From molecules to organs : Microscopy and multi-scale nature of development

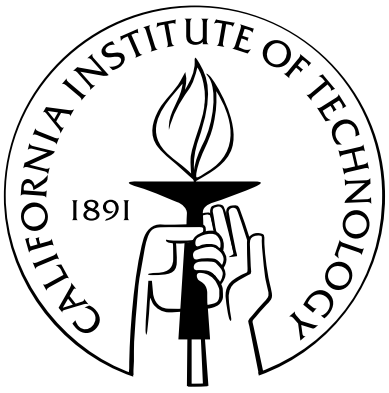
Thesis by

Vikas Trivedi

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2015

(Defended May 19, 2015)

Oats, peas, beans and barley grow,

Oats, peas, beans and barley grow,

Can you, or I, or anyone know

How oats, peas, beans and barley grow ?

- Old English Nursery Rhyme

Preface

The process of working towards and finally writing this thesis has definitely been very heart-warming and rewarding. The work presented here is an amalgamation of various projects that I have enjoyed and contributed to during my graduate career. This document draws on several manuscripts (mentioned below) that I have worked on and have either been published or submitted or are on the way to submission. Immense help has been generously extended by people along the way, especially Dr. Harry Choi for the work in Chapter 2, Dr. Yuwei Li for Chapter 4, Dr. Frederique Ruf-Zamojski for Chapter 5, Dr. Le A. Trinh for Chapter 6, Dr. Michael Liebling for Chapter 7 and Dr. Thai V. Truong for Chapter 8. I do feel that my words are inadequate to summarize my mental state while presenting this thesis to the scientific community, both present and future. Hence I shall resort to the past, borrowing from what Antonie van Leeuwenhoek wrote in a letter in 1716, coincidentally on the same day as commencement this year;*'my work, which I've done for a long time, was... pursued from a craving after knowledge... And therewithal, whenever I found out anything remarkable, I have thought it my duty to put down my discovery on paper, so that all ingenious people might be informed thereof.'*

Chapter 1

Trivedi, V. , Fraser S.E. *Imaging tools for the in vivo study of cell behavior during development. How imaging helps to understand the multi-scale nature of development* (to be submitted)

Chapters 2 and 3

Trivedi, V. , Choi, H.M.T., Fraser S.E., Pierce N.A. *Quantitative sub-cellular mapping of*

mRNA expression within intact vertebrate embryos. (to be submitted)

Chapter 4

Li,Y.*, Trivedi, V.* , Truong,T.V., Koos,D.S., Lansford,R., Chuong,C.M., Warburton,D., Moats,R.A., Fraser S.E. *Dynamic imaging of the growth plate cartilage reveals multiple contributors to skeletal morphogenesis.* Nature Communications (2015) 6:6798 doi:10.1038/ncomms7798

Chapter 5

Ruf-Zamojski,F.*, Trivedi, V.* , Fraser,S.E., Trinh,L.A., *Spatio-temporal differences in dystrophin dynamics at mRNA and protein levels revealed by a novel FlipTrap line.* PLOS ONE (2015), in press

Chapter 6

Trivedi, V. , Trinh, L.A., Fraser S.E. *Decoupling the role of contractility on cardiomyocyte growth in vivo.* (to be submitted)

Chapter 7

Trivedi, V. , Truong,T.V., Trinh,L.A., Holland D.B., Liebling, M., Fraser,S.E. *Dynamic structure and protein expression of the live embryonic heart captured by 2-photon light sheet microscopy and retrospective registration.* Biomed. Opt. Express (2015) 6, 2056-2066

Chapter 8

Trivedi, V. , Truong, T.V., Holland D.B., Trinh,L.A., Fraser,S.E. *Three-dimensional imaging of cellular shape changes and protein expression in the developing zebrafish heart.* Biophysical Journal (2014) 106 (2), 435a-436a

* equal contribution

Acknowledgments

As I reflect upon the past several years of my life as a graduate student at Caltech, I feel overwhelmed with emotions, especially when thinking of people who have made these years so memorable and whose association I shall cherish throughout my life. Caltech has been a very special place to me where I came first as a SURF student in 2009, for a brief period of 10 weeks in the lab of Dr. Marianne Bronner, who was a wonderful mentor. Very soon it was evident that the place had so much to offer that there could not be any doubt while choosing it as the destination for graduate studies.

I believe that among the students in the past, almost all superlatives must have been exhausted to describe the support, patience, inspiration, encouragement, and independence that both Scott and Niles, my advisors, have provided over the years to their mentees. But what has struck me, as the most extraordinary aspect of their role as mentors, is the ways I have observed them act, in a wide range of circumstances, as individuals who can make decisions of responsibility in the most human way. Their approach to situations, both academic and otherwise, has always come across, at least to me, as the line of thought that can be pursued with focused mind and committed heart. It is this simplicity of theirs, which has nurtured my confidence to be an independent researcher of my own right, a dream I started my graduate school with. I cannot thank Scott enough for bringing together a wonderful group of scientists in 'Fraserland', having perhaps the most diverse set of expertise I have ever seen in a single lab. His curiosity, his breadths of knowledge, patience and above all his jokes, have all made me feel at home through these years. I simply could not have been happier anywhere else and cannot ask for more. Niles'focused approach to science, along with the attention to details, has always made me appreciate the efforts needed for perfection. Long discussions with him, both in lab meetings as well as in person, have

always left me satisfied with the quality and quantity of inference that can be drawn from an experimental observation.

I am very thankful to my committee members, Mory, Ravi and Rob. Mory's excitement to image and understand heart has been very refreshing. He has been very generous to have me in his lab meetings and interact with lab members with regards to imaging. His insightful observations of cell motion in the movies of beating heart, that I have shown him over years, have kept my own enthusiasm at the top. Ravi has always motivated me to think about the biomechanical implications of my work. As my committee chair, discussions with him, about both grad school and life beyond grad school, have been very refreshing. Finally I am immensely thankful to Rob, not just for his comments on my work but also (and especially) for the courses he has taught, particularly the 'bioengineering boot camp'for incoming graduate students. This was a course that indeed changed my perspective on the processes in biology and this is something I shall always cherish and take with me as a scientist. I am also very grateful to Dr. Oscar Bruno and Dr. John Brady for all the interactions I have had with them.

Yuwei, Masahiro, Frederique, Carol, Francesco, Mimmo, Jason, Aura, Marcela, Daniel, Dan, Bill, Bertram, Ben, Marcus and Josh have made Fraser lab very lively. Particularly I am grateful to Le and Thai for their mentorship during my stay in the lab. Discussions as well as the fish room cleaning sessions with Le are hard to forget; especially the food she brought at troubled times. Thai has been the best *'partner in crimes'* (a phrase he himself mentioned) and a very dear friend. Everyone knows his love for devil as he always enjoys being the *'devil's advocate'* to emphasize that *'the devil is in the details'*! Harry and Aneesh have made working in Pierce lab the most enjoyable experience. Their friendship is invaluable. Harry's love for orderliness and Aneesh's crush on Apple products were constant source of enjoyment in the office.

I have also been very fortunate to enjoy friendship with so many at Caltech. Ankur and Abbie have always welcomed me at their home and Meemaw, Sir Xy, Katie and Elysse have been a family to me away from India. The friendship developed with Ninja, Bhargavi, Teja, Maha and Sri (*the Youtttth Group*) is an inseparable part of my life. Sujeet, Pinaky,

and Nisha (*the Winchell Group*); Subbu, Utkarsh, Gaurav, Vinamra, Panji, Paraj, and Srimoyee, I am glad to have these friends. I cannot forget the enjoyable rides in LA metro during my trips to USC; they were instrumental in introducing me to LA, to mention the least.

Lauren, Georgiann, Pat and Melinda in the labs, Linda in Bioengineering, Candace and Carol in SURF office, Laura in ISP, Lailin and Helen in Graduate studies office, Tess in Registrar's Office, Hakno from Caltech hiking club and Athena in Caltech Y have all made my stay in Caltech a memorable one. Felicia and Natalie have helped me a lot at different points in my grad school, especially during the transition to USC. Earnie's food has definitely been one constant source of happiness all throughout.

My parents, Purnima Trivedi and Dharmendra Nath Trivedi, have not only encouraged me to dream but have always supported my decisions to pursue them. My sister, Khushboo, has always defended me in all the mischief, even though when mistakes were mine. 'Thanks'is not the word that can convey my gratitude to them.

This thesis simply could not have been possible had I not been to an EMBO conference in Heidelberg. And the reason is not scientific but the coincidence of meeting Gopi, my fiance. Her patience and trust in me have strengthened me throughout the years. For all the love and care, for all the confidence in me, for all the Gujju advice she has given to this Bihari, for all the planning and taking away my trouble/tensions, for understanding me, I am privileged that you are there Gopi !

Finally I am thankful to the funding agencies, especially Rosen fellowship that supported me in my second year and also allowed me to visit Heidelberg! I am indebted to the education systems in both US and India that are responsible for what I am today. To all my teachers who encouraged the spirit of enquiry, I shall always be grateful.

Better than a thousand years lived,

With an unsettled [mind].

Devoid of insight,

Is one day lived. With insight...

Dhammapada (111)

Abstract

Morphogenesis is a phenomenon of intricate balance and dynamic interplay between processes occurring at a wide range of scales (spatial, temporal and energetic). During development, a variety of physical mechanisms are employed by tissues to simultaneously pattern, move, and differentiate based on information exchange between constituent cells, perhaps more than at any other time during an organism's life. To fully understand such events, a combined theoretical and experimental framework is required to assist in deciphering the correlations at both structural and functional levels at scales that include the intracellular and tissue levels as well as organs and organ systems. Microscopy, especially diffraction-limited light microscopy, has emerged as a central tool to capture the spatio-temporal context of life processes. Imaging has the unique advantage of watching biological events as they unfold over time at single-cell resolution in the intact animal. In this work I present a range of problems in morphogenesis, each unique in its requirements for novel quantitative imaging both in terms of the technique and analysis. Understanding the molecular basis for a developmental process involves investigating how genes and their products—mRNA and proteins—function in the context of a cell. Structural information holds the key to insights into mechanisms and imaging fixed specimens paves the first step towards deciphering gene function. The work presented in this thesis starts with the demonstration that the fluorescent signal from the challenging environment of whole-mount imaging, obtained by *in situ* hybridization chain reaction (HCR), scales linearly with the number of copies of target mRNA to provide quantitative sub-cellular mapping of mRNA expression within intact vertebrate embryos. The work then progresses to address aspects of imaging live embryonic development in a number of species. While processes such as avian cartilage growth require high spatial resolution and lower time resolution, dynamic events during

zebrafish somitogenesis require higher time resolution to capture the protein localization as the somites mature. The requirements on imaging are even more stringent in case of the embryonic zebrafish heart that beats with a frequency of $\sim$ 2-2.5 Hz, thereby requiring very fast imaging techniques based on two-photon light sheet microscope to capture its dynamics. In each of the hitherto-mentioned cases, ranging from the level of molecules to organs, an imaging framework is developed, both in terms of technique and analysis to allow quantitative assessment of the process *in vivo*. Overall the work presented in this thesis combines new quantitative tools with novel microscopy for the precise understanding of processes in embryonic development.

Contents

Preface **iv**

Acknowledgments **vi**

Abstract **ix**

List of Figures **xix**

List of Tables **xx**

1 Introduction **1**

1.1 The multi-scale nature of embryonic development 1

1.2 Microscopy and the foundations of modern biology 3

1.3 Imaging fixed specimens to understand genetic machinery *in situ* 6

1.4 Enriching spatial information with low time resolution to learn mechanisms of slow growth processes 7

1.5 Improving time resolution to capture events occurring over hours 8

1.6 Bridging multiple time-scales in growth processes 9

1.7 Pushing the speed of fast dynamic imaging to understand rapid mechanical time scales 10

1.8 References 12

2 Quantitative sub-cellular mapping of mRNA expression within intact vertebrate embryos **17**

2.1 Motivation 17

2.2 Experimental design 20
2.3 Relative quantification by redundant detection 23
2.4 Distinguishing homozygous and heterozygous embryos with HCR 24
2.5 Validating the slope estimates 25
2.6 Somitogenesis and genetic expression 28
2.7 Discussion 28
2.8 References 31

3 Fitting a line through data with *measurement errors* and *equation errors* in both variables 34

3.1 Motivation 34
3.2 Approach 35
3.3 Independent *measurement errors* with non-zero means in both axes 38
3.4 Correlated *measurement errors* (but independent of signal) with non-zero means in both axes 43
3.5 *Measurement errors* in both axes correlated with each other and with signal 46
3.6 Independent *measurement errors* with *equation errors* in both axes 49
3.7 Correlated *measurement errors* (but independent of signal) with *equation errors* in both axes 53
3.8 *Measurement errors* correlated with each other and the signal with *equation errors* in both axes 55
3.9 Slope estimation in cases where data deviates from normal distributions . . 59
3.10 References 62

4 Dynamic imaging of the growth plate cartilage reveals multiple contributors to skeletal morphogenesis 63

4.1 Motivation 63
4.2 4D imaging of cartilage growth using avian metacarpal culture 64
4.3 Anisotropic spreading of the PZ cells 66
4.4 Convergent-extension and cell division cannot account for cell spreading . . 68

4.5 Conservation of cell spreading behavior across avian embryos 70
4.6 ECM expansion and cell volume enlargement cause cell spreading 70
4.7 Discussion 72
4.8 References 77

5 Spatio-temporal differences in dystrophin dynamics at mRNA and protein levels 80
5.1 Motivation 80
5.2 Temporal differences in the localization of dmd transcript and Dmd protein during somite development 82
5.3 Dynamics of Dmd protein localization during somitogenesis 85
5.4 Discussion 87
5.5 References 89

6 Decoupling the role of contractility on cardiomyocyte growth *in vivo* 92
6.1 Motivation 92
6.2 Cell shape changes in cardiomyocytes over development 95
6.3 Heterogeneity in size of cardiomyocytes over development 96
6.4 Cardiomyocytes' shape and size correlate with the tissue curvature 98
6.5 Cardiac mutants with contractile defects indicate independence of cellular size and contraction 100
6.6 Cardiac looping is an effect of growth-induced buckling due to external spatial packing constraints 102
6.7 Discussion 102
6.8 References 106

7 Live 4D imaging of the embryonic zebrafish heart with 2-photon light sheet microscopy and retrospective registration 111
7.1 Motivation 111
7.2 2-photon light sheet microscope setup 114
7.3 Imaging setup characterization 115

7.4 Image reconstruction algorithm 117
7.5 4D visualization of cardiac tissue *in vivo* 118
7.6 Discussion 121
7.7 References 123

8 Macroscopic phase stamping (MaPS) for high resolution reconstruction of beating heart 126

8.1 Motivation 126
8.2 Microscope design 128
8.3 Image processing 129
8.4 Multi-color imaging of beating heart 131
8.5 Quantitative measurements of cardiomyocyte contraction *in vivo* 133
8.6 Discussion 135
8.7 References 138

9 Outlook 140

A Supplementary Information for Chapter 1 145

B Supplementary Information for Chapter 2 148

B.1 Zebrafish lines 148
B.2 Sample preparation for confocal imaging 148
B.3 Protocols 149
B.4 Probe penetration studies 152
B.5 Standard error in slope estimates and correlation between signals 154
B.6 Cooperativity between probes and hairpin polymers 155
B.7 Histograms of signal and background for data presented in Chapter 2 . . . 156
B.8 Replicates for data presented in Chapter 2 158
B.9 Imaging parameters on the microscope 162
B.10 HCR probe sequences 163

C Supplementary Information for Chapter 4 **167**

C.1 Sample preparation 167

C.2 Image analysis 167

C.3 Image based voxel analysis 168

D Supplementary Information for Chapter 5 **170**

D.1 Zebrafish lines 170

D.2 Sample preparation for live imaging 170

D.3 Protocols for *in situ* hybridizations 170

D.4 HCR probe sequences 172

E Supplementary Information for Chapter 6 **174**

E.1 Zebrafish lines 174

E.2 Sample preparation for live imaging 174

E.3 Imaging 175

E.4 Image analysis 175

E.5 Morpholino injections 176

E.6 Imaging parameters for presented data 177

F Supplementary Information for Chapter 7 **179**

F.1 Zebrafish lines 179

F.2 Sample preparation for live imaging 179

F.3 Imaging parameters for presented data 180

G Supplementary Information for Chapter 8 **181**

G.1 Zebrafish lines and sample preparation 181

G.2 Image processing 181

G.3 Synchronization of the cameras 186

G.4 Imaging parameters for presented data 187

G.5 MaPS for assessing health of heart 188

List of Figures

1.1 **From molecules to organs: Imaging and multi-scale nature of development** 5

2.1 **HCR mechanism** 18
2.2 **Relative quantification by redundant detection** 22
2.3 **Distinguishing homozygous and heterozygous embryos with HCR** . 26
2.4 **Validating the slope estimates using redundant detection with 3 colors** 27
2.5 **Using HCR to study somitogenesis quantitatively** 29

3.1 **Distribution of autofluorescence and signal for sub-cellular voxels** . 36
3.2 **Slope estimates for independent measurement errors** ($\sigma_{eu} = 0$) **in both axes** 42
3.3 **Fractional error in the estimated value of slope for independent measurement errors** ($\sigma_{eu} = 0$) **in both axes** 42
3.4 **Slope estimates for correlated measurement errors** ($\sigma_{eu} \neq 0$) **in both axes** 45
3.5 **Fractional error in the estimated value of slope for correlated *measurement errors*** ($\sigma_{eu} \neq 0$) **in both axes** 45
3.6 **Comparison of correlations** ($\sigma_{eu}, \sigma_{Xe}, \sigma_{Yu}$ **and** σ_{XY}) 48
3.7 **Slope estimates for measurement errors** ($\sigma_{eu} \neq 0$) **correlated with signal in both axes** ($\sigma_{Xe}, \sigma_{Yu} \neq 0$) 48
3.8 **Fractional error in the estimated value of slope for measurement errors** ($\sigma_{eu} \neq 0$) **correlated with signal in both axes** ($\sigma_{Xe}, \sigma_{Yu} \neq 0$) . . 49

3.9 **Slope estimates for independent measurement errors ($\sigma_{eu} = 0$) and heteroscedastic *equation errors* (p_i, q_i) in both axes** 51
3.10 **Slope estimates for correlated measurement errors ($\sigma_{eu} \neq 0$) and heteroscedastic *equation errors* (p_i, q_i) in both axes** 54
3.11 **Slope estimates for measurement errors ($\sigma_{eu} \neq 0$) correlated with signal ($\sigma_{Xe}, \sigma_{Yu} \neq 0$) and heteroscedastic *equation errors* (p_i, q_i) in both axes** 57
3.12 **Fractional error in the estimated value of slope for measurement errors ($\sigma_{eu} \neq 0$) correlated with signal ($\sigma_{Xe}, \sigma_{Yu} \neq 0$) and heteroscedastic *equation errors* (p_i, q_i) in both axes** 59
3.13 **Standard error in slope estimates for log-normal signal distribution** 60
3.14 **Slope estimates for log-normal signal distribution** 61

4.1 **Live imaging of cartilage growth** 65
4.2 **Cells undergo collective spreading displacement during cartilage growth** 67
4.3 **Convergent-extension and mitotic division cannot account for cell spreading** 69
4.4 **Quail PZ cells exhibit similar spreading displacement and orthogonal division** 71
4.5 **ECM expansion and cell volume enlargement account for cell trajectories** 73
4.6 **Quantitative model of embryonic cartilage elongation** 75

5.1 **Dmd-Citrine expression in zebrafish FlipTrap line** 81
5.2 **Differential expression of dmd-citrine in homozygous and heterozygous embryos via *in situ* HCR** 83
5.3 **Differential spatial expression of dmd mRNA and Dmd protein** . . 84
5.4 **Dmd Protein expression over development** 86

6.1 **Cardiac development in zebrafish** 93

6.2 **Cardiomyocytes alter their shape and become multi-nucleated over development** . . . 96
6.3 **Cardiomyocytes show heterogeneity in size over development** . . . 97
6.4 **Cardiac mutants with contractile defects show similarity in cellular growth to wild type embryos** . . . 99
6.5 **Cardiac looping is an effect of growth-induced buckling due to external spatial packing constraints** . . . 101

7.1 **Retrospective registration of beating heart with 2-photon light sheet microscope** . . . 112
7.2 **Live 4D imaging of the embryonic heart with 2p-SPIM** . . . 113
7.3 **Characterization of imaging performance of 1p-SPIM and 2p-SPIM** 116
7.4 **Results of 4D capture of the beating heart** . . . 119
7.5 **Trabeculating myocardium during cardiac beating cycle** . . . 120

8.1 **Overview of Macroscopic Phase Stamping (MaPS) for imaging beating heart** . . . 128
8.2 **Overview of image processing involved in MaPS** . . . 130
8.3 **4D reconstruction of beating myocardium and blood cells in 4.5 day old zebrafish embryo** . . . 132
8.4 **Quantitative measurements of cardiomyocyte contraction** . . . 134

B.1 **Probe penetration studies** . . . 152
B.2 **Replicates for probe penetration study in Fig. B.1** . . . 153
B.3 **Relationship between standard error in slope estimates and correlation between signals** . . . 154
B.4 **Cooperativity between probes and hairpin polymers** . . . 155
B.5 **Histograms of signal and background for *desmin* in Fig. 2.2** . . . 156
B.6 **Histograms of signal and background for *citrine* in Fig. 2.2** . . . 156
B.7 **Histograms of signal and background for *desmin* and *citrine* in Fig. 2.3** . . . 157

B.8 **Histograms of signal and background for *desmin* in Fig. 2.4** 157
B.9 **Replicates for redundant detection of *desmin* in Fig. 2.2** 158
B.10 **Replicates for redundant detection of *citrine* in Fig. 2.2** 159
B.11 **Replicates for homozygous-heterozygous study in Fig. 2.3** 160
B.12 **Replicates for slope multiplication study in Fig. 2.4** 161
B.13 **Limits on detector gain and laser power during imaging** 162

C.1 **Live imaging of cartilage growth** 168

E.1 **Widefield images of zebrafish embryos at 60hpf** 175
E.2 **Orthogonal views of 3D rendered confocal images of zebrafish heart** 176
E.3 **Comparison of cell size in 2D projections vs 3D renderings** 178
E.4 **Cardiac mutants with contractile defects show similar heterogeneity in cell shape as wild type embryos** 178

G.1 **Schematic of the 2p-SPIM optical setup used in MaPS** 181
G.2 **Effect of laser power on heart beat** 182
G.3 **Performance of MaPS** 183
G.4 **4D reconstruction of beating myocardium and blood cells in 5.5 days old zebrafish embryo** 185
G.5 **MaPS for assessing health of heart** 186

List of Tables

A.1 Spatio-temporal scales in growth and developmental processes 145

B.1 Imaging parameters for data presented in Chapter 2 162

E.1 Imaging parameters for data presented in Chapter 6 177

F.1 Imaging parameters for data presented in Chapter 7 180

G.1 Imaging parameters for data presented in Chapter 8 187

Chapter 1

Introduction

1.1 The multi-scale nature of embryonic development

Life is a phenomenon of intricate balance and dynamic interplay between processes occurring at a wide range of scales (spatial, temporal and energetic). Regarding life as a complex system composed of relatively simple but mutually interacting parts is an idea that has gained strength over the years. This has resulted from the ability of modern science to carefully separate parts from the whole [1]. As a result, studying life and its processes has attracted brilliant minds from a very wide range of disciplines, both fundamental and applied. It is no surprise that the foundations of such a bottom-up viewpoint of life were laid in the works of early organic chemists, whose efforts, as early as in the first half of 18th century, slowly abolished the widely held notion of vitalism [2]. Whether or not Friedrich Wöhler must be credited as instrumental towards this is a matter of historical debate [3], but it is beyond doubt that his synthesis of urea in 1828 proved that it is the complexity that differentiates the reactions within living organisms from those in inorganic processes [2,4]. At the molecular level there are no fundamental differences between the two systems.

Complexity in living systems arises primarily due to the nonlinear nature of the interactions between the different events that span different scales of length and time, and are often intertwined with several feedback loops and contingency pathways [5]. From the simple governing rules, a rich collection of dynamic properties and patterns emanate, which are distinct from the individual elements, thereby leading to a variety of emergent behaviors [6]. For example the molecular machinery within the cell, comprised of a sophisticated potage

of gene products, proteins and RNA, operates at time scales ranging from tens of minutes for production of proteins down to several seconds for their modification and diffusion. Interestingly the interplay of these many factors across species results in a spectrum of cell cycle times, ranging from several tens of minutes to tens of hours [7]. Such characteristic time scales, ranging from metabolic reaction times to the organism's lifetime, are linked to the differences in the length scales (both spatial and genomic) over which these processes are operating within different organisms (Table A.1).

Development and morphogenesis are the particular focus of this review, during which the formation of tissues or organs is robustly coordinated and repeated in each individual of a species [8]. During organogenesis, a wide range of physical mechanisms are employed by tissues to simultaneously pattern, move and differentiate based on information exchange between constituent cells, perhaps more than at any other time during an organism's life [9,10]. To fully understand such events, a combined theoretical and experimental framework is required to assist in deciphering the correlations at both structural and functional levels at scales that include the intracellular and tissue levels as well as organs and organ systems.

Over the span of last century several fields have converged to enable scientists to address life at various scales, facilitating its understanding and prediction at different levels [6]. Microscopy, especially diffraction-limited light microscopy, has emerged as a central tool to capture the spatio-temporal context of life processes. Perhaps it would't be an exaggeration to state that microscopy has transformed life sciences more any other discipline. While the efforts towards deciphering biological circuits through biochemical analyses, and more recently, systems biology approaches have offered great insights, both fall short of meeting the challenges posed by deciphering the physical mechanisms underlying development and organogenesis in the context of the entire organism. Imaging has the unique advantage of watching biological events as they unfold over time at single-cell resolution in the intact animal [11].

1.2 Microscopy and the foundations of modern biology

The profound impact of microscopy on biology can be traced back to the 1590s when the Dutch spectacle-makers Hans and Zacharias Janssen built what could be called the first compound microscope [12]. Although a rudimentary magnifying device consisting of two lenses placed at the end of a cylindrical structure, this was already a significant advance from single lens magnification devices. The invention was so well received that English and Italian instrument makers discovered a market for microscopes as early as the 1650s. With the improvements made using additional lenses, Robert Hooke was one of the first to look at the fine structure of cork and comment in 1665 that his *'microscope informs that the substance of Cork is altogether fill'd with Air, and that that Air is perfectly enclosed in little Boxes or Cells distinct from one another'* [13]. In fact, he had just discovered and coined the term 'cell'; later to be established as the structural and functional unit of life. The revolution those primitive microscopes had brought was seen all around in observations of plants reported by Nehemiah Grew, Malpighi, and others [14]; however the real paradigm shift in biology occurred when the Dutch tradesman Antony van Leeuwenhoek, in the late 1660s, started polishing lenses to improve the magnifying capabilities of microscopes up to 200 times. He became the first person to observe single cell organisms, protozoa, bacteria and spermatozoa, which he called animalcules [15]. Microscopy had thus made its presence felt in the field of life science. It had transformed biology into a science of the small in comparison to medicine, its earlier precursor, based primarily on macroscopic anatomy and physiology. The restrictions imposed by the inability to observe structures at the scale of microns, which had limited studies of anatomy to the organ level, had now been overcome.

Persistent efforts to improve microscope design in the early 18th century highlighted major causes of poor image quality, particularly spherical and chromatic aberrations. In spherical aberration light that passes through different portions of a lens is focused at different depths, and in chromatic aberration different wavelengths are focused at different depths. This causes blur and color fringes around the image respectively. Around 1830s, it was realized that a combination of lenses, one correcting for the aberration of another, was a potential solution to this problem, largely due to the work of Joseph Jackson Lister and

William Tulley [20]. This greatly enhanced the ability to image living samples 'sharper'and 'clearer'. While cells were widely observed and reported throughout the scientific world, it took almost the 150 years after Leeuwenhoek's discoveries to conceptualize the cell as the *'quantum minimum'* of life, capable of independent development, yet an integral part of the organism as a whole [16]. It took the creativity and boldness of Theodor Schwann and Mathias Jacob Schleiden in 1838-39 to conclude 'that there is one universal principle of development for the elementary part of organisms, however different, and that this principle is the formation of cells' [14,17]. Rudolf Virchow's aphorism *'Omnis cellula e cellula'* (All cells from pre-existing cells) provided further strength to the 'cell theory' [16,18]. Following this 'cell hypothesis', observations of neurons, such as those made by Santiago Ramón y Cajal in the developing chick brain using Golgi's staining techniques, resulted in the 'neuron hypothesis'; the idea that the unit of brain's structure and function are neurons, discrete, autonomous cells that send electrical signals that have a chemical basis for communication with each other [21]. This seminal work now provided the first unifying attribute to the study of life, the complexity of which, across different species, can now be understood by studying its fundamental structural and functional unit, the cell. In parallel microscopy emerged as the tool to 'observe' these units. Contemporaneously, the field of microscopy advanced by leaps and bounds, benefitting from the diverse abilities of Ernst Abbe, Friedrich Otto Schott, Carl Zeiss, Lord Rayleigh, Richard Zsigmondy, Frits Zernike, Ernst Ruska, August Köhler, and Georges Nomarski [19], thereby, rendering itself more and more amenable to the understanding of the multi-scale nature of embryonic development.

Today we have travelled a great distance along our scientific sojourn to a place where biological light microscopy attracts the best talents from optics, instrumentation, synthetic chemistry, genetics and cell biology, molecular programming, image analysis, computation, and related engineering fields. Together these disciplines are continuously pushing the limits of the spatio-temporal resolution required to understand how biological circuits work in specific physical contexts to produce three-dimensional (3D) forms and patterns. We now have the ability to not only quantitatively decipher the role of multiple players (via multiplexing), but also to probe into the separation and inter-dependence of scales (molecular,

mechanical and developmental) ranging from inter-cellular to organ level domains both *in situ* and *in vivo*. The abundance of both direct and correlative imaging techniques, coupled to image processing, are proving to be the *'open sésame'* for the rich variety of questions that we can now ask (Table A.1). In this thesis we present a range of problems in morphogenesis, each unique in its requirements for novel quantitative imaging both in terms of the technique and analysis. We show how the combination of new molecular and genetic tools with novel microscopies play a vital role in developing intuition about living systems via translating cellular events to organ level.

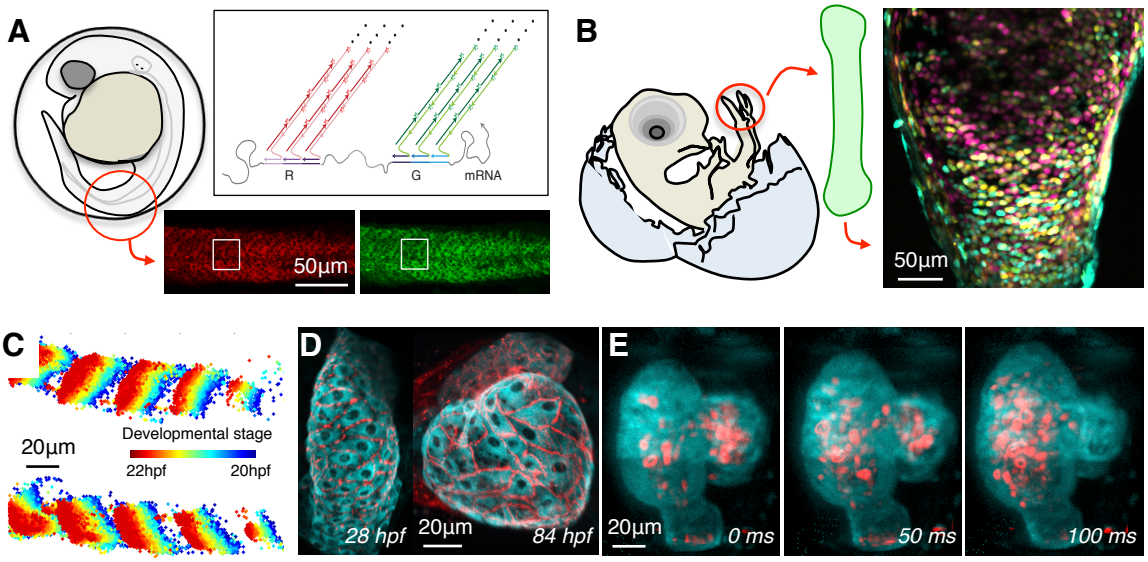


Figure 1.1: **From molecules to organs: Imaging and multi-scale nature of development**
This thesis addresses a range of problems in morphogenesis, each unique in its requirements for novel quantitative imaging both in terms of the technique and analysis. (A) Using hybridization chain reaction (HCR), quantitative sub-cellular mapping of mRNA expression within intact zebrafish embryos has been achieved. (B) Long term *in vivo* imaging of avian metacarpals allowed the capability to follow cells as they displace, divide, orient, enlarge, etc within intact cartilage during development. (C) Real-time monitoring of Dmd expression and localization *in vivo* at the transverse myosepta via live imaging of developing somites in zebrafish. (D) High-resolution confocal imaging of zebrafish heart at various stages, where tricaine was utilized to anesthetize embryos just prior to imaging, allows the study of changes in cell size and shape *in vivo* to understand events such as cardiac chamber formation. (E) Fast dynamic imaging of beating zebrafish heart with 2-photon light sheet microscope has achieved 4D (3D+time) reconstructions of beating heart at cellular resolution.

1.3 Imaging fixed specimens to understand genetic machinery *in situ*

Understanding the molecular basis for a developmental process involves investigating how genes and their products—mRNA and proteins—function in the context of a cell. Structural information holds the key to insights into mechanisms and imaging fixed specimens paves the first step towards deciphering gene function. Probing into the transcriptional state of cells involves characterizing when and where a gene (or a collection of genes) is expressed i.e. spatiotemporal changes in the expression of mRNA and proteins. Methods using transcriptional and translational reporter constructs, with fluorescent readouts, allow assessment of expression dynamics *in vivo*. However it must be noted that the time scales for RNA or protein dynamics at the molecular level (transcription, translation, posttranslational modifications, folding, etc) are much shorter than the time scales of their localization during developmental processes, ranging from minutes to hours. Detecting such expression patterns, in many cases, can be achieved by relaxing the need for live imaging through adopting fixed sample imaging (i.e. compromising on temporal resolution) yet achieving high spatial resolution, sub-cellular if possible. Immuno-histochemical techniques for proteins and *in situ* hybridization techniques for RNAs are useful toolkits in this respect.

A variety of mRNA *in situ* hybridization methodologies have been developed, over decades, that utilize different colorimetric or fluorimetric techniques. However their sequential contrast generation ability renders it difficult to visualize multiple mRNA targets within a single intact tissue/embryo. This functionality is greatly desired to decipher the relationships between different gene networks, thereby achieving a systems biology understanding of the regulatory circuits within the organism. Additionally their deposition chemistry renders them non-quantitative. Borrowing from the fields of molecular programming and nucleic acid nanotechnology, hybridization chain reaction (HCR) for conditional nucleic acid self-assembly [22] has been adapted in the context of mRNA *in situ* hybridizations and has been shown to provide durable, cost-effective, isothermal, and enzyme-free molecular signal amplification with high signal-to-background ratio *in situ* [23,24]. Orthogonality of the method allows easy multiplexing. In Chapter 2 and 3, I will discuss the method's mul-

tiplexing ability as a basis for evaluating quantitative aspects of the fluorescent signal using redundant detection of the same target mRNA with spectrally distinct fluorophores. I will elaborate on the efforts to establish that the fluorescent intensity, integrated within sub-cellular image voxels, scales linearly with the number of copies of target mRNA to provide quantitative sub-cellular mapping of mRNA expression within intact vertebrate embryos.

1.4 Enriching spatial information with low time resolution to learn mechanisms of slow growth processes

Cell migration and proliferation in the context of bone, cartilage or cell growth and proliferation during plant development, are slow developmental processes spread over multiple days [25,26]. This makes them special in their need for non-invasive imaging over long durations while keeping photo-toxicity and photo bleaching under control. This demands a balance between the total duration of imaging and the time resolution during the same. Understanding cellular events like changes in cell shape, size, number, position, etc. over long durations imposes a greater demand on spatial resolution than temporal resolution that can range from several tens of minutes to hours [27].

In the particular example of bone and cartilage development, the complex set of events that contribute to lengthening of skeletal elements viz. proliferation, daughter cell rearrangement, convergent extension, matrix deposition, and cell enlargement [28], occur at very long time scales and in several cases over multiple developmental phases [29]. For instance, chondrocytes migrate over a few microns, with slower division and rearrangements spread over hours [30]. Temporal aliasing in such a case can be successfully avoided even with time resolution of an hour. Previous studies based on static imaging do not provide direct evidence of the specific contributions of such processes, owing to non-uniform changes induced in inter- and intra-cellular volume during fixing. Additionally the thicknesses and optical properties of the samples pose further challenges for non-invasive live imaging. Imaging such processes not only requires novelty on the optical side, but also in terms of mounting the sample for imaging.

In Chapter 4, I will elaborate on the work on an avian metacarpal organ culture system

where time-lapse two-photon laser scanning microscopy was used to observe cell behaviors over more than two days with high spatial resolution and sufficient temporal resolution to trace cellular trajectories while avoiding photo-toxicity [31]. This approach allowed the capability to follow cells as they displace, divide, orient, enlarge, etc within intact cartilage during development. Subsequent quantitative analysis was utilized to show that oriented cell division, convergent extension, and daughter cell rearrangement do not contribute significantly to the observed growth process; instead, extracellular matrix deposition and cell volume enlargement are the key contributors to embryonic cartilage elongation.

1.5 Improving time resolution to capture events occurring over hours

Within the developing musculoskeletal system, a collection of tissue types develop at different rates, ranging from days, as described in case of skeletal elements, to hours, as in the case of rhythmically repeated morphogenesis of somites [32]. The process of somitogenesis in all vertebrates has been of great interest owing to its role in defining a segmented body plan that provides both rigidity and flexibility to the body [33]. Dynamic imaging provides the ability to capture cell movements and rearrangements; fixed imaging techniques can capture gene expression boundaries that can be defined with respect to the somite boundary. A combination of both live and fixed imaging facilitates the understanding of the coordination of such events across space and time [34].

The clarity and accessibility (optical and genetic) of zebrafish embryos make them ideal candidates for observing skeletal muscle development and somitogenesis [35]. As the somites increase in number, from three to around 22 over a period of 9 hours, each somite changes its shape from cuboidal to chevron shortly after its formation. These events span over 1mm of developing trunk and tail. This is concurrently accompanied by the production and localization of proteins like Dystrophin (Dmd) that act as structural elements, linking the extra-cellular matrix (ECM) and actin filaments in muscle fibers. In Chapter 5, I will elaborate on the work done on real-time monitoring of Dmd expression and localization *in vivo* via live imaging. The chapter describes the accumulation of Dmd protein at the

transverse myosepta prior to dmd mRNA localization but concurrent with the onset of myotome formation, a critical stage in muscle maturation.

1.6 Bridging multiple time-scales in growth processes

The great fun of watching organogenesis unfold within the embryonic body is observing the rich diversity of the growth patterns that are available both within an organism as well as across species. In a variety of organs, the shape is remodeled after the primary structure layout over several days and in a number of cases this slow process of morphogenesis is closely linked to cellular processes on shorter time scales. Of particular interest are organs like gut, lung and heart where a unique interplay of distinct time scales is displayed by their constituent cells; especially during their formation where the coordination of peristaltic or peristalsis-like contractions starts very early on in their development. In the case of the embryonic heart, the short timescale of beating governs the pumping ability of the organ and hence its functionality, whereas the long developmental time scale specifies the continuous morphological changes as the tissue is becoming functional. It is very natural to probe into the inter-dependence of the two time scales as crucial for determining the biomechanical implications on organogenesis [36]. The heart is the first organ to form and function (the heart tube forms around 22hpf in zebrafish and starts beating around 24hpf), driven by orchestrated cellular contractions at different scales (both length and time). The necessity of efficient cardiac output for embryonic development, further motivates understanding the dynamic behavior of the tissue in development, disease, and its own morphogenesis [37,38].

The fast rate of myocardial contractions over several hertz of frequency, however, poses challenges for sub-cellular imaging. Understanding cardiogenesis at the cellular level in terms of cell shape, size, number, etc has therefore remained unclear. As a result an alternate approach to image cardiogenesis has evolved that involves pharmacological manipulations to stop the heart and enable its observation at distinct time points. In Chapter 6, I discuss the work on high-resolution confocal imaging of zebrafish heart at various developmental stages, where tricaine was utilized to anesthetize embryos just prior to imaging, thereby halting myocardial contractions at a given stage. I will report the changes in cell size and shape, *in*

vivo and in 3D, that revealed a previously unappreciated heterogeneity in cardiomyocytes cell size and shape between the atrial and ventricular populations. Comparative analysis of cardiac mutants with contractile defects, argues for decoupling of contractile mechanics from cardiomyocytes growth, especially in case of early events like looping. With the aid of computational simulations based on quantitative imaging we further show that growth with spatial constraints can sufficiently explain early cardiogenesis events without invoking the requirements of contraction.

1.7 Pushing the speed of fast dynamic imaging to understand rapid mechanical time scales

The challenges posed by the beating motion of the developing heart have been the major roadblock in its fast sub-cellular resolution imaging. The zebrafish heart, around 5dpf, is a dynamic structure about 200 microns in size and moving with a 3D amplitude of 100 microns with each heartbeat (cardiac cycle). This repeats quasi-periodically at a frequency of few Hertz. Similar issues in other organisms have prevented volumetric imaging of the beating heart *in vivo*. Achieving the right balance between spatial and temporal resolutions is very important in the study of a dynamic tissue like heart. For example, to study the overall contraction, tissue level resolution might be sufficient, thereby allowing high temporal resolution. On the other hand, achieving sub-cellular resolution in imaging beating heart might compromise the temporal resolution. A plethora of imaging strategies has emerged to probe into the nature of contractions at different spatial scales. Wide field imaging with quantitative physiological measurements have shown that the embryonic vertebrate heart tube is a dynamic suction pump [39]. In studies aimed at measuring mechanical properties of tissues *in vivo*, where the need for sub-cellular resolution can be relaxed, imaging modalities like elastography have been demonstrated non-invasively in carotid arteries [40]. However the limitations in achieving direct 3D+time (4D) imaging with existing modalities have motivated prospective gating approaches where the 2D+time acquisitions are triggered at the same point (phase) in the cardiac cycle based on an independent signal such as cardiac conduction or respiration [41-43].

Retrospective registration techniques, on the other hand, rely on image similarity measures for rearranging asynchronously acquired 2D+time images to achieve 4D reconstructions of the beating heart. Due to requirements of additional hardware and real-time processing in the prospective approaches, retrospective registration approaches have gained popularity. Although the simplicity of the retrospective method in terms of image acquisition is its greatest advantage, it is inefficient in collecting data and therefore exposes the specimen to more light than necessary. Bleaching and photo-toxicity can occur during continuous illumination of the entire sample in confocal (laser scanning microscope,LSM) imaging. Single Plane illumination Microscopy (SPIM) [44] greatly helps in this regard as it generates optical sectioning by restricting the illumination to a single plane at a time. This drastically reduces the effects of bleaching within the sample. Added to this, the ability to illuminate the entire plane simultaneously, as opposed to pixel by pixel in LSM, allows imaging at very fast rates. As one spends less time on every plane generating a 2D image, this further reduces bleaching and photo-toxicity. Thus SPIM allows imaging biological samples in 3D for extended periods of time at fast frame rates [44,45]. Orthogonal geometry of SPIM provides further possibilities for improvements and robustness to reconstruction by recording an independent signal to accurately decide the spatio-temporal phase of the cardiac wall motion [46]. In Chapters 7 and 8, I discuss the work done with 2-photon SPIM to achieve 4D reconstruction of the beating heart at cellular resolution. These techniques open new avenues to linking the cellular shape changes due to contractile behavior of cardiomyocytes to the overall tissue bending during the cardiac cycle, thereby facilitating better understanding of cardiac development and function.

1.8 References

1. Uri Alon Connections simplicity in biology Nature 446, 497 (2007)

2. Bechtel, W. and Robert C. R. Vitalism. E. Craig (Ed.), Routledge Encyclopedia of Philosophy. London (1998)

3. Douglas M. Whler's 'Synthetic'Urea and the Rejection of Vitalism: A Chemical Legend Nature 153, 608-610 (1944)

4. Berzelius, J.J. (1836) 'Einige Ideen ber bei der Bildung organischer Verbindungen in die lebenden Naturwirksame ober bisher nicht bemerkte Kraft', Jahres-Berkcht ber die Fortschritte der Chemie 15: 237'45

5. John Gerhart , Marc Kirschner Cells, Embryo, and Evolution: Toward a Cellular and Developmental Understanding of Phenotypic Variation and Evolutionary Adaptability Blackwell Science Inc; 1 edition (1998)

6. Anteneodo, Celia and M. G. E. da Luz. Complex Dynamics of Life at Different Scales: From Genomic to Global Environmental Issues. Philosophical Transactions of the Royal Society A. 368/5561, 2010.

7. Rob Phillips, Jane Kondev, Julie Theriot, Hernan Garcia Physical Biology of the Cell: 2nd Edition (2012) Taylor and Francis

8. Pattern formation in morphogenesis Forgacs, Biological Physics of the developing embryos

9. Andrew C. Oates, Nicole Gorfinkiel, Marcos Gonzlez-Gaitn and Carl-Philipp Heisenberg Quantitative approaches in developmental biology Nature Reviews Genetics 10, 517-530 (2009)

10. Sean G. Megason ,Scott E. Fraser Imaging in Systems Biology Cell (2007) 130(5) , 784'795

11. http://micro.magnet.fsu.edu/primer/museum/bibliography.html#collecting

12. Hooke, Robert C. Micrographia: or Some Physiological Descriptions of Miniature Bodies Made by Magnifying Glasses. London, England: Jo. Martyn, and Ja. Allestry, 1665.

13. Wm. Turner The Cell Theory, Past and Present J Anat Physiol. Jan 1890; 24(Pt 2): 253'287.

14. Scientific Collections, Dr. W.A.W.Moll, University of Amsterdam

15. M. Tavassoli The cell theory: a foundation to the edifice of biology. Am J Pathol. 1980; 98(1): 44.

16. Theodor Schwann (Translated by Henry Smith) Microscopical researches into the accordance in the structure and growth of animals and plants, 1847, London, Sydenham Society.

17. Bagot, C. N., Arya, R. "Virchow and his triad: a question of attribution". British Journal of Haematology (2008). 143 (2): 180'189

18. http://www.nobelprize.org/educational/physics/microscopes/timeline/

19. Boris Jardine, 'The problems with lenses, and the 19th-century solution', Explore Whipple Collections, Whipple Museum of the History of Science, University of Cambridge, 2008

20. B. KolbIan, Q. Whishaw, 'Fundamentals of Human Neuropsychology'. Freeman/Worth; 6th edition (2008)

21. Triggered amplification by hybridization chain reaction R.M. Dirks and N.A. Pierce Proc Natl Acad Sci USA, 101(43): 15275-15278, 2004

22. Programmable *in situ* amplification for multiplexed imaging of mRNA expression H.M.T. Choi, J.Y. Chang, L.A. Trinh, J.E. Padilla, S.E. Fraser, and N.A. PierceNature Biotechnol, 28:1208-1212, 2010.

23. Next-generation *in situ* hybridization chain reaction: higher gain, lower cost, greater durability H.M.T. Choi, V.A. Beck, and N.A. PierceACS Nano, 8(5):4284-4294, 2014.

24. Lefebvre V, Bhattaram P. Vertebrate skeletogenesis. Curr Top Dev Biol. 90:291-317. (2010)

25. Kuchen EE, Fox S, de Reuille PB, Kennaway R, Bensmihen S, Avondo J, Calder GM, Southam P, Robinson S, Bangham A et al.: Generation of leaf shape through early patterns of growth and tissue polarity. Science 2012, 335:1092-1096.

26. Robinson S, Barbier de Reuille P, Chan J, Bergmann D, Prusinkiewicz P, Coen E: Generation of spatial patterns through cell polarity switching. Science 2012, 333:1436-1440.

27. Cooper KL, Oh S, Sung Y, Dasari RR, Kirschner MW, Tabin CJ. Multiple phases of chondrocyte enlargement underlie differences in skeletal proportions. Nature 7441:375-378. (2013)

28. Cole AG A review of diversity in the evolution and development of cartilage: the search for the origin of the chondrocyte. Eur Cell Mater. 2011 21:122-9.

29. Romereim, S. M., Conoan, N. H., Chen, B. and Dudley, A. T. A dynamic cell adhesion surface regulates tissue architecture in growth plate cartilage. Dev. Camb. Engl. 141, 2085-2095 (2014).

30. Li,Y., Trivedi, V., Truong,T.V., Koos,D.S., Lansford,R., Chuong,C.M., Warburton,D., Moats,R.A., Fraser S.E. Dynamic imaging of the growth plate cartilage reveals multiple con- tributors to skeletal morphogenesis Nature Communications (2015) 6:6798 doi: 10.1038/ncomms7798

31. Schrter, C., Herrgen, L., Cardona, A., Brouhard, G. J., Feldman, B. and Oates, A. C. (2008), Dynamics of zebrafish somitogenesis. Dev. Dyn., 237: 545'553

32. Miguel Maroto, Robert A. Bone and J. Kim Dale Somitogenesis 2012 Development 139, 2453-2456.

33. Hester SD, Belmonte JM, Gens JS, Clendenon SG, Glazier JA (2011) A Multi-cell, Multi-scale Model of Vertebrate Segmentation and Somite Formation. PLoS Comput Biol 7(10): e1002155.

34. Chen and Jenna L. Galloway The development of zebrafish tendon and ligament progenitors Jessica W. Development (2014) 141, 2035-2045

35. Bartman T., Hove, J. Mechanics and function in heart morphogenesis. Dev. Dynamics (2005), 233 373-381

36. Hove, R.J., Koster,R.H., Forouhar, A.S., Acevedo-Bolton, G., Fraser, S.E., Gharib, M., Intracardiac fluid forces are an essential epigenetic factor for embryonic cardiogenesis. Nature, 2003. 421: p. 172-177.

37. Bartman, T., Walsh, E.C., Wen, K.K., McKane, M., Ren, J., Alexander, J., Rubenstein, P.A., Stainier, D.Y.R., Early myocardial function affects endocardial cushion development in zebrafish. PloS Biology, 2004. 2(5): p. 673-681.

38. Forouhar, A.S., Liebling, M., Hickerson, A., Nasiraei-Moghaddam, A., Tsai, H.J., Hove, J.R., Fraser, S.E., Dickinson, M.E., Gharib, M., The embryonic vertebrate heart tube is a dynamic suction pump. Science, 2006. 312: p. 751-753.

39. Mercure E, Destrempes F, Roy Cardinal MH, Pore J, Soulez G, Ohayon J, Cloutier G. 'A local angle compensation method based on kinematics constraints for non-invasive vascular axial strain computations on human carotid arteries'Comput Med Imaging Graph. 2013 Aug 30

40. S. H. Bartling, J. Kuntz, and W. Semmler, 'Gating in small-animal cardio-thoracic CT,'Methods, vol. 50, no. 1, pp. 42'49, 2010.

41. M. W. Jenkins, O. Q. Chughtai, A. N. Basavanhally, M. Watanabe, and A. M. Rollins, '*in vivo* imaging of the embryonic heart using gated optical coherence tomography,'J. Biomed. Opt., vol. 12, no. 03, pp. 030505'030505, May 2007.

42. Yagel S, Cohen SM, Shapiro I, Valsky DV.3D and 4D ultrasound in fetal cardiac scanning: a new look at the fetal heart. Ultrasound Obstet Gynecol. 2007, 29(1):81-95.

43. Hung J, Lang R, Flachskampf F, Shernan SK, McCulloch ML, Adams DB, Thomas J, Vannan M, Ryan T 3D echocardiography: a review of the current status and future directions.J Am Soc Echocardiogr. 2007 20(3):213-33

44. Reynaud, Emmanuel G.; Peychl, Jan; Huisken, Jan; Tomancak, Pavel Guide to light-sheet microscopy for adventurous biologists. Nat. Methods, 12, no. 1, pp. 30-34, (2014)

45. Trivedi,V., Truong,T.V., Trinh,L.A., Holland, D.B., Liebling,M., Fraser,S.E. Dynamic structure and protein expression of the live embryonic heart captured by 2-photon light sheet microscopy and retrospective registration. Biomed. Opt. Express (2015) 6, 2056-2066

46. Truong,T.V. , Trivedi,V., Trinh,L.A., Holland D.B., Cutrale M., Choi J.M., Fraser,S.E. Live 4D Imaging of the Embryonic Vertebrate Heart with Two-Photon Light Sheet Microscopy and Simultaneous Optical Phase Stamping. Biophysical Journal (2014) 106 (2), 435a-436a

Chapter 2

Quantitative sub-cellular mapping of mRNA expression within intact vertebrate embryos

2.1 Motivation

Understanding the molecular basis for a developmental process involves investigating how genes and their products—mRNA and proteins—function in the context of cells and tissues. Structural information holds the key to insights into mechanisms and imaging fixed specimens paves the first step towards deciphering gene function. Probing into the transcriptional state of cells involves characterizing when and where a gene (or a collection of genes) is expressed. With the advancements made in sequencing technology, genomic diversity, at the levels of both sequence and function, is now better appreciated. Therefore the ability to visualize the nucleotide sequences in a morphological context is of utmost importance to identify their contributions towards to differences between species, between members of a species, and between cells within individual members. mRNA *in situ* hybridization techniques have been indispensable toolkits in this respect as they enable visualization of RNA molecules within the intact organism [1-10].

The method of mRNA *in situ* hybridization involves gentle fixing of tissues to retain their RNA in a form accessible by complementary DNA or RNA probes, labeled *a priori*, to facilitate visualization. In many cases the evolution of the spatio-temporal expression is slow enough to select discrete time points for studying the evolution of their expression patterns.

Visualizing multiple mRNA targets within a single intact tissue/embryo is greatly desired to decipher the relationships between different gene networks, thereby achieving a systems biology understanding of the regulatory circuits within the organism [11-13]. However with the existing methodologies, that necessitate sequential staining, this remains a challenge [1-6]. Recent advances have introduced direct-labeled fluorescent probes apt for multiplexing but the low signal-to-background, owing to no signal amplification, makes it difficult to use these within thick samples, especially in whole mount context [8-10].

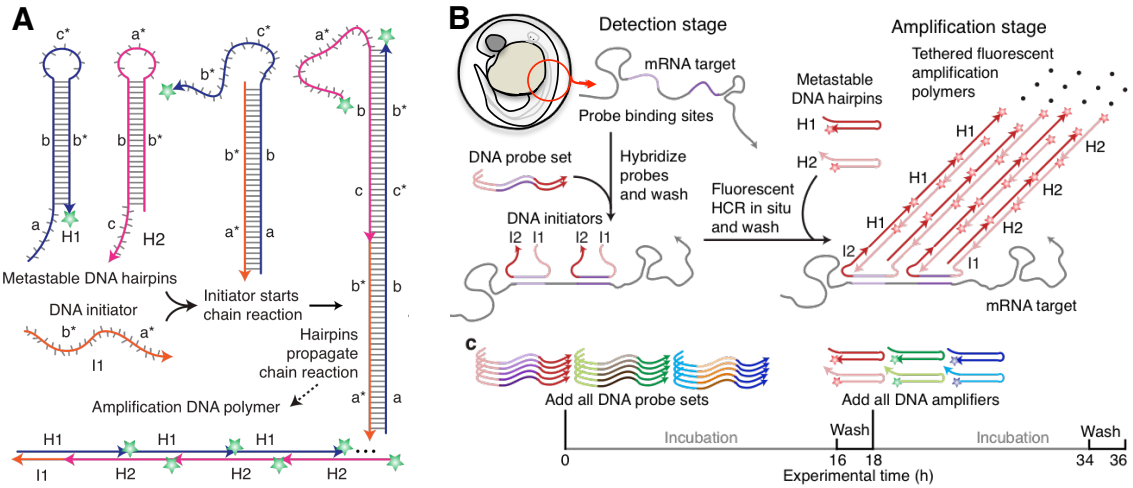


Figure 2.1: **HCR mechanism**
(A) Metastable fluorescent hairpins self-assemble into fluorescent amplification polymers upon detection of a cognate initiator. Initiator I1 nucleates with hairpin H1 via base-pairing to single-stranded toehold 'a', mediating a branch migration that opens the hairpin to form complex I1.H1 containing single-stranded segment 'c*-b*'. This complex nucleates with hairpin H2 by means of base-pairing to toehold 'c', mediating a branch migration that opens the hairpin to form complex I1.H1.H2 containing single-stranded segment 'b*-a*'. Thus, the initiator sequence is regenerated, providing the basis for a chain reaction of alternating H1 and H2 polymerization steps. Green stars denote fluorophores. (B) *in situ* hybridization protocol. Detection stage: probe sets are hybridized to mRNA targets, and unused probes are washed from the sample. Amplification stage: initiators trigger self-assembly of tethered fluorescent amplification polymers, and unused hairpins are washed from the sample. (lower) Experimental timeline. The same two-stage protocol is used independent of the number of target mRNAs. For multiplexed experiments (three-color example depicted), probe sets for different target mRNAs (five probes depicted per set) carry orthogonal initiators that trigger orthogonal HCR amplification cascades labeled by spectrally distinct fluorophores.

Despite their wide usage, mRNA *in situ* hybridization technique has predominantly

remained qualitative in nature, primarily as a contrast generating procedure. Quantitative information has been derived mainly from quantitative real time polymerase chain reaction (qPCR) and microarrays that compromise on the spatial context. Therefore a technique for rigorous quantitative measurements within the context of whole-mount embryos via imaging is best suited to benefit the field of developmental biology tremendously. Limited studies have been pursued along this line. Radioactive nucleotides incorporated into the cRNA probes to detect targets have been shown to provide a linear relation between signal intensity and targeted mRNA amounts, to allow for quantitative analysis [14-15]. However in absence of any signal amplification, contrast development requires long duration, ranging from days to weeks, depending upon target abundance [15]. Traditional colorimetric and fluorimetric *in situ* hybridization techniques (with haptens such as digoxigenin, biotin, fluorescein) have utilized catalytic deposition of reporter molecules (chromogens/fluorogens) for visualization. Herein antibodies conjugated to enzymes such as peroxidase or alkaline phosphatase produce different color precipitates [4-6], which essentially renders the method non-quantitative. Other issues amplify the problem as well; the time of contrast development (the time for precipitation) is target-dependent and thus requires an experienced hand [3,6]. The usage of antibodies in the reaction does not allow parallelization, making the non-orthogonal deposition chemistry a rate-limiting step towards multiplexing. The concerns of resolution while imaging conventionally stained samples have been overcome in recent efforts with tomographic techniques to a generate high-resolution 3D atlas of gene expression in the zebrafish embryo [16] but the method remains largely qualitative.

Borrowing from the fields of molecular programming and nucleic acid nanotechnology, hybridization chain reaction (HCR) for conditional nucleic acid self-assembly [17] has been adapted in the context of mRNA *in situ* hybridizations. The HCR mechanism has now been engineered for multiplexed, durable, cost-effective, isothermal, enzyme-free molecular signal amplification with high signal-to-background ratio, *in situ* [7,8]. Herein a pair of fluorescently labeled nucleic acid hairpins can hybridize with each other recursively to self-assemble into fluorescent polymers only in the presence of a cognate initiator that can trigger this amplification process (Fig. 2.1). Probes complimentary to a target mRNA are designed

with the initiator sequence attached. Upon the addition of hairpins in the following step, the initiator present at the site of the mRNA triggers a chain reaction such that the two hairpins sequentially nucleate and open to assemble into a long double-stranded amplification polymer [8]. As HCR is programmable, amplifiers recognizing different initiator sequences can be designed easily. Since they operate independently within the same sample, attaching spectrally distinct fluorophores to the hairpins allows easy multiplexing. Subcellular signal localization and co-localization, owing to the tethering of the amplification polymer to the initiating probe, provide the potential for correlative studies of multiple targets. But what makes the method's multiplexing ability even more remarkable, is that it provides a basis for evaluating quantitative aspects of the fluorescent signal using redundant detection of the same target mRNA with spectrally distinct fluorophores. In this chapter I will discuss the efforts to establish that the fluorescent intensity, integrated within sub-cellular image voxels, scales linearly with the number of copies of target mRNA to provide quantitative sub-cellular mapping of mRNA expression within intact vertebrate embryos.

2.2 Experimental design

The basic experimental framework is based upon the idea of imaging the HCR signal from the tissue/embryonic region of interest and examining whether HCR provides a quantitative readout of relative mRNA abundance within an image. For the purposes of this study, we decompose the image into *voxels* of $p \times p$ pixels. Let us assume that a voxel v_i contains m_i copies of the target mRNA and has an intensity recorded as x_i. The question to be addressed now becomes: is the HCR fluorescence (x_i) linear in m_i across *voxels*? This could have been an easily verifiable scenario if we had the knowledge of m_i in every *voxel*. However in the absence of this information we have designed an experimental strategy outlined in the paragraphs below.

