Transcriptome and Metabolic Profiling of Premalignant Progression in Barrett's Esophagus

by

Jia Zeng

A Dissertation Presented in Partial Fulfillment
Of the Requirement of the Degree
Doctor of Philosophy

Approved April 2014 by the
Graduate Supervisory Committee:

Deirdre R. Meldrum, Chair
Laimonas Kelbauskas
Michael T. Barrett
Kimberly J. Bussey
Weiwen Zhang

ARIZONA STATE UNIVERSITY

May 2014

ABSTRACT

Cell-cell interactions in a microenvironment under stress conditions play a critical role in pathogenesis and pre-malignant progression. Hypoxia is a central factor in carcinogenesis, which induces selective pressure in this process. Understanding the role of intercellular communications and cellular adaptation to hypoxia can help discover new cancer biosignatures and more effective diagnostic and therapeutic strategies.

This dissertation presents a study on transcriptomic and metabolic profiling of pre-malignant progression of Barrett's esophagus. It encompasses two methodology developments and experimental findings of two related studies.

To integrate phenotype and genotype measurements, a minimally invasive method was developed for selectively retrieving single adherent cells from cell cultures. Selected single cells can be harvested by a combination of mechanical force and biochemical treatment after phenotype measurements and used for end-point assays. Furthermore, a method was developed for analyzing expression levels of ten genes in individual mammalian cells with high sensitivity and reproducibility without the need of pre-amplifying cDNA. It is inexpensive and compatible with most of commercially available RT-qPCR systems, which warrants a wide applicability of the method to gene expression analysis in single cells.

In the first study, the effect of intercellular interactions was investigated between normal esophageal epithelial and dysplastic Barrett's esophagus cells on gene expression levels and cellular functions. As a result, gene expression levels in dysplastic cells were found to be affected to a significantly larger extent than in the normal esophageal epithelial cells. These differentially expressed genes are enriched in cellular movement,

TGFβ and EGF signaling networks. Heterotypic interactions between normal and dysplastic cells can change cellular motility and inhibit proliferation in both normal and dysplastic cells. In the second study, alterations in gene transcription levels and metabolic phenotypes between hypoxia-adapted cells and age-matched normoxic controls representing four different stages of pre-malignant progression in Barrett's esophagus were investigated. Through differential gene expression analysis and mitochondrial membrane potential measurements, evidence of clonal evolution induced by hypoxia selection pressure in metaplastic and high-grade dysplastic cells was found. These discoveries on cell-cell interactions and hypoxia adaptations provide a deeper insight into the dynamic evolutionary process in pre-malignant progression of Barrett's esophagus.

To my parents, Suping Chen and Jiping Zeng

and my husband, Yunze Yang

ACKNOWLEDGEMENTS

This work would not have been possible without the input from so many supportive people. I do not have language adequate to express my gratitude towards the innumerable people who helped me to complete my doctoral training.

I would like to acknowledge my advisor, Dr. Deirdre R. Meldrum, for her wisdom and support. I am grateful for the invaluable opportunities of growing up and working in her laboratory, Center for Biosignatures Discoveries Automation in the Biodesign Institute at the Arizona State University.

I owe a tremendous debt of gratitude to my mentor, Dr. Laimonas Kelbauskas, for his unwavering support and rigorous coaching at all stages of my PhD. He was always generous and mentored me wholeheartedly throughout this fun and challenging journey, in which I grow as a scientist.

I would like to thank my committee members, Dr. Kimberly J. Bussey, Dr. Michael T. Barrett and Dr. Weiwen Zhang, for their kindness, expertise, constructive suggestions and alternative viewpoints that greatly improved this work.

I would like to acknowledge all of the current and past members in Center for Biosignatures Discovery Automation, specifically Dr. Shih-Hui Chao, who reviewed this work and gave valuable feedback, Dr. Roger H. Johnson, who provided extremely helpful insights into this work, Dr. Weimin Gao, who kindly taught me numerous biology techniques, knowledge and way of thinking. Other people that directly contributed to this work include Dr. Jiangxin Wang, Aida Mohammadreza, Kristen B. Lee, Saeed Merza, and Dean Smith. I had the pleasure of working alongside Dr. Jieying Wu, Dr. Xu Shi, Bo

Wang and Jordan Yaron. I appreciate their encouragement and discussions. Thanks to Dr. Yanqing Tian, Dr. Fengyu Su and Dr. Liqiang Zhang for their help throughout these years. Thanks to Patti Senechal-Willis, Nanna Hansen, Benjamin Ueberroth and Craig Reading for experimental support. Thanks to Clifford Anderson, Carol A, Glaub and Jeffrey M. Robinson for administrative support.

I am grateful for the guidance I received from my program course instructor, Dr. JoAnn C. Williams and coordinator, Maria Hanlin and Laura Hawes. Thanks to all my classmates in the Biological Design Graduate Program for their camaraderie.

I would also like to thank Tong Fu from Center for Infectious Disease and Vaccinology for assistance in flow cytometry, Center for Personalized Diagnostics for Sanger and Illumina sequencing services and providing facilities for lentiviral infections, DNA lab at Arizona State University for Sanger sequencing and Agilent Bioanalyzer services, and University of Arizona Genetics Core for Covaris shearing services.

Of course, none of this would have ever happened without the support from my loving grandparents and other family members in China and United States.

Finally, I would like to thank National Human Genome Research Institute (Grant 3P50HG002360 and 3P50HG002360-10S1 to D. R. Meldrum) for funding this work and Graduate and Professional Student Association at Arizona State University for supporting my research projects.

TABLE OF CONTENTS

Page

LIST OF TABLES xiii

LIST OF FIGURES xiv

CHAPTER

1 OBJECTIVES AND CONTRIBUTION 1

1.1 Significance and objectives 1

1.2 Scientific contributions 4

2 INTRODUCTION 7

2.1 Cancer 7

2.1.1 Hallmarks of cancer 7

2.1.2 Neoplasia 7

2.1.3 Microenvironment 8

2.1.4 Metabolic reprogramming 13

2.2 Barrett's esophagus 18

2.2.1 Overview 18

2.2.2 Hallmarks of pre-malignant progression in Barrett's esophagus 20

2.2.3 Models of Barrett's esophagus neoplastic progression 26

2.2.4 Genome and transcriptome study of Barrett's esophagus 27

2.3 Single cell transcription analysis 29

2.3.1 Single cell analysis 29

2.3.2 Methods for transcript profiling at the single-cell level 31

CHAPTER Page

3 A MINIMALLY INVASIVE METHOD FOR RETRIEVING SINGLE ADHERENT CELLS OF DIFFERENT TYPES FROM CULTURES 48

3.1 Abstract 48

3.2 Introduction 49

3.3 Experiments 53

3.3.1 Microwell design 53

3.3.2 Cell culture 53

3.3.3 Cell loading into microwells 54

3.3.4 Direct lysis of single cells in microwells 55

3.3.5 Single-cell collection 55

3.3.6 Preservation of RNA at low temperatures 56

3.3.7 Flow rate and trypsinization time 57

3.3.8 Primer design for the shear flow stress response gene *CCL2* 59

3.3.9 On-chip lysis of single cells in microwells 60

3.3.10 RNA isolation, reverse transcription, and qPCR 61

3.3.11 Lentiviral transfection of cells 63

3.3.12 Fluorescence-assisted single cell harvesting 64

3.4 Results and discussion 65

3.4.1 Comparison between in-situ direct lysis and “pick-and-place” harvesting of single cells in microwells 65

3.4.2 Usage of cellular RNA preservation solution for single-cell harvesting 67

3.4.3 Low temperature harvesting for preserving cellular RNA 70

CHAPTER Page

3.4.4 The effects of flow-rate and trypsinization time on harvesting success rate and RNA preservation 71

3.4.5 Comparison between on-chip lysis of cells in microwells and harvesting cells from microwells 75

3.4.6 Fluorescence-assisted single cell harvesting 78

3.5 Conclusions 82

3.6 Acknowledgments 83

4 QUANTITATIVE SINGLE-CELL GENE EXPRESSION MEASUREMENTS OF MULTIPLE GENES IN RESPONSE TO HYPOXIA TREATMENT 84

4.1 Abstract 84

4.2 Introduction 86

4.3 Experiments 91

4.3.1 Cell culture 91

4.3.2 Primer design and selection of gene target 91

4.3.3 Cell staining and fluorescence activated cell sorting 93

4.3.4 Hypoxia treatment 94

4.3.5 Single cell collection 94

4.3.6 RNA isolation and reverse transcription 95

4.3.7 qPCR 95

4.4 Results and discussion 98

4.4.1 Two-step RT-qPCR analysis of single mammalian cells 98

4.4.2 Detection of up to ten genes from a single mammalian cell 104

CHAPTER Page

4.4.3 Gene expression under hypoxia 107

4.5 Conclusion 115

4.6 Acknowledgments 116

5 WHOLE TRANSCRIPTOME AND METABOLIC PROFILING OF INTERCELLULAR INTERACTIONS BETWEEN NORMAL AND PRE-MALIGNANT ESOPHAGEAL CELLS 117

5.1 Abstract 117

5.2 Introduction 118

5.3 Experiments 121

5.3.1 Cell lines 121

5.3.2 Fluorescence assisted cell sorting of co-culture and mono-culture of normal and neoplastic cells 122

5.3.3 RNA extraction 123

5.3.4 Whole transcriptome amplification 123

5.3.5 Library preparation and Illumina sequencing 124

5.3.6 Next-generation sequencing alignment 125

5.3.7 Differential gene expression 125

5.3.8 Functional annotations and pathway enrichment 126

5.3.9 RT-qPCR validation 127

5.3.10 Time lapse fluorescent microscopy 128

5.3.11 Image analysis 129

5.3.12 Statistical analysis 129

CHAPTER Page

5.4 Results and discussion 130

5.4.1 RNA-Seq analysis of the transcriptome in esophageal epithelial normal and dysplastic cells 130

5.4.2 Function enrichment of differentially expressed genes in heterotypic interactions 135

5.4.3 Upstream regulator analysis reveals *TGFβ* and *EGF* signaling networks are inhibited in co-cultured CP-D cells 139

5.4.4 Co-culture of CP-D and EPC-2 cells changed the proliferation and motility of both cell lines 144

5.5 Conclusion 148

5.6 Acknowledgments 149

6 ALTERATIONS IN GENE EXPRESSION LEVELS AND METABOLIC PHENOTYPE IN RESPONSE TO HYPOXIC SELECTION IN PRE-MALIGNANT BARRETT'S ESOPHAGUS CELLS 150

6.1 Abstract 150

6.2 Introduction 152

6.3 Experiments 154

6.3.1 Cell lines 154

6.3.2 Acute hypoxia treatment 154

6.3.3 RNA extraction 155

6.3.4 Whole transcriptome amplification 155

6.3.5 Library preparation and Illumina sequencing 155

CHAPTER Page

6.3.6 Next-generation sequencing alignment 156

6.3.7 Multiple condition gene expression comparison 156

6.3.8 Differential gene expression 157

6.3.9 Functional annotations and pathway enrichment 157

6.3.10 Mitochondrial membrane potential measurement using flow cytometry . 157

6.3.11 Statistical analysis 158

6.4 Results and discussion 159

6.4.1 Mitochondrial membrane potential measurement of hypoxia-adapted and age-matched control cells 159

6.4.2 RNA-Seq of hypoxia-adapted and age-matched control cells 161

6.4.3 Comparisons among hypoxia-adapted and control Barrett's esophagus cell lines 162

6.4.4 Differential gene expression analysis 168

6.4.5 Function enrichment and upstream analysis of differentially expressed genes between hypoxia-adapted and age-matched control cell lines 171

6.5 Conclusion 177

6.6 Acknowledgments 179

7 CONCLUSION AND FUTURE WORK 180

7.1 Conclusion 180

7.2 Future work 183

REFERENCES 185

APPENDIX

CHAPTER Page

A PERMISSIONS 215

LIST OF TABLES

Table Page

Table 1 Voltage-to-flow rate conversion 58

Table 2 Success rate of single-cell harvesting 59

Table 3 Primer sequences used in the study 60

Table 4 Two-way multivariate ANOVA of the effects of flow rate and trypsinization time on gene transcription analysis 73

Table 5 The effects of flow rate and trypsinization time on the mRNA preservation/retrieval efficiency 74

Table 6 Genes and corresponding primers 92

Table 7 Descriptive statistics of mitochondrial membrane potentials in eight cell lines at the single cell level 161

LIST OF FIGURES

Figure .....Page

Figure 1-1 Overview of experimental process 3

Figure 3-1 Microwell array design 53

Figure 3-2 The single cell manipulation platform used in this study 56

Figure 3-3 Flow rate conversion 57

Figure 3-4 Schematic view of the single cell manipulation platform 65

Figure 3-5 Single-cell harvesting steps 66

Figure 3-6 Comparison of RNA preservation levels between medium, RNALater and RNA Lysis Buffer 69

Figure 3-7 Comparison between low temperature and room temperature harvesting conditions for preserving cellular RNA 70

Figure 3-8 Comparison of gene transcription levels between the on-chip lysis of cells in microwells and harvesting cells from microwells 77

Figure 3-9 Harvesting single cells from a co-culture 79

Figure 3-10 RT-qPCR analysis of TurboGFP genes in CPD/TurboGFP and EPC2/FP635 cells 81

Figure 4-1 Amplification plots of gene transcripts using validated primers 101

Figure 4-2 Melt curves of gene transcripts using validated primers 103

Figure 4-3 Quantification of expression levels of ten genes in a single, hypoxia-treated CP-A cell 105

Figure Page

Figure 4-4 Gel electrophoresis analysis of RT-qPCR amplification products of 10 genes from a single CP-A cell 106

Figure 4-5 Oxygen depletion in the cell medium 108

Figure 4-6 qPCR results of long time Oxyrase treatment 109

Figure 4-7 Single-cell gene-expression profiling in control and hypoxia-treated CP-A cells 111

Figure 4-8 Gene expression levels of bulk cell samples under normoxic and hypoxic conditions 113

Figure 5-1 Workflow of transcriptome analysis of cell-cell interactions 130

Figure 5-2 Number of differentially expressed genes 132

Figure 5-3 Venn diagrams of differentially expressed genes 134

Figure 5-4 Top 5 functions enriched in the co-CPD vs. mono-CPD group 135

Figure 5-5 Functional annotations in the co-CPD vs. mono-CPD group 137

Figure 5-6 Top 5 functions enriched in the mono-CPD vs. mono-EPC2 group 138

Figure 5-7 Functional annotations in the mono-CPD vs. mono-EPC2 group 139

Figure 5-8 Upstream regulator analysis in the co-CPD vs. mono-CPD group 141

Figure 5-9 Differentially expressed genes in the networks 143

Figure 5-10 Fluorescence microscopy of Hoechst 33342 stained cells 145

Figure 5-11 Proliferation of CP-D and EPC-2 cells in mono-culture and co-culture 146

Figure 5-12 Migration of CP-D and EPC-2 cells in mono-culture and co-culture 147

Figure 6-1 Violin plot of mitochondrial membrane potentials 160

Figure 6-2 Sample cluster of differentially expressed genes 163

Figure Page

Figure 6-3 Gene ontology analysis of statistically significant genes 164

Figure 6-4 Hierarchical clustering of statistically significant genes in the glycolysis/gluconeogenesis pathway 165

Figure 6-5 Statistically significant genes in the glycolysis/gluconeogenesis pathway... 167

Figure 6-6 Hierarchical clustering of statistically significant genes in the oxidative phosphorylation pathway 168

Figure 6-7 Number of statistically significant genes found by three differential gene expression tests 169

Figure 6-8 Venn diagrams of differentially expressed genes identified by three statistical methods 170

Figure 6-9 Vascular system related functions enriched in differentially expressed genes in the HCPC vs. CCPC group 172

Figure 6-10 Other functions enriched in differentially expressed genes in the HCPC vs. CCPC group 173

Figure 6-11 Upstream regulator analysis of differentially expressed genes (part 1) 174

Figure 6-12 Upstream regulator analysis of differentially expressed genes (part 2) 175

Figure 6-13 TGF-β mediated hypoxia response is suppressed in HCPC cells 176

Figure 6-14 The proposed evolutionary process of Barrett's esophagus cells adapted to the hypoxic selection pressure 178

CHAPTER 1

OBJECTIVES AND CONTRIBUTION

1.1 Significance and objectives

Cell-cell interactions and metabolism are essential for growth and function of multicellular organisms. Aberrant intercellular communication plays a key role in carcinogenesis and tumor progression (Hanahan and Weinberg 2011, Hanahan and Coussens 2012). Metabolic reprogramming is one of the emerging hallmarks of cancer (Hanahan and Weinberg 2011). Understanding the role of intercellular communications and hypoxia-adapted metabolic reprogramming in pre-malignant progression could aid in early diagnosis and help discover more effective prognostic, diagnostic and management strategies of cancer.

Barrett's esophagus is a precancerous condition which predisposes to esophageal adenocarcinoma, a cancer type with high mortality rates (Shaheen and Ransohoff 2002). Barrett's esophagus is one of the best characterized models for studying pre-malignant progression due to its relative easy access to biopsies that can be taken at multiple time points from the same patient, following current standard-of-care procedures (Barrett, Yeung et al. 2002). Two important factors in the clonal evolution of Barrett's esophagus neoplastic progression are intercellular communications and hypoxia-selection pressure. Gene expression profiling (Helm, Enkemann et al. 2005) and metabolic phenotype measurements (Suchorolski, Paulson et al. 2013) indicate that Barrett's esophagus cells undergo a series of transcriptional and metabolic changes in the context of cell-cell interactions or hypoxia. However, the mechanisms underlying these changes remain little known. In this thesis, one hypothesis is that intercellular communications between

normal and dysplastic cells affects cellular proliferation and motility. Barrett's esophagus cells can reprogram their metabolisms to adapt to hypoxia selection pressure. Understanding the mechanisms and molecular signatures of cell-cell interactions and hypoxia adaptation in Barrett's esophagus will open new ways for designing diagnosis and treatment strategies for Barrett's esophagus and adenocarcinoma. The biosignatures discovered in this study can be used for early diagnosis, treatment and risk stratification for Barrett's esophagus. As cell-cell interactions and hypoxia -adaptations widely exist in other cancers, the pre-malignant evolutionary process characterized in this study will be transferrable to other cancer types and can help advance the field of cancer research.

The goal of this doctoral dissertation project is to study gene transcription and metabolic profiles of pre-malignant progression in Barrett's esophagus. The objectives of the project are:

i) Conceive, develop and optimize a minimally invasive method for retrieving single adherent cells of different types from cultures. This method will be used for harvesting single adherent cells from a glass substrate after metabolic phenotype measurement.

ii) Develop and optimize a reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) method for measuring expressions that does not require a pre-amplification step of more than five genes within a single cell. This single cell RT-qPCR method will detect heterogeneity in gene expression levels in a cell population and provide information on hypoxia response.

iii) Characterize differential gene expression profiles as a result of heterotypic cell-cell interactions in Barrett's esophagus using whole transcriptome RNA-Seq.

Identify functional relevance of gene expression changes and correlate with cellular functional phenotypes. This will allow unprecedented inquiries into pre-malignant progression in Barrett's esophagus.

iv) Study whole-transcriptome profiles and mitochondrial functions in pre-malignant Barrett's esophagus cell lines adapted to hypoxia and compare them with age-matched normoxic control cells. The result will lead to a better understanding of the pathways that regulate cell metabolism in pre-malignant stages, and potentially warrant deeper insight into cancer cell development and progression.

An overall study schematic goal is depicted in Figure 1-1. Single cell harvesting and RT-qPCR methods will be developed and optimized for implementing a multi-

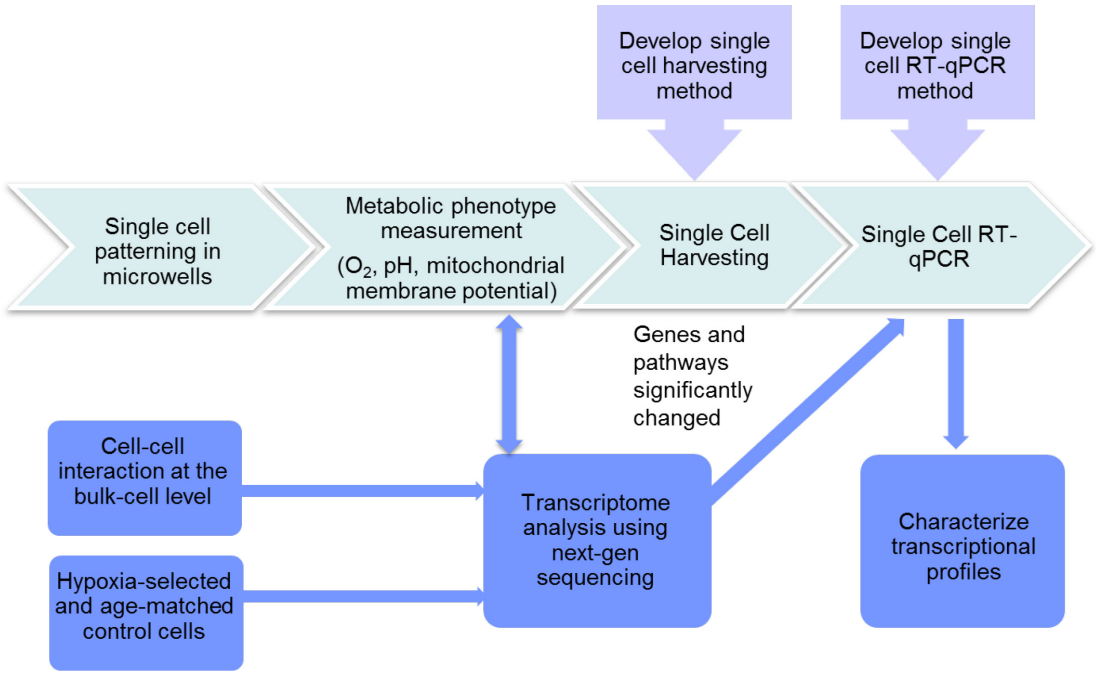


Figure 1-1 Overview of experimental process.

parameter single-cell analysis pipeline. Whole transcriptome analysis of cell-cell interactions and hypoxia responses in Barrett's esophagus cells will identify differentially expressed genes in the neoplastic progression. The findings will provide a set of gene candidates for gene expression profiling using RT-qPCR and be correlated with functional level alterations in this process.

1.2 Scientific contributions

Based on the results of this doctoral dissertation project, all objectives have been successfully achieved. The scientific contributions of the research are the following:

(i) A new method was developed for retrieving individual adherent cells with minimal perturbation using a combination of mechanical forces and biochemical treatment. A method to selectively harvest individual cells of different types from microwells with co-cultured cells was demonstrated using a fluorescence-assisted single-cell harvesting platform. Stress gene responses to varying levels of mechanical forces and biochemical treatment were measured. The findings of this part of work are useful for studies focused on single-cell analysis that involve any mechanical manipulation of live cells.

(ii) A SYBR green-based RT-qPCR method was developed for detecting expression levels of up to ten different genes without pre-amplification of cDNA. Gene expression heterogeneity can be detected in single cells with good reproducibility and specificity. This method is inexpensive and compatible with most commercially available RT-qPCR instrumentation. It can be easily integrated in many applications focused on gene expression analysis in single cells, which will provide further insights into the cellular mechanisms involved in physiological and pathological processes at the single-

cell level and has the potential of becoming a tool for future point-of-care medical applications.

(iii) In the transcriptome-wide study of effects of cell-cell interactions in pre-malignant progression of Barrett's esophagus a set of genes was found whose expression is down-regulated in dysplastic cells when they are co-cultured with normal esophageal epithelial cells. Upstream regulators, TGFβ and EGF, that act as "first responders" of cell-cell interactions are also identified. Normal cells are found to inhibit the growth of dysplastic cells which is mediated by growth factor signaling pathways. This study indicated that the fraction of normal to dysplastic cells can be used as risk stratification markers for Barrett's esophagus and esophageal adenocarcinoma. *TGFβ*, *EGF* and their downstream genes have great potential to become biosignatures for esophageal adenocarcinoma diagnosis and disease management.

(iv) Whole transcriptome analysis and metabolic phenotypic measurements of hypoxia response in Barrett's esophagus cells representing different stages of pre-malignant progression revealed a series of transcriptional alterations and changes in mitochondrial function. It was found that the mitochondrial functions as measured by mitochondrial membrane potentials are suppressed in hypoxia-selected Barrett's esophagus cells as compared to their age-matched normoxic controls. Hypoxia-adapted cells reprogrammed their metabolism and retained functional plasticity to survive and thrive under hypoxia stress. Adaptation to hypoxia can be used as a risk prediction marker for progression towards esophageal adenocarcinoma. Functional plasticity in hypoxia-adapted cells also suggested that physicians need to design multiple targets such

as oxidative phosphorylation and glycolysis to treat pre-malignant conditions in Barrett's esophagus.

With these results, one paper was published, one manuscript was submitted and two more manuscripts are in preparation for publication. All of them are listed below.

1. Zeng, J.*, Wang, J.*, Gao, W., Mohammadreza, A., Kelbauskas, L., Zhang, W., Johnson, R. H., and Meldrum, D. R. (2011). Quantitative single-cell gene expression measurements of multiple genes in response to hypoxia treatment. Analytical and Bioanalytical Chemistry 401, 3-13. (*Co-first authorship) PMID:21614642. *Published in a special, accelerated section, "Paper in Forefront."*

2. Zeng, J., Mohammadreza, A., Gao, W., Merza, S., Smith D., Kelbauskas, L., and Meldrum, D. R. A minimally invasive method for retrieving single adherent cells of different types from cultures. *Under review.*

3. Whole transcriptome and metabolic profiling of intercellular interactions between normal and pre-malignant esophageal cells. *In preparation.*

4. Alterations in gene expression levels and metabolic phenotype in response to hypoxic selection in pre-malignant Barrett's esophagus cells. *In preparation.*

CHAPTER 2

INTRODUCTION

2.1 Cancer

2.1.1 Hallmarks of cancer

Cancer is a major cause of death in the United States and many other parts of the world (Siegel, Ma et al. 2014). Although a great amount of effort has been devoted to cancer research and clinical practice, we still cannot eradicate or control the advanced disease of cancer. The complexity and diversity of cancer lay both roadblocks to, and stepping stones for successful treatment (Greaves and Maley 2012).

In 2000 and 2011, Hanahan and Weinberg proposed and updated cancer hallmarks: sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, activating invasion and metastasis, reprogramming of energy metabolism and evading immune destruction (Hanahan and Weinberg 2000, Hanahan and Weinberg 2011).Genome instability and mutation and tumor-promoting inflammation are two characteristics enabling cancer cells to acquire hallmark functional capabilities (Hanahan and Weinberg 2011). Tumor microenvironment – the signaling interactions between cancer cells and their supporting normal cells—also contribute to the acquisition of hallmark traits, further diversifying the mechanisms of cancer pathogenesis (Hanahan and Weinberg 2011).

2.1.2 Neoplasia

Neoplasia is unregulated growth of cells. Neoplastic cells can proliferate without the influence of any external stimuli while ignoring growth-controlling signals. Neoplasia usually results in an abnormal mass of tissue, known as neoplasm.

Before cells become neoplastic, they often first become metaplastic or dysplastic. Metaplasia is a transformation of one differentiated cell type into another. One typical example is Barrett's esophagus, in which the esophageal squamous epithelium transforms into intestinal-like columnar cells. Dysplasia is the fractional increase of immature cells compared to matured and differentiated cells in a sample. It is mainly due to genetic alterations, changes in gene expression, and dysregulation of cell maturation. Metaplastic or dysplastic cells may progress to cancer or regress to normal cells (Schlecht, Platt et al. 2003).

Neoplasms are usually caused by genetic mutations. They can be benign, pre-malignant or malignant (cancer). Recognized as an evolutionary process (Nowell 1976), neoplastic progression is characterized by genomic instability and clonal expansion. Neoplastic cells accumulate genetic and epigenetic alterations--which contribute to clonal heterogeneity--and undergo evolution by natural selection. Acquiring the hallmark traits of cancer provides an evolutionary advantage to neoplastic cells (Hanahan and Weinberg 2000). Based on the scope of this dissertation work, microenvironment and metabolic reprogramming in cancer are the primary areas of discussion.

2.1.3 Microenvironment

2.1.3.1 Intercellular interactions within the tumor microenvironment

Over the past several decades, the prevailing view towards neoplastic progression and carcinogenesis is that cancer cells act autonomously in isolation. Yet, more and more evidence has shown that a tumor is not merely a collection of homogenous cancer cells transforming autonomously. Its genesis and progression is rather an ecological process involving a dynamic interplay between malignant and non-malignant cells (Barcellos-

Hoff, Lyden et al. 2013). Tumors are recognized as complex organs consisting of various types of cells: cells at different stages of tumor progression, cancer-associated fibroblasts, angiogenic vascular cells and infiltrating immune cells (Hanahan and Weinberg 2011). Tumor cells are affected by reciprocal interaction between the parenchymal and stromal cells either through direct contact or through signaling molecules. Therefore, the tumor microenvironment is an integral part of cancer initiation, growth and progression. Intercellular communications between tumor cells and their microenvironment create a context that promotes tumor growth and evasion from immune attack. This new perspective will reveal more underlying organizing principles of tumorigenesis and progression. The functions of the tumor microenvironment that contribute to sustained growth and survival of neoplastic cells during neoplastic progression is illustrated here.

2.1.3.2 Sustaining proliferative signaling

An indispensable feature of the tumor microenvironment is the ability to support and promote the proliferation of cancer cells (Hanahan and Weinberg 2011, Hanahan and Coussens 2012). Normal epithelial cells are connected with each other to form cellular apical and basal surfaces and maintain the differentiated state. The basement membrane, a specialized form of extracellular matrix, provides both structural support and a variety of polarization cues to the epithelium. Genetic alterations can cause the loss of polarity in dysplastic or metaplastic cells (Alison, Hunt et al. 2002). Aberrant interactions between epithelial cells and the basement membrane can support genomic instability within the epithelium. Alterations in communications between tumor cells and their environment have been proposed to account for increased proliferation (Hanahan and Coussens 2012).

2.1.3.3 Evading Growth Suppressors

Tumor cells circumvent growth suppression in various ways, by inactivating or down-regulating tumor suppressor genes, or by evading cell-to-cell contact inhibition. Physical contacts among dense populations of normal cells can suppress further cell proliferation. This contact inhibition is an intrinsic mechanism for maintaining normal tissue homeostasis (Gatenby and Gillies 2008). Tumor cells, which carry mutations enabling neoplastic growth, can be suppressed when introduced into a context of normal connective tissue fibroblasts. The malignant phenotype of epithelial tumor cells could be reversed by the normal tissue microenvironment; normal fibroblasts hence act as an extrinsic epithelial growth suppressor at the early stage of cancer (Bissell and Hines 2011, Flaberg, Markasz et al. 2011). When the fibroblasts convert from normal tissue to cancer associated fibroblasts, the contact inhibition is abrogated (Bissell and Hines 2011, Flaberg, Markasz et al. 2011). The conversion might be driven by reprogramming of fibroblasts by dysplastic or metaplastic cells, or extrinsic conditions such as infection or fibrosis. The interplays between aberrant epithelial cells and their supporting fibroblasts might eventually relieve the inhibition of epithelial cell growth. As a result, neoplastic development takes place (Hanahan and Coussens 2012).

2.1.3.4 Resisting cell death

One barrier to tumorigenesis is programmed cell death, or apoptosis, which can be induced by various physiological stresses during tumorigenesis or anticancer therapy. Two machineries can trigger apoptosis: the extrinsic program, by which cells detect, receive and process extracellular death signals, and the intrinsic program, by which cells sense intracellular signals and initiate the apoptosis pathway.

When normal cells surrounding neoplastic cells sense the "invasion" or presence of foreign cell types, they will convey apoptotic signals to the neoplastic cells. Neoplastic cells have evolved diverse extrinsic and intrinsic strategies to avoid apoptosis and survive. In the microenvironment, cancer-associated fibroblasts synthesize molecules to form a neoplastic extracellular matrix for tumor cell survival (Lu, Weaver et al. 2012). Traditionally, normal immune cells combat pathogens and clean up the apoptotic cells. On the contrary, tumor-associated macrophages can bind to tumor cells, activate *PI3K/AKT* signaling, and suppress apoptosis (Chen, Zhang et al. 2011). Tumor-associated macrophages also guard tumor cells against chemotherapy-induced cell death in breast cancer (Shree, Olson et al. 2011). The interactions between cancer-associated fibroblasts and neoplastic cells facilitate evasion of apoptosis in the latter thus helping them evade the surveillance processes taking place in normal tissue. The resulting selective pressure for apoptosis-resistant tumor cells can lead to a significantly lower efficacy of cytotoxic and targeted therapy. Reprogramming energy metabolism

Metabolic reprogramming in tumor cells was proposed more than eighty years ago (Warburg, Wind et al. 1927) and is recognized as an emerging hallmark of cancer (Hanahan and Weinberg 2011). Neoplastic cells need to adjust their energy metabolism to support rapid growth and proliferation. They can use a variety of fuel sources to generate energy, synthesize biomaterials such as nucleotides, amino acids, and organelles for assembling new cells. The most prominent change is that cancer cells mainly rely on glycolysis rather than oxidative phosphorylation for their energy even in the presence of oxygen (Warburg, Wind et al. 1927). Metabolic reprogramming has been considered as the consequence of intrinsic mechanisms. However, more and more evidence is

substantiating the roles of reciprocal interactions between tumor cells and tumor microenvironment in metabolic reprogramming (Rattigan, Patel et al. 2012, Sotgia, Martinez-Outschoorn et al. 2012).

Cancer-associated fibroblasts can switch to glycolytic pathway once they are induced by reactive oxygen species released by breast cancer cells (Rattigan, Patel et al. 2012, Sotgia, Martinez-Outschoorn et al. 2012). Tumor cells can then uptake the lactate and pyruvate secreted by cancer-associated fibroblasts, to match their proliferation needs (Rattigan, Patel et al. 2012, Sotgia, Martinez-Outschoorn et al. 2012). In other cases, tumor cells utilize glucose and export lactate. Cancer-associated fibroblasts consume lactate, which acts as a paracrine modulator in driving tumor progression. Lactate shuffling suggests a symbiotic relationship between tumor cells and tumor microenvironment (Rattigan, Patel et al. 2012), which promotes the tumor growth.

2.1.3.6 Summary

The contributions from the tumor microenvironment to neoplastic progression have recently attracted substantial attention in the field. Understanding how multiple types of cells are co-opted to support different stages of carcinogenesis and progression is fundamental in studying the cross-talk between tumors and their microenvironment. One of the central challenges in tumor microenvironment studies is the delineation of intercellular communication signaling networks in greater detail and clarity. Cells at different stages of progression may activate different molecular pathways to recruit, adapt to and interact with their microenvironment. Subpopulations of cells might become more malignant or resistant in the shelter of their microenvironment. New discoveries in the tumor environment field will help physicians design innovative multi-target strategies for

both primary tumors and metastatic stages and revolutionize the treatment and management of cancer.

