth.2010.01.002.

204. Stenderup K, Justesen J, Clausen C, Kassem M. Aging is associated with decreased maximal life span and accelerated senescence of bone marrow stromal cells. *Bone*. 2003;33(6):919-926.

205. Stokes DG, Liu G, Dharmavaram R, Hawkins D, Piera-Velazquez S, Jimenez S a. Regulation of type-II collagen gene expression during human chondrocyte de-differentiation and recovery of chondrocyte-specific phenotype in culture involves Sry-type high-mobility-group box (SOX) transcription factors. *Biochem J*.

2001;360(Pt 2):461-470. doi:10.1042/0264-6021:3600461.

206. Strioga M, Viswanathan S, Darinskas A, Slaby O, Michalek J. Same or not the same? Comparison of adipose tissue-derived versus bone marrow-derived mesenchymal stem and stromal cells. *Stem Cells Dev*. 2012;21(14):2724-2752. doi:10.1089/scd.2011.0722.

207. Süel GM, Kulkarni RP, Dworkin J, Garcia-Ojalvo J, Elowitz MB. Tunability and noise dependence in differentiation dynamics. *Science*. 2007;315(5819):1716-1719. doi:10.1126/science.1137455.

208. Sun S, Liu Y, Lipsky S, Cho M. Physical manipulation of calcium oscillations facilitates osteodifferentiation of human mesenchymal stem cells. *FASEB J*. 2007;21(7):1472-1480. doi:10.1096/fj.06-7153com.

209. Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. Mammalian genes are transcribed with widely different bursting kinetics. *Science (80-)*. 2011;332(6028):472-474. doi:10.1126/science.1198817.

210. Sworder BJ, Yoshizawa S, Mishra PJ, et al. Molecular profile of clonal strains of human skeletal stem/progenitor cells with different potencies. *Stem Cell Res*. 2015;14(3):297-306. doi:10.1016/j.scr.2015.02.005.

211. Tajik A, Zhang Y, Wei F, et al. Transcription upregulation via force-induced direct stretching of chromatin. *Nat Mater*. 2016;15(12):1287-1296. doi:10.1038/nmat4729.

212. Taniguchi Y, Choi PJ, Li G-W, et al. Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science*. 2010;329(5991):533-538. doi:10.1126/science.1188308.

213. Tew S, McDermott B, Fentem R, Peffers M, Clegg P. Transcriptome-Wide Analysis of Messenger RNA Decay in Normal and Osteoarthritic Human Articular Chondrocytes. *Arthritis Rheumatol*. 2014;66(11):3052-3061.

214. Thattai M, van Oudenaarden A. Intrinsic noise in gene regulatory networks. *Proc Natl Acad Sci U S A*. 2001;98(15):8614-8619. doi:10.1073/pnas.151588598.

215. tom Dieck S, Muller A, Nehring A, et al. Metabolic labeling with noncanonical amino acids and visualization by chemoselective fluorescent tagging. *Curr Protoc Cell Biol*. 2012. doi:10.1002/0471143030.cb0711s56.

216. Tseng Q, Duchemin-Pelletier E, Deshiere A, et al. Spatial organization of the extracellular matrix regulates cell-cell junction positioning. *Proc Natl Acad Sci U S A*. 2012;109(5):1506-1511. doi:10.1073/pnas.1106377109.

217. Tuan RS, Boland G, Tuli R, et al. Adult mesenchymal stem cells and cell-based tissue engineering. *Arthritis Res Ther*. 2003;5(1):32. doi:10.1186/ar614.

218. Tyagi S, Kramer FR. Molecular Beacons: Probes that Fluoresce upon Hybridization. *Nat Biotechnol*. 1996;14(3):303-308. doi:10.1038/nbt0396-303.

219. UniProt Consortium. UniProt: a hub for protein information. *Nucleic Acids Res*. 2015;43(Database issue):D204-12. doi:10.1093/nar/gku989.

220. Uttamapinant C, Tangpeerachaikul A, Grecian S, et al. Fast, cell-compatible click chemistry with copper-chelating azides for biomolecular labeling. *Angew Chem Int Ed Engl*. 2012;51(24):5852-5856. doi:10.1002/anie.201108181.