Our approach is to utilize the multiplexing ability of HCR to redundantly detect the same target mRNA, using orthogonal amplifier systems and spectrally distinct fluorophores, to record two signal intensities (x_i, y_i) for each voxel v_i (Fig. 2.2A,B). For any given probe set consisting of n probes designed to target the mRNA of choice, we divide the probe sets

into two subsets consisting of n_1 and n_2 number of probes (such that $n_1+n_2 = n$) and detect them with orthogonal amplifier systems. Only in the case where signal intensity scales with mRNA abundance (m_i) the signal from two channels be linearly correlated. Conceptually, this comes down to evaluating whether the scatter of (x_i, y_i) *voxel* pairs shows a linear relationship between the x and y channels when detecting the same target mRNA with different probe sets and amplifiers. If the HCR fluorescence is linear in m_i across *voxels*, we would expect a linear relationship as follows :

- $y_i = \beta_0 + \beta_1 x_i.$

In a real experiment, the situation is complicated by several factors. Auto-fluorescence (AF) in the sample prevents us from observing the true intensities (x_i, y_i) directly in the experiments. Non-specific amplification (NSA) of hairpins and non-specific detection (NSD) of the targets, can also corrupt the true signal measurement. Therefore we always measure:

- $(X_i, Y_i) = (x_i, y_i) + (u_i, e_i).$

We interpret (u_i, e_i) as *measurement errors* [18,19] and perform independent control experiments to measure the background statistics. It must be emphasized that the microscope setting is kept identical between all experiments (Appendix B). One important aspect of our approach is that the signal intensity in every voxel is quantitative despite the polydispersity shown in polymer length at the site of hybridization [7.8,17] that should generate a range of fluorescent intensities *in situ*. This is due to averaging occurring at multiple levels. Typically multiple probes are used per target; as a result this causes the signal to be averaged over multiple polymers. Second, as we operate in the diffraction limited imaging regime, a typical pixel size of $\sim$0.3 microns averages signal from multiple target molecules. Further, post-acquisition analysis defines voxels, consisting of several pixels that are averaged together. Together these averaging mask out the effects of polydispersity to ensure that signal intensities, integrated within sub-cellular image voxels, robustly represent the biological distribution of target mRNA and hence can be used for relative quantification of its abundance.

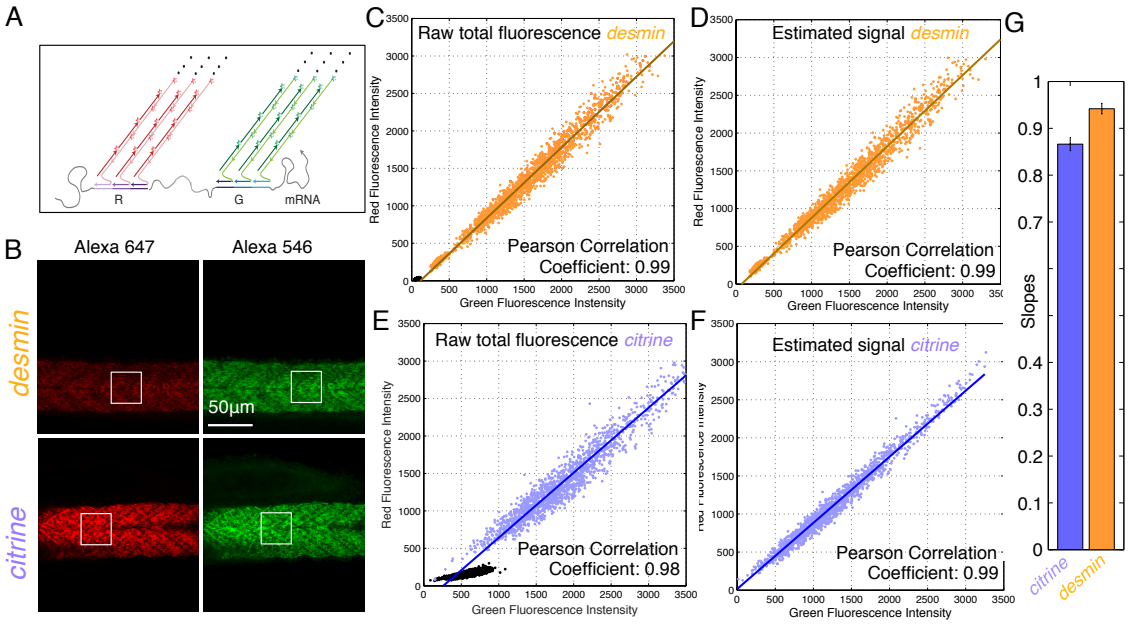


Figure 2.2: **Relative quantification by redundant detection**
(A) Schematic of the redundant detection experiment. For any given probe set designed to target the mRNA of choice, we divide the probe sets into two subsets (R and G) that use orthogonal amplifier systems and spectrally distinct fluorophores to detect the same mRNA. (B) Representative confocal images for two targets (*desmin* and *citrine*) in the red (Alexa 647) and green (Alexa 546) channels. White squares denote the representative regions of expression that are decomposed into voxels for subsequent analysis. Embryos were fixed at 26 hours post fertilization (hpf) (C,E) Plots of red vs green voxel intensity (raw total fluorescence) for *desmin* (orange) and *citrine* (purple) for six embryos combined together. Black dots denote the background NSA (+AF) and NSD (+AF +NSA) for *desmin* and *citrine* respectively, obtained from control embryos. (D,F) Scatter of estimated HCR signal for *desmin* (orange) and *citrine* (purple) with respect to an accurate linear model after correcting for the background. Note that in case of *desmin* the low background does not contribute significantly to the scatter, whereas in case of *citrine* the signal statistics are recovered after deconvolving the error in data polluted with the background, thereby giving a tight scatter. For the fit, sample size is 2400 voxels (400 voxels x 6 embryos). (G) Mean value of the six slopes obtained using maximum likelihood estimates for each of the six embryos. Error bars denote the standard deviation of six slopes from six embryos. For each fit per embryo, sample size is 400 voxels.

The final point of emphasis in our analysis is the acknowledgment of the fact that for any given probe set the actual number of probes bound to the target *in situ* will always be less than or equal to the total number of probes in the probe set (n), owing to the competition between mRNA secondary structures and other unknown factors. As a result such a fluctuation in hybridization yield and hence signal generation across voxels will cause the true HCR signal (x_i, y_i) to deviate from an exact linear model. We therefore add *equation errors* [18,19] in our model to account for this effect. These *equation errors* are modeled as a Gaussian around the ideal linear relationship such that the deviation from the linear model is simply a scatter around the line. Taking into account these issues, we have derived a rigorous maximum likelihood estimate for the slope of line (β_1), which is discussed in Chapter 3.

2.3 Relative quantification by redundant detection

Before proceeding to demonstrate the redundant detection approach for relative quantification of mRNA, we wanted to test whether the embryo has any systematic penetration issues. This is crucial to ensure that the changes in recorded intensities are due to relative abundance of target mRNA and not due to insufficient probe penetration. We tested this by detecting two targets *ntla*, expressed deep within the embryo in the notochord, and *desmin*, a more superficial target in the somites. Our data (Fig B.1) shows that penetration is not an issue; signal for both the targets saturates ∼8 hours. Therefore we always perform our quantitative experiments with overnight (∼ 16 hours) probe hybridization.

For the redundant detection studies, we chose two targets, *desmin* and *citrine*, in somites within zebrafish embryos. For the case of endogenous *desmin* target the best estimate of background comes from NSA (+ AF) in an independent sample, as the expression is ubiquitous within the embryonic trunk. In case of the transgenic line, targeting citrine allows us to estimate the NSD (+ AF + NSA) background from a wild type embryo that does not have citrine. First the signal for the two channels in both cases are strongly correlated as indicated by the Pearson correlation coefficients (Fig. 2.2C). This already provides the evidence for the quantitative nature of HCR where the signal intensities map

the relative abundance. We then apply our model to estimate the slope, taking into account all the background statistics. The standard errors on these estimated slopes are very small (Fig. 2.2D) which is consistent with the strong correlation between two channels. Our simulations for different correlation coefficients between the signal in the two channels (Fig B.3) has shown that standard errors in the estimated slope is inversely related to the correlation coefficient between the two signals. Therefore strong correlation between signal implies both accuracy of quantification as well as the quality of slope estimation. The redundant detection assay also provides us with the advantage of estimating the true signal by deconvolving the background from observed signal. This can be seen in the graphs to the right in Figures 2.2C and D. In the case that HCR is quantitative in the number of mRNA target molecules per sub-cellular voxel, the corrected data has a tighter scatter around the accurate linear model, as seen in the panels. This further allows us to characterize the precision of the relative quantification, even in the single color detections of the target of interest. Our results therefore show that redundant detection proves that fluorescent signal from a *voxel* indeed scales with the number of mRNA copies within the *voxel*.

Experimentally we have also observed evidences of cooperativity between probes or between polymers on a single target molecule (Fig B.4). Detecting the same mRNA molecule with different probes between different experiments has shown that such an intra-target cooperativity affects the signal intensities in such experiments. It must be noted though, that despite such effects (which appear to be mediated by mRNA), as long as the same probe/probe sets are used consistently throughout the experiment relative quantification remains unaffected. This allows us to compare *voxels* within an embryo as well as in different embryos exposed to the same probe set(s).

2.4 Distinguishing homozygous and heterozygous embryos with HCR

In order to validate the quantitative aspect of HCR further, we used transgenic lines to assess signal intensities for homozygous and heterozygous embryos. We used FlipTrap zebrafish lines [20] that express Citrine-fused proteins. This allows us to detect *citrine* as well as the

target mRNA to which it is fused (*desmin* in our case). In case of homozygous embryos all mRNA copies are expected to have the *citrine* sequence whereas the heterozygotes will have half (Figure 3A). Wild type transcripts are expected to have only the endogenous *desmin* sequence. Similar experimental procedure with relevant controls reveals that while the concentration of endogenous *desmin* (as assessed by mean signal intensity) remains the same across embryos (Fig. 2.3B), *citrine* intensity in homozygotes is approximately twice that in heterozygotes (Fig. 2.3C), as expected. Even at the voxel level (Fig. 2.3D), out estimates of the slopes (Fig. 2.3E) show that the number of *citrine* mRNA copies in homozygotes is approximately twice that in heterozygotes. It must be noted that owing to the effects of cooperativity between probes and hairpins (due to potential interaction of probes and polymers mediated by the target molecule), the observed intensities in the green (*desmin*) channel do differ between wild type and transgenic embryos (where the transcript differs due to presence or absence of the *citrine* sequence). Similar effect is expected for the red channel that contribute to deviations from the ideal ratio of 2. However for broad classification of genotypes, such as ones used in this section, HCR still provides a reasonable assay for distinction. It must be remembered again that such cooperativity issues do not affect the standard quantification experiment where a target is detected using the same probe set throughout an embryo.

2.5 Validating the slope estimates

Having established that HCR signal is quantitative not only on an average over multiple cells (mean signal intensity of citrine in previous section) but also at the level of sub-cellular voxels, we wanted to further demonstrate that the slope estimates we have are robust. Our approach is very similar to the redundant detection described before. Herein we detect endogenous *desmin* in three spectrally distinct channels (red (Alexa 647), green (Alexa 546) and blue (Alexa 488)) (Fig. 2.4A,B). We then calculate pairwise slopes for the three channels for each embryo based on the statistics available for the signal and the background. It can be seen that only in the case when the signals are linear in the number of detected mRNA copies, the product of the three slopes be unity. We show that the product of

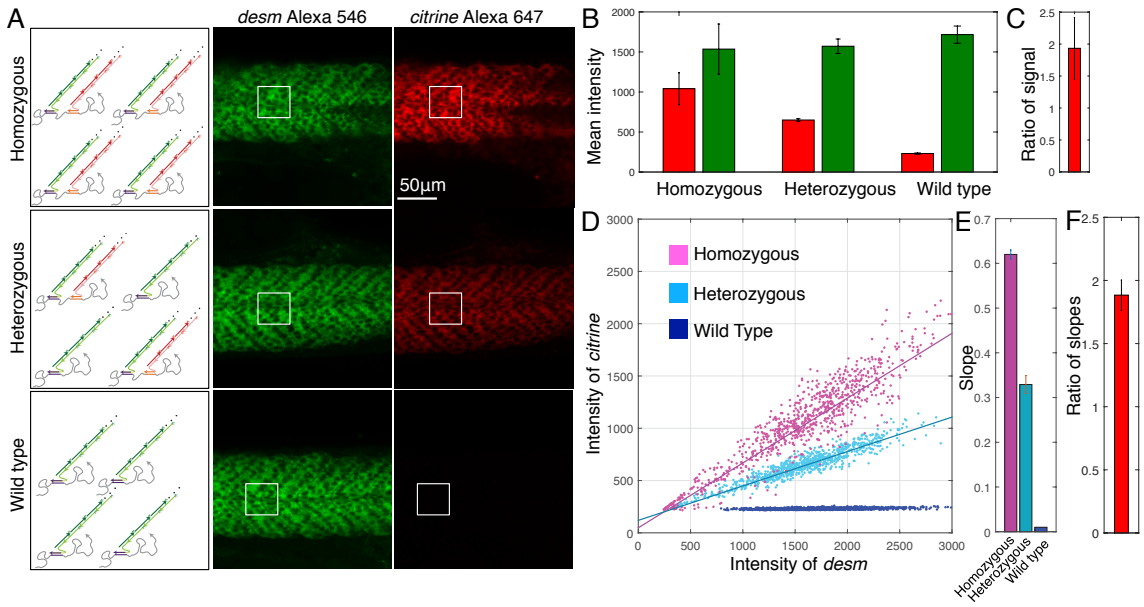


Figure 2.3: **Distinguishing homozygous and heterozygous embryos with HCR** (A) Schematic of the experiment using detection of endogenous *desmin* and fused *citrine* in different kinds of embryos. Endogenous *desmin* serves as an internal reference to permit comparison of citrine fluorescent intensities between voxels and between experiments. Representative confocal images for two targets (*desmin* and *citrine*) are shown in the red (Alexa 647) and green (Alexa 546) channels. White squares denote the representative regions of expression that are decomposed into voxels for subsequent analysis. Embryos were fixed at 26 hours post fertilization (hpf) (B) Mean signal intensity for different samples. Error bars denote the standard deviation of mean intensity between 4 embryos of each kind. (C) Ratio of the mean citrine signal intensity (after subtracting the background) between the homozygous and heterozygous embryos. Error bar denotes the standard deviation obtained by propagation of errors in the division. (D) Plots of red (citrine) vs. green (desmin) voxel intensity for four embryos. For the fit, sample size is 1600 voxels (400 voxels x 4 embryos). (E) Mean value of the four slopes obtained using maximum likelihood estimates for each of the four embryos. Error bars denote the standard deviation of six slopes from six embryos. For each fit per embryo, sample size is 400 voxels. (F) Ratio of slopes between the homozygous and heterozygous embryos. Error bar denotes the standard deviation obtained by propagation of errors in the division.

the three slopes for individual embryos is close to one taking into account the estimated standard errors in every slope and the propagation of errors in the product (Fig. 2.4C,D). We therefore establish that HCR signal is linear in the number of mRNA copies.

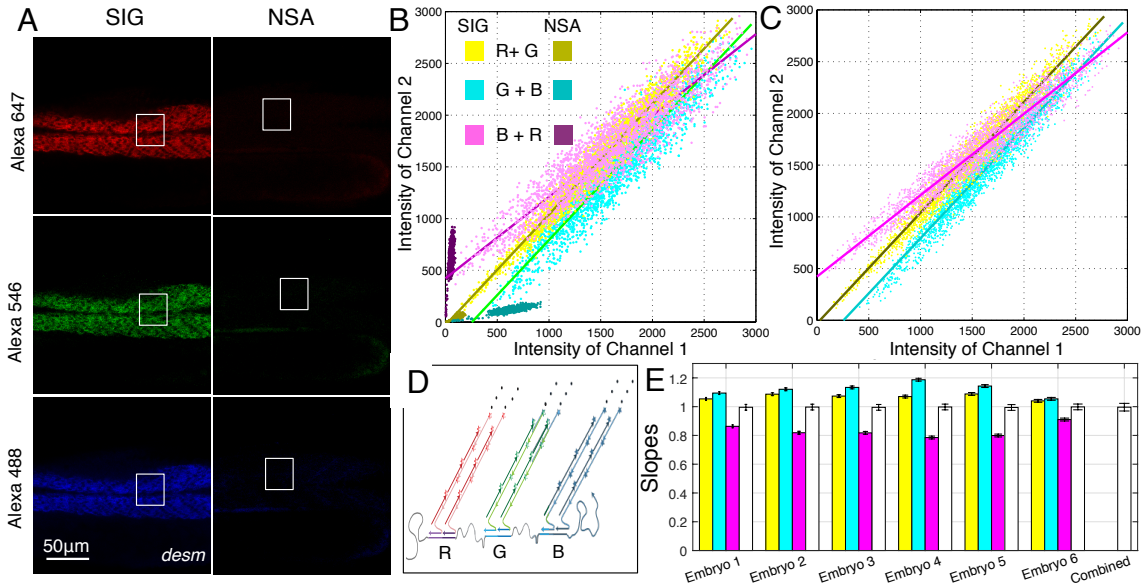


Figure 2.4: **Validating the slope estimates using redundant detection with 3 colors** (A) Representative confocal images for the endogenous *desmin* in the red (Alexa 647), green (Alexa 546) and blue (Alexa 488) channels. White squares denote the representative regions of expression that are decomposed into voxels for subsequent analysis. Embryos were fixed at 26 hours post fertilization (hpf) (C) Plots of voxel intensities for six embryos in different channels plotted in different combinations, as denoted by the legend in top left corner. For each combination the background is obtained from NSA (+AF) for the endogenous target from control embryos. For the fit, sample size is 2400 voxels (400 voxels x 6 embryos). (C) Scatter of estimated HCR signal with respect to an accurate linear model after correcting for the background. Note the reduction in the scatter in the estimated signal after deconvolving the background. (D) Schematic of the redundant detection experiment using 3 fluorophores. For any given probe set designed to target the mRNA of choice, we divide the probe sets into three subsets (R,G and B) that use orthogonal amplifier systems and spectrally distinct fluorophores to detect the same mRNA. (E) Slopes obtained for every embryo using maximum likelihood estimates. For each fit per embryo, sample size is 400 voxels. Error bars denote estimated standard error. White bars denote the product of the three slopes taking into account the error propagation in multiplication. The combined white bar denotes the average of individual products and remains close to 1. Error bar denotes the standard deviation over the 6 products.

2.6 Somitogenesis and genetic expression

In order to demonstrate the application of quantitative HCR in a biological setting, we used somitogenesis in early zebrafish embryos as a test case. The problem of embryonic body segmentation has been a challenge owing to the involvement of multiple genes and the inability of conventional *in situ* techniques to multiplex [11-13]. Using HCR we quantified the signal intensity of 4 different targets within intact zebrafish embryos to show their spatial expression pattern (Fig. 2.5). We can not only visualize and study these targets simultaneously but also exploit the relative quantification to form models of such gene regulatory networks within the spatial context. It must also be emphasized that such multi-channel data visualization is also a challenge and demand innovative data visualization techniques. Further analysis and representing such multi-dimensional quantitative information, which is now available due to quantitative nature of HCR, adds to the challenges in future.

2.7 Discussion

For decades, *in situ* hybridization techniques have provided biologists with a crucial tool for studying mRNA expression in a morphological context. Conventional approaches based on catalytic deposition of reporter molecules provide high contrast within intact vertebrate embryos. However, as outlined before, the staining is largely qualitative, multiplexing is cumbersome, and spatial resolution is typically compromised by diffusion of reporter molecules prior to deposition. *In situ* HCR enables straightforward mapping of multiple target mRNAs with high signal-to-background and sub-cellular resolution within a single intact vertebrate embryo. The current work has shown that the HCR signal scales linearly with the number of target mRNAs per subcellular *voxel* within the embryo, enabling accurate relative quantitation of mRNA abundance for each target within the multiplexed image. Instead of just answering the 'yes/no' question, 'is the transcript expressed in each *voxel*?' the observed accuracy and precision of HCR signal make it possible to ask the quantitative question 'what is the relative transcript expression level in each *voxel*?'. Such an unprecedented quantification is expected to provide newer insights into biological cir-

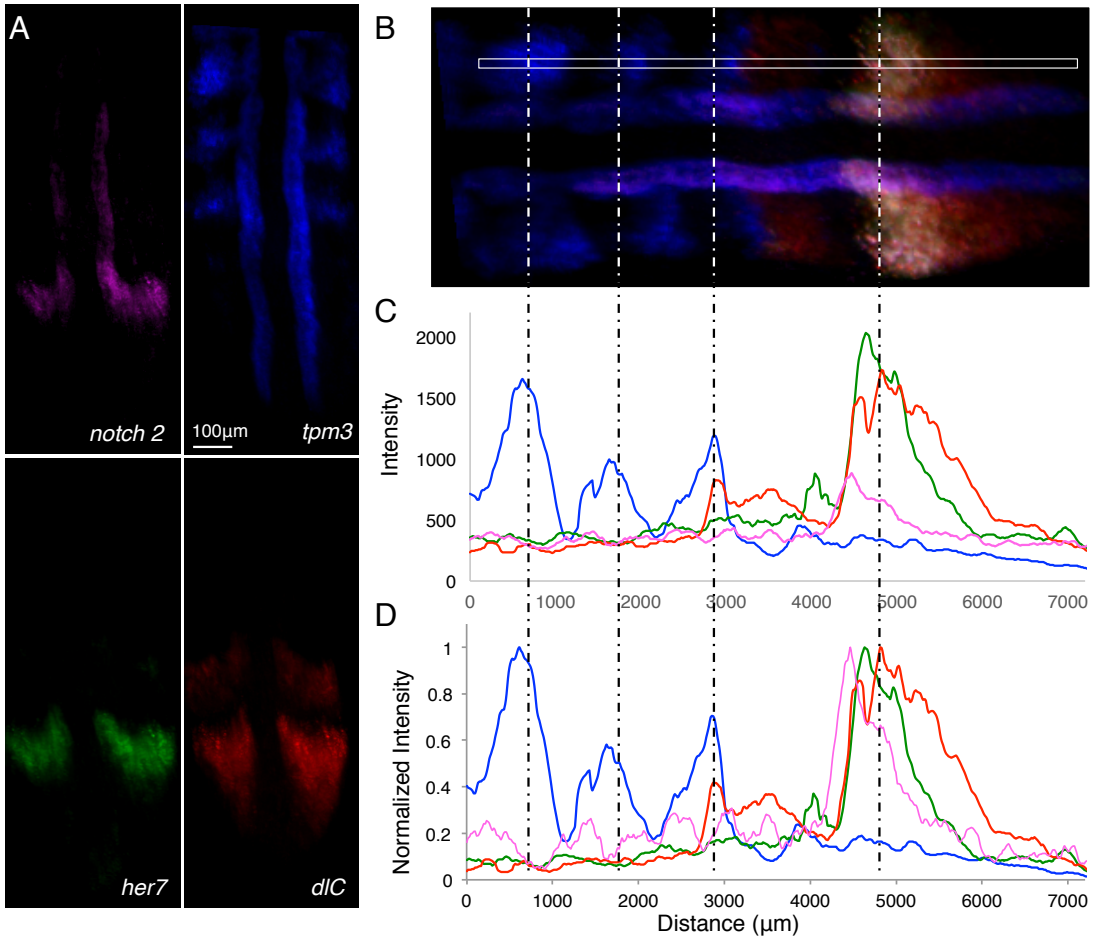


Figure 2.5: **Using HCR to study somitogenesis quantitatively**
(A) Representative confocal images for the four targets used in the study. Embryos were fixed at 10 hours post fertilization (hpf) (B) Mean intensity projection for all the channels shown as a merged image. White rectangle denotes the region of expression that is analyzed. Embryos were fixed at 26 hours post fertilization (hpf) and shown in the dorsal view (C) Plots of raw voxel intensities in different channels plotted as a function of the distance along the embryonic body axis as referenced in (B). (D) Plots of normalized voxel intensities in different channels. For normalization the average background signal (assessed by NSA for these endogenous targets) in every channel was subtracted individually and the signal was then normalized to the resultant maximum value.

cuits that work in the spatio-temporal context of cells and tissues to sculpt out various morphologies as the embryo develops.

2.8 References

1. N Denkers, P Garca?Villalba, CK Rodesch, KR Nielson, TJ Mauch FISHing for chick genes: Triple−label whole−mount fluorescence *in situ* hybridization detects simultaneous and overlapping gene expression in avian embryos Developmental dynamics 229 (3), 651−657

2. Acloque, H., Wilkinson, D.G. and Nieto, M.A. *in situ* hybridization analysis of chick embryos in whole−mount and tissue sections. in Avian Embryology, 2nd Edition vol. 87 (ed. Bronner−Fraser, M.) 169−185 (Elsevier, 2008).

3. Christine Thisse and Bernard Thisse High−resolution *in situ* hybridization to whole−mount zebrafish embryos Nature Protocols 3, 59 − 69 (2008)

4. Barroso−Chinea, P.; Aymerich, M. S.; Castle, M. M.; PerezManso, M.; Tunon, T.; Erro, E.; Lanciego, J. L. Detection of Two Different mRNAs in a Single Section by Dual *in situ* Hybridization: A Comparison between Colorimetric and Fluorescent Detection. J. Neurosci. Methods 2007, 162, 119−128.

5. Bruno L, Muto A, Spadafora ND, Iaria D, Chiappetta A, Van Lijsebettens M, Bitonti MB. Multi−probe *in situ* hybridization to whole mount Arabidopsis seedlings. Int J Dev Biol. 2011;55(2):197−203.

6. Frdrique Rozier,Vincent Mirabet, Teva Vernoux, Pradeep Das Analysis of 3D gene expression patterns in plants using whole−mount RNA *in situ*hybridization Nature Protocols 9, 2464−2475 (2014)

7. Programmable *in situ* amplification for multiplexed imaging of mRNA expression H.M.T. Choi, J.Y. Chang, L.A. Trinh, J.E. Padilla, S.E. Fraser, and N.A. PierceNature Biotechnol, 28:1208−1212, 2010.

8. Next−generation *in situ* hybridization chain reaction: higher gain, lower cost, greater durability H.M.T. Choi, V.A. Beck, and N.A. PierceACS Nano, 8(5):4284−4294, 2014.

9. Chan, P. M.; Yuen, T.; Ruf, F.; Gonzalez−Maeso, J.; Sealfon, S. C. Method for Multiplex Cellular Detection of mRNAs Using Quantum Dot Fluorescent *in situ* Hybridization. Nucleic Acids Res. 2005, 33, e161.

10. Raj, A.; van den Bogaard, P.; Rifkin, S. A.; van Oudenaarden, A.; Tyagi, S. Imaging Individual mRNA Molecules Using Multiple Singly Labeled Probes. Nat. Methods 2008, 5, 877−879.

11. Palmeirim, I., Henrique, D., Ish Horowicz, D. and Pourqui, O. (1997). Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. Cell 91, 639−648.

12. Holley SA, Jlich D, Rauch GJ, Geisler R, Nsslein−Volhard C. her1 and the notch pathway function within the oscillator mechanism that regulates zebrafish somitogenesis.Development. 2002;129(5):1175−83.

13. Vermot J, Pourqui O. Retinoic acid coordinates somitogenesis and left−right patterning in vertebrate embryos. Nature. 2005 May 12;435(7039):215−20.

14. Clayton, D. F., Huecas, M. E., Sinclair−Thompson, E. Y., Nastiuk, K. L., Nottebohm, F. Probes for rare mRNAs reveal distributed cell subsets in canary brain. Neuron. 1, 249−261 (1988).

15. Chen CC, Wada K, Jarvis ED. Radioactive *in situ* hybridization for detecting diverse gene expression patterns in tissue.J Vis Exp. 2012;(62). pii: 3764.

16. Junker, Jan Philipp; Nol, Emily S.; Guryev, Victor; Peterson, Kevin A; Shah, Gopi; Huisken, Jan; McMahon, Andrew P.; Berezikov, Eugene; Bakkers, Jeroen; Oudenaarden, Alexander van Genome−wide RNA Tomography in the Zebrafish Embryo. Cell, 159, no. 3, pp. 662−675, (2014)

17. Triggered amplification by hybridization chain reaction R.M. Dirks and N.A. Pierce Proc Natl Acad Sci USA, 101(43): 15275−15278, 2004

18. Fuller, W. A. (1987). Measurement Error Models. Wiley, New York.

19. Kendall, M.G. and Stuart, A. (1979). The Advanced Theory of Statistics, Vol 2. Griffin, London

20. L.A. Trinh, et al., A versatile gene trap to visualize and interrogate the function of the vertebrate proteome, Genes and development, 25(21), 2306−20 (2011)

Chapter 3

Fitting a line through data with *measurement errors* and *equation errors* in both variables

3.1 Motivation

As outlined in Chapter 2, redundant detection of the same target mRNA via Hybridization Chain Reaction (HCR) provides a basis for quantifying relative abundance of mRNA copies within intact vertebrate embryos. However in an actual experiment, the situation is complicated by several factors. I discussed the *measurement errors* and estimated the statistics on background from independent control experiments. *Equation errors* that cause our model to deviate from the linear model were also introduced. In fact for any given probe set the actual number of probes bound to the target *in situ* will always be less than or equal to the number of probes, owing to the competition between mRNA secondary structures and other unknown factors. This leads to fluctuations in the hybridization yield, which is further affected by amplification. As a result the signal generated is a convolution of the biological distribution and the signal due to HCR. Consequently, *voxels* with more targets will have more such convolutions, leading to broader distribution for the signal. In other words *equation error* will be signal dependent, such that at every *voxel* the sampling comes from a distribution that is proportional to the signal intensity in that *voxel*. Such an error is called *heteroskedastic* [1,3], as opposed to *homescedastic* where the error comes from the same distribution at every data point. Particularly in our case, since the two channels

are symmetric in terms of image acquisition (data measurement), such a *heteroscedastic equation error* occurs in both the variables. Therefore we have derived a rigorous maximum likelihood estimate for the slope of line based on independent control experiments to measure the background statistics. We use the following notation throughout the chapter for deriving the estimates for slope in various scenarios depending upon the relationship between the errors and the signal :

x_i = signal in x (Green) channel (assuming no errors)
y_i = signal in y (Red) channel (assuming no errors)
u_i = background (*measurement error*) present in x (Green) channel
e_i = background (*measurement error*) present in y (Red) channel
h_i = heteroscedastic *equation error* present in x (Green) channel
g_i = heteroscedastic *equation error* present in y (Red) channel
X_i = signal measured via the image in x (Green) channel (with all errors)
Y_i = signal measured via the image in y (Red) channel (with all errors).

This chapter presents the analytical derivations for the estimates of slopes and their validation with the help of simulated data where the statistics are known *a priori*. Our experiment data has shown that both the signal intensities as well as auto-fluorescence (or other backgrounds) are normally distributed (Figure 3.1). The maximum likelihood estimates for the slopes have been derived taking this into account. In section 3.9 we briefly discuss a few cases where the observed data deviates from normal distribution.

3.2 Approach

In order to treat both the axes symmetrically , we assume that x_i are random rather than fixed. Hence our model is *structural* and not *functional*[*]. As outline before, since the two channels are symmetric with respect to their treatment, classical least square estimates for the slope of the line do not hold true as both the axes suffer from errors. Maximum

[*][1], Page 2

likelihood estimates aim at finding the most likely value of a parameter (slope, β, in our case), given the data that is observed. Typically a likelihood function is defined in terms of the desired parameters (θ) and is maximized to obtain the estimates of the parameter ($\hat{\theta}$). We wish to prove that the application of maximum likelihood estimates (MLE) to the normal structural model produces estimators of the unknown parameters that are consistent and asymptotically normal. In all of our cases we know the measurement errors (u_i and e_i) from independent control experiments.

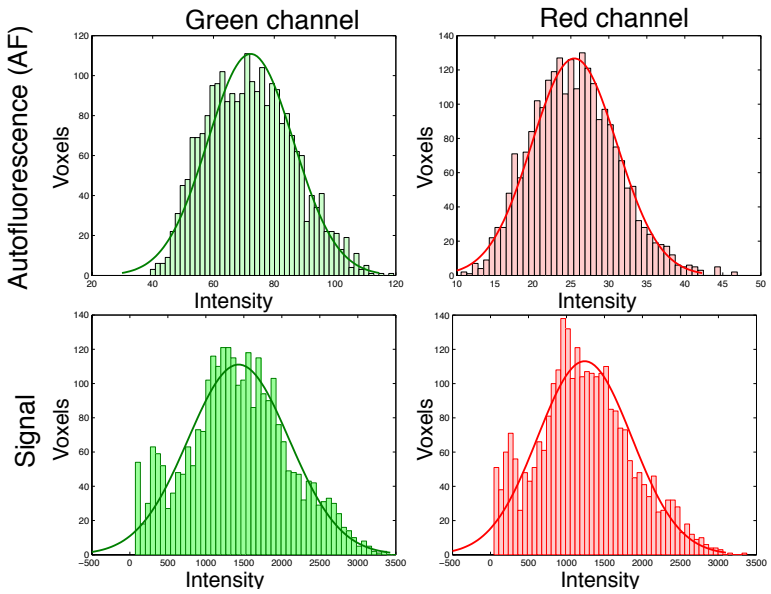


Figure 3.1: **Distribution of autofluorescence and signal for sub-cellular voxels** The autofluorescence and signal within sub-cellular *voxels* (5×5 pixels = $\sim$ 3×3 μm) are approximately normally distributed in both the channels.

Therefore under these assumptions, for a sample of n vectors $\mathbf{Z_i} = (Y_i, X_i)$, the log likelihood function can be written as follows[†]:

$$ln(\mathcal{L}(\theta)) \propto -(n-1)ln(|\mathbf{\Sigma_{ZZ}}|) - (n-1)tr(\mathbf{m_{ZZ}}\mathbf{\Sigma_{ZZ}^{-1}}) - n(\overline{\mathbf{Z}} - \mu_{\mathbf{Z}})\mathbf{\Sigma_{ZZ}^{-1}}(\overline{\mathbf{Z}} - \mu_{\mathbf{Z}})^T$$

where θ is the vector of all unknown parameters with respect to which the likelihood function $ln(\mathcal{L}(\theta))$ needs to be maximized. Based on the available population statistics we can write the estimates as follows :

[†][1], Page 105

$$\overline{\mathbf{Z}} = \begin{pmatrix} \overline{X} \\ \overline{Y} \end{pmatrix} = \begin{pmatrix} \frac{1}{n}\Sigma_{i=1}^{n} X_i \\ \frac{1}{n}\Sigma_{i=1}^{n} Y_i \end{pmatrix} \text{ and}$$

$$\mathbf{m_{ZZ}} = \begin{pmatrix} m_{XX} & m_{XY} \\ m_{XY} & m_{YY} \end{pmatrix} = \begin{pmatrix} \frac{1}{n}\Sigma_{i=1}^{n}(X_i - \overline{X})^2 & \frac{1}{n}\Sigma_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}) \\ \frac{1}{n}\Sigma_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}) & \frac{1}{n}\Sigma_{i=1}^{n}(Y_i - \overline{Y})^2 \end{pmatrix}$$

The estimator $\hat{\theta}$ of the parameters θ depend upon the nature of relationship between signal and the errors. In the most general case we can again write:

$$\mu_{\mathbf{Z}} = \mu_{\mathbf{Z}}(\theta) = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} = \begin{pmatrix} \mu_X(\theta) \\ \mu_Y(\theta) \end{pmatrix} \text{ and}$$

$$\mathbf{\Sigma_{ZZ}} = \mathbf{\Sigma_{ZZ}}(\theta) = \begin{pmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{XY} & \sigma_{YY} \end{pmatrix} = \begin{pmatrix} \sigma_{XX}(\theta) & \sigma_{XY}(\theta) \\ \sigma_{XY}(\theta) & \sigma_{YY}(\theta) \end{pmatrix}$$

Depending upon the number of parameters in θ and the number of elements in the vector of statistics $[\overline{\mathbf{Z}}, vech(\mathbf{m_{ZZ}})^T]$, following cases arise:

1. size(θ) = size($[\overline{\mathbf{Z}}, vech(\mathbf{m_{ZZ}})^T]$)

   $\hat{\mu_{\mathbf{Z}}} = \overline{\mathbf{Z}}$ and $\hat{\mathbf{\Sigma_{ZZ}}} = \mathbf{m_{ZZ}}$ maximize $ln(\mathcal{L}(\theta))$ with respect to $\mu_{\mathbf{Z}}$ and $\mathbf{\Sigma_{ZZ}}$ owing to the functional invariant property of MLE[‡]. Effectively MLE is equivalent to the method of moments in these cases.

2. size(θ) $\neq$ size($[\overline{\mathbf{Z}}, vech(\mathbf{m_{ZZ}})^T]$)

   In this case likelihood is maximized directly to obtain the required estimates [§]. Specific cases will be dealt in the following sections but it can be gereralized that :

   $tr(\mathbf{m_{ZZ}}\mathbf{\Sigma_{ZZ}^{-1}}) = \frac{\omega}{|\mathbf{\Sigma_{ZZ}}|}$

   where $\omega = m_{XX}\sigma_{YY} - 2m_{XY}\sigma_{XY} + m_{YY}\sigma_{XX}$ and

   $|\mathbf{\Sigma_{ZZ}}| = \sigma_{XX}\sigma_{YY} - \sigma_{XY}^2$

[‡][1], Page 105; [2], Page 14; [3], Chap 18

[§][2], Page 18; [4]

Therefore $ln(\mathcal{L}(\theta)) \propto -(n-1)ln(|\mathbf{\Sigma_{ZZ}}|) - (n-1)\frac{\omega}{|\mathbf{\Sigma_{ZZ}}|} - n(\overline{\mathbf{Z}} - \mu_{\mathbf{Z}})\mathbf{\Sigma_{ZZ}^{-1}}(\overline{\mathbf{Z}} - \mu_{\mathbf{Z}})^T$

It can be seen that for all values of $\theta, ln(\mathcal{L}(\theta))$ can be maximized when $\hat{\mu_{\mathbf{Z}}} = \overline{\mathbf{Z}}$ and $\frac{\partial ln(\mathcal{L}(\theta))}{\partial \theta}|_{\theta=\hat{\theta}} = 0$. Thus we obtain;

$$\boxed{\frac{\partial}{\partial \hat{\theta}}\{|\mathbf{\Sigma_{ZZ}}| + \omega\} = \frac{\omega}{|\mathbf{\Sigma_{ZZ}}|}\frac{\partial}{\partial \hat{\theta}}\{|\mathbf{\Sigma_{ZZ}}|\}}$$

It can be seen that evaluation of the above equation requires the knowledge of partial derivatives of the elements of $\mathbf{\Sigma_{ZZ}}$. Hence in of the cases that follow this section, we define and evaluate the matrix $\mathbf{D_{ZZ}}$ as follows :

$$\mathbf{D_{ZZ}} = \begin{bmatrix} \frac{\partial \sigma_{XX}}{\partial \theta_1} & \frac{\partial \sigma_{XY}}{\partial \theta_1} & \frac{\partial \sigma_{YY}}{\partial \theta_1} \\ \frac{\partial \sigma_{XX}}{\partial \theta_2} & \frac{\partial \sigma_{XY}}{\partial \theta_2} & \frac{\partial \sigma_{YY}}{\partial \theta_2} \\ \vdots & \vdots & \vdots \end{bmatrix}$$

and hence it follows
$$\frac{\partial \omega}{\partial \theta_k} = m_{YY}\frac{\partial \sigma_{XX}}{\partial \theta_k} - 2m_{XY}\frac{\partial \sigma_{XY}}{\partial \theta_k} + m_{XX}\frac{\partial \sigma_{YY}}{\partial \theta_k}$$
$$\frac{\partial |\mathbf{\Sigma_{ZZ}}|}{\partial \theta_k} = \sigma_{YY}\frac{\partial \sigma_{XX}}{\partial \theta_k} - 2\sigma_{XY}\frac{\partial \sigma_{XY}}{\partial \theta_k} + \sigma_{XX}\frac{\partial \sigma_{YY}}{\partial \theta_k}$$

It has been shown before[¶] that for asymptotic distribution of $\hat{\theta}$, the variance can be estimated as follows : $var(\hat{\theta}) = -1/(\frac{\partial^2 ln(\mathcal{L}(\theta))}{\partial \theta^2})|_{\theta=\hat{\theta}}$

3.3 Independent *measurement errors* with non-zero means in both axes

$$\begin{matrix} X_i = x_i + u_i \\ Y_i = \beta_0 + \beta_1 x_i + e_i \end{matrix} \quad ; \quad \begin{pmatrix} x_i \\ e_i \\ u_i \end{pmatrix} = \aleph \left[\begin{pmatrix} \mu_x \\ \mu_e \\ \mu_u \end{pmatrix}, \begin{pmatrix} \sigma_{xx} & 0 & 0 \\ 0 & \sigma_{ee} & 0 \\ 0 & 0 & \sigma_{uu} \end{pmatrix} \right]$$

Therefore we have

$$\mu_{\mathbf{Z}} = \begin{pmatrix} \mu_x + \mu_u \\ \beta_0 + \beta_1\mu_x + \mu_e \end{pmatrix};$$

$$\mathbf{\Sigma_{ZZ}} = \begin{pmatrix} \sigma_{xx} + \sigma_{uu} & \beta_1\sigma_{xx} \\ & \beta_1^2\sigma_{xx} + \sigma_{ee} \end{pmatrix}$$ where each entry in the matrix is calculated based on

[¶][3], Page 60

the expressions for (X_i, Y_i) defined above.

For example $\Sigma_{XX} = \sigma_{XX} = (n-1)^{-1}\Sigma_{i=1}^{i=n}(X_i - \bar{X})^2 = (n-1)^{-1}\Sigma_{i=1}^{i=n}(x_i + u_i - (\mu_x + \mu_u))^2$

$= (n-1)^{-1}\Sigma_{i=1}^{i=n}((x_i - \mu_x) + (u_i - \mu_u))^2$

$= (n-1)^{-1}\Sigma_{i=1}^{i=n}(x_i - \mu_x)^2 + (n-1)^{-1}\Sigma_{i=1}^{i=n}(u_i - \mu_u)^2 + 2(n-1)^{-1}\Sigma_{i=1}^{i=n}(x_i - \mu_x)(u_i - \mu_u)$

$= \sigma_{xx} + \sigma_{uu} + 2\sigma_{xu}$

As $\sigma_{xu} = 0$ in the present case , we get $\sigma_{XX} = \sigma_{xx} + \sigma_{uu}$

Similarly $\Sigma_{XY} = \sigma_{XY} = (n-1)^{-1}\Sigma_{i=1}^{i=n}(X_i - \bar{X})(Y_i - \bar{Y})$

$= (n-1)^{-1}\Sigma_{i=1}^{i=n}((x_i - \mu_x) + (u_i - \mu_u))(\beta_1(x_i - \mu_x) + (e_i - \mu_e))$

$= \beta_1(n-1)^{-1}\Sigma_{i=1}^{i=n}(x_i - \mu_x)^2 + (n-1)^{-1}\Sigma_{i=1}^{i=n}(x_i - \mu_x)(e_i - \mu_e) + \beta_1(n-1)^{-1}\Sigma_{i=1}^{i=n}(x_i - \mu_x)(u_i - \mu_u) + (n-1)^{-1}\Sigma_{i=1}^{i=n}(u_i - \mu_u)(e_i - \mu_e)$

Based on the nature of the correlation matrix, the last three terms on the right hand side of the above expression are zero. there fore we have :

$\sigma_{XY} = \beta_1(n-1)^{-1}\Sigma_{i=1}^{i=n}(x_i - \mu_x)^2 = \beta_1\sigma_{xx}$

Also $\Sigma_{YY} = \sigma_{YY} = (n-1)^{-1}\Sigma_{i=1}^{i=n}(Y_i - \bar{Y})^2$

$= (n-1)^{-1}\Sigma_{i=1}^{i=n}(\beta_1(x_i - \mu_x) + (e_i - \mu_e))^2$

$= \beta_1^2(n-1)^{-1}\Sigma_{i=1}^{i=n}(x_i - \mu_x)^2 + (n-1)^{-1}\Sigma_{i=1}^{i=n}(e_i - \mu_e)^2 + 2\beta_1(n-1)^{-1}\Sigma_{i=1}^{i=n}(x_i - \mu_x)(e_i - \mu_e)$

Again, since the last term is zero we get

$\sigma_{YY} = \beta_1^2(n-1)^{-1}\Sigma_{i=1}^{i=n}(x_i - \mu_x)^2 + (n-1)^{-1}\Sigma_{i=1}^{i=n}(e_i - \mu_e)^2 = \beta_1^2\sigma_{xx} + \sigma_{ee}$

Similar procedure is followed for the entries of $\mathbf{\Sigma_{ZZ}}$ in the sections to follow.

For the present case we know the values of $m_{XX}, m_{YY}, m_{XY}, \bar{X}$ and $\bar{Y}$ as measured by the sample with HCR signal and $\sigma_{uu}, \sigma_{ue}, \sigma_{ee}, \mu_u$ and μ_e as measured by the control experiment for estimating background. The remaining unknown parameters include $\beta_0, \beta_1, \sigma_{xx}$ and μ_x. Therefore the vector θ consisting of all the unknown parameters is : $\theta = \begin{bmatrix} \mu_x \\ \beta_0 \\ \beta_1 \\ \sigma_{xx} \end{bmatrix}$

By equating $\hat{\mu_{\mathbf{Z}}} = \overline{\mathbf{Z}}$ and $\hat{\mathbf{\Sigma_{ZZ}}} = \mathbf{m_{ZZ}}$ (using the method of moments) we have 5 possible equations to determine 4 entries of θ (i.e. size(θ) $\neq$ size($[\overline{\mathbf{Z}}, vech(\mathbf{m_{ZZ}})^T]$). Therefore for uniqueness of the parameters, likelihood function is maximized directly to obtain the required estimates (similar approach is adopted in all sections to follow). Working with the log likelihood function, we can write :

$$ln(\mathcal{L}(\theta)) \propto ln(\sigma_{ee}\sigma_{uu}) + ln(1 + \frac{\sigma_{xx}}{\sigma_{uu}} + \beta_1^2\frac{\sigma_{xx}}{\sigma_{ee}}) + \frac{m_{XX}(1+\beta_1^2\frac{\sigma_{xx}}{\sigma_{ee}})-2m_{XY}\beta_1\frac{\sigma_{xx}}{\sigma_{ee}}+m_{YY}(\frac{\sigma_{xx}}{\sigma_{ee}}+\frac{\sigma_{uu}}{\sigma_{ee}})}{\sigma_{uu}(1+\frac{\sigma_{xx}}{\sigma_{uu}}+\beta_1^2\frac{\sigma_{xx}}{\sigma_{ee}})}$$

Since the first term is constant, it can be dropped and we define $\frac{\sigma_{xx}}{\sigma_{ee}} = t$ and $\frac{\sigma_{ee}}{\sigma_{uu}} = \lambda$
Hence with the chance of variables the log likelihood function becomes

$$ln(\mathcal{L}(\theta)) \propto ln(1 + \lambda t + \beta_1^2 t) + \frac{m_{XX}(1+\beta_1^2 t)-2m_{XY}\beta_1 t+m_{YY}(t+\frac{1}{\lambda})}{\sigma_{uu}(1+\lambda t+\beta_1^2 t)}$$

Minimization with respect to t and β_1 respectively gives the following two equations
$(1+\lambda t+\beta_1^2 t)(\sigma_{uu}(\lambda+\beta_1^2)+\beta_1^2 m_{XX}-2\beta_1 m_{XY}+m_{YY}) = (\lambda+\beta_1^2)((1+\beta_1^2 t)m_{XX}-2\beta_1 m_{XY}t+ m_{YY}(t+\frac{1}{\lambda}))$
and
$(1 + \lambda t + \beta_1^2 t)(\sigma_{uu} + m_{XX} - \frac{m_{XY}}{\beta_1}) = ((1 + \beta_1^2 t)m_{XX} - 2\beta_1 m_{XY}t + m_{YY}(t + \frac{1}{\lambda}))$

Based on these two, we get the following expression
$m_{XX} - \frac{m_{XY}}{\beta_1} = \frac{\beta_1^2 m_{XX}-2\beta_1 m_{XY}+m_{YY}}{\lambda+\beta_1^2}$
This upon simplification gives
$m_{XY}\beta_1^2 - (m_{YY} - \lambda m_{XX})\beta_1 - m_{XY}\lambda = 0$

$\boxed{\Delta^2 = (m_{YY} - \lambda m_{XX})^2 + 4m_{XY}^2\lambda}$ where Δ^2 is the discriminant of the quadratic equation in β_1 above,

$\boxed{\hat{\beta_1} = \frac{(m_{YY}-\lambda m_{XX})+\Delta}{2m_{XY}}}$ where $\hat{\beta_1}$ is the estimated slope,

Similarily, we use the equations obtained after differentiating the maximum likelihood function to obtain

$t[\lambda(\sigma_{uu} + m_{XX} - \frac{m_{XY}}{\beta_1}) + \beta_1^2\sigma_{uu} + \beta_1 m_{XY} - m_{YY}] = \frac{m_{YY}}{\lambda} - \sigma_{uu} + \frac{m_{XY}}{\beta_1}$

$\Rightarrow t[\beta_1\sigma_{uu}(\lambda + \beta_1^2) + (m_{XY}\beta_1^2 - (m_{YY} - \lambda m_{XX})\beta_1 - m_{XY}\lambda)] = \frac{m_{YY}}{\lambda} - \sigma_{uu} + \frac{m_{XY}}{\beta_1}$

$\Rightarrow \beta_1\sigma_{uu}t(\lambda + \beta_1^2) = \frac{m_{YY}}{\lambda} - \sigma_{uu} + \frac{m_{XY}}{\beta_1}$

$\Rightarrow \hat{t} = \frac{\hat{\beta_1}(m_{YY} - \sigma_{ee}) + \lambda m_{XY}}{\hat{\beta_1}(\lambda + \hat{\beta_1}^2)\sigma_{ee}}$

$\boxed{\hat{\sigma_{xx}} = \frac{\hat{\beta_1}(m_{YY} - \sigma_{ee}) + \lambda m_{XY}}{\hat{\beta_1}(\lambda + \hat{\beta_1}^2)}}$ where $\hat{\sigma_{xx}}$ is the estimated variance of the signal in x-channel.

The estimates of the means can also be obtained from the population mean ($\hat{\mu_{\mathbf{Z}}} = \overline{\mathbf{Z}}$) as follows : $\boxed{\hat{\mu_x} = \bar{X} - \mu_u \text{ and } \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1(\bar{X} - \mu_u) - \mu_e}$

For estimating variance in $\hat{\beta}_1$, we know that $var(\hat{\beta}_1) = -1/(\frac{\partial^2 ln(\mathcal{L}(\beta_1))}{\partial\beta_1^2})|_{\beta_1 = \hat{\beta_1}}$

Therefore $\frac{\partial^2 ln(\mathcal{L}(\beta_1))}{\partial\beta_1^2} = \frac{-2(2\beta_1 t)^2}{\sigma_{uu}(1+\lambda t+\beta_1^2 t)^2}[\frac{\sigma_{uu}}{2} + m_{XX} - \frac{m_{XY}}{\beta_1} - \frac{((1+\beta_1^2 t)m_{XX} - 2\beta_1 m_{XY}t + m_{YY}(t+\frac{1}{\lambda}))}{(1+\lambda t+\beta_1^2 t)}] +$
$\frac{2t}{\sigma_{uu}(1+\lambda t+\beta_1^2 t)}[\sigma_{uu} + m_{XX} - \frac{((1+\beta_1^2 t)m_{XX} - 2\beta_1 m_{XY}t + m_{YY}(t+\frac{1}{\lambda}))}{(1+\lambda t+\beta_1^2 t)}]$

For $\beta_1 = \hat{\beta}_1$ we know that $\sigma_{uu} + m_{XX} - \frac{m_{XY}}{\hat{\beta_1}} - \frac{((1+\hat{\beta_1}^2 t)m_{XX} - 2\hat{\beta_1} m_{XY}t + m_{YY}(t+\frac{1}{\lambda}))}{(1+\lambda t+\hat{\beta_1}^2 t)} = 0$;

Therefore upon simplification we get

$\frac{\partial^2 ln(\mathcal{L}(\beta_1))}{\partial\beta_1^2}|_{\beta_1 = \hat{\beta_1}} = \frac{(2\hat{\beta_1} t)^2}{(1+\lambda t+\hat{\beta_1}^2 t)^2} + \frac{2t}{\sigma_{uu}(1+\lambda t+\hat{\beta_1}^2 t)}[\frac{m_{XY}}{\hat{\beta_1}}]$

$var(\hat{\beta}_1) = \frac{\sigma_{uu}(1+\lambda t+\hat{\beta_1}^2 t)^2}{\sigma_{uu}(2\hat{\beta_1} t)^2 + 2t(1+\lambda t+\hat{\beta_1}^2 t)[\frac{m_{XY}}{\hat{\beta_1}}]}$

For estimating the true values of signal $((\hat{y_i}, \hat{x_i}))$ from the observed $Z_i = (Y_i, X_i)$ under the structural model (i.e. random x_i), using an estimator of β_1 ($\hat{\beta}_1$) we know that[‖] :

$(\hat{y_i}, \hat{x_i}) = (Y_i - \mu_e, X_i - \mu_u) - (Y_i - \hat{\beta}_1 X_i)[(1, -\hat{\beta_1})\Sigma_{\epsilon\epsilon}(1, -\hat{\beta_1})^T]^{-1}(1, -\hat{\beta_1})\Sigma_{\epsilon\epsilon}$

The estimation of the true signal is obtained by subtracting a predictor of (e_i, u_i) from the observed signal (Y_i, X_i), which is constructed using the observed signal and the statistics of the background measured independently. For the present case, the above expression simplifies to $(\hat{y_i}, \hat{x_i}) = (Y_i - \mu_e, X_i - \mu_u) - \frac{Y_i - \hat{\beta_1} X_i}{\hat{\beta_1}^2\sigma_{uu} + \sigma_{ee}}(\sigma_{ee}, -\hat{\beta_1}\sigma_{uu})$

[‖] [1], Page 140

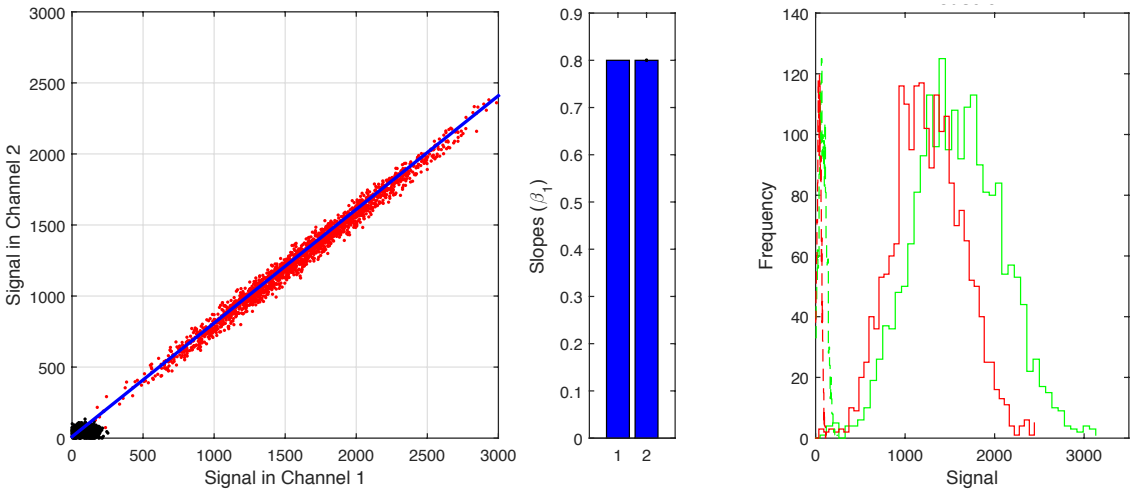


Figure 3.2: **Slope estimates for independent measurement errors** ($\sigma_{eu} = 0$) **in both axes**
(Left) Plots of simulated voxel intensities in two channels. For each fit, sample size is 2000 voxels. Red dots denote simulated signal and black represent the uncorrelated background. The values of statistics assumed were close to those measured in experiments ($\mu_x = 1500, \sigma_{xx} = 500^2, \mu_u = 100, \sigma_{uu} = 50^2, \mu_e = 75, \sigma_{ee} = 25^2$). (Middle) Comparison of the estimated slope (right bar) with its known value (left bar). Error bar on the right bar denotes estimated standard error. (Right) The distributions used for simulation. Solid red and green lines denote the observed signal (X_i, Y_i) in red and green channels respectively. The background distribution (u_i, e_i) in respective channels is denoted by dashed lines.

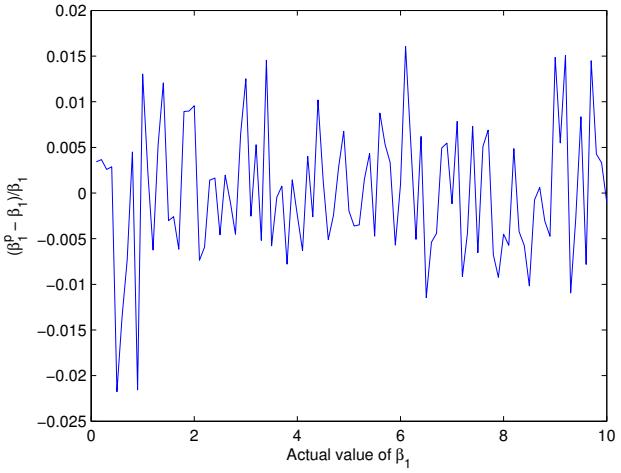


Figure 3.3: **Fractional error in the estimated value of slope for independent measurement errors** ($\sigma_{eu} = 0$) **in both axes**
Simulations were performed for known values of slopes (β_1; x-axis of the plot) The y-axis denotes the fractional error in the slope estimates using maximum likelihood method.

3.4 Correlated *measurement errors* (but independent of signal) with non-zero means in both axes

$$\begin{matrix} X_i = x_i + u_i \\ Y_i = \beta_0 + \beta_1 x_i + e_i \end{matrix} \; ; \; \begin{pmatrix} x_i \\ e_i \\ u_i \end{pmatrix} = \aleph \left[\begin{pmatrix} \mu_x \\ \mu_e \\ \mu_u \end{pmatrix}, \begin{pmatrix} \sigma_{xx} & 0 & 0 \\ 0 & \sigma_{ee} & \sigma_{eu} \\ 0 & \sigma_{eu} & \sigma_{uu} \end{pmatrix} \right]$$

Therefore we have

$$\mu_{\mathbf{Z}} = \begin{pmatrix} \mu_x + \mu_u \\ \beta_0 + \beta_1 \mu_x + \mu_e \end{pmatrix};$$

$$\mathbf{\Sigma_{ZZ}} = \begin{pmatrix} \sigma_{xx} + \sigma_{uu} & \beta_1 \sigma_{xx} + \sigma_{eu} \\ & \beta_1^2 \sigma_{xx} + \sigma_{ee} \end{pmatrix}$$

It can therefore be seen that $\theta = \begin{bmatrix} \mu_x \\ \beta_0 \\ \beta_1 \\ \sigma_{xx} \end{bmatrix}$

Here again we have 4 variables (since we know σ_{eu} by experiment) and 5 possible equations (i.e. size$(\theta) \neq$ size$([\overline{\mathbf{Z}}, vech(\mathbf{m_{ZZ}})^T])$ Therefore for uniqueness of the parameters likelihood is maximized directly to obtain the required estimates. Similar to the previous case we work with the log likelihood function, and define $\frac{\sigma_{xx}}{\sigma_{ee}} = t$; $\frac{\sigma_{ee}}{\sigma_{uu}} = \lambda$ and $\frac{\sigma_{eu}}{\sigma_{uu}} = \rho$

Hence with the change of variables the log likelihood function becomes

$ln(\mathcal{L}(\theta)) \propto ln(1 + \lambda t + \beta_1^2 t - \frac{\rho^2}{\lambda} - 2\beta_1 \rho t) + \frac{m_{XX}(1+\beta_1^2 t) - 2m_{XY}(\beta_1 t + \frac{\rho}{\lambda}) + m_{YY}(t + \frac{1}{\lambda})}{\sigma_{uu}(1 + \lambda t + \beta_1^2 t - \frac{\rho^2}{\lambda} - 2\beta_1 \rho t)}$

Minimization, as before, with respect to t and β_1 respectively gives the following two equations

$(1 + \lambda t + \beta_1^2 t - \frac{\rho^2}{\lambda} - 2\beta_1 \rho t)(\sigma_{uu}(\lambda + \beta_1^2 - 2\beta_1 \rho) + \beta_1^2 m_{XX} - 2\beta_1 m_{XY} + m_{YY}) = (\lambda + \beta_1^2 - 2\beta_1 \rho)((1 + \beta_1^2 t)m_{XX} - 2\beta_1 m_{XY} t + m_{YY}(t + \frac{1}{\lambda}) - 2m_{XY}\frac{\rho}{\lambda})$

and

$(1+\lambda t+\beta_1^2 t-\frac{\rho^2}{\lambda}-2\beta_1\rho t)(\sigma_{uu}(1-\frac{\rho}{\beta_1})+m_{XX}-\frac{m_{XY}}{\beta_1}) = (1-\frac{\rho}{\beta_1})((1+\beta_1^2 t)m_{XX}-2\beta_1 m_{XY}t+ m_{YY}(t+\frac{1}{\lambda})-2m_{XY}\frac{\rho}{\lambda})$

Based on these two ,we get the following expression

$(m_{XY}-\rho m_{XX})\beta_1^2-(m_{YY}-\lambda m_{XX})\beta_1-(m_{XY}\lambda-\rho m_{YY})=0$

$\boxed{\Delta^2=(m_{YY}-\lambda m_{XX})^2+4(m_{XY}-\rho m_{XX})(m_{XY}\lambda-\rho m_{YY})}$ where Δ^2 is the discriminant of the quadratic equation in β_1 above,

$\boxed{\hat{\beta_1}=\frac{(m_{YY}-\lambda m_{XX})+\Delta}{2(m_{XY}-\rho m_{XX})}}$ where $\hat{\beta_1}$ is the estimated slope.

Similarly

$\boxed{\hat{\sigma_{xx}}=\frac{\hat{\beta_1}(m_{YY}-\sigma_{ee})+(m_{XY}\lambda-\rho m_{YY})}{\hat{\beta_1}(\lambda+\hat{\beta_1}^2)}}$ where $\hat{\sigma_{xx}}$ is the estimated variance of the signal in x-channel.