2.1.4 Metabolic reprogramming

One of the emerging hallmarks of cancer is metabolic transformation, which supports cell growth and proliferation (Hanahan and Weinberg 2011). In 1927, Otto Warburg discovered that, in contrast to normal cells, cancer cells under normoxic conditions rely primarily on glycolysis instead of oxidative phosphorylation to support their energy needs (Warburg, Wind et al. 1927). This metabolic phenomenon is known as the "Warburg effect". The Warburg effect has been both supported and refuted during the past century. In normal cells, the primary role of mitochondria is to produce energy in the form of ATP through oxidative phosphorylation. The hypothesis that tumor mitochondria are damaged was challenged because tumor mitochondria do respire and produce ATP through oxidative phosphorylation (Weinhouse 1976). In certain cancer types, the reprogramming of energy metabolism does not correlate with mitochondrial defects (Gogvadze, Orrenius et al. 2008), although some aggressive tumors do display mitochondrial deterioration (Gogvadze, Zhivotovsky et al. 2010).

2.1.4.1 Hypoxia

Hypoxia is a condition when the oxygen level in tissue is reduced to lower than 80-100 Hg (Barer, Howard et al. 1970). It is implicated as a factor in a variety of diseases, including cancer. Hypoxic conditions are known to arise in intermediate stages of tumor growth. As tumor cells proliferate rapidly and massively, the distance between cells and the vasculature increases, which eventually creates a local microenvironment deficient in oxygen supply. Such hypoxic conditions become a selection pressure for tumor growth.

Tumor cells undergo genetic, transcriptional or metabolic level reprogramming in order to survive and proliferate under hypoxia (Harris 2002). Therefore, hypoxia is closely related to malignant phenotype and aggressive tumor behavior.

Even though hypoxia is recognized as instrumental in tumor progression, its role in earlier premalignant progression is poorly understood. Hypoxia can also occur in chronic inflammation, a risk factor for cancer, in several forms such as ulceration, scarring or burns. Acute tissue oxygen level changes between hypoxia and reoxygeneration can generate reactive oxygen species. Deep ulceration in Barrett's esophagus creates a periodic hypoxic environment for the esophageal epithelial cells (Suchorolski, Paulson et al. 2013). Barrett's esophagus develops as a protective mechanism to the acid-bile reflex, which can result in hypoxic conditions. It is poorly understood how cells alter their genotypes and phenotypes under selective pressure of hypoxia during premalignant progression.

In eukaryotic cells, the hypoxia-inducible factor (HIF) is a key molecular mediator under hypoxic conditions. HIF is a heterodimeric transcription factor, whose transcriptional activation function is inhibited under normoxic conditions. The inhibition is exerted via post-translational hydroxylation by oxygen-dependent oxygenases, prolyl hydroxylase domain proteins (PHD) (Masson, Willam et al. 2001) and factor inhibiting HIF (FIH) (Semenza 2002). Under hypoxia, HIF1-α translocates into the nucleus and interacts with co-activators p300/CBP to regulate a broad range of genes participating in adaptation to hypoxia (Semenza 2002). These proteins are involved in angiogenesis, metabolism, cell proliferation, apoptosis, immortalization and migration. The interplay between oncogenic pathways and hypoxia responses are discussed below.

2.1.4.2 Angiogenesis

HIF1-α triggers the expression of genes involved in blood-vessel formation, such as vascular endothelial growth factor-A (*VEGF-A*) and angiopoietin (*Ang-2*) (Oh, Takagi et al. 1999). VEGF can be secreted by cancer cells and normal cells in response to hypoxia. Consequently, new vasculatures can be formed to provide oxygen and nutrition to the cells. Nonetheless, the newly generated vessels are usually distorted, irregular, and leaky, which results in oxygen supply deficiencies.

2.1.4.3 Metabolism

Hypoxic cells usually switch from mitochondrial oxidative phosphorylation to glycolysis to meet their energy needs. This is promoted by HIF-mediated expression of both glucose transporters and enzymes in the glycolytic pathway. Glucose transporters GLUT1 and GLUT3 facilitate cellular glucose uptake (Vannucci, Seaman et al. 1996). Two enzymes in the pyruvate metabolism pathway are known to be mediated by HIF1-α. Lactate dehydrogenase A (LDH-A) is responsible for converting pyruvate to lactate. Pyruvate dehydrogenase kinase 1 (PDK1) is enzyme that phosphorylates and inactivates pyruvate dehydrogenase (PDH), thereby feeding pyruvate into the glycolysis rather than oxidative phosphorylation (Kim, Tchernyshyov et al. 2006).

The mammalian target of rapamycin (mTOR) pathway is inhibited by HIF downstream of O_2-sensing. Under normoxic conditions, diverse signals converge to the mTOR kinase and the latter transmits the signal to regulate cell survival and growth through mRNA translation, ribosomal biogenesis and metabolism. Hypoxia can inhibit mTOR pathway and thereby control protein synthesis, energy metabolism and cell survival under the selection pressure (Brugarolas, Lei et al. 2004).

2.1.4.4 Proliferation

Hypoxia affects cell proliferation by inducing the production of growth factors, such as transforming growth factor-β and platelet-derived growth factor (Lal, Peters et al. 2001). Some factors upstream of *HIF-1α* expression or function also regulate cell proliferation. Proteins in oncogene signaling pathways, such as p42/p44 mitogen-activated protein kinases (Berra 2000) and phosphatidylinositol 3-OH kinase (*PI3K*) (Zundel, Schindler et al. 2000) regulate *HIF-1α* either transcriptionally or post-transcriptionally and synergistically promote cell growth.

2.1.4.5 Apoptosis or Necrosis

Hypoxia can induce apoptotic and necrotic cell death. HIF-1α increases the expression of *NIX* (Sowter, Ratcliffe et al. 2001) and *NIP3* (Bruick 2000, Velde, Cizeau et al. 2000). Activated NIP3 can cause early plasma-membrane permeability, cytoplasmic vacuolation, mitochondrial damage and mitochondrial autophagy, which can eventually lead to necrosis.

The way HIF-1α interacts with p53-dependent apoptosis is determined by HIF-1α phosphorylation status (Suzuki, Tomida et al. 2001). When HIF-1α is dephosphorylated, there are two possible scenarios: HIF-1α stabilizes p53 and activates apoptosis, or p53 blocks HIF-1α and inhibits the transcriptional activation of anti-apoptotic genes. When HIF-1α is phosphorylated, it binds to ARNT (Bacon and Harris 2004) and regulates cytochrome c-independent apoptosis via NIX and NIP3.

Tumor cells have developed strategies to evade the apoptosis induced by HIF-1α. In the early stages of tumor progression, the activation of pro-apoptotic genes can be induced by hypoxia. Circumventing apoptosis due to hypoxia becomes a selective factor

for tumor cells. Cells that survive hypoxia would gain advantage in the clonal evolution of cancer and exhibit aggressive phenotypes that is associated with hypoxia.

2.1.4.6 Genomic instability

Hypoxia can alter cell cycle checkpoint control as well as sensing and repair of DNA damage. Acute hypoxia usually generates reactive oxygen species during the anoxia and re-oxygenation cycle. Acute hypoxia can activate cellular ataxia telangiectasia mutated (ATM) – ataxia telangiectasia- and Rad3-related kinase (ATR)-mediated cell cycle check points to arrest the cell cycle and repair DNA damage caused by reactive oxygen species (Bristow and Hill 2008). Failure in repairing DNA breaks will contribute to a markedly increased genomic instability.

Chronic hypoxia can cause genomic instability because translation of DNA repair proteins is slowed down. This will result in defective DNA repair, chromosomal aberrations, fragility, and aneuploidy in proliferating cells.

2.1.4.7 Migration

HIF-1α activation enhances esophageal adenocarcinoma migration and invasion (Jing, Wang et al. 2013). HIF-1α activation attenuates the expression of E-cadherin, a component of adherens junctions which bind cells within tissues together (Imai, Horiuchi et al. 2003). HIF-1α also enhances the expression matrix metalloproteinase-2 (*MMP-2)*, which disrupts cell-cell and cell-matrix interactions) (Semenza 2003), twist family bHLH transcription factor 1 (*TWIST1*), which regulates epithelial-mesenchymal transition) (Yang 2008), and *c-met*, which promotes invasive growth (Pennacchietti, Michieli et al. 2003).

2.2 Barrett's esophagus

2.2.1 Overview

Barrett's esophagus (BE) is a pre-malignant condition of esophageal adenocarcinoma. Barrett's esophagus develops when esophageal squamous epithelium is damaged by chronic acid-bile reflux and replaced by columnar epithelium. Barrett's esophagus is the strongest risk factor for esophageal adenocarcinoma (EAC). Patients with Barrett's esophagus have 30- to 60-fold greater risk of developing esophageal adenocarcinoma than the general population (Cameron, Ott et al. 1985, Drewitz, Sampliner et al. 1997, Kim, Weissfeld et al. 1997). Recent studies suggest that an approximate yearly rate of 0.5% of Barrett's esophagus patients who progress to adenocarcinoma (Shaheen, Crosby et al. 2000). The incidence of EAC has increased 7-fold from 1973 (3.6 cases per million) to 2006 (25.6 per million) in the United States (Spechler 2013). The increasing trend of esophageal adenocarcinoma is greater than that reported in melanoma, breast cancer and prostate cancer. The prognosis of esophageal adenocarcinoma is poor, with a five-year survival rate of only about 10% in most Western countries (Portale, Hagen et al. 2006). Introduction of screening and surveillance programs for early BE and EAC detection have resulted in greater survival rates.

Other than early identification and stratification purposes, Barrett's esophagus is considered as a valuable model for studying premalignant progression of cancer. Most premalignant conditions are difficult to follow, either because they are removed upon detection (colonic polyps) or hard to biopsy (pancreatic cancer) (Paulson and Reid 2004). In contrast, periodic biopsies of Barrett's esophagus patients can be taken from the same patient to test for dysplasia as a part of standard-of-care. This allows researchers to study

premalignant progression longitudinally and evaluate genetic alteration during clonal evolution in cancer. By studying the different steps during pre-malignant progression in Barrett's esophagus, researchers can gain insights into molecular mechanisms and potentially expand the findings to other types of cancers.

Pathological examination of tissue architecture has been the standard detection and classification method for Barrett's esophagus diagnosis and treatment. Dysplasia in Barrett's esophagus is the neoplastic epithelium which is still confined within the basement membrane and has not formed a mass. Dysplasia has been used as a marker for patient's stratification for the risk of developing adenocarcinoma (Reid, Levine et al. 2000, Weston, Sharma et al. 2000, Schnell, Sontag et al. 2001, Overholt, Lightdale et al. 2005). Barrett's esophagus ranges from metaplasia, low-grade dysplasia, high-grade dysplasia and invasive carcinoma. Challenges with using histology-based clinical care involve variation among pathologists' interpretations, high biological heterogeneity within the same grade of dysplasia and so forth (Ong, Lao-Sirieix et al. 2010). Therefore, identifying molecular and/or imaging markers for Barrett's esophagus neoplastic progression and esophageal adenocarcinoma risk prediction is urgent and necessary.

Besides tissue architecture, promising biomarkers for surveillance of Barrett's esophagus patients include DNA content abnormalities and loss of heterozygosity (9pLOH , 17pLOH) (Galipeau, Li et al. 2007), and markers of proliferation (minichromosome maintenance protein (*Mcm*) 2, 5 and *Ki67* (Sirieix, O'Donovan et al. 2003), *p53* positivity by immunohistochemistry (Weston, Banerjee et al. 2001, Murray, Sedo et al. 2006), cell cycle marker (*CDKN2A*, *cyclin A* and *cyclin D*) (Lao-Sirieix, Lovat et al. 2007), epigenetic changes (methylation markers) (Schulmann, Sterian et al. 2005,

Jin, Cheng et al. 2009). These markers have gone through the third or fourth stage of biomarkers development as defined by the Early Detection Research Network (EDRN) (Pepe, Etzioni et al. 2001, Ong, Lao-Sirieix et al. 2010). These markers, representing genomic instability, disruption of regulatory pathways and genetic divergence, also feature fundamental properties of neoplastic progression. However, most of these biomarkers are not clinically available because the biomarker lacks sufficient sensitivity and specificity (Pepe, Etzioni et al. 2001, Ong, Lao-Sirieix et al. 2010). This calls for large scale translational research to advance the field of biomarker discovery and clinical implementation.

2.2.2 Hallmarks of pre-malignant progression in Barrett's esophagus

Neoplastic progression in Barrett's esophagus arises from a series of genetic and epigenetic changes. These changes can be categorized according to the cancer hallmark traits that they affect. The hallmarks contribute to the evolution and progression from Barrett's metaplastic cells to esophageal adenocarcinoma (Morales, Souza et al. 2002).

2.2.2.1 Sustaining proliferative signaling

Genetic alterations in cell-cycle control genes usually can affect cell proliferation. Growth factors, hormones, and cytokines can activate transmembrane receptors and downstream signaling pathways that include pro-growth cyclin protein family. In Barrett's esophagus, cyclin D1 and cyclin E (Bani-Hani, Martin et al. 2000) are overexpressed. They form complexes with cyclin-dependent kinases and drive cell-cycle progression. Tumor cells can also promote their own growth via autocrine signaling with growth factors or by modifying the growth-factor receptors to achieve growth self-sufficiency. The pre-malignant progression in Barrett's esophagus is also associated with

upregulation of epidermal growth factor (Pande, Iyer et al. 2008), *Erb* family of tyrosine-kinase ligands and receptors (Miller, Moy et al. 2003), transforming growth factor (Rees, Onwuegbusi et al. 2006), and fibroblast growth factor (Lord, Park et al. 2003).

2.2.2.2 Evading growth suppressors

Mutations of TP53 (Gonzalez, Artimez et al. 1997) and *CDKN2A* (Klump, Hsieh et al. 1998, Wang, Guo et al. 2009), promoter hypermethylation of *CDKN2A* and *APC* (Wang, Guo et al. 2009), and non-random losses of heterozygosity (Barrett, Sanchez et al. 1999) are common inactivation mechanisms of tumor suppressor genes. Abnormalities in these genes block anti-growth signals. Hypermethylation of *p16* and *APC* is a frequent and early event during the progression from normal esophagus through Barrett's esophagus to esophageal adenocarcinoma. It is a strong predictor of progression to high-grade dysplasia or esophageal adenocarcinoma in Barrett's esophagus patients (Wang, Guo et al. 2009). Loss of p53 function also occurs at an early stage of progression, which inactivates cell cycle check point mechanisms and increases the fraction of 4N cells with 4N DNA amount (Galipeau, Cowan et al. 1996).

2.2.2.3 Resisting cell death

Inhibition of apoptosis occurs early in the dysplasia-carcinoma sequence of Barrett's esophagus (Katada, Hinder et al. 1997, Halm, Tannapfel et al. 1999). In both Barrett's metaplasia and esophageal adenocarcinoma, the overexpression of cyclooxygenase-2 (*COX-2*) (Wilson, Fu et al. 1998, Morris, Armstrong et al. 2001, Souza, Shewmake et al. 2004) and *Bcl-2* (Shimizu, Vallböhmer et al. 2006) inhibits the apoptotic pathways.

In normal cells, severe DNA damage causes an accumulation of p53 and apoptosis. But in Barrett's esophagus cells, loss of heterozygosity of p53 occur in 57% of patients with high-grade dysplasia. The loss of heterozygosity of p53 helps the cells with DNA damage evade apoptosis and thereby increases the genomic instability. These patients are at increased risk for neoplastic progression to esophageal adenocarcinoma (Reid, Prevo et al. 2001).

2.2.2.4 Enabling replicative immortality

To overcome the replicative limit and gain replicative potential, neoplastic cells need to stabilize their telomeres by reactivating telomerase (Shay and Bacchetti 1997). In the Barrett's metaplasia, dysplasia, and esophageal adenocarcinoma sequence, the expression of telomerase progressively increases (Morales, Lee et al. 1998, Lord, Salonga et al. 2000). Telomerase activation may be the main reason why Barrett's esophagus cells become immortal.

2.2.2.5 Inducing angiogenesis

Any neoplasm larger than a few milligrams requires sustained angiogenesis to survive (Ausprunk and Folkman 1977). Vascular endothelial growth factors (*VEGF*s) family stimulate endothelial cells of blood vessels to grow and migrate, thus inducing angiogenesis. Different members of *VEGF* family are expressed in epithelial and endothelial cells during the neoplastic progression of Barrett's esophagus, correlating with angiogenesis in the same process (Couvelard, Paraf et al. 2000). Expression levels of *VEGF* and fibroblast growth factor (*FGF*) were significantly increased in adenocarcinoma compared with in normal squamous mucosa or intestinal metaplasia (Lord, Park et al. 2003). Microvessel intensity and the percentage of immature blood

vessels also increase in this progression (Sihvo, Ruohtula et al. 2003). Therefore, angiogenesis and neovascularization may emerge early in Barrett's esophagus and develop progressively in the metaplasia-dysplasia-adenocarcinoma sequence.

2.2.2.6 Activating invasion and metastasis

Esophageal adenocarcinoma is prone to early metastasis (Dulak, Stojanov et al. 2013). Alterations in cell-cell adhesion as well as motility signals may contribute to invasion and metastasis. The expression of E-cadherin, which bridges two cells, is significantly down-regulated as the Barrett's metaplasia-dysplasia-adenocarcinoma sequence progresses (Bailey, Biddlestone et al. 1998). Recurrent mutations in *RAC1* cell motility signaling pathway are identified in esophageal adenocarcinoma (Dulak, Stojanov et al. 2013). Activated *RAC1* signaling can enhance cell invasiveness and motility. These genetic changes may alter cytoskeletal structure, increase invasive properties, induce mitosis, and thereby increasing tumor fitness (Dulak, Stojanov et al. 2013).

2.2.2.7 Reprogramming of energy metabolism

In esophageal adenocarcinoma, cells mainly rely on glycolysis rather than oxidative phosphorylation to produce ATP with damaged mitochondria, known as the Warburg effect (Warburg, Wind et al. 1927, Taylor, Smith et al. 2009). A panel of metabolic related or hypoxia response genes, such as *Glut-1* (Younes, Ertan et al. 1997), pyruvate kinase isoform M2 (*PKM2*) (Koss, Harrison et al. 2004), *VEGF* and erythropoietin (*EPO*) (Griffiths, Pritchard et al. 2007) have been reported in Barrett's esophagus tissue.

Barrett's metaplastic cells generate energy through normal mitochondrial phosphorylation. In the intermediate stages of Barrett's dysplasia, cells retained

functional mitochondria, employ oxidative phosphorylation to produce energy and increase glycolysis in response to substrate (Suchorolski, Paulson et al. 2013).

2.2.2.8 Evading immune destruction

FasL (death-promoting ligand) is a surface protein usually expressed by activated lymphocytes and binds to *Fas* (death receptor) on lymphocytes and gut-epithelial cells. The lymphocytes received the death signal and attack neoplastic cells. In Barrett's metaplastic tissue, neoplastic cells express an excess amount of *FasL*, which occupies *Fas* on the surface of lymphocytes and kills them (Younes, Schwartz et al. 1999, Younes, Lechago et al. 2000). This is mechanism how Barrett's esophagus cells evades immune surveillance.

2.2.2.9 Genomic instability

Chromosomal and genomic instability predicts the progression of Barrett's esophagus (Rabinovitch, Reid et al. 1989, Galipeau, Cowan et al. 1996, Barrett, Sanchez et al. 1999, Paulson, Maley et al. 2009). 17p (p53) allelic losses, methylation of *CDKN2A*/*p16*, loss of heterozygpsity at 5q, 9p, 13q, 17p and 18q, DNA-content aneuploidy or increased 4N (G2/tetraploid) populations are all associated with Barrett's esophagus neoplastic progression (Barrett, Sanchez et al. 1999). Genome-wide analysis shows the number of copy number alterations predicts the progression as well (Paulson, Maley et al. 2009).

2.2.2.10 Tumor-promoting inflammation

Tumor tissues are often infiltrated by immune cells that enable or promote tumor growth. Acid and bile reflux causes inflammation in esophagus, which potentially causes Barrett's metaplastic cells to progress (Fitzgerald, Abdalla et al. 2002). The inflammatory

cell infiltrate generates reactive oxygen species (Naya, Pereboom et al. 1997), which might contribute to DNA damage. Reactive oxygen species may also induce growth factors, survival factors or *Fas* ligand secretions (Younes, Schwartz et al. 1999). Inflammatory cell infiltrate themselves produce many cytokines, such as transforming growth factor β (*TGFβ*), interleukin one β (*IL-1β*), interferon γ (*IFNγ*), interleukin six (*IL-6*) (Zhang, Zhang et al. 2011) and *TNFα* (Tselepis, Perry et al. 2002). Persistence of Barrett's metaplasia and the development of dysplasia and adenocarcinoma are closely associated with *IL-1β* and *TNFα* (Jankowski, Harrison et al. 2000).

2.2.2.11 Tumor microenvironment

Barrett's esophagus cells evolve under the influence of their surrounding cells and other factors in the environment. Acid and bile in Barrett's luminal refluxate can induce double-stranded DNA breaks or promote oxidative DNA damage (Clemons, McColl et al. 2007).

In the stromal compartment of Barrett's esophagus, the gene expression profiles are different between different stages of progression (Lao-Sirieix and Fitzgerald 2010). Thrombospondin-1 (*TSP1*) is overexpressed in stroma from Barrett's esophagus biopsy samples. *TSP1* can activate *TGFβ*, which either controls proliferation or promotes epithelial-mesenchymal transition in Barrett's esophagus and esophageal adenocarcinoma (Rees, Onwuegbusi et al. 2006, Onwuegbusi, Rees et al. 2007). Furthermore, co-culture of squamous carcinoma and Barrett's carcinoma cells produces more pro-inflammatory cytokines compared with cells cultured individually (Fitzgerald, Abdalla et al. 2002).

2.2.3 Models of Barrett's esophagus neoplastic progression

Pre-clinical models of Barrett's esophagus neoplastic progression exist in both tissue culture and animal models. Tissue culture models include (1) squamous cell culture: EPC2 cells (Harada, Nakagawa et al. 2003), an esophageal epithelial cell line transformed with *hTERT*; (2) Barrett's esophagus cell culture: CP-A, derived from a patient with metaplastic BE, immortalized with *hTERT* transfection, with inactivated *CDKN2A* and wildtype *TP53*; CP-B, CP-C and CP-D derived from patients with high-grade dysplasia and display *CDKN2A* and *TP53* abnormalities (Palanca-Wessels, Barrett et al. 1998); BAR-T, immortalized with *hTERT*, initially show both functioning *CDKN2A* and *TP53*, and lost *CDKN2A* during adaptation to culture conditions (Jaiswal, Morales et al. 2007) (3) esophageal adenocarcinoma cell culture, 10 cell lines have been verified to be derived from human esophageal adenocarcinoma, including FLO-1, KYAE-1, SK-GT-4, OE19, OE33, JH-EsoAd1, OACP4C, OACM5.1, ESO26, and ESO51 (Boonstra, van Marion et al. 2010). Recently, Okawa T et al. (Okawa, Michaylira et al. 2007), Koskoff et al. (Kosoff, Gardiner et al. 2012) and Stairs et al. (Stairs, Nakagawa et al. 2008) created organotypic models of Barrett's esophagus and esophageal adenocarcinoma to mimic the tumor microenvironment in vivo. Rat, mouse and dog models exist for Barrett's esophagus study. Surgical models are created by inducing reflux of gastrodeudenal acid-bile reflex into the esophagus, or removing the mucosa of the distal esophagus and generating a hiatal hernia thus inducing columnar epithelium. Surgical models are difficult to implement and have not been popular. Genetic models for Barrett's esophagus are up-and-coming. One of the latest developments of genetic models include rat models expressing intestinal transcription factor *Cdx2* ectopically in esophageal squamous tissues,

which demonstrate a progression from squamous epithelium to Barrett's esophagus (Kong, Crissey et al. 2011). In 2011, Wang and Ouyang et al. reported that *p63*-deficient mice develop intestine-like metaplasia similar to Barrett's esophagus in a process that p63-null embryonic cells migrate towards epithelium and facilitate the proliferation of columnar epithelial cells (Wang, Ouyang et al. 2011).

In this study, cell lines representing normal squamous (EPC-2), early (CP-A) and late (CP-B, CP-C, and CP-D) in Barrett's esophagus neoplastic progression are used to characterize intracellular interactions and energy metabolism changes in pre-malignant progression that have persisted in culture.

2.2.4 Genome and transcriptome study of Barrett's esophagus

Genome-wide study of neoplastic progression in Barrett's esophagus started more than 10 years ago. In 2001, Riegman et al. reported an inventory of genetic aberration during the malignant transformation in Barrett's esophagus, using comparative genomic hybridization method to evaluate esophageal adenocarcinomas, as well as metaplasia, low-grade dysplasia and high-grade dysplasia in Barrett's esophagus (Riegman, Vissers et al. 2001). They identified losses of 5q21-q23, 9p21, 17p12–13.1, 18q21, and Y in low-grade dysplasia, loss of 7q33-q35 and gains of 7p12-p15, 7q21-q22, and 17q21 in high-grade dysplasia, and a variety of known and novel aberrations in adenocarcinoma as well. This study also revealed potential discriminators between different stages of neoplastic progression and adenocarcinoma.

Furthermore, Barrett et al. performed transcriptional profiling on biopsies obtained from Barrett's metaplasia and normal upper gastrointestinal mucosae, including gastric, duodenal, and esophageal squamous epithelium using oligonucleotide microarray

in 2002 (Barrett, Yeung et al. 2002). They identified tissue-specific cluster of genes whose expression was elevated in each of the four tissues. The Barrett's esophagus cluster showed genes are associated with a panel of different functional categories including cell cycle alteration, apoptosis, cellular movement, and stress responses, which are all associated with neoplasia. With the advancement of technology, researchers have performed differential gene expression profiling on normal, Barrett's esophagus and adenocarcinoma tissues, and found out genes involved in epidermal differentiation are suppressed during the progression to adenocarcinoma (Kimchi, Posner et al. 2005, Luthra, Wu et al. 2006). In 2009, Wang et al. analyzed three publically available microarray datasets and one public serial analysis of gene expression (SAGE) dataset using Gene Set Enrichment Analysis and immunohistochemistry methods. The results suggested that transcription factors such as *CDX1* and *CDX2*, as well as *BMP*/*TGFβ* pathways might be involved in the development of Barrett's esophagus (Wang, Qin et al. 2009).

Genome-wide single nucleotide polymorphisms (SNPs) studies provide a powerful tool to discover genetic cause and premalignant neoplastic progression of Barrett's esophagus. A study using high resolution array-comparative genomic hybridization by Lai et al. showed that copy number and allelic changes, corresponding to genomic instability, increase during the course of neoplastic progression (Lai, Paulson et al. 2007). Notably, the first genome-wide association study of Barrett's esophagus also employed a SNP test to analyze samples from more than 10,000 cases and controls (Su, Gay et al. 2012). The paper reported two SNPs on chromosomes 6p21 and 16q24 predispose to the development of Barrett's esophagus. The lead SNP 6p21 is close to the major histocompatability complex (MHC) region, where immune system, inflammation

and olfactory related genes are located. The other region 16q24 is close to *FOXF1*, which is involved in gastrointestinal tract development.

2.3 Single cell transcription analysis

2.3.1 Single cell analysis

Many traditional biochemical approaches rely on the presumption that cells within a population appear similar and behave like each other. Most assays are based on the lysis of hundreds or thousands of cells and the discernment of their component parts. However, cell-to-cell differences widely exist in multiple cell and tissue types, ranging from uni-cellular organisms to complex tissues (Lidstrom and Meldrum 2003, Raj and van Oudenaarden 2008). Cellular heterogeneity within genetically similar or identical cell populations usually arises from stochastic expression of genes, proteins and metabolites (Wang and Bodovitz 2010). The sources of such variability can be ascribed to extrinsic factors, such as subtle differences in the microenvironment (growth factors, oxygen and other environmental components), or intrinsic factors which are due to the inherently probabilistic and discrete nature (Snijder and Pelkmans 2011). Therefore, biochemical research at the bulk-cell level is prone to average out the cellular heterogeneity and mask the presence of functionally important subpopulation of cells.

Single cell analysis is a new frontier for gaining insights into cancer (Dalerba, Kalisky et al. 2011). A tumor is not merely a collection of malignant cells with identical behavior. Solid tumors are usually composed of molecularly distinct clones that differ in growth rates, metastatic potential and responses to drug treatment. Tumor heterogeneity adds complexity and difficulty in discovering diagnostic and prognostic cancer biomarkers, as well as finding effective target therapies of disease. Single cell analysis

can not only measure intratumor heterogeneity, but detect rare tumor cells at early stages (Russnes, Navin et al. 2011) and monitor circulating tumor cells. Investigating cancer at single-cell resolution has the potential to resolve fundamental biological questions and improve disease management.

Currently, single cell analysis has been challenging for several reasons: (1) the scarcity of the components, such as DNA, RNA and proteins, inside a single cell and (2) the lack of integrative analysis tools for "omics" data at the single-cell level. Over the past decade, significant efforts have been made towards developing tools to overcome the limitations and uncover the biology inside a single cell (Van Gelder, von Zastrow et al. 1990, Chiu and Lorenz 2009, Tang, Barbacioru et al. 2009, Zhu, Holl et al. 2009, Anis, Holl et al. 2010, Taniguchi, Choi et al. 2010, Tay, Hughey et al. 2010, Anis, Houkal et al. 2011, Bendall, Simonds et al. 2011, Fan, Wang et al. 2011, Flatz, Roychoudhuri et al. 2011, Gao, Zhang et al. 2011, Navin, Kendall et al. 2011, Schubert 2011, Zeng, Wang et al. 2011, Bartfai, Buckley et al. 2012, Kelbauskas, Ashili et al. 2012, Xu, Hou et al. 2012, Zong, Lu et al. 2012, Shi, Gao et al. 2013, Wang, Shi et al. 2013). Specifically, the ability of dissecting the transcriptome at the single-cell level, using next-generation sequencing (Tang, Barbacioru et al. 2009, Navin, Kendall et al. 2011, Xu, Hou et al. 2012, Zong, Lu et al. 2012, Lasken 2013), quantitative polymerase chain reaction (qPCR) (Flatz, Roychoudhuri et al. 2011, Zeng, Wang et al. 2011) and microscopy (Taniguchi, Choi et al. 2010), has greatly advanced our knowledge in understanding complex systems such as cancer and development.

2.3.2 Methods for transcript profiling at the single-cell level

Three major categories of transcript profiling methods exist for single-cell analysis: (1) whole-genome RNA-Seq, (2) real-time qPCR (RT-qPCR), and (3) image-based single molecule RNA fluorescence *in situ* hybridization (FISH). Both RNA-Seq and RT-qPCR take cell lysate as the input, while RNA FISH examines fixed and permeabilized cells. The former two approaches share common steps: single cell harvesting, cDNA synthesis and detection.

2.3.2.1 Harvesting

Prior to end-point analysis of transcripts, single cells need to be isolated from culture or tissue, or retrieved from experimental platforms. Single cells should be harvested with minimal perturbation to their original state.

2.3.2.1.1 Micromanipulation

Micromanipulation has been the gold standard method for handling single cells. There are two types of micromanipulation: mechanical and optical manipulations. A mechanical micromanipulation system usually has an inverted microscope, a motorized platform operated by a joy-stick, and a microcapillary pipette connected to a pump for aspirating and dispensing single cells. Using a mechanical micromanipulator, single cells can be individually captured from a cell population and transferred to other culture conditions or end-point experimental vials. A more advanced system with vision-based feedback control can help users verify the successful transfer of single cells into a new experimental condition or vial, adding a straight-forward quality control step (Anis, Holl et al. 2010, Anis, Houkal et al. 2011, Zeng, Wang et al. 2011, Kelbauskas, Ashili et al.

2012). Current mechanical micromanipulation platforms largely rely on the expertise of a well-trained experimentalist, have low throughput and are difficult to automate.

Optical tweezers are another micromanipulation method. Single cells can be trapped and manipulated by highly focused, near-infrared laser beams (Arai, Ng et al. 2005, Zhang and Liu 2008). Two forces applied by the light, the scattering force and the gradient force, can move the trapped cell and transfer it to another compartment for later analysis. The optical manipulation is sterile because it does not need physical contact. Fortunately, visual evaluation of single cells is an inherent advantage of optical tweezer manipulation. However, the throughput of optical micromanipulation is also limited and automation is hard to achieve. At present, micromanipulation is probably the most precise method for single-cell isolation and transfer. Nonetheless, implementation of high-throughput and automated platforms, with minimal damage to the cells, is needed urgently for single-cell analysis (Yun, Kim et al. 2013).

2.3.2.1.2 Laser capture microdissection

Laser capture microdissection is a method to obtain single cells from a fixed tissue or live cell culture under the guidance of microscopic visualization. A typical laser capture microdissection process includes the following steps: (1) visualizing cells via light microscopy, (2) transferring infrared laser energy to melt a thermolabile polymer and form a polymer-cell composite (capture method), or transferring ultraviolet laser energy via an ultraviolet laser to photovolatize a region of tissue (cutting method), and (3) removing single cells from the tissue or live culture (Emmert-Buck, Bonner et al. 1996, Suarez-Quian, Goldstein et al. 1999). One outstanding advantage of laser capture microdissection is that samples can be preserved in the condition closest to its original

one, without physical contact or contamination. However, even though laser capture microdissection has visual-feedback control during cell collection, it does not have this control when it places single cells into tubes or other vials. The throughput is also limited because it could only capture one cell at a time. Laser capture microdissection is a precise single cell sampling technique that could be used in DNA, RNA (Tietjen, Rihel et al. 2003), protein and metabolite level analysis (Suarez-Quian, Goldstein et al. 1999, Golubeva, Salcedo et al. 2013).

2.3.2.1.3 Fluorescent-activated cell sorting

Fluorescent-activated cell sorting (FACS) is an automated, high-throughput method for harvesting single cells. In FACS, cells from tissue or culture are dissociated, suspended and then mixed with a carrier fluid. A stream of cells is generated by hydrodynamic focusing. Only single cells could pass through an illumination zone where forward and side light scatter and several fluorescence parameters can be measured. An intense vibrator underneath the illumination zone creates droplets from the stream. Each droplet contains only one single cell. Droplets carrying individual cells are charged and deflected into a micro-titer plate. FACS can sort single cells with a rate up to 40 cells per second (El-Ali, Sorger et al. 2006). Nonetheless, because cells should be dissociated from the tissue or culture, the transcriptome changes due to microenvironment alteration may cover up the original subtle RNA profiles of the cell of interest. Moreover, FACS requires a relatively large volume of sample and reagent for analysis, which limits the full utilization of the FACS system in precious biological samples.