221. Vertel BM. The ins and outs of aggrecan. *Trends Cell Biol*. 1995;5(12):458-464. doi:10.1016/S0962-8924(00)89115-1.

222. Vigfúsdóttir ÁT, Pasrija C, Thakore PI, Schmidt RB, Hsieh AH. Role of Pericellular Matrix in Mesenchymal Stem Cell Deformation during Chondrogenic Differentiation. *Cell Mol Bioeng*. 2010;3(4):387-397. doi:10.1007/s12195-010-0135-x.

223. Vogel C, Marcotte EM. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet*. 2012;13(4):227-232. doi:10.1038/nrg3185.

224. Wang J, Liao L, Wang S, Tan J. Cell therapy with autologous mesenchymal stem cells-how the disease process impacts clinical considerations. *Cytotherapy*. 2013;15(8):893-904. doi:10.1016/j.jcyt.2013.01.218.

225. Wernet MF, Mazzoni EO, Celik A, Duncan DM, Duncan I, Desplan C. Stochastic spineless expression creates the retinal mosaic for colour vision. *Nature*. 2006;440(7081):174-180. doi:10.1038/nature04615.

226. Whitfield MJ, Lee WCJ, Van Vliet KJ. Onset of heterogeneity in culture-expanded bone marrow stromal cells. *Stem Cell Res*. 2013;11(3):1365-1377. doi:10.1016/j.scr.2013.09.004.

227. Wilusz RE, Sanchez-adams J, Guilak F. The structure and function of the pericellular matrix of articular cartilage. *Matrix Biol*. 2014;39:25-32. doi:10.1016/j.matbio.2014.08.009.

228. Wilusz RE, Zauscher S, Guilak F. Micromechanical mapping of early osteoarthritic changes in the pericellular matrix of human articular cartilage. *Osteoarthritis Cartilage*. 2013;21(12):1895-1903. doi:10.1016/j.joca.2013.08.026.

229. Wu YJ, La Pierre DP, Wu J, Yee AJ, Yang BB. The interaction of versican with its binding partners. *Cell Res*. 2005;15(7):483-494. doi:10.1038/sj.cr.7290318.

230. Yang C, Tibbitt MW, Basta L, Anseth KS. Mechanical memory and dosing influence stem cell fate. 2014;(March):1-8. doi:10.1038/NMAT3889.

231. Yang MT, Fu J, Wang Y-K, Desai RA, Chen CS. Assaying stem cell mechanobiology on microfabricated elastomeric substrates with geometrically modulated rigidity. *Nat Protoc*. 2011;6(2):187-213. doi:10.1038/nprot.2010.189.

232. Yang R, Chen M, Lee CH, Yoon R, Lal S, Mao JJ. Clones of ectopic stem cells in the regeneration of muscle defects in vivo. *PLoS One*. 2010;5(10):e13547. doi:10.1371/journal.pone.0013547.

233. Ylöstalo J, Bazhanov N, Prockop DJ. Reversible commitment to differentiation by human multipotent stromal cells in single-cell-derived colonies. *Exp Hematol*. 2008;36(10):1390-1402. doi:10.1016/j.exphem.2008.05.003.

234. Yoon BC, Jung H, Dwivedy A, O'Hare CM, Zivraj KH, Holt CE. Local translation of extranuclear lamin B promotes axon maintenance. *Cell*. 2012;148(4):752-764. doi:10.1016/j.cell.2011.11.064.

235. Yoon HI, Yhee JY, Na JH, et al. Bioorthogonal Copper Free Click Chemistry for Labeling and Tracking of Chondrocytes In Vivo. *Bioconjug Chem*. 2016;27(4):927-936. doi:10.1021/acs.bioconjchem.6b00010.

236. Zenklusen D, Larson DR, Singer RH. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat Struct Mol Biol*. 2008;15(12):1263-1271. doi:10.1038/nsmb.1514.

237. Zopf CJ, Quinn K, Zeidman J, Maheshri N. Cell-cycle dependence of transcription dominates noise in gene expression. Kondev J, ed. *PLoS Comput Biol*. 2013;9(7):e1003161. doi:10.1371/journal.pcbi.1003161.