The estimates of the means are : $\boxed{\hat{\mu_x}=\bar{X}-\mu_u \text{ and } \hat{\beta}_0=\bar{Y}-\hat{\beta}_1(\bar{X}-\mu_u)-\mu_e}$

Following similar approach as before we obtain

$var(\hat{\beta}_1)=\frac{\sigma_{uu}(1+\lambda t+\hat{\beta_1}^2 t-\frac{\rho^2}{\lambda}-2\hat{\beta_1}\rho t)^2}{\sigma_{uu}(2\hat{\beta_1}t)^2+2t(1+\lambda t+\hat{\beta_1}^2 t-\frac{\rho^2}{\lambda}-2\hat{\beta_1}\rho t)[\frac{m_{XY}}{\beta_1}]}$

It can be seen that in the situation where $\rho=0$, the expression reduces to the one in the previous case with uncorrelated errors. Once again for estimating the true values of signal $((\hat{y_i},\hat{x_i}))$ from the observed $Z_i=(Y_i,X_i)$ we follow similar approach as before to obtain

$(\hat{y_i},\hat{x_i})=(Y_i-\mu_e,X_i-\mu_u)-\frac{Y_i-\hat{\beta_1}X_i}{\hat{\beta_1}^2\sigma_{uu}+\sigma_{ee}-2\hat{\beta_1}\sigma_{eu}}(\sigma_{ee}-\hat{\beta}_1\sigma_{eu},\sigma_{eu}-\hat{\beta}_1\sigma_{uu})$

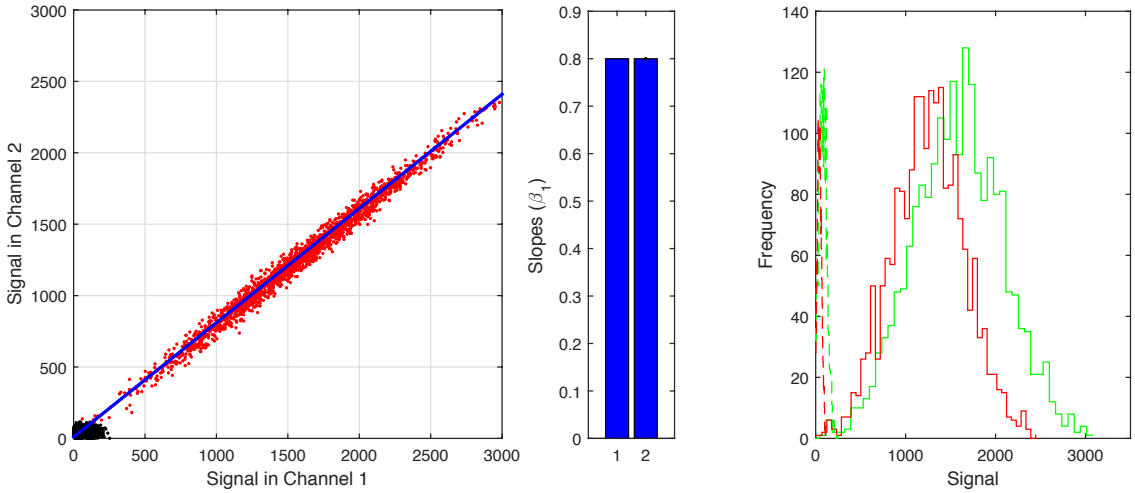


Figure 3.4: **Slope estimates for correlated measurement errors** ($\sigma_{eu} \neq 0$) **in both axes**
(Left) Plots of simulated voxel intensities in two channels. For each fit, sample size is 2000 voxels. Red dots denote simulated signal and black represent the correlated background. The values of statistics assumed were close to those measured in experiments ($\mu_x = 1500, \sigma_{xx} = 500^2, \mu_u = 100, \sigma_{uu} = 50^2, \mu_e = 75, \sigma_{ee} = 25^2, \sigma_{eu} = 20^2$). (Middle) Comparison of the estimated slope (right bar) with its known value (left bar). Error bar on the right bar denotes estimated standard error. (Right) The distributions used for simulation. Solid red and green lines denote the observed signal (X_i, Y_i) in red and green channels respectively. The background distribution (u_i, e_i) in respective channels is denoted by dashed lines.

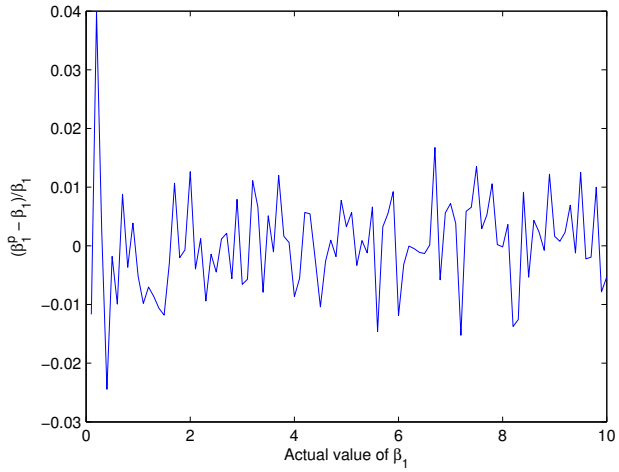


Figure 3.5: **Fractional error in the estimated value of slope for correlated *measurement errors*** ($\sigma_{eu} \neq 0$) **in both axes**
Simulations were performed for known values of slopes (β_1; x-axis of the plot) The y-axis denotes the fractional error in the slope estimates using maximum likelihood method.

3.5 *Measurement errors* in both axes correlated with each other and with signal

$$\begin{matrix} X_i = x_i + u_i \\ Y_i = \beta_0 + \beta_1 x_i + e_i \end{matrix} \quad ; \quad \begin{pmatrix} x_i \\ e_i \\ u_i \end{pmatrix} = \aleph \left[\begin{pmatrix} \mu_x \\ \mu_e \\ \mu_u \end{pmatrix}, \begin{pmatrix} \sigma_{xx} & \sigma_{xe} & \sigma_{xu} \\ \sigma_{xe} & \sigma_{ee} & \sigma_{eu} \\ \sigma_{xu} & \sigma_{eu} & \sigma_{uu} \end{pmatrix} \right]$$

Therefore we have

$$\mu_{\mathbf{Z}} = \begin{pmatrix} \mu_x + \mu_u \\ \beta_0 + \beta_1 \mu_x + \mu_e \end{pmatrix};$$

$$\mathbf{\Sigma_{ZZ}} = \begin{pmatrix} \sigma_{xx} + \sigma_{uu} + 2\sigma_{xu} & \beta_1 \sigma_{xx} + \sigma_{eu} + \sigma_{xe} + \beta_1 \sigma_{xu} \\ & \beta_1^2 \sigma_{xx} + \sigma_{ee} + 2\beta_1 \sigma_{xe} \end{pmatrix}$$

Additionally we also have two measurements of m_{Xe} and m_{Yu} available from experiments as follows :

$m_{Xe} = \sigma_{xe} + \sigma_{eu}$

$m_{Yu} = \beta_1 \sigma_{xu} + \sigma_{eu}$

It can therefore be seen that $\theta = \begin{bmatrix} \mu_x \\ \beta_0 \\ \beta_1 \\ \sigma_{xx} \\ \sigma_{xe} \\ \sigma_{xu} \end{bmatrix}$

Now we have 6 variables and 7 possible relations. Therefore for uniqueness of the parameters likelihood is maximized directly to obtain the required estimates. Similar to the situations before we work with the log likelihood function, and minimize it with respect to σ_{xx} and β_1 respectively gives the following two equations after defining the following terms (which are known experimentally) :

$s_{xe} = m_{Xe} - \sigma_{eu}$

$s_{xu} = m_{Yu} - \sigma_{eu}$

And with the same definition of $\frac{\omega}{|\mathbf{\Sigma_{ZZ}}|}$ as before , we get

$(\sigma_{ee} - 2\beta_1\sigma_{eu} + \beta_1^2\sigma_{uu})(\frac{\omega}{|\mathbf{\Sigma_{ZZ}}|} - 1) = m_{YY} - 2m_{XY}\beta_1 + m_{XX}\beta_1^2$

and

$(\sigma_{xx}\sigma_{eu} - \sigma_{uu}(s_{xe} + \beta_1\sigma_{xx}) + \frac{s_{xu}}{\beta_1}(2s_{xe} + \frac{\sigma_{ee}}{\beta_1}))(\frac{\omega}{|\mathbf{\Sigma_{ZZ}}|} - 1) = \frac{m_{YY}s_{xu}}{\beta_1^2} - m_{XX}(\beta_1\sigma_{xx} + s_{xe}) + m_{XY}\sigma_{xx}$

Using $\frac{\sigma_{xx}}{\sigma_{ee}} = t$; $\frac{\sigma_{ee}}{\sigma_{uu}} = \lambda$ and $\frac{\sigma_{eu}}{\sigma_{uu}} = \rho$

$|\mathbf{\Sigma_{ZZ}}|(\sigma_{uu}(\lambda + \beta_1^2 - 2\beta_1\rho) + \beta_1^2 m_{XX} - 2\beta_1 m_{XY} + m_{YY}) = (\lambda + \beta_1^2 - 2\beta_1\rho)(\omega\sigma_{uu})$

and

$(\frac{\beta_1\rho t}{\lambda} - (\frac{s_{xe}\beta_1}{\sigma_{uu}\lambda} + \beta_1^2 t) + \frac{s_{xu}}{\lambda\sigma_{xx}^2}(2s_{xe} + \frac{\sigma_{ee}}{\beta_1}))(\frac{\omega}{|\mathbf{\Sigma_{ZZ}}|} - 1) = \frac{1}{\sigma_{uu}}((\frac{s_{xe}\beta_1}{\lambda\sigma_{uu}} + \beta_1^2 t)m_{XX} - \beta_1 m_{XY}t - \frac{m_{YY}s_{xu}}{\beta_1\lambda\sigma_{xx}})$

It can be seen that upon division, unlike the previous cases we will not get a quadratic equation in β_1 and this is due to non-zero terms s_{xe} and s_{xu}. Therefore analytical solution is not possible in this case and hence we can approach this case by numerically solving the two equations simultaneously within MATLAB. However in our experiments we have seen that the correlation of the background to signal is negligible as compared to the correlation between signals (Figure 3.6). Therefore for the real experimental data we ignore these correlations. Using simulated data we show that owing to correlation between signal dominating over other cross correlation terms (i.e.$\sigma_{XY} \gg \sigma_{Xe}, \sigma_{Yu}$), our estimated slope remains unaffected by neglecting the terms m_{Xe} and m_{Yu}. It can be seen that these equations reduce to the previous case when $s_{xe} = s_{xu} = -\sigma_{eu}$ implying $m_{Xe} = m_{Yu} = 0$.

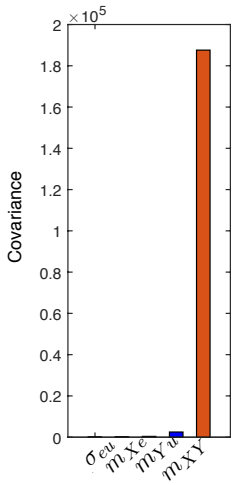


Figure 3.6: **Comparison of correlations ($\sigma_{eu}, \sigma_{Xe}, \sigma_{Yu}$ and σ_{XY})**
Covariances estimated in a real experiment based on background and signal (using *citrine*) from 4 types of samples: signal in both channels, signal in x and background in y, background in x and signal in y and background in both the channels. The correlation of background to the signal is negligible and is therefore ignored in estimating slope of the line from observed data.

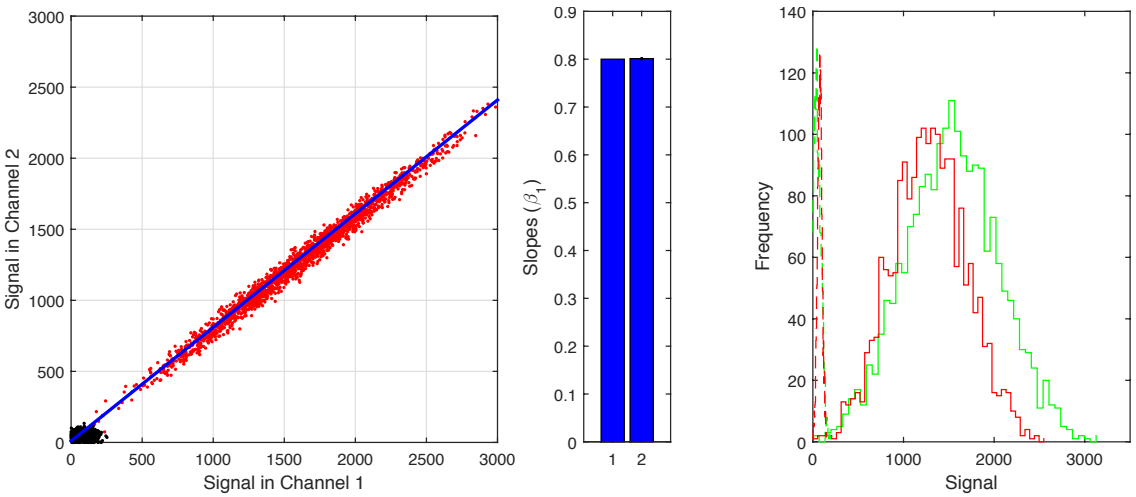


Figure 3.7: **Slope estimates for measurement errors ($\sigma_{eu} \neq 0$) correlated with signal in both axes ($\sigma_{Xe}, \sigma_{Yu} \neq 0$)**
(Left) Plots of simulated voxel intensities in two channels. For each fit, sample size is 2000 voxels. Red dots denote simulated signal and black represent the correlated background. The values of statistics assumed were close to those measured in experiments ($\mu_x = 1500, \sigma_{xx} = 500^2, \mu_u = 100, \sigma_{uu} = 50^2, \mu_e = 75, \sigma_{ee} = 25^2, \sigma_{eu} = 20^2, \sigma_{Xe} = 20^2, \sigma_{Yu} = 100^2$). (Middle) Comparison of the estimated slope (right bar) with its known value (left bar). Error bar on the right bar denotes estimated standard error. (Right) The distributions used for simulation. Solid red and green lines denote the observed signal (X_i, Y_i) in red and green channels respectively. The background distribution (u_i, e_i) in respective channels is denoted by dashed lines.

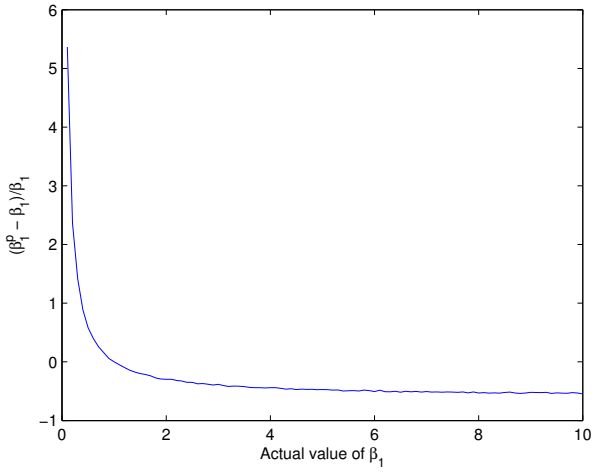


Figure 3.8: **Fractional error in the estimated value of slope for measurement errors** ($\sigma_{eu} \neq 0$) **correlated with signal in both axes** ($\sigma_{Xe}, \sigma_{Yu} \neq 0$)
Simulations were performed for known values of slopes (β_1; x-axis of the plot) The y-axis denotes the fractional error in the slopes estimated by assuming ($\sigma_{Xe}, \sigma_{Yu} = 0$). Such a trend is observed given that with higher values of slopes, correlation between signal dominates over other cross correlation terms (i.e.$\sigma_{XY} \gg \sigma_{Xe}, \sigma_{Yu}$)

3.6 Independent *measurement errors* with *equation errors* in both axes

$$\begin{aligned} X_i &= x_i + u_i + p_i \\ Y_i &= \beta_0 + \beta_1 x_i + e_i + q_i \end{aligned} \quad ; \quad \begin{pmatrix} x_i \\ e_i \\ u_i \\ p_i \\ q_i \end{pmatrix} = \aleph \left[\begin{pmatrix} \mu_x \\ \mu_e \\ \mu_u \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{xx} & 0 & 0 & 0 & 0 \\ 0 & \sigma_{ee} & 0 & 0 & 0 \\ 0 & 0 & \sigma_{uu} & 0 & 0 \\ 0 & 0 & 0 & \sigma_{pp} = \alpha_1 x_i & 0 \\ 0 & 0 & 0 & 0 & \sigma_{qq} = \alpha_2 y_i \end{pmatrix} \right]$$

Here we have assumed the variance in the *heteroskedastic equation error* to be a linear function of the signal ($\sigma_{pp} = \alpha_1 x_i, \sigma_{qq} = \alpha_2 y_i$) . This is based on the argument, mentioned in section 3.1, that the signal generated is a convolution of the biological distribution and the signal due to HCR (probe binding and amplification). Consequently, voxels with more targets will have more such convolutions. Conceptualizing signal generation from every target molecule as independent random events, the sum total of signal in a *voxel* will be normally distributed when all the target molecules in that *voxel* are considered (by central limit theorem). Consequently the average value of these events in a *voxel* (i.e. the signal

in that *voxel*) will come from a distribution whose variance will be proportional to the signal in the *voxel*.In other words *equation error* will be signal dependent, such that this dependence is linear in nature. This forms the basis for our choice of the linear functional form for the *heterskedastic equation error*.Therefore we have

$$\mu_{\mathbf{Z}} = \begin{pmatrix} \mu_x + \mu_u \\ \beta_0 + \beta_1\mu_x + \mu_e \end{pmatrix};$$

$$\mathbf{\Sigma_{ZZ}} = \begin{pmatrix} \sigma_{xx} + \sigma_{uu} + \sigma_{pp} & \beta_1\sigma_{xx} \\ & \beta_1^2\sigma_{xx} + \sigma_{ee} + \sigma_{qq} \end{pmatrix}$$

It can therefore be seen that $\theta = \begin{bmatrix} \mu_x \\ \beta_0 \\ \beta_1 \\ \sigma_{xx} \\ \alpha_1 \\ \alpha_2 \end{bmatrix}$

Here we have 6 variables and 5 possible equations . However this situation is different from previous cases where the number of equations were more than the number of variables. Hence we need at least one more piece of information to estimate the parameters uniquely. In the absence on any measures of α_1 and α_2, we assume that their ratio α (i.e. $\alpha = \frac{\alpha_1}{\alpha_2}$ is a parameter to be estimated. Our approach now is to assume different values of α and to show with numerical simulations for different values of α that the estimated slope $\hat{\beta_1}$ is insensitive to α (Figure 3.12). Once we assume a value for α (we assume $\alpha = 1$, for all estimations), we have 6 variables and 6 equations which takes us to the case where MLE is equivalent to the method of moments owing to the functional invariant property of MLE, as pointed out before. Now we solve the set of simultaneous equations to derive :

$$\alpha m_{XY}\beta_1^2 - (\alpha(m_{YY} - \sigma_{ee}) - (m_{XX} - \sigma_{uu}))\beta_1 - m_{XY} = 0$$

$\boxed{\Delta^2 = (\alpha(m_{YY} - \sigma_{ee}) - (m_{XX} - \sigma_{uu}))^2 + 4\alpha m_{XY}^2}$ where Δ^2 is the discriminant of the quadratic

equation in β_1 above,

$\boxed{\hat{\beta}_1 = \frac{\alpha(m_{YY}-\sigma_{ee})-(m_{XX}-\sigma_{uu})+\Delta}{2\alpha m_{XY}}}$ where $\hat{\beta}_1$ is the estimated slope, and

$\boxed{\hat{\sigma_{xx}} = \frac{m_{XY}}{\hat{\beta}_1}}$ where $\hat{\sigma_{xx}}$ is the estimated variance of the signal in x-channel.

The estimates of the means are : $\boxed{\hat{\mu_x} = \bar{X} - \mu_u \text{ and } \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1(\bar{X} - \mu_u) - \mu_e}$

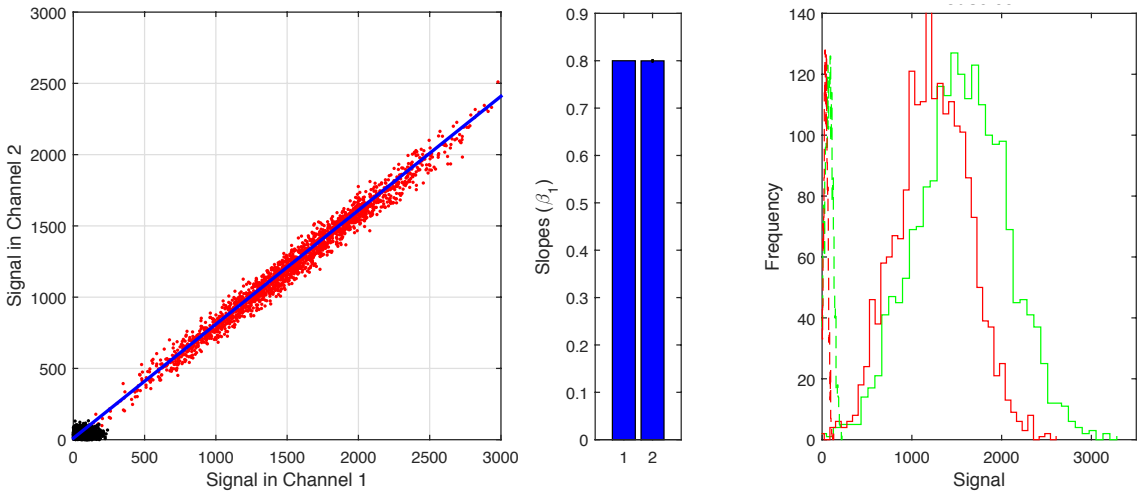


Figure 3.9: **Slope estimates for independent measurement errors** ($\sigma_{eu} = 0$) **and heteroscedastic *equation errors*** (p_i, q_i) **in both axes**
(Left) Plots of simulated voxel intensities in two channels. For each fit, sample size is 2000 voxels. Red dots denote simulated signal and black dots represent the uncorrelated background. The values of statistics assumed were close to those measured in experiments ($\mu_x = 1500, \sigma_{xx} = 500^2, \mu_u = 100, \sigma_{uu} = 50^2, \mu_e = 75, \sigma_{ee} = 25^2$). (Middle) Comparison of the estimated slope (right bar) with its known value (left bar). Error bar on the right bar denotes estimated standard error. (Right) The distributions used for simulation. Solid red and green lines denote the observed signal (X_i, Y_i) in red and green channels respectively. The background distribution (u_i, e_i) in respective channels is denoted by dashed lines.

Similar to the approach used in section 3.3, we can derive the variance in $\hat{\beta}_1$ as follows :

$$\frac{\partial^2 ln(\mathcal{L}(\beta_1))}{\partial \beta_1^2} = \frac{-4[(2m_{XY}(m_{XX}-\sigma_{xx}))^2+(m_{XX}(\beta_1^2\sigma_{xx}+\sigma_{ee}+\sigma_{qq})-2m_{XY}\beta_1\sigma_{xx}+m_{YY}(\sigma_{xx}+\sigma_{uu}+\sigma_{pp}))\sigma_{xx}^2]}{(m_{XX}(\beta_1^2\sigma_{xx}+\sigma_{ee}+\sigma_{qq})-2m_{XY}\beta_1\sigma_{xx}+m_{YY}(\sigma_{xx}+\sigma_{uu}+\sigma_{pp}))^2}$$

For $\beta_1 = \hat{\beta}_1$ we can simplify the above expression as

$$\frac{\partial^2 ln(\mathcal{L}(\beta_1))}{\partial \beta_1^2} = \frac{-2[2m_{XY}^2(m_{XX}-\hat{\sigma_{xx}})^2+\hat{\sigma_{xx}}^2(m_{XX}m_{YY}-m_{XY}^2)]}{m_{XX}m_{YY}-m_{XY}^2}$$

Therefore upon simplification we get

$$\frac{\partial^2 ln(\mathcal{L}(\beta_1))}{\partial \beta_1^2}\Big|_{\beta_1=\hat{\beta}_1} = \frac{-2m_{XY}[2\hat{\beta_1}^2 m_{XX}^2+m_{XX}m_{YY}-4\beta_1 m_{XX}m_{XY}+m_{XY}^2]}{\hat{\beta_1}^2(m_{XX}m_{YY}-m_{XY}^2)}$$

$var(\hat{\beta}_1) = \frac{\hat{\beta_1}^2(m_{XX}m_{YY}-m_{XY}^2)}{2m_{XY}[2\hat{\beta_1}^2 m_{XX}^2+m_{XX}m_{YY}-4\beta_1 m_{XX}m_{XY}+m_{XY}^2]}$

For estimating the true values of signal $((\hat{y_i}, \hat{x_i}))$ from the observed $Z_i = (Y_i, X_i)$ under the structural model (i.e. random x_i), with the assumption that the covariance matrix of the background is known (through an independent control experiment), we know that** : $(\hat{y_i}, \hat{x_i}) = (\bar{Y} - \mu_e, \bar{X} - \mu_u) + (Y_i - \bar{Y}, X_i - \bar{X})(I - \frac{n-4}{n-1}\mathbf{m_{ZZ}}^{-1}\Sigma_{\epsilon\epsilon})$, where n is the total number of voxels in an embryo (typically ≥ 400). Similar to the approaches used before, the estimation of the true signal is obtained again by subtracting a predictor of (e_i, u_i) from the observed signal (Y_i, X_i), which is constructed using the observed signal and the statistics of the background measured independently. Rearranging the terms in the above expression gives us :

$(\hat{y_i}, \hat{x_i}) = (Y_i - \mu_e, X_i - \mu_u) + \frac{n-4}{n-1}(Y_i - \bar{Y}, X_i - \bar{X})\mathbf{m_{ZZ}}^{-1}\Sigma_{\epsilon\epsilon}$

Unlike the previous three cases where only *measurement errors* were present , in the present case (and the next two to follow) we also have *equation errors*. However under the assumption that *measurement errors* and *equation errors* are independent (i.e. $\sigma_{up} = \sigma_{uq} = \sigma_{ep} = \sigma_{eq} = 0$), the covariances between the total errors in two axes $(u_i + p_i, e_i + q_i)$ remain equal to the covariances between the *measurement errors*. Thus through independent control experiments the background statistics can be known again and signal estimation can be done. This provides an equivalence between the expressions used in sections 3.6 through 3.8 to sections 3.3 through 3.5, the exact forms of which will differ depending upon the measured statistics.

For the present case, the above expression simplifies to

$$(\hat{y_i}, \hat{x_i}) = (\bar{Y} - \mu_e, \bar{X} - \mu_u) + \frac{1}{m_{YY}m_{XX}-m_{XY}^2}\begin{pmatrix}(Y_i - \bar{Y})[m_{XX}(m_{YY} - \frac{n-4}{n-1}\sigma_{ee}) - m_{XY}^2] + \frac{n-4}{n-1}(X_i - \bar{X})[m_{XY}\sigma_{ee}] \\ (X_i - \bar{X})[m_{YY}(m_{XX} - \frac{n-4}{n-1}\sigma_{uu}) - m_{XY}^2] + \frac{n-4}{n-1}(Y_i - \bar{Y})[m_{XY}\sigma_{uu}]\end{pmatrix}^T$$

For $n \geq 400$ we have $\frac{n-4}{n-1} \geq 0.99 \sim 1$. Thus we can approximate the above expression

** [1], Page 119

as $(\hat{y}_i, \hat{x}_i) = (\bar{Y} - \mu_e, \bar{X} - \mu_u)$

$$+ \frac{1}{m_{YY}m_{XX} - m_{XY}^2} \begin{pmatrix} (Y_i - \bar{Y})[m_{XX}(m_{YY} - \sigma_{ee}) - m_{XY}^2] + (X_i - \bar{X})[m_{XY}\sigma_{ee}] \\ (X_i - \bar{X})[m_{YY}(m_{XX} - \sigma_{uu}) - m_{XY}^2] + (Y_i - \bar{Y})[m_{XY}\sigma_{uu}] \end{pmatrix}^T$$

3.7 Correlated *measurement errors* (but independent of signal) with *equation errors* in both axes

$$\begin{matrix} X_i = x_i + u_i + p_i \\ Y_i = \beta_0 + \beta_1 x_i + e_i + q_i \end{matrix} \quad ; \quad \begin{pmatrix} x_i \\ e_i \\ u_i \\ p_i \\ q_i \end{pmatrix} = \aleph \left[\begin{pmatrix} \mu_x \\ \mu_e \\ \mu_u \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{xx} & 0 & 0 & 0 & 0 \\ 0 & \sigma_{ee} & \sigma_{eu} & 0 & 0 \\ 0 & \sigma_{eu} & \sigma_{uu} & 0 & 0 \\ 0 & 0 & 0 & \sigma_{pp} = \alpha_1 x_i & 0 \\ 0 & 0 & 0 & 0 & \sigma_{qq} = \alpha_2 y_i \end{pmatrix} \right]$$

Therefore we have

$$\mu_{\mathbf{Z}} = \begin{pmatrix} \mu_x + \mu_u \\ \beta_0 + \beta_1 \mu_x + \mu_e \end{pmatrix};$$

$$\mathbf{\Sigma_{ZZ}} = \begin{pmatrix} \sigma_{xx} + \sigma_{uu} + \sigma_{pp} & \beta_1 \sigma_{xx} + \sigma_{eu} \\ & \beta_1^2 \sigma_{xx} + \sigma_{ee} + \sigma_{qq} \end{pmatrix}$$

It can therefore be seen that $\theta = \begin{bmatrix} \mu_x \\ \beta_0 \\ \beta_1 \\ \sigma_{xx} \\ \alpha_1 \\ \alpha_2 \end{bmatrix}$

With similar assumptions as before ($\alpha = \frac{\alpha_1}{\alpha_2}$, known), we have 6 unknowns and 6 equations (since we know σ_{eu} by experiment). Therefore we have

$\alpha(m_{XY} - \sigma_{eu})\beta_1^2 - (\alpha(m_{YY} - \sigma_{ee}) - (m_{XX} - \sigma_{uu}))\beta_1 - (m_{XY} - \sigma_{eu}) = 0$

$\boxed{\Delta^2 = (\alpha(m_{YY} - \sigma_{ee}) - (m_{XX} - \sigma_{uu}))^2 + 4\alpha(m_{XY} - \sigma_{eu})^2}$ where Δ^2 is the discriminant

of the quadratic equation in β_1 above,

$\boxed{\hat{\beta}_1 = \frac{\alpha(m_{YY}-\sigma_{ee})-(m_{XX}-\sigma_{uu})+\Delta}{2\alpha(m_{XY}-\sigma_{eu})}}$ where $\hat{\beta}_1$ is the estimated slope, and

$\boxed{\hat{\sigma_{xx}} = \frac{m_{XY}-\sigma_{eu}}{\hat{\beta}_1}}$ where $\hat{\sigma_{xx}}$ is the estimated variance of the signal in x-channel.

The estimates of the means are : $\boxed{\hat{\mu_x} = \bar{X} - \mu_u \text{ and } \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1(\bar{X} - \mu_u) - \mu_e}$

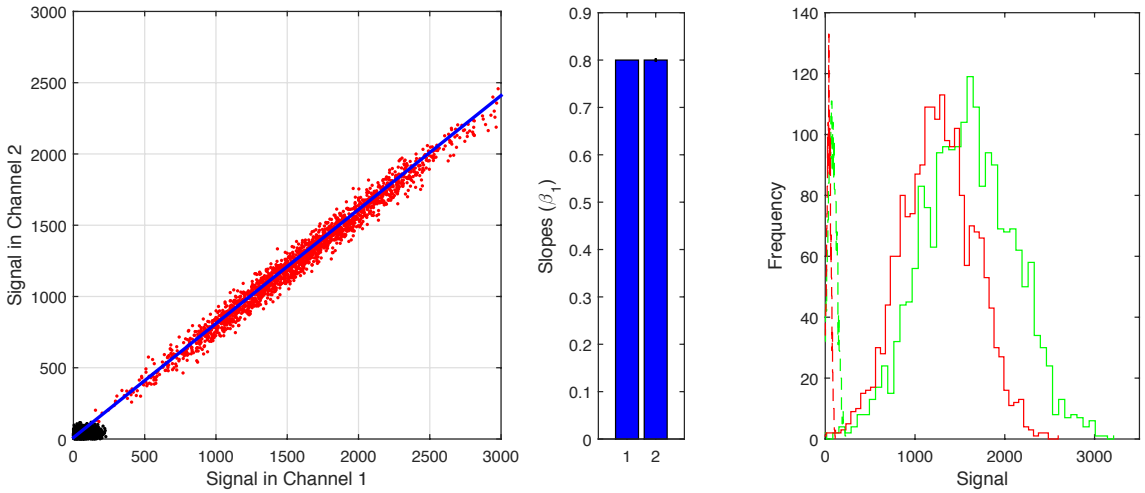


Figure 3.10: **Slope estimates for correlated measurement errors** ($\sigma_{eu} \neq 0$) **and heteroscedastic *equation errors*** (p_i, q_i) **in both axes**
(Left) Plots of simulated voxel intensities in two channels. For each fit, sample size is 2000 voxels. Red dots denote simulated signal and black dots represent the correlated background. The values of statistics assumed were close to those measured in experiments ($\mu_x = 1500, \sigma_{xx} = 500^2, \mu_u = 100, \sigma_{uu} = 50^2, \mu_e = 75, \sigma_{ee} = 25^2, \sigma_{eu} = 20^2$). (Middle) Comparison of the estimated slope (right bar) with its known value (left bar). Error bar on the right bar denotes estimated standard error. (Right) The distributions used for simulation. Solid red and green lines denote the observed signal (X_i, Y_i) in red and green channels respectively. The background distribution (u_i, e_i) in respective channels is denoted by dashed lines.

Again, we can derive the variance in $\hat{\beta}_1$ as follows :

$$\frac{\partial^2 ln(\mathcal{L}(\beta_1))}{\partial \beta_1^2} = \frac{-2[2m_{XY}^2(m_{XX}-\hat{\sigma_{xx}})^2+\hat{\sigma_{xx}}^2(m_{XX}m_{YY}-m_{XY}^2)]}{m_{XX}m_{YY}-m_{XY}^2}$$

Therefore upon simplification we get

$$\frac{\partial^2 ln(\mathcal{L}(\beta_1))}{\partial \beta_1^2}\Big|_{\beta_1=\hat{\beta_1}} = \frac{-2[m_{XY}^2(2m_{XX}^2+\hat{\sigma_{xx}}^2)-2m_{XX}\hat{\sigma x}(4m_{XY}^2+\hat{\sigma_{xx}}m_{YY})]}{m_{XX}m_{YY}-m_{XY}^2}$$
$$= \frac{-2[m_{XY}^2(2m_{XX}^2+(\frac{m_{XY}-\sigma_{eu}}{\hat{\beta}_1})^2)-2m_{XX}(\frac{m_{XY}-\sigma_{eu}}{\hat{\beta}_1})(4m_{XY}^2+(\frac{m_{XY}-\sigma_{eu}}{\hat{\beta}_1})m_{YY})]}{m_{XX}m_{YY}-m_{XY}^2}$$

$$= \frac{-2[m_{XY}^2(2\hat{\beta_1}^2 m_{XX}^2+(m_{XY}-\sigma_{eu})^2)-2m_{XX}(m_{XY}-\sigma_{eu})(4\hat{\beta_1}m_{XY}^2+(m_{XY}-\sigma_{eu})m_{YY})]}{\hat{\beta_1}^2(m_{XX}m_{YY}-m_{XY}^2)}$$

$$var(\hat{\beta_1}) = \frac{\hat{\beta_1}^2(m_{XX}m_{YY}-m_{XY}^2)}{2[m_{XY}^2(2\hat{\beta_1}^2 m_{XX}^2+(m_{XY}-\sigma_{eu})^2)-2m_{XX}(m_{XY}-\sigma_{eu})(4\hat{\beta_1}m_{XY}^2+(m_{XY}-\sigma_{eu})m_{YY})]}$$

Similar to the approach in previous section we estimate the true values of signal $((\hat{y_i}, \hat{x_i}))$ from the observed $Z_i = (Y_i, X_i)$ under the structural model (i.e. random x_i), such that we have $(\hat{y_i}, \hat{x_i}) = (\bar{Y} - \mu_e, \bar{X} - \mu_u)$

$$+ \frac{1}{m_{YY}m_{XX}-m_{XY}^2} \begin{pmatrix} (Y_i-\bar{Y})[m_{XX}(m_{YY}-\frac{n-4}{n-1}\sigma_{ee})-m_{XY}(m_{XY}-\frac{n-4}{n-1}\sigma_{eu})]+ \\ \frac{n-4}{n-1}(X_i-\bar{X})[m_{XY}\sigma_{ee}-m_{YY}\sigma_{eu}] \\ (X_i-\bar{X})[m_{YY}(m_{XX}-\frac{n-4}{n-1}\sigma_{uu})-m_{XY}(m_{XY}-\frac{n-4}{n-1}\sigma_{eu})]+ \\ \frac{n-4}{n-1}(Y_i-\bar{Y})[m_{XY}\sigma_{uu}-m_{XX}\sigma_{eu}] \end{pmatrix}^T$$

Again for $n \geq 400$ we can approximate the above expression as $(\hat{y_i}, \hat{x_i}) = (\bar{Y} - \mu_e, \bar{X} - \mu_u)$

$$+ \frac{1}{m_{YY}m_{XX}-m_{XY}^2} \begin{pmatrix} (Y_i-\bar{Y})[m_{XX}(m_{YY}-\sigma_{ee})-m_{XY}(m_{XY}-\sigma_{eu})]+ \\ (X_i-\bar{X})[m_{XY}\sigma_{ee}-m_{YY}\sigma_{eu}] \\ (X_i-\bar{X})[m_{YY}(m_{XX}-\sigma_{uu})-m_{XY}(m_{XY}-\sigma_{eu})]+ \\ (Y_i-\bar{Y})[m_{XY}\sigma_{uu}-m_{XX}\sigma_{eu}] \end{pmatrix}^T$$

3.8 *Measurement errors* correlated with each other and the signal with *equation errors* in both axes

$$\begin{matrix} X_i = x_i + u_i + p_i \\ Y_i = \beta_0 + \beta_1 x_i + e_i + q_i \end{matrix} ; \begin{pmatrix} x_i \\ e_i \\ u_i \\ p_i \\ q_i \end{pmatrix} = \aleph \left[\begin{pmatrix} \mu_x \\ \mu_e \\ \mu_u \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{xx} & \sigma_{xe} & \sigma_{xu} & 0 & 0 \\ \sigma_{xe} & \sigma_{ee} & \sigma_{eu} & 0 & 0 \\ \sigma_{xu} & \sigma_{eu} & \sigma_{uu} & 0 & 0 \\ 0 & 0 & 0 & \sigma_{pp} = \alpha_1 x_i & 0 \\ 0 & 0 & 0 & 0 & \sigma_{qq} = \alpha_2 y_i \end{pmatrix} \right]$$

Therefore we have

$$\mu_{\mathbf{Z}} = \begin{pmatrix} \mu_x + \mu_u \\ \beta_0 + \beta_1 \mu_x + \mu_e \end{pmatrix};$$

$$\mathbf{\Sigma_{ZZ}} = \begin{pmatrix} \sigma_{xx} + \sigma_{uu} + \sigma_{pp} + 2\sigma_{xu} & \beta_1\sigma_{xx} + \sigma_{eu} + \sigma_{xe} + \beta_1\sigma_{xu} \\ & \beta_1^2\sigma_{xx} + \sigma_{ee} + \sigma_{qq} + 2\beta_1\sigma_{xe} \end{pmatrix}$$

As before, we also have two measurements of m_{Xe} and m_{Yu} available from experiments as follows :

$m_{Xe} = \sigma_{xe} + \sigma_{eu}$

$m_{Yu} = \beta_1\sigma_{xu} + \sigma_{eu}$

It can therefore be seen that $\theta = \begin{bmatrix} \mu_x \\ \beta_0 \\ \beta_1 \\ \sigma_{xx} \\ \sigma_{xe} \\ \sigma_{xu} \\ \alpha_1 \\ \alpha_2 \end{bmatrix}$

With similar assumptions as before ($\alpha = \frac{\alpha_1}{\alpha_2}$, known), we have 8 unknowns and 8 equations (since we know σ_{eu} by experiment). Therefore upon solving the simultaneous equations, we get

$\alpha((m_{XY} - \sigma_{eu}) + (m_{Xe} - \sigma_{Yu}))\beta_1^2 - ((\alpha(m_{YY} - \sigma_{ee}) - (m_{XX} - \sigma_{uu}))\beta_1 - ((m_{XY} - \sigma_{eu}) - (m_{Xe} - \sigma_{Yu})) = 0$

$\boxed{\Delta^2 = ((\alpha(m_{YY} - \sigma_{ee}) - (m_{XX} - \sigma_{uu}))^2 + 4\alpha((m_{XY} - \sigma_{eu})^2 - (m_{Xe} - \sigma_{Yu})^2)}$ where Δ^2 is the discriminant of the quadratic equation in β_1 above,

$\boxed{\hat{\beta_1} = \frac{\alpha(m_{YY} - \sigma_{ee}) - (m_{XX} - \sigma_{uu}) + \Delta}{2\alpha((m_{XY} - \sigma_{eu}) + (m_{Xe} - \sigma_{Yu}))}}$ where $\hat{\beta_1}$ is the estimated slope.

$\boxed{\hat{\sigma_{xx}} = \frac{1}{\hat{\beta_1}}[(m_{XY} + \sigma_{eu}) - (m_{Xe} + m_{Yu})]}$ where $\hat{\sigma_{xx}}$ is the estimated variance of the signal in x-channel.

The estimates of the means are : $\boxed{\hat{\mu_x} = \bar{X} - \mu_u \text{ and } \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1(\bar{X} - \mu_u) - \mu_e}$

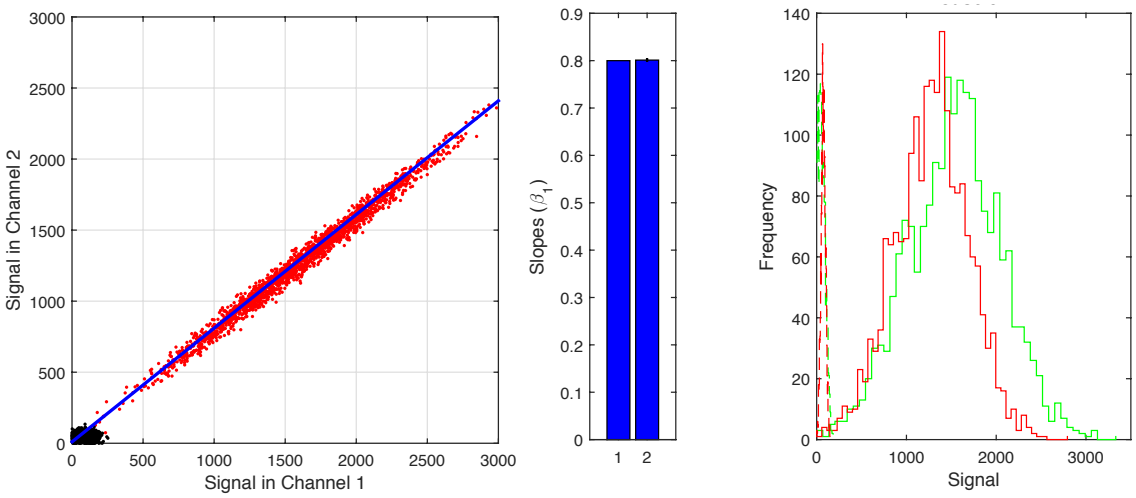


Figure 3.11: **Slope estimates for measurement errors** ($\sigma_{eu} \neq 0$) **correlated with signal** ($\sigma_{Xe}, \sigma_{Yu} \neq 0$) **and heteroscedastic *equation errors*** (p_i, q_i) **in both axes**
(Left) Plots of simulated voxel intensities in two channels. For each fit, sample size is 2000 voxels. Red dots denote simulated signal and black represent the correlated background. The values of statistics assumed were close to those measured in experiments ($\mu_x = 1500, \sigma_{xx} = 500^2, \mu_u = 100, \sigma_{uu} = 50^2, \mu_e = 75, \sigma_{ee} = 25^2, \sigma_{eu} = 20^2, \sigma_{Xe} = 20^2, \sigma_{Yu} = 100^2$). (Middle) Comparison of the estimated slope (right bar) with its known value (left bar). Error bar on the right bar denotes estimated standard error. (Right) The distributions used for simulation. Solid red and green lines denote the observed signal (X_i, Y_i) in red and green channels respectively. The background distribution (u_i, e_i) in respective channels is denoted by dashed lines.

Again, we can derive the variance in $\hat{\beta}_1$ as follows : $\frac{\partial^2 ln(\mathcal{L}(\beta_1))}{\partial \beta_1^2}$

$$= \frac{-[4(m_{XX}(\hat{\beta_1}\sigma_{xx}+\sigma_{xe})-m_{XY}(\sigma_{xx}+\sigma_{xu}))^2+(m_{XX}m_{YY}-m_{XY}^2)\{(\sigma_{xx}(\sigma_{uu}+\sigma_{pp}+\sigma_{xu})-\sigma_{xu}^2)^2-(m_{XX}\sigma_{xx})^2\}]}{m_{XX}m_{YY}-m_{XY}^2}$$

$$= \frac{-[4(m_{XX}(m_{XY}-m_{Yu})-m_{XY}(\frac{m_{XY}-m_{Xe}}{\hat{\beta_1}}))^2+(m_{XX}m_{YY}-m_{XY}^2)\{(\sigma_{xx}(m_{XX}-\sigma_{xu})-\sigma_{xu}^2)^2-(m_{XX}\sigma_{xx})^2\}]}{m_{XX}m_{YY}-m_{XY}^2}$$

$$= -\frac{4(m_{XX}(m_{XY}-m_{Yu})-m_{XY}(\frac{m_{XY}-m_{Xe}}{\hat{\beta_1}}))^2}{m_{XX}m_{YY}-m_{XY}^2} - \frac{(m_{XX}m_{YY}-m_{XY}^2)\{(\sigma_{xx}(m_{XX}-\frac{m_{Yu}-\sigma_{eu}}{\hat{\beta_1}})-(\frac{m_{Yu}-\sigma_{eu}}{\hat{\beta_1}})^2)^2-(m_{XX}\sigma_{xx})^2\}}{m_{XX}m_{YY}-m_{XY}^2}$$

$$= -\frac{4(m_{XX}\hat{\beta_1}(m_{XY}-m_{Yu})-m_{XY}(m_{XY}-m_{Xe}))^2}{\hat{\beta_1}^2(m_{XX}m_{YY}-m_{XY}^2)} - \frac{(m_{XX}m_{YY}-m_{XY}^2)\{(\hat{\beta_1}\sigma_{xx}(\hat{\beta_1}m_{XX}-(m_{Yu}-\sigma_{eu}))-(m_{Yu}-\sigma_{eu})^2)^2-(\hat{\beta_1}m_{XX}\sigma_{xx})^2\}}{\hat{\beta_1}^2(m_{XX}m_{YY}-m_{XY}^2)}$$

Therefore we get $var(\hat{\beta_1})$

$$= \frac{\hat{\beta_1}^2(m_{XX}m_{YY}-m_{XY}^2)}{[4(m_{XX}\hat{\beta_1}(m_{XY}-m_{Yu})-m_{XY}(m_{XY}-m_{Xe}))^2+(m_{XX}m_{YY}-m_{XY}^2)\{(\hat{\beta_1}\sigma_{xx}(\hat{\beta_1}m_{XX}-(m_{Yu}-\sigma_{eu}))-(m_{Yu}-\sigma_{eu})^2)^2-(\hat{\beta_1}m_{XX}\sigma_{xx})^2\}]}$$

As mentioned before, using simulated data we demonstrate that the estimate of slopes is insensitive to α (Figure 3.12). For estimating true signal, we know that as the *measurement errors* and *equation errors* are independent . Therefore

$$\mathbf{\Sigma_{Z\epsilon}} = \begin{pmatrix} \beta_1\sigma_{xe}+\sigma_{ee} & \beta_1\sigma_{xu}+\sigma_{ue} \\ \sigma_{xe}+\sigma_{ue} & \sigma_{xu}+\sigma_{uu} \end{pmatrix} = \begin{pmatrix} \beta_1\sigma_{xe} & \beta_1\sigma_{xu} \\ \sigma_{xu} & \sigma_{xe} \end{pmatrix} + \mathbf{\Sigma_{\epsilon\epsilon}}$$

$$\Rightarrow \mathbf{\Sigma_{Z\epsilon}} = \sigma_{xe}\begin{pmatrix} \beta_1 & 0 \\ 0 & 1 \end{pmatrix} + \sigma_{xu}\begin{pmatrix} 0 & \beta_1 \\ 1 & 0 \end{pmatrix} + \mathbf{\Sigma_{\epsilon\epsilon}}$$

For all previous cases , the first two terms are zero as signal is uncorrelated to the background ($\sigma_{xe} = \sigma_{xu} = 0$). Therefore for all previous cases we have $\mathbf{\Sigma_{Z\epsilon}} = \mathbf{\Sigma_{\epsilon\epsilon}}$ and hence we use the background statistics ($\mathbf{\Sigma_{\epsilon\epsilon}}$) to estimate true signal $(\hat{y_i}, \hat{x_i})$. In the current case we have $(\hat{y_i}, \hat{x_i}) = (Y_i - \mu_e, X_i - \mu_y) + \frac{n-4}{n-1}(Y_i - \bar{Y}, X_i - \bar{X})\mathbf{m_{ZZ}}^{-1}\mathbf{\Sigma_{Z\epsilon}}$.

For the present case this yields:

$$(\hat{y_i}, \hat{x_i}) = (Y_i - \mu_e, X_i - \mu_y) +$$
$$\frac{1}{m_{YY}m_{XX}-m_{XY}^2}\frac{n-4}{n-1}(Y_i-\bar{Y}, X_i-\bar{X})\begin{pmatrix} m_{XX}m_{Ye}-m_{XY}m_{Xe} & m_{XX}m_{Yu}-m_{XY}m_{Xu} \\ -m_{XY}m_{Ye}+m_{YY}m_{Xe} & -m_{XY}m_{Yu}-m_{YY}m_{Xu} \end{pmatrix}$$

$$= (Y_i - \mu_e, X_i - \mu_y) + \frac{1}{m_{YY}m_{XX}-m_{XY}^2}(\frac{n-4}{n-1})\begin{pmatrix} (Y_i-\bar{Y})(m_{XX}m_{Ye}-m_{XY}m_{Xe}) + \\ (X_i-\bar{X})(-m_{XY}m_{Ye}+m_{YY}m_{Xe}) \\ (Y_i-\bar{Y})(m_{XX}m_{Yu}-m_{XY}m_{Xu}) + \\ (X_i-\bar{X})(-m_{XY}m_{Yu}-m_{YY}m_{Xu}) \end{pmatrix}^T$$

Finally: $(\hat{y_i}, \hat{x_i}) = (Y_i - \mu_e, X_i - \mu_y) +$

$$\frac{1}{m_{YY}m_{XX}-m_{XY}^2}(\frac{n-4}{n-1})\begin{pmatrix} (Y_i-\bar{Y})(m_{XX}(\hat{\beta_1}(m_{Xe}-\sigma_{eu})+\sigma_{ee})-m_{XY}m_{Xe}) + \\ (X_i-\bar{X})(-m_{XY}(\hat{\beta_1}(m_{Xe}-\sigma_{eu})+\sigma_{ee})+m_{YY}m_{Xe}) \\ (Y_i-\bar{Y})(m_{XX}m_{Yu}-m_{XY}(\frac{m_{Yu}-\sigma_{eu}}{\hat{\beta_1}}+\sigma_{uu})) + \\ (X_i-\bar{X})(-m_{XY}m_{Yu}-m_{YY}(\frac{m_{Yu}-\sigma_{eu}}{\hat{\beta_1}}+\sigma_{uu})) \end{pmatrix}^T$$

It can be seen that all the terms on the right hand side are known through control experiments. Therefore the true signal can be estimated. However in real experiments we do not

need to use this expression as the correlation of the background to signal is negligible as compared to the correlation between signals (Figure 3.6). Therefore for the real experimental data we ignore these correlations. It can be seen that in such a scenario the estimate of β_1 derived here, is identical to the one in section 3.7 when $m_{Xe} = m_{Yu} = 0$. Therefore for simplicity of actual experiment, we use the model in section 3.7 to estimate the slope of the line.

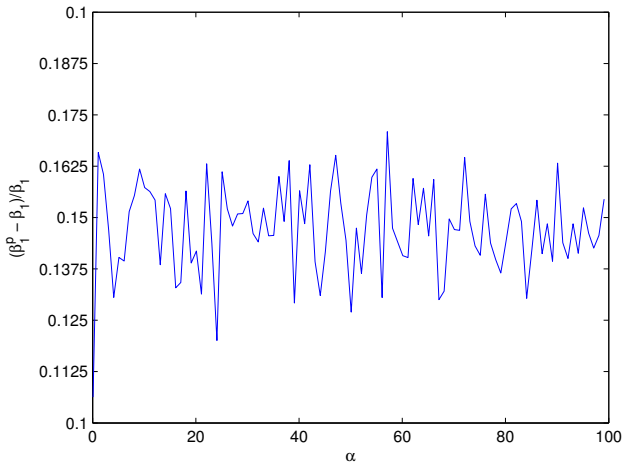


Figure 3.12: **Fractional error in the estimated value of slope for measurement errors** ($\sigma_{eu} \neq 0$) **correlated with signal** ($\sigma_{Xe}, \sigma_{Yu} \neq 0$) **and heteroscedastic *equation errors* (p_i, q_i) in both axes**
Simulations were performed for known values of slopes (β_1; x-axis of the plot) The y-axis denotes the percentage error in the slope estimates using maximum likelihood method for different values of α

3.9 Slope estimation in cases where data deviates from normal distributions

We have shown that the observed data (signal and background) in our experiments follows normal distribution. However to understand cases where the distribution might be non-normal (such as log normal), we evaluate the performance of the MLE analysis via simulations of the statistics for a few cases of log-normal distribution of signal. This shows that the method performs reasonably well for log-normal distributions with a range of scale factors (Figure 3.13). The absolute value of percentage error in the expected value of the

estimated slope is very low for the range tested here, though there is a a reduction in error with increase in scale factor. However the standard error in the estimates is high at lower scale factors corresponding to very narrow signal distributions (Figure 3.13 A and 3.14C). This is primarily because of the very small value of variances for signal (m_{XX} and m_{YY}) in such cases, despite the strong correlation between channels ($m_{XY} \gg m_{XX}, m_{YY}$), that causes large fluctuations owing to divisions of small numbers. Overall the MLE framework performs well for broad distributions that can be reasonably approximated with Gaussian distributions, as indicated through the limited number of cases presented here. Further investigations into dependence of estimates on relative values of variances of and between signals is needed to provide a general framework for understanding the failure modes of current approach better.

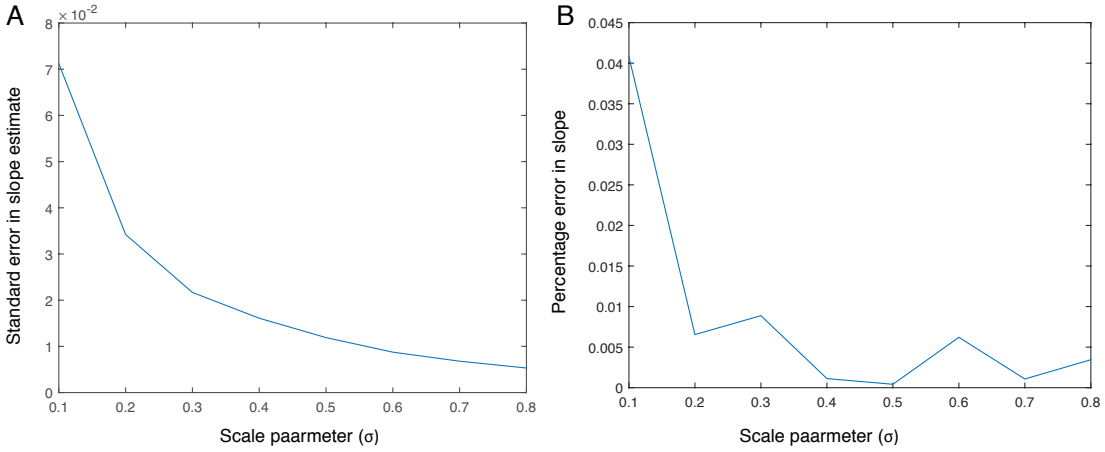


Figure 3.13: **Standard error in slope estimates for log-normal signal distribution** (A) Standard error in slope estimates decrease with increasing value of scale factor for log normal signal distribution. (B) Percentage error in predicated value of slopes for different values of scale factor

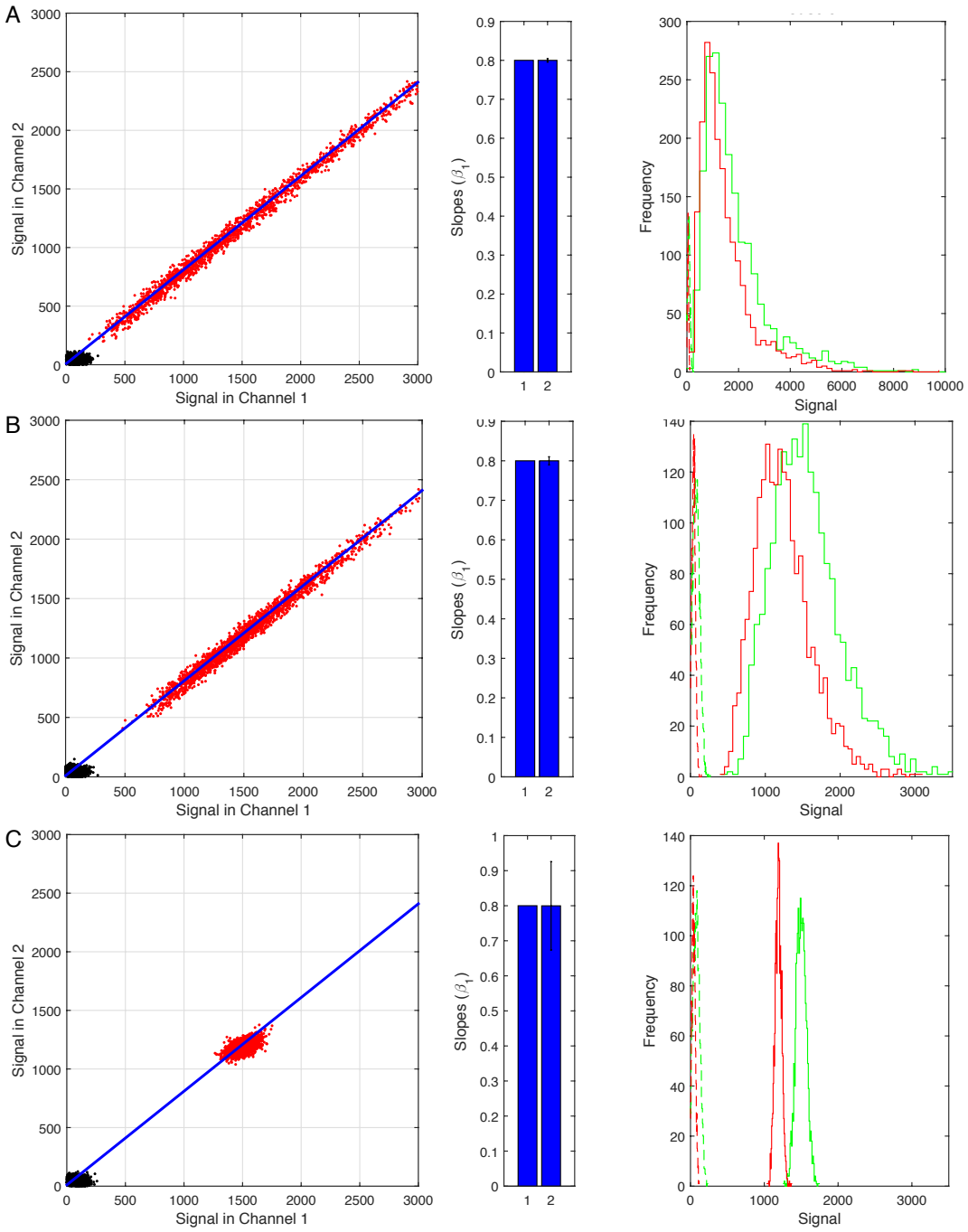


Figure 3.14: **Slope estimates for log-normal signal distribution**
Three representative examples of slope estimates for log-normal signal distribution. Red dots denote simulated signal using location parameter, $\mu = 7.2605$, and scale parameter, $\sigma = 0.6492$ (A), 0.325 (B), 0.0325 (C). For each fit, sample size is 2000 voxels. (Left) Plots of simulated voxel intensities in two channels. (Middle) Comparison of the estimated slope (right bar) with its known value (left bar). Error bar on the right bar denotes estimated standard error. (Right) The distributions used for simulation. Solid red and green lines denote the observed signal (X_i, Y_i) in red and green channels respectively. The background distribution (u_i, e_i) in respective channels is denoted by dashed lines.