2.3.2.1.4 Microfluidics

The microfluidic lab-on-a-chip platform has emerged as a powerful technology to analyze microscale to nanoscale biological systems (Zare and Kim 2010). It can handle rare cells and streamline multiple procedures on a single chip with high-throughput (Chiu and Lorenz 2009). A plethora of microfluidic single cell sorting/isolation techniques have been developed based on various forces, including optical (Krishnan and Erickson 2012), magnetic (Liu, Lien et al. 2009), electrical (Prinz, Tegenfeldt et al. 2002, Peitz and van Leeuwen 2010) and mechanical force (Di Carlo, Wu et al. 2006).

In optical manipulation, target cells in a microfluidic channel can be simultaneously measured and sorted based on their physical properties, such as size, light absorption and refractive index (Zhang and Liu 2008). Optical trapping can be combined with a microfluidic system, which helps investigators manipulate the microenvironment and study cell behavior and response under physical and chemical stimulation.

Magnetic-activated cell sorting techniques sort single cells based on their membrane-conjugated magnetic nanoparticles or internalized magnetic nanoparticles (Schmitz, Radbruch et al. 1994, Tseng, Judy et al. 2012). Microfluidic channels or microwells can be integrated with magnets to generate a magnetic field and trap single cells attached to magnetic beads (Tseng, Judy et al. 2012).

Dielectrophoresis is a noncontact, noninvasive method to trap and isolate cells from a heterogeneous culture (Pohl and Pohl 1978). A microfluidic cytometer utilizing dielectrophoresis was developed to trap cells against fluid flows (Hu, Bessette et al. 2005, Hunt and Westervelt 2006). Dielectrophoresis operates on the motion of dielectric cells,

focuses on cells within the fluid stream and then traps single cells into the center of the microchannel.

Hydrodynamic manipulation is the most common mechanical approach for isolating cells (Yang, Li et al. 2002, Wheeler, Throndset et al. 2003, Lee, Hung et al. 2005). Cells can be hydrodynamically focused into a narrow stream inside a microchannel. This is then followed with accurate positioning of cells by balanced flows.

Microfluidic single-cell isolation techniques hold great potential for engineering and biotechnological studies because they are handy, inexpensive and high throughput. When combined with other upstream (microenvironment monitoring, drug treatment, imaging and so on) or downstream biochemical assays, microfluidics can become more versatile in the blooming field of single-cell analysis.

2.3.2.1.5 Summary of harvesting methods

Each of the current single cell harvesting techniques has its own advantages and drawbacks. The micropipette manipulation method has almost the highest precision but its throughput is low. Laser capture microdissection does not perturb the original states of the cell but the cost is expensive. Fluorescent-activated cell sorting is a traditional method for sorting cells. However, non-cellular particles or multiple cells can be collected as a single cell in the process. The achievements of microfluidic single-cell isolation techniques are encouraging in mimicking tissue or organ-like network environments *in vitro*. The above-mentioned techniques will become more useful when combined with other phenotype measurement techniques to explore various aspects of single cells.

2.3.2.2 Next-generation sequencing

Next-generation sequencing has become a revolutionary approach for acquiring genome-scale data since 2005 (Margulies, Egholm et al. 2005). It allows for more qualitative and quantitative DNA to be acquired than the automated Sanger sequencing method (Sanger, Nicklen et al. 1977) which is considered "first-generation" technology. Next-generation sequencing brought paradigm shifts in many perspectives including template preparation, sequencing, and data analysis (Reis-Filho 2009).

2.3.2.2.1 Template Preparation

Next-generation sequencing does not require a cloning step as needed in the Sanger sequencing approach (Metzker 2010, Mardis 2011, Mardis 2013). DNA are randomly broken into smaller sizes and covalently linked with adapters. The DNA library fragments are clonally amplified by emulsion PCR or on a solid-phase, either on a bead or a flat glass microfluidic channel instead of microplate wells used in Sanger sequencing (Mardis 2013). The digital nature of clonal amplification results in a population of identical templates, each of which originates from a single fragment and undergoes the sequencing reaction.

2.3.2.2.2 Sequencing

The process of next-generation sequencing is massively parallel (Mardis 2013). After fragment amplification, the nucleotide addition reaction and imaging detection happen simultaneously. Repeated cycles of nucleotide synthesis and sequence extensions can yield hundreds of megabases to gigabases of output in one single instrument run. Four major platforms based on different sequencing mechanisms are reviewed below.

2.3.2.2.2.1 Pyrosequencing

The first commercially available next-generation sequencing platform is based on pyrosequencing technology developed by Roche/454 in 2005. It relies on detecting pyrophosphate released during nucleotide incorporation. Each cycle of pyrosequencing consists of (a) the addition of a single nucleotide, (b) the introduction of substrate and enzymes (adenosine 5'-phosphosulphate, sulphurylase, luciferase, and luciferin), (c) the generation of ATP from PPi and the emission of light during the luciferin reaction driven by ATP, (d) a wash step by apyrase to remove unincorporated nucleotide (Shendure and Ji 2008). Two of the most outstanding characteristics of pyrosequencing by Roche/454 are reading length and speed (Liu, Li et al. 2012). The 454 GS FLX sequencer can read up to 1,000 bp in length and the run time is 23 hours.

2.3.2.2.2.2 Sequencing by synthesis

The cyclic reversible termination sequencing method, developed by Solexa (acquired by Illumina in 2007), dominates the current next generation sequencing market. After the adaptor-flanked fragment is PCR amplified, a dense array of sequences will form on the solid platform. A typical sequencing cycle usually consists of the following steps: (a) the addition of a mixture of four dideoxynucleotides, each labeled with a different dye, (b) a wash step of unincorporated nucleotides, (c) four-color imaging detection of the incorporated nucleotide, (d) cleavage of both the fluorescent labels and terminating groups (Mardis 2013).

2.3.2.2.2.3 Sequencing by ligation

Sequencing by Oligo Ligation Detection (SOLiD) is a sequencing platform based on two-base-encoded probes in the sequencing by ligation process. It is also a cyclic

method with the following steps: (a) the hybridization of a fluorescent probe to the target sequence and the ligation between the probe and the sequence primer, (b) a wash step to remove non-ligated probes, (c) four-color imaging and (d) the cleavage of the fluorescent probe. SOLiD sequencing has extraordinary sequencing accuracy. The SOLiD 5500xl has the accuracy of 99.99% when it came to the market in 2010 (Barba, Czosnek et al. 2014).

2.3.2.2.2.4 pH change monitoring

In 2010, Ion Torrent commercialized a semiconductor-based sequencing instrument. It detects the pH change upon the incorporation of a new nucleotide during the synthesis process (Rothberg, Hinz et al. 2011). During sequencing, different types of nucleotides flow across the chip in a systematic order. When a nucleotide is incorporated, a proton is released. The pH change can be detected by the hydrogen ion detector and the signal can be translated into a quantitated readout of nucleotide bases. In Ion Torrent sequencing, incorporating nucleotides are not modified with fluorescence dyes or blocking groups. A semiconductor sensor replaced a camera in recording signals. Both factors speed up the sequencing process and lower the cost (Mardis 2013). Ion Torrent has the potential of becoming a fast sequencer in detecting new pathogens.

2.3.2.2.2.5 Single-molecule real-time sequencing

Single-molecule real-time sequencing is one of emerging third-generation DNA sequencing technologies (Eid, Fehr et al. 2009, Schadt, Turner et al. 2010). It is distinct from aforementioned next-generation sequencing methods because it detects single molecules without washing steps during DNA synthesis. It directly observes kinetics of DNA polymerase as it incorporates nucleotides. The DNA polymerase is modified to lower the rate of polymerization and can incorporate fluorescently-tagged nucleotides.

This approach requires neither PCR amplification nor synchronized synthesis and detection, thereby reducing amplification biases and dephasing errors. Since no scanning and washing steps are needed, the rate of sequencing could be accelerated from days to minutes. The read lengths of single-molecule real-time sequencing are 1,000 bp on average, reaching 10,000 bp at maximum. It is an unmatched capability so far, enabling de novo assembly (Schadt, Turner et al. 2010).

2.3.2.2.3 Sequencing alignment and assembly

The first step to understand the sequence reads is to align them to a reference or to assemble them de novo (Flicek and Birney 2009).

Alignment programs for Sanger sequencing, such as BLAST (Altschul, Gish et al. 1990), were designed to search for homologous sequences in large databases. In contrast, short-read alignment algorithms for next-generation sequencing tend to assume the sources of expected mismatches are species polymorphisms and the technology error rate (Flicek and Birney 2009, Li and Homer 2010). Two major categories of alignment algorithms are (a) hash table-based implementations, which build a hash table of either the input reads or the reference genome assembly, and (b) alignment based on suffix/prefix tries, such as Burrows Wheeler Transform, which enables fast string matching (Li and Durbin 2009, Li and Homer 2010).

The framework for assembly is based on the notion of a k-mer in a de Bruijn graph data structure. In this graph, k-mers are collected into nodes and the nodes are adjacently combined to form continuous linear stretches and combined further into larger nodes. Finally, the sequencing errors will be corrected to create a final graph structure representing the original genome sequence (Miller, Koren et al. 2010).

2.3.2.3 RNA-Seq

The advent of next-generation sequencing technologies brought unprecedented changes in characterizing and quantifying transcriptome by introducing RNA (cDNA) sequencing at massive scale (Wang, Gerstein et al. 2009). A population of RNA can be reverse-transcribed into cDNA and fragmented, followed with library construction which adds adaptors to both ends of the cDNA for sequencing. The sequence reads are aligned with the reference genome or transcriptome. The aligned reads are classified as rRNA, snRNA, exon, intron, junction and intergenic reads. Exonic reads and junction reads are used to generate a base-resolution expression profile for each gene. The applications of RNA-Seq are versatile: mapping transcriptional structure, splicing patterns and other post-transcriptional modifications, cataloguing mRNAs, non-coding RNAs and small RNAs, and identifying differentially expressed genes between different conditions (Pareek, Smoczynski et al. 2011).

One of the advantages of RNA-Seq is its accuracy in quantifying RNA expression levels. Compared with hybridization-based approaches such as microarrays, RNA-Seq can detect novel transcripts that do not exist in current genome databases. The sequences throughout the exon are uniformly covered. The gene expression levels can be quantified by normalizing the total number of reads that fall into the exons of a gene against the length of the exons that can be uniquely mapped (RPKM). In principle, by increasing coverage, depth and amplification capability, RNA-Seq can capture the presence of every RNA molecule from a cell (Wang, Gerstein et al. 2009). It also does not have an upper limit of quantification, either. RNA-Seq greatly expands the dynamic range of

quantifying gene expression level--more than 8,000-fold (Nagalakshmi, Wang et al. 2008).

2.3.2.4 Single cell RNA-Seq

The total RNA from one single cell is about 1-10 picograms (Iscove, Barbara et al. 2002). Many researchers are working on lowering RNA input requirements for next-generation sequencing. Particularly, real-time single-molecule sequencing holds promise in analyzing RNA at the single cell level. Nevertheless, the amount of RNA from a single cell does not meet the microgram level input on any current platforms. Therefore, amplification is an indispensable and the most critical step for RNA-Seq at the single-cell resolution. So far, four main amplification approaches have been reported for amplifying RNA from single cells for RNA-Seq (Hebenstreit 2012).

The initial mRNA-Seq method was reported in 2009. (Tang, Barbacioru et al. 2009). Tang et al. described an amplification method to generate as long as 3 kilobases of cDNA without bias. The mRNAs obtained from a single cell are reverse-transcribed into cDNA, using a poly-T primer anchoring on its 3' end. Excess primers are digested with exonuclease. Poly-A sequence is added to the 3' end of the first-strand cDNAs and then second-strand cDNAs are synthesized using another poly-T primer with another anchor sequence. The cDNA is amplified using two anchor sequence primers by PCR. After that, the cDNAs are fragmented, ligated with adaptors and further amplified (Tang, Barbacioru et al. 2010). However, the strand information is not preserved since PCR amplification occurs before fragmentation.

Another major category of amplification uses the "template-switching" method (Cocquet, Chong et al. 2006). This method is recently commercialized by Clontech as the

SMARTer (Switching mechanism at the 5' end of the RNA transcript) Ultra Low RNA Kit for Illumina sequencing. The SMART-Seq method starts with lysing individual cells in a hypotonic solution containing a high concentration of RNase inhibitors. It also uses poly-T primer to synthesize the first strand from the poly-A RNA. It employs Moloney murine leukemia virus (MMLV) reverse transcriptase to add a few nontemplated C nucleotides to the 3' end of the cDNA. The oligonucleotide with ribonucleotide G forms base-pairs with the additional C nucleotides and become an extended template for the second strand synthesis. The MMLV reverse transcriptase switches templates and synthesizes DNA until the end of the sequence. The oligonucleotide with ribonucleotide G and poly-T serve as the priming sites for PCR amplification. This ensures only the cDNAs with ribonucleotide G priming site can be amplified and sequenced, improving the read coverage across full-length transcripts (Goetz and Trimarchi 2012, Ramsköld, Luo et al. 2012, Picelli, Björklund et al. 2013). RNA-Seq data generated using SMART-Seq showed a substantial increase in the number of alternative transcript isoforms and identification of single-nucleotide polymorphisms.

In 2012, Tamar Hashimshony et al. presented Cell Expression by Linear amplification and Sequencing (CEL-Seq) method using linear in vitro transcription (IVT) (Hashimshony, Wagner et al. 2012). This strategy begins with reverse transcription of RNA from individual cells. The primer is designed with an anchoring poly-T, a sequencing adaptor, a unique barcode for individual cells, and a T7 promoter. After the second-strand synthesis, the cDNA is pooled from multiple cells with unique barcodes for RNA synthesis. RNA is then fragmented and ligated with another adapter. Finally,

RNA is reverse transcribed to DNA. The linear amplification method is claimed to outperform single cell RNA-Seq with PCR amplification in sensitivity and accuracy.

Recently, Xinghua Pan et al. developed two methods for amplifying cDNA for single cell RNA-Seq (Pan, Durrett et al. 2013). One approach uses Phi29 DNA polymerase, an active player in multiple displacement amplification. After full-length cDNAs are synthesized from RNA, cDNAs are circularized at the intramolecular level for efficient amplification by Phi29 DNA polymerase. Another approach is called semi-random primed PCR-based mRNA transcriptome amplification. RNAs are also reverse-transcribed into cDNA first. Semi-random primers are then used to generate a library of overlapping cDNA fragments. Both methods can cover RNA (cDNA) in the full-length, as long as 23 kb. The Phi29-based method uses an isothermal reaction and produces longer products; while the semi-random primed PCR method detects more genes when the transcript level is low.

Amplification for single-cell RNA-Seq faces several challenges including deepening transcript coverage, reducing amplification bias and increasing reproducibility. Besides amplification, single-cell RNA-Seq technology also needs to meet several requirements (Hebenstreit 2012). Quantification of transcripts at the single cell level lacks controls. Spike-in controls or endogenous RNA controls can help calibrate the original number of transcripts given the information on sequencing reads. Another issue is parameter estimation for mathematical models of transcript regulation. At least a few hundred cells in each experimental condition should be analyzed for modeling purposes. Amplification methods with multiplexing of different cells, such as SMART-Seq and CEL-Seq, will help accelerate the process.

2.3.2.5 Single-cell real-time quantitative polymerase chain reaction

Real-time quantitative polymerase chain reaction (RT-qPCR) is a classical method for analyzing transcript levels. It is characterized by a large dynamic range, excellent reproducibility and sufficient sensitivity to detect a single transcript. Single-cell RT-qPCR includes several sequential experimental steps: cell collection, cell lysis and RNA extraction, reverse transcription, pre-amplification (if needed), real-time PCR and data analysis (Stahlberg and Bengtsson 2010).

Current RT-qPCR systems depend on the detection and quantification of fluorescent reporters. Two of the commonly used reporters are SYBR Green and Taqman Probe (Kubista, Andrade et al. 2006). SYBR Green dye binds to all double-stranded DNA. As the double-stranded DNA is synthesized throughout the cycle, the fluorescent signal increase can be measured and is proportional to the amount of PCR products. The SYBR Green probe is inexpensive and easy to use. However, it could bind to primer-dimers and other non-specific amplification products, which become background signals. Taqman probes are oligonucleotides with a fluorescent dye on the 5' base and a quenching dye on the 3' base. Once a Taqman probe hybridizes to an internal region of a PCR product, its fluorescent and quenching dyes are separated and the former one emits light. Fluorescence signals increases each cycle, proportional to the rate of cleaving fluorescent and quenching dyes (Parashar, Chauhan et al. 2006). Taqman probe-based qPCR has significantly higher specificity, because its probe is targeting specific amplicons. It also provides feasible multiplexing strategy, because multiple primers with different fluorescent probes could be designed to target different genes from a single cell.

Since transcript variability is intrinsic to different cells, special consideration is required in both experimental and data analysis processes. Theoretically, no gene could be used as a reference gene for normalization in single cells (Bustin, Benes et al. 2009). In this scenario, an RNA spike-in control can be used to separate experimental variability from biological variability at the single-cell level (Shi, Gao et al. 2013). The spike-in RNA can be a RNA sequence that does not exist in the species being studied. One challenge with spike-in RNA is that it should be injected into the cell so that it will go through the same lysis process as the endogeneous RNAs do.

High-content microfluidic RT-qPCR platforms have made great strides in analyzing transcripts at the single-cell level (Marcus, Anderson et al. 2006, White, VanInsberghe et al. 2011). Microfluidic platforms can limit the loss of the picogram amounts of RNA present in single cells (Marcus, Anderson et al. 2006). They also get rid of variances brought by manual handling. Most importantly, they increase the reaction throughput and minimize batch-to-batch variations. Commercially available high-throughput systems, such as Fluidigm's BioMark™ and Applied Biosystems' OpenArray® can measure transcript levels of up to 96 genes in 96 samples and 48 genes for 48 samples, respectively.

2.3.2.6 Single-molecule RNA fluorescence *in situ* hybridization

Single-molecule RNA fluorescence *in situ* hybridization (FISH) is based on the use of fluorescence labeled oligonucleotides, each targeting RNA of interest (Taniguchi, Choi et al. 2010). Hybridization of fluorescent probes to fixed single cells enables direct counting of mRNA abundance at the single-cell level.

Standard RNA FISH can quantify up to three genes simultaneously within single cells (Tischler and Surani 2013). Using spectral barcoding, an approach dividing individual probe sets into subsets and coupling each subset with a different fluorescent probe, researchers can simultaneously detect up to seven RNAs in individual cells (Femino, Fay et al. 1998, Levsky, Shenoy et al. 2002). For RNA FISH, researchers do not need to lyse individual cells and extract RNA. RNA FISH acts directly on mRNA within a fixed cell, thereby preserving the spatial information of the transcripts (Tischler and Surani 2013). Because RNA FISH directly counts the number of transcripts, it does not need internal references for quantification. It holds the potential to be combined with genomic (DNA FISH) or proteomic (protein labeling) approaches for studying "omics" within single cells. One limitation of RNA FISH is that it only measures the copy number of known transcripts (Tischler and Surani 2013).

2.3.2.7 Summary

The approaches mentioned above can provide gene transcription information at the single-cell level. The information will be revolutionary in revealing discrete molecular states within heterogeneous cell populations and identifying aberrant subpopulations. In this study, RNA-Seq and RT-qPCR were integrated to explore a little-known area in tumor neoplastic progression. The transcriptomes of cell-cell interactions and hypoxia-adaptations using RNA-Seq were analyzed. After statistical filtering and gene ontology study, changes were correlated with metabolic phenotype measurements and transcript states were analyzed using RT-qPCR. The lower cost of RT-qPCR allows analysis of a larger number of individual cells, providing higher statistical power. By combining different approaches, insights were gained into previously unknown cellular decision-

making processes in normal and aberrant states. It is expected that the progress of single-cell transcription profiling will advance personalized medicine.

CHAPTER 3

A MINIMALLY INVASIVE METHOD FOR RETRIEVING SINGLE ADHERENT CELLS OF DIFFERENT TYPES FROM CULTURES

3.1 Abstract

The field of single-cell analysis has gained significant momentum over the last decade. Separation and isolation of individual cells is an indispensable step in almost all currently available single-cell analysis technologies. However, stress levels introduced by such manipulations remain largely unstudied. In this thesis a method is presented for minimally invasive retrieval of selected individual adherent cells of different types from cell cultures. The method is based on a combination of mechanical (shear flow) force and biochemical (trypsin digestion) treatment. Alterations in the transcription levels of stress response genes in individual cells exposed to varying levels of shear flow and trypsinization were quantified. Optimal temperature, RNA preservation reagents, shear flow rate and trypsinization conditions necessary to minimize changes in the stress-related gene expression levels are reported. The method and experimental findings are broadly applicable and can be used by a broad research community working in the field of single cell analysis.

3.2 Introduction

The field of single cell analysis has experienced a tremendous growth over the last decade owing to the intense interest in intercellular heterogeneity and its functional role at the tissue level and disease states in vivo (Eberwine, Yeh et al. 1992, Elowitz, Levine et al. 2002, Lidstrom and Meldrum 2003, Irish, Hovland et al. 2004). New technological advancements have enabled the exploration of biological phenomena with single-cell resolution (Arai, Ng et al. 2005, Anis, Holl et al. 2010, Wang and Bodovitz 2010, Kelbauskas, Ashili et al. 2012). Almost all existing methods for single-cell analysis that require isolation of (Arai, Ng et al. 2005) individual cells involve some type of mechanical transportation or manipulation of single cells for sample preparation and/or analysis purposes. One of the current technological challenges is the minimization of perturbation to the cells as a result of such transportation to make biologically relevant inferences about cell function possible. If the resulting stress to the cell is significant it can alter cellular profiles at the physiological, gene transcription and/or expression levels and confound experimental results. Although widely used, stress levels introduced to cells by manipulation and, more importantly, their potential effects on cell function remain largely unknown. Mechanical cues and mechanical stress have been found to strongly affect most cellular functions and critically influence gene transcription during embryogenesis, organogenesis (Mammoto, Mammoto et al. 2012) and embryonic vasculature development (Roman and Pekkan 2012). Mechanical stress also exhibits a direct effect on the nuclear architecture-mediated gene transcription regulation (Martins, Finan et al. 2012), oncogenesis (Jean, Gravelle et al. 2011), stem cell differentiation, cancer metastasis and the immune response (Hynes 2009) among others. It is thus likely

that mechanical stress introduced during cell manipulation can significantly alter gene expression profile in cells resulting in atypical both gene expression profile and cellular function. Therefore, characterization of stress levels that can significantly perturb cell function is necessary for studies that utilize single-cell analysis techniques.

In the context of single-cell analysis methods, perturbations can be divided into two major categories with regard to time scales. One category is perturbations that cause reversible alterations that occur on a timescale that is much shorter than the time between the perturbation and analysis. By definition, perturbations of this type do not result in significant changes in the cell at the time of analysis and thus can be considered negligible. The second category is perturbations that induce a long-lasting (on timescales comparable or longer than the time between stress administration and analysis) response in the form of a modified gene expression profile. These perturbations can introduce modifications to the cell function, mRNA or protein expression levels or all of them simultaneously and thus need to be properly assessed before reaching any conclusions about experimental findings. It is likely that adherent cell types should be affected by manipulation more than non-adherent cells simply due to the fact that the former need to be detached from the growth substrate or dissociated from tissue before any kind of manipulation can be performed. Owing to changes in cellular tension, the detachment step itself could cause the cell to respond with an altered gene expression profile mediated by mechanosensing through e.g. integrin-actin linkages and mechanostransduction via downstream signaling cascades such as receptor-type tyrosine-protein phosphatase alpha (*RPTP-α*), Src family kinases (*SFK*s) (von Wichert, Jiang et al. 2003, Jiang, Huang et al. 2006, Yu, Law et al. 2011), focal adhesion kinase (*FAK*) (Wang,

Dembo et al. 2001, Schober, Raghavan et al. 2007) and others. In addition, any type of manipulation can induce additional cellular responses at biomolecular and/or organelle levels. Epithelial cells adhere to the extracellular matrix through transmembrane adhesion protein complexes. At the basal membrane, the adhesion of epithelial cells to the extracellular matrix is built upon different types of cell-ECM adhesions, including focal adhesions and hemidesmosomes, both of which are mediated by integrin connections (Balda and Matter 2003), nascent adhesions, focal complexes, focal adhesions, podosomes and others (Dubash, Menold et al. 2009). These protein complexes, including integrin-actin networks and integrin-intermediate filament networks, regulate the adhesion but also mediate mechanosensing and signal mechanotransduction into the cell (Roca-Cusachs, Iskratsch et al. 2012). To remove cells from a given culture substrate, various mechanical and chemical methods have been employed. For instance, proteolytic enzymes, such as trypsin, or chelators, can break the integrin-ligand bonds that mediate cell attachment to the substrate (Phelan 2007). However, enzymatic dissociation can damage cells, especially the cell surface. Moreover, alterations of gene expression levels in cells treated with trypsin were discovered using global gene expression profiling on the microarray platform (Chaudhry 2008). Therefore, trypsinization should be performed with caution, by optimizing both the duration of trypsinization and the concentration of trypsin.

Mechanical means such as scraping or shear flow were employed to remove cells from substrates (Zhang, Jones et al. 2008). However, mechanical methods are usually disruptive to the cells and potentially result in a loss of cellular contents. When combined with chemical force, shear flow can remove cells from the surface and transfer them

without causing damage to the plasma membrane. Still, even under relatively mild conditions, shear stress can affect signal transduction pathways, especially in endothelial cells (Chen, Li et al. 1999). Because epithelial cells experience much less shear stress in the body, gene expression alterations within those cells in response to shear stress could be significant and need to be explored. So far, expression levels of chemokine (C-C motif) ligand 2 (*CCL2*) have been shown to be upregulated in epithelial cells in response to shear flow stress (Flores, Battini et al. 2010).

In this study a new method is developed for retrieval of individual adherent cells from cell cultures that is based on a combination of mechanical forces and biochemical treatment. Stress response induced by the method in terms of gene expression levels in individual cells is investigated. Alterations in expression levels of stress response-related genes in single cells are quantified as a function of varying mechanical stress while employing both shear flow and trypsin digestion to detach single cells from glass substrates. Treatment conditions were optimized in a highly controllable manner to minimize the effects of gene expression changes that could be induced by stresses. An optimal range for mechanical force needed to efficiently detach single cells with no detectable change in expression levels of the studied stress-response genes is reported. The utility of this technique is expanded to distinguish and harvest co-cultured cells from microwells using a fluorescence-assisted single-cell harvesting method. These findings can be useful for studies focused on single-cell analysis that involve any mechanical manipulation of live cells.

3.3 Experiments

3.3.1 Microwell design

Arrays of 3 × 3 microwells were used for culturing single-cells. Each microwell has an inner diameter of 50 µm and is 20 µm deep (Figure 3-1) (Zhu, Holl et al. 2009). The microwells were fabricated in fused silica substrates using wet-etch lithography. The dimensions of these microwells were optimized for cell-cell interaction studies to monitor cellular gene expressions in a controlled microenvironment.

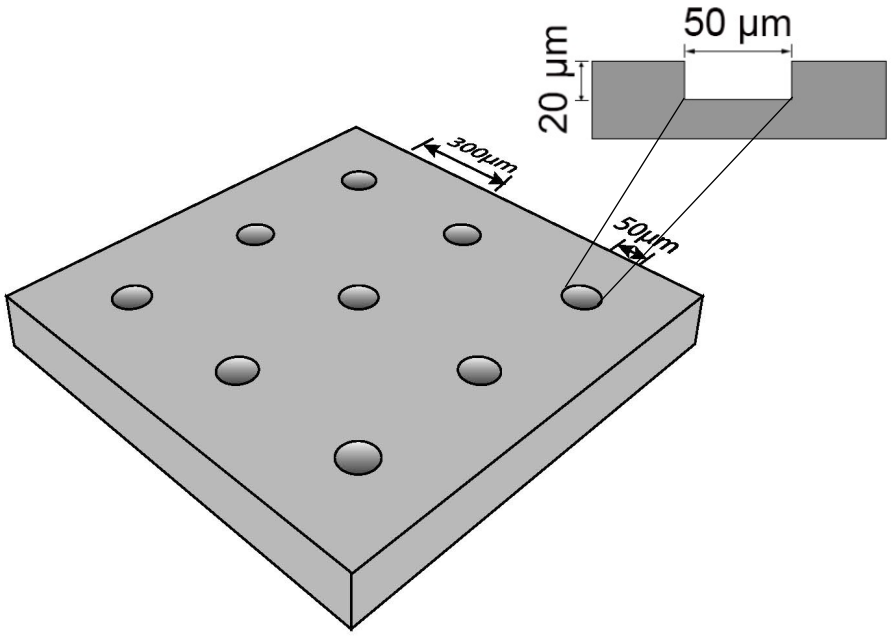


Figure 3-1 Microwell array design
3 × 3 arrays of wells with 300-µm center-to-center spacing were fabricated on fused silica wafers using hydrofluoric acid (HF) deep wet etch lithography. Each well is 20-µm deep and has a diameter of 50 µm.

3.3.2 Cell culture

The normal cell line, EPC-2 (Harada, Nakagawa et al. 2003), the metaplastic cell line, CP-A, and the pre-malignant cell line, CP-D, were derived from a healthy,

metaplastic and dysplastic human esophageal region in Barrett's esophagus (Palanca-Wessels, Barrett et al. 1998), respectively. Cells were cultured using Gibco keratinocyte serum-free cell growth medium (Invitrogen, Carlsbad, CA, USA), supplemented with hEGF (Peprotech, Rocky Hill, NJ, USA) at 2.5 μg/500 mL, BPE (bovine pituitary extract) at 25 mg/500 mL and penicillin–streptomycin solution (Invitrogen) at 100 units/100 μg/mL. Cells were grown at 37 °C under 5% CO_2. Prior to experimentation, cells were cultured in a 75 cm^2 flask to approximately 80% confluency.

3.3.3 Cell loading into microwells

Individual cells were loaded into microwells using a single-cell manipulation platform as previously described (Anis, Holl et al. 2010, Kelbauskas, Ashili et al. 2012). Briefly, the platform is built around a diaphragm micropump that can aspirate and dispense sub-nanoliter volumes of liquid. Single cells in suspension were aspirated and dispensed into microwells using a 40-μm diameter glass capillary micropipette utilizing closed-loop microscopic vision-based feedback. Cells can be aspirated from a Petri dish by a drag force generated through a negative pressure applied to the micropipette capillary. Cells can be dispensed through the micropipette capillary into the microwell by applying a positive pressure to the capillary, generating an ejection force on the cell (Kelbauskas, Ashili et al. 2012). Glass substrates containing 3 × 3 arrays of microwells were glued to the bottom of a Petri dish with a pre-cut hole using medical-grade epoxy glue (K45-S-14ML, Chemical Concepts, Huntingdon Valley, PA). Loading 9 single-cells in a 3x3 array of microwells with one cell per well takes approximately 5-8 minutes, while loading two cells per well requires 20-25 minutes. After loading into the microwells, the cells were incubated in Keratinocyte SFM, at 37 °C under 5% CO_2 for

16-24 hours to allow for cell adhesion and recuperation from potential stress caused by manipulation.

3.3.4 Direct lysis of single cells in microwells

The microwell substrate was placed on the stage of a pick-and-place single-cell manipulation platform. The micropipette was filled with Cell Lysis Buffer (Zymo Research, Orange, CA, USA) and lowered until it touched the substrate. The micropipette orifice was aligned with a selected cell inside a microwell and encapsulated by the micropipette tip. The Cell Lysis Buffer was dispensed for 1 minute with enough volume to thoroughly coat the cell. After this, the lysate was immediately aspirated into the micropipette for 1 minute to minimize the diffusion of the lysate from the microwell. The lysate was then dispensed through the micropipette into the cap of a 1.5 mL microcentrifuge tube for 1 minute.

3.3.5 Single-cell collection

The pick-and-place single-cell manipulation platform (Anis, Holl et al. 2010, Kelbauskas, Ashili et al. 2012) (Figure 3-2) was used to collect single-cells from microwells. The microwell substrate containing individual cells was washed three times with 1 mL of warmed 1× PBS, and exposed to 1 mL of 0.05% v/v trypsin-EDTA for 8-12 minutes at 37 $^{\circ}$C. Subsequently, 1 mL of trypsin inhibitor (DTI) was added to the trypsinized cells to deactivate trypsin, which was followed by adding 1 mL of Keratinocyte SFM medium. The selected cell was first aligned with the micropipette orifice and then aspirated into the micropipette capillary. The aspirated cell was dispensed from the microcapillary tip into the cap of a 1.5-mL microcentrifuge tube

(VWR, West Chester, PA, USA) containing 200 µL of Keratinocyte SFM, Cell Lysis Buffer, or RNALater (Life Technologies, Austin, TX, USA).

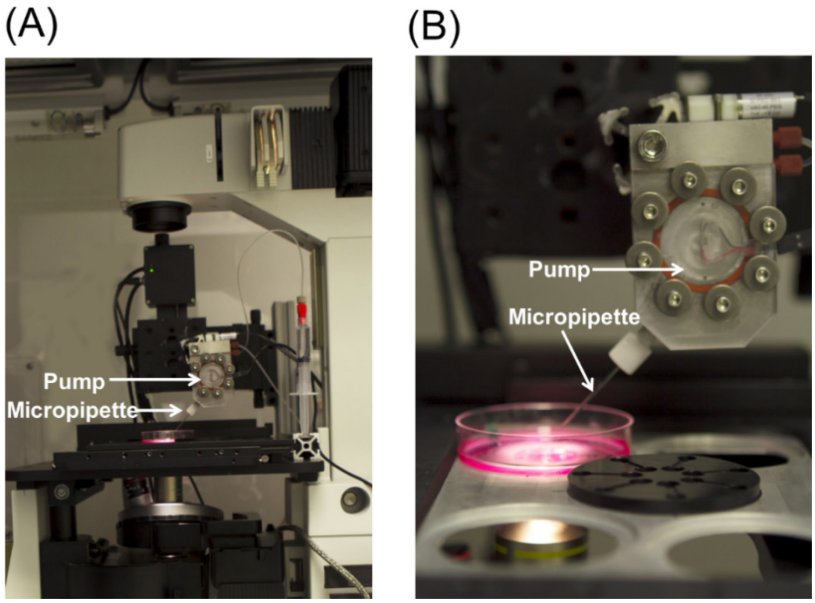
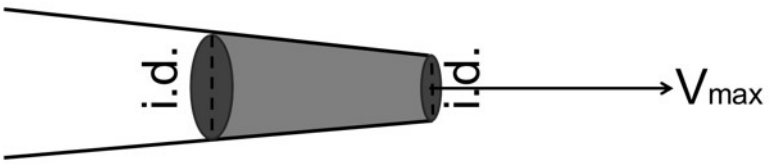


Figure 3-2 The single cell manipulation platform used in this study
(A) A side view of the platform built on a microscope. (B) Micropipette controlled by a piezoelectric pump is used for aspirating cells from and dispensing them into microwells located in the petri-dish.