3.10 References

1. Fuller, W. A. (1987). Measurement Error Models. Wiley, New York.

2. Cheng, C. and Van Ness, J. W. (1999). Statistical Regression with Measurement Error. Arnold, London.

3. Kendall, M.G. and Stuart, A. (1979). The Advanced Theory of Statistics, Vol 2. Griffin, London.

4. Birch, M.W. (1964).A note on maximum likelihood estimation of a linear structural relationship. J.Amer. Statist. Assoc., 59, 1175-1178

5. Madansky, A. (1959) The Fitting of Straight Lines When both Variables are Subject to Error. J.Amer. Statist. Assoc., 54 (285), 173-205

Chapter 4

Dynamic imaging of the growth plate cartilage reveals multiple contributors to skeletal morphogenesis

4.1 Motivation

Among the diverse skeletal elements, the growth plate cartilage of long bones is ideal for 4D (xyz and t) analyses of cell behaviors due to its largely unidirectional growth and cylindrical shape. Growth is accompanied by the progressive differentiation of chondrocytes along the proximo-distal axis (PDA; Fig 4.1A) [1-3]. Cells residing close to articular cartilage in the resting zone (RZ), deposit collagen II while differentiating into proliferative chondrocytes (PZ), which further differentiate sequentially into the enlarged chondrocytes in the prehypertrophic (PHZ) and hypertrophic zone (HZ). Processes such as the deposition of extracellular matrix (ECM), provide the necessary the scaffold for the invasion of osteoblast cells, which in turn provide a calcified bony matrix for long bone maturation [1-3]. As these are slow but continuous process they require long-term imaging with detailed monitoring. Multiple signaling pathways, including Indian hedgehog (Ihh), Parathyroid hormone related peptide (Pthrp), bone morphogenetic proteins (Bmp), fibroblast growth factors (FGFs), and wingless/int-1 molecules (Wnt) [4-8], form a molecular regulatory circuit controlling cartilage elongation. Significant work in the past, based on static imaging, has suggested the importance of convergent-extension (CE) like behaviors of the PZ cells, volume increase

of the HZ cells and deposition of ECM, to adult cartilage elongation [9-13]. However, the specific contributions and relationships of the various cellular processes to embryonic tissue elongation have yet to be fully defined. Static imaging studies face two general challenges for dissecting tissue morphogenesis: first, fixing and processing of tissues can alter cell sizes and the spaces between them in a non-uniform fashion; second, dynamic cellular processes can only be inferred from static studies. In the absence of tools for visualizing the various cellular processes and systematically dissecting their roles, it is difficult to understand how individual cell behaviors are translated into the collective cell behavior that underlies tissue morphogenesis.

To meet these challenges, we performed quantitative imaging on live tissue and generate predictive models to extract information on the underlying cellular dynamics. In this chapter I describe the work on quantitative live optical imaging for visualization of individual cells and the quantitative description of their behaviors in 4D. In silico modeling based on measures of PZ cell movement inferred from images created testable predictions for future studies. Our closed loop analysis revealed that embryonic cartilage elongation is highly coordinated, with critical contributions from two types of cell morphogenesis in the PZ: ECM deposition and cell volume enlargement.

4.2 4D imaging of cartilage growth using avian metacarpal culture

To permit our quantitative imaging analyses of skeleton shaping and the underlying cellular processes in the PZ, we established an organ culture system that supported normal growth and permitted long term imaging of the live specimen over 56 hours [14]. We started with chick metacarpals of the forelimb at the embryonic day 8 (E8) as it is largely PZ and is sufficiently thin that nutrients can penetrate to the chondrocytes resulting in normal growth when isolated in culture. The cells were labeled globally with cytoplasmic green fluorescent protein (cytoplasmic-GFP) to facilitate visualization in vivo.

To non-invasively visualize cells in multiple layers of the explanted bone, we used two-photon laser scanning microscopy (2p-LSM, [15,16]), which can image cells deeper in the

tissue than conventional confocal laser scanning microscopy. Image stacks were collected of one end of the cultured metatarsal hourly for 55 hours, to a depth that reached half the thickness (Fig 4.1 B,C). The illuminated and unexposed half of the metacarpal exhibited similar length extension, suggesting there is little if any photo-toxicity or detrimental effects of being cultured on the microscope stage [14]. We identified the PZ cells based on their positions in the tissue.

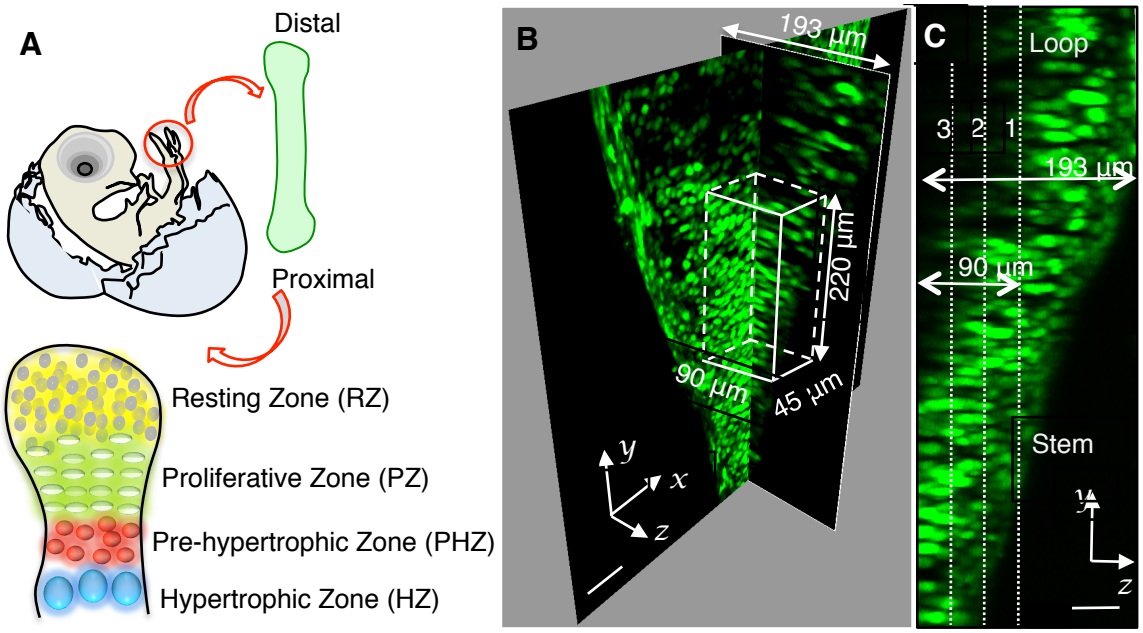


Figure 4.1: **Live imaging of cartilage growth**
(A)The growth plate cartilage displays a stem-loop shape and contains four zones along the PDA: resting (RZ), proliferative (PZ), prehypertrophic (PHZ) and hypertrophic (HZ) zones. (B) Two orthogonal sections shown in the 3D rendered image of the GFP (green) expressing metacarpal at t=29 hours (chosen for the best depiction) of live imaging. The white cuboid drawn shows the approximate region selected for PZ cell tracking. (H) Side view of the same sample. The middle 90 μm region spanning both stem and loop parts was digitally sliced into 3 sections along the imaging axis (Methods), to facilitate cell segmentation. Scale bars: (B) 50 μm, (C) 20 μm

To quantitatively define these cell behaviors, we performed 3D image segmentation, so that individual PZ cells could be tracked over time, and mapped the 4D trajectories of 472 PZ cells over a 22090x90 μm^3 region (Fig 4.1 B, C; Fig 4.2 A,B). We estimated that our segmentation and tracking is able to accurately track 95% of the cells in the volume (this was validated by labeling the post-imaging sample with phalloidin, and finding that

97% of the cells identified in this way were GFP labeled [14]). In addition, all computed trajectories were manually checked and corrected for possible errors. Our processing was able to identify 481 cells; out of which we eventually analyzed 472 cells (98%). Thus, we conclude that our segmentation and tracking is able to accurately cover 95% (98% of 97% labeled cells) of all the cells.

4.3 Anisotropic spreading of the PZ cells

For analyzing and displaying the 4D trajectories, we defined a Cartesian coordinate system aligned with the shape of the metacarpal (Fig 4.1B; 4.2A): The PDA was defined as the y axis, with the middle of the tissue taken as the origin; the axis perpendicular to the long axis in the imaging plane was defined as the x axis; the axis coming out (along the light path) and orthogonal to the image plane was defined as the z axis. Decomposition of cellular trajectories using this Cartesian coordinate system revealed a clear bias to the spreading behavior, with the most significant movement along the y-axis (Fig 4.2C). This y-axis motion occurred in two phases, with many cellular trajectories moving proximally in the first 10 hours of imaging; thereafter, all trajectories coursed distally. Color-coding the trajectories according to the initial positions of the cells revealed an inherent order, with the displacements increasing progressively from the proximal to the distal end in the second phase (Fig 4.2C). Such a coordinated pattern was not observed in the x and z directions, suggesting that the cell displacements are mainly determined by their initial positions along the PDA.

The orderly spreading of cells suggests some cooperation between cells, the nature of which needed further investigation. By analyzing the total y displacement of each cell with respect to its initial y position, we found a linear relationship (R = 0.986) (Fig 4.2D), implying that the displacement of a cell is contributed additively by all the cells proximal to it. Such an linear displacement with single slope implies that all the cells contribute equally to the overall displacement. Importantly, the equal contribution from each cell to the overall displacement means that the relative speed between nearest-neighbor cells is the same, independent of their positions along the PDA.To explore the nature of the

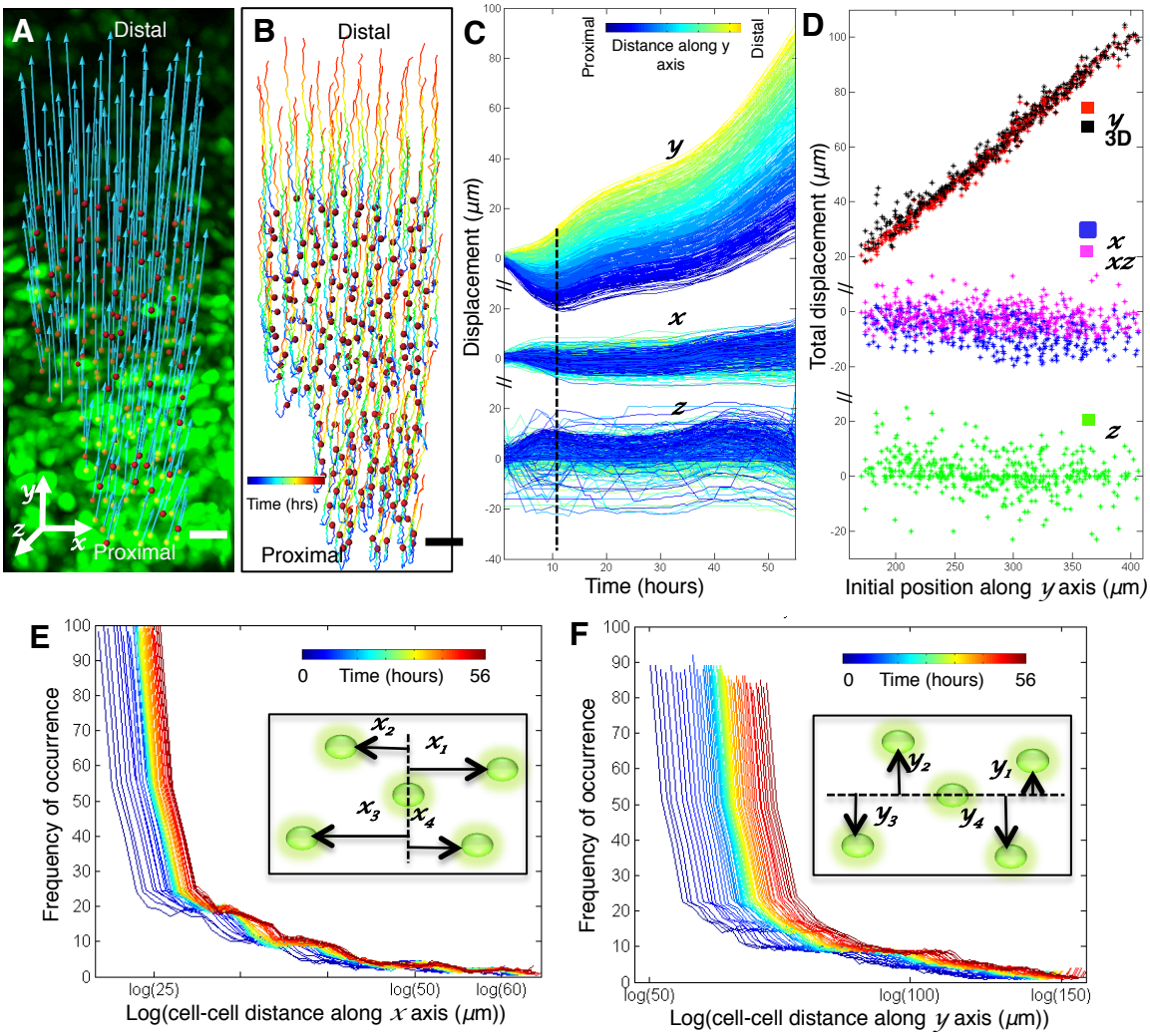


Figure 4.2: **Cells undergo collective spreading displacement during cartilage growth**

(A, B) PZ cells were segmented (red dots), and their net displacement vectors (blue lines, A) and trajectories (colored lines, B) were mapped, showing strong orientation of the cell displacement towards the distal end of the tissue. (C) Cell displacement over time relative to its initial y position (color-coded as shown in the top inset), showing the largest displacement along the y axis and smaller along the x and z axes. (D) Total cell displacement (t=55 hours) along different axes and planes relative to their initial y positions. Cell displacements along the y axis account for most of the displacements in 3D with linear increase in magnitude according to their initial y positions ($R^2 = 0.986$) (E, F) Analysis of cell-cell distance change. The mean of center-to-center distance between all possible pairs of cells at any given time was measured (as indicated in the inserts), and the distributions of those means for all cells over time (color-coded) were plotted along the x (E) and the y axes (F) on a semi-log scale (to amplify the increase in mean over time graphically). More increase in cell-cell distance along the y-axis as compared to the x-axis implies an anisotropic spreading behavior of PZ cells. (n=472 cells in c-f). Scale bar: $50\mu m$.

cooperative behavior suggested by the trajectory analysis at the individual cell level, we employed a cell-based reference system to analyze the distance change between neighboring cells. The mean of the distance between a cell and every other cell was calculated at each time point. This yielded unimodal distributions of the mean distances, with similar shapes at all time points, strongly suggesting that cells were displaced away from each other at a global level (Fig 4.2E, F). Similar to cell displacement behavior, cell-cell distance change showed a larger spread along the y-axis, direction of tissue elongation.

4.4 Convergent-extension and cell division cannot account for cell spreading

What, then, are the cellular processes accounting for the observed anisotropic cell spreading and metacarpal elongation? Previous qualitative studies have proposed an important role of convergent extension (CE) in controlling cartilage morphology [10]. CE achieves elongation by cell-cell intercalations orthogonal to the growth axis, resulting in the narrowing and lengthening of the embryonic axis of the frog embryo, for example [17]. Based on the results in previous section (Fig 4.2E, F), it is clear that there is extension in all directions. CE could only be a significant contributor during metacarpal elongation if changes in other cellular morphogenesis compensates for the convergence. To definitively test the contribution of CE in metacarpal elongation, we performed a polygon analysis, in which the relative motions of all cells with respect to their immediate neighbors are tracked over time (Fig 4.3A) [18]. If CE takes place, the increase in height (along the PDA) of the polygon should accompany the decrease in its width; therefore we counted the number of cells displaying the diverse set of neighbor behaviors (Fig. 4.3B). Majority of cells underwent extension in both x and y directions (green), while only approximately 10% of cells underwent CE (red), according to polygon analysis.

We next tested the ability of mitotic division to explain the growth observed by our quantitative imaging. We identified the cells that underwent mitosis in chick with the cytoplasmic-GFP (Fig 4.3 E, F) and in quail with the nuclear H2B-mCherry (Fig 4.4A, B) [19]. In both systems, the mitotic rate was below 7%, and oriented orthogonal to

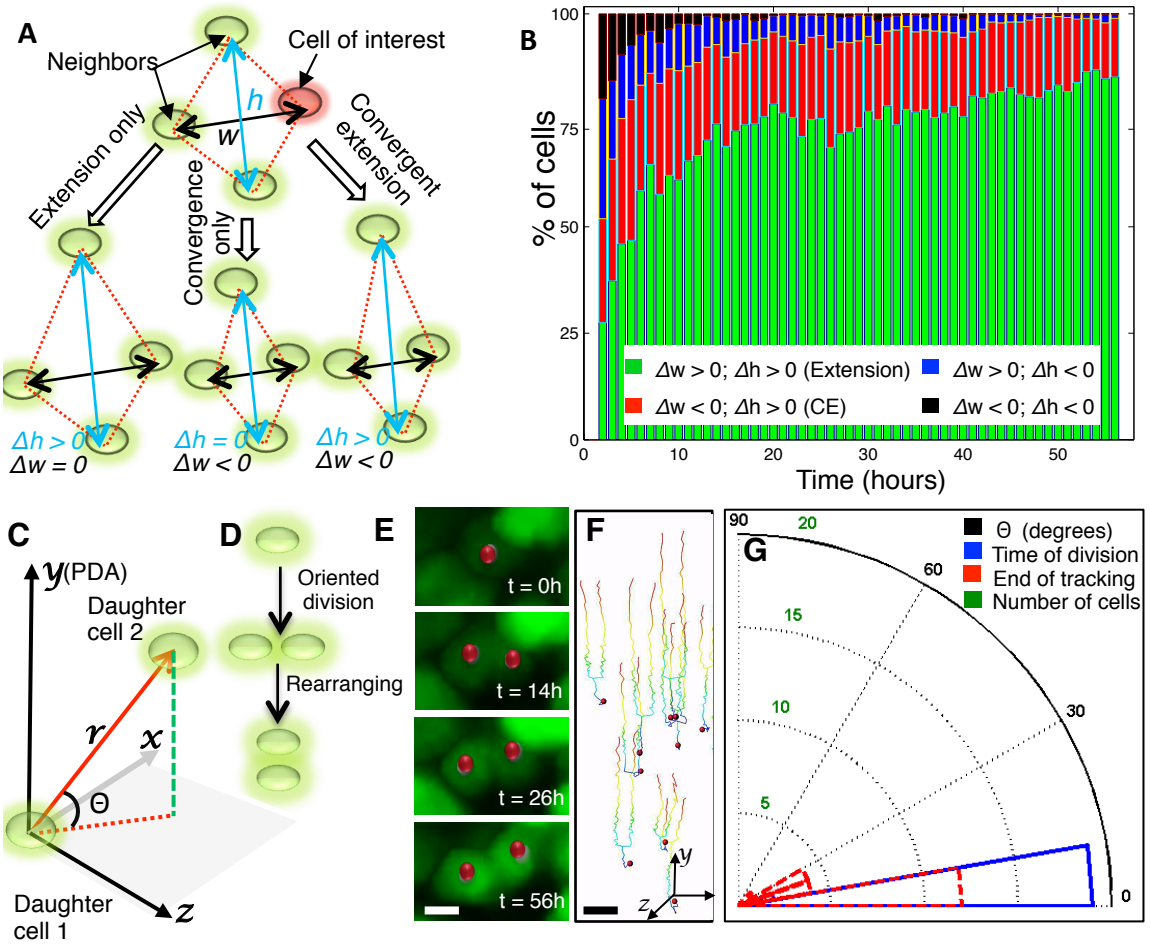


Figure 4.3: **Convergent-extension and mitotic division cannot account for cell spreading**

(A) Schematic diagrams of polygon analysis. A polygon was drawn for every cell (red) by choosing three nearest neighbors (green). Change in the height (δh) and the width (δw) of the polygon depends on the nature of cell motion with respect to its neighbors. (B) Percentage count of cells undergoing different types of cell motion. Majority of cells underwent extension in both x and y directions (green), while only approximately 10% of cells underwent convergent-extension (CE) (red), according to polygon analysis (C) Polar coordinate system defined for analyzing the distance (r) and angle(Θ) with respect to x-z plane between two daughter cells. If a cell divides along or orthogonal to the PDA, Θ at the time of division is 90 degrees or 0, respectively. (D) Schematic of oriented division and daughter cell rearrangement (E-F) Several time frames of one representative dividing cell expressing GFP (red dots) are presented (E), and the trajectories of all dividing cells were mapped (F). (G) Polar histogram showing that Θ measured for all dividing cells were below 15 degrees at the time of division (blue), ruling out the possibility of oriented cell division along the PDA. The fact that Θ for all dividing cells were below 30 degrees by the end of tracking (red) further excluded the possibility of daughter cell rearranging afterwards (n=17 cells). Scale bars: (E,F) 10μm

the PDA (Fig 4.3 C, E-G), which is insufficient to explain the large and homogenous cell displacements. These findings make it unlikely that mitotic division is a major contributor to the observed cell displacements during the time window for growth in this study. It must be mentioned that since ultimately the number of cells in a tissue is expected to contribute to its size, further studies are needed to fully elucidate the precise spatio-temporal relationship between cell division and cartilage growth, through extending the analysis over longer time windows. Owing to the small number of mitotic events, daughter cell rearrangement (Fig. 4.3 D) also does not account for the growth. This is further validated by our analysis where we measured the angles and the distances between daughter cells over time (Fig 4.3 C, G; 4.4 E) and showed that only 10% of the dividing cells underwent significant rearrangement.

4.5 Conservation of cell spreading behavior across avian embryos

To determine whether the PZ cell displacement pattern is conserved across avian embryos, the same methodology was applied to analyze metacarpal elongation in transgenic quail embryos constitutively expressing H2B-mCherry in their cell nuclei (Fig. 4.4). We observed similar cellular trajectories (Fig. 4.4 A-D) and cell-cell distance change (Fig. 4.4 F, G), with displacement lengths (Fig. 4.4 D) increasing in an additive and homogeneous fashion in the y direction. In quail, the two distinct phases of cell displacement occurred with slightly different timing, with cell trajectories that were first displaced distally, then bi-directionally, followed by another phase of distal displacement (Fig. 4.4 C). These subtle differences in timing between quail and chick may be the product of species-dependent differences in the timing or patterning of long bone growth, in a largely conserved elongation mechanism.

4.6 ECM expansion and cell volume enlargement cause cell spreading

Having ruled out major roles for CE, mitotic cell division and daughter cell rearrangement, we addressed the possible roles of cell volume enlargement and intercellular space expansion

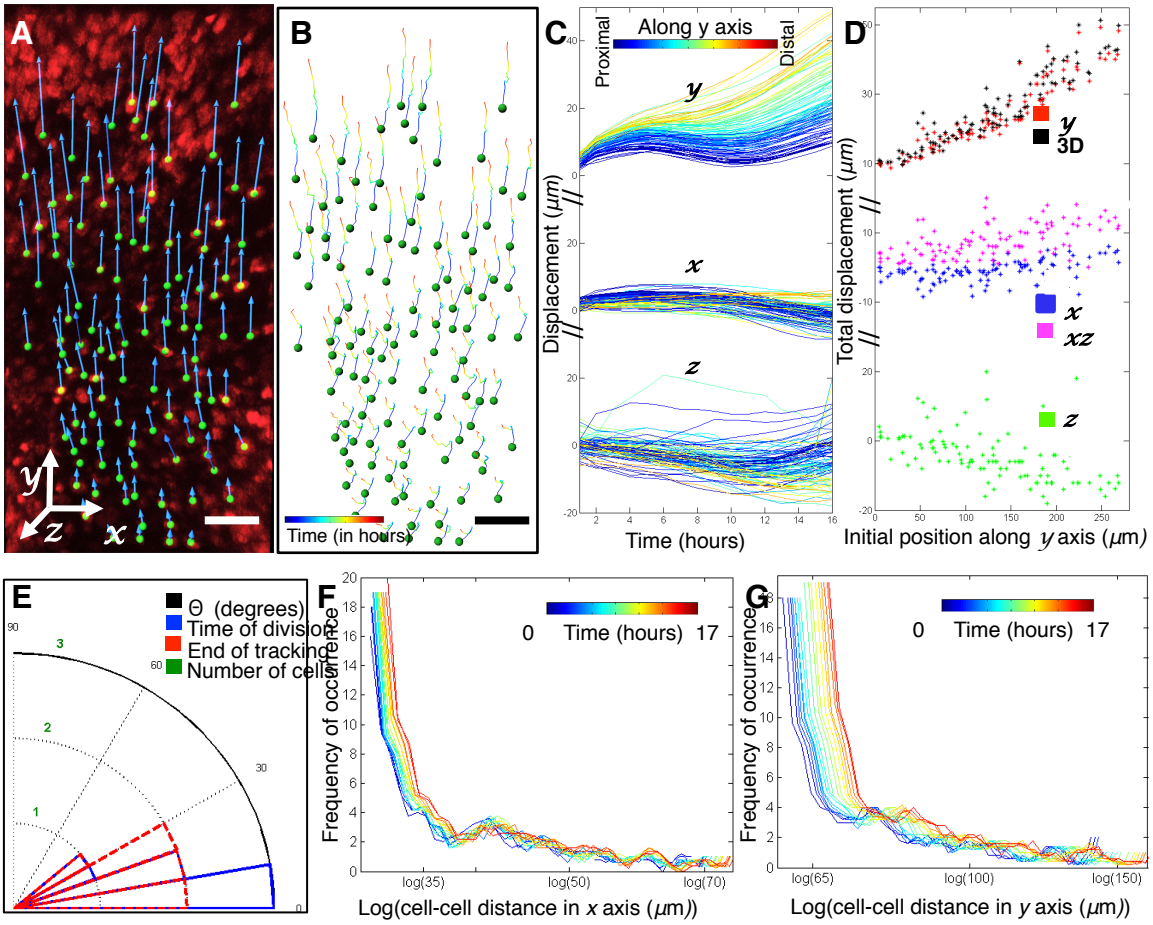


Figure 4.4: **Quail PZ cells exhibit similar spreading displacement and orthogonal division**

(A,B) Nuclei in the quail metacarpal expressing H2B-mCherry (red) were segmented (green dots), and their net displacement vectors (blue lines, A) and trajectories (colored lines, B) were mapped, showing strong orientation of the cell displacement towards the distal end of the tissue similar to observations in chick. (C) Cell displacement over time relative to its initial y position (color- coded as shown in the top inset), showing the largest displacement along the y axis and smaller along the x and z axes (n = 74 nuclei). (D) Total cell displacement (t = 16 hours) along different axes and planes relative to their initial y positions. Cell displacements along the y-axis account for most of the displacements in 3D with similar trend according to their initial y positions as in chick (n = 74 nuclei). (E) Angles measured for all dividing nuclei (based on the coordinate system introduced in Fig 4.3 C). As before Θ measured for all dividing cells were below 15 degrees at the time of division (blue), ruling out the possibility of oriented cell division along the PDA with no daughter cell rearrangement (n = 4 nuclei).(F,G) Analysis of cell-cell distance change. The mean of center-to-center distance between all possible pairs of nuclei at any given time was measured and the distributions of those means for all cells over time (color-coded) were plotted along the x (F) and y axes (G) on a semi-log scale (to amplify the increase in mean over time graphically). More increase in cell-cell distance along the y-axis as compared to the x-axis implies an anisotropic spreading behavior of cells (n = 74 nuclei). Scale bar (a): $50\mu m$

[11-13]. One of the major components of the ECM filling the intercellular space in the PZ is col2a (Fig 4.5 A, B) [6]. Since 97% of the PZ cells in our imaged chick metacarpals were labeled with cytoplasmic-GFP, the dark area between green cells offers a clean means to recognize the space occupied by ECM, thus allowing us to simultaneously determine the contributions of both cell and ECM volumes. We selected a region in the distal PZ (white box in Fig 4.5C) and segmented the image into cell (GFP bright) and ECM (GFP dark) (Fig 4.5D) based on a threshold equal to the average image intensity; the results showed a 62% increase in ECM volume, and a 27% increase in cell volume (Fig 4.5D). Thus, the increase in the total volume of the selected region results from increases in both cell and ECM volumes, with ECM expansion playing the more significant role.

To test the sufficiency of the observed increases in cellular and ECM volumes to account for the experimental observed cell trajectories, we performed computer simulations based on a model where the cartilage tissue growth and the accompanying cell displacement depends solely on ECM and cellular volume increases (Fig. 4.6). Using the measured rates of cell volume and ECM volume increases (the slopes of the graphs in Fig 4.5 E) as inputs into the model along with the initial positions of all the cells determined by our cell tracking, we were able to accurately predict the positions of the cells over time (Fig. 4.5 F, G). The resultant simulated cell trajectories overlapped well with the experimentally-observed trajectories within less than 3% absolute error, for all tracked cells over the experimental time window (inset of Fig. 4.5 F), suggesting that the measured cell spreading along the PDA can be sufficiently explained by the expansion of ECM and cell volumes.

4.7 Discussion

Our dynamic analyses offer several novel insights into PZ cells behavior during embryonic cartilage elongation: first, the PZ cells show progressive anisotropic spreading mainly along the PDA; second, the relative speed between individual cells are constant and this homogenous behavior additively contribute to the orderly cell motions observed at a global level; third, mitotic cell divisions, daughter cell rearrangements and motions of CE all take place, but are minor contributors to cellular trajectories ; fourth, cell spreading is largely due to

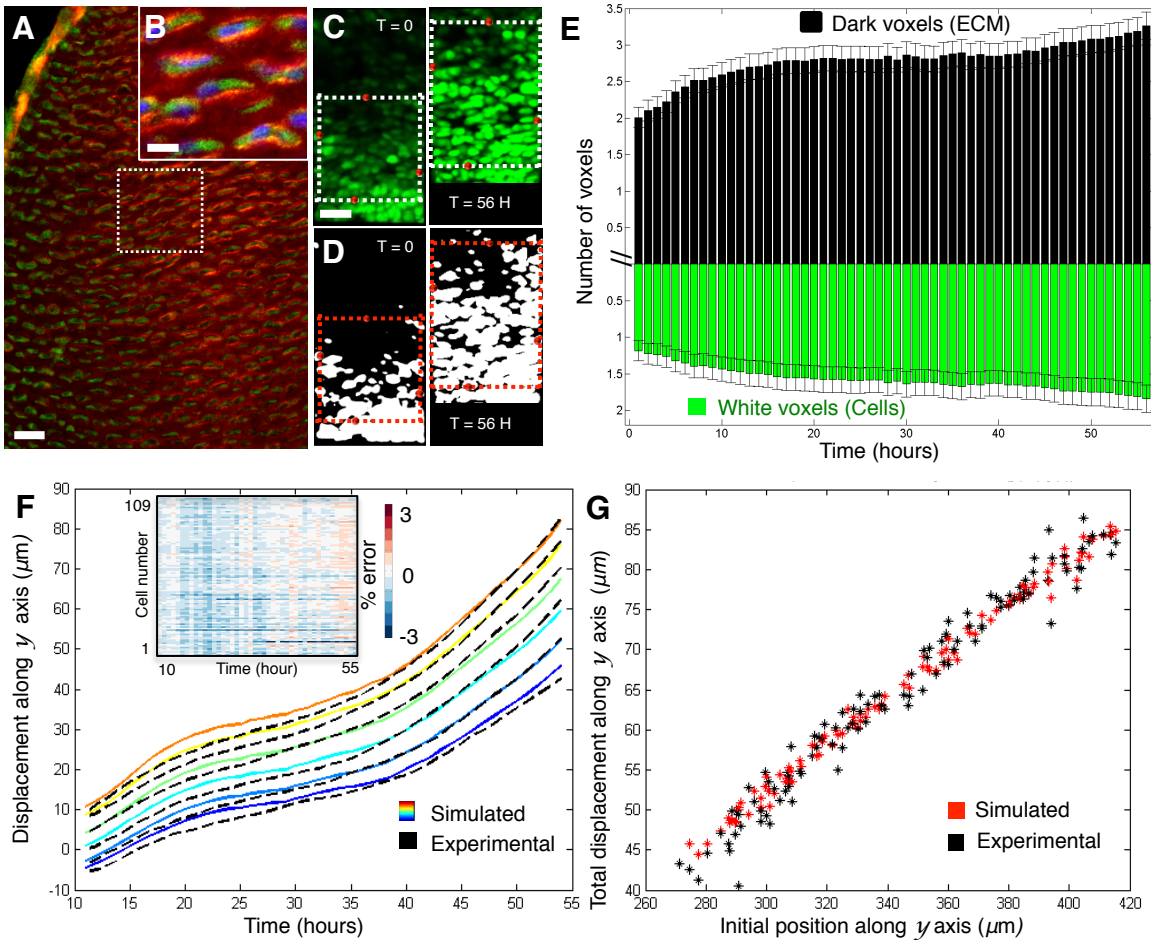


Figure 4.5: **ECM expansion and cell volume enlargement account for cell trajectories**

(A) Representative confocal images of Col2a antibody staining (red) of the frozen sections from the chick metacarpal expressing cytoplasmic-GFP (green) at low magnification (B) Enlarged image of the region in the white box in (A) with nuclei counterstained with DAPI (blue) (C) Maximum intensity projections of the first and last time points in the 4D live imaging of a chick metacarpal expressing GFP; the region enclosed within the expanding white box selected for voxel analysis, based on the same 4 cells on the boundaries (red dots). (D) Corresponding binary images to classify voxels as either belonging to ECM (black) or cell (white). (E) Total count of the number of dark and white voxels, denoting the volume occupied by ECM and cells respectively, shows the expansion of both ECM and cell volume. (F) Overlapping simulated and experimental (tracked) cell trajectories along the y axis depicted for 6 randomly chosen cells, suggests that the change in cell volume and ECM are sufficient to account for cell displacement. Heat map in the insert depicts the errors of all simulated trajectories as a percentage of the experimental values showing that the absolute errors are always below 3%. (G) Total cell displacement length (t=55 hours) along the y axis of all simulated cells are plotted against their initial y positions, displaying similar distribution pattern to the experimental ones. Scale bar: (A,C) 15μm, (B) 50μm

the expansion of the ECM, with a secondary contribution from increase in cell volume. A central task of modern biology is to understand how cells build a tissue in a 4D context [15]. Live imaging provides a unique approach to meet this challenge due to its noninvasive ability to collect longitudinal data from cells embodied in their normal context of the live tissues [15,18]. Live imaging generates large datasets, which can be unwieldy to analyze, and biological interpretation of such complex numerical datasets requires quantitative and statistically robust toolsets. We developed a suit of custom software tools (based on MATLAB) to meet this challenge, providing a means for analyzing relative distance change between individual cells (Fig 4.2), CE, mitotic division orientation and daughter cell rearranging (Fig 4.3), the rate of ECM deposition and cell volume change (Fig 4.5). We showed that cell behavior during skeleton shaping could be systematically dissected using these quantitative approaches, which can also be adapted for studies of diverse morphogenetic events with minor modifications.

Previous work analyzing retroviral-based clones of cells with static imaging suggested that PZ cells undergo a stereotyped rearrangement after division to create elongated ranks of related cells [20, 21]. Here we observed such cell rearrangement at a very low frequency.This low frequency of cell rearrangement out of the 472 cells we analyzed might be due to the shorter timescale in the current experiments. The virally-labeled clones of cells in chick humerus [20] were analyzed five days after viral infection. In that study, the typical labeled patch consisted of about four cells, suggesting a slow cell division rate, and permitting the subsequent rearrangement to be slow. Our live imaging of the isolated metacarpal lasted for only 55 hours. Alternatively, it may be that the formation of ranks of cells takes place earlier in the metacarpal, while cell movement at later stages are restricted due to the different microenvironment. In the future, it will be important to examine the expression of col2a and other ECM proteins during different developmental stages. A recent report of live imaging analysis in a craniofacial cartilage showed that cell rearrangements take place by two daughter cells spreading along their interface [21]. However, the slow elongation rate of the cranial cartilage and the mosaic nature of the GFP-labeled tissue used in that study make it difficult to determine the quantitative contributions of this mode of cell

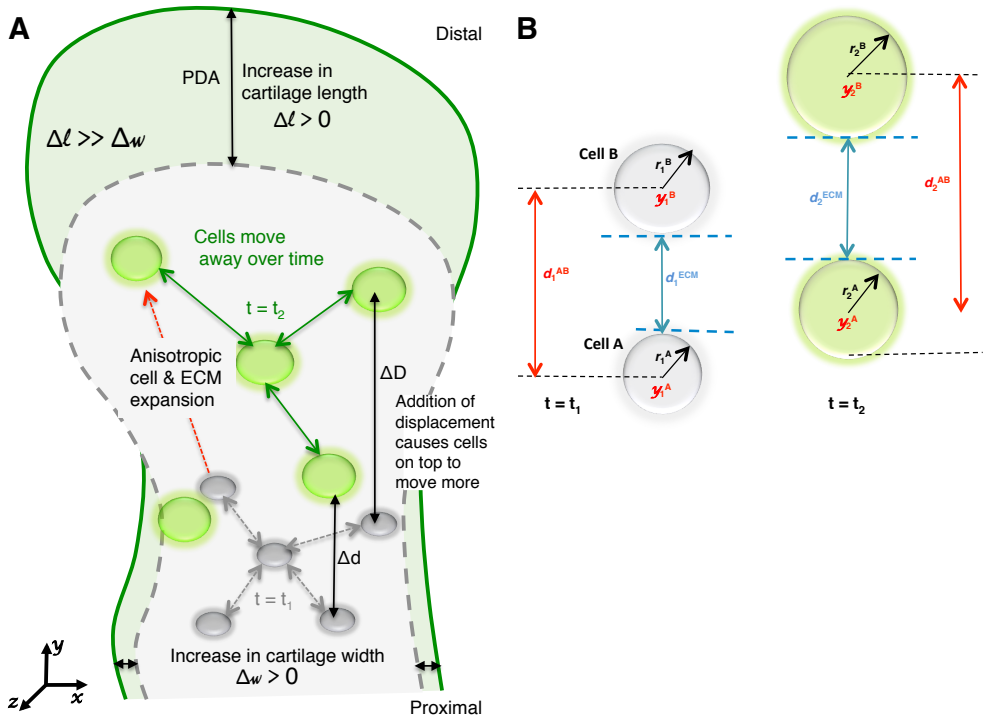


Figure 4.6: **Quantitative model of embryonic cartilage elongation**
(A) The overall picture of cartilage growth: cell spreading contributed by the anisotropic expansion of ECM and cell volume (cells depicted in gray at t_1 and in green at t_2) results in more increase in length than width of the tissue (from the dashed grey outline at t_1 to the solid green line at t_2). (B) Cartilage elongation is a cumulative effect from ECM and cell size increases along the y axis, as depicted for a pair of nearest-neighbor cells A and B, at two different time points t_1 and t_2. At time t_1 (t_2), cell A and B are at positions $y_1^A(y_2^A)$ and $y_1^B(y_2^B)$, and have radii $r_1^A(r_2^A)$ and $r_1^B(r_2^B)$, respectively. The distance between their surfaces, $d_1^{ECM}(d_2^{ECM})$, represents the span of ECM between the two cells. As the cells transition from t_1 to t_2, the increase in intercellular distance (d_1^{AB} to d_2^{AB}) is the sum of the corresponding increases in ECM (d_1^{ECM} to d_2^{ECM}) and the increase in cell sizes (r_1^A to r_2^A, r_1^B to r_2^B).

rearrangement to skeleton shaping. Thus, even in systems reporting cell rearrangements in limb skeleton or in cranial skeleton [22-26], it seems unlikely to be the driving force for skeleton elongation.

Immunohistochemistry shows that col2a fills the intercellular space in the PZ (Fig. 4.5 A,B). In the case of Col2a mutant mice cells are highly misoriented and the cartilage elements are significantly shorter in the PZ [22]. Thus, based on both our circumstantial and published genetic evidence, we propose that col2a deposition, cell morphology and cell spreading coordinate with each other to regulate cartilage elongation. The causal relationship remains uncertain. One attractive hypothesis is that the elongated shape of the PZ cells might encourage more col2a deposition, which in turn might provide the environment that encourages the elongated cell shape. Non-canonical Wnt signaling has been shown to play a major role in the molecular decision making during cartilage growth [12, 20, 25, 26]. Future studies, examining the consequences of such alteration in col2a and wnt signaling, using combined molecular perturbations with the analytical toolsets described in this chapter, may help shed light on the relative roles played by physical interactions, the ECM and chemical signaling in the creation of proper tissue architecture and morphology during skeletogenesis.

4.8 References

1. Karsenty G. Transcriptional control of skeletogenesis. Annu Rev Genomics Hum Genet. 9:183-96. Review. (2008)

2. Lefebvre V, Bhattaram P. Vertebrate skeletogenesis. Curr Top Dev Biol. 90:291-317. (2010)

3. Rafiq K, Cheers MS, Ettensohn CA. The genomic regulatory control of skeletal morphogenesis in the sea urchin. Development. 139(3):579-90. (2012)

4. Shum L, Coleman CM, Tuan RS. Morphogenesis and dysmorphogenesis of the appendicular skeleton. Birth Defects Res C Embryo Today. 69(2):102-22. (2003)

5. Summerbell D. Evidence for regulation of growth, size and pattern in the developing chick limb bud. J Embryol Exp Morphol. 65:129?150. (1981)

6. Kronenberg HM. Developmental regulation of the growth plate. Nature 423, 332-336 (2003)

7. Romereim SM, Dudley AT. Cell polarity: The missing link in skeletal morphogenesis? Organogenesis. 7(3):217-28. Review. (2011)

8. Vortkamp A et al., Regulation of rate of cartilage differentiation by Indian hedgehog and PTH-related protein. Science 273, 613-622. (1996)

9. Lanske B et al. PTH/PTHrP receptor in early development and Indian hedgehog-regulated bone growth. Science 273, 663-666. (1996)

10. De Luca F et al., Regulation of growth plate chondrogenesis by bone morphogenetic protein-2. Endocrinology 142, 430-436. (2001)

11. Mancilla E. E, De Luca F, Uyeda J. A, Czerwiec F. S, Baron J. Effects of fibroblast growth factor-2 on longitudinal bone growth. Endocrinology 139, 2900-2904. (1998)

12. Yang Y, Topol L, Lee H, Wu J. Wnt5a and Wnt5b exhibit distinct activities in coordinating chondrocyte proliferation and differentiation. Development 130, 1003-1015. (2003)

13. Dodds, G. Row formation and other types of arrangement of cartilage cells in endochondral ossification. Anat Rec 46, 385-399 (1930)

14. Li,Y., Trivedi, V., Truong,T.V., Koos,D.S., Lansford,R., Chuong,C.M., Warburton,D., Moats,R.A., Fraser S.E. Dynamic imaging of the growth plate cartilage reveals multiple con- tributors to skeletal morphogenesis Nature Communications (2015) 6:6798 doi: 10.1038/ncomms7798

15. Megason SG, Fraser SE. Digitizing life at the level of the cell: high-performance laser-scanning microscopy and image analysis for in toto imaging of development. Mech Dev 120, 1407-20. (2003)

16. Denk W, Strickler JH, Webb WW. Two-photon laser scanning fluorescence microscopy. Science 6;248(4951):73-6. (1990)

17. Li Y, Ahrens M, Wu A, Liu J, Dudley AT. Calcium-calmodulin dependent protein kinase II activity regulates the proliferative potential of growth plate chondrocytes. Development 138, 359-70. (2011).

18. Supatto W, McMahon A, Fraser SE, and Stathopoulos A: Quantitative imaging of collective cell migration during Drosophila gastrulation: multiphoton microscopy and computational analysis. Nature Protocols, 4(10), 1397-1412 (2009).

19. Sato Y, Poynter G, Huss D, Filla MB, Czirok A, Rongish BJ, Little CD, Fraser SE, Lansford R. Dynamic analysis of vascular morphogenesis using transgenic quail embryos. PLoS One. 14;5(9):e12674. (2010)

20. Li Y, Dudley AT. Noncanonical frizzled signaling regulates cell polarity of growth plate chondrocytes. Development 136, 1083-92. (2009)

21. Romereim, S. M., Conoan, N. H., Chen, B. and Dudley, A. T. A dynamic cell adhesion surface regulates tissue architecture in growth plate cartilage. Dev. Camb. Engl. 141, 2085-2095 (2014).

22. Barbieri O, Astigiano S, Morini M, Tavella S, Schito A, Corsi A. Depletion of cartilage collagen fibrils in mice carrying a dominant negative Col2a1 transgene affects chondrocyte differentiation. Am J Physiol Cell Physiol. 285(6):C1504-12. (2003)

23. Aszodi, A., Hunziker, E.B., Brakebusch, C., and Fassler, R. Beta1 integrins regulate chondrocyte rotation, G1 progression, and cytokinesis. Genes Dev 17, 2465-2479. (2003).

24. Saito A, Hino S, Murakami T, Kanemoto S, Kondo S, Saitoh M. Regulation of endoplasmic reticulum stress response by a BBF2H7-mediated Sec23a pathway is essential for chondrogenesis. Nat Cell Biol. (10):1197-204. (2009)

25. Gao B1, Song H, Bishop K, Elliot G, Garrett L, English MA, Andre P, Robinson J, Sood R. Wnt signaling gradients establish planar cell polarity by inducing Vangl2 phosphorylation through Ror2. Dev Cell. 20(2):163-76. (2011)

26. Kuss P, Kraft K, Stumm J, Ibrahim D, Vallecillo-Garcia P, Mundlos S. Regulation of cell polarity in the cartilage growth plate and perichondrium of metacarpal elements by HOXD13 and WNT5A. Dev Biol. 385(1):83-93. (2013)

Chapter 5

Spatio-temporal differences in dystrophin dynamics at mRNA and protein levels

5.1 Motivation

Dystrophin (Dmd) is an essential structural protein in muscle development and homeostasis, responsible for the maintenance of skeletal muscles. It links the extracellular matrix to the cytoskeletal actin filaments in muscle fibers through the Dystrophin Associated Protein Complex (DAPC) [1, 2]. Mutations in components of the DAPC result in muscular dystrophy, a condition characterized by progressive muscle weakness and degeneration. Dmd mutations are the cause of approximately 90% of all human forms of muscular dystrophies. In particular, Dmd is the gene that is mutated in most cases of Becker Muscular Dystrophy and Duchenne Muscular Dystrophy (Bmd and Dmd, respectively) [3-6]. Dmd is an X-linked recessive, fatal disorder. It has been estimated that 1 in 3,500 to 5,000 boys suffers from Dmd [7, 8], and that approximately three to six of every 100,000 live births are affected by Bmd [9, 10]. Several vertebrate models have been developed to study Dmd [11], ranging from mdx mice [12, 13] and dystrophic golden retriever dogs [14], to Dmd-deficient cats [15], and mutant zebrafish [16-18]. The clarity and accessibility (optical and genetic) of zebrafish embryos make them ideal for following problems such as skeletal muscle development and somitogenesis [16]. Several zebrafish alleles of *dmd* mutations were identified by the Tubingen screen [17]. *Sapje* and *Sapje*-like mutants recapitulate the muscle de-

generation phenotypes observed in humans [18,19], providing a good model for the study of muscular dystrophies [20]. These genetic models have provided many insights into the molecular lesions that can lead to muscular dystrophy; however, real-time monitoring of Dmd expression and localization in vivo, in order to assess both wildtype muscle development and progression of muscle dystrophy at the cellular and protein levels, has not been accomplished. Although multiple studies demonstrate that Dmd is an important player in muscle development, structure, maintenance and signaling, all analyses of Dmd localization have been based on fixed specimens. For example, in the live imaging study of muscle fiber attachment in *Sapje* mutants [21], the Dmd protein dynamics could not be assessed. As muscular dystrophy is a progressive disorder, the ability to follow Dmd and the progression of the disease in real time would be informative; however, there has been no animal model with a fluorescently tagged Dmd protein that would permit such studies.

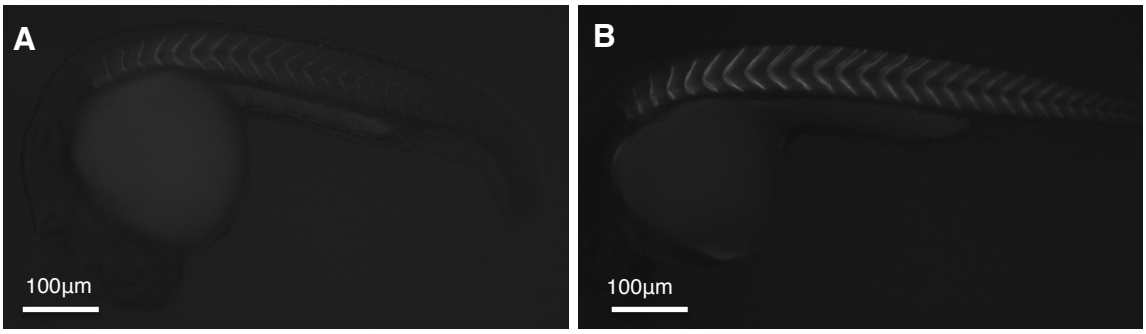


Figure 5.1: **Dmd-Citrine expression in zebrafish FlipTrap line**
Wide-field fluorescent images of zebrafish embryos at (A) 24hpf (B) 32hpf. The expression becomes stronger with development.

In case of zebrafish, as the somites increase in number, from 3 to around 22 over a period of 9 hours (at 28.5^oC), each somite changes its shape from cuboidal to chevron shortly after its formation. These events span over 1mm of developing trunk and tail. This is concurrently accompanied by production and localization of Dmd that act as structural elements, linking the ECM and actin filaments in muscle fibers. The modular organization of the Dmd protein into large, well-conserved protein domains allows a compact insertion of Citrine fluorescent protein, without disrupting the function of Dmd in FlipTrap zebrafish lines [22]. In this chapter I describe the characterization of a gene trap line in which

full-length Dmd is fused to the fluorescent protein Citrine (*Gt(dmd-citrine)*^{ct90a}) (Figure 5.1) and demonstrate a spatial difference of *dmd* transcript localization during development that is not reflected at the protein level. Real-time analysis in (*Gt(dmd-citrine)*^{ct90a} embryos show that, unlike its transcript, Dmd protein is expressed at the myosepta as soon as it is detected, and this protein localization becomes more defined and sharper as the skeletal muscle develops.

5.2 Temporal differences in the localization of dmd transcript and Dmd protein during somite development

Dmd transcript has been reported to be spatially distinct in both mature and developing muscles.In developing zebrafish muscle, dmd mRNA is initially detected throughout the cytoplasm of myoblasts and localizes to the transverse myosepta by 24hpf [23-25]. To assess the localization of *dmd* mRNA and determine if the *dmd-citrine* transcript recapitulates the endogenous localization of *dmd*, we employed in situ Hybridization Chain Reaction (HCR) [48] to simultaneously detect *dmd* and *dmd-citrine* transcripts in FlipTrap embryos. Similar to the experimental procedures outlined in chapter 2, we found that the expression of the fusion transcript is same as that of the endogenous transcript (Figure 5.2). Our imaging shows that *dmd* localizes predominantly to the transverse myosepta in 48hpf embryos. Additionally, *dmd* and *dmd-citrine* probes localize as dots in the nuclei suggesting that the HCR probes are labeling nascent transcripts (Fig 5.2 I-J, N-O). The citrine probe labels a single dot in heterozygous embryos and two dots in homozygous embryos; the dmd probe labels two dots in all embryos (Fig 3 I,N inset, Fig 3 J, O). These data are consistent with heterozygous embryos having a single copy of the gene trap and homozygous embryos having two copies. To further characterize the localization of *dmd* transcript, we performed a time-course of *dmd* expression using chromogenic *in situ* hybridization. Assessing the expression of *dmd* in one hour developmental time intervals starting at 19.5 hpf, we found that *dmd* transcript switches from a non-localized diffused pattern in the cytoplasm of the muscle cells at 21.5hpf, to a restricted pattern at the transverse myosepta at 22.5 hpf (Fig 5.3 4 A-D), where it remains throughout development. The same switch from a diffused to

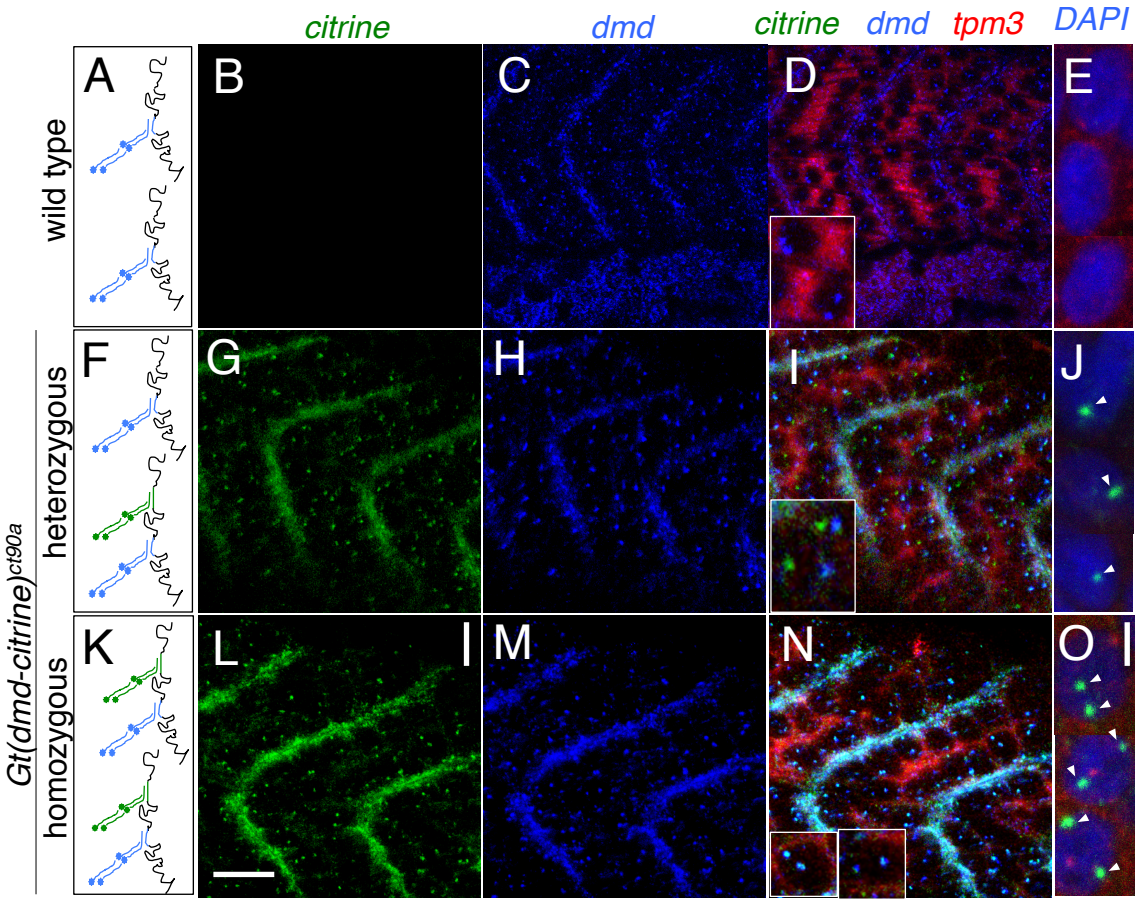


Figure 5.2: **Differential expression of dmd-citrine in homozygous and heterozygous embryos via *in situ* HCR**
(A, F, K) Schematic of *in situ* HCR experiment. (B-E, G-J, L-O) Representative confocal images showing *dmd* (blue) and *dmd-citrine* (green) transcripts in wild type (B-E), heterozygous (G-J) and homozygous (L-O) $Gt(dmd\text{-}citrine)^{ct90a}$ embryos. Tmp3 transcripts counter-stain muscles in red. Insets in (D),(I) and (N) show zoomed in view of transcription sites in the nuclei. Two colocalized cyan dots are detected in homozygous embryos (N) whereas only one dot is detected in the nuclei of heterozygotes (I) while no *citrine* dots are detected in wild types embryos (B,D), consistent with the copy number of *citrine* inserted in the respective embryos. (E, J, O) Magnified images of the three types of embryos with nuclei stained in DAPI and the *citrine* transcripts indicated by arrowheads. Scale bar (B-D, G-I, L-N) 20μm (E,J,O)10μm

localized pattern is observed when detecting *dmd* transcript by HCR in situ hybridization with the added advantage of detecting the nuclear dots at different developmental stages (Fig 5.3 F, J).

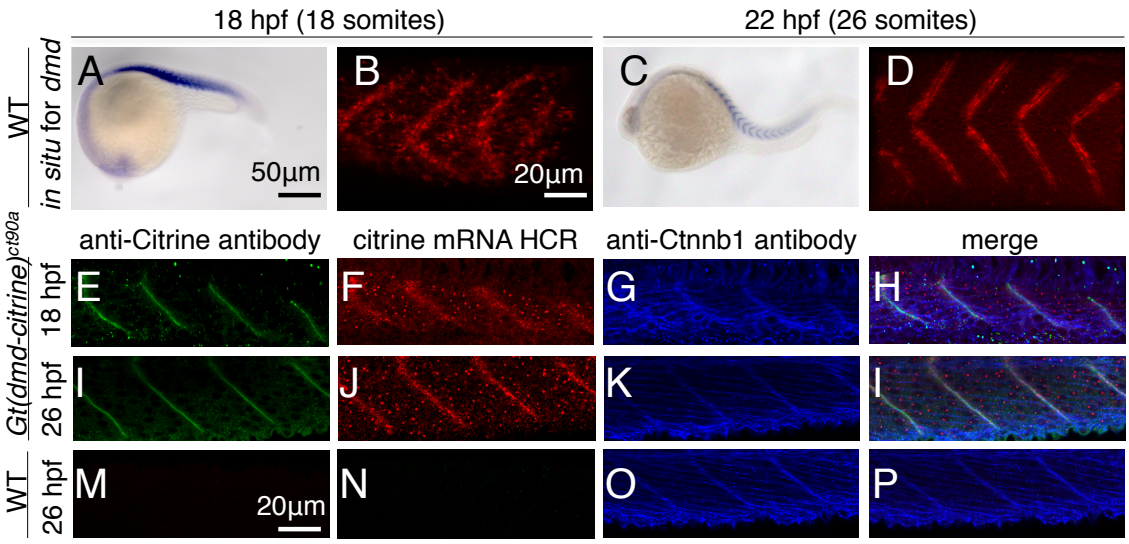


Figure 5.3: **Differential spatial expression of dmd mRNA and Dmd protein** (A-D) Representative widefield (A,C) 3D rendered confocal images (B,D) for chromogenic *in situ* hybridization for *dmd* in wild type embryos at 18hpf (A,B) and 22hpf (C,D). (E-P). Representative confocal images of *dmd-citrine* transcript (F,J,N) detected by *in situ* HCR (red) and antibody staining for Dmd-citrine protein (green) in FlipTrap and wildtype embryos at 18hpf and 26hpf. Counter-stain with antibody to Ctnnb1 in shown in blue (G,K,O). Comparison of protein and transcript expression at 18hpf and 26 hpf shows that Dmd protein localize exclusively to the myosepta while dmd-citrine mRNA is expressed in the cytoplasm at 18hpf and becomes more localized to the myosepta at 26hpf. *dmd-citrine* expression in the nucleus can be seen as dots that appear similar in distribution between 18hpf and 26hpf. (M-P) Dmd-citrine protein (M) and transcript (N) are not detected in wildtype embryos.

The differential distribution of *dmd* transcript led us to ask whether Dmd protein exhibits a similar change in localization during development. To assess mRNA and protein in the same sample, we performed in situ HCR to detect *dmd* transcript followed by immunohistochemistry to detect Dmd protein with an antibody to Citrine (Fig 5.3 E-P). This dual detection approach revealed that Dmd protein is restricted to the myosepta from the onset of Dmd protein expression, at 18 hpf, when dmd transcript is diffusely dispersed in the cytoplasm (Fig 5.3 E-H). By 26 hpf both dmd transcript and protein are localized to the myosepta with nascent dmd transcripts detected as small dots in nuclei of muscle fibers

(Fig 5. 3 I-L).

5.3 Dynamics of Dmd protein localization during somitogenesis

The Citrine fluorescent signature in the *ct90a*^{GT} line enables real-time analysis of Dmd localization during somitogenesis (Fig 5.4). Dmd-citrine could be detected at low levels in the somite boundaries starting at 18-18.5 hpf (Fig 5.4 A, E). The expression increases in intensity from the anterior to posterior somites. In vertebrates, the somatic muscles develop from the paraxial mesoderm in an anterior-to-posterior wave. In zebrafish, this anterior-to-posterior developmental pattern initiates after gastrulation (12 hpf) during which *dmd* transcripts have been detected [24]. However, we detect Dmd-citrine protein localization at the somite boundaries 6 hours after transcripts can be detected, at 18 hpf, indicating that there is a significant delay between transcriptional activation and protein maturation.

To quantify the spatio-temporal dynamics of Dmd-citrine expression, we mapped the pixel intensity of the fluorescence signal over time (Fig 5.4 I-L). Kymographs of the Dmd-citrine expression show that the Dmd-citrine protein shifts in space over time (Fig 5.4 I). However, fluorescent images indicate that Dmd-citrine expression remains at the myoseptum (Fig 5.4 A-H) suggesting that the shift in space may be due to movement of the tissue or growth of the embryo. The shift in Dmd-citrine expression is unlikely due to an increase in the overall size of the somites, as the width of each somite remains constant and the distance between each myoseptum remains the same (Fig 5.4 A-H). The shift in Dmd-citrine protein expression is accompanied by an increase in the total area in which Dmd-citrine can be detected in each developing somite (Fig 5.4 I, L), indicating that more Dmd-citrine proteins are accumulating at the myoseptum over time. To quantify the relative levels of Dmd-citrine in time, we mapped the fluorescent intensity from the time-lapse analysis. We find that the relative mean pixel intensity increases over developmental time, indicating that more fluorescent signal can be detected per unit area (Fig 5.4 K). Additionally, more pixels have fluorescent signature over developmental time (Fig 5.4 L), indicating that as the somites mature more Dmd-citrine proteins are localizing to the myosepta.