3.3.6 Preservation of RNA at low temperatures

After trypsin deactivation, the substrate containing the 3 × 3 array of microwells with cells was placed inside a 4 °C refrigerator for 15 minutes. Simultaneously, the pick-and-place manipulation platform was cooled using four ice-packs for 15 minutes. The substrate was placed on the cooled stage, and the temperature was monitored with a thermometer, ranging from 2 °C to 10 °C on the station. The pick-and-place cell manipulation system was used to aspirate the harvested cells and dispense them into caps

of 1.5-mL microcentrifuge tubes as described in the Single-cell Collection Section above. A control group of cells was kept at room temperature during the harvesting process.



$$V_{\max} = Cell\ Velocity$$

$$V_{avg} \cdot Area = Flow\ Rate$$

$$V_{\max} = 2V_{avg}$$

$$\frac{\pi}{4}(40\ \mu m)^2 = Area$$

$$Flow\ Rate = \frac{V_{\max}}{2} \cdot \frac{\pi}{4}(40\ \mu m)^2$$

$$450(x\frac{V}{s}) = Cell\ Velocity(V_{\max})$$

$$Flow\ Rate = \frac{450(\frac{\mu m}{V})}{2} \cdot (x\frac{V}{s}) \cdot \frac{\pi}{4}(40\ microns)^2$$

$$Flow\ Rate(\mu m^3/s) = 2513.27 \times (x\frac{V}{s})$$

$$(10^{-6}\ \mu m^3/s) = 1\ nL/s$$

Figure 3-3 Flow rate conversion
Flow rate values of 1, 5, and 10 V/s were converted to units of nL/s. The inner diameter (i.d.) of the tip orifice is 40 µm. They were based on the assumption that the diaphragm of the piezoelectric pump is a fixed-fixed bending beam and that the piezoelectric actuator acts on the center of the beam.

3.3.7 Flow rate and trypsinization time

The two major determinants of successful cell detachment are the trypsinization time and the applied shear flow rate. Each trypsinization time point and flow rate was selected

based on preliminary experiments that showed a range of possible values that demonstrated noticeable changes in detachment, from no detachment, partial detachment to full detachment. Flow rate was converted from the voltage change rate of the piezoelectric pump to volume change rate (Table 1, Figure 3-3).

Table 1 Voltage-to-flow rate conversion

V/s	Flow rate (um^3/s)	Flow rate (nL/s)
1	2513.27	0.002513
5	12566.35	0.012566
10	25132.7	0.025133

To measure the success of each combination of flow rate and trypsinization time (Table 2), the micropipette tip was positioned vertically at a 90 degree angle with the substrate surface so that it enclosed the cell without making direct contact with the cell. A selected flow rate was then applied to aspirate the cell from the substrate. A successfully harvested cell is considered one that detaches without visible tearing of the plasma membrane and leaving remnants on the substrate. After a visual inspection and confirmation, the harvested cells were dispensed through the micropipette capillary into the caps of 1.5-mL microcentrifuge tubes as described in the Single-cell Collection Section above. The success rate of each combination of parameters is summarized in Table 2.

Table 2 Success rate of single-cell harvesting

Flow rate (nL/s)	Trypsinization Time (minutes)	Success rate of cell harvesting
2.5	5	6.7%
2.5	8	13.3%
2.5	10	40.0%
12.5	5	20.0%
12.5	8	46.7%
12.5	10	66.7%
25	5	80.0%
25	8	73.3%
25	10	93.3%

3.3.8 Primer design for the shear flow stress response gene *CCL2*

The chemokine (C-C motif) ligand 2 (*CCL2*) gene is upregulated in epithelial cells under shear flow stress (Flores, Battini et al. 2010). The qPCR primer was designed using the Primer-BLAST tool (*www.ncbi.nlm.nih.gov/tools/primer-blast/*). Multiple primer pairs were designed with a length range of 100-400 bp and appropriate GC content to provide a sufficient thermal window for efficient annealing. The selected primers were first evaluated at the bulk sample level. Optimized primer oligos for single-cell analysis of the target genes were obtained from Fisher Scientific (Pittsburgh, PA, USA). The selected genes and their corresponding primers are listed in Table 3.

Table 3 Primer sequences used in the study

Gene	Description	Sequence	GenBank Access No.
28S	Homo sapiens RNA, 28S ribosomal 1 (*RN28S1*)	F: CCGCTGCGGTGAGCCTTGAA R: TCTCCGGGATCGGTCGCGTT	NR_003287.2
ACTB	Homo sapiens actin, beta mRNA (*ACTB*)	F: CATGTACGTTGCTATCCAGGC R: CTCCTTAATGTCACGCACGAT	NM_001101.3
GAPDH	Homo sapiens glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*)	F: TGTTGCCATCAATGACCCCTT R: CTCCACGACGTACTCAGCG	NM_002046.3
Hsp70	Homo sapiens heat shock 70kDa protein 1A (*HSPA1A*), 1B (*HSPA1B*), 2 (*HSPA2*)	F: CGACCTGAACAAGAGCATCA R: AAGATCTGCGTCTGCTTGGT	NM_021979.3 NM_005346.4 NM_005345.5
CCL2	Homo sapiens chemokine (C-C motif) ligand 2	F: CATCTGGCTGAGCGAGCCCT R: GCTGAGCGAGCCCTTGGGGA	NM_002982.3
Turbo GFP	From TurboGFP vector (Sigma Aldrich)	F: AGGACAGCGTGATCTTCACC R: CTTGAAGTGCATGTGGCTGT	N/A

3.3.9 On-chip lysis of single cells in microwells

For on-chip lysis of single cells, individual cells were loaded into microwells using the pick-and-place single-cell manipulation platform (Anis, Holl et al. 2010, Kelbauskas, Ashili et al. 2012). Glass substrates containing 3 × 3 arrays of microwells

were simply placed in a Petri dish, rather than glued to the bottom of the dish. A total of 9 cells were loaded into the 3 × 3 array of microwells (1 cell/well). The cells were incubated in Keratinocyte SFM, at 37 °C under 5% CO_2 for 16-24 hours. The cells in the microwells were visually inspected under the microscope prior to the experiment for occupancy and/or morphological abnormalities. It was observed that after incubation, an average of 7 out of 9 cells were in the microwells due to cell motility, with the remaining 2 cells usually located in the interstitial area just outside of the microwells. The microwell chips were picked up and placed into a 1.5-mL Eppendorf tube containing 200 μL of RNA Lysis Buffer (Zymo Research, Orange, CA, USA) using tweezers. The cells on the chip were lysed in the tube for 30 seconds. After that, the microwell chip was taken out of the tube, and RNA isolation from the lysis buffer was immediately performed.

3.3.10 RNA isolation, reverse transcription, and qPCR

Single cell RNA isolation, reverse transcription, and qPCR were performed as previously described (Gao, Zhang et al. 2011, Zeng, Wang et al. 2011). The 1.5-mL microcentrifuge tubes, each containing an individual cell in the cap, were spun down at 4 °C and 17,000 g for 10 minutes. After 160 μL of medium was taken out, 320 μL of RNA Lysis Buffer from the ZR RNA MicroPrep Kit (Zymo Research, Orange, CA, USA) was added into the tube. The RNA from each cell was extracted using the ZR RNA MicroPrep Kit according to the manufacturer's instructions. A total volume of 6 μL of extracted RNA solution was obtained from the isolation steps and immediately used for the reverse transcription step. A total volume of 10 μL of the cDNA synthesis mixture contained the following reagents: 2 μL of 5 × VILO Reaction Mix (Invitrogen, Carlsbad,

CA, USA), 1 μL of 10 × SuperScript Enzyme Mix (Invitrogen, Carlsbad, CA, USA), including SuperScript III RT, RNaseOUT Recombinant Ribonuclease Inhibitor, and a proprietary "helper" protein (Invitrogen, Carlsbad, CA, USA), 6 μL of total RNA from a single cell, and 1 μL of DEPC-treated water (Ambion, Austin, TX, USA). The contents in each tube were gently mixed and spun down, and the cDNA synthesis was performed in the following thermal steps: a) 25 °C for 10 min, b) 42 °C for 60 min, and c) 85 °C for 5 min to inactivate the reverse-transcriptase. The cDNA obtained from these reactions was stored at -20 °C until further use.

Prior to qPCR, 10 μL of cDNA obtained from each single cell was diluted by adding 20 μL of DEPC-treated water (Ambion, Austin, TX, USA) for detecting multiple genes from a single cell with technical replicates. The qPCR runs were conducted using the following reagent mixtures: 5 μL of EXPRESS SYBR GreenER qPCR SuperMix Universal (Invitrogen, Carlsbad, CA, USA), 1 μL of each primer (4 μM), 0.1 μL of ROX Reference Dye (25 μM) (Invitrogen, Carlsbad, CA, USA), 2 μL of diluted cDNA ($1/15^{th}$ of the cDNA in 30 μL obtained from each single cell) and 0.9 μL of DEPC-treated water (Ambion, Austin, TX, USA). For negative controls, 2 μL of DEPC-treated water was used instead of cDNA. The thermal cycling profile was set up as follows: one cycle at 95 °C for 10 min; 40 cycles consisting of 95 °C for 15 s; 60 °C for 1 min; and 80 °C for 10 s with signal detection; melt-curve analysis at 60 °C for 1 min and the temperature increased in 0.3 °C increments to 95 °C, then at 95 °C for 15 s. The experiments were run on a StepOne Real Time PCR System (Applied Biosystems, Carlsbad, CA, USA). Data analysis was carried out using the StepOne software (Applied Biosystems, Carlsbad, CA, USA).

3.3.11 Lentiviral transfection of cells

The CP-D and EPC-2 cell lines were tagged with *TurboGFP* and *TagFP635*, respectively, to distinguish different cell types in one microwell using the single-cell manipulation platform equipped with an epi-fluorescence imaging mode. CP-D cells were transfected with Lentiviral MISSION® pLKO.1-puro-UbC-*TurboGFP*™ Positive Control Transduction Particles (Sigma-Aldrich, St Louis, MO, USA), which contained a gene encoding *TurboGFP* under the control of the UbC promoter at a multiplicity of infection (MOI) of 2 following manufacturer's instructions. Similarly, EPC-2 cells were transfected with lentiviral MISSION® pLKO.1-puro-UbC-*TagFP635*™ Positive Control Transduction Particles (Sigma-Aldrich, St Louis, MO, USA), which contained a gene encoding *TagFP635* under the control of the UbC promoter at a MOI of 2. 96 hours post-infection, the cells were imaged using a Nikon C1si (Nikon Inc., Melville, NY, USA) confocal microscope to inspect the expression of cytosolic *TurboGFP* in CP-D cells and *TagFP635* in EPC-2 cells. After the culture was expanded into 75 cm^2 flasks, a puromycin kill curve experiment was performed to determine the minimum concentration of puromycin to cause 0% viability of treated cells. 1.0 μg/mL and 0.5 μg/mL of puromycin were found to effectively kill the CP-D and EPC-2 cells, respectively, in which the *TurboGFP* or *TagFP635* was not successfully expressed after 96 hours. CP-D and EPC-2 cells were grown in Keratinocyte serum-free medium containing puromycin (1.0 μg/mL for CP-D cells, 0.5 μg/mL for EPC-2 cells) for five passages, and then grown in normal Keratinocyte SFM. The expression of cytosolic *TurboGFP* or *TagFP635* was retained in CP-D and EPC-2 cells, respectively, when checked under a Nikon C1si confocal microscope.

3.3.12 Fluorescence-assisted single cell harvesting

To distinguish and harvest different types of cells (CP-D/*TurboGFP* and EPC-2/*FP635*) from the co-culture in microwells, a fluorescence-assisted single cell harvesting platform was developed. To this end, a mercury arc epi-fluorescence illumination lamp and a cooled CCD camera were installed onto the pick-and-place single-cell manipulation platform equipped with appropriate excitation/emission filters. A LabVIEW program (National Instruments, Austin, TX, USA) was written for adjusting exposure time and gain settings of the fluorescent microscope. After cells were trypsinized as described in the Single-cell Collection Section, they were visualized under transmitted light illumination. Co-cultured cells were then imaged with the camera in epi-fluorescence mode using the different filter cubes. CP-D/*TurboGFP* and EPC-2/*FP635* cells can be easily distinguished in green and red channels, respectively. The two types of cells were then collected separately into 1.5-mL Eppendorf tubes as described in Section 3.3.5 Single-cell Collection.

3.4 Results and discussion

3.4.1 Comparison between in-situ direct lysis and "pick-and-place" harvesting of single cells in microwells

In order to perform end-point gene transcription analysis of single cells, the total RNA from individual cells needs to be harvested with minimal loss. First, the ability of different single-cell harvesting methods to preserve and recover the maximum amount of total RNA using several different buffers was tested. The in-situ direct lysis of single cells in microwells (Figure 3-1) contained three steps: (1) lysing single cells in microwells; (2) aspirating the lysate using a micropipette and a custom high-precision pump; (3) dispensing the lysate into the cap of a microcentrifuge tube for RT-qPCR analysis. The in-situ direct lysis method was initially used to eliminate the steps of cell

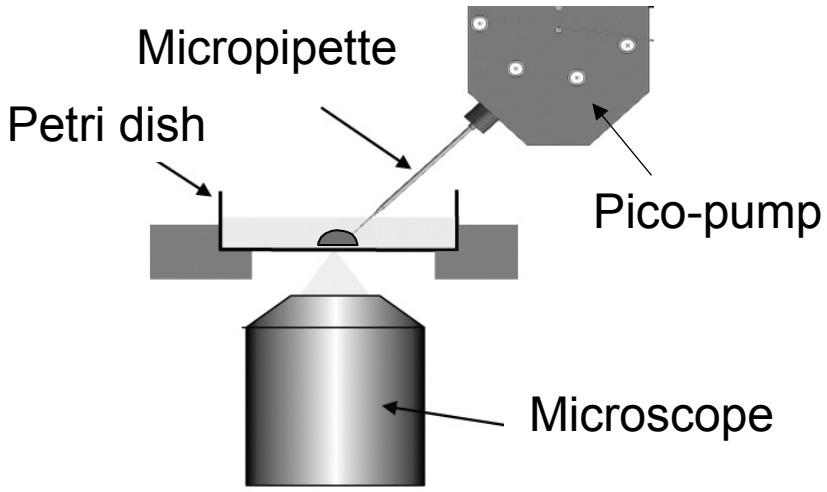


Figure 3-4 Schematic view of the single cell manipulation platform
The single-cell manipulation platform uses a micropipette controlled by a piezoelectric pump. Single cells can grow in microwells which are glued to the bottom of the Petri dish. The micropipette is used for aspirating and dispensing single cells.

detaching and transferring from microwells to analysis vials and avoid potential stress during the cell detachment process. As a comparison to the first method, a "pick-and-place" cell harvesting method (Zeng, Wang et al. 2011, Kelbauskas, Ashili et al. 2012) using the single-cell manipulation platform (Anis, Holl et al. 2010) developed in the Center for Biosignatures Discovery Automation (D. Meldrum, Director) was tested (Figure 3-2, Figure 3-4). The method combines trypsinization and shear flow to detach single cells from the bottom of the microwells and collect them for downstream analysis, e.g. RT-qPCR (Figure 3-5). Esophageal epithelial cells (CP-A cell line) in microwells were treated with 0.05% Trypsin for 6 minutes until they were partially detached as judged from the change in the cellular morphology from fully stretched to a more spherical shape. Single cells were then collected from the microwells using the micropipette in the single-cell manipulation platform using 12.5 pL/s flow rate.

Using the in-situ direct lysis method, the single cells were immediately lysed in microwells by adding the RNA lysis buffer into the microwell. The addition of the lysis

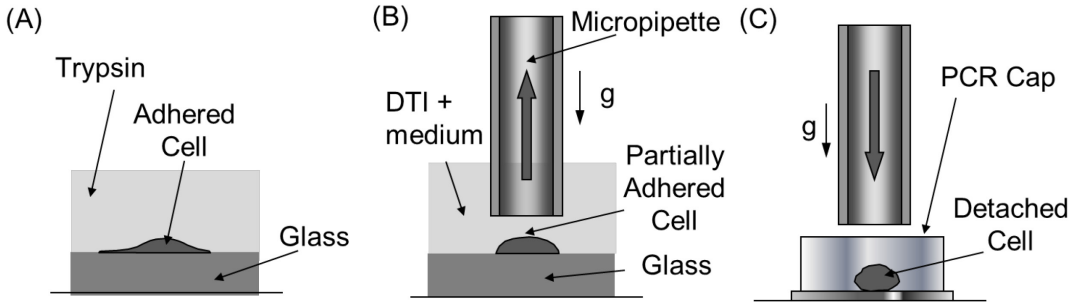


Figure 3-5 Single-cell harvesting steps
The single-cell harvesting procedure contains three major steps: (A) trypsinization to partially detach the cell from the substrate; (B) trypsin deactivation with trypsin inhibitor (DTI) and cell aspiration into the micropipette tip; (C) transfer of the cell into the cap of a PCR tube for downstream analysis.

buffer resulted in substantial morphological changes. It was observed that the lysate diffused out of the microwell prior to being collected rendering this method inappropriate for collecting total RNA from single cells. Compared with in-situ direct lysis, single-cells can be harvested without physical tearing or morphological changes to the cell using the pick-and-place harvesting method. Therefore, the pick-and-place harvesting was further developed and optimized in later experiments.

3.4.2 Usage of cellular RNA preservation solution for single-cell harvesting

Stabilization of the total RNA is an important aspect of gene transcription assays. To preserve the RNA quality, RNA degradation by RNase should be minimized during cell harvesting and until RNA extraction. Three different solutions - RNA Later®, RNA Lysis Buffer and Keratinocyte serum-free medium - were tested for their ability to preserve RNA in single cells harvested at room temperature. The RNA Later® solution (Ambion, Austin, TX, USA) is a concentrated salt solution (25 mM Sodium Citrate, 10 mM EDTA, 70 g ammonium sulfate/100 ml solution, pH 5.2) that rapidly permeates tissues to stabilize and protect cellular RNA (Lader 2001). RNA Later® was mixed with Keratinocyte serum-free medium at a 5:1 v/v ratio in my experiments. RNA Lysis Buffer (Zymo Research, Orange, CA, USA) is composed of guanidinium thiocyanate, which can both lyse cells and deactivate RNases by denaturing them (Forman and Jia 2012). Keratinocyte serum-free medium was used as a control for RNA preservation experiments. In the harvesting experiments, 200 μL of the corresponding solution was added into the cap of a microcentrifuge tube. After the single cell was aspirated into the micropipette tip, it was directed into the cap and dispensed into the liquid solution in the cap. The tube was immediately closed and placed on dry ice.

First, the total RNA was extracted from single cells using ZR RNA MicroPrep Kit (Zymo Research) and expression levels of the *28S* and actin β (*ACTB*) housekeeping genes were measured. Both genes are highly expressed in human cells and show low cell-to-cell copy number variability. They can easily be detected using qPCR and used as references for total RNA extraction efficiency in comparison studies. The amount of *28S* rRNA and *ACTB* mRNA was measured with reverse transcription quantitative PCR (RT-qPCR) using the Superscript VILO cDNA synthesis kit and EXPRESS SYBR GreenER qPCR SuperMix Kit (both kits are from Invitrogen, Carlsbad, CA, USA). Compared with the quantification cycle (C_q) values of the two genes in the group where only cell medium was used for storing the harvested cells, the C_q values of the two genes in both the RNA Later® and the Lysis buffer groups are higher. The differences in C_q values of *28S* and *ACTB* genes are shown in Figure 3-6. The C_q of both genes in the lysis buffer groups are significantly different from that of the medium group ($p = 0.02$ for *28S*, $p = 0.04$ for *ACTB*, calculated using the non-parametric Mann-Whitney statistical significance test, $\alpha = 0.05$, two-tailed). Higher amounts of *28S* and *ACTB* mRNA in the medium group indicate that the medium outperforms RNA Later® and lysis buffer in preserving RNA. RNA Later® contains ammonium sulfate which can permeate the cellular membrane and lead to a leakage of cellular components including RNA (Park, Yu et al. 2006). In addition, it was not possible to spin down picogram levels of single cell RNA from RNA Later® solution, which may have further reduced the RNA extraction efficiency. RNA lysis buffer, on the other hand, lyses cells almost instantaneously, potentially exposing the picogram levels of single cell RNA to environmental RNases. Another RNA-preserving agent, RNAstable (Biomatrica, San Diego, CA, USA), was tested by adding

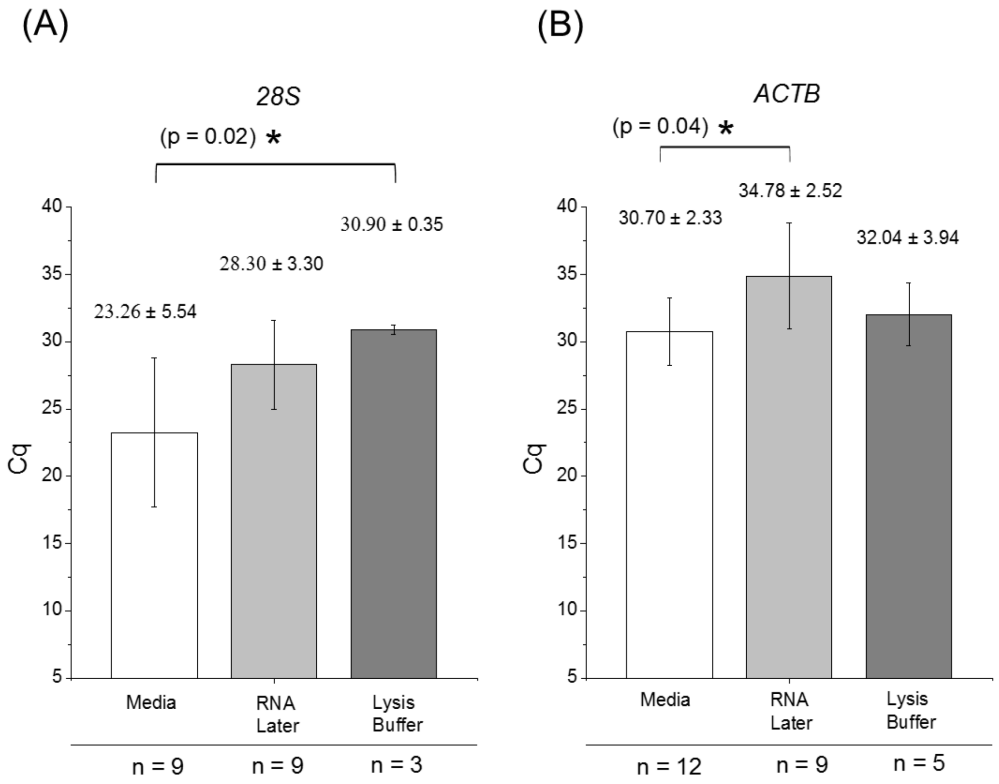


Figure 3-6 Comparison of RNA preservation levels between medium, RNALater and RNA Lysis Buffer

The quantification cycle (Cq) values are higher in the lysis buffer group for the 28S (Panel A) and higher in the RNALater group for the ACTB gene (Panel B) as compared with the medium group. The corresponding Cq mean values are shown on top of the bars to compare the three conditions. Errors and error bars are corresponding standard deviations. The number of cells analyzed (n) for each harvesting condition is shown at the bottom of the graphs. Each qPCR reaction was run with three technical replicates. The difference between the lysis buffer and medium preservation group is statistically significant for the 28S gene and between the RNALater and medium group for the ACTB gene (both tested with the two-tailed Mann-Whitney test, $\alpha = 0.05$). Therefore, dispensing cells in cell culture medium can better preserve mRNA.

200 µL of RNAstable LD into the cap prior to harvesting. Results (data not shown) indicated that RNAstable LD is inferior in preserving cellular RNA as compared with the medium. Based on these results it was found that the Keratinocyte cell growth medium

exhibits the highest level of protection against RNA degradation/loss among the four solutions. Single-cell collection method was optimized by using 200 μL of Keratinocyte SFM in the 1.5-mL microcentrifuge tubes.

3.4.3 Low temperature harvesting for preserving cellular RNA

Low temperatures are known to inhibit enzymatic activity of cellular proteins, including RNase. To determine whether lowering the temperature can facilitate RNA preservation due to RNase inactivation, two ice packs were placed on the stage of the single-cell manipulation platform for 15 min before the harvesting experiment to pre-cool

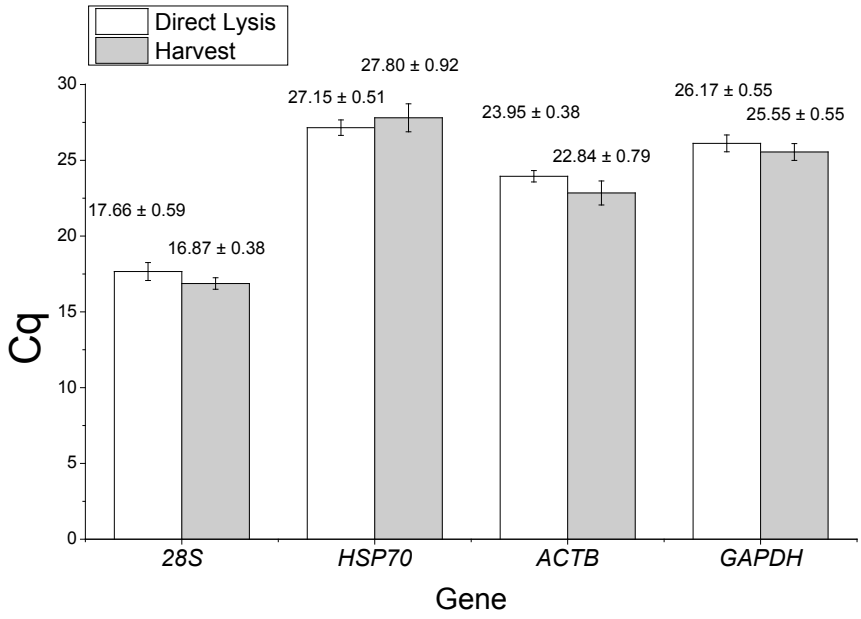


Figure 3-7 Comparison between low temperature and room temperature harvesting conditions for preserving cellular RNA
Three individual cells were harvested for each condition and each qPCR reaction was run with three technical replicates. The differences between the Cq mean values (shown above each box with S.D. as error) between the two groups in the *28S*, *ACTB* and *HSP70* genes are not statistically significant (Mann-Whitney test, $\alpha = 0.05$, two-tailed). Error bars represent the standard deviations.

the platform. During the entire harvesting procedure, the Petri dish containing cells in the microwells was surrounded by the ice packs. Room temperature harvesting without ice pack cooling was used as the control. RT-qPCR analysis was performed on three genes- *28S*, *ACTB*, and *HSP70* from the harvested cells to assess the preservation efficiency. The *HSP70* gene was chosen due to its role in cold response (Fujita 1999) and cellular stress response in general (Guzhova and Margulis 2006, Murphy 2013). Gene transcription levels of all three genes in the ice cooled harvesting group do not exhibit statistically significant changes compared with room temperature harvesting control (Figure 3-7). This indicates that the mRNA of the studied genes is stable under both cool and room temperature conditions, while other genes, such as *ACTB*, may be more sensitive to degradation at room temperature. Furthermore, the unchanged level of *HSP70* mRNA indicated that the low temperature condition did not introduce an additional stress factor to the cells. Even though no statistically significant differences were observed between the room temperature and the ice-cooled conditions for the studied genes, the ice-cooled condition was used throughout the study to avoid potential degradation of the total RNA at room temperature.

3.4.4 The effects of flow-rate and trypsinization time on harvesting success rate and RNA preservation

In order to detach adherent cells from the microwell glass surface with minimal perturbation, a combination of mechanical force and enzymatic digestion was used. Shear flow through the micropipette tip can aspirate single cells out of the microwell; however, applying shear flow can result in damages to the plasma membrane and loss of the cellular RNA. Trypsin is a traditional protease which can cleave membrane adhesion

proteins, primarily integrins, and detach the cells from the substrate. Excessive trypsinization, however, especially with regard to trypsinization time, can alter the gene expression profile and affect cell viability (Wassarman and Keller 2003, Rasmussen, Frøbert et al. 2011). To minimize potential perturbations to the gene transcription profile and the loss of total RNA during harvesting, a combination of both methods was employed. First, trypsinization followed by applying a shear flow was performed to harvest single cells. Because either step can damage the cell integrity and/or mRNA profile, the approach was optimized by testing different combinations of mechanical, chemical, and temporal parameters. The purpose was to examine how the trypsinization time, trypsin concentration, and flow rate affect the harvesting success rate (HSR) and gene transcription levels. The main goal was to identify the lowest flow-rate and shortest trypsinization time needed to achieve reliable detachment of cells with the highest HSR while causing minimal changes in gene transcription levels.

A total of nine different conditions (Table 2) were tested, including three rates of the shear flow (2.5 pL/s, 12.5 pL/s, 25 pL/s) and three trypsinization times (5 min, 8 min, 10 min).

Fifteen cells were treated under each condition and the successfully harvested cells were used for RT-qPCR analysis (Table 4). The success rate of cells harvested under each condition without taking into account cell stress levels demonstrates that trypsinizing cells for 10 minutes and harvesting at a flow rate of 25 pL/s has the highest success rate of 93.3% (Table 2).

Table 4 Two-way multivariate ANOVA of the effects of flow rate and trypsinization time on gene transcription analysis

(* Significant at p = 0.05 level)

Gene	Flow rate (FR)	Trypsinization time (TT)	FR and TT combined
HSP70	p = 0.367	p = 0.000*	p = 0.493
ACTB	p = 0.014*	p = 0.082	p = 0.048*
28S	p = 0.000*	p = 0.000*	p = 0.000*
CCL2	p = 0.022*	p = 0.000*	p = 0.018*

To assess the stress levels resulting from the different harvesting conditions, the mRNA levels of *28S*, *ACTB*, *HSP70* and *CCL2* genes were analyzed in harvested cells using RT-qPCR. The cDNA extracted from single cells obtained for each harvesting condition was pooled and qPCR was performed on them. The results show that the transcription levels of the *HSP70*, *28S* and *CCL2* genes are generally higher in cells harvested after 5 min trypsinization as compared to cells harvested after 8 or 10 min of trypsinization. The difference in their expression levels in cells between 5 min, 8 min and 10 min trypsinization groups is statistically significant (Table 4, Table 5). Moreover, the transcription levels of the *ACTB*, *28S* and *CCL2* genes are markedly affected by the flow rate. In general, the flow-rate of 12.5 pL/s yields the lowest C_q values for these genes. Different combinations of the flow rate and trypsinization time also significantly affect the gene transcription levels of the *ACTB*, *28S* and *CCL2* genes. The two-way multivariate ANOVA test ($\alpha = 0.05$) performed on the Cq values as a function of

trypsinization time (TT) or flow rate (FR) shows (Table 4) that both the trypsinization time and flow rate have a statistically significant impact on the transcription levels of these genes. If the cells are trypsinized for a shorter time, a higher flow rate needs to be applied to retrieve the cells and vice versa. Interestingly, the p-values of the two factors (Table 4, combined FR and TT p-values) indicate that enzymatic cleavage and mechanical shear flow have a collective effect on the transcription of the *ACTB*, *28S* and *CCL2* genes. Since shorter trypsinization times and low flow rates are desired for preserving cellular RNA, the trypsinization time was limited to 5-7 minutes and a flow rate of 12.5 pL/s was utilized for single cell harvesting.

Table 5 The effects of flow rate and trypsinization time on the mRNA preservation/retrieval efficiency

Flow Rate (nL/s)	Trypsini-zation Time (minutes)	*ACTB*	*28S*	*HSP70*	*CCL2*
		Cq ± S.D.	Cq ± S.D.	Cq ± S.D.	Cq ± S.D.
2.5	5	32.27 ± 0.19	13.52 ± 0.27	29.64 ± 0.38	30.16 ± 0.37
2.5	8	34.97 ± 3.04	15.01 ± 0.05	36.05 ± 2.03	36.69 ± 2.51
2.5	10	35.33 ± 2.10	13.60 ± 0.90	31.00 ± 1.79	33.90 ± 0.56
12.5	5	33.60 ± 1.05	18.23 ± 0.62	29.30 ± 0.88	31.43 ± 1.13
12.5	8	33.22 ± 3.13	19.45 ± 0.03	33.52 ± 0.40	36.50 ± 0.80
12.5	10	not detected	18.94 ± 0.59	31.45 ± 1.45	34.38 ± 0.00
25	5	26.69 ± 0.74	13.81 ± 0.14	29.30 ± 0.62	32.79 ± 0.83
25	8	34.87 ± 0.31	29.10 ± 1.20	34.21 ± 1.77	36.49 ± 0.62
25	10	27.93 ± 0.39	14.81 ± 0.13	30.03 ± 1.80	37.07 ± 0.15

3.4.5 Comparison between on-chip lysis of cells in microwells and harvesting cells from microwells

With the preservation medium, temperature, trypsinization time and flow-rate conditions optimized, it was important to determine whether the harvesting procedure itself under those conditions causes any observable RNA loss or changes in gene transcription. To this end a comparison was made between the amounts of mRNA extracted from cells that were harvested with those that did not undergo harvesting. Individual cells were directly lysed in microwells on the chip and this condition was used as the untreated control group. If no significant differences in gene transcription levels of the selected genes can be detected between the control group and the harvested group, the harvesting procedure can be considered suitable to collect cells for gene transcription analysis.

In the on-chip lysis experiment, 9 single cells were loaded into a 3 × 3 array of microwells made in fused silica chips, and the chips were placed in a Petri dish. The cells on the chip were incubated overnight. After being visually inspected under the microscope for cell occupancy in the microwells, the glass chips with cells adhered to the well bottom were placed into a 1.5-mL microcentrifuge tube containing 200 μL of RNA Lysis Buffer. In the harvesting experiment, single-cells were harvested from the microwells on the cooled stage, using 12.5 pL/s flow-rate and 6 min trypsinization time, and placed into the cap of a microcentrifuge tube containing 200 μL of cell growth medium. The mRNA levels of the *28S*, *ACTB*, *GAPDH* and *HSP70* genes in cells lysed on the chip or harvested were analyzed using RT-qPCR.

C_q values of the four genes were compared between the on-chip lysis and harvesting groups to assess the capability of the two methods to preserve cellular RNA. The C_q values of *28S*, *ACTB*, *GAPDH* and *HSP70* genes in harvested cells were found to be very close to those measured in on-chip lysed cells (Figure 3-8). The C_q values of the *28S*, *ACTB* and *GAPDH* genes in harvested cells are slightly lower (ΔC_q = -0.79 for *28S*, -1.10 for *ACTB* and -0.57 for *GAPDH*), indicating a larger amount of mRNA in the harvested group; while the C_q values of the *HSP70* gene in harvested cells are slightly higher than in on-chip lysed cells (ΔC_q = 0.65), indicating a smaller amount of RNA in harvested cells. However, the Mann-Whitney test showed that the C_q differences between the two groups in all of the genes are not statistically significant. This suggests that the parameters used for harvesting preserve cellular RNA efficiently and do not induce detectable changes in the mRNA profile or in the amount of the total RNA extracted from the cell.

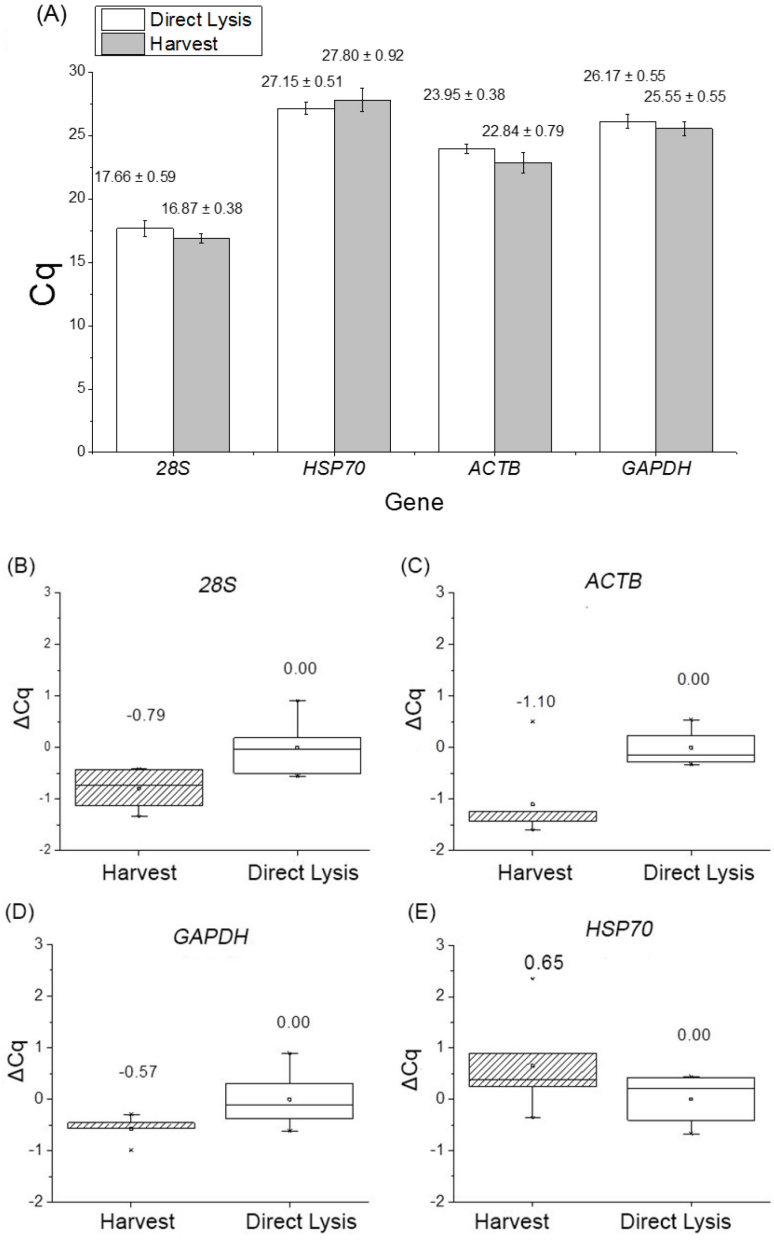


Figure 3-8 Comparison of gene transcription levels between the on-chip lysis of cells in microwells and harvesting cells from microwells

(A) C_q values of the *28S*, *ACTB*, *GAPDH* and *HSP70* genes. The corresponding C_q mean values are shown on top of the bars. Error bars show the standard deviations. ΔC_q between the two groups in all of the genes (Panel B-E) are not statistically significant (Mann-Whitney test, n=6, $\alpha = 0.05$, two-tailed).

3.4.6 Fluorescence-assisted single cell harvesting

Cell-cell communications are critical to regulating various core cellular responses, such as metabolism and homeostasis. Understanding gene transcription changes as a result of intercellular interactions in cells at different stages of pre-malignant progression may help discover new cancer biosignatures.

To enable harvesting of individual cells of different types from co-cultured cell populations, a fluorescence-assisted single-cell harvesting modality was added to the method. To perform cell harvesting with minimal effect on gene transcription levels, the harvesting parameters optimized in the cell stress as presented above were used. Using this approach, one can distinguish co-cultured cells of different types utilizing cell-type specific fluorescent markers and separately collect individual cells from a single microwell.

To this end, two cell lines were produced, each expressing a different fluorescent protein. Normal epithelial EPC-2 cells were transfected with Lentiviral vectors expressing cytosolic FP635 to establish the EPC-2/FP635 cell line. Dysplastic Barrett's esophagus CP-D cells were transfected with Lentiviral vectors expressing cytosolic TurboGFP to establish the CP-D/TurboGFP cell line. Co-cultured EPC-2/FP635 and CP-D/TurboGFP cells could easily be distinguished under the fluorescence microscope (Figure 3-9).

The single-cell manipulation platform (Anis, Holl et al. 2010, Kelbauskas, Ashili et al. 2012) was equipped with an epi-fluorescence illumination source and excitation/emission filter sets for TurboGFP and FP635. With the help of a custom-written program (LabVIEW) for fluorescence-assisted cell harvesting, EPC-2/FP635 cells

and CP-D/TurboGFP cells can be identified within the co-culture in the microwells. They can be detached and collected separately using the optimized harvesting conditions.

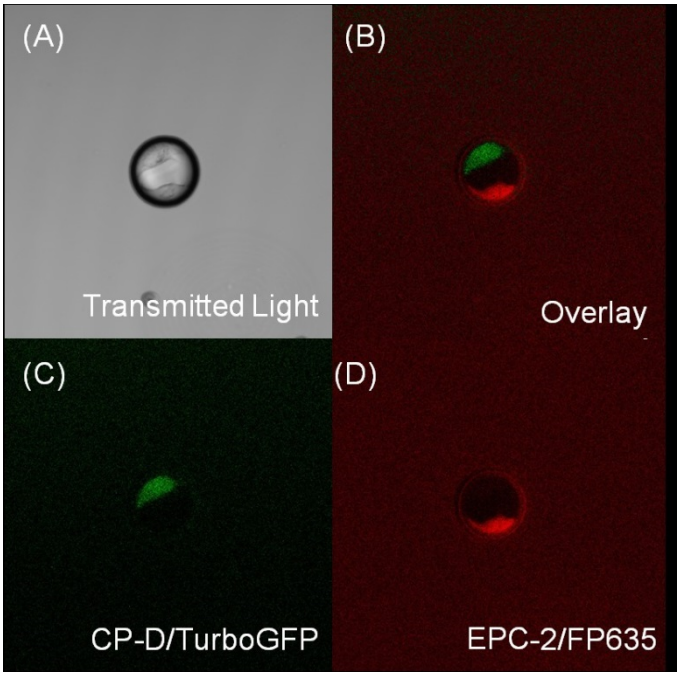


Figure 3-9 Harvesting single cells from a co-culture
One CP-D/TurboGFP cell and One EPC-2/FP635 cell were co-incubated in microwells. Micrographs of: (A) bright field; (B) overlay of green and red channel with bright field channel; (C) CP-D/TurboGFP cell in green spectral channel, (D) EPC-2/FP635 in red spectral channel. Co-cultured EPC-2/FP635 and CP-D/TurboGFP cells can be distinguished and harvested using the fluorescence-assisted single cell harvesting platform.

In order to demonstrate the ability of the technique to distinguish specific signatures of single cells, the presence of TurboGFP transcripts in harvested cells was measured using RT-qPCR. It was expected that the Cq values of the Turbo-GFP gene in CP-D/TurboGFP cells would be significantly lower compared with those in EPC-2/FP635 cells.

The amplification plots of RT-qPCR showed the presence of TurboGFP in CP-D/TurboGFP cells and marked differences in Cq in EPC-2/FP635 cells (Figure 3-10A). Normalized Cq values using CP-D/TurboGFP cells as the reference demonstrated a significant difference in signal between two types of cells as determined by the Wilcoxon test ($\alpha = 0.05$, two-tailed) with $p = 0.0009$ (Figure 3-10B).

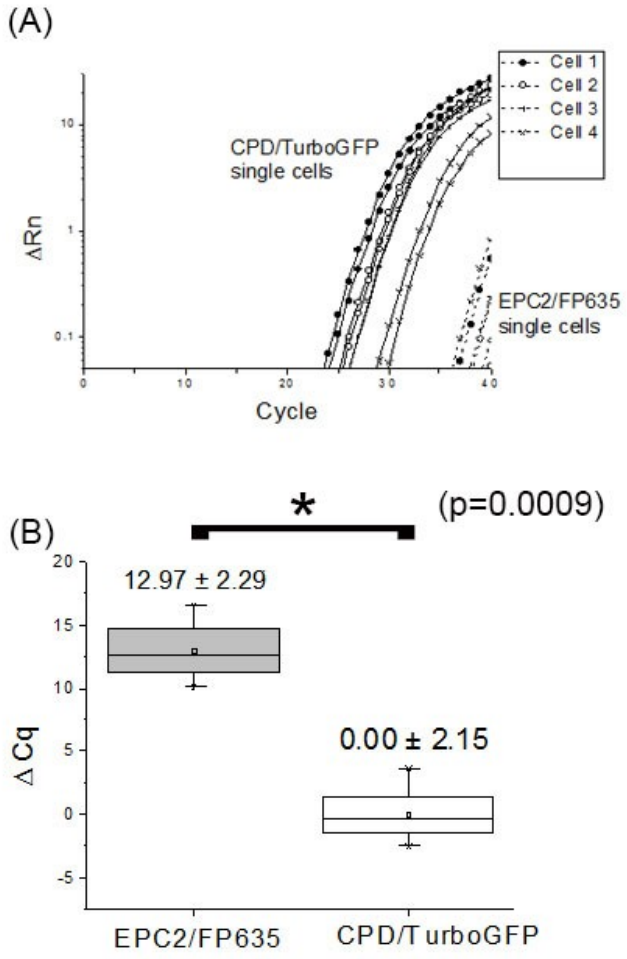


Figure 3-10 RT-qPCR analysis of TurboGFP genes in CPD/TurboGFP and EPC2/FP635 cells

Different cell types can be reliably distinguished using the fluorescent assisted harvesting method as proven by the RT-qPCR results. Quantitative cycle values on the amplification plots (A) and ΔCq analysis using CPD/TurboGFP cells as the control (B) both demonstrate the presence of TurboGFP in harvested CPD/TurboGFP cells but not in EPC2/FP635 cells.

3.5 Conclusion

In this study, a method was developed for retrieving adherent cells from substrates with minimal disruption and perturbation. The method features combined enzymatic cleavage and mechanical force applied by a piezo-pump on a single cell manipulation station. Reagents for RNA preservation, temperature settings, flow rate, and trypsinization time were also optimized to minimize gene transcription profile changes brought by harvesting. Using this method, gene transcription levels were analyzed in both harvested and on-chip directly-lysed single cells. The results showed conclusively that the harvesting method that was developed and optimized can preserve the RNA profiles in the cells retrieved from microwells. The application was expanded to fluorescence-assisted single cell harvesting from a co-culture of different cell types. This method provides an approach to transfer adherent single cells from cell culture to any downstream end-point analysis or re-culture. It enables researchers to retrieve adherent single cells without perturbation and thus has the potential to become a broadly applicable tool in the growing field of single-cell analysis.

3.6 Acknowledgments

I thank Dr. Brian Reid's group from Fred Hutchinson Cancer Research Center for providing cell lines, Jordan Yaron for providing TurboGFP primers and Patti Senechal-Willis for providing cell-culture. This work was supported by the following grants to Dr. Deirdre R. Meldrum from the NIH National Human Genome Research Institute, Centers of Excellence in Genomic Sciences, Grant Numbers 5 P50HG002360 and 3P50HG002360-10S1.

CHAPTER 4

QUANTITATIVE SINGLE-CELL GENE EXPRESSION MEASUREMENTS OF MULTIPLE GENES IN RESPONSE TO HYPOXIA TREATMENT

4.1 Abstract

Cell-to-cell heterogeneity in gene transcription plays a central role in a variety of vital cell processes. To quantify gene expression heterogeneity patterns among cells and to determine their biological significance, methods to measure gene expression levels at the single-cell level are highly needed. Reported here is an experimental technique based on the DNA-intercalating fluorescent dye SYBR green for quantitative expression level analysis of up to ten selected genes in single mammalian cells. The method features a two-step procedure consisting of a step to isolate RNA from a single mammalian cell, synthesize cDNA from it, and a qPCR step. The method was applied to cell populations exposed to hypoxia, quantifying expression levels of seven different genes spanning a wide dynamic range of expression in randomly picked single cells. In the experiment, 72 single Barrett's esophageal epithelial (CP-A) cells, 36 grown under normal physiological conditions (controls) and 36 exposed to hypoxia for 30 min, were randomly collected and used for measuring the expression levels of 28S *rRNA*, *PRKAA1*, *GAPDH*, *Angptl4*, *MT3*, *PTGES*, and *VEGFA* genes. The results demonstrate that the method is sensitive enough to measure alterations in gene expression at the single-cell level, clearly showing heterogeneity within a cell population. Presented are technical details of the method development and implementation, and experimental results obtained by use of the procedure. The advantages of this technique are expected to facilitate further developments and advances in the field of single-cell gene expression profiling on a

nanotechnological scale, and eventually to serve as a tool for future point-of-care medical applications.

4.2 Introduction

Cell-to-cell heterogeneity in gene transcription plays a central role in a variety of vital cell processes, including differentiation (Losick and Desplan 2008, Furusawa and Kaneko 2009), stimulus response (Levsky, Shenoy et al. 2002), survival (Cohen, Geva-Zatorsky et al. 2008, Fraser and Kaern 2009), and carcinogenesis (Elowitz, Levine et al. 2002). Therefore, studies of the molecular mechanisms responsible for intercellular variability in gene expression levels at the single-cell level are expected not only to provide critical insights into core cellular processes but also to pave the way for new, more effective disease prevention and treatment strategies. The vast majority of currently existing experimental techniques for gene expression profiling are based on analysis of bulk samples containing 10^5–10^7 or more cells. Inherent to samples of that size is the ensemble-averaging of the results over a large number of cells, hiding key information emanating from individual cells. Thus, bulk techniques are rendered unsuitable for intercellular heterogeneity studies. The emerging importance of cell population heterogeneity imposes a demand for reliable gene transcription profiling techniques specifically tailored for individual cells. Measuring gene expression at the single-cell level is challenging because of the small amounts of total available mRNA (~1 pg). The large dynamic range of expression levels among genes is another hurdle that must be overcome by developing highly sensitive, specific, and reproducible detection strategies. The most reliable quantitative approach suitable for single-cell studies is based on reverse transcription (RT) without pre-amplification followed by quantitative polymerase chain reaction (qPCR). Besides enabling highly quantitative measurement of mRNA copy numbers, RT-qPCR theoretically allows detection of a single copy of mRNA. Several

experimental techniques based on RT-qPCR and specialized for quantitative gene expression profiling in individual cells have been reported in the literature (Hartshorn, Anshelevich et al. 2005, Hartshorn, Eckert et al. 2007, Bengtsson, Hemberg et al. 2008, Diercks, Kostner et al. 2009, Taniguchi, Kajiyama et al. 2009, Stahlberg and Bengtsson 2010). Several of these techniques use a single-tube RT-qPCR approach (Bengtsson, Hemberg et al. 2008, Diercks, Kostner et al. 2009, Gong, Ogunniyi et al. 2010), in which all steps, including cell lysis, cDNA synthesis by reverse transcription, and quantitative PCR, are performed in one tube. This reduces the probability of mRNA loss and possible contamination during the sample-handling process. Although advantageous in terms of sample conservation, these single-tube methods can generate only one measurement for each cell, making it impossible to distinguish biological variation from measurement variation. In addition, the single-tube operation limits the number of genes that can be detected per single cell to less than five (Taniguchi, Kajiyama et al. 2009, Gong, Ogunniyi et al. 2010). To overcome these hurdles, multiple-step RT-qPCR procedures were recently developed (Taniguchi, Kajiyama et al. 2009, Joglekar, Wei et al. 2010). One of these protocols is based on a reusable single-cell cDNA library immobilized on beads for measuring the expression of multiple cDNA targets (from several copies to several hundred thousand copies) in a single mammalian cell. The results showed that the measurement error of this method is less than 15.9% among replicates (Bengtsson, Hemberg et al. 2008). Another procedure has been successful in assessing ten mRNA transcripts from a single cell, and each with one technical replicate (Hartshorn, Anshelevich et al. 2005). The procedure could potentially be extended to analyze twenty different mRNAs from a single cell by use of duplex PCR (Joglekar, Wei et al. 2010).

Despite some initial successes (Bengtsson, Hemberg et al. 2008, Taniguchi, Kajiyama et al. 2009, Joglekar, Wei et al. 2010, Stahlberg and Bengtsson 2010), analyzing multiple genes from one single cell by RT-qPCR remains challenging because the total amount of cDNA must be divided into multiple portions, limiting the sensitivity of RT-qPCR product detection.

Molecular oxygen is required for energy production in aerobic organisms. A shortage of oxygen (hypoxia) creates significant stress in cells, to which they respond by several different molecular mechanisms, including reduction in energy demand, cell cycle arrest, production and secretion of survival and angiogenic factors, and so forth. Hypoxia plays a pivotal role in cancer, causing alterations in cellular metabolism, increased resistance to radiation and chemotherapy (Bertout, Patel et al. 2008), and possibly increasing cells' metastatic potential (Lopez-Lazaro 2007). Oxygen deprivation has been shown to cause alterations in stem cell proliferation, differentiation, and pluripotency (Mazumdar, Dondeti et al. 2009). Despite recent advances in understanding cells' responses to hypoxia, the underlying molecular mechanisms remain unclear. More specifically, there is a lack of studies at the single-cell level that could provide deeper insights into hypoxia-driven selection and survival among different cell populations.

To gain better understanding of the molecular mechanisms contributing to the hypoxia response pathways, an experimental technique based on the DNA-intercalating fluorescent dye SYBR green was developed and evaluated. The method enables quantitative expression level analysis of up to ten genes of interest in single mammalian cells.

The technique is a modification of one initially developed for gene expression analysis in single bacterial cells (Gao, Zhang et al. 2011) and has so far not been tested in mammalian cells. The choice of SYBR green for qPCR is based on its relative ease of use, low cost, and suitability for development of high throughput, lab-on-chip procedures for RT-qPCR in single mammalian cells. The procedure presented here enables detection of up to ten genes, each with three technical replicates, from a single cell, with high repeatability (i.e. low standard deviations of quantification cycle, C_q). In addition, the SYBR-based chemistry used in the protocol provides greater flexibility to measure more genes in the future than Taqman-based technology.

The method features a two-step procedure consisting of RNA isolation from a single mammalian cell followed by cDNA synthesis and a qPCR step. The expression levels of multiple genes of interest can be quantified simultaneously in single mammalian cells. The primers were designed for selected gene targets known, on the basis of bulk-cell studies, to be involved in hypoxia response (Arany, Huang et al. 1996, Yoshiji, Gomez et al. 1996, Zhong, Agani et al. 1998, Mu, Brozinick et al. 2001, Wang, Wood et al. 2008, Murata, Yudo et al. 2009, Lee, Natsuizaka et al. 2010). The primers were evaluated at both bulk-cell and single-cell levels. The method was applied to cell populations exposed to hypoxia and the expression levels of ten genes spanning a wide dynamic range of expression in randomly picked, single cells was quantitatively measured. The technical details of method development and implementation are presented along with experimental results obtained by use of the procedure. The results showed significant gene expression level heterogeneity among the analyzed cells for each of the target genes. Detailed interpretation of the observed heterogeneity among cells

from an isogenic cell population and its biological significance requires more experimental data. However, the results presented here demonstrate that the method is sensitive enough to quantify cellular responses at the single-cell level and to reveal gene expression heterogeneity in a cell population. The advantages of this technique will facilitate further developments and advances in single-cell gene-expression profiling.

4.3 Experiments

4.3.1 Cell culture

The Barrett's esophageal epithelial cell line CP-A was cultured using Gibco keratinocyte serum-free medium (SFM) cell growth medium (Invitrogen, Carlsbad, CA, USA), supplemented with hEGF (Peprotech, Rocky Hill, NJ, USA) at 2.5 μg/500 mL, BPE (bovine pituitary extract) at 25 mg/500 mL and penicillin–streptomycin solution (Invitrogen) at 100 units/100 μg/mL. Cells were grown at 37 °C under 5% CO_2. Before experimentation, cells were cultured in a 75-mL flask to approximately 80% confluency. Cells were washed with 1× PBS (Cellgro, Manassas, VA, USA) and detached from the flask with 0.05% (*v*/*v*) trypsin–EDTA (Invitrogen). The trypsinization was blocked by Dulbecco's modified Eagle medium (DMEM) (Invitrogen) supplemented with 5% fetal bovine serum (FBS) (Invitrogen). After trypsinization, cells were centrifuged at 900 rpm for 3 min then resuspended in 1 mL cell-growth medium.

4.3.2 Primer design and selection of gene target

Thirteen genes were chosen for RT-qPCR expression level analysis. *18S rRNA*, *28S rRNA*, *ACTB*, and *GAPDH* were selected as reference genes because they are highly expressed housekeeping genes in mammalian cells. The *28S rRNA* gene is a reliable internal control for comparative transcription analyses under hypoxic conditions (Zhong and Simons 1999). *HSP70* and *HSC70* genes were chosen because *HSP70* mRNA undergoes dramatic changes under stress conditions, whereas *HSC70* does not and can be used as a matched reference gene (Mayer and Bukau 2005). *HIF1α* (Zhong, Agani et al. 1998), *VEGFA* (Yoshiji, Gomez et al. 1996), *PRKAA1* (Mu, Brozinick et al. 2001), *p300* (Arany, Huang et al. 1996), *MT3* (Wang, Wood et al. 2008), *Angptl4* (Murata, Yudo et al.

2009) and *PTGES* (Lee, Natsuizaka et al. 2010) are involved in hypoxia response signaling pathways. Gene sequences were retrieved from GenBank. qPCR primers were designed using the Primer-BLAST, Primer 3 or PrimerExpress V2.0 software packages, or retrieved directly from PrimerBank (*http://pga.mgh.harvard.edu/primerbank/*). For each gene, multiple primer pairs were designed on the basis of their coding regions with amplification product lengths between 100 and 700 bp and annealing temperatures mostly between 60 and 65 °C. The amplification efficiencies of the primers were evaluated at bulk-cell and single-cell levels. Optimized primer oligos for single-cell analysis of the target genes were obtained from Fisher Scientific (Pittsburgh, PA, USA). The selected genes and their corresponding primers are listed in Table 6.

Table 6 Genes and corresponding primers

Gene	GenBank Access No. and Description	Sequence
28S rRNA	NR_003287.2, Homo sapiens RNA, 28S ribosomal 1 (*RN28S1*)	F: CCGCTGCGGTGAGCCTTGAA R: TCTCCGGGATCGGTCGCGTT
18S rRNA	NR_003286.2, Homo sapiens RNA, 18S ribosomal 1 (*RN18S1*)	F: CCCGACCCGGGGAGGTAGTG R: GCCGGGTGAGGTTTCCCGTG
ACTB	NM_001101.3, Homo sapiens actin, beta mRNA (*ACTB*)	F: CATGTACGTTGCTATCCAGGC R: CTCCTTAATGTCACGCACGAT
GAPDH	NM_002046.3, Homo sapiens glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*)	F: TGTTGCCATCAATGACCCCTT R: CTCCACGACGTACTCAGCG
HIF1A	NM_001530.3 and NM_181054.2, Homo sapiens hypoxia inducible factor 1, alpha subunit, transcript variants 1 and 2 (*HIF1A*)	F: CGTTCCTTCGATCAGTTGTGA R: CTTCCATACGGTCTTTTGTG
VEGFA	NM_001171622-630, NM_001025366-370, NM_001025376,	F: GCTACTGCCATCCAATCGAG R: TGGTGATGTTGGACTCCTCA

	NM_001033756, Homo sapiens vascular endothelial growth factor A (*VEGFA*) transcript variants 1 to 8	
p300	NM_001429.3, Homo sapiens E1A binding protein p300 (*EP300*)	F: GCTTCAGACAAGTCTTGGCAT R: GCCTGTGTCATTGGGCTTTTG
PRKAA1	NM_006251.5, NM_206907.3, Homo sapiens protein kinase, AMP-activated, alpha 1 catalytic subunit (*PRKAA1*), transcript variants 1and 2	F: AACCATGATTGATGATGAAGCCT R: GGTGTTTCAGCAACCAAGAATG
Hsc70	NM_153201.1, NM_006597.3, Homo sapiens heat shock 70kDa protein 8 (*HSPA8*), transcript variants 1 and 2	F: TGTGGCTTCCTTCGTTATTGG R: GCCAGCATCATTCACCACCAT
Hsp70	NM_021979.3, NM_005346.4, NM_005345.5, Homo sapiens heat shock 70kDa protein 1A (*HSPA1A*), 1B (*HSPA1B*), 2(*HSPA2*)	F: CGACCTGAACAAGAGCATCA R: AAGATCTGCGTCTGCTTGGT
Angptl4	NM_001039667.1 NM_139314.1, Homo sapiens angiopoietin-like 4 (*ANGPTL4*, *ANGPTL2*)	F: ACCTCCCGTTAGCCCCTG R: CATGGTCTAGGTGCTTGTGGTC
MT3	NM_005954.2, Homo sapiens metallothionein 3	F: ATGGACCCTGAGACCTGCC R: TTGCACACACAGTCCTTGGC
PTGES	NM_004878.4, Homo sapiens prostaglandin E synthase	F: TCAAGATGTACGTGGTGGCC R: GAAAGGAGTAGACGAAGCCCAG

4.3.3 Cell staining and fluorescence activated cell sorting

Cells were stained by incubating in cell medium containing 10 μmol L^{-1}Hoechst 33342 dye (Invitrogen, Eugene, OR, USA) at 37 °C for 30 min. After staining the cells were trypsinized using 0.05% (*v*/*v*) trypsin (Invitrogen, Carlsbad, CA, USA) solution and

resuspended in keratinocyte SFM containing 10 μmol L^{-1} Hoechst 33342. Cells were kept on ice before sorting with a BD FACS Aria cell sorter (BD Biosciences, San Jose, CA, USA). Cells in G1 phase were used in gene expression assays.

4.3.4 Hypoxia treatment

CP-A cells at 80% confluency were incubated in keratinocyte SFM cell-growth medium containing 2% (*v*/*v*) Oxyrase (Oxyrase, Mansfield, OH, USA) at 37 °C for up to 30 min (see below). Cells were subsequently trypsinized in 0.05% (*v*/*v*) trypsin solution containing 2% (*v*/*v*) Oxyrase at 37 °C for 9 min. The trypsinization was blocked by Dulbecco's modified Eagle medium (DMEM) (Invitrogen) supplemented with 5% fetal bovine serum (FBS) (Invitrogen) containing 2% (*v*/*v*) Oxyrase. The oxygen concentration was determined by use of an optical sensor (Tian, Shumway et al. 2010) calibrated with a Clark electrode. To determine the optimum Oxyrase treatment time, the expression levels of selected genes in bulk cells exposed to hypoxia for 10, 20, 30, 60, 180, and 360 min were measured.

4.3.5 Single cell collection

Single cells were collected using a pick-and-place single-cell manipulation robot (Anis, Holl et al. 2010). Single cells were aspirated and dispensed using a 40-μm diameter glass capillary micropipette under closed-loop microscopic vision-based feedback. After a selected cell was aligned with the micropipette orifice, the cell was aspirated by applying a negative pressure to the micropipette capillary, generating a drag force on the cell and pulling it into the capillary. The micropipette tip containing an aspirated cell was directed into the cap of a 1.5-mL microcentrifuge tube (VWR, West Chester, PA, USA) containing 200 μL keratinocyte SFM. The cell was dispensed by

applying a positive pressure to the micropipette capillary, generating an ejection force on the cell. Fluidic aspiration and dispensing of the cells were accomplished with minimal shear force on the cells so as to not cause physical damage to the cell. A total of 36 hypoxia-treated single cells and 36 control single cells were collected and analyzed.

4.3.6 RNA isolation and reverse transcription

Each single cell in the cap of a 1.5-mL microcentrifuge tube was centrifuged at 4 °C and 17,000 *g* for 20 min. Medium (160 μL) was taken out and 200 μL RNA lysis buffer from the ZR RNA MicroPrep Kit (Zymo Research, Orange, CA, USA) was added into the tube. The RNA extraction step was carried out using the ZR RNA MicroPrep Kit following the manufacturer's instructions. A total volume of 6 μL RNA was eluted from the column matrix and immediately used in reverse transcription reactions. A total volume of 10 μL cDNA synthesis reaction mixture contained the reagents: 2 μL of 5 × VILO Reaction Mix (Invitrogen), 1 μL of 10 × SuperScript Enzyme Mix, including SuperScript III RT, RNaseOUT Recombinant Ribonuclease Inhibitor, and a proprietary "helper" protein (Invitrogen), 6 μL of total RNA from a single cell, and 1 μL of DEPC-treated water (Ambion, Austin, TX, USA). The tube contents were gently mixed and then cDNA synthesis was performed at 25 °C for 10 min, 42 °C for 60 min, followed by 85 °C for 5 min to inactivate the reverse-transcriptase. The cDNA obtained was stored at −20 °C until further use.

4.3.7 qPCR

The Express SYBR GreenER qPCR SuperMix Kit (Invitrogen) was used for qPCR analysis. Each qPCR reaction for method development, primer testing, and optimization, and the hypoxia response experiment was performed using 5% ($1/20^{th}$) of

the cDNA obtained from a single cell. Reactions were conducted in 0.1-mL MicroAmp Fast 8-Tube Strips (Applied Biosystems, Foster City, CA, USA). Hypoxia treatment expression assays were conducted in adhesive-sealed, clear 384-well PCR plates (BioExpress, Kaysville, UT, USA). DEPC-treated water (30 μL) was added to 10 μL of cDNA solution obtained from the reverse transcription step. The total reaction volume was 10 μL and comprised the reagents: 5 μL of Express SYBR GreenER qPCR SuperMix Universal, 1 μL of each primer (4 μmol L^{-1}), 0.1 μL of ROX reference dye (25 μmol L^{-1}), 2 μL of cDNA solution ($1/20^{th}$ of a total of 40 μL cDNA solution obtained from a single cell) and 0.9 μL of DEPC-treated water (Ambion). In negative control reactions, the 2 μL of cDNA solution was replaced with DEPC-treated water. The thermal cycling profile was set up as follows: one cycle at 95 °C for 10 min; 40 cycles at 95 °C for 15 s, 60 °C for 1 min, and 80 °C for 10 s with signal detection; melt-curve analysis at 60 °C for 1 min and the temperature increased in 0.3 °C increments to 95 °C, then at 95 °C for 15 s. The method development experiments were run in a StepOne Real Time PCR System (Applied Biosystems). The gene expression profiling of hypoxia-treated cells was run in an Applied Biosystems 7900 Real-Time PCR System.

In order to push the limit on the number of genes whose expression levels could be quantified from a single mammalian cell, qPCR reactions were run on $1/45^{th}$ of the cDNA obtained from individual cells. DEPC-treated water (80 μL) was added to the 10 μL of cDNA obtained from the reverse transcription step. Each reaction was conducted in a 0.1-mL PCR tube (Applied Biosystems). The reaction was set up in a total volume of 10 μL and contained the reagents: 5 μL of Express SYBR GreenER qPCR SuperMix Universal, 1 μL of each primer (4 μmol L^{-1}), 0.1 μL ROX reference dye

(25 μmol L^{-1}), 0.1 μL of *Taq* DNA Polymerase (5 U μL^{-1}; Fermentas, Glen Burnie, MD, USA), 2 μL of cDNA solution (1/45th of a total 90 μL of cDNA solution obtained from a single cell), and 1.8 μL of DEPC-treated water. The thermal cycling profile was: 1 cycle at 95 °C for 10 min; 50 cycles at 95 °C for 15 s, 60 °C for 1 min, and 80 °C for 10 s with signal detection; melt-curve analysis at 60 °C for 1 min and the temperature increased in 0.3 °C increments to 95 °C then at 95 °C for 15 s. These experiments were run in a StepOne Real Time PCR System (Applied Biosystems). Data analysis was carried out using the StepOne software (Applied Biosystems). ANOVA *t*-test was used for statistical significance analysis.

4.4 Results and discussion

4.4.1 Two-step RT-qPCR analysis of single mammalian cells

A procedure was developed for quantifying the expression levels of multiple genes in a single mammalian cell using SYBR green-based qPCR. The procedure comprises six steps:

1. fluorescence-activated cell sorting to obtain cells in a particular phase of the cell cycle, e.g. G1;

2. single-cell collection;

3. RNA extraction;

4. reverse transcription;

5. qPCR;

6. data analysis.

Adapting technology for RNA isolation and reverse transcription from single bacterial cells (Gao, Zhang et al. 2011), a new procedure was conceived, developed, and optimized for the isolation, purification, and reverse transcription of the total RNA from a single mammalian cell. The ZR RNA MicroPrep kit (Zymo Research) was used, followed by cDNA synthesis using the SuperScript VILO cDNA Synthesis Kit (Invitrogen). Because of picogram levels of available cDNA templates for qPCR in single-cell experiments, the probability of primer dimer formation increases. In this case, fluorescence signal originating from double-stranded primer dimers confounds quantification of the target gene amplification product. To eliminate such interference, primer pairs whose amplification product lengths are between 100 bp and 700 bp were designed. Therefore, amplification products can be distinguished from primer dimers,

which are typically less than 100 bp, in both melt-curve analysis and agarose gel electrophoresis. Also, the qPCR thermal cycling program was modified to enable a signal-detection step (80 °C for 10 s) after the annealing/amplification step (60 °C for 1 min). The melting temperatures of primer dimers are usually less than 75 °C, whereas those of the target gene amplification products are more than 80 °C. Thus, when the fluorescence signal is detected at 80 °C, only the target gene amplification products will remain intact and the double-stranded primer dimers will dissociate. A total volume of 10 μL SYBR GreenER qPCR reagent kit was used for each qPCR reaction (as described in Section 4.3.7 qPCR). The total amount of cDNA obtained from a single CP-A cell was divided into twenty equal portions and each portion was used for one qPCR reaction. Three technical replicates were run for each pair of primers to assess the sample-handling error. cDNA obtained from a bulk cell sample was diluted to a level corresponding to the amount of cDNA from 10 cells and used as the positive control. One negative control reaction was run for each pair of primers, in which cDNA was replaced with DEPC-treated water. Experiments were conducted in a StepOne Real Time PCR System (Applied Biosystems). After qPCR, the contents of each PCR reaction were subjected to agarose gel electrophoresis and sequence analysis.