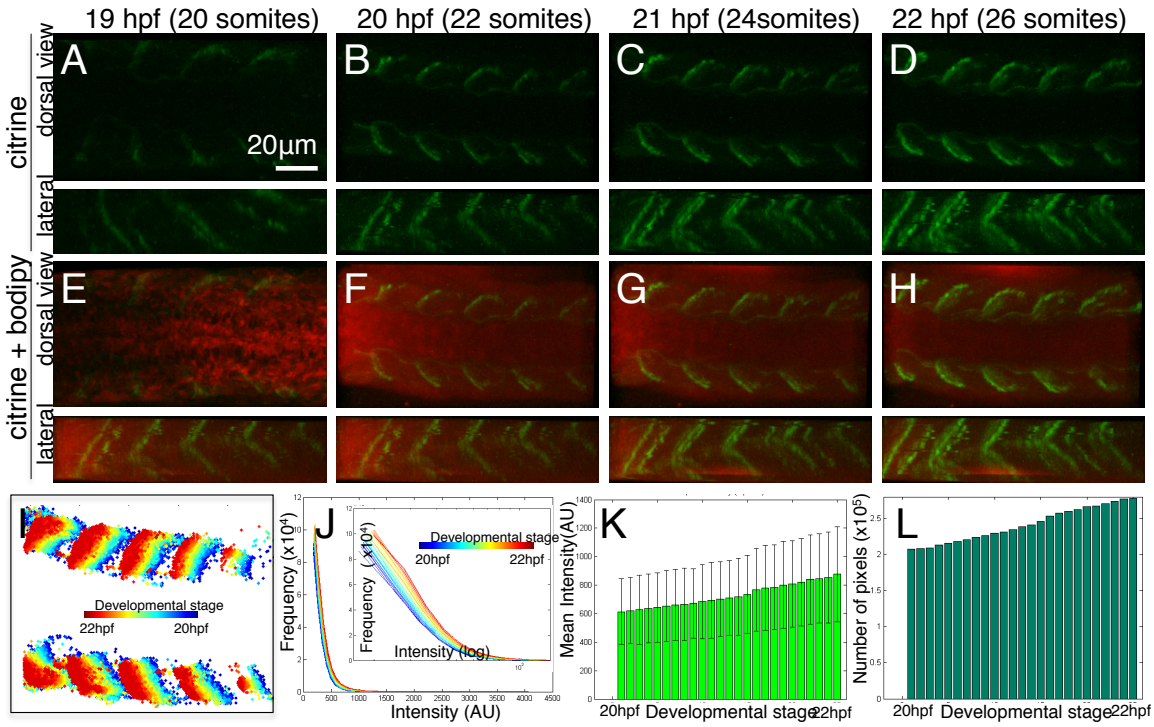


Figure 5.4: **Dmd Protein expression over development**
(A-H). 3D projections of confocal Z-stacks images in developing trunk of Gt(dmd-citrine)ct90a embryos from the 20 to 26 somites stages (19-22 hpf), showing expression of Dmd-citrine (a-d, green) and the vital stain BodipyTR methyl ester (e-h, red) from dorsal and lateral view.(i-l) Quantitation of Dmd-citrine fluorescent signal in confocal time-lapse. (i). Kymograph of Dmd-citrine expression color-coded according to developmental stage. The spatial shift in expression appears to be due to tissue growth and expression enhancement over time. (j). Histogram of fluorescent pixel intensities color-coded according to developmental stage with intensity in X-axis and frequency in Y-axis. Plot shows Dmd-citrine expression increasing over time resulting in shift slope along the X-axis (blue to red). The inset shows the same plotted on a log scale to enhance observation. (k). Mean pixel intensity plotted against development stage showing increase in fluorescent intensity of Dmd-citrine as development progresses. The error bars denote the variation among all the pixels at that particular stage. (l). Total pixel count of Dmd-citrine fluorescent plotted against developmental time showing that Dmd-citrine expression increase in spatial area over time. Scale bar= $20\mu m$.

5.4 Discussion

In this chapter the work on the zebrafish gene trap line, *Gt(dmd-citrine)*^{ct90a} that expresses a Dmd-citrine protein fusion was presented. The line enabled the dynamic study of muscular dystrophy. We have shown that the expression of Dmd-citrine in these embryos recapitulates endogenous Dmd protein and mRNA localization. The tagging of Dmd with citrine did not seem to affect the interaction of Dmd with its DAPC partners. Defects in protein interactions within the DAPC would have lead to dystrophic phenotypes, and we did not observe any phenotypes in homozygous embryos. This transgenic line enables spatio-temporal study of Dmd dynamics in a live vertebrate animal model.

dmd transcript has been described to exhibit spatial differences in localization during development. We have found that *dmd* transcript transitions from a cytoplasmic localization in the myoblast to a discrete localization pattern in the myoseptum between 21.5 hpf and 22.5 hpf (Fig. 5.3). This transition in localization pattern coincides with a developmental stage when myoblasts fuse to form myotubes [26]. The syncytial nature of the myotubes may allow *dmd* transcripts to be discretely localized to the myoseptum. In the individual myoblast, the cell membrane may serve as a physical barrier for localization to the myosepta. Consistent with this hypothesis, we observe localization of *dmd* transcript in the myoseptum in 21.5 hpf embryos when cytoplasmic *dmd* transcript could be detected, indicating that localization of the transcript can occur at these earlier stages in development. Quantitative analysis of the time-lapse data shows that not only does the mean signal intensity of Dmd-citrine increase (by 45%) as development proceeds but the total expression area near the myoseptum increases as indicated by a 40% increase in the total number of pixels (Fig 5.4 I-L). These results suggest that, as the myoblasts transitions to syncytial fibers, more Dystrophin is recruited to the myosepta. The accumulation of Dmd to the myosepta, as the muscle fibers mature, hints towards its importance in providing the mechanical anchoring for contracting muscle.

Interestingly we found that, during early zebrafish embryo development, *dmd* mRNA transcripts first appeared diffused in the cytoplasm before being localized at the myosepta. We did not observe such a differential pattern for Dmd protein. Dmd-citrine is initially

detected at the myoseptum between the 18-19 somite stages prior to the localization of the mRNA at the myoseptum. This temporal difference in the localization of the mRNA and protein to the myosepta suggests that Dmd-citrine is translated or matures into detectable fluorescence at the myoseptum. Immunohistochemistry with antibodies to Dmd suggests that the temporal maturation of Dmd and Dmd-citrine are similar as both proteins are detected at the myoseptum starting at the 18-19 somite stage. At this stage cell fusion should be completed in the anterior somites, but is not yet complete in the posterior somites [27]. Consistent with this anterior to posterior wave of maturation of the somites, we observe a similar anterior to posterior wave of localization of Dmd-citrine to the myoseptum in the time-lapse microscopy (Fig 5.4).

As development progresses, we see significant decrease in cytoplasmic distribution of *dmd* transcript over 4 hours from 18hpf to 22hpf. In situ HCR shows the nascent transcript sites in the nucleus with nearly no localization of *dmd-citrine* or *dmd* mRNA in the cytoplasm by 26hpf when most of the myoblasts have completed fusion (Fig 5.3). There is a possibility that Dmd protein is present in the cytoplasm at levels below our detection limit; however, our data indicate that the majority of the Dmd proteins are localized to the myosepta. Previous work on dynamics of mRNA-protein (mRNP) nucleocytoplasmic transport in mammalian cells has shown that the cytoplasmic diffusion coefficient of dystophin mRNAs is low [28]. It would be of interest to assess the dynamics of Dmd protein that is predicted to be slower given it?s large size (427 KDa protein). Therefore increasing localization of the transcript itself at the myoseptal junctions might facilitate effective translation or maturation of Dmd-citrine into detectable fluorescence at the site of its activity. This localization might be helpful in bypassing the time required for protein diffusion from the cytoplasm of fusing myoblasts or mature myotomes to the somite boundaries, thereby making the availability of Dystrophin, for anchoring, limited only by the rate of on-site translation. The *Gt(dmd-citrine)*^{ct90a} line provides opportunities for future studies to test these hypotheses.

5.5 References

1. Arahata, K., et al., Immunostaining of skeletal and cardiac muscle surface membrane with antibody against Duchenne muscular dystrophy peptide. Nature, 1988. 333(6176): p. 861-3.

2. Zubrzycka-Gaarn, E.E., et al., The Duchenne muscular dystrophy gene product is localized in sarcolemma of human skeletal muscle. Nature, 1988. 333(6172): p. 466-9.

3. Hoffman, E.P., R.H. Brown, Jr., and L.M. Kunkel, Dystrophin: the protein product of the Duchenne muscular dystrophy locus. Cell, 1987. 51(6): p. 919-28.

4. Kunkel, L.M., et al., Analysis of deletions in DNA from patients with Becker and Duchenne muscular dystrophy. Nature, 1986. 322(6074): p. 73-7.

5. Kunkel, L.M., et al., Molecular genetics of Duchenne muscular dystrophy. Cold Spring Harbor symposia on quantitative biology, 1986. 51 Pt 1: p. 349-51.

6. Monaco, A.P., et al., Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. Nature, 1986. 323(6089): p. 646-50.

7. Emery, A.E.H., Duchenne Muscular Dystrophy. . Oxford Monographs on Medical Genetics (2nd ed.). Vol. xv. 1993, Oxford, UK: Oxford Univ. Press.

8. Emery, A.E., Muscular dystrophy into the new millennium. Neuromuscul Disord, 2002. 12(4): p. 343-9.

9. Mostacciuolo, M.L., et al., Reappraisal of the incidence rate of Duchenne and Becker muscular dystrophies on the basis of molecular diagnosis. Neuroepidemiology, 1993. 12(6): p. 326-30.

10. Emery, A.E., Population frequencies of inherited neuromuscular diseases–a world survey. Neuromuscul Disord, 1991. 1(1): p. 19-29.

11. Collins, C.A. and J.E. Morgan, Duchenne's muscular dystrophy: animal models used to investigate pathogenesis and develop therapeutic strategies. International journal of experimental pathology, 2003. 84(4): p. 165-72.

12. Bulfield, G., et al., X chromosome-linked muscular dystrophy (mdx) in the mouse. Proceedings of the National Academy of Sciences of the United States of America, 1984. 81(4): p. 1189-92.

13. McArdle, A., R.H. Edwards, and M.J. Jackson, How does dystrophin deficiency lead to muscle degeneration?–evidence from the mdx mouse. Neuromuscular disorders : NMD, 1995. 5(6): p. 445-56.

14. Cooper, B.J., et al., The homologue of the Duchenne locus is defective in X-linked muscular dystrophy of dogs. Nature, 1988. 334(6178): p. 154-6.

15. Winand, N.J., et al., Deletion of the dystrophin muscle promoter in feline muscular dystrophy. Neuromuscular disorders : NMD, 1994. 4(5-6): p. 433-45.

16. Chen and Jenna L. Galloway The development of zebrafish tendon and ligament progenitors Jessica W. Development (2014) 141, 2035-2045

17. Granato, M., et al., Genes controlling and mediating locomotion behavior of the zebrafish embryo and larva. Development, 1996. 123: p. 399-413.

18. Bassett, D. and P.D. Currie, Identification of a zebrafish model of muscular dystrophy. Clinical and experimental pharmacology and physiology, 2004. 31(8): p. 537-40.

19. Guyon, J.R., et al., Genetic isolation and characterization of a splicing mutant of zebrafish dystrophin. Human molecular genetics, 2009. 18(1): p. 202-11.

20. Gibbs, E.M., E.J. Horstick, and J.J. Dowling, Swimming into prominence: the zebrafish as a valuable tool for studying human myopathies and muscular dystrophies. The FEBS journal, 2013. 280(17): p. 4187-97.

21. Berger, J., et al., Dystrophin-deficient zebrafish feature aspects of the Duchenne muscular dystrophy pathology. Neuromuscul Disord, 2010. 20(12): p. 826-32.

22. Trinh, L.A., et al., A versatile gene trap to visualize and interrogate the function of the vertebrate proteome. Genes and development, 2011. 25(21): p. 2306-20.

23. Bohm, S., et al., Dystrobrevin and dystrophin family gene expression in zebrafish. Gene Expr Patterns, 2008. 8(2): p. 71-8.

24. Bolanos-Jimenez, F., et al., Dystrophin and Dp71, two products of the DMD gene, show a different pattern of expression during embryonic development in zebrafish.

25. Mitsui, T., et al., Preferential subsarcolemmal localization of dystrophin and beta-dystroglycan mRNA in human skeletal muscles. J Neuropathol Exp Neurol, 1997. 56(1): p. 94-101.

26. Snow, C.J. and C.A. Henry, Dynamic formation of microenvironments at the myotendinous junction correlates with muscle fiber morphogenesis in zebrafish. Gene expression patterns : GEP, 2009. 9(1): p. 37-42

27. Srinivas, B.P., et al., A conserved molecular pathway mediates myoblast fusion in insects and vertebrates. Nature genetics, 2007. 39(6): p. 781-6.

28. Mor, A., et al., Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells. Nature cell biology, 2010. 12(6): p. 543-52.

Chapter 6

Decoupling the role of contractility on cardiomyocyte growth *in vivo*

6.1 Motivation

Vertebrate heart is a specialized bi-layered organ that is assembled to function very early on during embryonic development; and consists of endocardium, forming the lumen and myocardium that generates contraction (Figure 6.1A). The process of cardiogenesis starts with changes at cellular level in terms of shape, size, number, organization and their physical environment, where each step is driven by considerable changes in the molecular state of the cells, viz. expressions of RNA and proteins [1,2]. The specialty of cardiac tissue lies in the fact that its constituent cells display a unique interplay of two distinct time scales: the short timescale of beating governs the pumping ability of the organ and hence its functionality, whereas the long developmental time scale dictates its 3D structure through the continuous morphological changes as the tissue is becoming functional. It is therefore very natural to probe into the inter-dependence of the two time scales as crucial for determining the biomechanical implications on cardiogenesis [12].

The clarity and accessibility (optical and genetic) of zebrafish embryos make them ideal for studying cardiogenesis. A growing body of work demonstrates that the insights gained from zebrafish offers important insights into cardiovascular development in amniotes (avian and mouse) [3,4,8]. With the beating starting around 24 hours post-fertilization (hpf), within the two-layered heart tube, the sequences of molecular and physical events induce looping [5] and ballooning [6] to carve out the basic pattern of mature heart, consisting of

two chambers, atrium and ventricle [7] (Figure 6.1A-D).

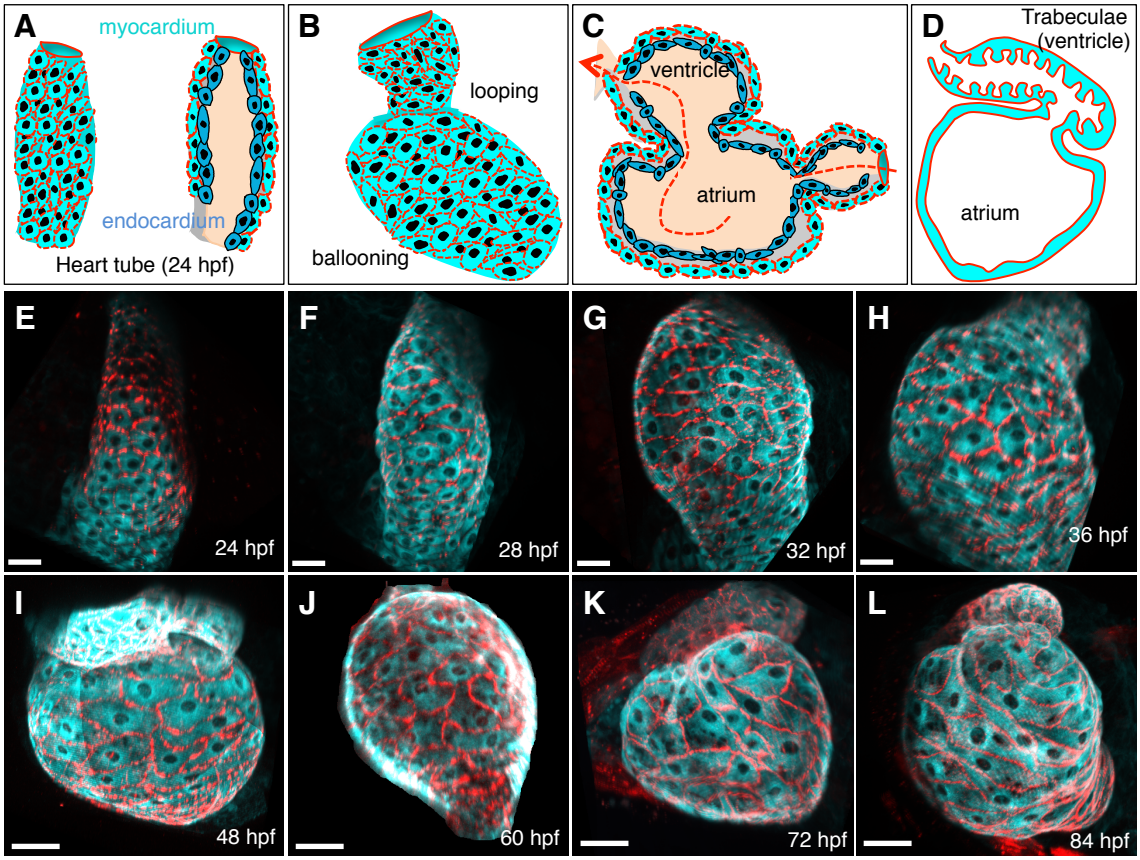


Figure 6.1: **Cardiac development in zebrafish**
(A-D) Schematics depicting the development of heart. Starting with the assembly of a two-layered heart tube (A) at 24hpf, heart consists of endocardium (blue), forming the lumen, and myocardium (cyan) that generates contraction. Subsequent bending and twisting of the organ, starting 36hpf, induces looping and localized bulges along the tube, ballooning (B) to form the two chambers, atrium and ventricle, around 48 hpf (C). Morphological differences throughout the organ become more pronounced via formation of ventricular trabeculae (D), starting 60hpf. (E-L) Representative high-resolution 3D rendered confocal images of developing heart in zebrafish embryos *in vivo* at different stages. Tpm4a-Citrine expression is cytoplasmic (cyan) and restricted to myocardium with a very clean outline of cell boundary provided by Desmin-mCherry fusion protein (red). Nuclei, visible by the negative contrast, can be identified throughout the tissue. Scale bars: $20\mu m$

Given its dynamic nature, the form and function of the organ have been implied to be tightly coupled throughout its different phases of development [8-11], though it still remains unclear how the changes at the cellular level translate into changes in the organ shape *in vivo*. Anatomical defects observed in cases of blocking blood flow [8] or contractility [13,14]

in zebrafish or very recently the dependence of abnormal atrial conduction patterns on hemodynamic forces in other amniotes [15-18] have all supported the role of cardiac contractions in governing the organ architecture. However evidences of normal looping and chamber differentiation observed in silent heart mutants even in absence of contraction [25] present a perplexing scenario for the role of physical forces in cardiac development. In case of cardiac looping proposals of several cellular mechanisms have emerged over the years. In case of zebrafish ventricular development, previous studies have elucidated regionally confined cell shape change as a cellular mechanism for the formation of chamber curvatures [6]. Further, the role of cardiomyocyte hypertrophy has also been examined in the context of ventricular formation wherein increases in surface area and myofibril content were shown to be correlated with chamber formation [21]. Similar studies in chick embryos have proposed differential hypertrophic growth in the myocardium as the main force that cause the heart tube to bend [22]. Retrospective clonal approach within the murine heart has shown evidences of oriented cell growth within myocardium resulting in cardiac chamber morphogenesis [23]. Parallel approaches using computational and physical simulation models have demonstrated that compressive loads resulting from unequal growth of the heart and pericardial cavity result in growth-induced buckling within a confined space inducing the looping in embryonic heart [9,24]. Although informative in understanding morphogenetic mechanisms, such cellular- and organ-level approaches have been largely correlative in nature, if not isolated from each other. Especially in case of cardiomyocytes, the role of their contractility in affecting their own shape, size, organization and thus in turn the overall architecture of myocardium has been poorly understood. Particularly, since the myocardium does not experience hemodynamic forces directly, all previous studies on the effects of perturbed blood flow on cardiomyocytes have been have been correlative in nature and primarily in 2D with image projections [10,21]. Understanding cardiomyocytes morphogenetic events in terms of forces driving and emerging from local changes in cell shape still remains a challenge.

In this chapter I address the question of inter-dependence of the hitherto mentioned dual scales of beating and development and its role in cardiogenesis in terms of cellular changes

in shape and size in myocardium. Based on high-resolution confocal imaging of zebrafish heart at various developmental stages, the changes in cell size and shape *in vivo* in 3D are reported. Based on the distribution of cell sizes, a previously unappreciated heterogeneity in cardiomyocytes between the atrial and ventricular populations is discovered. We have been able to detect signs of mutli-nucleation in zebrafish hearts as early as 48hpf. Utilizing the mutagenetic approaches, easily available for zebrafish embryos, we also examine the role of contraction on the cellular morphology and the overall tissue geometry. Additionally we model cardiac looping computationally, based on the cellular parameters inferred from our imaging results, to bridge the gap between cellular- and organ-level information. Together, the work presented in this chapter provides a model to re-evaluate the role of contraction in cardiogenesis and possibly decouple the dependence on contractile forces in case of cardiomyocyte growth, which eventually define the organ shape and function.

6.2 Cell shape changes in cardiomyocytes over development

Using double transgenic FlipTrap lines [26], pioneered in our lab, that label cardiomyocytes with cytoplasmic citrine and membrane localized mCherry, we can visualize myocardium with sub-cellular resolution. In Figure 6.1E-L we show the changes in the myocardial tissue as the linear heart tube transforms into a two-chambered organ. Structures such as ventricular trabeculae are identified in 3D as late as 84 hpf (Fig. E.2 A-C,C(inset)). Nuclei, visible by the negative contrast, can be identified throughout the tissue (Fig. 2B,C, arrowheads) and in all cross sections (Fig. E.2, arrowheads). We show that as early as 48hpf, bi-nucleated cardiomyocytes can be seen in the atrium (Figure 6.2B) and the number progressively increases with several multi-nucleated cells around 84hpf (Fig 6.2C, D). Simultaneously the cells change their shapes as well with development. In Fig. 6.2G, we show several representative cells at different stages. The cells remain approximately circular only at very early stages (around 24hpf) but elongate along one axis starting 28hpf. As the heterogeneity in cell shape becomes pronounced with development, quantitative measures of cell shape based on 2D projections can bias the estimates. Especially as the cells at late stages (72hpf onwards) curve following the overall tissue topology, 2D projections incorrectly

estimate the relative differences between cell sizes (Fig E.3 A-E). Our transgenic label allows the segmentation of individual cardiomyocytes to quantify shape and size in 3D (Figure 6.2 A-C).

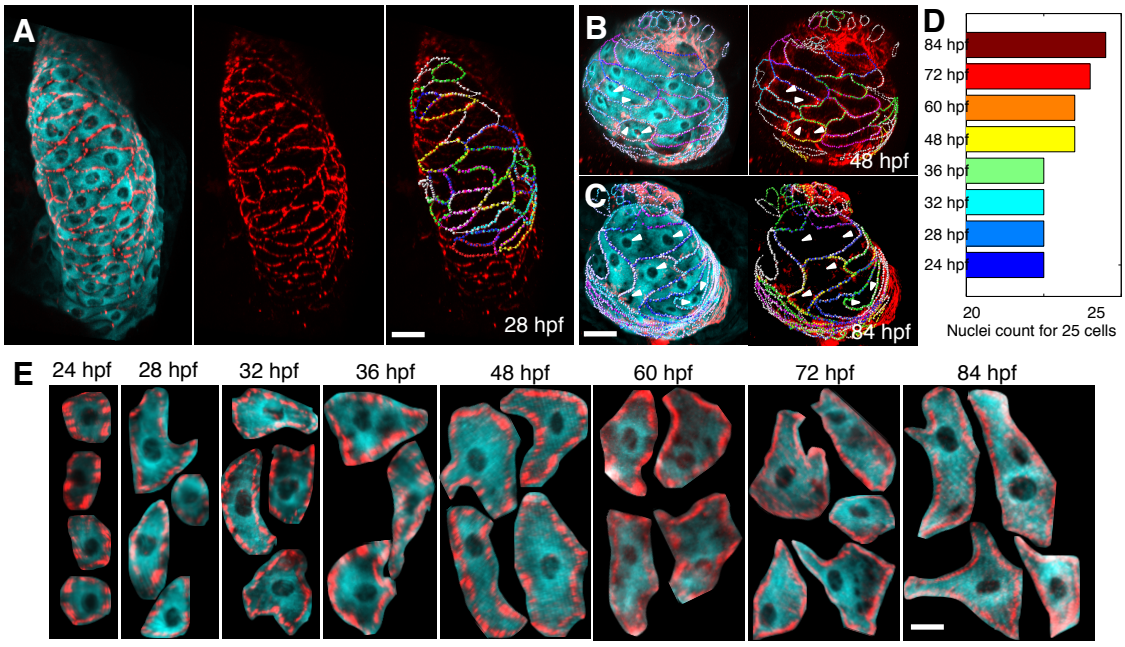


Figure 6.2: **Cardiomyocytes alter their shape and become multi-nucleated over development**
(A-D) Representative 3D rendered confocal images of zebrafish heart at 28 hours post fertilization (hpf) shows cardiomyocytes (cyan, left) with clear membrane (red, middle) to allow for segmentation (right). (B-C) Similar images as (A), depicting two-chambered heart at 48 hpf (B) and 84 hpf (C). Arrowheads denote the position of nuclei within the cell body. (D) Nuclei count for 25 randomly chosen cells at every stage show that cardiomyocytes become multi-nucleated over development (E) Representative cells shown at different stages reveal that cells remain approximately cellular only at very early stages (around 24hpf) but elongate along one axis starting 28hpf. The heterogeneity in cell shape becomes pronounced with development. Scale bars: (A-C) $20\mu m$; (E) $5\mu m$

6.3 Heterogeneity in size of cardiomyocytes over development

We calculated the perimeter of cardiomyocytes over development by segmenting our imaging data in 3D. Similar to observations made before [10,21] the cells increase in both perimeter

and area over time (Fig 6.3A). We further calculated the variance in the perimeter of these cells as shown as error bars in the figure. This variance remains low around 24hpf, indicating that the cells are similar in size throughout the organ. As development proceeds, starting 36hpf, the variance in cell size increases to further imply that cells differ from each other in size more and more with development. As these changes in size and variance are concurrent with the emergence of the two chambers, we plotted cellular perimeters in the atrium and ventricle separately (Fig 6.3B). Following the process of looping (36 hpf, Figure 6.1H), when the basic architecture for the emergence of the two chambers is evident, the cells in ventricle do not change their size (perimeter) from 36hpf to 84hpf (Figure 6.3B). In fact, the average ventricular cell size ($60\mu m$) is very close to that of the undifferentiated cell population at 24hpf ($70\mu m$). Additionally the small error bar indicates that cells remain similar to each other in terms of their size. During this period the atrial cells change their perimeter by approximately 2 fold, on average, at 84hpf and also show large variation across the cell population at each stage.

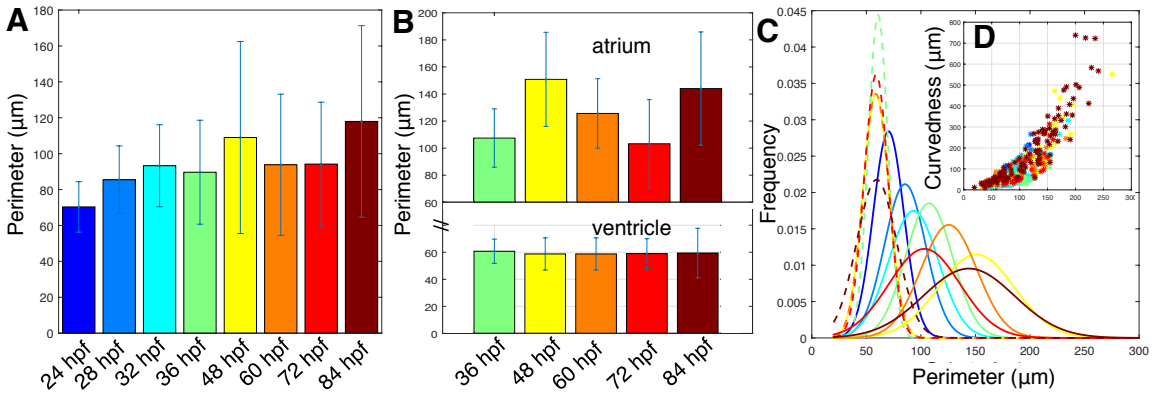


Figure 6.3: **Cardiomyocytes show heterogeneity in size over development** (A) Perimeter of cardiomyocytes at various stages of development (n $\geq$ 60 cells, n = 4 embryos at every stage). Error bars denote standard deviation in perimete. It can be seen that the cells not only increase in size but also in heterogeneity over development. (B) Cell perimeter plotted for two chambers separately over development starting 36 hpf. (C) Distribution of cell perimeters (Approximated to Gaussian distributions) reported in (A). Starting around 24 hours post fertilization (hpf) the distribution remains largely unimodal but as the cells become specified as ventricle and atrial, bimodality becomes apparent. Solid lines denote atrial cells and corresponding broken lines denote ventricular cells starting 36hpf. (D) Scatter plot for curvedness (defined as deviation from a plane) and cell perimeter suggests that cardiomyocyte size correlates with the deviation from the plane.

In order to further understand the nature of cell size distributions between the chambers, we plotted the distribution of perimeters in Fig 6.3C. Around 24hpf the distribution remains largely unimodal indicating that all cells belong to the same population. As the atrium and ventricle develop and mature, bimodality starts becoming apparent (Figure 6.3C, solid and dashed lines). This is indicative of the size difference between the cells in the two chambers. This implies that cells undergo further specialization throughout the organ as the heart matures. Together our results display a previously unappreciated heterogeneity in cell size of cardiomyocytes.

6.4 Cardiomyocytes' shape and size correlate with the tissue curvature

Following the onset of multi-nucleation, the cell shapes vary largely across the cell population at a given stage, with shapes that cannot be classified as regular geometrical objects, like strictly circular or elliptical or otherwise (Figure 2 B,C,G). Therefore to understand the shape changes in cardiomyocytes we define new metrics based on 3D segmentations already achieved. We define two principle curvatures of the cells, one along the direction of tissue axis and the other orthogonal to it (Fig E.3 F). Qualitative visual inspection of the images (Fig 6.2E), already shows that the cells become more oblong as development proceeds. It can also be seen that cells at 24hpf are circular and smaller as compared to the more oblong ones at later stages implying that larger cells deviate from circular shape. For example the cells in the atrium become oblong and large at the same time, where as cells in ventricle remain largely circular owing to their small size (Fig. 6.2B,C). Based on these results deviation from a planar topology seems to be suggestive of cell size. We estimated the local curvature in tissue as measured by the deviation of cell from a plane and termed it as curvedness (Fig E.3G). We show that cardiomyocyte size correlates with the deviation from the plane (Fig 6.3 D). It can be appreciated that irrespective of the developmental stage and independent of the chamber-specificity of a cell at a given stage, smaller cells are likely to be located in regions where the local curvature is smaller, alternatively the local topology is flatter.

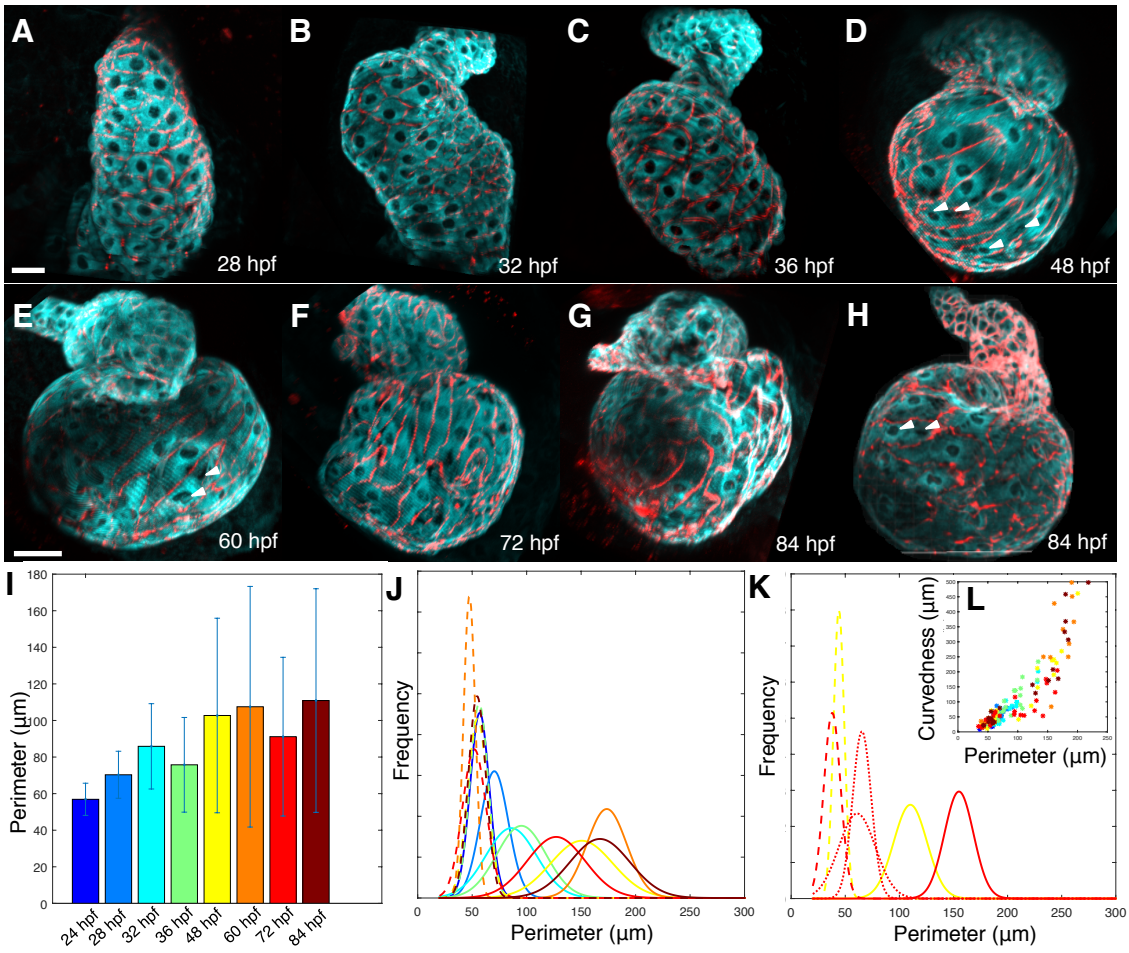


Figure 6.4: **Cardiac mutants with contractile defects show similarity in cellular growth to wild type embryos**
(A-H) Representative 3D rendered confocal images of developing heart *in vivo* at different stages (n = 4 each), in mutant zebrafish embryos, *Gt(tpm4a-citrine)*$^{ct31a/ct31a}$*;Gt(desm-mCherry)* $^{ct122aR/+}$ (A-G) and anti-*amhc* morpholino injected embryos *Gt(tpm4a-citrine)* $^{ct31a/+}$*;Gt(desm-mCherry)* $^{ct122aR/+}$ (*wea* mutant) (H). As before cytoplasmic Tpm4a-Citrine expression (cyan) and membrane Desmin-mCherry expression (red) allows us to visualize cardiomyocytes with sub-cellular resolution. Scale bars: $20\mu m$ (I) Perimeter reported for cardiomyocytes at various stages of development (n $\geq$ 60 cells, n = 4 embryos at every stage) in homozygotes. Error bars denote standard deviation in perimeter at corresponding stages.(J) Distribution of cell perimeters reported in (I), with the same color-coding, and approximated to Gaussian distributions. Solid lines denote atrial cells and corresponding broken lines denote ventricular cells starting 36 hours post fertilization (hpf). (K) Distribution of cell perimeters, with approximations to Gaussian distributions, reported for *wea* mutant at 48hpf (yellow, solid and broken) and 72 hpf (red, solid and broken), and *Gt(tpm4a-citrine)* $^{ct31a/ct31a}$*;Gt(desm-mCherry)* $^{ct122aR/+}$ embryo with arrested cell growth at 72hpf (red, dotted). Solid lines denote atrial cells and corresponding broken lines denote ventricular cells. (L) Scatter plot for curvedness and cell perimeter, for cells in (I), suggests dependence of cardiomyocyte size on curvedness, similar to wild type embryos

6.5 Cardiac mutants with contractile defects indicate independence of cellular size and contraction

Based on the results obtained with wild type embryos, we probed into the role of contraction on cell shape and size using two types of contractility defect mutants. First we make use of the FlipTrap line *Gt(tpm4a-citrine)* $^{ct31a/+}$ that happens to have mutant phenotypes in homozygous embryos where cardiomyocytes do not contract (Fig. E.1 C,D). Secondly we use weak atrium (*wea*) mutation that disrupts the atrial myosin heavy chain (*amhc*) gene, and thus atrium loses its ability to contract [13] (Fig E.1 E,F) while the ventricle retains the normal contractile machinery. Similar to wild type embryos, we segment the cardiomyocytes in 3D to analyze the perimeter and shape in the mutant embryos (Figure 6.4, E.2 E.4). As the morphological differences between mutant and wild type hearts in case of *amhc* mutants are not apparent until 48hpf, the analysis for these embryos was done starting 48 hpf.

In Fig. 6.4 (A-C) we show that while homozygous hearts do not contract, they do start ballooning, very early on, as compared to the wild type embryos at comparable stages. The process of looping seems unperturbed in these embryos (Figure 6.4D), with distinct chamber formation (Figure 6.4 E-H). It is also evident by visual inspection that even in the absence of contraction the cardiomyocytes show the signs of both multi-nucleation (Figure 6.4 D, E, arrow heads) as well as distinct cell population based on cell size and shape in the two chambers. Further we show in Fig 6.4 (I) that the cellular heterogeneity in cell size remains unaffected. Fig 6.4 (J) shows that myocardial cells from atrium and ventricle have a have a clear size difference. As before, while the cells in ventricle remain closer to the ones at 24 hpf in size and are similar to each other (Fig. E.4), atrial cardiomyocytes increase in perimeter by nearly 2 fold, on average, at 84hpf, very similar to the observation made in wild type embryos. Remarkably we also show that the curvedness metric determines the size of the cells irrespective of developmental stage and chamber position (Figure 6.4L).

In case of *amhc* mutants (Fig. 6.4 H; E.4A), despite the defects in contractile machinery, the differences between cell sizes in the two chambers remains similar to the observations made in both wild type as well as mutant embryos with no contraction. Figure 6.4K shows

that while the ventricular cell perimeter distribution remain similar in terms of mean and spread between 48 hpf and 72 hpf (yellow and red broken lines respectively), the atrial population increases in size over this period (yellow and red solid lines respectively). In the special case of a chance mutant that we obtained from the batch of homozygotes, where in addition to their inability to contract, cells had stunted growth (Fig E.1H; E.2G-I, E.4D) the heart remains closer to its initial tubular geometry. With no difference in size between the cells in this case (Figure 6.4K, red dotted lines; E.5 J-L), events like ballooning and looping do not take place. Together these mutant analyses suggest that the changes in cardiomyocytes based on cell shape and size, is independent of their contractile ability. The results hint towards inherent molecular mechanisms that control cell growth and draw on a strong correlation with the local curvature to determine cell shape and size.

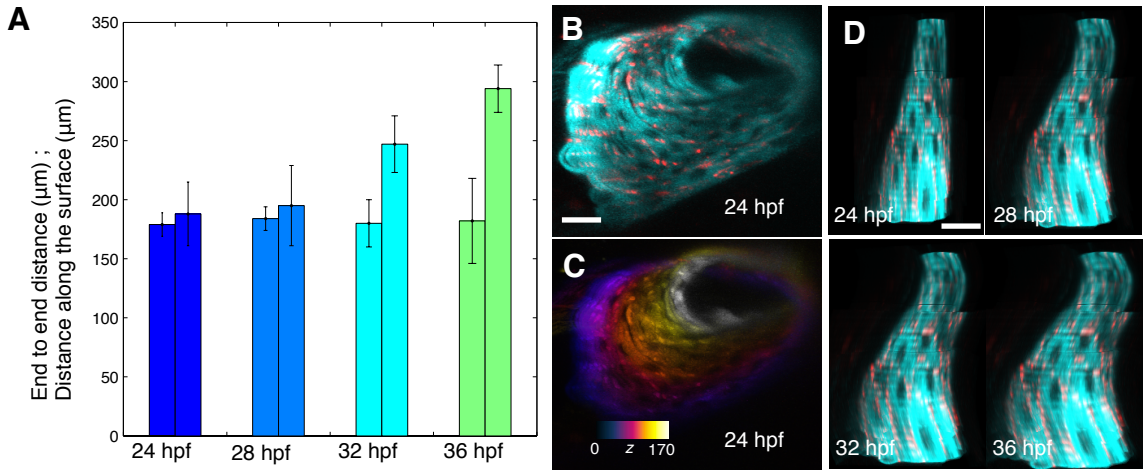


Figure 6.5: **Cardiac looping is an effect of growth-induced buckling due to external spatial packing constraints**
(A) Bar graphs showing the end-to end distances for hearts (left bars at every stage) that remains constant while the tissue length (as measured along its surface, right bars at every stage), increases with development (n = 4 each) Scale bars: $20\mu m$ (B-C) Heart tube at 24 hours post fertilization (hpf) (D-G) Computer simulation of cardiac growth from 24hpf to 36hpf, where the geometry of the cardiac tube at 24hpf (B-D) is taken as the initial geometric input to the model. Changes in geometry are based on the measurements of cell size during this period (Figure 3). Growth at cellular level when translated to tissue level indicates regional bulging (ballooning). Bending (looping) is also induced to accommodate the growth along the tissue axis within the confined space of the pericardium (constant end-to-end distance).

6.6 Cardiac looping is an effect of growth-induced buckling due to external spatial packing constraints

In order to test the possible role of cell shape changes responsible for looping, we performed computational simulations based on the quantitative information derived from the images. To start we measured the end-to-end distance for hearts between 24hpf and 36hpf (Fig. 6.5A), which remained nearly constant. We also measure the tissue length along its surface that increased nearly 1.5 times during this period. This suggests growth-induced buckling due to spatial packing constraints as a possible candidate for inducing looping. To simulate the process of looping, we took the geometry of the cardiac tube at 24hpf (Fig 6.5 B, C) as the initial geometric input to the model. We further allowed growth of the tissue based on the observed increase in cell sizes. To account for the differential growth observed between atrial and ventricular populations, we restrict the growth to the region of the heart tube that will eventually form atrium based on our imaging results. Fig. 6.5 (D) show that as the cells increase in size, this causes a corresponding increase in organ size over time. Regional bulging and bending qualitatively capture the form changes as observed during ballooning and looping in embryonic heart. We therefore show that the growth at cellular level when translated to tissue level provides a strong evidence of growth-induced buckling leading to observed changes at organ level.

6.7 Discussion

As mentioned before, with previous work implying the onset of hemodynamic forces very early on heart development [36,37], the idea of blood flow impacting chamber formation has pursued at great length [10,13,15,17-21]. However the role of contractility on cardiomyocyte shape and size determination has not been fully understood. This has limited our understanding of the biology and biophysics of cardiomyocytes *in vivo*. Our work elucidates how changes at cellular level manifest them-selves at overall tissue level.

Following the regional differences in heart?s surface, the myocardial cells on it become curved as they enlarge with development (Fig. 6.2). 2D projections therefore not only

fail to estimate cell size, but also depend largely on the projection angle thereby possibly skewing the statistics on cell shape and size towards uniformity (Fig E.3). The analogy can be seen as the difference in the distance between two points on earth measured on its spherical surface versus the one on Mercator map projection, which by its very nature cannot be conformal and area preserving simultaneously.

Formation of multinucleated cardiomyocytes in mammals has been reported as readout of normal differentiation [31]. However the first evidences of bi- and multi-nucleated cells have been reported only in post-natal progenies [32,33]. While multi-nucleation in zebrafish has been discussed in the context of skeletal muscles and osteoclasts [34,35], the onset of multi-nucleation in cardiomyocytes has been, to the best of our knowledge, unreported. Our results have shown that cells show signs of multi-cellularity as early as 48hpf. Even in case of mutants with contractility defect, multi-nucleation remains unaffected.

Our results have shown that cardiomyocytes show differential increase in size between atrium and ventricle both in wild type and mutant embryos. It appears to be the case where chamber formation can be seen as a balance between retention of and deviation from initial cell shape and size. The region along the heart axis that retains the same cell shape and size as those at 24hpf becomes ventricle in future while the region with cell growth forms atrium. While we do not understand the exact reason for the observed cellular size differences, it can be argued that such differences might play a vital role in providing the mechanical strength to the two chambers necessary for their proper function. Ample evidences from previous studies in humans and rabbits have shown that atrial myocytes have both lower and shorter action potentials, as well as less active tension than ventricular myocytes [27-29]. It can be proposed that the time scale of needed for the functionality of these cells is crucially dependent on their length scale and hence the limit on the size of these cells. However these lines of enquiry need to be pursued in future studies. One possibility is the role of cell junction proteins in affecting cell shape and size *in vivo*. Recent studies have provided evidences of cell adhesion and shape being guided by coordination of cell junction formation through RhoU/Arhgef7b/Pak signaling in the atrioventricular cardiomyocytes [30].

Our comparative analysis of cardiac mutants with contractile defects, further argues for

decoupling of contractile mechanics from the growth of cardiomyocytes, especially in case of early events like looping. Developing heart in zebrafish embryo is a multilayered $200\mu m$ length scale structure, with several tens of microns thick myocardial wall, confined within the pericardial space of several hundreds of microns. This imposes external packing constraints that force the growing heart tube to twist and bend. Our simulations further show that growth with spatial constraints can sufficiently explain early cardiogenesis events without invoking the requirements of contraction. It must be mentioned that our simulation, based on kinematic model, is oblivious to changes within the endocardial tissue, and the mechanical properties of the two layers. Therefore future efforts to incorporate such parameters will allow better modeling approaches. However our current approach, incorporating changes at cellular level, already provides strong evidence for growth-induced buckling.

While there have been evidences for defective cardiac morphology in perturbed blood flows [36-37], previously shown mutations like *silent heart* [25] have normal looping and chamber differentiation despite the absence of contraction. Even in case of mutation of large protein, such as Titin in the *pik* mutant where sarcomere assembly is affected, development has been reported to be normal during the first three days [38]. These remain unexplained in light of blood flow as an essential epigenetic factor for cardiogenesis. Also in case of *heart of glass* (*heg*) mutations in the endocardium [39], early events in cardiogenesis such as chamber-specific myocardial differentiation occur normally and the cardiomyocyte numbers remain same as those in wild type embryos as late as 72hpf. Similarly evidences of distinguishable chambers are observed in mutants with cardiac bifidia, *miles apart* (*mil*) [40] and *bonnie and clyde* (*bon*) [41]. Our imaging also shows that for mutants with no or partial contraction, cardiomyocyte growth remains unaffected, although the overall tissue morphology clearly deviates from that in wild type embryos. We propose that this results from changes affecting the endocardium, similar to the case in *sih* embryos where it peels away from the myocardium with development resulting in embryonic death around 7dpf [25]. The effect on myocardium at later stages appears to be a secondary effect resulting primarily from the malformation of endocardium, governed through several possibilities of physical and molecular mechanisms. Numerous previous studies [10,13,15,17-21,36,37] have

elucidated the effect of blood flow on endothelial tissue. Such studies will necessitate the ability to image the endocardial motion as well as blood flow during beating. This demands very fast imaging of beating embryonic heart and will be the focus of chapters 7 and 8. Based on the work presented in the current chapter, we propose that while contraction and growth are decoupled in case of cardiomyocytes, absence of flow affects the endocardium thereby altering the overall organ geometry in case of mutants. Overall our novel findings reveal previously un-noticed properties of cardiomyocytes *in vivo* to provide new insights into heart development and disease.

6.8 References

1. Heisenberg, C.P., Bellaiche , Y. Forces in tissue morphogenesis and Pattening, Cell (2013)

2. Forgacs, Biological Physics of the developing embryos

3. Stainier, D.Y.,Lee, R.K., Fishman, M.C., Cardiovascular development in zebrafish. Development, 1993. 119: p. 31−40.

4. Stainier, D.Y., Zebrafish genetics and vertebrate heart formation. Nature Review Genetics, 2001. 2(1): p. 39−48.

5. Taber, L.A., et al The role of mechanical forces in the torsional component of cardiac looping (2010)

6. Auman, H.J., et al , Functional Modulation of Cardiac Form through regionally confined cell shape changes (2007)

7. High, F.A., Epstein, J.A. The multifaceted role of Notch in cardiac development and disease Nature Review Genetics, 2008. 9(1): p. 49−61.

8. Santhanakrishnan, A., Miller, LA, Fluid Dynamics of Heart Development. Cell Biochem Biophys, 2011.

9. Freund, J.B., Goetz, J.G., Hill, K.L., and Vermot, J.Fluid flows and forces in development: functions, features and biophysical principles. Development. 2012; 139: 1229?1245

10. Harvey, R.P. Patterning the vertebrate heart. Nature Review Genetics, 2003. 3(1): p. 544−556.

11. Hove, R.J., Koster,R.H., Forouhar, A.S., Acevedo−Bolton, G., Fraser, S.E., Gharib, M., Intracardiac fluid forces are an essential epigenetic factor for embryonic cardiogenesis. Nature, 2003. 421: p. 172−177.

12. Bartman T., Hove, J. Mechanics and function in heart morphogenesis. Dev. Dynamics (2005), 233 373−381

13. Berdougo, E., Coleman, H., Lee, D.H., Stainier, D.Y., and Yelon, D. Mutation of weak atrium/atrial myosin heavy chain disrupts atrial function and influences ventricular morphogenesis in zebrafish. Development, 2003. 130(24):6121−6129

14. Rottbauer, W., Wessels, G., Dahme, T., Just, S., Trano, N., Hassel, D., Burns, C.G., Katus, H.A., and Fishman, M.C. Cardiac Myosin Light Chain−2. A Novel Essential Component of Thick−Myofilament Assembly and Contractility of the Heart Circ. Res. 2006 99(3):323−331

15. Bressan MC1, Louie JD1, Mikawa T1. Hemodynamic forces regulate developmental patterning of atrial conduction. PLoS One. 2014;9(12)

16. Bressan M1, Yang PB1, Louie JD1, Navetta AM1, Garriock RJ1, Mikawa T2. Reciprocal myocardial−endocardial interactions pattern the delay in atrioventricular junction conduction. Development. 2014 ;141(21):4149−57

17. William J. Kowalski,1 Nikola C. Teslovich,1 Prahlad G. Menon,1 Joseph P. Tinney,1,2 Bradley B. Keller,1,2 and Kerem Pekkan1,* Left Atrial Ligation Alters Intracardiac Flow Patterns and the Biomechanical Landscape in the Chick Embryo Dev Dyn. May 2014; 243(5): 652?662

18. A. M. Markl, F. P. Chan, M. T. Alley, K. L. Wedding, M. T. Draney, C. J. Elkins, D. W. Parker, R. Wicker, C. A. Taylor, R. J. Herfkens, and N. J. Pelc, ?Time−resolved three−dimensional phase−contrast MRI,? J. Magn. Reson Imaging 17(4), 499?506 (2003).

19. S. H. Bartling, J. Kuntz, and W. Semmler, ?Gating in small−animal cardio−thoracic CT,? Methods, vol. 50, no. 1, pp. 42?49, 2010.

20. Yagel S, Cohen SM, Shapiro I, Valsky DV.3D and 4D ultrasound in fetal cardiac scanning: a new look at the fetal heart. Ultrasound Obstet Gynecol. 2007, 29(1):81−95.

21. Lin YF1, Swinburne I, Yelon D. Multiple influences of blood flow on cardiomyocyte hypertrophy in the embryonic zebrafish heart. Dev Biol. 2012 15;362(2):242−53

22. Shi Y, Yao J, Young JM, Fee JA, Perucchio R, Taber LA. Bending and twisting the embryonic heart: a computational model for c−looping based on realistic geometry. Front Physiol. 2014 12;5:297.

23. Sigolne M. Meilhac, Milan Esner, Michel Kerszberg, Julie E. Moss, and Margaret E. Buckingham Oriented clonal cell growth in the developing mouse myocardium underlies cardiac morphogenesis J Cell Biol 2004 164:97−109

24. Bayraktar M, Mnner J Cardiac looping may be driven by compressive loads resulting from unequal growth of the heart and pericardial cavity. Observations on a physical simulation model. Front Physiol. 2014 ;5:112.

25. Amy J. Sehnert, Anja Huq, Brant M. Weinstein, Charline Walker, Mark Fishman and Didier Y. R. Stainier Cardiac troponin T is essential in sarcomere assembly and cardiac contractility. Nature Genet. 2002 31, 106?110

26. Trinh LA, Hochgreb T, Graham M, Wu D, Ruf−Zamojski F, et al. (2011) A versatile gene trap to visualize and interrogate the function of the vertebrate proteome. Genes and development 25: 2306−2320.

27. Giles WR, Imaizumi Y. Comparison of potassium currents in rabbit atrial and ventricular cells. J Physiol 405: 123?145, 1988.

28. Sze Ying Ng, Chun Kit Wong and Suk Ying Tsang Differential gene expressions in atrial and ventricular myocytes: insights into the road of applying embryonic stem cell−derived cardiomyocytes for future therapies Am J Physiol Cell Physiol 299: C1234?C1249, 2010.

29. MARIANNE J. LEGATO Ultrastructure of the Atrial, Ventricular, and Purkinje Cell, with Special Reference to the Genesis of Arrhythmias. Circulation.1973; 47: 178−189

30. Michael Dickover, Jeffrey M. Hegarty, Kim Ly, Diana Lopez, Hongbo Yanga, Ruilin Zhang, Neil Tedeschi, Tzung K. Hsiai, Neil C. Chi The atypical Rho GTPase, RhoU, regulates cell−adhesion molecules during cardiac morphogenesis Developmental Biology 2014; 389 (2): 182−191

31. Kevin Bersell, Shima Arab, Bernhard Haring, Bernhard Khn Neuregulin1/ErbB4 Signaling Induces Cardiomyocyte Proliferation and Repair of Heart Injury Cell (2009) 138 (2); 257−270

32. Liu Z, Yue S, Chen X, Kubin T, Braun T Regulation of cardiomyocyte polyploidy and multinucleation by CyclinG1. Circ Res. 2010; 106(9):1498−506.

33. Mollova M, Bersell K, Walsh S, Savla J, Das LT, Park SY, Silberstein LE, Dos Remedios CG, Graham D, Colan S, Khn B.Cardiomyocyte proliferation contributes to heart growth in young humans. PNAS 2013; 110(4): 1446−51.

34. Bettina C. Kirchmaier, Kar Lai Poon, Thorsten Schwerte, Jan Huisken, Christoph Winkler, Benno Jungblut, Didier Y. Stainier, Thomas Brand The Popeye domain containing 2 (popdc2) gene in zebrafish is required for heart and skeletal muscle development Developmental Biology 363 (2012) 438−450

35. Faiza Sharif, Merijn A.G. de Bakker, Michael K Richardson, Osteoclast−like Cells in Early Zebrafish Embryos Cell J. 2014 Summer; 16(2): 211−224.

36. Forouhar, A.S., Liebling, M., Hickerson, A., Nasiraei−Moghaddam, A., Tsai, H.J., Hove, J.R., Fraser, S.E., Dickinson, M.E., Gharib, M., 2006. The embryonic vertebrate heart tube is a dynamic suction pump. Science 312, 751?753.

37. Vermot, J., Forouhar,A.S., Liebling, M., Wu, D., Plummer, D., Morteza, G., Fraser, S.E., Reversing blood flows act through klf2a to ensure normal valvulogenesis in the developing heart. PloS Biology, 2009. 7(11): p. e1000246.

38. Xu X, Meiler SE, Zhong TP, Mohideen M, Crossley DA, Burggren WW, Fishman MC. Cardiomyopathy in zebrafish due to mutation in an alternatively spliced exon of titin. Nat Genet. 2002;30(2):205−9.

39. Mably JD, Mohideen MA, Burns CG, Chen JN, Fishman MC. heart of glass regulates the concentric growth of the heart in zebrafish.Curr Biol. 2003 Dec 16;13(24):2138−47.

40. Chen JN, Haffter P, Odenthal J, Vogelsang E, Brand M, van Eeden FJ, Furutani−Seiki M, Granato M, Hammerschmidt M, Heisenberg CP, Jiang YJ, Kane DA,Kelsh RN, Mullins MC, Nsslein−Volhard C. Mutations affecting the cardiovascular system and other internal organs in zebrafish.Development. 1996 Dec;123:293−302

41. Kikuchi Y, Trinh LA, Reiter JF, Alexander J, Yelon D, Stainier DY. The zebrafish bonnie and clyde gene encodes a Mix family homeodomain protein that regulates the generation of endodermal precursors.Genes Dev. 2000;14(10):1279−89.

Chapter 7

Live 4D imaging of the embryonic zebrafish heart with 2-photon light sheet microscopy and retrospective registration

7.1 Motivation

As described in the previous chapter, the vertebrate heart is the first organ to form and function during development [1,2]. As the heart undergoes morphogenesis toward its mature 3-dimensional (3D) shape, its beating is critical for the survival of the embryo and the further development of the heart. The intrinsically dynamic and 3D nature of the embryonic heart presents a challenge to any study of its biophysics and biology, as it necessitates a means to capture fast, high 3D-resolution images as it beats. The zebrafish is an ideal model organism for studying vertebrate embryonic cardiac development owing to its optical accessibility, robust transgenic techniques, fast development, and small size - its embryonic heart begins beating as a tube about 200 μm long and 100 μm in diameter [2]. Because imaging modalities typically capture 2D images at only one instant of time, it is a challenge to capture the 3D structure of the zebrafish heart at sub-cellular resolution as it beats quasi-periodically a few beats per second and amplitude of fifty to a hundred μm [3]. The challenge of imaging a moving 3D biological structure has been addressed previously by prospective synchronization approaches, using an independent information channel (e.g. cardiac conduction or respiratory motion) to gate the 2D image acquisitions [4-7]. Alter-

natively, retrospective temporal registration of image stacks has been used to reconstruct the dynamic motion of cardiac tissues, taking advantage of the quasi-periodic nature of the heart motion [8]. This latter approach, relying on post-acquisition image processing, does not require the sophistication of real-time image processing and specialized hardware of the former prospective gating approach.

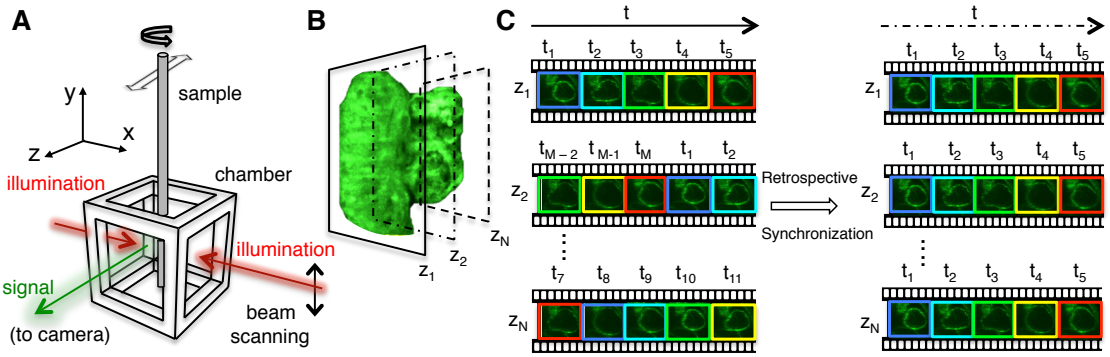


Figure 7.1: **Retrospective registration of beating heart with 2-photon light sheet microscope**
(A) Schematic of the orthogonal illumination and detection pathways of the imaging setup. (B) Successive movies of the beating heart were collected with the light sheet parked at different z positions. (C) Schematic overview of the retrospective synchronization algorithm using the movies collected at successive z positions. Time points $t_1, t_2, .., t_M$ denote distinct states of contractions of the heart as captured by our imaging protocol.

In this chapter I present the work done towards capturing the 4D (3D plus time) dynamics of the developing vertebrate heart, by combining an optimized retrospective image registration protocol with two-photon excitation selective plane illumination microscopy (2p-SPIM) [10]. SPIM, also known as light sheet microscopy, combines excitation of an optical section from the side of the sample and a camera in orthogonal direction to collect light in order to generate fast optical sectioning, while minimizing photo-damage. By using two- photon excitation in light sheet mode (Fig. 7.1A), we achieve a balance of high penetration depth and high imaging speed [10] (also see Section 5). At successive optical sections (z-planes), covering the entire depth of the organ, we acquired short movies of the beating heart, spanning several cardiac cycles. Temporal registration of the image sequences acquired from consecutive z-planes were retrospectively registered in time to produce a 4D movie of the beating heart, using a temporal registration algorithm based on maximizing

the mutual information of the adjacent z-slices [8] (Fig. 7.1B, C). This created an imaging and image reconstruction pipeline that can generate 4D image series of the beating heart, which we have applied to imaging zebrafish protein trap lines, FlipTraps [11], described before (Fig. 7.2A). The combination of FlipTraps lines that express in cardiac tissue and our imaging pipeline allowed us to capture for the first time endogenous protein expressions in a beating zebrafish heart.