For method-validation purposes, qPCR analyses were conducted of highly expressed genes (*18S rRNA*, *28S rRNA*, *ACTB*, and *GAPDH* genes) in single cells. To increase primer binding specificity, several iterations of primer optimization were conducted. So far, 13 pairs of primers targeting 13 different genes (Table 6) have been validated. Specifically, *28S rRNA* is a reliable internal control for comparative transcription analyses under hypoxic conditions (Zhong and Simons 1999). Therefore,

28S rRNA was used as an internal reference for comparing gene expression levels within one cell and among multiple single cells. The choice of *Hsp70* and *Hsc70* genes is based on the fact that *Hsp70* expression levels change significantly in response to stress whereas *Hsc70* is a constitutively expressed cognate gene whose levels remain constant. Thus, *Hsc70* can be used as a reference gene (Mayer and Bukau 2005) for *Hsp70*. *HIF1α* (Zhong, Agani et al. 1998), *VEGFA* (Yoshiji, Gomez et al. 1996), *PRKAA1* (Mu, Brozinick et al. 2001), *p300* (Arany, Huang et al. 1996), *MT3* (Wang, Wood et al. 2008), *Angptl4* (Murata, Yudo et al. 2009), and *PTGES* (Lee, Natsuizaka et al. 2010) were chosen because of their roles in hypoxia response-signaling pathways.

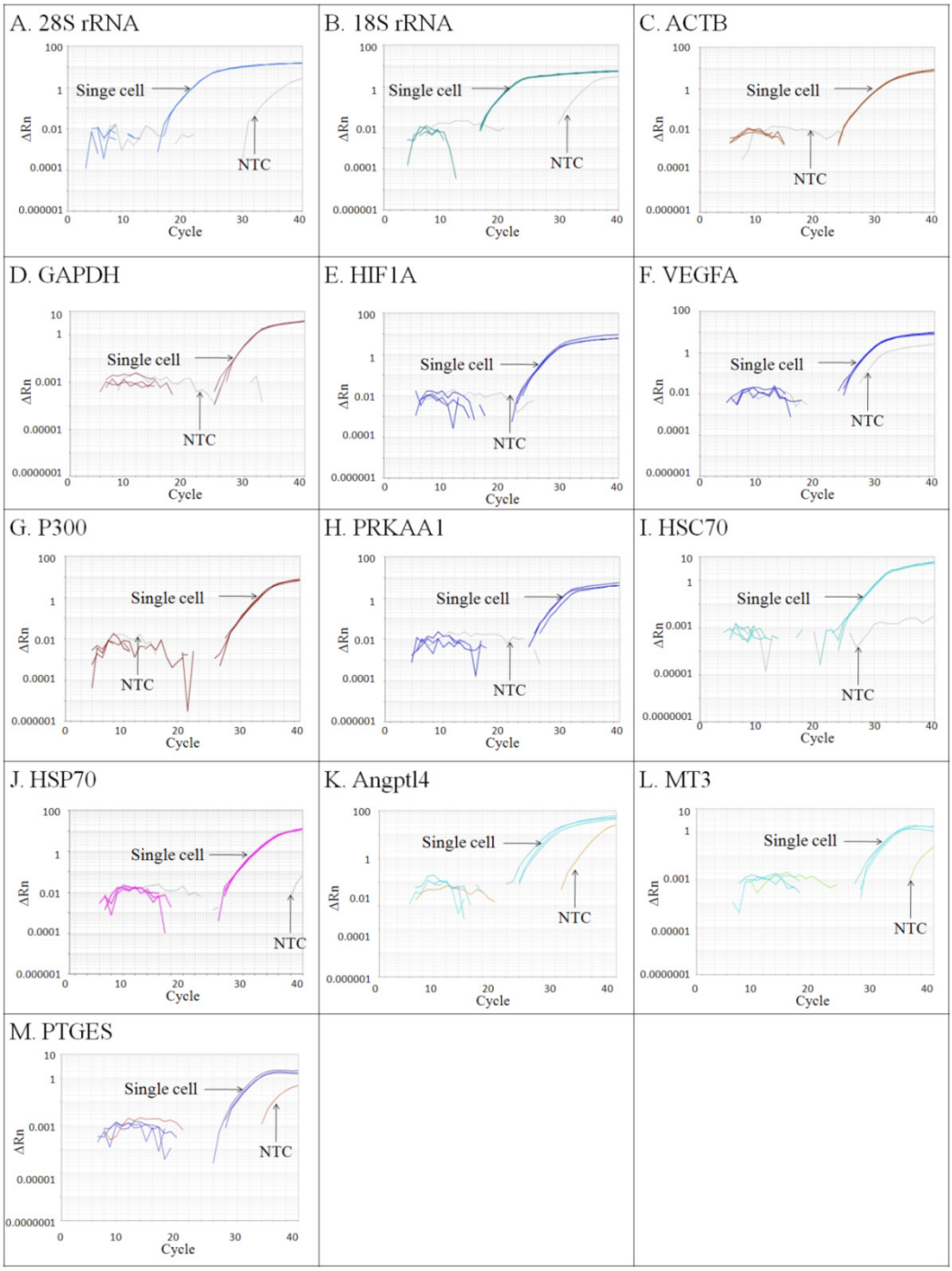


Figure 4-1 Amplification plots of gene transcripts using validated primers
$1/20^{th}$ of the total cDNA obtained from a single CP-A cell was used for each qPCR reaction shown. This includes three technical replicates and the no-template controls (NTC). Each panel shows real-time amplification signal curves obtained from a single cell.

Using these primers, the expression levels of genes of interest were quantified by RT-qPCR using 1/20th of the cDNA obtained from a single CP-A cell. Representative amplification plots for the 13 genes are shown in Figure 4-1. The C_q ratio of *28S rRNA* to *18S rRNA* has been found to be typically approximately 1 in mammalian cells (Fleige and Pfaffl 2006). The results here showed that the C_q values for *28S rRNA* and *18S rRNA* are 19.73 and 20.71, respectively, indicating that the method yields intact mRNA and is reproducible and quantitative. The C_q values of 28S and 18S rRNAs are approximately 10 cycles lower than those of ACTB and GAPDH mRNAs. This is consistent with the fact that rRNA concentrations are 100 to 1000 times higher than mRNAs in cells, as found in bulk cell studies. Reactions with C_q values lower than 37 and standard deviations of technical replicates smaller than 1% were regarded as successful. In most reactions, no-template negative controls (NTC) were not detected (Figure 4-1). In some reactions residual NTC signals, most likely emanating from primer dimers, were detected at cycle numbers significantly higher than the product (at least 3 or 4 cycle difference). The spurious NTC signals were clearly distinguishable from those of the amplification products (cDNA) because of distinctly higher C_q values (Figure 4-1) and different melting peak temperatures (Figure 4-2). The amplification products from the reactions containing cDNA showed clear bands of the correct sizes in agarose gel electrophoresis (data not shown). The gel bands containing products of the right size were cut, and the product was purified for sequencing. Sequences were confirmed using BLAST annotation against the NCBI GenBank database.

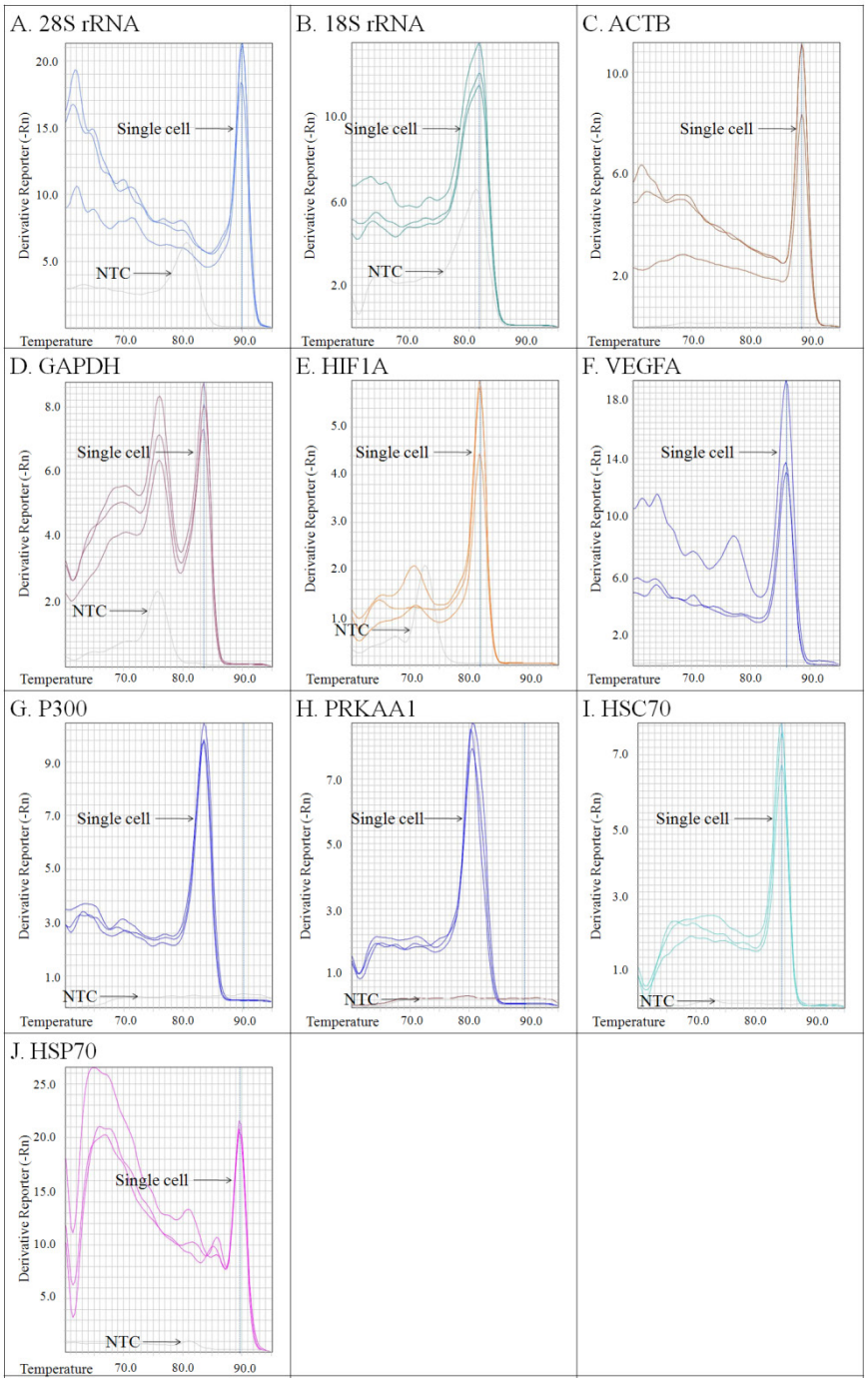
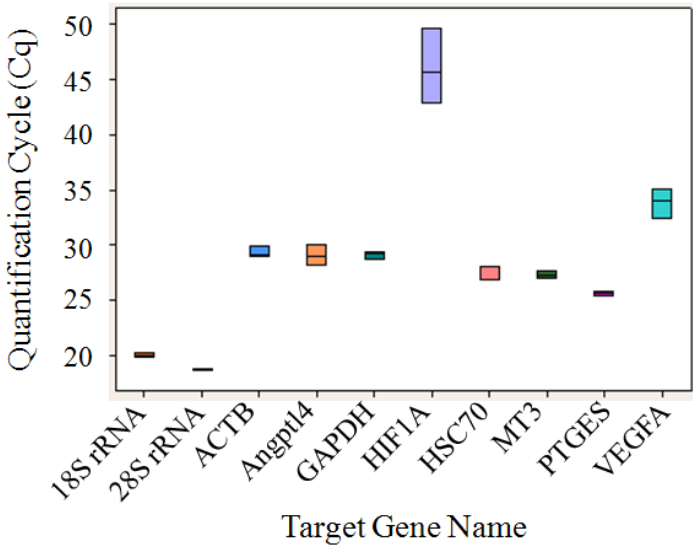


Figure 4-2 Melt curves of gene transcripts using validated primers
Each panel represents derivative reporter values of real-time fluorescence signals in melt curve analysis.

4.4.2 Detection of up to ten genes from a single mammalian cell

A method was developed which uses SYBR green-based RT-qPCR for detecting expression levels of up to ten different genes (*18S rRNA*, *28S rRNA*, *ACTB*, *Angptl4*, *GAPDH*, *HIF1A*, *HSC70*, *MT3*, *PTGES*, and *VEGFA*) in a single CP-A cell. These genes span a broad range of copy numbers in cells. A population of CP-A cells was exposed to hypoxia, and randomly selected, single cells were collected by use of a single-cell manipulation workstation. After RNA extraction and reverse transcription, cDNA template obtained from a single CP-A cell was divided into 45 equal portions of 2-μL volume each. This dilution enabled up to 45 qPCR reactions to be performed on the cDNA obtained from a single cell, enabling triplicate analyses of multiple genes from the same cell. Each portion of the template was added to an SYBR GreenER qPCR reagent mixture, resulting in a total reaction volume of 10 μL. For each gene one RT-qPCR NTC was run in parallel. To increase the amplification level, 0.1 μL of *Taq* DNA Polymerase (Fermentas) was added to each reaction mixture. As in the previous experiments, reactions for each of the genes were run in triplicate. To achieve sufficient amplification levels of low-abundance transcripts, the thermal cycling profile was modified by extending the amplification cycle number to 50. The reactions were run utilizing the StepOne Real Time PCR instrument.



Target Gene Name	C_q Mean	C_q SD	C_q SD%
18S rRNA	19.96	0.20	0.98%
28S rRNA	18.71	0.06	0.30%
ACTB	29.33	0.55	1.88%
Angptl4	29.08	0.90	3.09%
GAPDH	29.09	0.38	1.31%
HIF1A	46.09	3.37	7.30%
HSC70	27.65	0.66	2.40%
MT3	27.31	0.34	1.24%
PTGES5	25.62	0.16	0.64%
VEGFA	33.83	1.32	3.90%

Figure 4-3 Quantification of expression levels of ten genes in a single, hypoxia-treated CP-A cell

$1/45^{th}$ of the total cDNA obtained from a single CP-A cell was used for each qPCR reaction. Each box includes quantification cycle (C_q) values from three technical replicates for one gene. Average C_q values and standard deviations among three technical replicates are summarized in the table.

Using this method, the transcription levels of up to ten genes with C_q values lower than 50 can be simultaneously quantified (Figure 4-3). The standard deviations of the qPCR C_q value of eight genes (*18S rRNA*, *28S rRNA*, *ACTB*, *GAPDH*, *Angptl4*, *MT3*, *PTGES* and *HSC70*) were less than 1 (3%) among the technical replicates (Figure 4-3).

As expected, larger variations between the technical replicates was observed for low-abundance transcripts, for example *HIF1A* and *VEGFA* (3.4 (7%) and 1.3 (4%), respectively). It is most likely that this is because of picogram levels of available cDNA templates, which makes the amplification events at the beginning less probable and introduces more noise in C_q values. No amplification signal was detected for the NTC reactions of *HIF1A*, *VEGFA*, and *HSC70*, and for NTC reactions of *18S rRNA*, *28S rRNA*,

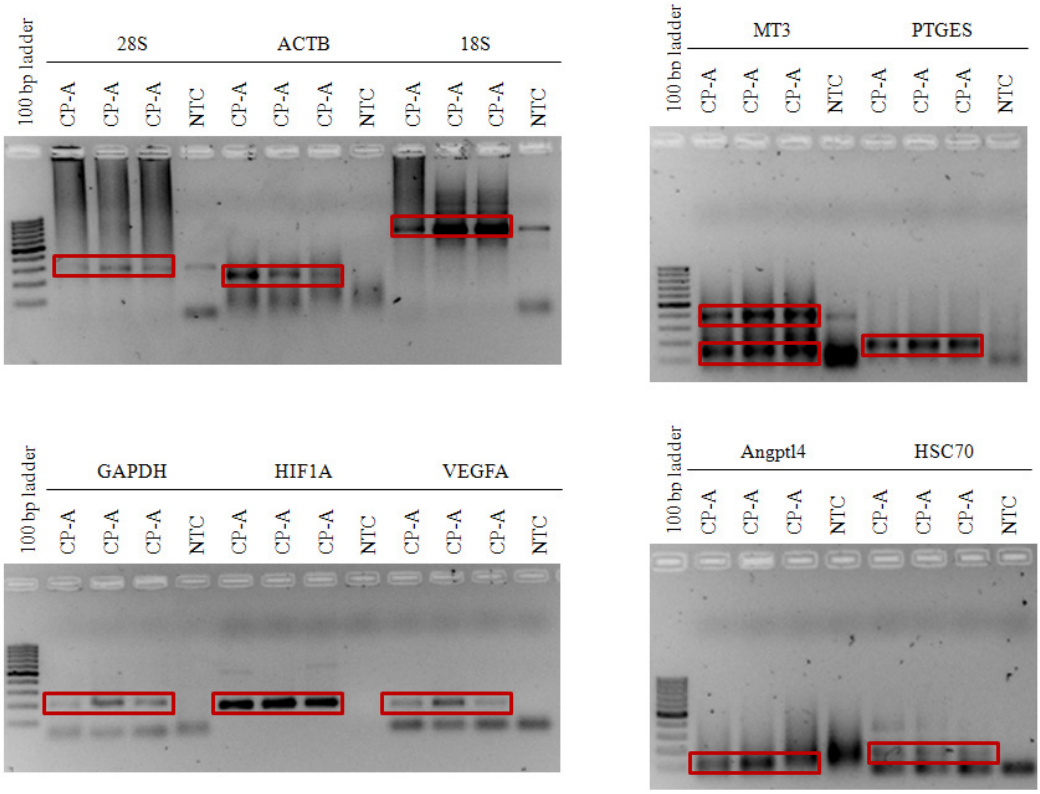


Figure 4-4 Gel electrophoresis analysis of RT-qPCR amplification products of 10 genes from a single CP-A cell

2.5 µL of the qPCR products were loaded on a 2% agarose gel. The bands in the frame indicate the correct amplification product. In several cases the correctness of the PCR products was confirmed by sequencing the contents of the corresponding bands.

ACTB, *GAPDH*, *Angptl4*, *MT3*, and *PTGES* amplification signals were observed after a large number of cycles. However, the melting temperatures (T_m value) were 0.4–10.3 °C different from the T_m values of the corresponding products in cDNA sample reactions. Therefore, amplification products resulting from unavoidable minute contamination and/or random, non-specific amplification can be identified by their characteristic T_m values. Both length and sequence of the amplification products were confirmed by agarose gel electrophoresis (Figure 4-4) and sequencing.

The C_q values for *28S* and *18S rRNA*s are approximately 10 cycles lower than those of the housekeeping genes *ACTB* and *GAPDH* (100–1000 times higher in copy number), and 7–17 cycles lower than those of the genes of interest (100–100,000 times higher in copy number).

4.4.3 Gene expression under hypoxia

One of the research objectives was to understand epithelial cells' responses to hypoxic stress. To address the role of intercellular gene expression heterogeneity under different physiological conditions, gene expression levels in CP-A cells in response to hypoxia were studied. Oxyrase, an oxygen-scavenging enzyme system, was used to produce hypoxic conditions in cell growth medium. Oxygen concentration in the cell medium was reduced to <0.01 ppm within 20 min after Oxyrase addition (Figure 4-5).

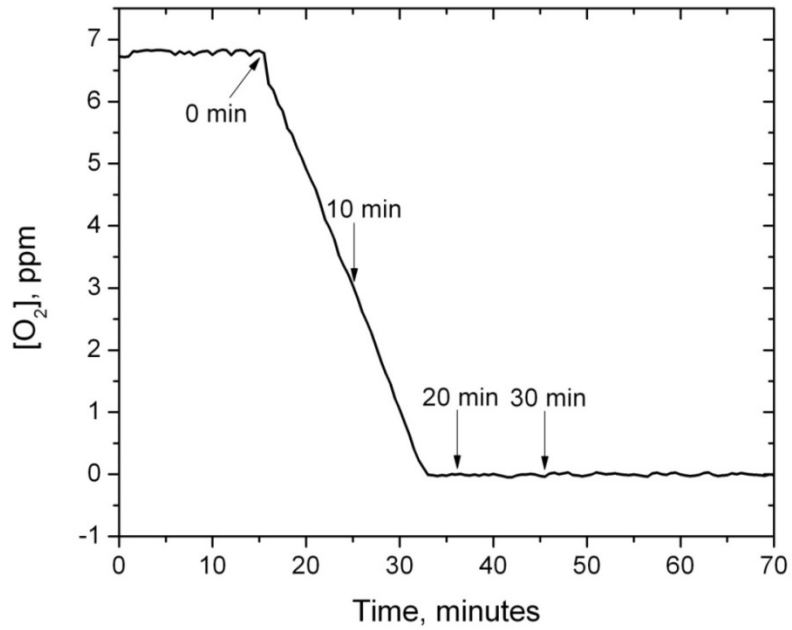


Figure 4-5 Oxygen depletion in the cell medium
Time course of the oxygen concentration in the cultivation medium before and after addition of 2.0% (v/v) Oxyrase. *Arrows* indicate the time points of gene expression profiling experiments. Oxyrase was added at time zero.

To determine the time course of changes in gene expression, bulk cell samples (~10^4 cells per sample) were analyzed at 0, 10, 30, 60, 180, and 360 min after adding Oxyrase to the cell medium, and the expression levels were measured of six genes (*Angptl4*, *PTGES*, *MT3*, *PRKAA1*, *VEGFA*, and *GAPDH*) known to be involved in hypoxia response, using *28S rRNA* as the internal control. A 1:1000 dilution of the total synthesized cDNA was used for RT-qPCR. The results showed that mRNA levels of *Angptl4*, *PRKKA1*, and *MT3* increased and that of *GAPDH* decreased in response to the Oxyrase treatment, whereas mRNA levels of *VEGFA* and *PTGES* were not significantly affected (Figure 4-6).

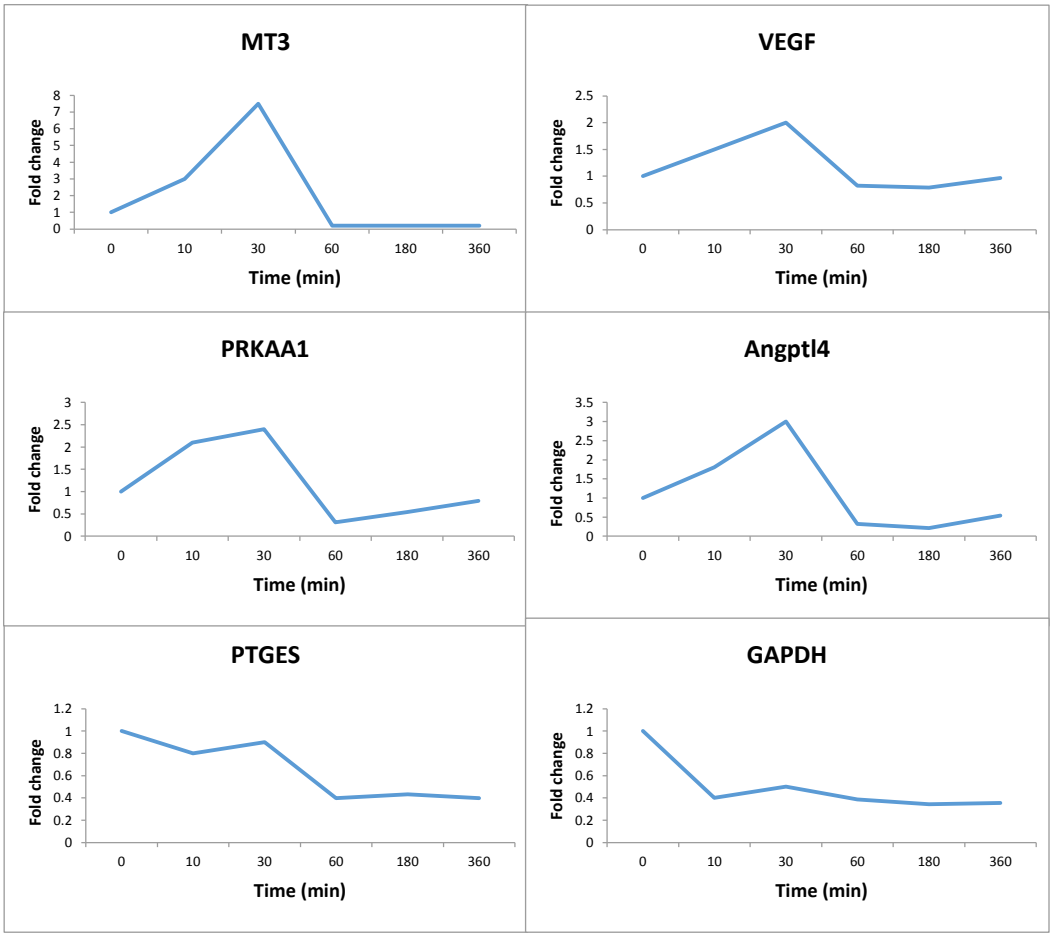


Figure 4-6 qPCR results of long time Oxyrase treatment
Bulk CP-A cell samples were used to determine the treatment time when the largest changes in gene expression levels occur. Expression levels of seven genes, *28s rRNA*, *MT3*, *VEGFA*, *PRKAA1*, *Angptl4*, *PTEGS*, and *GAPDH*, were measured using qPCR at bulk cell levels as a function of time at 0, 10, 30, 60, 180, and 360 minutes after addition of Oxyrase. The gene expression levels were indicated as fold-changes normalized against 28s rRNA. The fold-changes of gene expression were calculated as $2\text{-}\Delta(\Delta C_q)$, where $\Delta C_q = C_q$, target - C_q, 28s, and $\Delta(\Delta C_q) = C_q$, stimulated - ΔC_q, control.

The absence of significant changes in the expression levels of *VEGFA* and *PTGES* genes in bulk cell samples may be cell type-specific or a result of the relatively short hypoxia exposure (30 min). Among the genes studied, *MT3* seems to be most

sensitive to hypoxia, with 3 and 7-fold increases in expression level after 10 and 30 min treatments, respectively. At least twofold increases in transcription levels were observed for two other hypoxia-response genes, *Angptl4* and *VEGFA*, in response to 30-min Oxyrase treatment. On the basis of the higher expression observed in bulk cell experiments after 30 min treatment with Oxyrase, this treatment time was chosen for single-cell analysis.

A total of 72 single CP-A cells were collected, 36 grown under normal physiological conditions (controls) and 36 exposed to hypoxia for 30 min, the expression levels of seven selected genes in each cell were measured. To limit cell-to-cell variability that may result from differences in cell cycle phase, cells in G1 phase sorted by means of fluorescence-activated cell sorting were used. In addition to six genes (*PTGES* (Lee, Natsuizaka et al. 2010), *Angptl4* (Cazes, Galaup et al. 2006, Galaup, Cazes et al. 2006, Gentil, Le Jan et al. 2006, Gustavsson, Mallard et al. 2007, Murata, Yudo et al. 2009), *MT3* (Wang, Wood et al. 2008), *GAPDH* (Liu, Cox et al. 1995, Iyer, Kotch et al. 1998, Foldager, Munir et al. 2009), *PRKAA1* (Mu, Brozinick et al. 2001), and *VEGFA* (Liu, Cox et al. 1995, Yoshiji, Gomez et al. 1996, Hu, Fan et al. 2009)) known to be involved in the hypoxia response signaling pathway in various bulk cell-based studies, the highly-expressed *28S rRNA* was included as the reference gene. The number of cells with all seven target genes detected was higher in hypoxia-treated cells. For example, *PTGES* transcripts were detected in all 36 hypoxia-treated cells and in 27 control cells (Figure 4-7A). This can be attributed to higher mRNA copy numbers of these genes in cells exposed to hypoxia than in control cells. The *28S rRNA* gene was detected in all 72 cells, with small cell-to-cell variations (standard deviations of $C_q < \pm 0.5$).

The differences between ΔC_q values ($\Delta C_q = C_{q,\ target\ gene} - C_{q,\ 28s}$ (Golding, Paulsson et al. 2005)) of the target genes as measured under normoxic and hypoxic

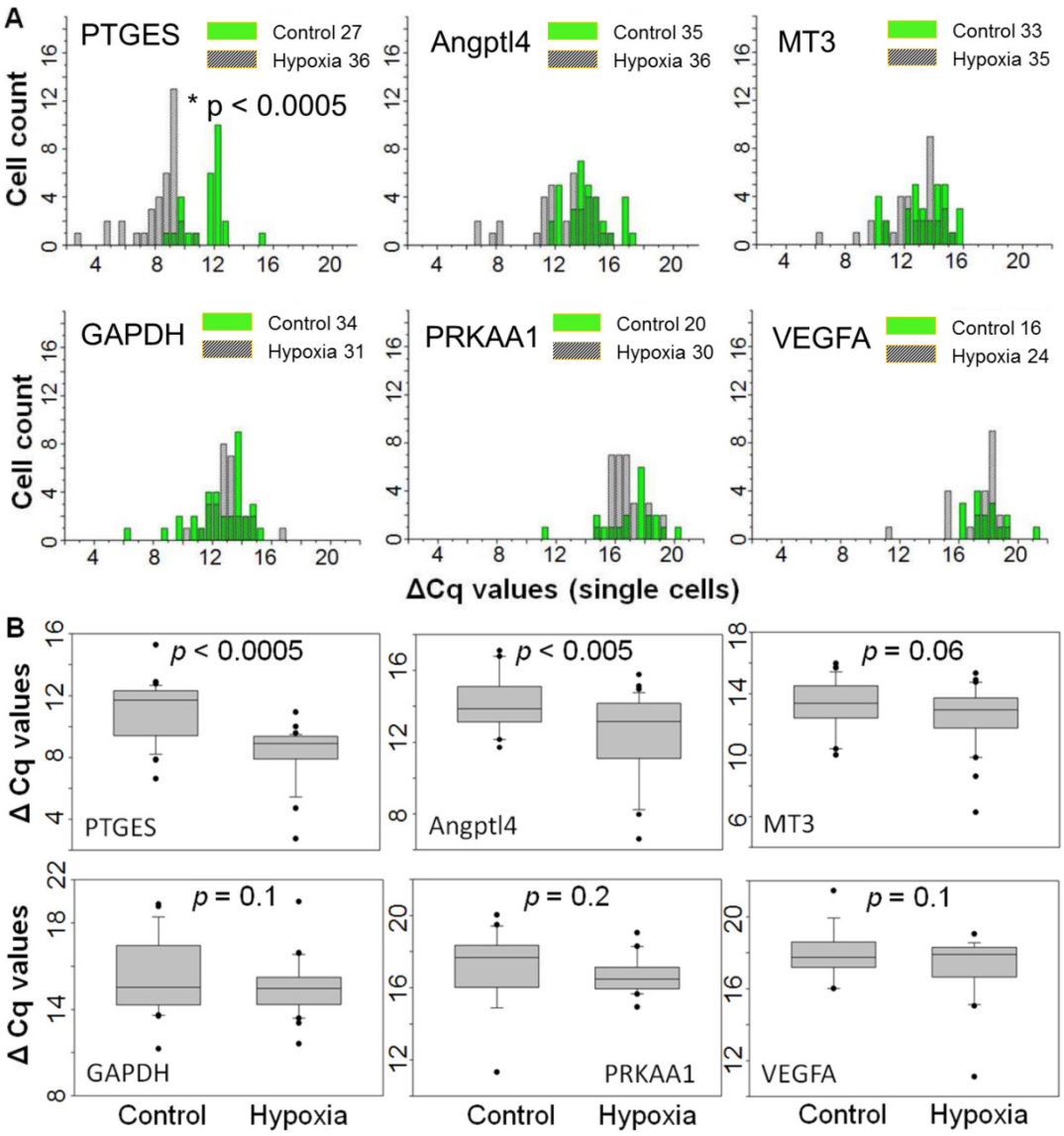


Figure 4-7 Single-cell gene-expression profiling in control and hypoxia-treated CP-A cells

(A) Histograms of gene expression levels in control (*green bars*) and hypoxia-treated (30 min, *hatched bars*) single CP-A cells in G1. (B) Box plots of single-cell gene-expression levels and *p*-values associated with differences between in untreated controls and hypoxia-treated CP-A cells.

conditions were highly variable among cells, indicating significant cell-to-cell heterogeneity in the cell population. Distribution histograms of gene expression levels in control and hypoxia-treated CP-A cells indicate significant cell-to-cell heterogeneity in all six studied genes (Figure 4-7A). PTGES and Angptl4 genes showed the largest differences in ΔC_q values between control and hypoxia-treated cells. The ΔC_q values for these genes were lower in the treated cells, indicating up-regulation of PTGES and Angptl4 in response to hypoxia. Statistical analysis of gene expression levels in hypoxia-treated vs. control single cells confirmed a significant reduction of ΔC_q values for *PTGES* ($p < 0.0005$) and *Angptl4* ($p < 0.005$) genes, whereas changes in ΔC_q for the other four genes were not statistically significant (Figure 4-7B). In some early bulk cell studies, expression levels of several of the target genes used in this study, for example the *MT3* gene encoding a metal-binding protein metallothionein 3 and the *PTGES* gene encoding a prostaglandin E synthase, were dramatically increased under hypoxia (Wang, Wood et al. 2008). The hypoxia treatment times used in the studies were 24 h (Lee, Natsuizaka et al. 2010), much longer than those used in this study (0.5 h) and, therefore, the findings are not directly comparable.