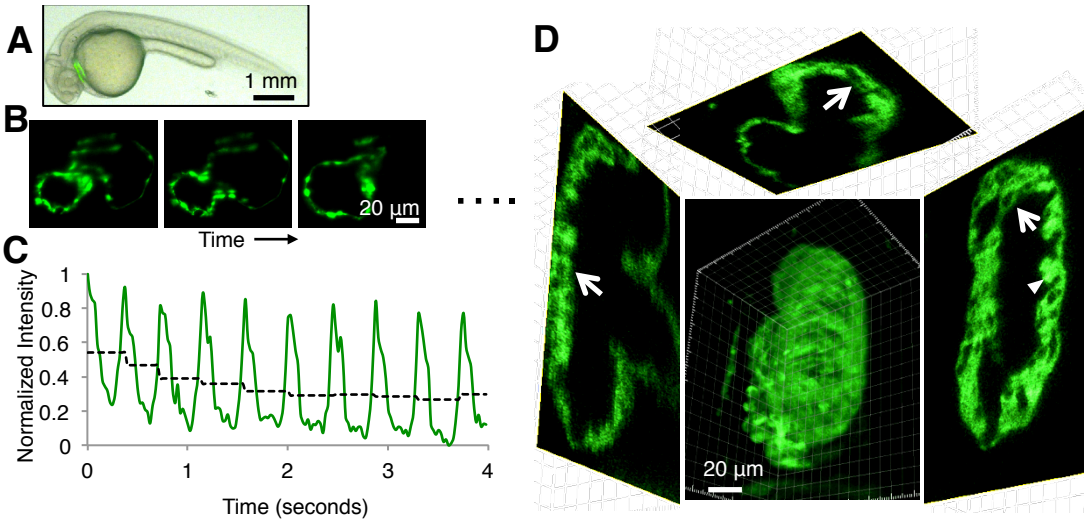


Figure 7.2: **Live 4D imaging of the embryonic heart with 2p-SPIM**
(A) Wide field fluorescent image of *GtTpm4-citrine)*$^{ct31a/+}$. (B-C) demonstration of photobleaching for a typical movie of the beating heart at a certain z-depth (A), where the normalized signal intensity integrated over the entire image is plotted as a function of time (B). Dashed black line segments show the signal intensity averaged over the time window of one local maximum to the next, revealing a decrease of $\sim 35\%$ over the first ~ 1 second, and remaining nearly constant after that. We use the near-constant intensity portion of each movie for subsequent image processing and reconstruction. (D) 3D-rendered view of a single time point in the 4D reconstructed image series of the beating myocardium of *GtTpm4-citrine)*$^{ct31a/+}$ at 84 hpf. The arrows indicate nuclei visible by negative contrast, in different cross-sections, demonstrating the quality of resolution achieved with our 2p light sheet imaging. Trabeculating myocardium (arrowhead) can be easily visualized owing to the image quality and synchronization efficiency.

7.2 2-photon light sheet microscope setup

Our custom-built 2p-SPIM setup, which could also be used in linear 1-photon (1p) excitation mode, has been described in detail previously [10]. Briefly, femtosecond-pulsed near-infrared laser light (Chameleon Ultra II, Coherent) was spherically focused by illumination objectives (Plan-Apochromat, 10X, 0.3 NA, water immersion, Nikon) from both sides of the specimen; the positions of the beams were scanned at 1KHz by galvanometers (H6215, Cambridge Technologies) and aligned to create a bidirectional Gaussian-beam excitation light sheet. This provided a thin excited optical section through the specimen, which was mounted in a liquid-filled sample chamber (Fig.7.1A). The sample could alternately be excited in 1p-SPIM mode, utilizing a continuous-wave visible-wavelength laser system (Sole-6, Omicron), aligned along the same excitation light paths. The fluorescence signal from the excited optical section was captured by a high numerical aperture (NA) water-immersion detection objective (Plan-Apochromat 20X, 1.0 NA, Carl Zeiss; or Plan-Apochromat 25X, 1.1 NA, Nikon), tube lens (focal length = 164.5 mm, Carl Zeiss; or focal length = 200 mm, Mitutoyo), optional 0.63X de-magnifying image adapter, and camera (iXON 885 or Zyla, Andor). The sample position is controlled by stages (Sutter Instrument, Newport) that allow adjustments in x (optical axis of excitation light), y, z (optical axis of detection direction), and θ (rotation angle about the y axis).

Software control of imaging setup for standard tasks (e.g. view finding, z-stacks) was done through Micro-Manager [12]. Image collection for retrospective reconstruction was performed using the Solis software (Andor) with a custom routine, written within the Solis scripting feature, that collected short 70-frames-per-second movies of the beating heart at each z-plane (duration of 4 seconds; $\sim$10 heart beats). The procedure was repeated approximately 120-140 times, collecting image sequences from the entire z-depth of the heart, with step size of 1 μm. 125 mW of laser light from each side provided the bidirectional illumination at 920 nm light. We found photobleaching of the specimen from the excitation; there was a decrease of the average signal intensity, typically by 30-40%, over the first 0.5-1 second at each z-position, after which the signal intensity remained nearly constant (Fig. 7.2 B-C). We used the near-constant intensity portion of each movie for subsequent image

processing and reconstruction.

7.3 Imaging setup characterization

We estimated the resolution performance of our imaging setup in both 1p- and 2p-SPIM modes, by imaging 170-nm fluorescent beads (Invitrogen) embedded in 1.5% agarose (Fig. 7.3A). The light sheet thickness was adjusted for the 1p and 2p laser light to achieve the same axial resolution of $\sim 1.5\ \mu m$ full width half maximum (FWHM). Lateral resolution was found to be $\sim 0.5\ \mu m$ FWHM for both imaging modes, as expected from the image pixel size of 260 nm and Nyquist?s criterion (which states that the effective resolution is at best equal to 2 times the pixel size). Note that to make full use of the detection NA of ~ 1, the pixel size would have to be smaller than $\sim 140\ \mu m$. For imaging the cardiac samples, to enhance signal to noise ratio (SNR) and reduce image size to maximize imaging speed, we used pixel size of 0.520 or 0.635 μm, which then should yield lateral resolution of ~ 1 to 1.3 μm. Summary of imaging parameters of all presented data is provided in Table F.1.

With 1p-SPIM and 2p-SPIM modes performing similarly with the ideal beads sample, we compared how they performed in imaging the live zebrafish heart sample. For the comparison to be independent from possible confounding issues related to the reconstruction algorithm, we stopped the heart beating by using a Tricaine concentration of 0.18% in the imaging chamber (while still keeping the embryo alive), to capture 3D structure via a conventional z-stack. Fig. 7.3B and C show the comparison of 1p-SPIM and 2p-SPIM in imaging the non-beating heart. We see that 1p- and 2p-SPIM perform equally well in capturing the endocardium wall structures near the outer surface of the embryo, while 2p-SPIM does better at resolving the endocardium deeper into the heart. The higher penetration depth of 2p-SPIM comes at a cost: fluorescence signal rate for 2p excitation is typically lower than for 1p excitation; thus to maintain similar SNR, 2p-SPIM cannot achieve imaging speeds that could be possible with the 1p counterpart. Note that 1p-SPIM has been used to image the dynamic structure of zebrafish hearts [17]; in our study we chose to use 2p-SPIM for its higher penetration depth, and achieved imaging speed of 70 fps, which has been shown to be sufficiently fast for capturing the beating of the zebrafish

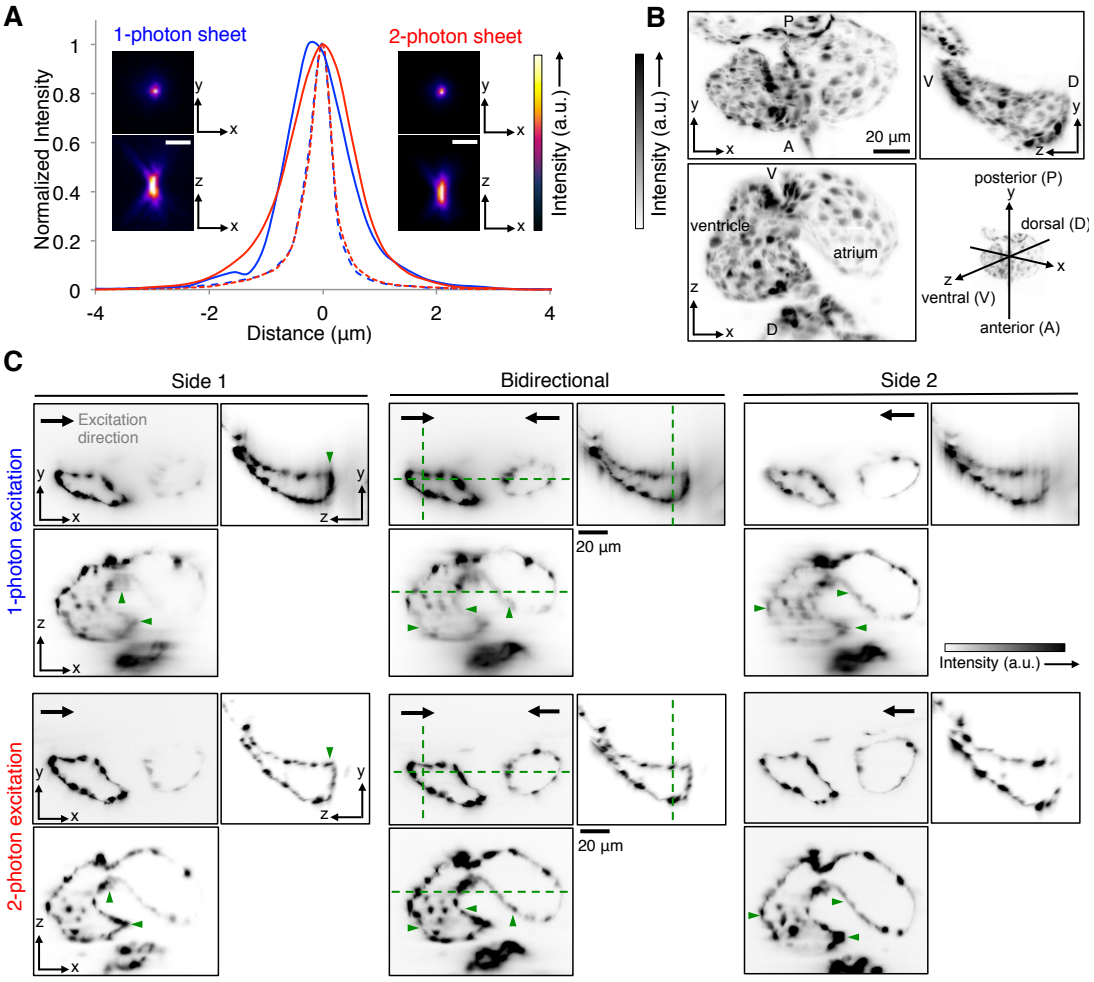


Figure 7.3: **Characterization of imaging performance of 1p-SPIM and 2p-SPIM** (A) Representative maximum intensity projection (MIP) images of 170nm-diameter fluorescent beads along the xy and xz planes,imaged by the two modalities. Plot shows the signal line profiles along the axial z-axis (solid line) and lateral x-axis (dashed line) over the bead images, showing FWHM axial resolution of $\sim 1.5\ \mu m$ and lateral resolution of $\sim 0.5\ \mu m$. (B, C) Comparison of penetration depth of 1p-SPIM and 2p-SPIM in imaging the stopped endocardium of a 60-hpf *Tg(kdrl:GFP)*^{s843} (B) MIPs of the 2p-SPIM dataset along the 3 planes as shown (C) Orthogonal views of the 3D datasets, showing particular planes (designated by dashed green lines in the central image panels), captured with either uni-directional (light sheet coming from side 1 or side 2) or bidirectional (light sheets coming from both sides simultaneously) excitation. Best effort was made to show the same planes within the heart for each dataset for fair comparison (with the constraint that the live sample did shift slightly between datasets). Linear display contrast ranges were set between the minimum and maximum pixel value in each of the images. Overall, we see that 1p-SPIM and 2p-SPIM performs equally well in resolving the endocardium wall closer to the outer surface of the embryo, but 2p-SPIM performs better at capturing cardiac structures deeper inside the animal, as highlighted by structures marked with green arrowheads. Such a difference is expected to be more pronounced as the development of the embryo proceeds

heart [8].

We have used Gaussian-beam light sheets for our imaging because of its simplicity; further improvement to the sectioning capability of the 2p excitation could be achieved by utilizing Bessel-beam light sheets [18-20]. Since the side lobes of Bessel beams do not contribute to useful signal generation in 2p excitation, more average laser power would be required for using Bessel- versus Gaussian-beam 2p-SPIM - care then must be exercised to ensure that the extra laser power does not inflict unacceptable photodamage to the live samples under study.

7.4 Image reconstruction algorithm

The movies were acquired at consecutive focal planes through the heart, where each sequence began at an arbitrary time point during the beating cycle. The retrospective registration then aims to synchronize these movies in time to assemble a full 4D reconstruction of the beating heart. From the set of movies acquired at different z-planes, we selected one corresponding to the plane in the middle of the z-stack as the starting reference sequence (*reference movie*) to which the other movies (*test movies*) were synchronized recursively. We selected a time interval in this reference movie that contained two to four periods with minimal photo bleaching, as described above. Following this, a warping function (cardiac phase function) is determined that pairs each frame in the reference movie with a similar frame in the test movie under the constraint that movie frames cannot be brought out of order. This warping function is computed by minimizing a cost function that has two terms: a similarity term that favors solutions in which the frames of the warped test movie match the frames of the reference movie; and a regularizing term that penalizes candidate warping functions that would excessively shrink or dilate the time-axis in the warped test movie. The optimal warping function is obtained by solving a minimization problem by use of a dynamic time warping algorithm [22]. Once the optimal warping function is determined, the test movie is interpolated (possibly at time-points in between frames if the local cardiac phase in the reference advances faster than in the test, or leaving out frames if it advances slower). This interpolated test movie then becomes the new reference movie for synchronizing the

neighbor at the following z-depth. By proceeding recursively, the entire stack of movies is synchronized. The above algorithm relies on high signal to noise ratio to ensure reliable registration. However noise in the image (photon shot noise, bleaching, optical scattering) could affect the similarity metric and render it unreliable. Therefore the images within a movie were transformed into wavelet domain to take advantage of the noise de-correlation properties of the wavelet-decomposition [8]. Using a mixed Matlab-C implementation within Matlab 2012b (The Mathworks, Natick, MA) we produced 4D data sets that we rendered with Imaris 7.7.2 (Bitplane AG, Zurich). Synchronization of our dataset where each movie had $\sim$ 200 frames of size 300 X 400 pixels and $\sim$ 120 z-slices, required approximately 10 hours on a 2.3GHz Intel(R) Xeon(R) processor with 64GB RAM.

7.5 4D visualization of cardiac tissue *in vivo*

Our imaging and image reconstruction pipeline not only enables observations of the anatomy but also of the molecular markers with the endogenous protein expression (Fig. 7.2D). The high axial resolution of the images obtained with 2p-SPIM allows us to resolve cellular structures in the beating 84 hpf heart in the FlipTrap line *GtTpm4-citrine)*$^{ct31a/+}$ in 3D (Fig. 7.2D, 7.4A, 7.5A). As seen in Fig. 7.1E, cytoplasmic expression of the Citrine fusion protein distinctly highlights the nuclei as darkened regions. The cytoplasmic labeling allows high-resolution visualization of the convoluted spongiform tissue, called trabeculae, in the inner surface of the ventricular myocardium. Iso-surface rendering of the image data further highlights the 3D extent of these structures (Fig. 7.4F; 7.5B). Recent studies using 3D time-lapse imaging of beating zebrafish hearts have investigated the initiation of cardiac trabeculation from 60-78 hpf, with the novel finding of myocardial protrusions that invade luminally into the trabecular layer to form new cell-cell contacts [25]. Our results demonstrate that 4D imaging of the developing heart can be performed into later stages of trabeculation, as heart structure becomes thicker and more complex, and thus would enable investigation into the further development of the trabeculae.

The myocardium generates the contraction of the heartbeat; the endocardium lining the inner surface of the heart conveys this motion to the blood, thereby pumping blood

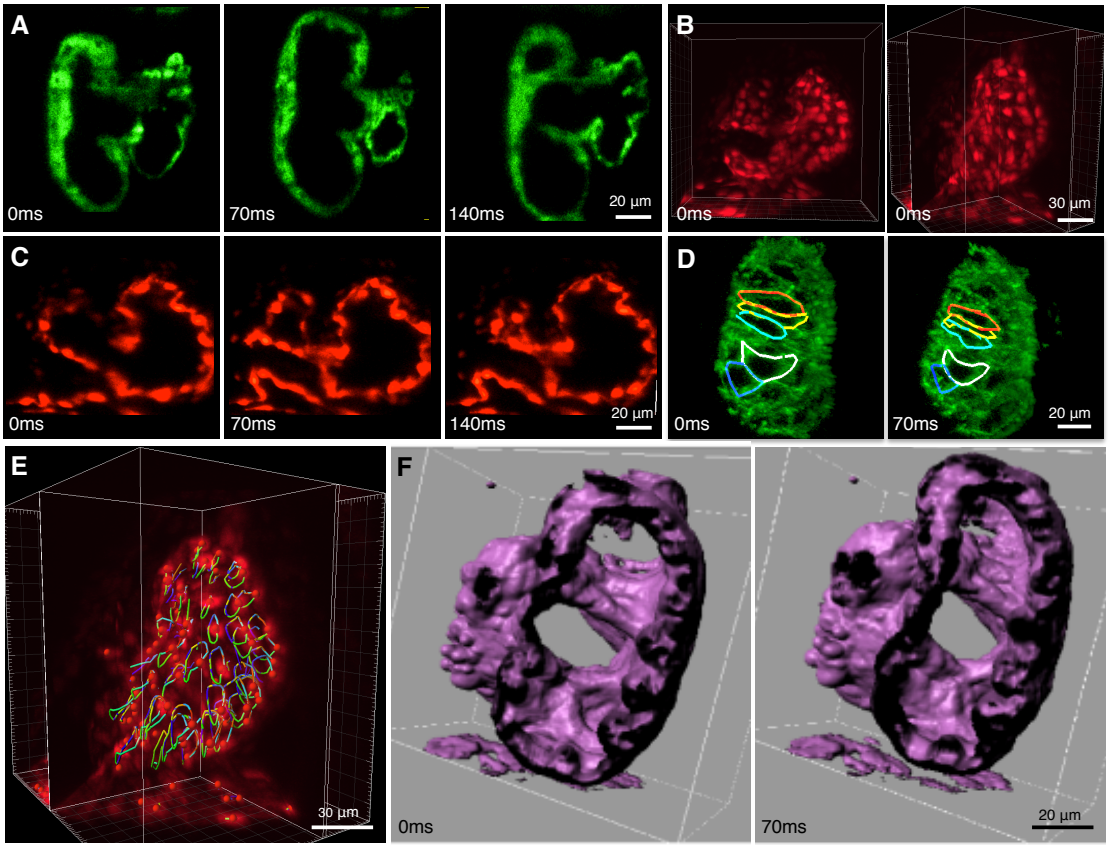


Figure 7.4: **Results of 4D capture of the beating heart**
(A-D) Fluorescent images of the heart in different transgenic lines at representative time points during its beating cycle. Images are from a single optical section in (A) *Gt(Tpm4-citrine)*$^{ct31a/+}$ and (C) *Tg(kdrl:GFP)*^{s843}. (B) 3D-rendered images of *Tg(kdrl:GFP)*^{s843} heart, at 84 hours post-fertilization (hpf), from two viewing angles demonstrating the resolution achieved with 2p-SPIM. (D) 3D rendered image of an *Gt(ctnna-citrine)*$^{ct3a/ct3a}$ embryonic heart at 84 hpf permits visualization of cellular boundaries at two time points in the heart beating cycle. The high image quality allows segmentation of cells through the entire cardiac cycle. (E) Endocardial cells in *Tg(kdrl:GFP)*^{s843}, when tracked over time (right), return to their initial position (due to periodicity of the heart motion), verifying the accuracy of the reconstruction algorithm. (F) Iso-surface rendering of *Gt(Tpm4-citrine)*$^{ct31a/+}$ at 84 hpf and digitally cropped showing the heart tubular topology, shown in 2 different time points during the beating. The surface morphology of the inside of the myocardium is clearly visualized, revealing the trabeculating ventricle.

for circulation throughout the body. Applying our imaging and reconstruction pipeline to a transgenic zebrafish line that fluorescently labels the endothelial cells with eGFP, *Tg(kdrl:GFP)*^{s843}, allowed us to visualize the motion of the endocardium of the beating heart (Fig. 7.4B, C). We characterized the endocardium motion in 3D by tracking the trajectories of individual endocardial cells during the beating cycle, the trajectories are closed paths with each cell returning to its initial positions due to the periodic nature of the heart beat (Fig. 7.4E).

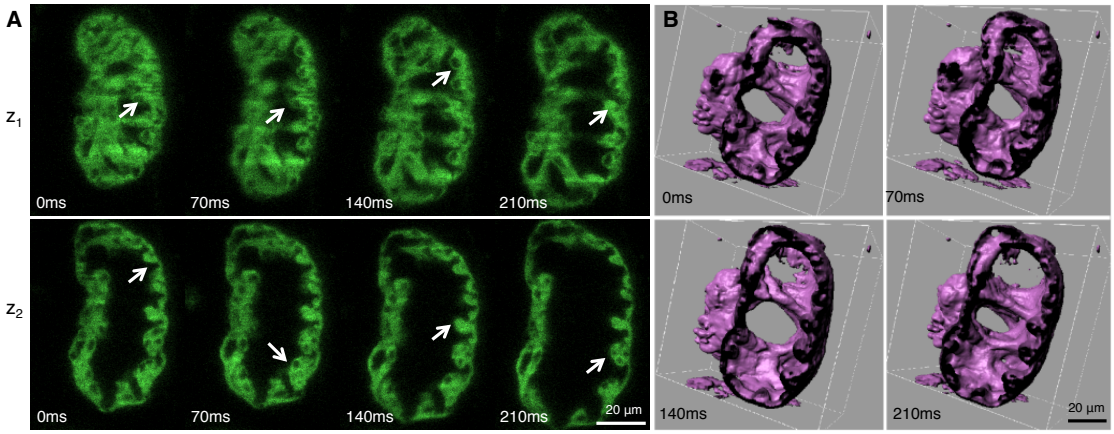


Figure 7.5: **Trabeculating myocardium during cardiac beating cycle**
(A) Optical sections of the reconstructed beating heart of the transgenic line *GtTpm4-citrine)*$^{ct31a/+}$, at 84 hpf, shown at 4 distinct time points during the beating cycle, and at 2 different z-sections. The fluorescent fusion construct labels the cytoplasm of myocardial cells. The arrows indicate the spongiform structure (trabeculae) formed by the inner lining of the ventricular myocardium. The nuclei are visible by negative contrast, in different cross-sections. (B) Iso-surface rendering of the same reconstructed dataset, shown at 4 distinct phase points and digitally cropped showing that the heart remains topologically equivalent to the cylindroid it started as. Not only can trabeculating ventricle be easily seen in the isosurface view but also the surface morphology of the inside of the myocardium is clearly visualized.

Beyond the visualization of the myocardial and endocardial tissue layers of the beating heart, the ability to see single cardiac cells undergoing contraction in full 3D provides a needed linkage between the changes of individual cell shapes and the overall motion of the *in vivo*. Progress in this direction in the context of the beating murine heart is providing new windows into cardiac anatomy and physiology [19]. Using the imaging and image reconstruction pipeline described here, we were also able to image the cellular anatomy of myocardial cells as they undergo contractions using the transgenic line *Gt(ctnna-citrine)*$^{ct3a/ct3a}$, which

labels the endogenous expression of alpha-catenin and highlights the boundaries of all cells in the developing heart (Fig. 7.4D). Our imaging of endogenous proteins expressed at their normal level expression compliments recent imaging efforts [18,20,21] where the fluorescence signal came from increasing the protein levels through the use of non-native over-expression transgenic systems. Even though our labeling is of lower intensity, the reconstructed image quality is high enough for us to carry out cell boundary segmentation of representative myocardial cells in the ventricle, over the entire beating cycle (Fig. 7.4D). Further studies into the nature of contractions in the 3D context of developing embryos will facilitate better understanding of forces and motions involved in cardiac function, thereby opening avenues for studying diseases and possible therapeutic interventions.

7.6 Discussion

The work done in this chapter focused on an imaging and image reconstruction pipeline that captures the dynamic three-dimensional beating motion of the live embryonic vertebrate heart at subcellular resolution. Live, intact zebrafish embryos were imaged using 2-photon light sheet microscopy, which offers deep and fast imaging at 70 frames per second, and the individual optical sections were assembled into a full 4D reconstruction of the beating heart using an optimized retrospective image registration algorithm. Although the simplicity of the retrospective method in terms of image acquisition is its greatest advantage, it is inefficient in collecting data and therefore exposes the specimen to more light than necessary. Bleaching and photo-toxicity can occur during continuous illumination of the sample for making movies. Although we avoided the effects of photobleaching at present, it can be envisioned that long-term imaging using this method might start to have adverse effects on the sample. Additionally it must also be noted that owing to the finite frame rate used for image acquisition, local variations in heart rates exhibit deviations from periodicity [16]. Therefore we non-uniform deformation of the time axis is needed to account for heart phase that would provide maximum similarity with the reference movie [17]. The accuracy of this method is limited by the frame rate (higher frame rates lead to better registration accuracy), any measurement noise, and the recursive nature of the procedure [23,24]. Interpolation

or mis-registration artifacts are most visible during the fastest phases of the cardiac cycle, and can be partially avoided by acquiring long movies (here ∼10 cardiac cycles long) to offer multiple cycles as candidates for matching. Balancing between the requirements of faithful reconstruction, photo-damage and long-term imaging can therefore be challenging. Alternate strategies for reconstruction are desired. One possibility is to take advantage of the orthogonal geometry of SPIM for improvements and robustness to reconstruction by utilizing different views of beating heart for faithful reconstruction. Chapter 8 discusses a novel strategy that was devised during the work on thesis that has allowed us to reconstruct images of beating heart with minimal photo-damage and image-processing artifacts. Such a pipeline has allowed us to study the nature of contraction in cardiomyocytes at several developmental stages. Together the work presented in these two chapters paves the way to open new venues for linking the cellular shape changes due to contractile behavior of cardiomyocytes to the overall tissue bending during cardiac cycle, thereby facilitating better understanding of cardiac development and function.

7.7 References

1. R.J. Hove, R.H. Koester, A.S. Forouhar, G. Acevedo−Bolton, S.E. Fraser and M. Gharib, ?Intracardiac fluid forces are an essential epigenetic factor for embryonic cardiogenesis,? Nature 421, 172−177 (2003).

2. D.Y. Stainier, ?Zebrafish genetics and vertebrate heart formation,? Nature Review Genetics 2(1), 39−48 (2001).

3. A.S. Forouhar, M. Liebling, A. Hickerson, A. Nasiraei−Moghaddam, H.J. Tsai, R.J. Hove, S.E. Fraser, M.E. Dickinson and M. Gharib, ?The embryonic vertebrate heart tube is a dynamic suction pump,? Science 312, 751−753 (2006).

4. M. Markl, F. P. Chan, M. T. Alley, K. L. Wedding, M. T. Draney, C. J. Elkins, D. W. Parker, R. Wicker, C. A. Taylor, R. J. Herfkens, and N. J. Pelc, ?Time−resolved three−dimensional phase−contrast MRI,? J. Magn. Reson Imaging 17(4), 499?506 (2003).

5. R. Jerecic, M. Bock, S. Nielles−Vallespin, C. Wacker, W. Bauer, and L. R. Schad, ?ECG−gated Na−23−MRI of the human heart using a 3D−radial projection technique with ultra−short echo times,? Magn. Reson. Mater. Phys. Biol. Med. 16(6), 297?302 (2004).

6. J.M. Taylor, C.D. Saunter, G.D. Love, J.M. Girkin, D.J. Henderson and B.C. Chaudhry, ?Real time optical gating for three−dimensional beating heart imaging,? J. Biomed. Opt 16(11), 116021 (2011).

7. Taylor JM (2014) Optically gated beating−heart imaging. Front. Physiol. 5:481

8. M. Liebling, A. S. Forouhar, M. Gharib, S. E. Fraser and M. E. Dickinson, ?Four−Dimensional Cardiac Imaging in Living Embryos via Postacquisition Synchronization of Nongated Slice Sequences,? J. Biomed. Opt. 10(5), (2005)

9. J. Ohn, H.J. Tsai and M. Liebling, ?Joint Dynamic Imaging of Morphogenesis and Function in the Developing Heart,? Organogenesis 5(4), (2009).

10. T.V. Truong, W. Supatto, D.S. Koos, J.M. Choi and S.E. Fraser, ?Deep and Fast Live Imaging with Two−photon Scanned Light−sheet Microscopy,? Nat. Methods 8(9), 757−60 (2011).

11. L.A. Trinh, et al., ?A versatile gene trap to visualize and interrogate the function of the vertebrate proteome,? Genes and development, 25(21), 2306−20 (2011).

12. A.D. Edelstein, M.A. Tsuchida, N. Amodaj, H. Pinkard, R.D. Vale and N. Stuurman, ?Advanced methods of microscope control using ?Manager software,? Journal of Biological Methods. 1(2), e10 (2014)

13. W.C. Skarnes, H. von Melchner, W. Wurst, G. Hicks, A.S. Nord, T. Cox, S.G. Young, P. Ruiz, P. Soriano, M. Tessier−Lavigne, et al., ?A public gene trap resource for mouse functional genomics,? Nature Genet. 36, 543?544 (2004).

14. W.L. Stanford, J.B. Cohn and S.P. Cordes, ?Gene−trap mutagenesis: past, present and beyond,? Nat Rev Genet. 2(10), 756−68 (2001).

15. A. Lisette, Maddison, L. Mingyu, and C. Wenbiao, ?Conditional Gene−Trap Mutagenesis in Zebrafish Methods,? Mol Biol. 1101, 393?411 (2014).

16. M. Westerfield, ?The zebrafish book,? University of Oregon Press, Eugene, OR (1994)

17. M. Mickoleit, B. Schmid, M. Weber, F. Fahrbach, S. Hombach, S. Reischauer and J. Huisken, ?High−resolution reconstruction of the beating zebrafish heart,? Nat. Methods 11(9), 919−922 (2014).

18. T. A. Planchon, L. Gao, D. E. Milkie, M. W. Davidson, J. A. Gailbraith, C. G. Gailbraith, and E. Betzig, ?Rapid three−dimensional isotropic imaging of living cells using Bessel beam plane illumination?, Nature Methods 8, 417−423 (2011).

19. O. E. Olarte, J. Licea−Rodriguez, J. A. Palero, E. J. Gualda, D. Artigas, J. Mayer, J. Swoger, J. Sharpe, I. Rocha−Mendoza, R. Rangel−Rojo, and P. Loza−Alvarez, ?Image formation by linear and nonlinear digital scanned light−sheet fluorescence microscopy with Gaussian and Bessel beam profiles?, Biomedical Optics Express 3, 1492?1505 (2012).

20. F. O. Fahrbach, V. Gurchenkov, K. Alessandri, P. Nassoy, and A. Rohrbach, ?Light−sheet microscopy in thick media using scanned Bessel beams and two−photon fluorescence excitation?, Optics Express 21, 13824?13839 (2013).

21. M. Liebling, A.S. Forouhar, R. Wolleschensky, B. Zimmermann, R. Ankerhold, S.E. Fraser, M. Gharib and M.E. Dickinson, ?Rapid Three−Dimensional Imaging and Analysis of the Beating Embryonic Heart Reveals Functional Changes During Development,? Developmental Dynamics 235(11), 2940−2948 (2006).

22. M. Liebling, J. Vermot, A.S. Forouhar, M. Gharib, M.E. Dickinson and S.E. Fraser, ?Nonuniform temporal alignment of slice sequences for four−dimensional imaging of cyclically deforming embryonic structures,? Proceedings of the 3rd IEEE International Symposium on Biomedical Imaging: Macro to Nano (ISBI'06), Arlington, VA, USA, April 6−9, 1156−1159 (2006)

23. S. Bhat, I. V. Larina, K. V. Larin, M. E. Dickinson, and M. Liebling, ?4D Reconstruction of the Beating Embryonic Heart From Two Orthogonal Sets of Parallel Optical Coherence Tomography Slice−Sequences," IEEE Transactions on Medical Imaging, 32(3), (2013).

24. I.V. Larina, K.V. Larin, M.E. Dickinson, and M. Liebling, "Sequential Turning Acquisition and Reconstruction (STAR) method for four−dimensional imaging of cyclically moving structures," Biomedical Optics Express, 3(3), 650−660, (2012).

25. D.W. Staudt, J. Liu, K.S. Thorn, N. Stuurman, M. Liebling, and D.Y.R. Stainier, ?High−resolution imaging of cardiomyocyte behavior reveals two distinct steps in ventricular trabeculation,? Development 141(3), 585−593 (2014).

26. A.D. Aguirrea, C. Vinegonia, M. Sebasa and R. Weissledera, ?Intravital imaging of cardiac function at the single−cell level,? PNAS 111(31), 11257−62 (2014).

27. P. Mahou, J. Vermot, E. Beaurepaire and W. Supatto, ?Multicolor two−photon light−sheet microscopy,? Nature Methods 11, 600?601 (2014).

Chapter 8

Macroscopic phase stamping (MaPS) for high resolution reconstruction of beating heart

8.1 Motivation

As discussed in the previous chapter, retrospective registration techniques have emerged as a popular method for 3D+time (4D) imaging of beating heart owing their simplicity in terms of image acquisition. However concerns of photo-bleaching and image processing artifacts, as discussed before, necessitate alternate strategies. As shown in Fig. 8.1A, the process of obtaining 3D+time (4D) imaging of beating heart can be decomposed as a systematic acquisition of 2D images at all positions (z-depths) within the tissue and at all spatio-temporal configuration of the contractile motion (phase,ϕ) within the cardiac cycle, at given z-resolution and frame rate. It can therefore be seen that to populate the phase-position (ϕ-z) matrix, the time required to scan the volume at a given phase must be smaller than the duration of a phase at a given frame rate. Retrospective registration approaches populate the matrix relying on image similarity measures for rearranging asynchronously acquired 2D+time input data at different depths (Fig 8.1A) within the heart volume either in a single direction [1,2] or from orthogonal (or multiple) directions [3,4].

As an alternative to the retrospective registration techniques, prospective gating approaches have also been widely used where the 2D+time acquisitions are triggered at the same phase within the cardiac cycle based on an independent signal (Fig. 8.1A). Sev-

eral strategies to this end, in the context of magnetic resonance imaging (MRI) [5,6], X-ray/Optical computed tomography (CT/OCT) [7,8] or ultrasonography [9], have been implemented over the years. These include measurement of cardiac conduction or respiration [10], local measures of heart motion based either on kymograms from raw data [11] or intrinsic gating in small-animal CT based on acquired projection data [7]. Real-time optical gating approaches have been recently demonstrated in the context of zebrafish to form a 3-D model of the heart [12,13]. However it can be seen that to populate the entire ϕ-z matrix using the prospective approach, that waits for the right trigger at appropriate phase at every z, the total acquisition time is very large. Liquid lens based high-speed volume recordings [3] have recently been used to image imaging fast movements when such synchronization based approaches fail, although the distance translated by the lens limits the effective frame rate.

Therefore the work described in this chapter aims at achieving high-resolution 4D reconstruction of beating zebrafish heart by learning from the pros and cons of both the retrospective and prospective strategies, while utilizing the orthogonal geometry of Single Plane Illumination Microscopy (SPIM) to our advantage. Our current approach, Macroscopic Phase Stamping (MaPS), introduces an additional acquisition path to obtain a low-resolution macroscopic bright-field image of the beating heart captured synchronously with the high-resolution images (fluorescent or otherwise). This provides us with an independent signal to accurately decide the spatio-temporal phase of the heart while continuously acquiring high-resolution images at different z-positions within the sample. Effectively this approach translates into repetition of such an acquisition scheme to fill in the ϕ-z matrix diagonally (Fig. 8.1A). Here we demonstrate this method in the context fluorescent images of developing zebrafish heart, starting 3 days post-fertilization (dpf) and upto 5.5 dpf, acquired with our custom built two-photon (2p-) SPIM [14]. We harvest the orthogonal geometry of the set up to acquire fluorescent images of the heart ventrally while simultaneously obtaining macroscopic bright-field images laterally via an additional imaging path (Fig. 8.1B). Coupled to this the chapter describes a novel image registration technique to obtain high-fidelity 4D reconstructions of the heart as it beats, with subcellular resolution,

reduced photo-toxicity, and minimal post-processing artifacts.

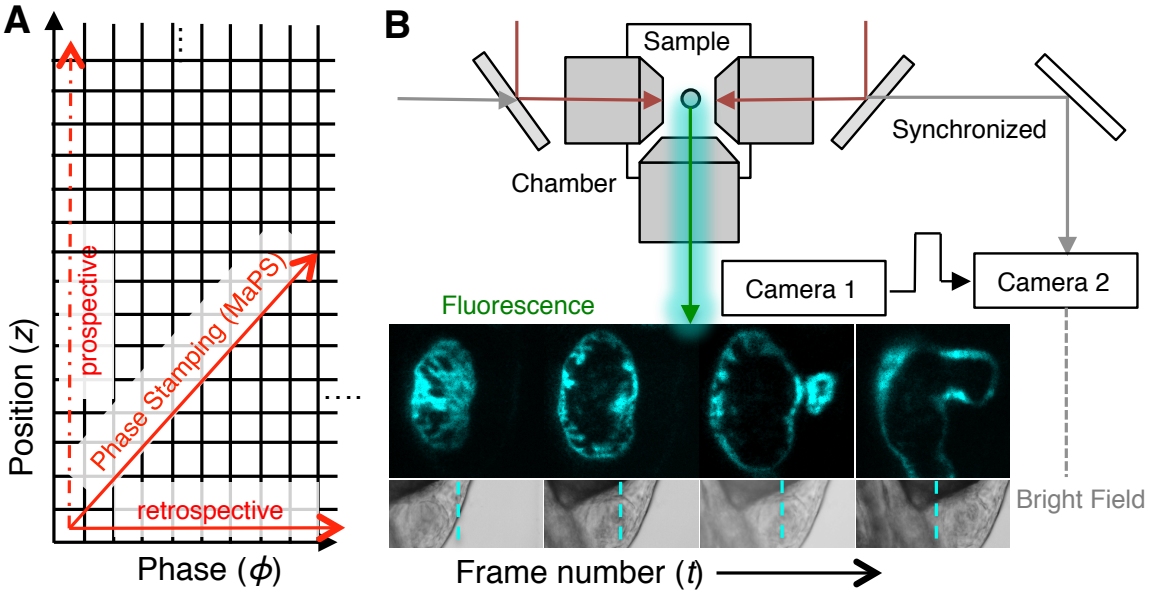


Figure 8.1: **Overview of Macroscopic Phase Stamping (MaPS) for imaging beating heart**
(A)The process of obtaining 3D+time (4D) imaging of beating heart can be decomposed as a systematic acquisition of 2D images at all positions (z-depths) within the tissue and at all spatio-temporal configuration of the contractile motion (phase,ϕ) within the cardiac cycle, at given z-resolution and frame rate. Retrospective approaches rely on capturing all ϕ's at a given z-position. Prospective approaches capture an image at every z but at a pre-determined ϕ. Our approach, macroscopic phase stamping (MaPS) acquires both z and ϕ simultaneously. (B) Two-photon light sheet setup (2p-SPIM) showing the light paths for image acquisition, where the two cameras (fluorescence and bright-field) are synchronized. 2p-laser light (deep red arrows) is used to obtain fluorescence (green arrows) from thin optical (z-) sections. Another illumination-detection light path, with a far-red light emitting diode (LED) (grey arrows) is introduced and signal is collected in the transmission mode onto a different camera.

8.2 Microscope design

The 2p-SPIM setup is similar to the one described in section 7.2. Briefly, we excite the sample by scanned 2p-laser light (Fig. 8.1B, deep red arrows) to obtain fluorescence (Fig. 8.1B, green arrows) from thin optical (z-) sections, with bi-directional illumination to increase the field of view [14]. The fluorescence signal is collected on to an electron multiplying charge-

coupled device (EM-CCD) camera. The real difference from SPIM setup in section 7.2 is the addition of another illumination-detection light path, with a far-red light emitting diode (LED) (Fig. 8.1B, grey arrows) focused via one of the low NA illumination objectives and collected via another in the transmission mode onto a scientific complementary metal oxide semiconductor (sCMOS) camera. This provides us with wide-field images of the beating heart, and hence the macroscopic phase of the tissue, from an orthogonal direction such that the plane of focus remains constant for the wide-field images despite the stage motion to obtain z sectioning in the fluorescent mode (Fig. 8.1B, image panel). The details of the imaging set up are provided in Fig. G.1. At the core of our technique lies the synchronization of these two images, one rich in spatial information (fluorescence) and the other rich in phase information (wide field) to obtain mutual information about ϕ and z for 4D reconstruction of beating heart. During the acquisition process we move the sample continuously such that the light sheet never illuminates a single plane continuously but spans through the thickness of the heart. Restricting the laser sheet to one focal plane for long durations not only causes photobleaching (described before, Fig 7.2C) but also increases in the local temperature (Fig. G.2), thereby leading to changes in heart rate, much sensitive to temperature fluctuations [15]. We operate both the cameras at 80 frames per second (fps), to ensure sufficient sampling of distinct phase points for reconstruction. Thickness of the light sheet and stage motion, synchronized to camera, provide an optical sectioning of $1\mu m$.

8.3 Image processing

During the processing pipeline, we use the fluorescent images to obtain z-position (Fig. 8.2A, z-stamping), which has one to one correspondence with the frame, given the uniformly monotonous nature of the stage movement during acquisition. The bright field images are also used to determine the exact axial position of sample in orthogonal direction (bright field image frame) (Fig. 8.2B, z-motion correction). We subtract this motion from the bright field images to obtain a region of interest, comprising of only beating heart, across all images. These images are used to determine heart period and the number of

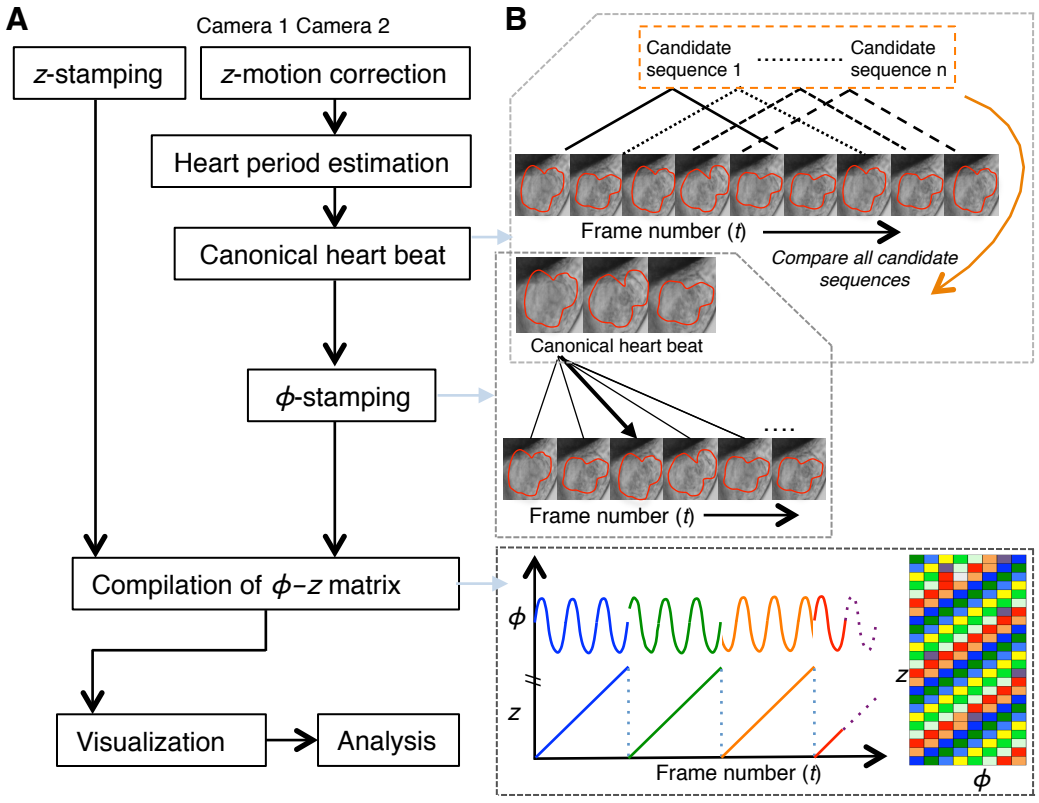


Figure 8.2: **Overview of image processing involved in MaPS**
(A) Flowchart showing various steps in the image processing involved in MaPS. (B) Pictorial depiction of some of the steps involved. Briefly, the fluorescent camera (Camera 1) is used for z-stamping and wide field camera (Camera 2) is used to determine the canonical heart beat after comparing all possible heart beats (from $\sim$ 20,000 frames analyzed). Once decided,the canonical beat sequence is used to ϕ-stamp all frames. This information is then relayed over to the fluorescent images to compile 4D reconstruction of beating heart for visualization and analysis.

distinct phases (ϕ) (Fig. 8.2A). A large number of continuous image sequences, comprising of the number of phases thus determined, can be obtained from our data set and each can be thought of as representative of a heart beat (Fig. 8.2A, B). However it is known that natural variations between heart beats, even in wild type embryos, is innate to the biology of the organ. Therefore we compare all possible candidate sequences (within the first 10% of image frames) to obtain a canonical heart beat (Fig. 8.2A, B) such that this sequence of images provides the most uniform heart behavior in terms of the number of occurrences of every phase point. Utilizing this canonical sequence we can faithfully determine the phase of every frame (ϕ-stamping) (Fig. 8.2A, B). This information is then compiled into the ϕ-z matrix and is relayed over to the high-resolution fluorescent images to allow visualization and analysis. Depending upon the duration of acquisition, not only can multiple fluorescent images be stamped for the same ϕ-z coordinates, but also different coordinates are populated non-uniformly owing to quasi-periodic nature of the heart motion. By considering these possibilities our method allows better sampling of images rendering it statistically rich. Detailed outline of the implementation of this algorithm in MATLAB is provided in Appendix G.

8.4 Multi-color imaging of beating heart

Coupled with 2p-SPIM, we use FlipTrap zebrafish lines [16] (also described before, chapters 2,6,7) that express Citrine-fused proteins in the heart to visualize endogenous expressions during beating. As mentioned before the higher penetration depth of 2p-SPIM comes at a cost: fluorescence signal rate for 2p excitation is typically lower than for 1p excitation; thus to maintain similar SNR, 2p-SPIM cannot achieve imaging speeds that could be possible with the 1p counterpart. However in our study we chose to use 2p-SPIM for its higher penetration depth, especially for later developmental stages (4dpf and beyond) and achieved imaging speed of 80 fps, which was sufficiently fast for capturing the beating of the zebrafish heart. The reconstructed image is of high quality such that it allows us to resolve cellular structures in the beating heart starting from 4dpf (Fig. 8.3, Fig. G.4). As seen in Fig. 8.3 (A-D), with the double transgenic line *Gt(tpm4-citrine)*$^{ct31a/+}$*;Tg(gata5:dsRed)*,

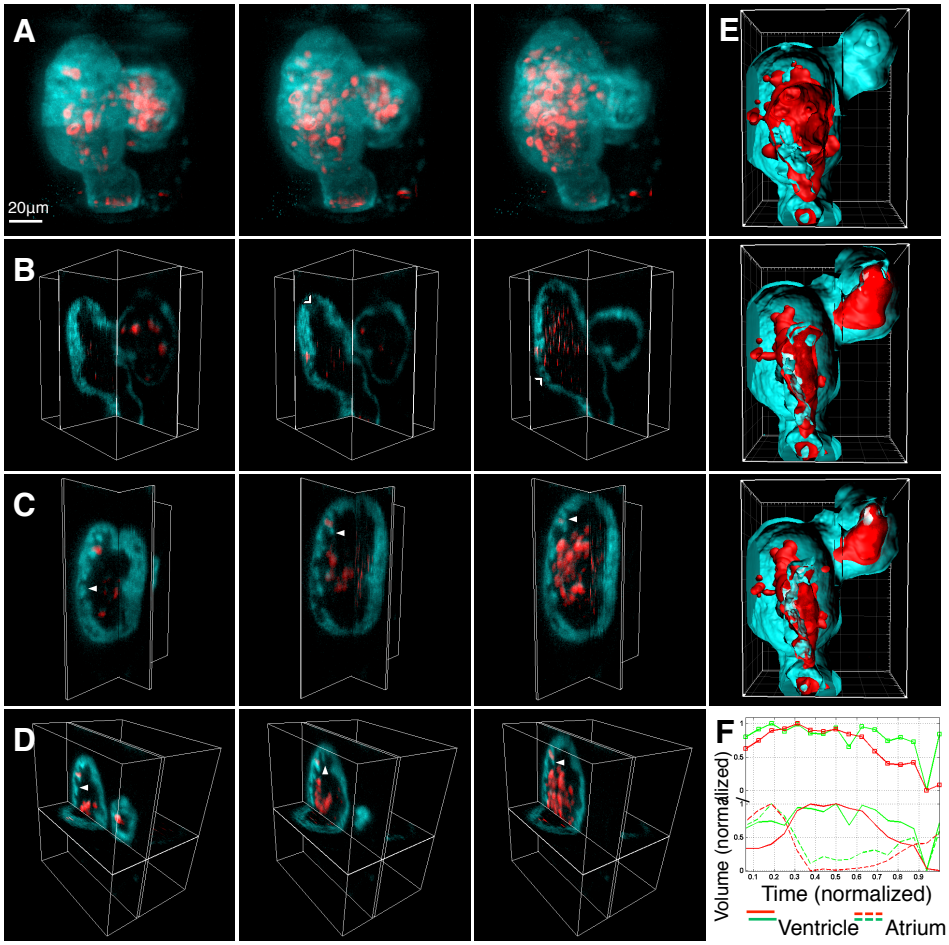


Figure 8.3: **4D reconstruction of beating myocardium and blood cells in 4.5 day old zebrafish embryo**
(A-D) 3D rendered fluorescent images of the heart in *Gt(tpm4-citrine)*$^{ct31a/+}$*;Tg(gata5:dsRed)* at 4.5dpf at representative time points during its beating cycle. Cyan shows the myocardial wall and RBCs are depicted in red. Arrowheads denote trabeculating ventricle in different cross-sections(B-D). (E) Iso-surface rendering of the same data set (F) Normalized volumes for blood and myocardial space during heart beat quantified from the image data set.

cytoplasmic expression of the Citrine fusion protein in myocardium distinctly highlights the nuclei as darkened regions (arrowheads) as well as the expression of dsRed in blood cells resolves individual blood cells during beating. MaPS allows an ensemble representation of the position of blood cells, indicative of their statistical occurrence at that point in space (See Discussion). The myocardial labeling also allows high-resolution visualization of the convoluted spongiform tissue, called trabeculae, in the inner surface of the ventricular myocardium. The high axial resolution of the images obtained allows us to resolve such structure in different planes at different phases during the heartbeat (Fig. 8.3 B-D, Fig. G.4). We also obtain Iso-surface renderings of the image data to quantify the volume of the two chambers, atrium and ventricle, during the cardiac cycle (Fig. 8.3 E-F). Our data shows an out-of-phase relationship between the ventricular and atrial volumes, as expected. Image reconstruction with MaPS therefore serves as indicative of both the anatomy as well as the functional behavior of the organ.

8.5 Quantitative measurements of cardiomyocyte contraction *in vivo*

Changes in tissue shape and form, in the context of development, have been implied to be dependent on local changes in cell shape [17]. Despite the contractile behavior of cardiomyocytes, the link between cellular shape changes and tissue bending during cardiac cycle remains unclear primarily due to the inability of cellular imaging *in vivo*. Only recently has there been a progress in visualizing single cardiomyocyte contractile cycle in beating murine heart [18] thereby facilitating possible probing into disease and therapeutic intervention. Using MAPS we demonstrated the behavior of multiple cardiomyocytes as they undergo contraction in zebrafish heart (Fig. 8.4). We report here the imaging of the cellular anatomy of myocardial cells as they undergo contractions using the transgenic line $Gt(ctnna\text{-}citrine)^{ct3a/ct3a}$, which labels the endogenous expression of alpha-catenin and highlights the boundaries of all cells in the developing heart (Fig. 8.4A). We show that our spatio-temporal resolution is good enough to follow cellular contractions for several representative cells over time (Fig. 8.4B). We can track these cells over time during the beating

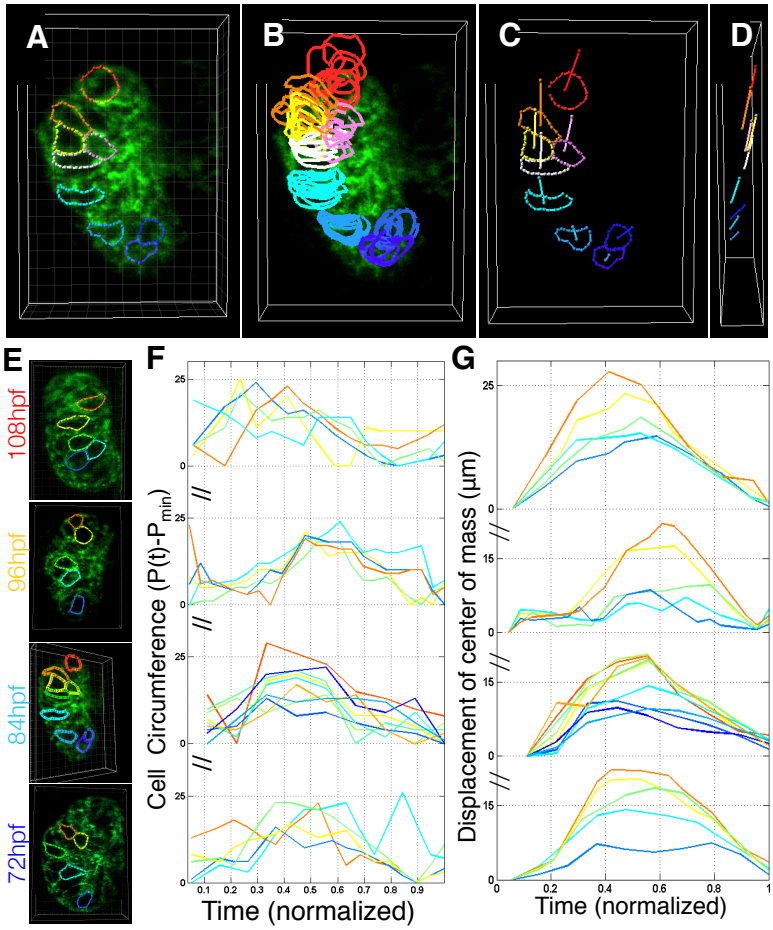


Figure 8.4: **Quantitative measurements of cardiomyocyte contraction** (A-B) 3D rendered images of an *Gt(ctnna-citrine)*$^{ct3a/ct3a}$ embryonic heart at 84 hpf permits visualization of cellular boundaries at all points in the heart beating cycle. The high image quality allows segmentation of cells through the entire cardiac cycle.(C-D) Straight line segments denote the maximum displacement of center of mass of the cells over the beating cycle (C,ventral view; D, lateral view). (E) 3D rendered fluorescent images of the heart in *Gt(ctnna-citrine)*$^{ct3a/ct3a}$ at 4 developmental stages at representative time points during its beating cycle. (F) Changes in cell circumference over beating cycle (D) Displacement of center of mass over beating cycle

cycle to determine their displacement in 3D within the intact tissue (Fig. 8.4C,D). We have been able to segment and quantify the nature of contraction of individual cardiomyocytes, spread throughout the ventricular tissue, over different developmental stages (Fig. 8.4E-G). Our analysis shows that individual cardiomyocytes (Fig. 8.4F, colored curves) display a range of heterogeneous behavior with varying amount of time lags between each other. The center of mass, however, follows an orderly displacement with strong correlation to anatomy such that the cells near to the outflow tract displace less, owing to the anchoring of this region to the pericardial cavity, as compared to the cells farther away (Fig. 8.4G).

8.6 Discussion

Image registration techniques for imaging beating zebrafish heart have been in existence for several decades now. Achieving the right balance between spatial and temporal resolutions is very important in the study of a dynamic tissue like heart. For example, to study the overall contraction, tissue level resolution might be sufficient thereby allowing high temporal resolution. On the other hand, achieving sub-cellular resolution in imaging beating heart might compromise the temporal resolution. Both retrospective and prospective techniques have been successful in providing windows of observation during heart development in various organisms [1-13]. But as mentioned in sections 7.1 and 8.1, both the techniques come with compromises. While the prospective gating approaches allow a robust phase stamp on the image, the additional hardware and total acquisition time needed are major roadblocks. Retrospective registration is simple but inefficient in terms of data-acquisition where concerns of photo-damage start to become serious, especially for long-term imaging. Our approach was to take the best of the both worlds and march diagonally in the ϕ-z matrix (Fig 8.1A) rather than along an axis (which is the case for the two approaches). This has been possible especially because of the straightforward hardware modifications we made to the conventional SPIM set up; to create an illumination-detection path in addition to the existing bidirectional 2-photon illumination SPIM that provides deep and fast optical sectioning. This has permitted us to fully harvest the orthogonal geometry of SPIM to provide an independent signal to from macroscopic movements of beating heart to accurately

decide the spatio- temporal phase in 4D (space + time). Therefore the name Macroscopic Phase Stamping (MaPS). The generality of the MaPS approach, beyond imaging embryonic zebrafish heart, can be easily seen. The phase stamping channel can be both optical (in case of transparent samples like zebrafish embryos) or OCT (in case of opaque tissues like mouse embryos). The detection position of the phase camera can range from anywhere between 0 to 4π steradians depending upon the accessibility to the sample for faithful acquisition of the position independent phase-rich information. The method is applicable for any moving structure where distinct phase points can be identified. This can range from the movement generated due to breathing in lungs (in higher vertebrates) to periodic peristalsis in the gut, provided two imaging directions are possible: one that captures the macroscopic movements and the other provides the desired optical sectioning.

One of the crucial requirements for MaPS to work successfully is the synchronization of the two cameras (phase and fluorescent) as we achieve both z- and ϕ-stamping based on the wide field images. We have verified the accuracy of these stamps based on both static structures (such as pericardial wall) as well as moving structures. A detailed description is provided in Appendix G. Another unanticipated advantage of the bright field images towards accurate phase stamping is the presence of blood cells. Bright field images consist of cardiac structure moving with different velocities that can be broadly classified into static, transient and periodic structures. Pericardium, blood cells and cardiac wall can respectively be categorized as above. While the static pericardium allows inferring z stage motion and periodic myocardial walls are needed for phase inference, transient blood cells are needed for providing directionality to wall motion. During diastole and systole heart walls pass through similar spatial positions that cannot be faithfully distinguished from each other when only the heart walls are analyzed. Therefore it's the blood cells, that actually respond differently to diastolic and systolic pressures, provide an overall directionality based on their bulk motion.

Our imaging protocol usually involves acquiring $\sim$ 20,000 frames to populate the phase-z matrix sufficiently. Typically a z-stack has 150 frames and heartbeat consists of 20-25 phase points, given the frame rate we use currently. Thus we are clearly overpopulating

the ϕ-z matrix at least 5 times. Heart motion being quasi-periodic leads to non-uniform filling of different phase-z coordinates in the matrix thereby leading to different numbers of possible images for different coordinates. This provides us with a dual advantage. First in case of wild type embryos, given the statistical nature of our reconstruction assay, all such images at a given ϕ-z coordinate must be equivalent. However one can still chose the best among the possibilities based on the arguments of physical continuity of the tissue between consecutive phases at the same z and consecutive z?s at the same phase, very similar to the retrospective approach described in chapter 7. This is a memory intensive step that we have bypassed currently by imaging at frame rates fast enough such that the issues due to finite frame rate can be successfully avoided. In other words our stamping approach is robust enough to ensure that all the images in the set of possible images at any ϕ-z coordinate are equivalent to each other. The second advantage of using MaPS is to provide a hold on the health of the heart. The stream of wide field images contains phase-rich information that can be utilized very simply. In case of wild type embryos, phase jumps between consecutive image frames is expected to be minimal. However within mutant/diseased hearts where the rhythmicity is affected, such jumps in phases are expected to be more frequent. A discussion on this based on simulated data is provided in the Appendix G.

Overall the work presented in this chapter demonstrates that a combination of novel imaging tools, both in terms of technique and analysis, can facilitate pushing the speed of fast dynamic imaging to understand behavior of cells, such as beating cardiomyocytes, *in vivo*. With long term imaging and analysis, it will now be possible to understand how the shape of heart is remodeled, after the primary structure layout, over several days and how this slow process of morphogenesis is closely linked to cellular processes on shorter time scales.

8.7 References

1. M. Gargesha, M. Jenkins, D. Wilson, and A. Rollins, ?High temporal resolution OCT using image−based retrospective gating,? Opt. Exp., vol. 17, pp. 10786?10799, 2009.

2. Peterson LM, Jenkins MW, Gu S, Barwick L, Watanabe M, Rollins AM. 4D shear stress maps of the developing heart using Doppler optical coherence tomography.Biomed Opt Express. 2012 ;3(11):3022−32.

3. Bhat S, Larina I, Larin KV, Dickinson ME, Liebling M. 4D reconstruction of the beating embryonic heart from two orthogonal sets of parallel optical coherence tomography slice−sequences.IEEE Trans Med Imaging. 2013 Mar;32(3):578−88

4. I.V. Larina, K.V. Larin, M.E. Dickinson, and M. Liebling, "Sequential Turning Acquisition and Reconstruction (STAR) method for four−dimensional imaging of cyclically moving structures," Biomedical Optics Express, Vol. 3, Issue 3, pp. 650−660, 2012.

5. A. M. Markl, F. P. Chan, M. T. Alley, K. L. Wedding, M. T. Draney, C. J. Elkins, D. W. Parker, R. Wicker, C. A. Taylor, R. J. Herfkens, and N. J. Pelc, ?Time−resolved three−dimensional phase−contrast MRI,? J. Magn. Reson Imaging 17(4), 499?506 (2003).

6. R. Jerecic, M. Bock, S. Nielles−Vallespin, C. Wacker, W. Bauer, and L. R. Schad, ?ECG−gated Na−23−MRI of the human heart using a 3D−radial projection technique with ultra−short echo times,? Magn. Reson. Mater. Phys., Biol., Med. 16(6), 297?302 (2004).