Interestingly, it was found that the gene-expression patterns obtained from bulk-cell (Figure 4-8) and single-cell samples (Figure 4-7) differed in terms of response to hypoxia. Only two genes, *Angptl4* and *VEGFA*, showed the same trend in both bulk and single-cell analyses: *Angptl4* was up-regulated and *VEGFA* did not change significantly in response to hypoxia treatment. Whereas no significant changes at the single-cell level

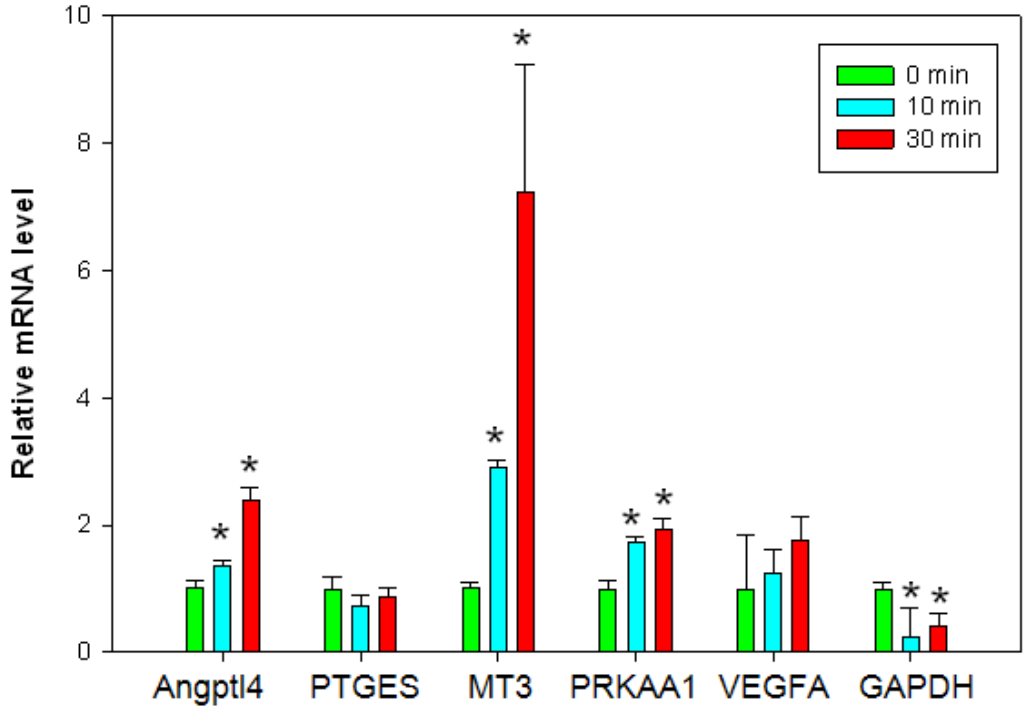


Figure 4-8 Gene expression levels of bulk cell samples under normoxic and hypoxic conditions

Results from three different experiments are shown. *Green bars* represent control cells that were not exposed to hypoxia. *Light blue bars* and *red bars* depict gene expression levels from cells exposed to hypoxia for 10 and 30 min, respectively. The changes of gene expression were calculated as $2^{-\Delta(\Delta C_q)}$, where $\Delta C_q = C_{q,\ \text{target}} - C_{q,\ 28s,}$ and $\Delta(\Delta C_q) = \Delta C_{q,\ \text{stimulated}} - \Delta C_{q,\ \text{control}}$. *Asterisks* indicate statistical significance at $p < 0.05$ on the basis of ANOVA *t*-test analysis.

were observed for *MT3* and *PTGES*, both genes were significantly up-regulated in bulk-cell samples, consistent with other bulk-cell studies (Arany, Huang et al. 1996, Lee, Natsuizaka et al. 2010). *GAPDH* was not significantly changed in single cells but significantly down-regulated in bulk cells whereas *PTGES* was significantly up-regulated in single cells whereas no significant change was observed in bulk cells. Recent studies on mRNA levels in individual cells suggest that cell-to-cell alterations in gene expression levels seem to be a result of variations at the bulk mRNA stability and/or translational level (Golding, Paulsson et al. 2005, Siegal-Gaskins and Crosson 2008, Le and Cheng 2009, Valencia-Burton, Shah et al. 2009, Lidstrom and Konopka 2010). A recent study showed that single-cell gene expression has a log-normal distribution, reflecting true biological variability (Stahlberg and Bengtsson 2010). This finding indicates that average gene expression levels quantified in a population of cells may be substantially different from expression levels measured in individual cells from the same population (Bengtsson, Hemberg et al. 2008). This result also emphasizes the importance of developing and applying microfluidics-based instrumentation for high-throughput single-cell gene expression measurements with improved statistical power.

4.5 Conclusion

In summary, a qPCR-based method was developed and applied for single-cell gene expression analysis, enabling measurement of multiple gene targets in a single mammalian cell. The method is based on separate RNA isolation, cDNA synthesis, and qPCR steps. Using this method, gene transcription levels were quantified in control and hypoxia-treated cells at both bulk and single-cell levels. The results show that quantitative analysis of gene expression of multiple genes can be achieved in single cells with good reproducibility and specificity. In addition, significant gene-expression heterogeneity was observed among the sorted cell population.

The procedure can be further improved by performing absolute mRNA abundance determination in single cells, using a gene sequence cloned into a plasmid as a reference to calculate absolute mRNA copy numbers. *ACTB* mRNA transcribed in vitro using T7 RNA Polymerase can be used to validate the efficiency and reliability of the reverse transcription step of the procedure. Given the compatibility of this method with most commercially available RT-qPCR instrumentation and its relatively low cost, it should be amenable to many applications focused on gene expression analysis in single cells, for example high-throughput, chip-based techniques (Kelly and Woolley 2005), which will provide further insights into the cellular mechanisms involved in physiological and pathological processes at the single-cell level.

4.6 Acknowledgments

This work was supported by a grant from the NIH National Human Genome Research Institute, Centers of Excellence in Genomic Sciences (Grant Number 5 P50 HG002360 to Deirdre Meldrum). I thank Patti Senechal-Willis for the assistance with cell culture. I thank Tong Fu for help with fluorescence-activated cell sorting.

CHAPTER 5

WHOLE TRANSCRIPTOME AND METABOLIC PROFILING OF INTERCELLULAR INTERACTIONS BETWEEN NORMAL AND PRE-MALIGNANT ESOPHAGEAL CELLS

5.1 Abstract

Intercellular communications are critical to the growth and function of multicellular organisms. The interplay between aberrant cells and their microenvironment can drive pathogenesis and progression of many diseases including cancer. Solid tumors are organ-like structures with complex and dynamic interactions among different clones. Understanding the role of cell-cell communications in the tumor microenvironment could lead to new cancer biosignatures and more effective prognostic, diagnostic and management strategies of the disease. This study is about the effects of intercellular interactions between normal and late-dysplastic Barrett's esophagus cells on gene transcription levels and metabolic functions. It was investigated how homotypic and heterotypic intercellular interactions affect gene expression changes using next-generation sequencing. Analysis of next-generation sequencing results identified differentially expressed genes. These genes were enriched in cellular movement and other cancer-related pathways. Heterotypic interactions suppressed downstream genes of *TGFβ* and *EGF* pathways in late-dysplastic cells, compared with homotypic interactions. Gene transcription changes were correlated with cellular proliferation and motility phenotypes and their functional relevance during pre-malignant progression was explored.

5.2 Introduction

Cell-cell interactions are essential for growth and function of multicellular organisms. Aberrant intercellular communication plays a key role in carcinogenesis and tumor progression. More and more evidence has shown that a tumor is not merely a collection of homogenous cancer cells undergoing transformation by themselves (Hanahan and Weinberg 2011). Tumor genesis and progression is an ecological process involving dynamic interplays between malignant and non-malignant cells (Barcellos-Hoff, Lyden et al. 2013). Tumor cells are affected by reciprocal interaction between the parenchymal and stromal cells in two ways: physically through direct contact or intervening extracellular matrix, or biologically through soluble ligands. The communication and signaling create a context that promotes tumor growth and helps it acquire hallmark traits of cancer (Hanahan and Weinberg 2011).

Esophageal adenocarcinoma is a highly lethal cancer type (Spechler 2013). It arises from metaplasia and dysplasia of the esophageal epithelium (Barrett's esophagus) (Shaheen, Crosby et al. 2000). Recent studies and clinical practices have identified and used dysplasia and genomic markers for predicting the risk of developing adenocarcinoma (Jin, Cheng et al. 2009). However, how the transformation of metaplasia-dysplasia-adenocarcinoma occurs still remains to be clarified.

Neoplastic cells in Barrett’s esophagus collect genetic and epigenetic alterations as they undergo evolution by natural selection. This process is influenced by their surrounding cells and other factors in the environment. Acid and bile luminal refluxate can induce double-stranded DNA breaks or promote oxidative DNA damage (Clemons, McColl et al. 2007). Inflammatory cell infiltrate can generate reactive oxygen species

(Naya, Pereboom et al. 1997), which might contribute to DNA damage in neoplastic cells. Reactive oxygen species might induce growth factors, survival factors or *Fas* ligand secretions (Younes, Schwartz et al. 1999). Inflammatory cell infiltrate can also produce cytokines, such as transforming growth factor β (*TGFβ*), interleukin one β (*IL-1β*), interferon γ (*IFNγ*), IL-6 (Zhang, Zhang et al. 2011) and tumor necrosis factor α (*TNFα*) (Tselepis, Perry et al. 2002). In the stromal compartment of Barrett's esophagus, gene expression profiles are distinctive among different stages of progression (Lao-Sirieix and Fitzgerald 2010). Thrombospondin-1 (*TSP1*) is overexpressed in stroma from Barrett's esophagus biopsy samples. *TSP1* can activate *TGFβ*, which either controls proliferation or promotes epithelial-mesenchymal transition in Barrett's esophagus and esophageal adenocarcinoma (Rees, Onwuegbusi et al. 2006). Furthermore, co-culture of squamous carcinoma and Barrett's carcinoma cells produces more pro-inflammatory cytokines compared with cells cultured individually (Fitzgerald, Abdalla et al. 2002). The current results suggest that cell-cell interactions in the tumor microenvironment can change epithelial cell behavior in Barrett's esophagus. More systematic research is required to elucidate the molecular components contributing to esophageal adenocarcinoma development and enhance understanding of this dynamic process.

Presented here is a study on gene expression and metabolic phenotype profiling of heterotypic cell-cell interactions between normal and late-dysplastic Barrett's esophagus cells. It investigated how homotypic and heterotypic intercellular interactions affect gene expression changes using next-generation sequencing. Using gene ontology and pathway enrichment analysis, a set of genes related to cellular movement and other cancer-related pathways is identified. Heterotypic interactions suppressed downstream genes *of TGF-β*

and *EGF* pathways in late-dysplastic cells. It was also found that heterotypic interactions between normal and dysplastic cells inhibited cellular proliferation and changed motility in both dysplastic and normal cells. Normal cells are found to inhibit the growth of dysplastic cells via *TGF-β* mediated growth factor pathways. The fractions of normal to dysplastic cells can be used as risk stratification markers for Barrett's esophagus and esophageal adeonocarcinoma. *TGF-β*, *EGF* and their downstream genes can become potential biosignatures for early diagnosis, early detection and risk prediction in pre-malignant progression of Barrett's esophagus.

5.3 Experiments

5.3.1 Cell lines

One normal cell line, EPC-2 (Harada, Nakagawa et al. 2003), and one cell line derived from an esophageal region of high-grade dysplasia in Barrett's esophagus, CP-D (Palanca-Wessels, Barrett et al. 1998), were used for cell-cell interaction studies. Cells were cultured using Gibco keratinocyte serum-free cell growth medium (Invitrogen, Carlsbad, CA, USA), supplemented with hEGF (Peprotech, Rocky Hill, NJ, USA) at 2.5 μg/500 mL, BPE (bovine pituitary extract) at 25 mg/500 mL and penicillin–streptomycin solution (Invitrogen, Carlsbad, CA, USA) at 100 units/100 μg/mL. Cells were grown at 37 °C in a humidified atmosphere containing 5% CO_2.

CP-D and EPC-2 cell lines were tagged with *TurboGFP* and *TagFP635*, respectively, to distinguish different cell types using fluorescence microscopy and fluorescence activated cell sorting. For lentiviral infection of both cell lines, 1.3 x 10^4 CP-D cells and 1.6 x 10^4 EPC-2 cells were seeded into individual wells of a Costar® 96-well-plate (Corning, Corning Life Sciences, Corning, NY, USA) containing 100 μL of Gibco Keratinocyte serum-free medium. After 24 hours of incubation, the cell culture reached about 80% confluency. 100 μL of Keratinocyte serum-free medium containing 8 mg/mL of hexadimethrine bromide was added to each well. Lentiviral MISSION® pLKO.1-puro-UbC-*TurboGFP*™ Positive Control Transduction Particles (Sigma-Aldrich, St Louis, MO, USA), containing a gene encoding *TurboGFP* under the control of the UbC promoter, were added to the well of CP-D cells at a multiplicity of infection of 2. Lentiviral MISSION® pLKO.1-puro-UbC-*TagFP635*™ Positive Control Transduction Particles (Sigma-Aldrich, St Louis, MO, USA), containing a gene encoding *TagFP635*

under the control of the UbC promoter, were added to the well of EPC-2 cells at a multiplicity of infection of 2 as well. The plate was gently stirred and centrifuged at 1000 rpm, 37° C for 30 minutes. After 18-20 hours, the medium containing lentiviral particles was replaced with 120 μL of Keratinocyte serum-free medium in each well. Four days after infection, the cells were imaged using a Nikon C1si (Nikon Inc., Melville, NY, USA) confocal microscope to inspect the expression of cytosolic *TurboGFP* in CP-D cells and *TagFP635* in EPC-2 cells.

After the culture was expanded into 75 cm^2 flasks (Corning, Corning, NY), a puromycin kill curve experiment was performed to determine the minimum concentration of puromycin to cause 0% viability ration in puromycin treated cells. 1.0 μg/mL and 0.5 μg/mL of puromycin were found to effectively kill the CP-D and EPC-2 cells, respectively, in which *TurboGFP* or *TagFP635* was not successfully expressed after 96 hours. CP-D and EPC-2 cells were grown in keratinocyte serum-free medium containing puromycin (1.0 μg/mL for CP-D cells and 0.5 μg/mL for EPC-2 cells) for five passages, and then grown in normal keratinocyte serum-free medium. The expression of cytosolic *TurboGFP* or *TagFP635* was retained in CP-D and EPC-2 cells, respectively, when checked under a Nikon C1si confocal microscope.

5.3.2 Fluorescence assisted cell sorting of co-culture and mono-culture of normal and neoplastic cells

For CP-D and EPC-2 mono-cultures, 2×10^6 cells of each type were seeded into a 75 cm^2 flask. For CP-D and EPC-2 cells co-culture, 1×10^6 CP-D and 1×10^6 EPC-2 cells suspension were mixed and seeded into a 75 cm^2 flask. After 24 hours, cells were treated with 0.05% (v/v) trypsin–EDTA (Invitrogen, Carlsbad, CA, USA). The

trypsinization was blocked by Dulbecco's modified Eagle medium (DMEM) (Invitrogen, Carlsbad, CA, USA) supplemented with 5% fetal bovine serum (FBS) (Invitrogen, Carlsbad, CA, USA). After trypsinization, cells were centrifuged at 900 rpm for 3 min then re-suspended in 300 µL of PBS (Gibco, Carlsbad, CA, USA) and kept on ice. CP-D and EPC-2 cells were sorted out from co-culture on a BD FACS Aria (BD Biosciences, San Jose, CA, USA), using a 488 nm laser to excite *TurboGFP* in CP-D cells and *FP635* in EPC-2 cells. Manual compensation was performed to correct for spectral cross-over of fluorescent proteins. Mono-cultured CP-D and EPC-2 cells were also sorted using the same *TurboGFP* or *FP635* gates and regions.

5.3.3 RNA extraction

Total RNA was extracted from the sorted cells using RNeasy mini kit (Qiagen, Valencia, CA, USA) according to the manufacturer's protocol. The silica-gel membrane and the spin-column technology removed the majority of the DNA. After proprietary buffer RW1 treatment, residual DNA in the RNA samples was digested on the column using RNase-free DNase set (Qiagen, Valencia, CA, USA) for 15 min at 20–30°C to remove DNA more completely. The RNase-free DNase set consists of 10 µL of DNase 1 stock solution in 70 µL of proprietary buffer RDD. RNA was eluted by adding 30 µL of RNase-free water into the silica-gel membrane and stored at -80°C.

5.3.4 Whole transcriptome amplification

The quantity and purity of RNA obtained from FACS sorted cells was measured using spectrophotometry on a Nanodrop instrument (Agilent Technologies, Santa Clara, CA, USA). The concentration of RNA was adjusted to 10 ng/µL. 50 ng of RNA was amplified using Nugen Ovation RNA-Seq V2 kits (Nugen Technologies, San Carlos, CA,

USA) per manufacturers' instruction on an Apollo 324 Library Preparation System (IntegenX, Pleasanton, CA, USA). Briefly, the RNA was reverse transcribed to the first-strand cDNA by using a combination of random hexamers and poly-T oligomer. The RNA template was fragmented by RNA-dependent DNA polymerase and double-stranded DNA was generated by the same polymerase. The dsDNA was amplified linearly using a single primer isothermal amplification (SPIA) process: RNase H degraded RNA in RNA in the DNA/RNA heteroduplex; the SPIA primer bound to the cDNA; the polymerase synthesized new cDNA strand replacing the RNA; random hexamers amplified the second-strand cDNA linearly. This technology covered non-coding RNA and non-polyadenylated RNA besides mRNA and reduced the conversion of ribosomal RNA to cDNA (Kurn, Chen et al. 2005).

Amplified DNA was measured by a Nanodrop instrument for quality control purposes. cDNA was sheared to about 250 bp using a Covaris S2 instrument (Covaris, Woburn, MA, USA) and checked again using a Nanodrop instrument.

5.3.5 Library preparation and Illumina sequencing

Illumina sequencing libraries were prepared on an Apollo 324 Library Preparation System using the PrepX™ ILM DNA Library Preparation kit (IntegenX, Pleasanton, CA, USA) with four different barcoded adapters for multiplexing. The adapter-ligated libraries were amplified by 10 cycles of PCR using a KAPA HiFi Library Amplification Kit (Kapa Biosystems, Woburn, MA, USA). PCR amplified libraries were qualified using high sensitivity DNA assay on an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) and quantified using a KAPA Library Quantification Kit - Illumina/Universal (Kapa Biosystems, Woburn, MA, USA).

Clusters were generated using the cBot platform (Illumina, San Diego, CA, USA). Four samples were multiplexed per lane with two lane replicates. Single-end sequencing with 50 base reads was performed on an Illumina HiSeq 2000 (Illumina, San Diego, CA, USA) following the manufacturer's guidelines.

5.3.6 Next-generation sequencing alignment

The sequenced reads were parsed based on the index to allow analysis of the data on the individual sample basis. Raw reads were analyzed using the GeneSifter® Analysis Edition pipeline (PerkinElmer, Inc., Seattle, WA, USA), a cloud-based software architecture. After quality assessment, raw reads were aligned to the *Homo sapiens* genome reference build 37.2 (GRCh37.p2) using Burrows-Wheeler Aligner (Li and Durbin 2009) with Genome Analysis Toolkit (McKenna, Hanna et al. 2010) for variant calling.

5.3.7 Differential gene expression

Three methods were used for identifying differential gene/transcript expressions:

In the GeneSifter pairwise analysis pipeline, the raw read count was normalized by total mapped million reads (RPM) and reported as log2 values. Welch's t-test (does not require equal variance between two groups) was run on the log transformed RPM between two conditions to test whether transcript levels were changed due to intracellular interactions. A Benjamini–Hochberg correction was performed for multiple testing adjustments. Genes were considered as differentially expressed when logarithmic ratio of fold change ≥ 2 and false discovery rate (FDR) < 0.05.

DESeq (Anders and Huber 2010) of R/Bioconductor (Gentleman, Carey et al. 2004), an analysis based on the negative binomial distribution, was also used for

differential expression analysis. The count values from different samples were normalized to the library size factors so that they were on a common scale. Genes with logarithmic ratio of fold change ≥ 2 and FDR < 0.05 (Benjamini-Hochberg correction) were declared significant.

EdgeR package (Robinson, McCarthy et al. 2010) of R/Bioconductor (Gentleman, Carey et al. 2004), another testing based on negative binomial model, was used for differential analysis as well. The raw count data was normalized using trimmed mean of M-values (TMM) between each pair of samples as the scale factors. After the inter-library dispersions were estimated, an exact test (Robinson and Smyth 2008) was performed to identify differentially expressed genes. Genes with logarithmic ratio of fold change ≥ 2 and FDR < 0.05 (Benjamini-Hochberg correction) were identified as significant.

Venn diagrams showing the overlaps of gene candidates from four different statistical tests were drawn using Venn Diagram Plotter (PNNL; http://omics.pnl.gov).

5.3.8 Functional annotations and pathway enrichment

Functional and pathway analysis of statistically significant gene expression changes (candidate genes) were performed using Ingenuity Pathway Analysis (Qiagen, Redwood City, CA, USA). Fisher's Exact Test was used to calculate p-values, which determined the probability of biological functions/pathways enriched in the candidate genes was due to the random effects. A Benjamini–Hochberg correction was performed for multiple testing corrections. Functions or pathways with p-values (Fisher's Exact Test) less than 0.05 were considered to be significantly relevant. The activation Z-score was calculated by using information about the direction of gene regulation (Kramer, Green et

al. 2014). Comparison analysis of pathway enrichment from different statistical tests was also carried out using the Ingenuity Pathway Analysis.

5.3.9 RT-qPCR validation

Quantitative real-time PCR was used to validate changes in gene expression. Primers for each of the target sequences were selected from PrimerBank (Wang and Seed 2003) or designed using the Primer-BLAST tool (*www.ncbi.nlm.nih.gov/tools/primer-blast/*). Multiple primer pairs were designed and evaluated at the bulk cell RNA level. Optimized primer oligos were obtained from Integrated DNA Technologies (Integrated DNA Technologies, Coralville, IA, USA).

RNA extracted from FACS sorted co-culture and mono-culture cells was used for reverse transcription and qPCR. A total volume of 20 µL of the cDNA synthesis mixture contained the following reagents: 4 µL of 5 × VILO Reaction Mix (Invitrogen, Carlsbad, CA, USA), 2 µL of 10 × SuperScript Enzyme Mix (Invitrogen, Carlsbad, CA, USA), including SuperScript III RT, RNaseOUT Recombinant Ribonuclease Inhibitor, and a proprietary "helper" protein (Invitrogen, Carlsbad, CA, USA), and 14 µL of total RNA. The contents in each tube were gently mixed and spun down, and the cDNA synthesis was performed in following thermal steps: (a) 25 °C for 10 min, (b) 42 °C for 60 min, and (c) 85 °C for 5 min to inactivate the reverse-transcriptase. The cDNA obtained from these reactions was stored at -20 °C until further use.

Prior to qPCR, cDNA obtained from each single cell was diluted 500 times by adding DEPC-treated water (Ambion, Austin, TX, USA). The qPCR runs were conducted using the following reagent mixtures: 5 µL of EXPRESS SYBR GreenER qPCR SuperMix Universal (Invitrogen, Carlsbad, CA, USA), 1 µL of each primer (4 µM), 0.1

μL of ROX Reference Dye (25 μM) (Invitrogen, Carlsbad, CA, USA), 2 μL of diluted cDNA, and 0.9 μL of DEPC-treated water (Ambion, Austin, TX, USA). For negative controls, 2 μL of DEPC-treated water was used instead of cDNA. The thermal cycling profile was set up as follows: one cycle at 95 °C for 10 min; 40 cycles consisting of 95 °C for 15 s, 60 °C for 1 min, and 80 °C for 10 s with signal detection; melt-curve analysis at 60 °C for 1 min and the temperature increased in 0.3 °C increments to 95 °C, then at 95 °C for 15 s. The experiments were run on a StepOne Real Time PCR System (Applied Biosystems, Carlsbad, CA, USA). Data analysis was carried out using the StepOne software (Applied Biosystems, Carlsbad, CA, USA). *28S* were used to normalize samples for comparison.

5.3.10 Time lapse fluorescent microscopy

For CP-D and EPC-2 cells co-culture, 3.75×10^4 CP-D and 3.75×10^4 EPC-2 cells suspension were mixed and seeded into individual wells of a Costar® 24-well-plate (Corning Life Sciences, Corning, NY, USA). For CP-D and EPC-2 mono-culture, 7.5×10^4 cells of each type were seeded into individual wells in a Costar® 24-well-plate. After 24 hours of growth, cells were stained with nucleus stain Hoechst 33342 (Life Technologies, Carlsbad, CA, USA) at 1 μg/mL at 37 °C for 15 min. Time lapse images were taken every 5 minutes for 1 hour using a Nikon Eclipse TE2000-E microscope (Nikon Inc., Melville, NY, USA). The microscope was equipped with a 20 × phase contrast objective (Nikon Inc., Melville, NY, USA) and a cooled CCD camera (CoolSNAP HQ, Photometrics, Tucson, AZ, USA), controlled by NIS-Elements imaging software (Nikon Inc., Melville, NY, USA).

5.3.11 Image analysis

Images were analyzed using custom-written MATLAB (MathWorks Inc., Natick, MA, USA) code and a Fiji TrackMate plugin (Schindelin, Arganda-Carreras et al. 2012).

5.3.12 Statistical analysis

Nonparametric Mann-Whitney test and Kruskal-Wallis test were performed to determine significance of differences. P-values of less than 0.05 were considered statistically significant.

5.4 Results and discussion

5.4.1 RNA-Seq analysis of the transcriptome in esophageal epithelial normal and dysplastic cells

Over several decades, genetic models of the neoplastic progression in Barrett's esophagus and esophageal adenocarcinoma have been discovered and proposed (Barrett, Sanchez et al. 1999, Morales, Souza et al. 2002, Maley, Galipeau et al. 2006, Merlo, Pepper et al. 2006), including evolution of neoplastic cell lineages (Barrett, Sanchez et al. 1999), epigenetically regulated alterations of *HOXB* genes (di Pietro, Lao-Sirieix et al. 2012), susceptibility loci (Levine, Ek et al. 2013), and recurrent driver events (Dulak, Stojanov et al. 2013). During pre-malignant progression, neoplasms are heterogeneous and consist of interactions between normal and neoplastic cells (Anderson, Weaver et al. 2006, Merlo, Kosoff et al. 2011, Greaves and Maley 2012). In this thesis it was hypothesized that heterotypic interactions in the tumor microenvironment can down-regulate gene transcription, slow down cellular proliferation in dysplastic cells, and thus inhibit neoplastic progression.

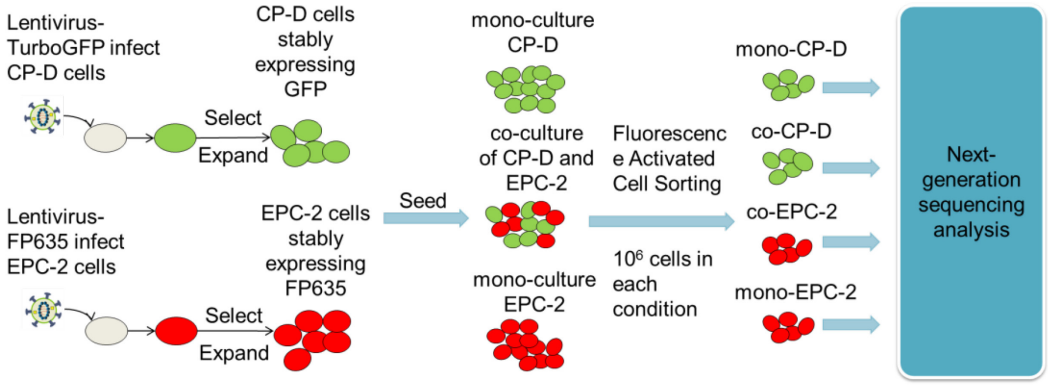


Figure 5-1 Workflow of transcriptome analysis of cell-cell interactions

To test this hypothesis, a co-culture system containing high-grade dysplastic cells (CP-D) and esophageal epithelial squamous cells (EPC-2) was constructed. In order to distinguish different cell types in the co-culture, CP-D cells were transfected with lentiviral-*TurboGFP* to stably express cytosolic *TurboGFP*; EPC-2 cells were transfected with lentiviral-*FP635* to stably express cytosolic *FP635*. After co-culturing CP-D and EPC-2 cells for 24 hours, fluorescent activated cell sorting (FACS) was used to separate two cell types and performed whole transcriptome sequencing, using mono-cultured CP-D and EPC-2 cells as controls (Figure 5-1). Each condition—co-cultured CP-D cells, mono-cultured CP-D cells, co-cultured EPC-2 cells and mono-cultured EPC-2 cells—contained three biological replicates. Four samples were multiplexed per lane with two lane replicates on an Illumina HiSeq 2000 sequencer. The raw reads were aligned to the *Homo sapiens* genome reference build 37.2 (GRCh37.p2) using Burrows-Wheeler Aligner (Li and Durbin 2009) with Genome Analysis Toolkit (McKenna, Hanna et al. 2010) for variant calling. From approximately 100 million single-end 50-bp sequencing reads, a median of 72 million mapped reads per sample were recovered. The majority of mapped reads are annotated gene features (exon-intron regions) and rRNA or snRNA, followed by intergenic regions.

Three methods were applied to identify differentially expressed genes: DESeq (Anders and Huber 2010), EdgeR (Robinson and Smyth 2008, Robinson, McCarthy et al. 2010) and welch's t-test (Dudoit, Yang et al. 2002). For multiple testing corrections, Benjamini–Hochberg correction (Hochberg and Benjamini 1990) was performed to obtain the false discovery rate (FDR). The pairwise comparisons included three pairs: (1) co-cultured CP-D vs. mono-cultured CP-D, (2) co-cultured EPC-2 vs. mono-cultured EPC-2, and (3) mono-cultured EPC-2 vs. mono-cultured CP-D. A threshold of

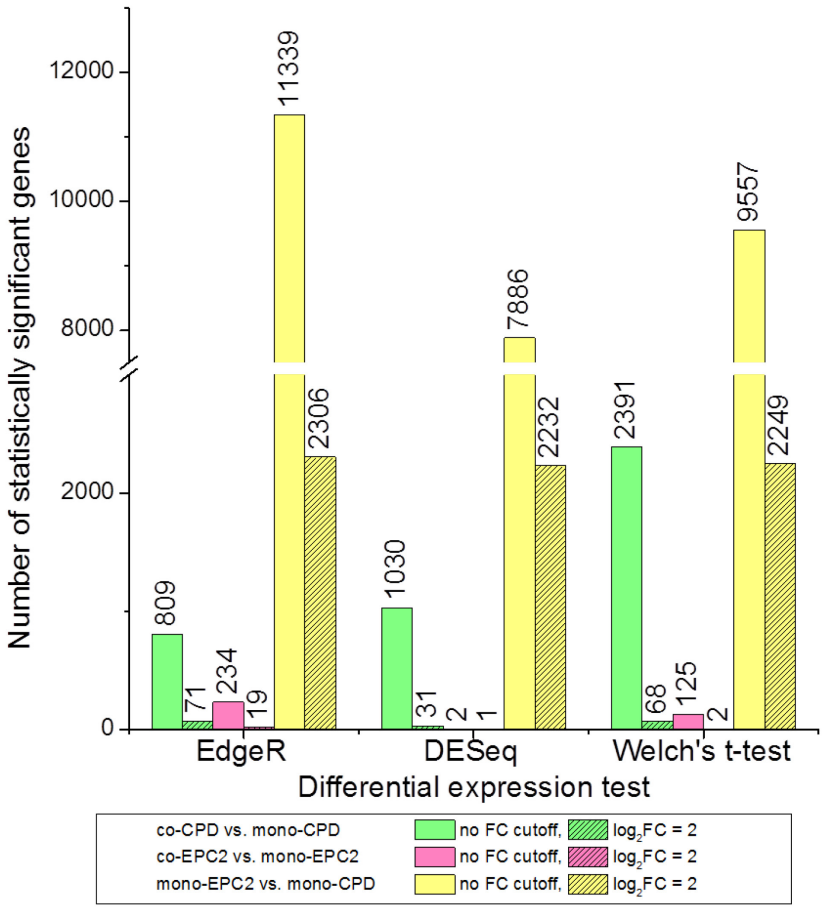


Figure 5-2 Number of differentially expressed genes

Benjamini-Hochberg correction at FDR < 0.05 and another cutoff of logarithmic transformed fold change values (log_2FC) at 2 were set. Identified by all three methods at FDR < 0.05, the numbers of differentially expressed genes with $log_2FC = 2$ cutoff were much lower than those without $log_2FC = 2$ cutoff (Figure 5-2). EdgeR found the highest number of differentially expressed genes, while DESeq returned relatively few. The number of differentially expressed genes determined by all three methods showed the same trend in different pairwise comparison groups: more differentially expressed genes were found in the co-cultured CP-D vs. mono-cultured CP-D group than the co-cultured EPC-2 vs. mono-cultured EPC-2 group; the mono-cultured EPC-2 vs. mono-cultured CP-D group had the highest number among all three (Figure 5-2). This indicated that gene expression profiles in CP-D cells changed more due to heterotypic interactions than those in EPC-2 cells. The gene expression difference between CP-D and EPC-2 cells was the largest.

The differentially expressed genes in three pairwise groups determined by different tests with FDR < 0.05 and $log_2FC \geq 2$ (Figure 5-3) were also compared. All three methods were concordant with each other, showing a number of overlapping genes in the co-cultured CP-D vs. mono-cultured CP-D group and the mono-cultured EPC-2 vs. mono-cultured CP-D groups. Particularly, a large portion of genes with differential expression found by DESeq were also identified by EdgeR. Both EdgeR and welch's t-test found unique differentially expressed genes in the co-cultured CP-D and mono-cultured CP-D group, which were not shared by other methods. Previous evaluations showed that DESeq was often conservative, while EdgeR was too liberal and yields potential false positives (Soneson and Delorenzi 2013). In order to retain the ability to

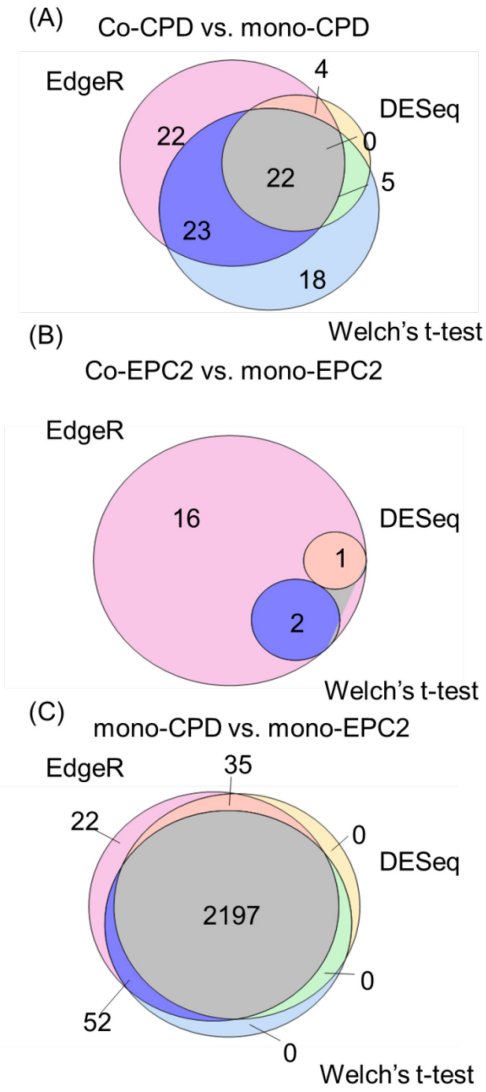


Figure 5-3 Venn diagrams of differentially expressed genes

detect the truly differentially expressed genes, the primary step of functional enrichment included all the differentially expressed genes found by three methods. However, to control false discovery rates, further functional annotations focused mainly on the genes identified by DESeq.