7. S. H. Bartling, J. Kuntz, and W. Semmler, ?Gating in small−animal cardio−thoracic CT,? Methods, vol. 50, no. 1, pp. 42?49, 2010.

8. M. W. Jenkins, O. Q. Chughtai, A. N. Basavanhally, M. Watanabe, and A. M. Rollins, ?In vivo imaging of the embryonic heart using gated optical coherence tomography,? J. Biomed. Opt., vol. 12, no. 03, pp. 030505?030505, May 2007.

9. Yagel S, Cohen SM, Shapiro I, Valsky DV.3D and 4D ultrasound in fetal cardiac scanning: a new look at the fetal heart. Ultrasound Obstet Gynecol. 2007, 29(1):81−95.

10. Hung J, Lang R, Flachskampf F, Shernan SK, McCulloch ML, Adams DB, Thomas J, Vannan M, Ryan T 3D echocardiography: a review of the current status and future directions.J Am Soc Echocardiogr. 2007 20(3):213−33

11. G. M. Kachelriess, D. A. Sennst, W. Maxlmoser, and W. A. Kalender, ?Kymogram detection and kymogram−correlated image reconstruction from subsecond spiral computed tomography scans of the heart,? Med. Phys., vol. 29, no. 7, pp. 1489?1503, Jul. 2002.

12. Taylor,J.M., Saunter,C.D., Love,G.D., Girkin, J.M., Henderson, D.J., Chaudhry, B.C. Real time optical gating for three−dimensional beating heart imaging. J. Biomed. Opt., 2011. 16(11), 116021

13. Taylor JM (2014) Optically gated beating−heart imaging. Front. Physiol. 5:481

14. Thai V Truong, Willy Supatto, David S Koos, John M Choi and Scott E Fraser deep and fast live imaging with two−photon scanned light−sheet microscopy Nat Methods. 2011 ;8(9):757−60

15. Denvir,M.A., Tucker,C.S., Mullins,J.J. Systolic and diastolic ventricular function in zebrafish embryos: Influence of norepenephrine, MS−222 and temperature BMC Biotechnology 2008, 8(21): p

16. L.A. Trinh, et al., ?A versatile gene trap to visualize and interrogate the function of the vertebrate proteome,? Genes and development, 25(21), 2306−20 (2011)

17. Lecuit,T., Lenne,P. Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis Nature Reviews, 2007 8: p 633−644

18. Aguirrea,A.D., Vinegonia,C., Sebasa, M., Weissledera,R. Intravital imaging of cardiac function at the single−cell level PNAS , 2014

Chapter 9

Outlook

The work done in this thesis addresses a variety of questions in developmental biology and all throughout the work imaging has been central to the approach. Through the eight chapters, I have tried to highlight that the diverse nature of processes in a developing embryo, ranging from molecular level to the organ level, present a rich diversity of questions to be asked; therefore diverse approaches are needed as well to answer such questions. With novel microscopy tools, both for imaging as well as analysis, it is now possible to connect processes across spatio-temporal scales. A glimpse of such an attempt can be seen in Chapter 4, where the local movements of chondrocytes is translated to the tissue scale to understand cartilage growth. In the work done on cardiac *looping* (Chapter 6) the effort was to understand how the shorter time scale of beating in embryonic heart affects the developmental time scale over which the organ undergoes morphogenesis. The limits of fast imaging were finally pushed in Chapters 7 and 8 , where light sheet microscope was used to capture the dynamics of beating heart at the cellular resolution. This has provided us with the ability to observe individual cells of myocardium as they contract or the cells of endocardium as they constantly engage in a tug of war to bear the shear due to blood flow. Ultimately linking such macroscopic events to the cellular machinery demands the ability to probe into the molecular state of the cells. One aspect of this question, directed towards quantitative mapping of mRNA expression within intact embryos, has been the focus of Chapter 2. And there again, imaging was used as the primary tool to infer the information about relative molecular abundance. As pointed out by physicist Philip W. Anderson, in his 1972 article in *Science* [1], the twin difficulties of scale and complexity render the behavior

of large aggregates to be more than a simple extrapolation of properties of constituent elements. In an ensemble like living systems, entirely new properties might at each level of complexity and microscopy as a tool provides the hold on this behavior at various scales to enable its fundamental understanding.

The field of microscopy has been built and rebuilt on a rich legacy of scientists across disciplines; hence lessons from the past matter. In a number of cases techniques established in the past for a purpose quite different have been adapted decades later in the context of biological imaging. A classic example is the ultramicroscope, designed by Henry Siedentopf and Richard Zsigmondy in 1902 to study colloidal particles [2], which was almost forgotten for six decades before it was rediscovered in rheology and fluid mechanics as particle image velocimetry (PIV) and in mineralogy as deep-field microscope [3]. The utility of the orthogonality, of illumination and detection axes, in this technique took another 20 years to be realized in the context of biological imaging [4, 5]. Through a series of demonstrations with a variety of samples like pig cochlea, drosophila embryos, oceanic bacteria, the method transitioned through various names like orthogonal-plane fluorescence optical sectioning [6] and thin laser light sheet microscope [7] to ultimately attain its current name selective plane illumination microscopy (SPIM) in 2004 [8]. Since then the technique has been evolving as the cutting edge technology for dynamic imaging *in vivo*. SPIM is not the only example where concepts from other disciplines and sub-fields of optics, were borrowed for the first time in the context of biological imaging. In fact even in the first discussion ever of the possibility of dark-field microscopy with epi-illumination, found in the 1836 writings of the English clergyman J. B. Reade at Cambridge [9], he says, *'to illustrate the two methods (Bright-field and dark-field) by reference to the telescope it may be observed that the discomfort of viewing spots on the sun not unaptly corresponds with the view of microscopic objects on an illuminated field; while the removal of all inconvenient and ineffective light from the field of the microscope corresponds with the clear and quiet view of stars on the dark blue vault of the firmament.'* It is only prudent to acknowledge that, in our efforts to develop better quantitative tools for biological imaging, for an inter-disciplinary field such as microscopy, understanding both the successes and failures of past efforts is very fruit-

ful. As elegantly put forth by Frits Zernike: *'how quick we are to learn-that is, to imitate what others have done or thought before-and how slow to understand -that is, to see the deeper connections. Slowest of all, however, are we in inventing new connections or even in applying old ideas in a new field'* [10].

With the advent of more sophisticated techniques probing into molecular composition and dynamics is quite possible now. Using fluorescence suppression bases techniques like Frster resonance energy transfer (FRET) and fluorescence recovery after photobleaching (FRAP), molecular diffusion and interactions can now be studied with sub-micron precision. With image correlation techniques, fluctuations in fluorescence, which would have otherwise been discarded as noise, have been optimally utilized to understand the even shorter time scales of molecular dynamics *in vivo*. Even with regard to fixed imaging, better fixing techniques like CLARITY can now transform an intact tissue into an optically transparent hydrogel-hybridized form [11]. With better and better labeling techniques and optically coupled dyes, microscopy is answering far more variety of questions than what could have been imagined just 30 years ago, not to mention the days of Hooke or Abbe. What was once remarked by Charles P Shillaber, in his classic *Photomicrography* in 1944, that *'the microscope with its accessories is by far the least understood, the most inefficiently operated, and the most abused of all laboratory instruments'* [12], does no longer hold as microscopy is now the inevitable 'gadget'to study life. The future is full of possibilities and perhaps we now stand at the cusp of time to re-realize, in a new light of course, what Hooke realized when he first saw through the microscope [13];

'For the limits to which our thoughts are confind, are small in respect of the vast extent of Nature itself; some parts of it are too large to be comprehended and some too little to be perceived. And from thence it must follow, that not having a full sensation of the Object, we must be very lame and imperfect in our conceptions about it, and in all the propositions which we build upon it...by the help of Microscopes, there is nothing so small, as to escape our inquiry; hence there is a new visable World discovered to the understanding'

9.0.1 References

1. 1. P. W. Anderson More Is Different Science 1972: 393-396.

2. Richard A. Zsigmondy Properties of colloids Nobel Lecture, December 11, 1926

3. McLachlan Jr., D.The use of deep-field microscopy in crystal morphologyJ. Cryst. Growth 1971 Vol. 8(4), pp. 363-371

4. Voie, A.H., Spelman, F.A., Sutton, D. and Burns, D.H. Quantitative measurements of the three-dimensional anatomy of theguinea pig cochlea using optical ?dots Engineering in Medicine and Biology Society 1989

5. Shaw, P.J., Agard, D.A., Hiraoka, Y. and Sedat, J.W. Tilted view reconstruction in optical microscopy. Three-dimensional reconstruction of Drosophila melanogaster embryo nuclei Biophys. J. 1989 55(1), pp. 101-110

6. Voie AH, Burns DH, Spelman FA. Orthogonal-plane fluorescence optical sectioning: three-dimensional imaging of macroscopic biological specimens.J Microsc. 1993 ;170(Pt 3):229-36.

7. Fuchs E, Jaffe J, Long R, Azam F. Thin laser light sheet microscope for microbial oceanography. Opt Express. 2002 Jan 28;10(2):145-54.

8. Huisken J, Swoger J, Del Bene F, Wittbrodt J, Stelzer EH. Optical sectioning deep inside live embryos by selective plane illumination microscopy.Science. 2004;305(5686):1007-9.

9. C. R. Goring, Andrew Pritchard Micrographia: Containing Practical Essays on Reflecting, Solar, Oxy-hydrogen Gas Microscopes; Micrometers; Eye-pieces, &c. &c Whitaker and Company, 1837

10. F. Zernike How I Discovered Phase Contrast Science 1955: 345-349.

11. K Chung, J Wallace, S-Y Kim, S Kalyanasundaram, AS Andalman, TJ Davidson, JJ Mirzabekov, KA Zalocusky, J Mattis, AK Denisin, S Pak, H Bernstein, C Ra-

makrishnan, L Grosenick, V Gradinaru, and K Deisseroth. Structural and molecular interrogation of intact biological systems. Nature (2013) 497: 332-337.

12. Hooke, Robert C. Micrographia: or Some Physiological Descriptions of Miniature Bodies Made by Magnifying Glasses. London, England: Jo. Martyn, and Ja. Allestry, 1665.

Appendix A

Supplementary Information for Chapter 1

Events/Processes	Length scale	Time sacle	Reference
Subcellular			
Diffusion in cytoplasm	1-10 μm	10 ms - 10 s	[1-4]
Vesicular transport	5-10 μm	1-3 μm/s	[1-4]
Membrane potential	1-10 μm	1-5 ms	[1-4]
Cytoskeletal assembly	3-10 μm	1-5 min	[1-4]
Cellular/multi-cellular			
Cell division			
in bacteria	2-5 μm	20-40 min	[1-4]
in yeast	3-6 μm	70-140 min	[1-4]
in gastrulating fish	500 - 700 μm	15-30 min	[5,6]
in gastrulating drosophila	300 - 500 μm	15-30 min	[7]
in growing mouse cartilage	5-10 μm	1-2 hrs	[8]
Cell migration			
in gastrulating fish	500 - 700 μm	3-4 hrs	[5,6]
in gastrulating drosophila	300 - 500 μm	3-4 hrs	[7]
in growing chicken cartilage	5-10 μm	1-2 hrs	[9]
Flagelated cell motion	0.5-1mm	1-2 s	[1,3]
Embryonic heart			
in zebrafish embryo 26hpf	50 μm	0.4-0.5 s	[10-12]
in zebrafish embryo 120hpf	100 μm	0.5s	[10-12]
in chicken embryo HH15	200 μm	0.5 s	[12,13]
in mouse embryo E8.5	75-100 μm	0.3-0.4 s	[12, 14]
in mouse embryo E10.5	150-200 μm	0.4 s	[12, 14]
Somitogenesis			
in zebrafish embryo	40-50 μm	30 min	[15,16]
in chicken embryo	80-100 μm	90 min	[16]
Calcium waves	50-100 μm	50-100 ms	[17]

Table A.1: Spatio-temporal scales in growth and developmental processes

A.0.2 References

1. Rob Phillips, Jane Kondev, Julie Theriot, Hernan Garcia Physical Biology of the Cell: 2nd Edition (2012) Taylor & Francis

2. Bruce Alberts , Alexander Johnson, Julian Lewis , Martin Raff , Keith Roberts , Peter Walter Molecular Biology of the Cell (2007) Garland Science; 5 edition

3. Harvey Lodish Arnold Berk, Chris A. Kaiser , Monty Krieger , Anthony Bretscher , Hidde Ploegh, Angelika Amon ,Matthew P. Scott Molecular Cell Biology (2012) W. H. Freeman; Seventh Edition edition

4. www.BioNumbers.org

5. Tomer, R., Khairy, K., Amat, F. & Keller, P. J. Quantitative high−speed imaging of entire developing embryos with simultaneous multiview light−sheet microscopy. Nat. Meth. 9, 755?763 (2012).

6. Schmid, B.; Shah, G.; Scherf, N.; Weber, M.; Thierbach, K.; Campos, C.; Roeder, I.; Aanstad, P.; Huisken, J. High−speed panoramic light−sheet microscopy reveals global endodermal cell dynamics. Nat. Commun., 4, (2013)

7. Krzic, U., Gunther, S., Saunders, T. E., Streichan, S. J. & Hufnagel, L. Multiview light−sheet microscope for rapid in toto imaging. Nat. Meth. 9, 730?733 (2012).

8. Romereim, S. M., Conoan, N. H., Chen, B. & Dudley, A. T. A dynamic cell adhesion surface regulates tissue architecture in growth plate cartilage. Dev. Camb. Engl. 141, 2085−2095 (2014).

9. Li,Y., Trivedi, V., Truong,T.V., Koos,D.S., Lansford,R., Chuong,C.M., Warburton,D., Moats,R.A., Fraser S.E. Dynamic imaging of the growth plate cartilage reveals multiple con− tributors to skeletal morphogenesis Nature Communications (2015) 6:6798 doi: 10.1038/ncomms7798

10. Hove, R.J., Koster,R.H., Forouhar, A.S., Acevedo−Bolton, G., Fraser, S.E., Gharib, M., Intracardiac fluid forces are an essential epigenetic factor for embryonic cardiogenesis. Nature, 2003. 421: p. 172−177.

11. Forouhar, A.S., Liebling, M., Hickerson, A., Nasiraei−Moghaddam, A., Tsai, H.J., Hove, J.R., Fraser, S.E., Dickinson, M.E., Gharib, M., The embryonic vertebrate heart tube is a dynamic suction pump. Science, 2006. 312: p. 751−753.

12. Santhanakrishnan, A., Miller, LA, Fluid Dynamics of Heart Development. Cell Biochem Biophys, 2011.

13. Vennemann, P., Kiger, K. T., Lindken, R., Groenendijk, B. C. W., Stekelenburg−de Vos, S., ten Hagen, T. L. M., et al. (2006). In vivo micro particle image velocimetry measurements of blood−plasma in the embryonic avian heart. Journal of Bio− mechanics, 39(7), 1191?1200.

14. Jones, E. A. V., Baron, M. H., Fraser, S. E., & Dickinson, M. E. (2004). Measuring hemodynamics during development. Ameri− can Journal of Physiology. Heart and Circulatory Physiology, 287, H1561?H1569.

15. Miguel Maroto, Robert A. Bone and J. Kim Dale Somitogenesis 2012 Development 139, 2453−2456.

16. Hester SD, Belmonte JM, Gens JS, Clendenon SG, Glazier JA (2011) A Multi−cell, Multi−scale Model of Vertebrate Segmentation and Somite Formation. PLoS Comput Biol 7(10): e1002155.

17. J. Vermot, S. E. Fraser, M. Liebling "Fast fluorescence microscopy for imaging the dynamics of embryonic development," HFSP Journal, vol 2, pp. 143−155, 2008

Appendix B

Supplementary Information for Chapter 2

B.1 Zebrafish lines

Adult fish were raised and maintained as described before* in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals by the respective bodies at California Institute of Technology and University of Southern California. The protocol was approved by Institutional Animal Care and Use Committee (IACUC) at both universities. Wild-type embryos were obtained from AB strains. FlipTrap *Gt(desm-citrine)*$^{ct122a/+}$ transgenic embryos were obtained from a previously described screen†. Transgenic line *Tg(kdrl:eGFP)* was provided by the Stainier lab. Upon crossing appropriate adult lines, the embryos obtained were raised in Egg Water (60 μg/ml of stock salts in distilled water) at 28.5°C with addition of 0.003% (w/v) 1-phenyl-2-thiourea (PTU) around 18hpf to reduce pigmentation, until the desired stage before fixation.

B.2 Sample preparation for confocal imaging

Imaging was performed on a Zeiss LSM710 inverted confocal microscope. Embryos were placed into agarose mould and imaged with the appropriate lasers and a water- immersion 25X 0.8NA Apochromat objective using the Zeiss LSM software. To permit accurate comparisons, all embryos in a given experiment were imaged using the same laser power; gain detector settings and scanning speed. Custom written codes in MATLAB 2013b were used for all post-acquisition analysis. A detailed description of all the microscope setting for all presented data is provided in Table B.1

*Westerfield, M., The Zebrafish Book. Edition 2.1 ed. 1994, Eugene, OR: University Oregon Press

†Trinh LA, Hochgreb T, Graham M, Wu D, Ruf−Zamojski F, et al. (2011) A versatile gene trap to visualize and interrogate the function of the vertebrate proteome. Genes and development 25: 2306−2320

B.3 Protocols

B.3.1 Preparation of fixed whole-mount zebrafish embryos

1. **Fixation**:
   Add 500μL of freshly prepared 4% paraformaldehyde (PFA)[‡] to 80 dechorionated zebrafish embryos in 2mL eppendorf tubes and incubate for 2 hours at room temperature (RT) or overnight @ 4°C

2. **Dehydration**:
   Wash 1-3: 1× PBS, 5 min each @ RT
   Wash 4-7: 100% MeOH, 10 min each @ RT
   Wash 8: 100% MeOH, 50 min @ RT[§]

3. **Rehydration**:
   Wash 1: 75% MeOH + 25% 1× PBST, 5 min @ RT
   Wash 2: 50% MeOH + 50% 1× PBST, 5 min @ RT
   Wash 3: 25% MeOH + 75% 1× PBST, 5 min @ RT
   Wash 4-8: 100% 1× PBST, 5 min @ RT [¶‖]

B.3.2 Multiplexed *in situ* hybridization using DNA HCR

1. **Pre-hybridization for probes**[**]:
   (a) Incubate in 500μL of HB-1, 2hrs @65°C[††]
   (b) Incubate fresh 500μL of HB-1, 1.5hrs @45°C

2. **Probe hybridization**:
   (a) Mix 1 pmol of each probe in 500μL of HB-1 @45°C for 30 minutes.[‡‡]
   (b) Add this probe solution to embryos for overnight (16hrs) incubation @45°C

3. **Probe Washes**
   (a) Wash in 500μL HB-2, 15 min @45°C.
   (b) Wash in 500μL (75% of HB-2 + 25% of 5× SSC), 15 min @45°C.
   (c) Wash in 500μL (50% of HB-2 + 50% of 5× SSC), 15 min @45°C.
   (d) Wash in 500μL (25% of HB-2 + 75% of 5× SSC), 15 min @45°C.
   (e) Wash 2 times, in 500μL 5× SSC, 15 min @45°C.
   (f) Wash 2 times, in 500μL 5× SSC, 30 min @45°C.

4. **Pre-hybridization for Hairpins**:

[‡]Use fresh PFA and cool to 4 °C before use to avoid increased autofluorescence.
[§]Store samples in MeOH @ -20°C or proceed directly to next step
[¶]Store samples in 1X PBST @ 4°C or proceed directly to next step
[‖]Prepare embryos every two weeks to avoid increased autofluorescence.
[**]For each sample 7-8 embryos can be added to 2 mL eppendorf tube
[††]Caution: The solution is very viscous and samples become transparent. Also, reduce the temperature depending upon the fragility of tissue
[‡‡]This step should be coordinated with step 2(a) so that it ends near simultaneously with the same

(a) Incubate samples in 500μL of HB-3, 30 min @RT
(b) Incubate fresh 500μL of HB-3, 30 min @RT
(c) Aliquot 30pmol (for 500μL of HB-3) of each hairpin in a separate PCR tube
(d) Heat the hairpins to 95°C for 90 seconds in PCR machine (or otherwise)
(e) Store the heated hairpins in dark for 30 minutes @RT. *(Keep hairpins unmixed)*
(f) Equilibrate fresh 500μL of HB-3 to reach RT

5. **Hairpin Amplification**:
   (a) Mix all hairpins in 500μL of pre-equilibrated HB-3 @RT *(Final concentration of each hairpin is 60nM)*
   (b) Add this hairpin solution to samples for overnight (16hrs) incubation @RT
6. **Hairpin Washes**
   (a) Wash 4 times, in 500μL 5× SSCT for 5 min each @RT
   (b) Wash 2 times, in 500μL 5× SSCT for 30 min each @RT
7. **Storage/Imaging**:
   Samples can be stored in PBST @4°C or imaged directly in 5× SSCT

B.3.3 Buffer recipes for embryo preparation

10× PBS	For 1 L of solution
1.37 M NaCl	80 g NaCl
27 mM KCl	2 g KCl
80 mM Na_2HPO_4	11.4 g Na_2HPO_4 anhydrous
20 mM KH_2PO_4	2.7 g KH_2PO_4 anhydrous
pH 7.4	Adjust pH to 7.4 with HCl
	Fill up to 1 L with ultrapure H_2O

4% Paraformaldehyde (PFA)	For 25 mL of solution
4% PFA	1 g of PFA powder
1× PBS	25 mL of 1× PBS

PBST	For 50 mL of solution
1× PBS	5 mL of 10× PBS
0.1% Tween 20	500 μL of 10% Tween 20
	Fill up to 50 mL with ultrapure H_2O

HB-1: Probe hybridization buffer	For 40 mL of solution
50% formamide	20 mL formamide
5× sodium chloride sodium citrate (SSC)	10 mL of 20× SSC
9 mM citric acid (pH 6.0)	360 μL 1 M citric acid, pH 6.0
0.1% Tween 20	400 μL of 10% Tween 20
50 μg/mL heparin	200 μL of 10 mg/mL heparin
1× Denhardt's solution	800 μL of 50× Denhardt's solution
10% dextran sulfate	8 mL of 50% dextran sulfate
	Fill up to 40 mL with ultrapure H_2O

HB-2: Probe wash buffer	For 40 mL of solution
50% formamide	20 mL formamide
5× sodium chloride sodium citrate (SSC)	10 mL of 20× SSC
9 mM citric acid (pH 6.0)	360 μL 1 M citric acid, pH 6.0
0.1% Tween 20	400 μL of 10% Tween 20
50 μg/mL heparin	200 μL of 10 mg/mL heparin
	Fill up to 40 mL with ultrapure H_2O

HB-3: Amplification buffer	For 40 mL of solution
5× sodium chloride sodium citrate (SSC)	10 mL of 20× SSC
0.1% Tween 20	400 μL of 10% Tween 20
10% dextran sulfate	8 mL of 50% dextran sulfate
	Fill up to 40 mL with ultrapure H_2O

5× SSCT	For 40 mL of solution
5× sodium chloride sodium citrate (SSC)	10 mL of 20× SSC
0.1% Tween 20	400 μL of 10% Tween 20
	fill up to 40 mL with ultrapure H_2O

B.3.4 Reagents and supplies

Alexa Fluor 488 carboxylic acid, 2,3,5,6-tetrafluorophenyl ester (Molecular Probes Cat. # A-30005)
Alexa Fluor 514 carboxylic acid, succinimidyl ester (Molecular Probes Cat. # A-30002)
Alexa Fluor 546 carboxylic acid, succinimidyl ester (Molecular Probes Cat. # A-20002)
Alexa Fluor 647 carboxylic acid, succinimidyl ester (Molecular Probes Cat. # A-20006)
Alexa Fluor 700 carboxylic acid, succinimidyl ester (Molecular Probes Cat. # A-20010)
N,N-Dimethlyformamide (DMF) (Alfa Aesar Cat. # 43997-AE)
Paraformaldehyde (PFA) (Sigma Cat. # P6148)
Formamide (Deionized) (Ambion Cat. # AM9342)
20× sodium chloride sodium citrate (SSC) (Invitrogen Cat. # 15557-044)
Heparin (Sigma Cat. # H3393)
50% Tween 20 (Invitrogen Cat. # 00-3005)
50× Denhardt's solution (Invitrogen Cat. # 750018)
Dextran sulfate (Sigma Cat. # D6001)
25 mm × 75 mm glass slide (VWR Cat. # 48300-025)
22 mm × 22 mm No. 1 coverslip (VWR Cat. # 48366-067)
SYBR Gold nucleic acid gel stain (Invitrogen Cat. # S-11494)
SlowFade Gold antifade reagent with DAPI (Molecular Probes Cat. # S36938)

B.4 Probe penetration studies

Systematic probe penetration can lead to decreased signal. While it does not affect the quantitative nature of HCR, for better signal , it is desired to work in the regime where insufficient penetration of probes is not an issue.

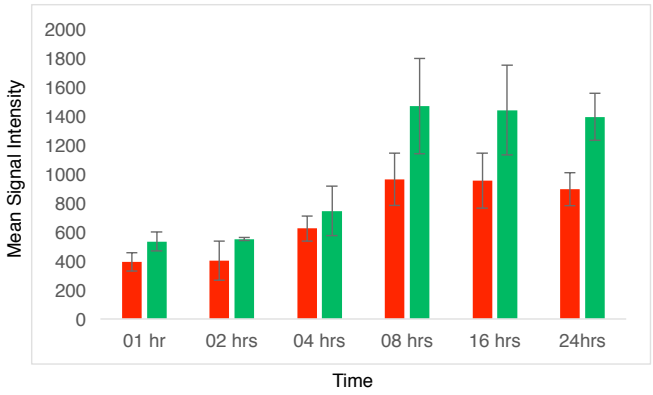


Figure B.1: **Probe penetration studies**
Graph showing mean signal intensity for various durations of probe hybridization (in hours) for *ntla*(green) and *citrine*(red). Error bars indicate standard deviation for 3 replicates

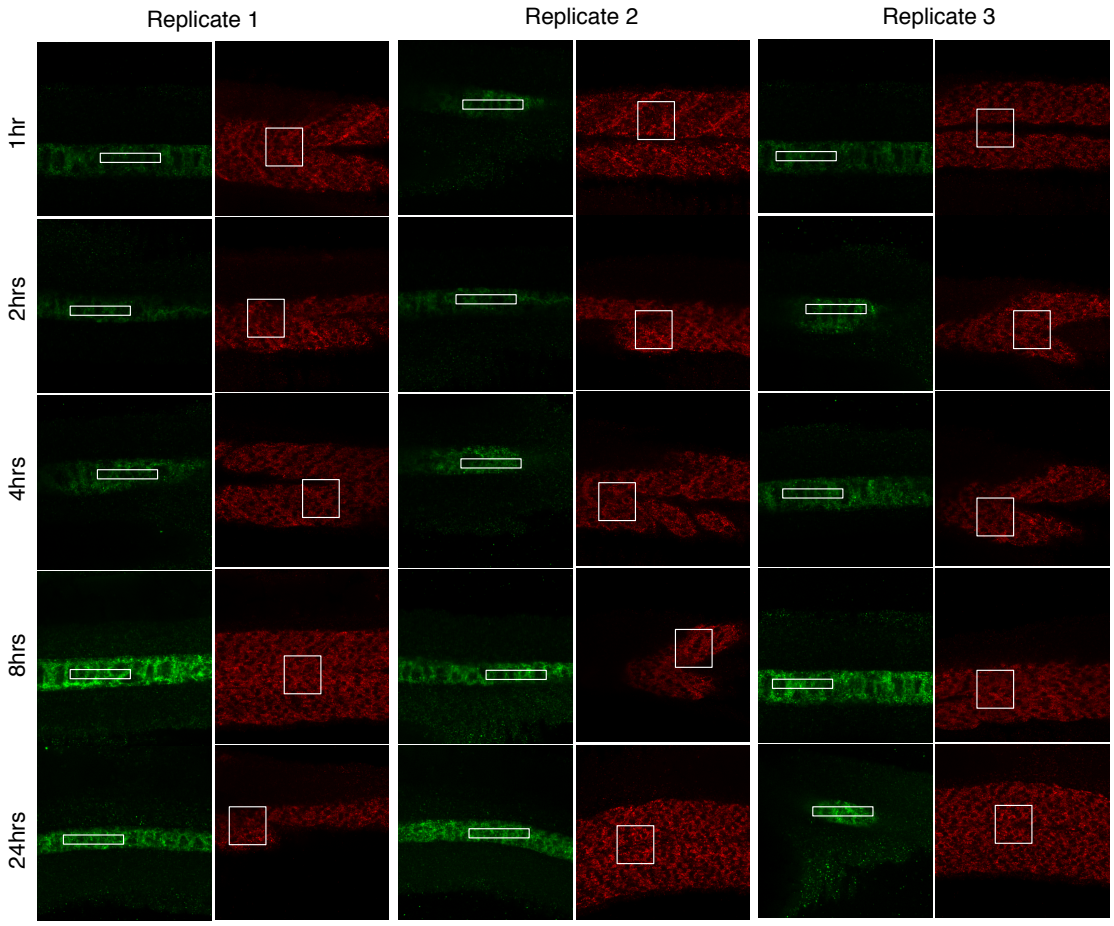


Figure B.2: **Replicates for probe penetration study in Fig. B.1**

B.5 Standard error in slope estimates and correlation between signals

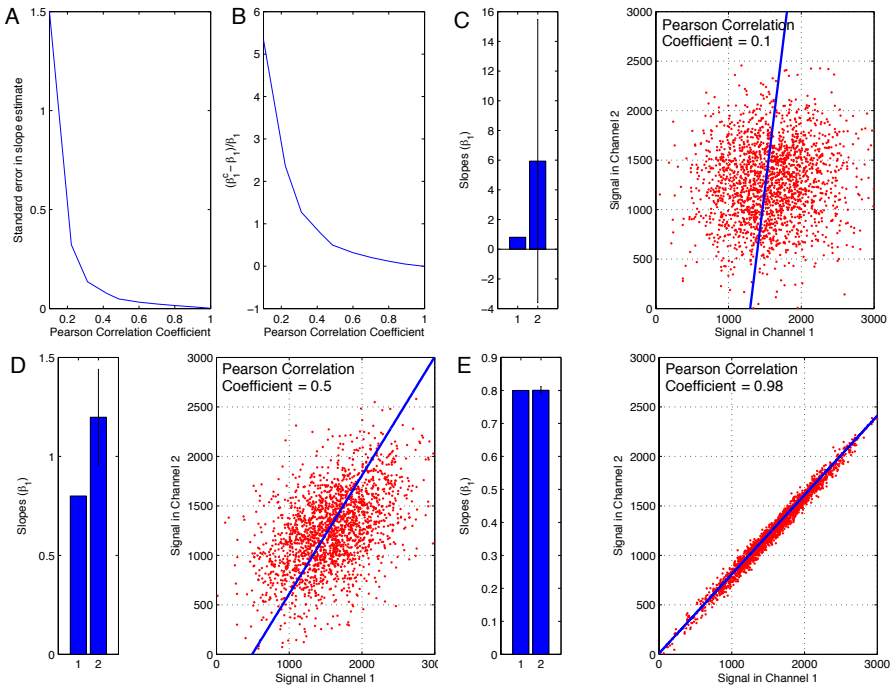


Figure B.3: **Relationship between standard error in slope estimates and correlation between signals**

(A) Standard error in slope estimates decrease with increasing value of Pearson correlation coefficient between signals thereby implying that correlation between signal provides a hold on accuracy of slope prediction. (B) Fractional error in predicated value of slopes for different values of Pearson correlation coefficient between signals (C-E) Representative simulation results for 3 different values of correlation coefficient between signals in the two channels. (Left) Comparison of the estimated slope (right bar) with its known value (left bar). Error bar on the right bar denotes estimated standard error. (Right) Plots of simulated voxel intensities in two channels. For each fit, sample size is 2000 voxels. Red dots denote simulated signal. The values of statistics assumed were close to those measured in experiments ($\mu_x = 1500, \sigma_{xx} = 500^2$, correlation coefficients are indicated).

B.6 Cooperativity between probes and hairpin polymers

Experimentally we have also observed evidences of cooperativity between probes or between polymers on a single target molecule (Fig B.4). Detecting the same mRNA molecule with different probes between different experiments has shown that such an intra-target cooperativity affects the signal intensities in such experiments. In Figure B.4 below we show an example (with *desmin* as a target) where the number of probes in the green channel is kept constant while the number of probes in the red channel is varied between different samples. As expected the mean signal intensities in the red channel are different between experiments Figure B.4B. However the mean signal intensity in green channel also changes between experiments and we attribute it to the cooperatively between different probes and the growing polymers which are different between these experiments. It must be noted though, that despite such effects (which appear to be mediated by mRNA), as long as the same probe/probe sets are used consistently throughout the experiment relative quantification remains unaffected. This allows us to compare *voxels* within an embryo as well as in different embryos exposed to the same probe set(s).

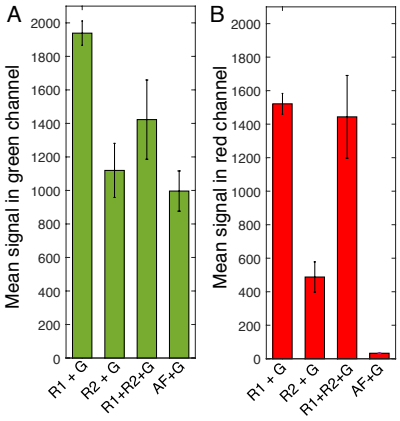


Figure B.4: **Cooperativity between probes and hairpin polymers**
Bar graphs showing mean signal intensities (for *desmin*) in the green and red channels for experiments where the number of probes in the green channel (probe set G) is kept constant while the number of probes in the red channel (probe sets R1, R2 , R1 and R2, and no probes (AF)) is varied between different samples. Cooperativity between probes and polymers change green signal intensities between samples where the mean intensities are expected to be constant.

B.7 Histograms of signal and background for data presented in Chapter 2

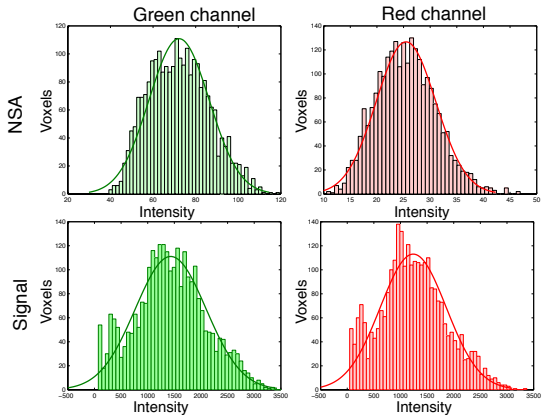


Figure B.5: **Histograms of signal and background for *desmin* in Fig. 2.2**

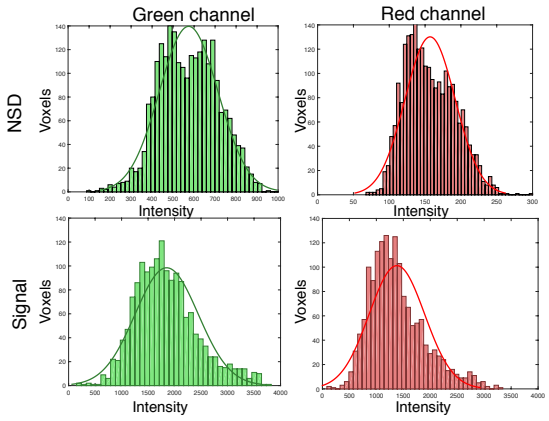


Figure B.6: **Histograms of signal and background for *citrine* in Fig. 2.2**

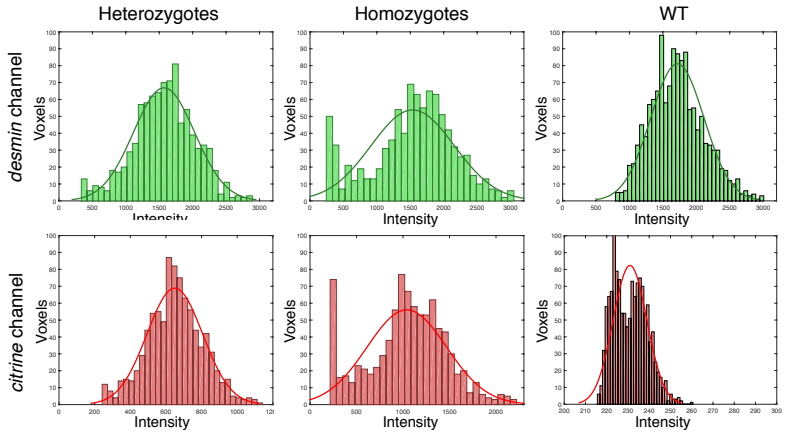


Figure B.7: **Histograms of signal and background for *desmin* and *citrine* in Fig. 2.3**

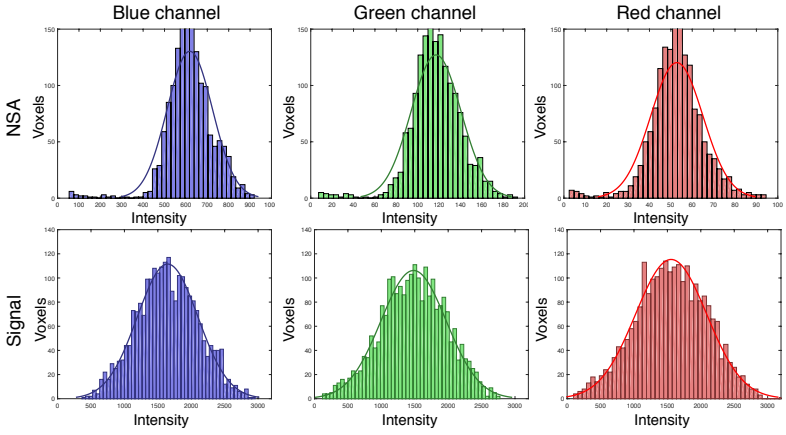


Figure B.8: **Histograms of signal and background for *desmin* in Fig. 2.4**

B.8 Replicates for data presented in Chapter 2

B.8.1 Replicates for redundant detection of *desmin* in Fig. 2.2

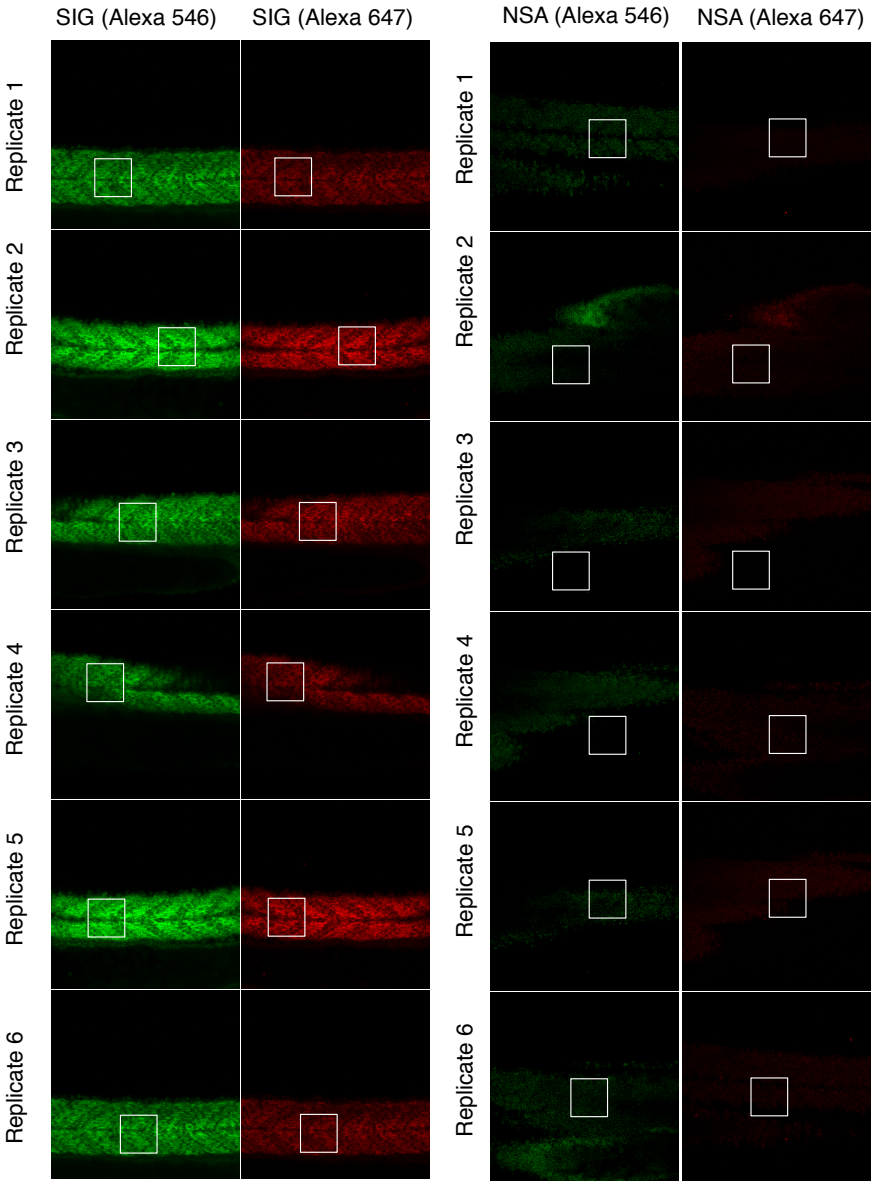


Figure B.9: **Replicates for redundant detection of *desmin* in Fig. 2.2**

B.8.2 Replicates for redundant detection of *citrine*in Fig. 2.2

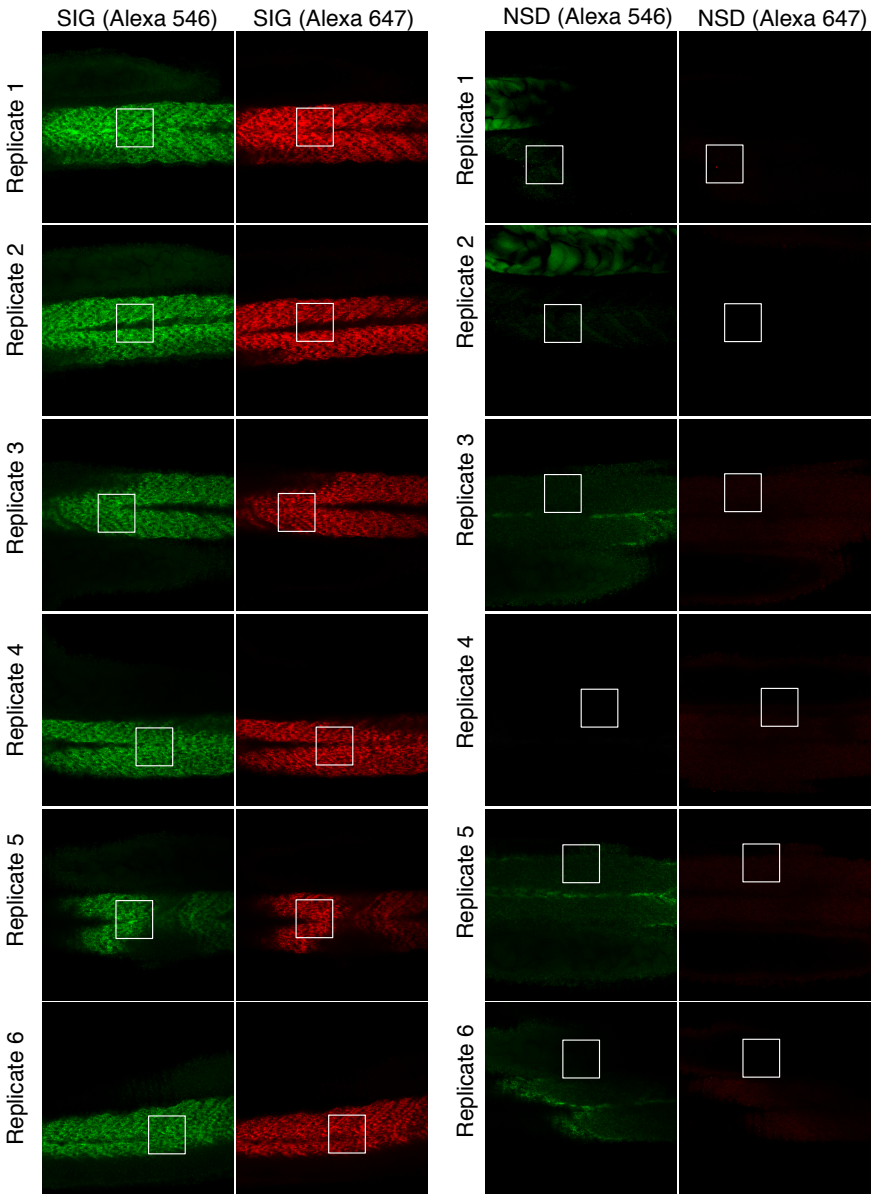


Figure B.10: **Replicates for redundant detection of *citrine* in Fig. 2.2**

B.8.3 Replicates for homozygous-heterozygous study in Fig. 2.3

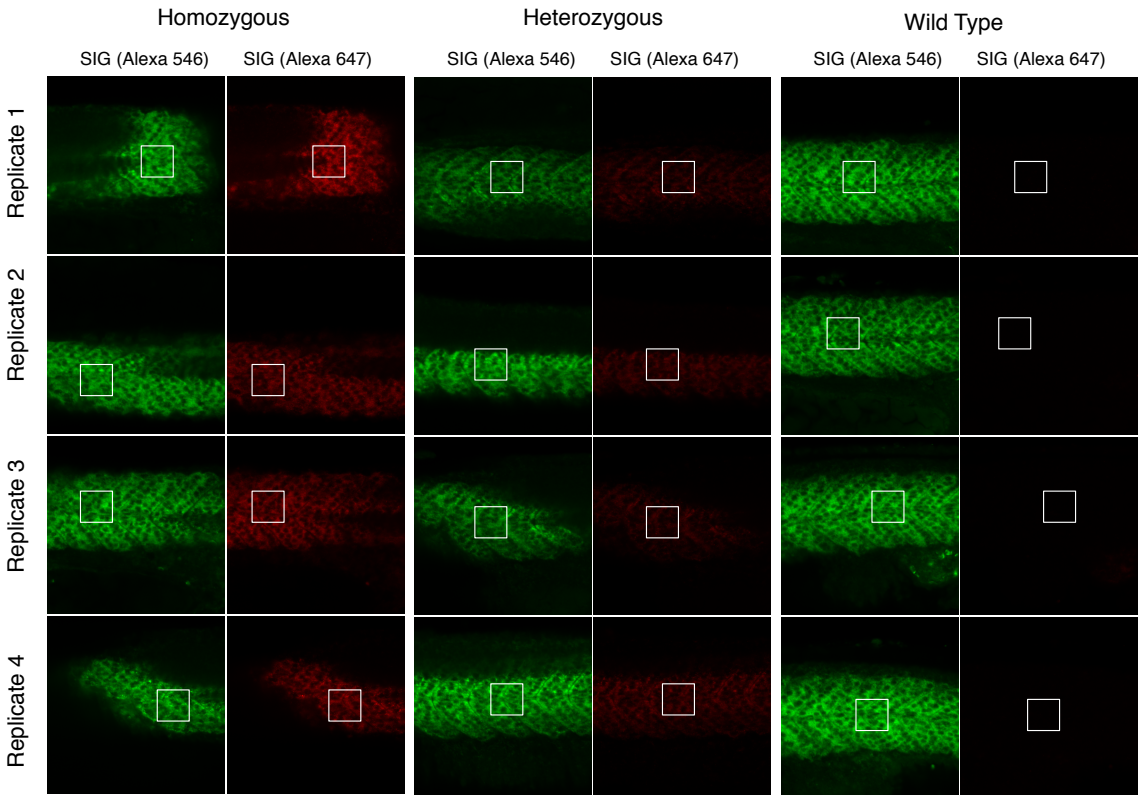


Figure B.11: **Replicates for homozygous-heterozygous study in Fig. 2.3**

B.8.4 Replicates for slope multiplication study in Fig.2.4

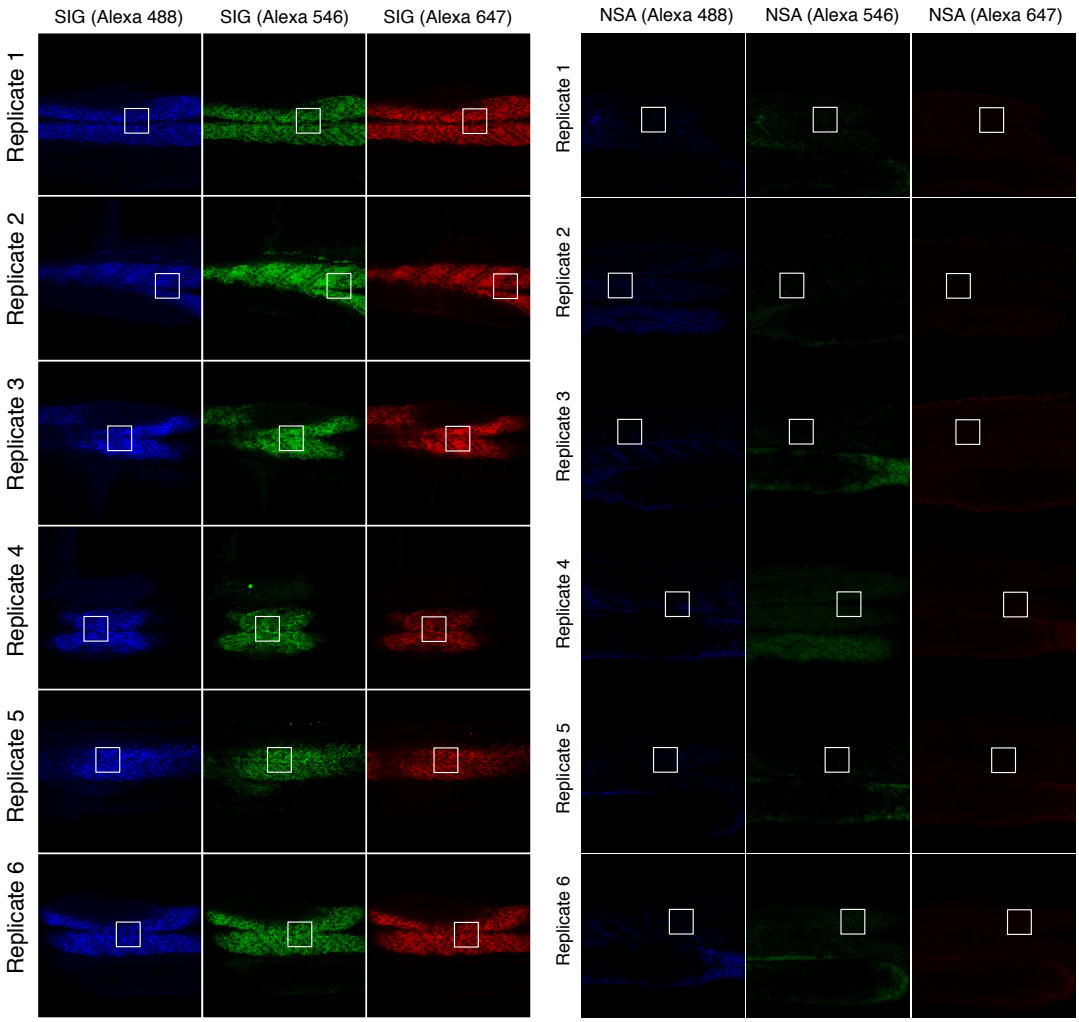


Figure B.12: **Replicates for slope multiplication study in Fig. 2.4**

B.9 Imaging parameters on the microscope

B.9.1 Detector Gain and Laser Power

It is crucial for our quantitative analysis that the imaging is done at the microscope settings such that the detector gain and laser power do not cross the manufacturer specified linear regimes. In Fig. B.13 an example is shown where we adjusted one of the microscope parameters in the red channel while keeping the others constant. We use the redundant detection experiment to show that beyond the specified detector gain limit (Fig. B.13A), we get non-linear relationship between the two channels which is incorrect. The absolute numbers on these parameters vary from one microscope to another and hence the result shown here must be interpreted as indicative.

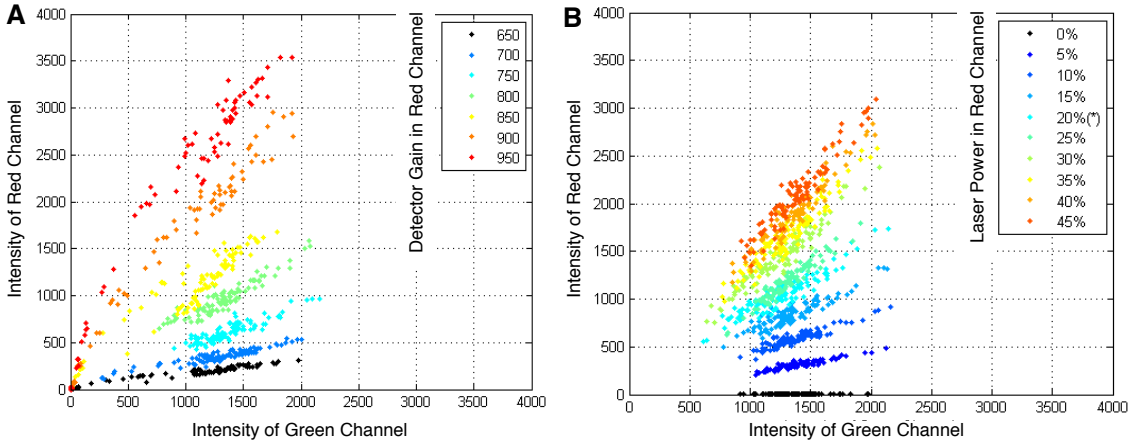


Figure B.13: **Limits on detector gain and laser power during imaging**

B.9.2 Imaging parameters for presented data

Fig.	**Stage** (hpf)	**Imaged volume (pixels)**	**Lateral resolution** (μm)	**Axial resolution** (μm)	**Pixel time** (μs)	**Pinhole size** (μm)	**488nm %**	**541nm %**	**633nm %**
2.2, B.3	26	512x 512x 1	0.35	1.5	3.15	70	N/A	10	4.5
2.2, B.4	26	512x 512x 1	0.35	1.5	3.15	70	N/A	12	12
2.3, B.5	26	512x 512x 1	0.35	1.5	3.15	70	N/A	10	6.9
2.4, B.6	26	512x 512x 1	0.35	1.5	3.15	70	15	10	4.5
2.5	15	512x 512x 70	0.35	1.5	3.15	70	10	13	11
B.2	26	512x 512x 1	0.35	1.5	3.15	70	N/A	12	4.5

Table B.1: Imaging parameters for data presented in Chapter 2

B.10 HCR probe sequences

B.10.1 HCR amplifier sequences

Initiator and hairpin sequences for the five DNA HCR amplifiers used in the present work. All sequences are listed 5′ to 3′.

`/5′-dye-C12/:` 5′ Alexa Fluor modification with a C12 spacer
`/C9-dye-3′/:` 3′ Alexa Fluor modification with a C9 spacer

DNA HCR B1

I1	gAggAgggCAgCAAACgggAAgAgTCTTCCTTTACg
I2	gCATTCTTTCTTgAggAgggCAgCAAACgggAAgAg
H1	CgTAAAggAAgACTCTTCCCgTTTgCTgCCCTCCTCgCATTCTTTCTTgAggAgggCAgCAAACgggAAgAg /C9-dye-3'/
H2	/5'-dye-C12/ gAggAgggCAgCAAACgggAAgAgTCTTCCTTTACgCTCTTCCCgTTTgCTgCCCTCCTCAAgAAAgAATgC

DNA HCR B2

I1	CCTCgTAAATCCTCATCAATCATCCAgTAAACCgCC
I2	AgCTCAgTCCATCCTCgTAAATCCTCATCAATCATC
H1	ggCggTTTACTggATgATTgATgAggATTTACgAggAgCTCAgTCCATCCTCgTAAATCCTCATCAATCATC /C9-dye-3'/
H2	/5'-dye-C12/ CCTCgTAAATCCTCATCAATCATCCAgTAAACCgCCgATgATTgATgAggATTTACgAggATggACTgAgCT

DNA HCR B3

I1	gTCCCTgCCTCTATATCTCCACTCAACTTTAACCCg
I2	AAAgTCTAATCCgTCCCTgCCTCTATATCTCCACTC
H1	CgggTTAAAgTTgAgTggAgATATAgAggCAgggACAAAgTCTAATCCgTCCCTgCCTCTATATCTCCACTC /C9-dye-3'/
H2	/5'-dye-C12/ gTCCCTgCCTCTATATCTCCACTCAACTTTAACCCggAgTggAgATATAgAggCAgggACggATTAgACTTT

DNA HCR B4

I1	CCTCAACCTACCTCCAACTCTCACCATATTCgCTTC
I2	CACATTTACAgACCTCAACCTACCTCCAACTCTCAC
H1	gAAgCgAATATggTgAgAgTTggAggTAggTTgAggCACATTTACAgACCTCAACCTACCTCCAACTCTCAC /C9-dye-3'/
H2	/5'-dye-C12/ CCTCAACCTACCTCCAACTCTCACCATATTCgCTTCgTgAgAgTTggAggTAggTTgAggTCTgTAAATgTg

DNA HCR B5

I1	CTCACTCCCAATCTCTATCTACCCTACAAATCCAAT
I2	CACTTCATATCACTCACTCCCAATCTCTATCTACCC
H1	ATTggATTTgTAgggTAgATAgAgATTgggAgTgAgCACTTCATATCACTCACTCCCAATCTCTATCTACCC /C9-dye-3'/
H2	/5'-dye-C12/ CTCACTCCCAATCTCTATCTACCCTACAAATCCAATgggTAgATAgAgATTgggAgTgAgTgATATgAAgTg

B.10.2 Probes for *desmin, citrine, ntla, tpm3, notch2, her7* and *dlC*

All probes are double initiator probes and have have the following format :

`Initiator-Spacer1-(Probe Sequence)-Spacer2-Initiator`

The sequence of spacers for different systems is as follows :

System	Spacer 1	Spacer 2
B1	ATATT	ATATA
B2	AAAAA	AAAAA
B3	TACAA	TAAAA

In the tables below probe sequences for the targets used in Chapter 2 are provided.

Target mRNA: **desmin (*desmin*)**

Probe #	Probe Sequence
1	TTggCTTCTCTgAgAgCCTCgTTATTCTTgTTCACTgCCTggTTCAAATC
2	TCTCgCAggTgTAggACTggAgCTggTgACggAACTgCATggTCTCCTgC
3	TgAAACCTTAgACTTATACCAgTCCTCggCCTCgCTgATATTCTTggCAg
4	CgATAgCCTCgTACTgCAggCgAATgTCTCTgAgggCCgCAgTCAggTCT
5	ggTTTggACATgTCCATTTggATCTgCACCTgACTCTCCTgCATCTggTT
6	CTgCAgCTCACggATCTCCTCCTCATgAATCTTCCTgAggAATgCAATCT
7	CTTCgTgAAgACCCTCgATACgTCTTTCCAggTCCAgCCTggCCAgAgTg
8	gCAgCATCgACATCAgCTCTgAAAgCAgAAAggTTgTTTTCAgCTTCCTC
9	CTCACTCATTTgCCTCCTCAgAgACTCATTggTgCCCTTgAgAgAgTCAA
10	AGGCTGAATATTTCGTGCTCATGACTGGGCCTGTTGGGTTTGTACGCTGT
11	TCTTGGTCACCTCGTAAACTCTGGAGGTCAGTCTTGAGGAGCCAGAGGAA
12	TCATTGAGGTGCTGGAGCTCGGCCTTCTCATTAGTACGCGTGTTGAGGAA
13	CACCTGTCCGCGCAGCTCTCTCATCTCCTCCTCGTACAGCTCTGCAATAC
14	TAGGTTGTCCCTCTCGATCTCCACACGGGATCTCTGATTGGTCAGTGCCT
15	TCTGGTGGATCTCCTCTTGAAGTCTGAGCTTTAGTTTCTGTAGGTCATCG
16	CATCTTCACATTCAGCAGATCCTGGTACTCGCGGAGGTGGCGGGCCATCT
17	AGCCTGAGGCCATCAGTGAAAATGCAAGGCATTCTGTGAAACTCGATCTG
18	GTCATCCACATGCGGTAATGGGTCACGGCAATACTGGGTGTGTGCTTAAA
19	CCACTTTTCCTCCCTCTGGTTTTTCAGCTCACTTTCCTTTTGTCTCCATG
20	TCAGAAACTCTGAGCTCATCTGGCCACTTTCTGAACGCTCAGCATCACAC
21	TTGTTGCTGTGGAACAGTGTCTCAGATATGACCCAACCACACACACACGC

Target mRNA: **citrine (*citrine*)**

Probe #	Probe Sequence
1	GTTCTTCTGCTTGTCGGCCATGATATAGACGTTGTGGCTGTTGTAGTTGT
2	TTCAgCTCgATgCggTTCACCAgggTgTCgCCCTCgAACTTCACCTCggC
3	ACGCTGCCGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCC
4	gCgggTCTTgTAgTTgCCgTCgTCCTTgAAgAAgATggTgCgCTCCTggA
5	CgTAgCCTTCgggCATggCggACTTgAAgAAgTCgTgCTgCTTCATgTgg
6	GCGGTCACGAACTCCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGTC

Target mRNA: **tropomyosin 3 (*tpm3*)**

Probe #	Probe Sequence
1	TggATCTCCTgCAgCTCCATCTTCTCCTCATCCTTCAgAgCCCTgTTCTC
2	CTTCATATTTgCggTCAgCCTCCTCAgCAATgTgCTTggCCTCCTTAAgC

Target mRNA: **ntla (*ntla*)**

Probe #	Probe Sequence
1	TCAAATAAAGCTTGAGATAAGTCCGACGATCCTACTAAATCCCGTTGGAT
2	CATTCGTCTCCCAGTCTTGGTGACAATCATTTCATTGGTGAGCTCTTTAA
3	CGTATTTCCACCGATTATTATCGGCCGCCACAAAATCCAGCAGGACCGAG
4	GTGTATCCTGGGTTCGTATTTGTGCAATGAGTTTAACATAATCTGTCCTC
5	TGTAGTTATTGGTGGTAGTGCTGCGGTGGGAGTAATGGCTGGGATATGGA
6	GTGTTTGTGGTGTGGGCCAGGGTTCCCATCCCGCTGGAGTTGGGGATCTG
7	TGTCAGGCCACCTGTAATGGAGCCCGATGCTGAGCCTGATGGGGTGAGAG
8	CCTCGCTTAGGCCTGGATCGTACATTGAGGAGGGAGAGGACACAGGCAGC
9	TCCTTAAATGTGAAGCGATCTCAGTAGCTCTGAGCCACAGGCGCCCATGA
10	gCAgCTCTgTggTTCCTCAAgCTggAgTATCTCTCACAgTACgAACCCgA
11	CAgggCTgACCAgCTgTCATgAgACgCAAgACTTCCggAAgAgTTgTCCA
12	TCgTCCCTgCAACTgACCACAgACTTgggTACTgACTggTgTTggAggTA
13	gAggAggTCAgACCCgAgTAggACATCgAAgAACCgCgTAggAACTgAgA
14	TgCTgTgAgCCgggCgATggAgCTCTCgAACTgggCATCTCCAACgCCAA
15	TTCTAgATTTCCTCCTgAAgCCAAgATCAAgTCCATAACTgCAgCATCAg
16	TggggCCgTTACTgggCAggAACCAgCCACCgAgTTgTgAATATCCAgAT

Target mRNA: **notch2 (*notch2*)**

Probe #	Probe Sequence
1	TACAAGGGTCACTGTGATGACAGTACTCCCCTAAGAAACCAGGAGCACAC
2	TGGTGCAAGGGTTATTGGGGTAGCAGGTGGAGTTTTTCTGGATTTCGCAG
3	ATTGCATATACCCTTGTTGGCACATGGGCTGGACAGACATGCGTCCTCTT
4	ACCATGAAAACCAGGCGCACAGGTACAAGAGTATTCGCGATTTGGCAGAG
5	TAAAACCTGGCAGGCAGTGGCATACATAGGTGGTTTCTGTTGTTTGCCGG

Target mRNA: **her7 (*her7*)**

Probe #	Probe Sequence
1	tctgtgccaggattttcattgcacgtgtactccaatagttgatatcaatg
2	agtctgatggtggggtgagcaagacgttgacacagcatggatgtgaccga
3	tgctgtgggttggattctgcgtgctgctttctggagtgacctctgacagg
4	gggaggttggttgagctgtcgaatgctttgtgttggagtctgtgcttctg
5	gcagaagtttcaggttttctagactccggttcatcctttctcttctacgt
6	ttcctccaagagaaagcgggctgctttttgaagacaggaagaaaagcctt
7	tcagttaaggccaaggtctccaaacagtctggctggctggttctgggtgg
8	ggatgtcctgctttggagacagtgtttttgtggtgtggacgacgcgcatg
9	cgcttcattcgccttttgcaggaacagaactgtgtattccaggatctctg

Target mRNA: **deltaC (*dlC*)**

Probe #	**Probe Sequence**
1	AGTGCTTCAGGCAAACACGGAAAAAGATCTGGCAGTCGCTGGACCCTTTA
2	AGTATTTCTGACATTCCGGTGCCGTAGGTGCACGGCGGCTCAGGTAATAT
3	CTCGCCGTGGTAGAATTCATCGCACACGACACGATAAGAAAAGCGCAGTT
4	AATAATCGCCTTTCCATCCAGGAAGGCAAATTCTGTTGCCAGCGGCGTCA
5	CGGTGTTGGTACACGTGGCGTCATTTCTACAGGGCTTGTGATTAGTGCAA

Appendix C

Supplementary Information for Chapter 4

C.1 Sample preparation

The viral-infected chick metacarpals were dissected at E8 (HH 32?33) for organ culture. Molten agarose was poured into the fluorodish (World Precision Instruments) and the custom-designed mold was immediately inserted into it. When the agarose was solidified, the mold was pulled out, leaving grooves in the agarose for holding the metacarpals. The metacarpals were submerged in DMEM/F12 growth medium containing 10 mM β-glycerophosphate, 0.2% bovine serum albumin, 50 mM ascorbate acid (Sigma-Aldrich) and 1% glutamine-penicillin-streptomycin (Invitrogen) in a humidified chamber at 37°C on the stage of the inverted laser scanning microscope (LSM 710, Carl-Zeiss). For imaging chick metacarpal expressing GFP, two-photon laser excitation was used with 26% relative power at wavelength of 900 nm. Optical sectioning was achieved at intervals of 1 μm. For imaging quail metacarpal expressing H2B-mCherry, one-photon laser excitation was used with 26% relative power at wavelength of 561 nm and same optical sectioning.

C.2 Image analysis

C.2.1 Cell segmentation and tracking

The 4D images were imported into IMARIS 7.6.4. The chick metacarpal expressing GFP had a thickness of 193 μm. The middle 90-mm region spanning both stem and loop parts was digitally sliced into three sections along the imaging axis (z axis) (Fig. 4.1C). Three-dimensional spot segmentation of cells and subsequently 4D tracking of cells were performed for each section individually. The cell trajectories from all three sections were combined for quantitative analysis in MATLAB R2013b. Four-dimensional image of quail metacarpal expressing H2B-mCherry had sparse enough signal due to nuclear labels (as opposed to ubiquitous cytoplasmic label in chick). In addition the relatively low penetration of the one-photon excitation only generated images with good resolution for a thickness of $80\mu m$. Therefore, digital slicing of this data set was not required, and all the nuclei were segmented and tracked using IMARIS 7.6.4.

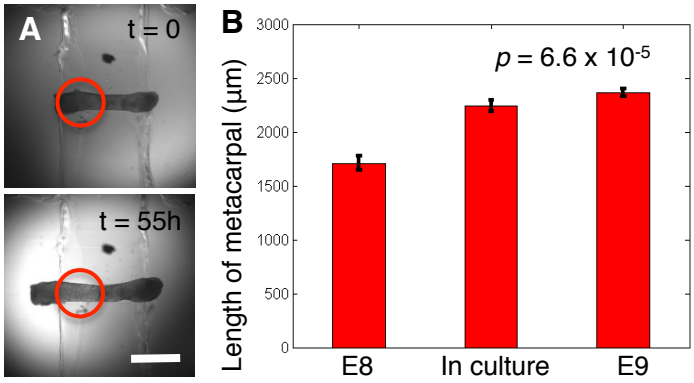


Figure C.1: **Live imaging of cartilage growth**
(A) Top view of the metacarpal sitting in the groove. (Below) The same metacarpal exhibiting similar extension on both sides after live imaging for 55 h. The red circles depict the region imaged. (B) Chick metacarpals were dissected at E8 and cultured for 24 hours. The cultured tissue and the tissue dissected from E9 embryos showed similar elongation (2 experiments, n = 21 metacarpals per experiment). Scale bars: (A) 500 μm,

C.2.2 Cell displacement based measurements

Total displacement lengths in a given direction were calculated by subtracting the relevant coordinates at time t = 0 and a time of interest. On the basis of the coordinates of the tracked cells, center-to-center distance between a cell and every other cell was measured, the distribution of which was used for calculating the mean distance along different orientations. This process was repeated for all the cells and a distribution of average distances was plotted for each time point to depict the shifts in distances over time. The mean of each distribution was further calculated and plotted over time to show statistically average behavior of all the cells over time.

C.2.3 Analysis of daughter cell rearrangement

Based on the definition of the 3D distance (r) and the 3D angle (Θ) between two daughter cells, as described in Fig. 4.3C. The axis direction is the same at all time points, but the origin is fixed on one of the daughter cells. Changes (increase or decrease) in r denote relative displacement of daughter cells (away or towards) over time. Changes in Θ denote rearrangement of the vector joining the center of daughter cells with respect to the xz plane.

C.3 Image based voxel analysis

C.3.1 Image binarization

The images were imported as Tiff files in MATLAB 2013b and cropped based on the coordinates of the four cells (in the upper region of PZ) on the boundaries used as the references for drawing the edges of the box for analysis at all times (Fig. 4.5C). Cropped images were binarized frame-by-frame based on 10 different values of thresholds around the average voxel intensity of that frame. On the basis of the above thresholds, every voxel was assigned an identity as either belonging to cell or to ECM based on the value of one or zero in the binary

image, respectively. This process was repeated for all time points and the total numbers of bright and dark voxels were calculated.

C.3.2 Computer simulation of cellular trajectories

Given that the majority of cartilage growth happens along the PDA, y axis (Fig. 4.2), we assume that measured rates of change of cell volume and ECM volume (the slopes of the graphs in Fig. 5.5 E) come entirely from the changes along the y axis. We thus took these measured rates of change and the initial positions of all the cells determined by our imaging and tracking as inputs into a custom-written MATLAB code to simulate the expected cell trajectories. As depicted in Fig. 4.6B, at every time point the intercellular distance was increased assuming an anisotropic but uniform change along the y axis. New cellular positions were predicted at every time point, which were then iteratively used as inputs into the next time step to construct the trajectory over the total duration of imaging. Note that our simulation did not contain any freely adjustable parameters-the only inputs are the measured rates of change in cell and ECM size, and the imaged and tracked cellular positions.

Appendix D

Supplementary Information for Chapter 5

D.1 Zebrafish lines

This study was carried out in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the California Institute of Technology. The protocol was approved by the Institutional Animal Care and Use Committee (IACUC) of the California Institute of Technology (Permit Number: 1227). Wild-type embryos were obtained from AB and TL strains. Adult fish were maintained as described before*

D.2 Sample preparation for live imaging

Imaging was performed on a Zeiss LSM510 inverted confocal microscope. Embryos were placed into agarose mould and imaged with the following lasers: 514nm to excite Citrine; 543nm to excite Bodipy-TR, Alexa-546 and membrane-mCherry; 458nm to excite H2B-cerulean. Images were taken using a water- immersion 40X 1.1NA Apochromat objective using the Zeiss LSM software and time-lapse imaging was performed using the Zeiss multi-time macro. To permit accurate comparisons, all embryos in a given experiment were imaged using the same laser power; gain detector settings and scanning speed. For live imaging, BODIPY-TR (Life Technology, catalogue #C34556) was used as a counterstain. 3D reconstructions and projections were performed using Imaris 6.4 (Bitplane). Custom written codes in MATLAB 2013b were used to quantify the number of pixels and their intensities. The kymograph was generated by assigning all the pixels of a given time point one single color.

D.3 Protocols for *in situ* hybridizations

D.3.1 Chromogenic *in situ* hybridization (ISH)

Chromogenic ISHs were performed and imaged as described before †. A digoxigenin (DIG) probe was synthesized using a plasmid template that contained a 2Kb fragment of the 3?

*Westerfield, M., The Zebrafish Book. Edition 2.1 ed. 1994, Eugene, OR: University Oregon Press

†Trinh, L.A., et al., Fluorescent in situ hybridization employing the conventional NBT/BCIP chromogenic stain. Biotechnique, 2007. 42(6): p. 756-759

end of the dmd gene. Embryos were dechorionated and fixed in 4% PFA overnight at 4°C before being dehydrated in a series of methanol and phosphate buffered saline with Tween (PBST; 1X PBS, 0.1% Tween-20). Proteinase K was used to permeabilize embryos before washing and pre-hybridized at 70°C for 1hour. Hybridization was performed overnight at 70°C and DIG probe was detected with 1:5000 anti-DIG alkaline phosphatase. Stained embryos were imaged with both wide-field and confocal microscopes. To improve the 3D localization and resolution of the ISHs, we used the technique described previously to detect the far-red fluorescence of the NBT/BCIP chromogenic stain using a 510 LSM Zeiss confocal microscope.

D.3.2 *in situ* Hybridization Chain Reaction (HCR)

The protocols for *in situ* HCR is the same as described in Appendix B.

D.4 HCR probe sequences

D.4.1 Probes for *dmd, citrine* and *tpm3*

Target mRNA: **dystrophin (*dmd*)**
Amplifier: **RNA HCR A1**
Fluorophore: **Alexa Fluor 647**

Initiator	Spacer	Probe Sequence
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	AgCTgCTCCTCAgAgAAATgCTgCCA gCACAgCAggATCTTTTgAAggAg
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	ggATggACTggACCAACTCCTCCTTT TgAgTCAgCCAAgAATCAAACAgC
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	CTTCCgCAggTTggCggCTACTTCATT TgggTCA TTTgTgCCACTgCTTT
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	TTgTCCATggTCTgCCTCCTCAgCTC CAggTCTgCTTTTAAAATCgCCAg
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	CCTCTTTgCTCTTgATgTTAgTCAgCA gATCCTggACgAgCgAgCAAAgC
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	ATTTgTCCCAgCgCTgAgCAAACCTC TCCAgCTTTgCTTCCAgCTTCCCA
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	ggTgACAATAgTTgAAATCTTggTgCT ggTgAgCTgAAgCgCCTgCACCA
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	ACCTTggTgACAgTTgCCATggTTgCg TgCgTgA TCTCCgACTgggA TgT
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	TggACATgCCCTCCTTAgTATgCTTC ACCA TCTTCTTCTggTTCgTggTC
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	TCCgCAgTTCTgAATCCACgACAAT CTgTCTCTTTTTCTgAggTggAggg
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	gAgTTCTgCCATCTTTgggTggTCCC AACATgTTgTTTgggTCTggTgAT
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	AAACAgAgggCTTTCTgCATTCgTCT gAgCTTCA TTgCCgTCCTgT A TgC
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	TTTgAgATTgTgCTgCTCAAAggCTT CACAggCTgCAggCATgCTCAgAA
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	ATgCTggTCAgACAgTTgATCACCTg CACgATgTCCATgAACTgCTCgTT
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	TAAACgTTgAgCAgCCAgTTgAgACA CA TgTCCACACAgAgAggCACgTT
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	TgATTCAgACTCTggCAgTAgTgCTg gA TCAgCAgA TgCTCA TCA TCCCT
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	TCTCTCCAgCTCTCCCTTCTCTTCA gTCTCCA TTgAgA TgAggA TCTgAg
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	ACgATCgTACTCCgCTTgCAgCTTC CTgTTTTCCTgCTCCAgA TCA TTgA
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	AATCgTCCCTTgTgTTgCCgCAgTAg TTTggCTTCTgCgA TT AgTTCTgC
CCgAAUACAAAgCAUCAACgACUAg	AAAAAA	CTgCgATTCCAgCTgCTTgTTgTgAT CCTCCAggA TTTgCA TTCTTgCCT

Target mRNA: **citrine (*citrine*)**
Amplifier: **RNA HCR A2**
Fluorophore: **Alexa Fluor 488**

Initiator	Spacer	Probe Sequence
CCgAAUACAAAgCAUCAACgACUAg	AAAAA	UCCUCAACCAgCUggAUACgCCUgUUCAgAgAAgCCACCUCUgCCUCAgC
CCgAAUACAAAgCAUCAACgACUAg	AAAAA	CCAgCUUUUgCAgggCUgUggCCAg UCUCUCCUgAgCACgAUCCAACUCC
CCgAAUACAAAgCAUCAACgACUAg	AAAAA	AAUCACCUUCAUCCCUCUCUCgC UCUCAUCUgCggCCUUCUCggCUUCCU
CCgAAUACAAAgCAUCAACgACUAg	AAAAA	UggAUCUCCUgCAgCUCCAUCUUC UCCUCAUCCUUCAgAgCCCUgUUCUC
CCgAAUACAAAgCAUCAACgACUAg	AAAAA	CUUCAUAUUUgCggUCAgCCUCCU CAgCAAUgUgCUUggCCUCCUUAAgC
CCgAAUACAAAgCAUCAACgACUAg	AAAAA	CUCUgUACgCUCCAACUCUCCCUC AACgAUCACCAgCUUACgAgCCACCU
CCgAAUACAAAgCAUCAACgACUAg	AAAAA	UggUUUUCUCCAgUUUggCCACAgA CCUCUCAgCAAACUCUgCACgggUC

Target mRNA: **tropomyosin 3 (*tpm3*)**
Amplifier: **RNA HCR A3**
Fluorophore: **Alexa Fluor 546**

Initiator	Spacer	Probe Sequence
TACgCCCTAAgAATCCgAACCCTATg	AAAAAA	gAACAgCTCCTCgCCCTTgCTCACTC CggATAACTTCgTATAATgTATgC
TACgCCCTAAgAATCCgAACCCTATg	AAAAAA	TTgCCggTggTgCAgATgAACTTCAgg gTCAgCTTgCCgT AggTggCA TC
TACgCCCTAAgAATCCgAACCCTATg	AAAAAA	TTgAAgAAgTCgTgCTgCTTCATgTgg TCggggTAgCgggCgAAgCACAT
TACgCCCTAAgAATCCgAACCCTATg	AAAAAA	TCTTgTAgTTgCCgTCgTCCTTgAAgA AgA TggTgCgCTCCTggACgT Ag
TACgCCCTAAgAATCCgAACCCTATg	AAAAAA	ATgTTgCCgTCCTCCTTgAAgTCgATg CCCTTCAgCTCgA TgCggTTCAC
TACgCCCTAAgAATCCgAACCCTATg	AAAAAA	TgATATAgACgTTgTggCTgTTgTAgTT gT ACTCCAgCTTgTgCCCCAgg
TACgCCCTAAgAATCCgAACCCTATg	AAAAAA	TgTTgTggCggATCTTgAAgTTCACCT TgA TgCCgTTCTTCTgCTTgTCg
TACgCCCTAAgAATCCgAACCCTATg	AAAAAA	TTCTCgTTggggTCTTTgCTCAgggCgg ACTggT AgCTCAggT AgTggTT

Appendix E

Supplementary Information for Chapter 6

E.1 Zebrafish lines

Adult fish were raised and maintained as described before* in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals by the respective bodies at California Institute of Technology and University of Southern California. The protocol was approved by Institutional Animal Care and Use Committee (IACUC) at both universities. Wild-type embryos were obtained from AB strains. FlipTraps *Gt(tpm4a-citrine)* $^{ct31a/+}$ and *Gt(desm-mCherry)*$^{ct122aR/+}$ transgenic embryos were obtained from a previously described screen†. The wild-type transcript knockdown in embryos heterozygous for *Gt(desm-mCherry)*$^{ct122aR/+}$ is not sufficient enough for lethality and the embryos develop normally. Upon crossing appropriate adult lines, the embryos obtained were raised in embryo medium (Instant Ocean) at 28.5°C with addition of 0.003% (w/v) 1-phenyl-2-thiourea (PTU) around 18hpf to reduce pigmentation.

E.2 Sample preparation for live imaging

5-6 embryos at appropriate stage were placed into 1% agarose (Catalog No. 16500-100, Invitrogen^{TM}) moulds created in a petridish, using a custom designed negative plastic mould‡. In order to arrest heart beat for imaging, an immobilization solution was simultaneously prepared by adding 2ml of 0.01% Tricaine to 8ml of 1% UltraPure^{TM} Low Melting Point Agarose (Catalog No. 16520-050, Invitrogen^{TM}) solution prepared in 30% Danieau (17.4mM $NaCl$, 210μM KCl, 120μM $MgSO_4$?7H_2O, 180μM $Ca(NO_3)_2$, 0.1mM HEPES buffer in water, pH 7.6) with 0.003% PTU. This solution was then added on top of the embryos already placed in the mould. Following solidification of agarose at room temperature (1-2 minutes), the petridish was filled with 30% Danieau solution and 0.01% Tricaine, at 28.5°C. Subsequent imaging was performed on an inverted confocal microscope by positioning the petridish appropriately on the microscope stage.

*Westerfield, M., The Zebrafish Book. Edition 2.1 ed. 1994, Eugene, OR: University Oregon Press

†Trinh LA, Hochgreb T, Graham M, Wu D, Ruf−Zamojski F, et al. (2011) A versatile gene trap to visualize and interrogate the function of the vertebrate proteome. Genes and development 25: 2306−2320

‡ Megason, S.G., In toto imaging of embryogenesis with confocal time-lapse microscopy. Methods in molecular biology, 2009. 546: p. 317-32

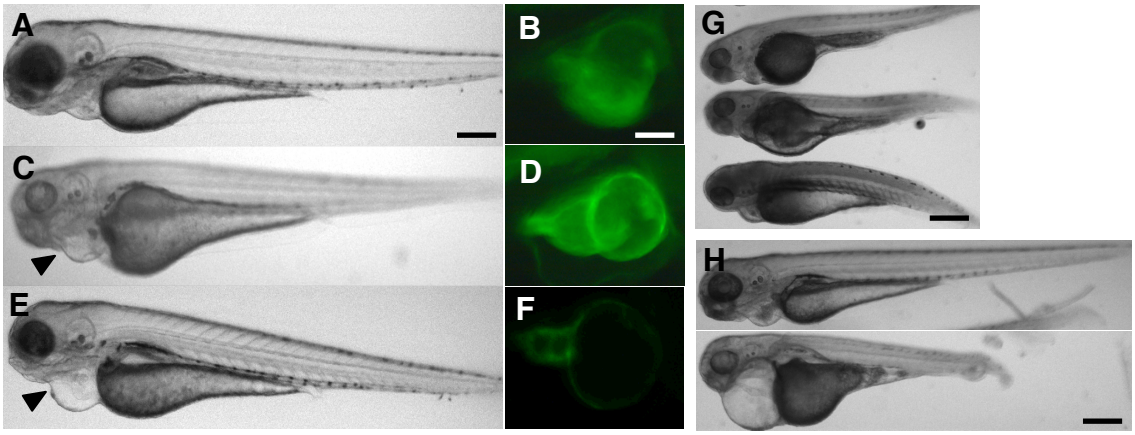


Figure E.1: **Widefield images of zebrafish embryos at 60hpf**
(A-B) *Gt(tpm4a-citrine)* $^{ct31a/+}$ (C-D) *Gt(tpm4a-citrine)* $^{ct31a/ct31a}$ (E-F) *Gt(tpm4a-citrine)* $^{ct31a/+}$ injected with anti-*amhc* morpholinos. Arrowheads denote edema in mutant embryos. Scale bars: (A,C,E,G,H) 500μm; (B,D,F) 20μm

E.3 Imaging

Imaging was performed on a Zeiss LSM710 inverted confocal microscope. Samples were imaged with the following lasers: 488nm to excite Citrine and 555nm to excite membrane-mCherry. Images were taken using a water-immersion 40X 1.1NA Apochromat objective using the Zeiss Zen software. To permit accurate comparisons, all embryos at a given stage were imaged using the same laser power, detector gain settings and scanning speed. Table E.1 provides the detailed description of the imaging parameters used for all images presented in the chapter. Wide field imaging was performed on Olympus Fluorescent microscopes.

E.4 Image analysis

E.4.1 Image segmentation

3D reconstructions and projections were performed using Imaris 7.7.2 (Bitplane). We segmented the cells manually by marking distinct points (spots) along the cardiomyocytes? boundary, following the Desmin-mCherry signal, using the ?Measurement Points? tab. Sufficient number of these nearly-equally spaced points were connected to represent cell boundary. ?Measurement Points? within IMARIS provides the position of points in 3D as well as angle between adjacent pairs of points. These statistics were then extracted as used as input to custom written codes in MATLAB 2014b. The sum of distances between points along the boundary provides the estimate of cell perimeter. This gives us an accurate estimate of the area of every cell. Curvedness, defined as deviation from a plane, is calculated by fitting the points to a plane. The sum of absolute distances between the point and its ideal position within the plane is used.

E.4.2 Computer Simulation of cardiac looping

Our simulation model assumes a simplistic approach, where we take the geometry of the cardiac tube at 24hpf as the initial geometric input to the model. We further allow growth based on the observed increase in cell sizes. To take into account for the differential growth observed between atrial and ventricular populations, we restrict the growth to the region of the heart tube that will eventually form atrium based on our imaging results.

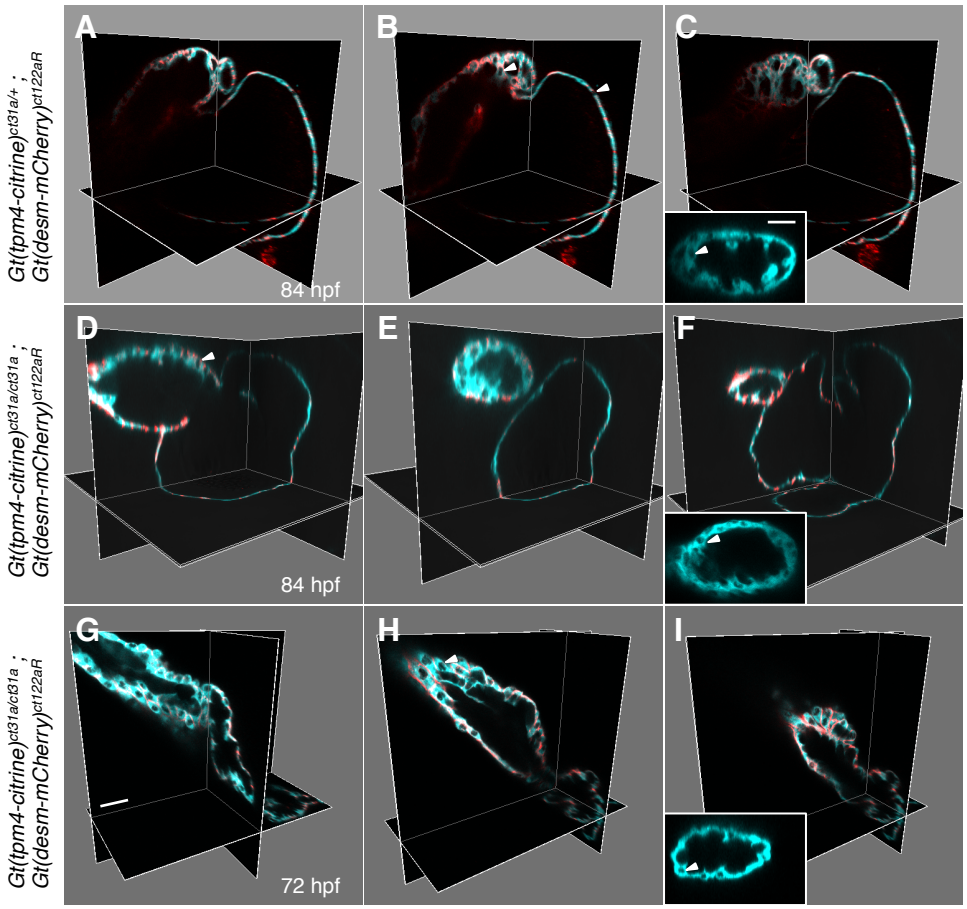


Figure E.2: **Orthogonal views of 3D rendered confocal images of zebrafish heart** (A-C) *Gt(tpm4a-citrine)* $^{ct31a/+}$ (D-F) *Gt(tpm4a-citrine)* $^{ct31a/ct31a}$ (G-I) *Gt(tpm4a-citrine)* $^{ct31a/ct31a}$ with arrested cell growth. Arrowheads denote nuclei visible in different cross-sections. Insets in panels (C,F,I) show trabeculation in ventricle. Scale bars: $10\mu m$

E.5 Morpholino injections

anti-*amhc* Morpholino (5?-ACTCTGCCATTAAAGCATCACCCAT-3?) were ordered from Gene Tools and diluted following manufacturer instructions. Embryos, obtained by crossing FlipTraps *Gt(tpm4a-citrine)*$^{ct31a/+}$ and *Gt(desm-mCherry)*$^{ct122aR/+}$ zebrafish adults, were injected at the 1-cell stage with 1.25ng of morpholino into the yolk. Embryos were then raised, at 28.5°C and sorted for both Citrine and mCherry expression. Doubly positive embryos were raised further for imaging at appropriate stages.

E.6 Imaging parameters for presented data

Fig.	**Stage** (hpf)	**Imaged volume (pixels)**	**Lateral resolution** (μm)	**Axial resolution** (μm)	**Pixel time** (μs)	**Pinhole size** (μm)	**488nm %**	**555nm %**
6.1E	24	512x 512x 104	0.31	1.5	3.15	70	3.3 - 10	11.8
6.1F	28	512x 512x 128	0.35	1	3.15	70	3.3 - 10	11.8
6.1G	32	512x 512x 79	0.39	1.5	3.15	70	4.1 - 13	16.6
6.1H	36	512x 512x 141	0.31	1.5	3.15	70	4.1 - 13	16.6
6.1I	48	512x 512x 137	0.39	1.5	1.58	70	6.9	8.3
6.1J	60	512x 512x 70	0.63	2	3.15	70	7	11
6.1K	72	512x 512x 91	0.31	1.5	3.15	70	6 - 9	11
6.1L	84	512x 512x 109	0.45	1.5	3.15	70	2.7-16.7	8 -14.2
6.4A	28	512x 512x 84	0.31	1.5	3.15	70	1.2	16.6
6.4B	32	512x 512x 95	0.31	1.5	3.15	70	0.7 -5.6	10.8 -14.3
6.4C	36	512x 512x 109	0.31	1.5	3.15	70	0.4-2.6	6.2 -8.2
6.4D	48	512x 512x 132	0.39	1.5	3.15	70	3	4.2
6.4E	60	512x 512x 118	0.39	1.5	3.15	70	0.7 -11.8	4.2
6.4F	72	512x 512x 113	0.39	1.5	3.15	70	2.1 - 10	6.5 -7.2
6.4G	84	512x 512x 118	0.39	1.5	3.15	70	3.3-11.3	7.2
6.4H	72	512x 512x 71	0.39	1.5	3.15	70	5.2 -10.4	8.6
E.4B	48	512x 512x 119	0.35	1.5	3.15	70	3.1-14.1	-
E.4C	72	512x 512x 79	0.35	1.5	3.15	70	7 - 9	11

Table E.1: Imaging parameters for data presented in Chapter 6

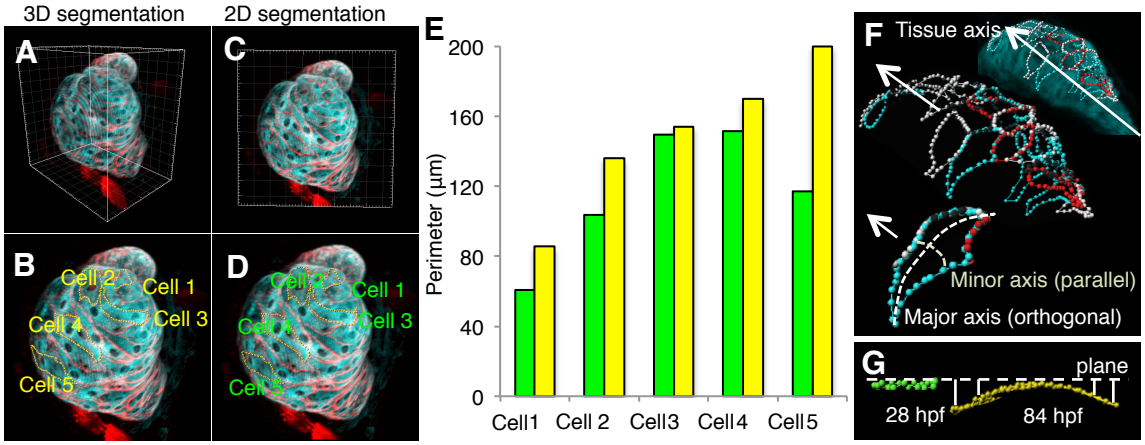


Figure E.3: **Comparison of cell size in 2D projections vs 3D renderings** (A-D) *Gt(tpm4a-citrine)* $^{ct31a/+}$ at 84hpf (E)Plot showing cell size measured with the two approaches.It can be seen that 2D measurements depend on the angle of projection leading to incorrect measurements. (F)Major and minor axes in heart. (G) Schemtic for measuring deviations from a plane.

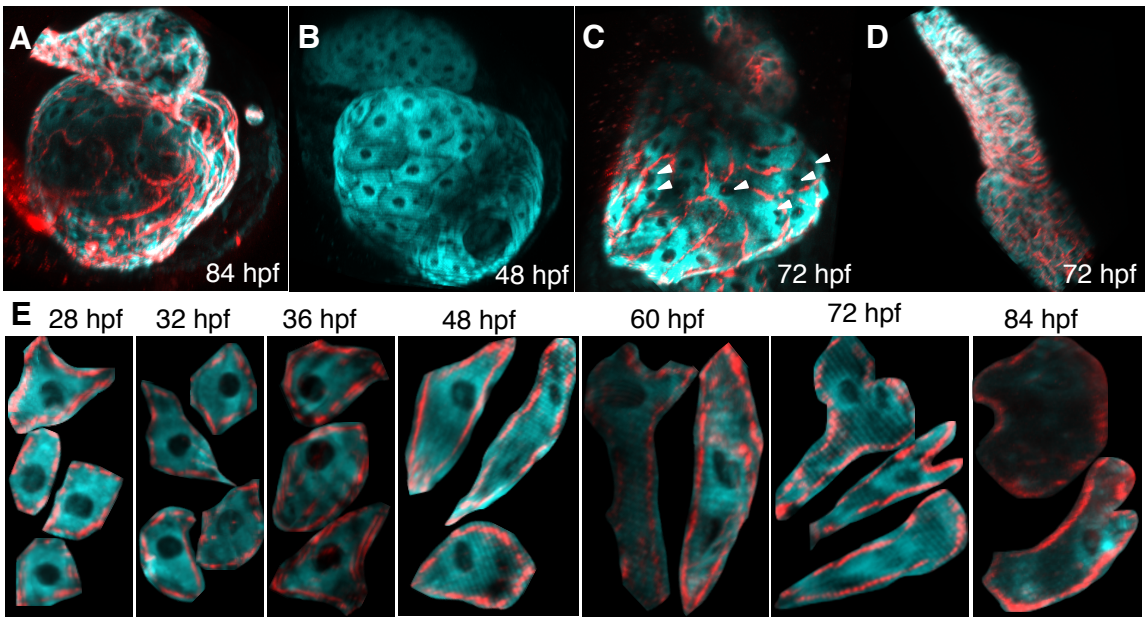


Figure E.4: **Cardiac mutants with contractile defects show similar heterogeneity in cell shape as wild type embryos**
(A-D)Representative high-resolution 3D rendered confocal images of developing heart in zebrafish embryos *in vivo* at different stages in homozygous embryo (A). anti-*amhc* morpholino injected embryos (B,C) and homozygous embryo with stunted cell growth(D). (E) Representative cells in *Gt(tpm4a-citrine)* $^{ct31a/ct31a}$ shown at different stages display similar heterogeneity in cell shape as in wild type embryos

Appendix F

Supplementary Information for Chapter 7

F.1 Zebrafish lines

Adult fish were raised and maintained as described before* in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals by the respective bodies at California Institute of Technology and University of Southern California. The protocol was approved by Institutional Animal Care and Use Committee (IACUC) at both universities. Wild-type embryos were obtained from AB strains. FlipTraps *Gt(tpm4a-citrine)* $^{ct31a/+}$ and *Gt(desm-mCherry)*$^{ct122aR/+}$ transgenic embryos were obtained from a previously described screen†. The wild-type transcript knockdown in embryos heterozygous for *Gt(desm-mCherry)*$^{ct122aR/+}$ is not sufficient enough for lethality and the embryos develop normally. Transgenic line *Tg(gata5:dsRed)* was obtained from Zebrafish International Resource Center (ZIRC, Eugene, OR), and line *Tg(kdrl:eGFP)* was provided by the Stainier lab. Upon crossing appropriate adult lines, the embryos obtained were raised in Egg Water (60 μg/ml of stock salts in distilled water) at 28.5°C with addition of 0.003% (w/v) 1-phenyl-2-thiourea (PTU) around 18hpf to reduce pigmentation.

F.2 Sample preparation for live imaging

The samples were immobilized in 0.85mm-inner-diameter glass capillaries (Wiretrol I, Drummond) with 1% UltraPure^{TM} Low Melting Point Agarose (Catalog No. 16520-050, Invitrogen^{TM}) solution prepared in 30% Danieau (17.4mM $NaCl$, 210μM KCl, 120μM $MgSO_4$?7H_2O, 180μM $Ca(NO_3)_2$, 0.1mM HEPES buffer in water, pH 7.6) with 0.003% PTU and 0.01% (100mg/L) Tricaine. The embryos were transferred to the agarose solution, and gently swirled to bring them into direct contact with the agarose, and then pulled into the capillary with a stainless steel plunger (VWR). Following solidification of agarose at room temperature (1-2 minutes), the capillary was transferred to the sample chamber filled with 30% Danieau solution and 0.01% Tricaine, at 24.0°C. After being mounted in the imaging chamber, the embryo embedded in agarose, was extruded out of the capillary and allowed

*Westerfield, M., The Zebrafish Book. Edition 2.1 ed. 1994, Eugene, OR: University Oregon Press

†Trinh LA, Hochgreb T, Graham M, Wu D, Ruf−Zamojski F, et al. (2011) A versatile gene trap to visualize and interrogate the function of the vertebrate proteome. Genes and development 25: 2306−2320

to equilibrate for 15 minutes before imaging. The extrusion allows full optical access for imaging without lensing effects/reflections from light passing through the glass capillary. Table F.1 provides the detailed description of the imaging parameters used for all images presented in the chapter.

F.3 Imaging parameters for presented data

	Fig. 7.2, 7.4, 7.5	Fig. 7.3A	Fig. 7.3B,C
Embryonic stage (hpf)	84	n/a	60
Camera used	iXon 885	Zyla	Zyla
Use of 0.63X adapter	Yes	No	No
Binning	1	1	2
Detection objective	20X (Zeiss)	25X (Nikon)	25X (Nikon)
Imaged volume (xyz) (voxels), typical	400 x 500 x150	100 x 100 x 100	400 x 500 x 170
Lateral resolution (μm)	0.635	0.26	0.52
Axial resolution (μm)	1	0.26	1
Excitation wavelength(nm)	920	920 for 2p 488 for 1p	920 for 2p 488 for 1p
Unidirectional laser power (mW)	125 for 2p	40 for 2p 0.5 for 1p	125 for 2p 2 for 1p
Frame rate (fps)	70	N/A	N/A
Exposure time (ms)	12.5	40	40
Excitation rejection filter	SP 750	SP 750	SP 750
Detection band pass filter	BP 525/50 (GFP-labeled samples) BP 535/50 (Citrine-labeled samples)	BP 525/50	BP 525/50
Number of frames imaged (for 4D reconstruction)	22500	N/A	N/A

Table F.1: Imaging parameters for data presented in Chapter 7

Appendix G

Supplementary Information for Chapter 8

G.1 Zebrafish lines and sample preparation

Same as described in sections F.1and F.2

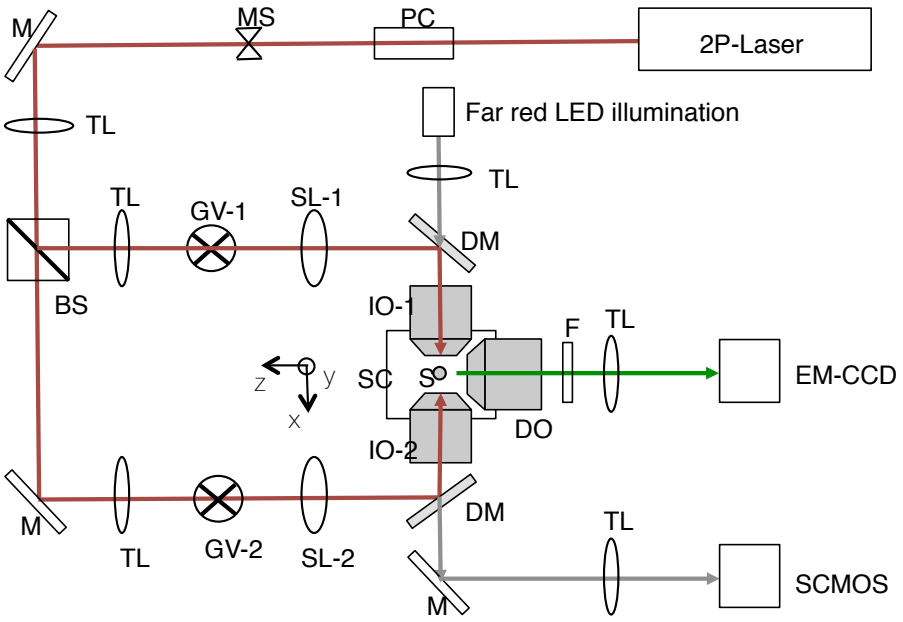


Figure G.1: **Schematic of the 2p-SPIM optical setup used in MaPS**

G.2 Image processing

Majority of the image processing was accomplished via custom written codes in MATLAB 2013b. The code is designed to infer the phase and position data from bright field (BF) images and assemble the information in $\phi - z$ matrix. Once this is achieved the information is relayed over to fluorescent (FL) images, which are synchronized to the BF images

during acquisition. The entire processing pipeline can be divided into following steps that correspond to different modules in our custom written code.

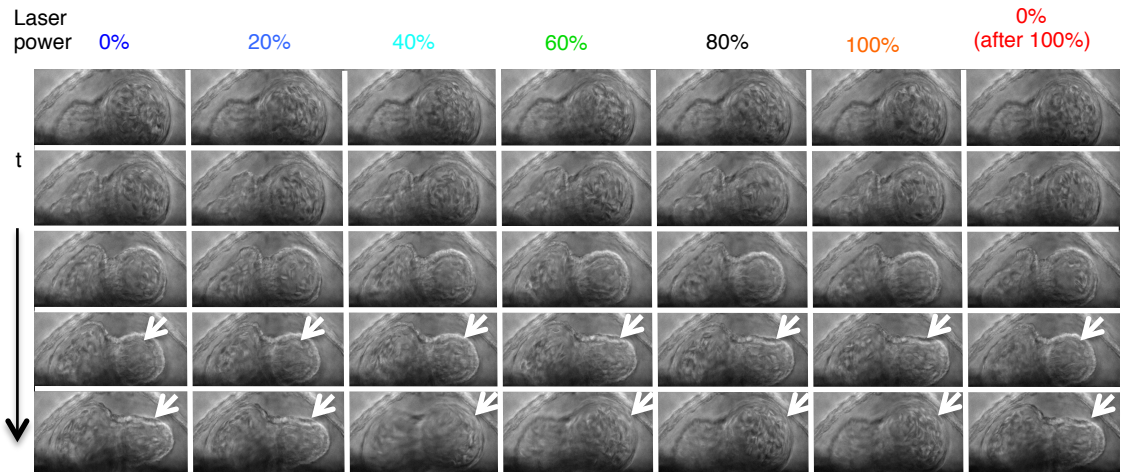


Figure G.2: **Effect of laser power on heart beat**
Bright-field images of beating zebrafish heart at 60hpf. Arrowheads denote differences in beating pattern when illuminated (via a sheet of laser light at a given z-section) at different laser powers. The differences are due to the localized heat generated at the illuminated z-plane

G.2.1 Pre-processing of files for input to the MATLAB code

The .sif files (for bright field (BF) images) from the EMCCD camera were converted into individual tiff files using the Batch Conversion option in SOLIS software. During the conversion it is ensured that the full dynamic range option (0-65525 for 16-bit images) is selected to obtain gray scale tiff images. The individual images were then stacked using ImageJ (US National Institutes of Health). The .tif files (for fluorescent (FL) images) from micro-manager software used for operating the sCMOS camera might be a collection of tif stacks, possibly due to a size limit on the file size in the operating system (Windows in our case). All these need to be compiled into one .tif file. Once this is done, field of views in both the BF and FL image files are cropped to contain only the relevant region and ensure that once the file is saved the size is $\leq$4GB as this is an upper limit on the size of a single file that MATLAB can read. This is currently the case as the program is written to take inputs from one single file for BF and FL images each.

G.2.2 Inferring z stage motion in pixels using kymograph

A static structure (pericardium) was identified in bright field (BF) images and a line drawn across it was used to generate kymograph (Orthogonal views in Stack, ImageJ) that was then binarized. The kymograph was used as an input to MATLAB code that extracted the distance moved by the stage (in pixels), by identifying the maxima's and minima's during the forward stage motion, thereby accomplishing z-stamping. Motion in the z-direction was estimated by counting the number of pixels moved between successive frames.

G.2.3 Stage Motion subtraction using a user-defined region of interest

A user-defined rectangular area is used as the region of image that contains beating heart such that it has the smallest possible frame size with heart motion clearly visible at all phases. This rectangle was treated as a sliding box that was moved between successive frames by the amount specified by the z-motion in pixels to account for the stage motion. Consequently the stage motion was subtracted to obtain the same region of image with beating heart for all frames. This step ensures that the subsequent image correlations done, to infer phases during heartbeat, are always based on motion of heart walls and not corrupted by the periodic stage motion.

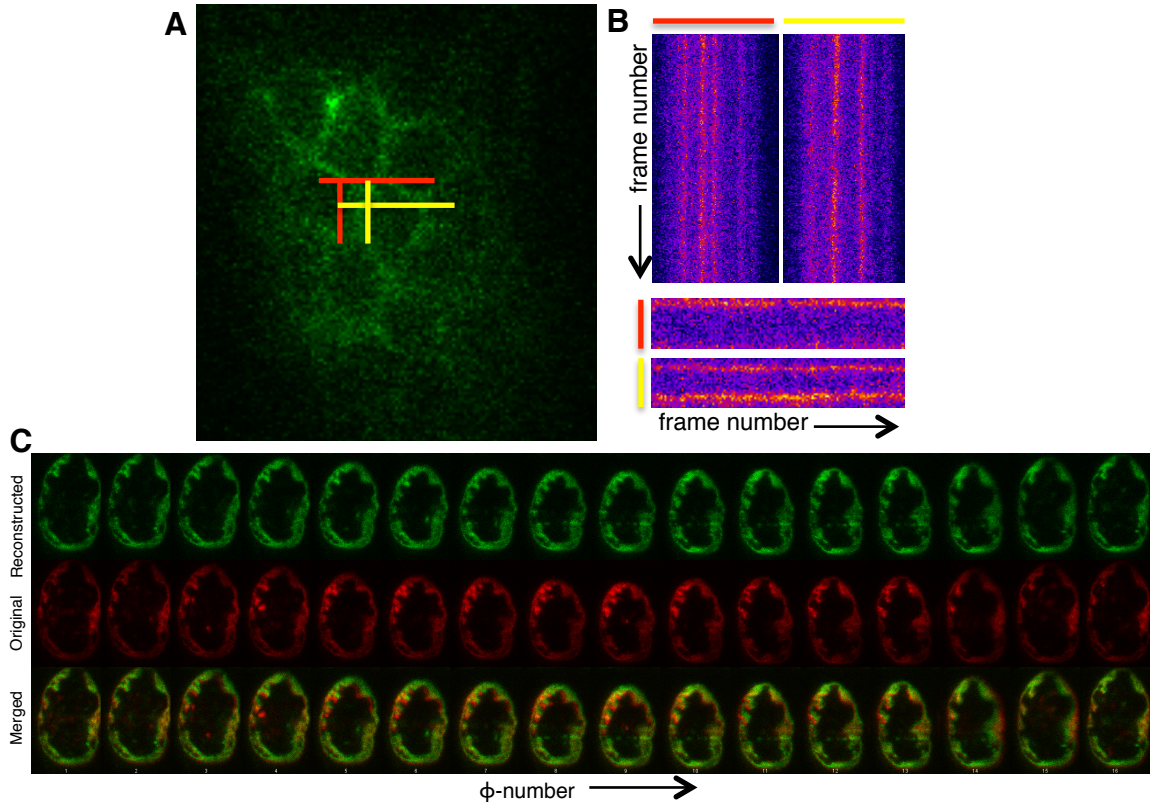


Figure G.3: **Performance of MaPS**
(A-B) Reconstruction of the static pericardial structure in beating heart. Kymographs (B) for the lines drawn (A) show that the reconstruction does not introduce motion artifacts. (C) Reconstructed movie (green) at a given z-section compared with the real time movie acquired by imaging the same plane (similar to retrospective approach). The overlap shows that reconstruction pipeline faithfully captures all phases of heart motion at a given frame rate.

G.2.4 z-stamping of image frames

The pixel size in images from sCMOS camera is $\sim 0.65 \mu m$ which is different from the z-stage motion of the stage (as the horizontal motion seen in BF images), defined to be $1 \mu m$

per frame. As a result it is crucial to define the movement in terms of pixels for stage motion subtraction In BF images. However the FL camera, acquiring images in orthogonal direction, is programmed to acquire 1 fluorescent image per um movement of the stage, thereby having a constsnt z resolution of $1\mu m$. Therefore the z-stamping is done taking into account the z-resolution in FL images where every frame, between (and including) the beginning and end of the forward stage motion, is assigned z-values in increments of 1 for successive frames.

G.2.5 Heartbeat frequency estimation

In order to do robust phase stamping, we first determine the number of BF frames that can represent a heartbeat. As the processing time scales exponentially with the number of frames analyzed, we analyze only the first 10% of image frames ($\sim$2000 frames) to determine the number of frames in a representative heartbeat. To start, the first frame among the stage-motion-corrected images is chosen to be the reference frame and all the first 10% image frames are correlated with it in order to obtain the oscillations of correlation values during heart beat. It is ensured that the jumps at the beginning and at the end of forward stage motion are taken into account in order to calculate these oscillations. This step is repeated by assigning every frame within these 10% images as the reference frame Fourier transforms of these oscillations are recorded and the number of frames corresponding to the dominant frequency in Fourier domain is chosen as the heartbeat frequency. This gives the number of frames, say N, needed to be present in a canonical heartbeat.

G.2.6 Canonical heartbeat determination

As the number of phases in a heart beat need to be defined based on BF images, a canonical heartbeat needs to be established first to deal with variations between heartbeats. To this end, the number of frames in a heartbeat, previously defined as N, is treated as N clusters which are populated with frames having the corresponding phase point. The canonical heartbeat is chosen to be the one that minimizes the spread between number of occurrences between clusters. This is achieved by correlating each of the first 10% image frames with a set of images having N continuous frames to calculate the phase value of the frame and thus populate the clusters. This process is then repeated for the next set of N frames shifted from the previous by 1 frame till all possible sets of N continuous images, in the first 10% frames, have been used, at least once, as reference set. Based on this the mean and standard deviation of the number of frames in a given cluster is calculated and the reference set that yields minimum dispersion index (defined as the ratio of standard deviation to mean) is chosen as the canonical heart beat. These N frames within the canonical heart beat are now treated as the canonical phase points

G.2.7 Phase stamping

Using the canonical image sequence every frame is assigned a phase number based upon the phase image with which it has the maximum correlation value. It is ensured that the threshold used for phase stamping is above 0.98.

G.2.8 phase-z matrix compilation

Upon successful z- and ϕ-stamping, the information is assembled into the $\phi - z$ matrix. If a position in the matrix has more than one entry, a new matrix was initialized calling it as the next 'brother'.

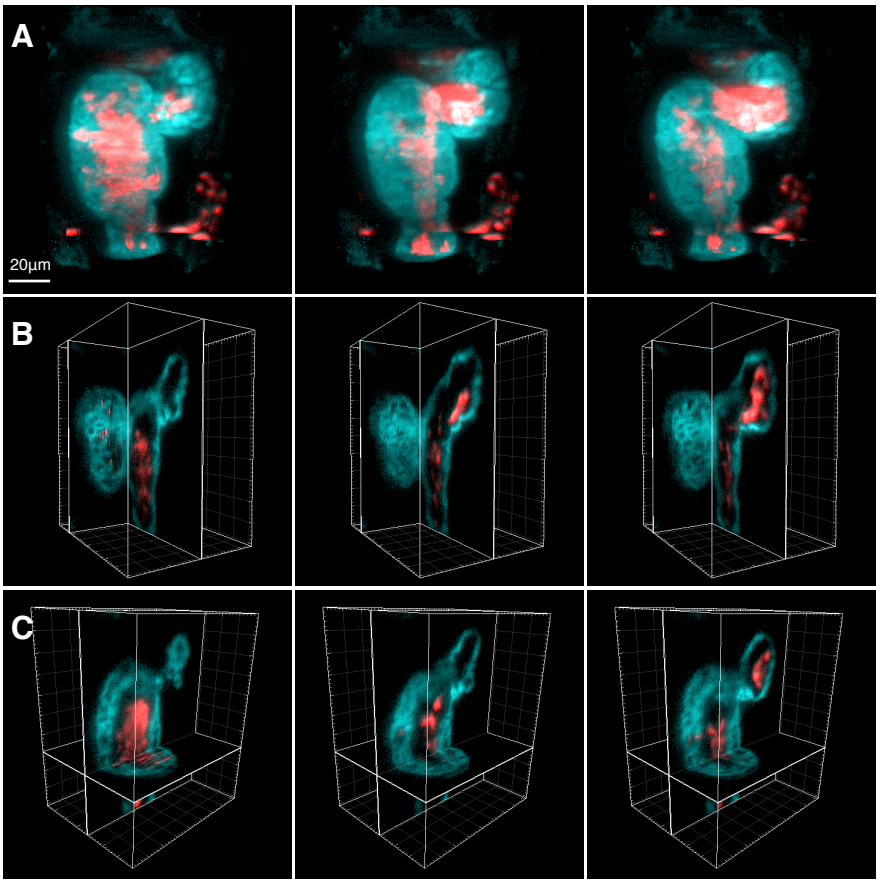


Figure G.4: **4D reconstruction of beating myocardium and blood cells in 5.5 days old zebrafish embryo**
(A-C) 3D rendered fluorescent images of the heart in *Gt(tpm4-citrine)*$^{ct31a/+}$*;Tg(gata5:dsRed)* at 5.5dpf at representative time points during its beating cycle. Cyan shows the myocardial wall and RBCs are depicted in red. Arrowheads denote trabeculating ventricle in different cross-sections(B-D).

G.2.9 Reconstruction

The information contained in the $\phi - z$ matrix of the first brother is then relayed over to the FL images and the corresponding frames are assembled to allow 3D reconstruction of beating heart. Images were rendered in 3D using IMARIS 7.2. Despite acquiring large number of frames and over population of certain coordinates, certain coordinates remain completely unfilled, being reflective of the cardiac beating signature, thereby giving missing slices in 3D reconstruction. Such occurrences are less than 5% and do not affect accuracy

3D reconstruction. However for the sake of better visualization of cross sectional images, we fill in the missing slices by taking an average of images from the surrounding coordinates in the phase-z matrix, invoking the argument of physical continuity of tissue between slices. It must be emphasized here that RBCs appear abnormal and discontinuous between slices. This is not because of error in reconstruction but the non-periodic nature of their motion, specially exaggerated when images at different z sections (but at the same phase) are compiled from different heartbeats. The position of blood cells must be understood as indicative of their statistical occurrence at that point in space. For better rendering smoothening the signal in blood channel with a Gaussian blur can be done.

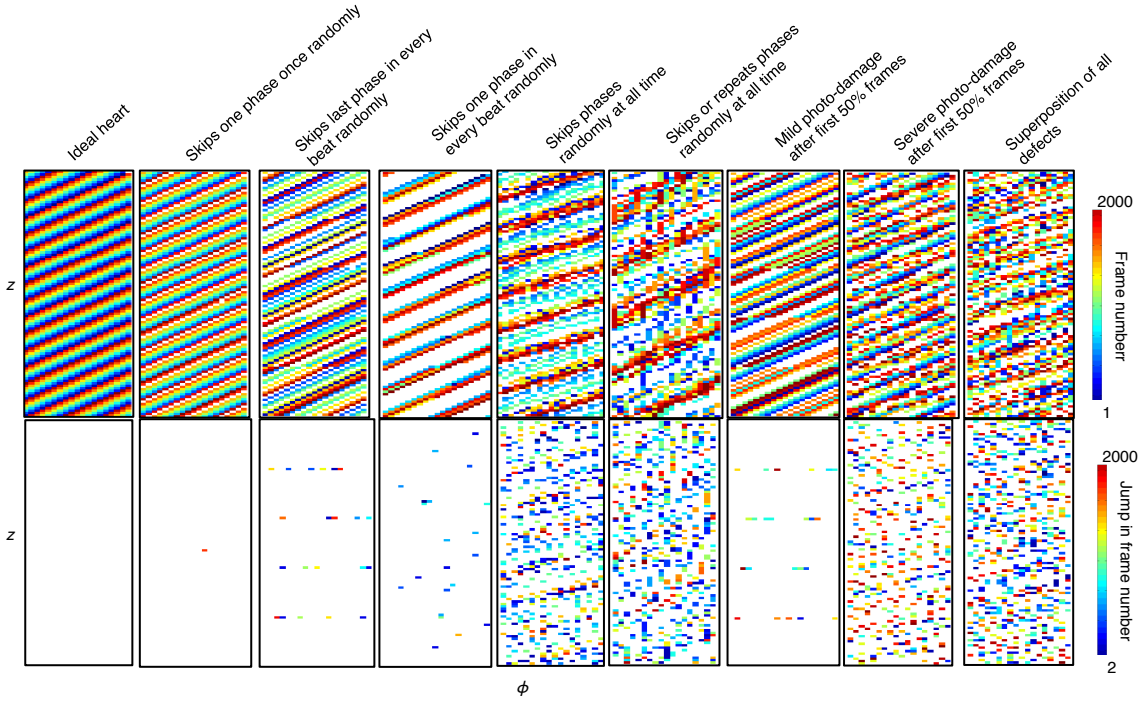


Figure G.5: **MaPS for assessing health of heart**
The stream of wide field images contains phase-rich information that can be utilized to assess the health of heart. In an ideal situation, the order of occurrence of phases of heart motion must be sequential with the frame number for the images assigned to a diagonal in the $\phi - z$ matrix. Any deviation (quantified as jumps between frame numbers) indicates deviation from wild type behavior. The figure shows the results on jumps as evaluated from simulated data.

G.3 Synchronization of the cameras

With the Bright Field and Fluorescent cameras completely synchronized to take images at the same frame rate, we have been able to achieve distinct phases for accurate reconstruction of beating hearts. Accurate z stamping is ensured by counting movement of stage in pixels as measured by the Bright Field camera. Fluorescent images at different optical sections consist of structures moving with different signatures. z-sections consisting of completely static structure i.e. pericardium, serve as a critical test for accuracy of z stamping (Fig.

G.3). If the z stamping is accurate, images of static structures with same z coordinate must overlap irrespective of their phase position, which indeed is the case as shown in Fig. G.3 (A,B). Kymographs for different x-y sections consist of straight lines denoting minimal stage motion, thereby also verifying the accuracy of piezo stage motion (Fig. G.3B). Accuracy of phase stamping has been assessed by comparing the sequence of images, compiled into phase-z matrix, from a given optical section (z-coordinate) against images acquired by parking the light sheet at the same position without any stage motion (Fig. G.3C). The signal from the two images are well co-localized. This must be considered in light of the fact that the difference comes largely from the background noise in the two image sets and the actual signal has nearly complete overlap. It can be argued that while pixel count for z stamping already ensures super resolution in z, time resolution is limited by the frame rate of the two cameras. We propose that this can also be improved by running the bright field camera (sCMOS) at a rate which is an integer multiple of the frame rate of fluorescent camera (EM-CCD). While this will increase the number of phase points, a stage of diminishing returns can be reached when minimal motion is observed between the frames and the random motion of blood cells leads to majority of correlation differences between phases. Our preliminary experiments (Data not shown), shows that with bright field camera operating at twice the frequency of fluorescent camera (80fps), we reach an upper limit in terms of identifying distinct phase points.

G.4 Imaging parameters for presented data

	Fig. 8.1B, 8.3, G.3C	Fig. 8.4, G.3A	Fig. G.4
Embryonic stage (hpf)	108	72,84,96,108	132
FL Camera used	iXon 885	iXon 885	iXon 885
BF Camera used	Zyla	Zyla	Zyla
Binning	2	2	2
Detection objective	25X (Nikon)	25X (Nikon)	25X (Nikon)
Imaged volume (xyz) (voxels), typical	800 x 500 x150	400 x 500 x150	800 x 500 x150
Lateral resolution (μm)	0.635	0.635	0.635
Axial resolution (μm)	1	1	1
Excitation wavelength(nm)	930	920	930
Unidirectional laser power (mW)	125	125	125
Frame rate (fps)	80	80	80
Exposure time (ms)	12.5	12.5	12.5
Excitation rejection filter	SP 750	SP 750	SP 750
Detection band pass filter	BP 535/50 (Citrine-labeled samples) BP 600/50 (dsRed-labeled samples)	BP 535/50	BP 535/50
Number of frames imaged (for 4D reconstruction)	20000	20000	20000

Table G.1: Imaging parameters for data presented in Chapter 8

G.5 MaPS for assessing health of heart

As mentioned in Chapter 8 , one of the unanticipated advantages of MaPS is the fact that the stream of wide field images contains phase-rich information that can be utilized to assess the health of heart. In an ideal situation , the order of occurrence of phases of heart motion must be sequential and periodic. The frame numbers of the image sequence must therefore be a sufficient predictor of the order of occurrence go phases. Viewed differently, the frame number for the images assigned to a diagonal in the $\phi - z$ matrix must therefore be sequential and monotonously increasing with a step size of 1. Any deviation (quantified as jumps between frame numbers) indicates that heart motion deviates from periodicity. We have simulated a variety of scenarios in Fig. G.5 to this effect. In case of wild type embryos, phase jumps between consecutive image frames is expected to be minimal. However within mutant/diseased hearts where the rhythmicity is affected, such jumps in phases are expected to be more frequent. Therefore using this approach MaPS can be used to comment on the health of heart in the input data set. Such an approach not only provides a way to characterize mutant/diseased hearts, while acquiring high resolution images of the same simultaneously but also a way to understand the effect of noise/variability between heartbeats even in the case of wild type animal.