5.4.2 Function enrichment of differentially expressed genes in heterotypic interactions

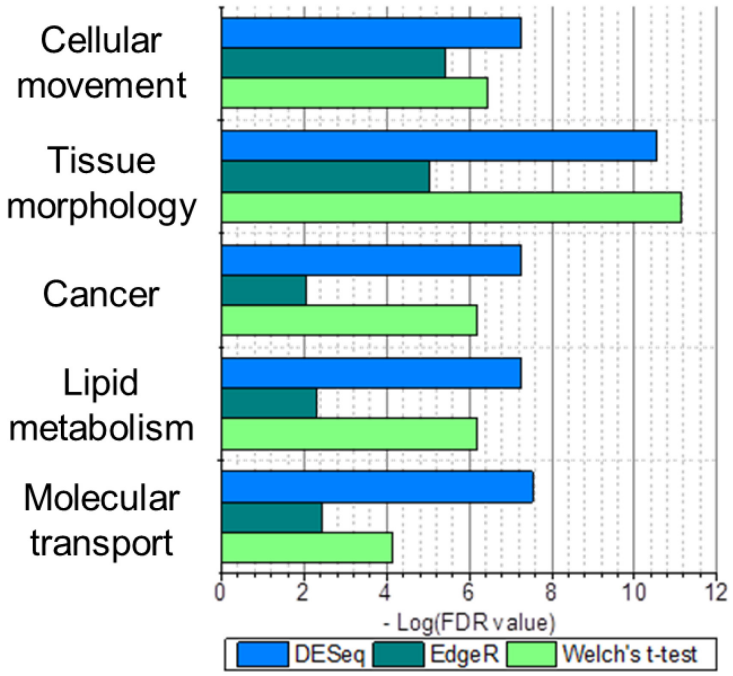


Figure 5-4 Top 5 functions enriched in the co-CPD vs. mono-CPD group

To discover pathways related to transcriptome alterations, Ingenuity Pathway analysis (IPA) was performed to identify functional categories associated with differentially expressed genes. The list of differentially expressed genes identified independently by DESeq, EdgeR and welch's t-test was used for IPA core analysis.

In the co-cultured CP-D vs. mono-cultured CP-D group, 67 significant bio-function terms in genes were found by DESeq, 74 in genes by EdgeR, and 70 in genes by welch's t-test (Fisher's Exact Test and Benjamini-Hochberg correction, FDR < 0.05). Among the significant bio-functions enriched from genes found by all three methods, cellular movement ranked among the top 5 functions in –log(FDR) (Figure 5-4). Most of

the other bio-function terms and pathways were also closely related to neoplasia and tumorigenesis, such as tissue morphology, cancer, lipid metabolism and molecular transport. How those functions are regulated in co-cultured CP-D cells in the DESeq identified genes (threshold: activation z-score = 2, Fisher's Exact Test, Figure 5-5) was investigated. Interestingly, most of the bio-functions were suppressed in co-cultured CP-D. The biggest category of down-regulated features was cellular movement, which included a panel of movement related functions, such as invasion and migration. Other cellular movement associated functions—organization of cytoplasm, organization of cytoskeleton and microtubule dynamics—were also inhibited in co-cultured CP-D cells. Co-culturing with EPC-2 cells decreased cancer related functions (metastasis, neoplasia and cell production) as well.

In the co-cultured EPC-2 vs. mono-cultured EPC-2 group, only a few differentially expressed genes (1 by DESeq, 20 by EdgeR and 2 by welch's t-test) were found (Figure 5-2, Figure 5-3B). Most of the enriched functional categories were not significantly changed (threshold: activation z-score = 2, p = 0.05, Fisher's Exact Test).

The transcriptomes in CP-D and EPC-2 cells when they were in mono-culture were compared. A total of 209 significant bio-function terms in genes was determined by DESeq, 69 by EdgeR, and 67 by welch's t-test (Fisher's Exact Test and Benjamini-

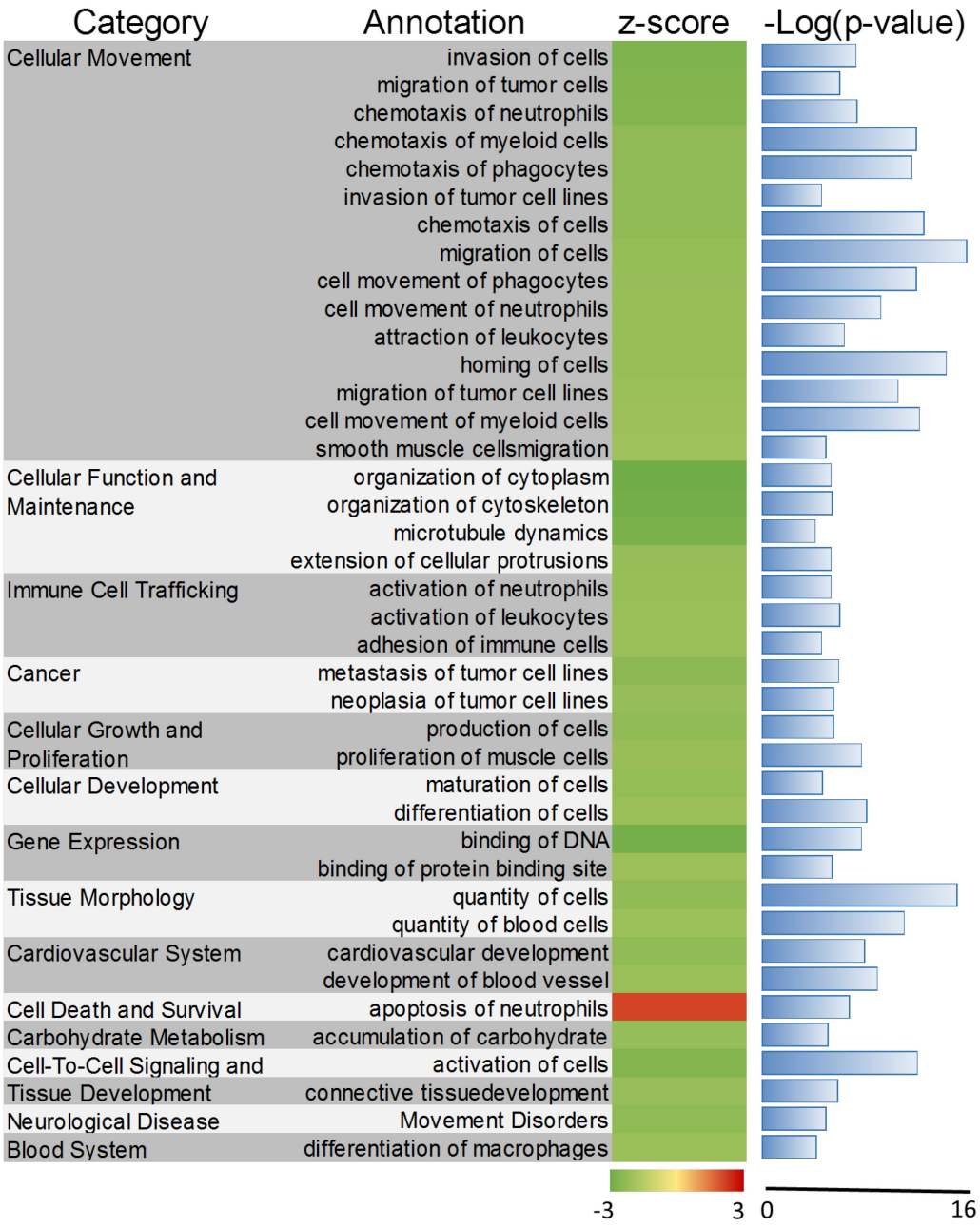


Figure 5-5 Functional annotations in the co-CPD vs. mono-CPD group

Hochberg correction, FDR < 0.05). Cellular movement, cancer, cellular development, cell growth and proliferation, as well as cell death and survival ranked as top 5 significant functions (Figure 5-6). Most of them were involved in neoplasia and tumorigenesis,

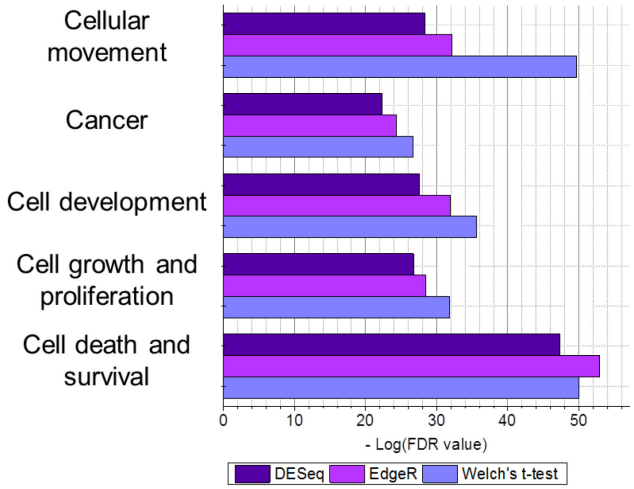


Figure 5-6 Top 5 functions enriched in the mono-CPD vs. mono-EPC2 group

which emphasized that the transcriptional difference is due to the cell type difference. The majority of significantly altered functions were activated rather than inhibited (threshold: activation z-score = 2, Fisher's Exact Test, Figure 5-7). Functions involved in cancer, cellular functions and maintenance, cellular development, tissue development, tissue morphology and cellular movement were more active in CP-D cells than in EPC-2 cells, all of which may participate in neoplastic progression. Morphology of cells was less active in CP-D cells, reflecting the changes in cellular shape and size (cytology) as well as cell cohesion (architecture) and polarity in the CP-D cells.

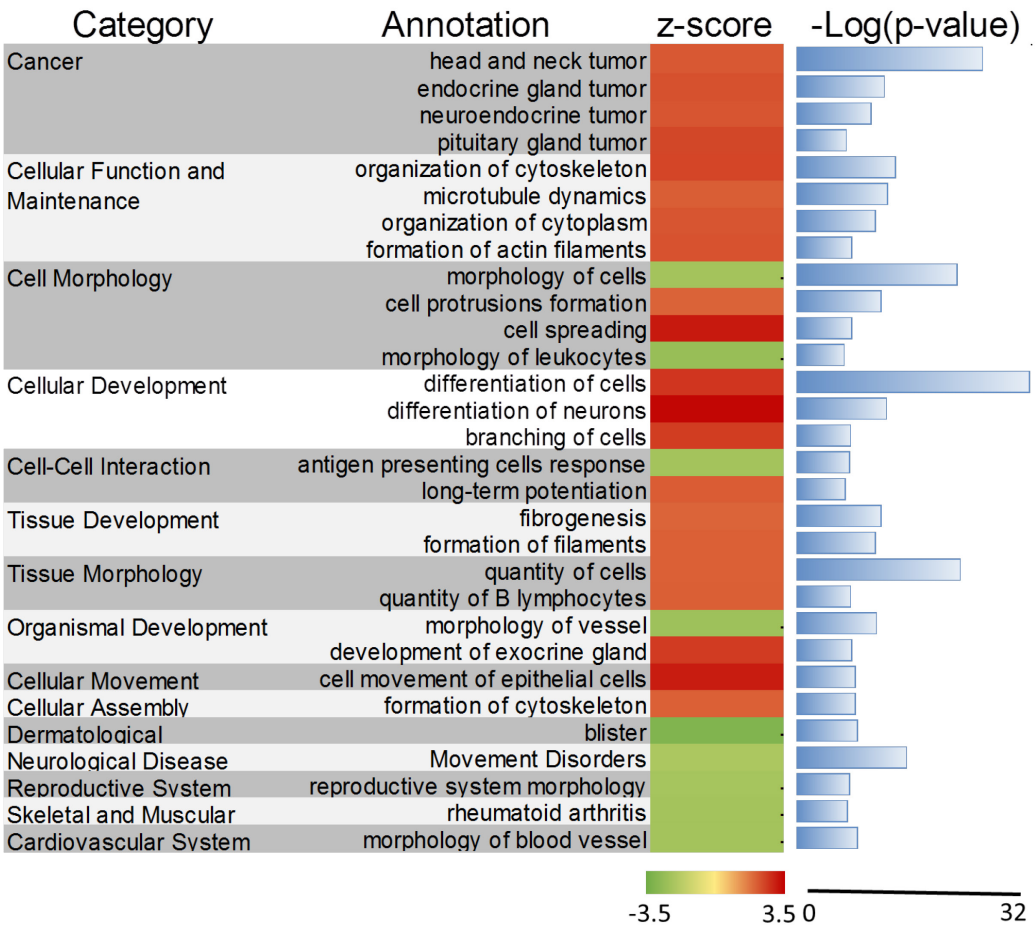


Figure 5-7 Functional annotations in the mono-CPD vs. mono-EPC2 group

5.4.3 Upstream regulator analysis reveals *TGFβ* and *EGF* signaling networks are inhibited in co-cultured CP-D cells

To examine the upstream molecules that trigger the transcriptional changes, the upstream regulators and their networks were analyzed. Ingenuity Pathway Analysis was used to predict the status of upstream molecules and their biological roles in heterotypic interactions.

A total of 40 upstream regulators activated or inhibited (threshold: activation z-score = 2, $p < 0.05$, Fisher's Exact Test) with expression level (log ratio) changes were found. These regulators include growth factors, cytokines, transmembrane receptors,

transcription regulators, enzymes, kinases, ion channel, ligand-dependent nuclear receptors, peptidase and other signal transducers (Figure 5-8). The regulators were sorted based on their enrichment p-value within each category and identified regulators of pivotal importance in heterotypic interactions.

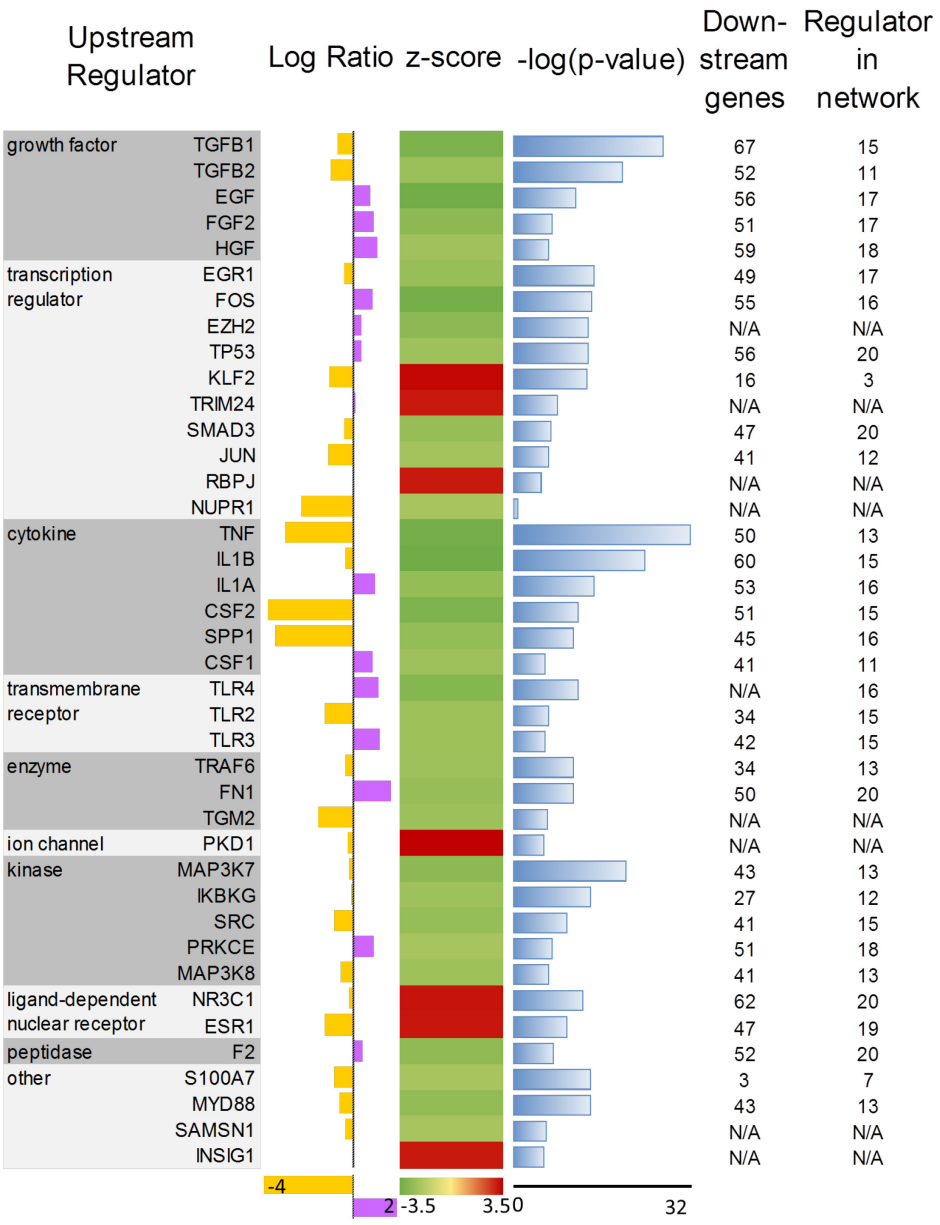


Figure 5-8 Upstream regulator analysis in the co-CPD vs. mono-CPD group
Log ratio: Log ratio of gene expressions in each regulator. Downstream genes: downstream genes in the data-set. Regulators in network: downstream regulators

Growth factors *TGFβ1* (transforming growth factor β) and *TGFβ2* (transforming growth factor β2) were predicted to be down-regulated in co-cultured CP-D cells. They appeared in the upstream of 19 regulators and affected the expression of 70 genes in their

mechanistic network. Among 37 genes which were targeted by *TGFβ1* and *TGFβ2* in the dataset (Figure 5-9), 32 genes were involved in epithelial neoplasia (p = 3.08E11) and 28 involved in cellular movement (p = 2.24E-17, both p-values calculated by Fisher's Exact Test). *TGFβ* has dual effects in neoplasia and later stages tumor progression. It is involved in many aspects of the communications between cancer cells and non-neoplastic cells in the tumor microenvironment (Ungefroren, Sebens et al. 2011). As an early tumor suppressor, *TGFβ* inhibits proliferation and induces apoptosis. During esophageal adenocarcinoma progression, *TGFβ* loses its anti-proliferative function; instead, it mediates epithelial-to-mesenchymal transition by several mechanisms and promotes esophageal carcinogenesis (Jankowski, Harrison et al. 2000, Rees, Onwuegbusi et al. 2006). Intriguingly, *DOCK2*, a target gene of *TGFβ1*, is down-regulated in co-cultured CP-D cells with single nucleotide variants in one intron region. The splice-site mutant human *DOCK2* gene has been identified as a recurrent drive event in esophageal adenocarcinoma in an exome and whole-genome sequencing study of esophageal adenocarcinoma. Mutant *DOCK2* family members can enhance cellular motility and invasion (Dulak, Stojanov et al. 2013). These results suggest that heterotypic interaction can affect neoplastic progression by changing cellular motility via *TGFβ*-mediated networks.

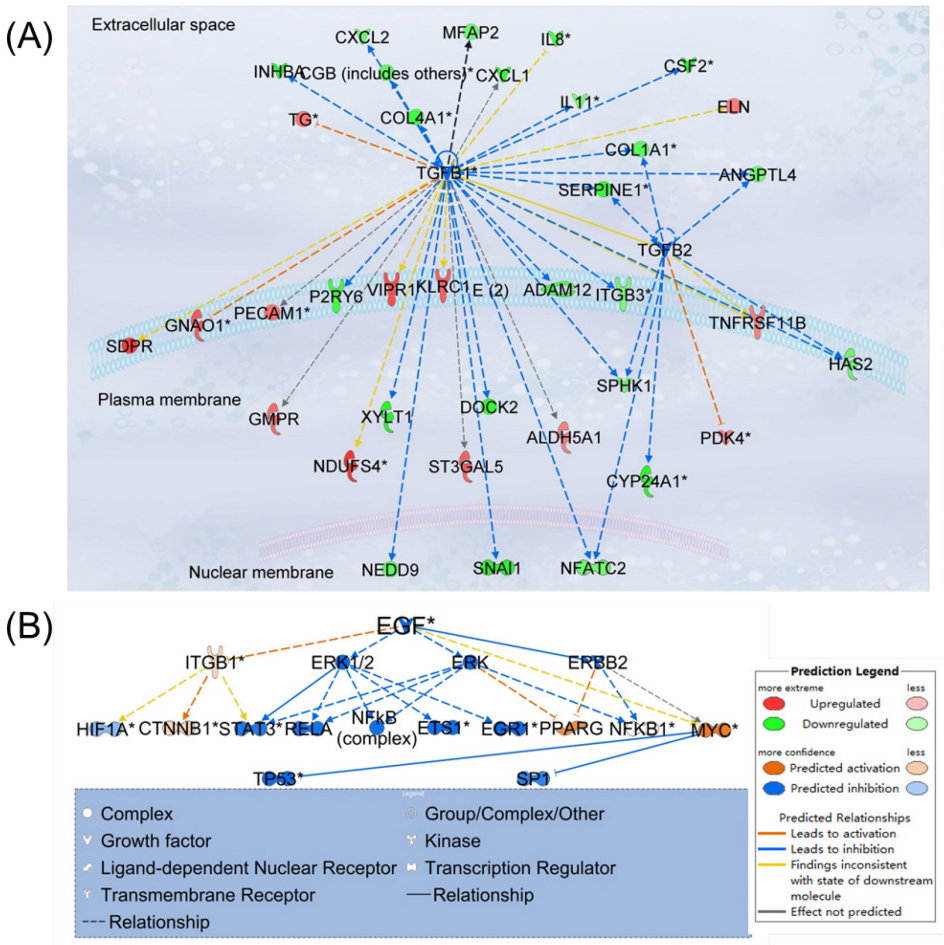


Figure 5-9 Differentially expressed genes in the networks
(A) *TGFB1* and *TGFB2* network, (B) *EGF* mechanistic network shows the upstream regulators regulated by *EGF*.

It was also observed that *EGF* (epidermal growth factor) was predicted to be down-regulated in co-cultured CP-D cells. It affected 57 genes in its mechanistic network and acted in the upstream of 10 transcription regulators as well as 7 other regulators (Figure 5-9). All of these regulators in the network were involved in apoptosis (p = 1.50E-15), proliferation (p = 1.95E-14), differentiation (p = 7.32E-14) and migration (p = 9.65E-14, all p-values calculated by Fisher's Exact Test). Downstream of *EGF*, a panel of genes was down-regulated, including *Erb* and *TP53*. Increased expression of *EGF*

receptors has been implicated in Barrett's esophagus progression (Jankowski 1993, Jankowski, Wright et al. 1999, Pande, Iyer et al. 2008). Abnormal *Erb* family of tyrosine-kinase receptors (Miller, Moy et al. 2003) and *TP53* (Neshat, Sanchez et al. 1994) signaling have been previously implicated in the evolution of Barrett's metaplasia-dysplasia-adenocarcinoma. Therefore, inhibited *EGF* signaling networks indicated that *EGF* can potentially slow down the neoplastic progression in CP-D cells.

Taken together, the inhibition of growth factors and their downstream transcription factors affected cellular motility and proliferation. Given that dysplastic cells have relaxed proliferation controls (Jankowski, Wright et al. 1999) and esophageal adenocarcinoma is an early invasive cancer (Dulak, Stojanov et al. 2013), heterotypic interactions may down-regulate growth factor signaling and contribute to these phenotypes.

5.4.4 Co-culture of CP-D and EPC-2 cells changed the proliferation and motility of both cell lines

To further investigate whether cell motility and proliferation were affected by heterotypic interactions, fluorescence microscopy was performed to count cell numbers and track cell movements. CP-D and EPC-2 co-culture at a 1:1 ratio, CP-D mono-culture and EPC-2 mono-culture in a Costar® 12-well-plate (Corning, Corning Life Sciences, Corning, NY, USA) were seeded with the same density in four replicates. After growing the cells for 24 hours, the first replicate group (four wells) of cells was stained with Hoechst 33342 (Invitrogen, Carlsbad, CA, USA) and they were imaged (Figure 5-10) using time-lapse fluorescence microscopy. After 48, 72 and 96 hours, the second, third and fourth groups of cells were stained with Hoechst and visualized.

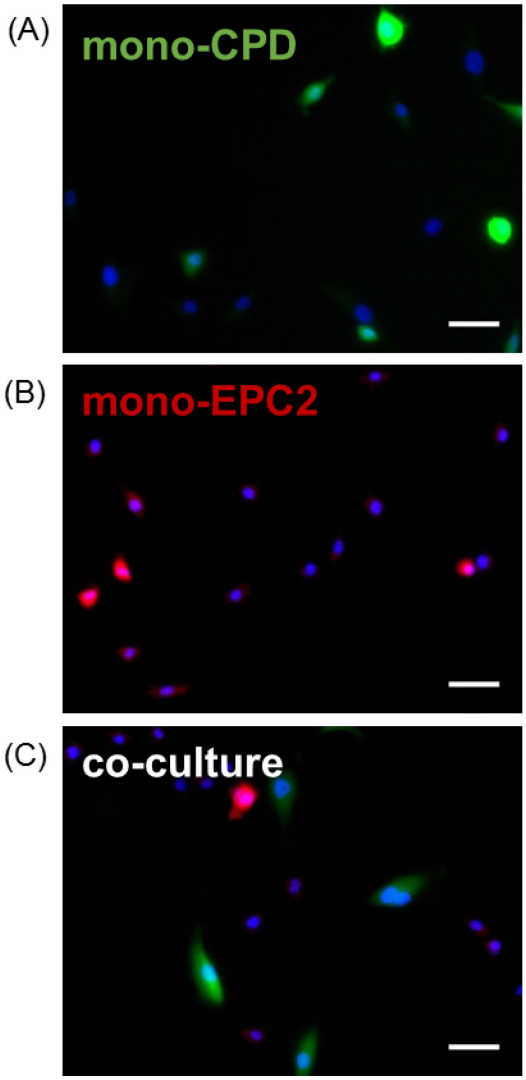


Figure 5-10 Fluorescence microscopy of Hoechst 33342 stained cells (A) mono-CPD cells, (B) mono-EPC2 cells, (C) co-culture of CPD and EPC2 cells. Scale bar: 50 μm.

After counting the cells (Figure 5-11) it was found that the proliferation of CP-D cells was slowed down in the co-culture groups compared with the mono-culture groups. Interestingly, the proliferation of EPC-2 cells in co-culture was also slower than mono-culture, although relevant gene expression did not change significantly according to the RNA-Seq analysis. This suggests that heterotypic interaction suppressed the proliferation of both dysplastic and normal cell lines.

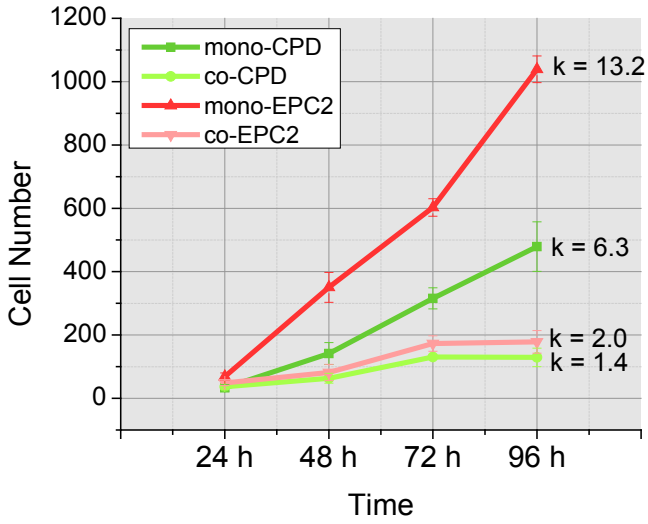


Figure 5-11 Proliferation of CP-D and EPC-2 cells in mono-culture and co-culture
N = 3. Slopes (k) are shown after the line.

Cell motility was analyzed by comparing the speed, displacement and directionality ratio (Euclidean distance/displacement) of co-culture and mono-culture CP-D and EPC-2 cells, respectively. After 24 hours, co-culture CP-D cells traveled significantly further than mono-culture CP-D cells during the one-hour recording ($p = 5E-6$, Mann-Whitney test, $\alpha = 0.05$, Figure 5-12A). They also moved in a significantly more directional manner than mono-culture CP-D cells ($p = 5E-7$, Mann-Whitney test, $\alpha = 0.05$, Figure 5-12C).

On the contrary, EPC-2 cells moved significantly slower in co-culture than in mono-culture ($p = 2E-4$, Mann-Whitney test, $\alpha = 0.05$, Figure 5-12A). The displacement of EPC-2 cells in co-culture was significantly shorter compared with mono-culture ($p = 2E-2$, Mann-Whitney test, $\alpha = 0.05$, Figure 5-12B). Therefore, heterotypic interactions altered the motility of both cell lines in several ways: increased the displacement and

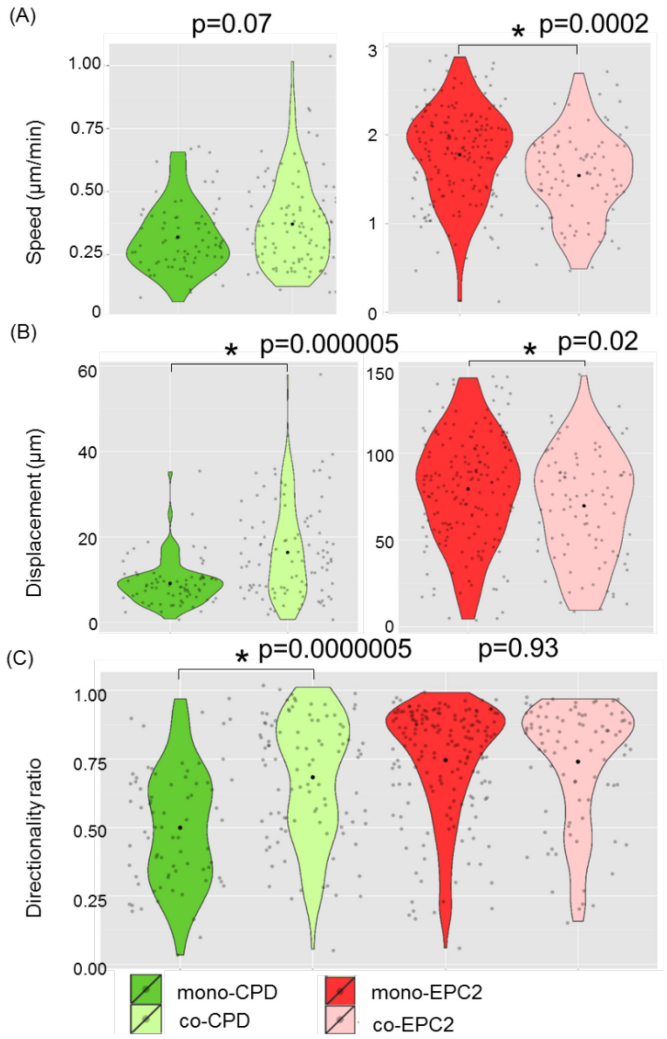


Figure 5-12 Migration of CP-D and EPC-2 cells in mono-culture and co-culture
Violin plots of migration speed (A), displacement (B) and directionality ratio (C). *Black dot*: mean value, *grey dot*: individual data point.

directionality of CP-D cells, slowed down the speed of EPC-2 cells and decreased the displacement of EPC-2 cells.

5.5 Conclusion

In this study, the effects of intercellular interactions between normal and late-dysplastic Barrett's esophagus cells on transcriptome and phenotypic levels were examined. It was investigated how homotypic and heterotypic intercellular interactions alter gene expressions using next-generation sequencing. Bioinformatic analysis of the next-generation sequencing results identified differentially expressed genes in high-grade dysplastic cells as a result of heterotypic and homotypic interactions. These genes were enriched in cellular movement and other cancer-related pathways. Heterotypic interactions suppressed downstream genes of *TGFβ* and *EGF* pathways in late-dysplastic cells. Taken the results of proliferation and motility studies together, the inhibition of *EGF* signaling networks seem to slow down the proliferation in co-cultured CP-D cells. *TGFβ* signaling networks may be undergoing the transition from anti-proliferation to promoting motility. Normal cells seem to inhibit the growth of dysplastic cells via growth factor signaling pathways. The fraction of normal to dysplastic cells can be used as the risk stratification marker for pre-malignant progression in Barrett's esophagus. *TGFβ*, *EGF* and their downstream genes have great potential to become biomarkers for early diagnosis and early treatment of esophageal adenocarcinoma.

5.6 Acknowledgments

I thank Tong Fu for help with fluorescence-activated cell sorting, Jason Steel and Kristina Buss from Center for Personalized Diagnostic for sequencing services. This work was supported by the following grants to Dr. Deirdre R. Meldrum from the NIH National Human Genome Research Institute, Centers of Excellence in Genomic Sciences, Grant Numbers 5 P50HG002360 and 3P50HG002360-10S1.

CHAPTER 6

ALTERATIONS IN GENE EXPRESSION LEVELS AND METABOLIC PHENOTYPE IN RESPONSE TO HYPOXIC SELECTION IN PRE-MALIGNANT BARRETT'S ESOPHAGUS CELLS

6.1 Abstract

Hypoxia acts as important selective pressure in the clonal evolution of Barrett's esophagus neoplastic progression. It can trigger metabolic reprogramming, one of the emerging hallmarks of cancer. Metabolic phenotype measurements indicated that Barrett's esophagus cells undergo a series of changes under hypoxic conditions. However, the mechanisms underlying these changes remain less clear. This study provides a comparison of the alterations on transcriptome and metabolic phenotype in hypoxia-adapted cells and age-matched normoxic control cell lines in Barrett's esophagus. Gene expression differences between hypoxia and control cell lines revealed alterations in the metabolic processes, such as glycolysis and oxidative phosphorylation related genes. Differential gene expression analysis uncovered drastic differences between one pair of hypoxia-adapted high-grade dysplastic cells and age-matched control cells. It was discovered that a dynamic evolutionary process of cells adapt to hypoxia. Compared with CPA and CPD cells, dysplastic CPB and CPC cells are at transitional states and make lots of gene expression changes to adapt to hypoxia. These findings open new ways for designing diagnosis and treatment strategies for Barrett's esophagus and adenocarcinoma. Physicians will need to tackle the functional plasticity of hypoxia adapted dysplastic cells when designing metabolic targets for treatment. Also, hypoxia adaptation can be used as

early diagnosis, treatment and risk stratification biosignatures in pre-malignant progression of Barrett's esophagus.

6.2 Introduction

Hypoxia, the lack of oxygen supply, is a risk factor for cancer. It arises as cells grow far away from the vasculature and live in a local hypoxic environment. Acute oxygen level changes between hypoxia and re-oxygeneration can generate reactive oxygen species. Such hypoxic conditions constitute a selective pressure for cellular growth and proliferation. To survive and thrive under hypoxia, cells need to reprogram their gene expression and phenotypic profiles (Harris 2002) in various aspects: angiogenesis (Oh, Takagi et al. 1999), metabolism (Vannucci, Seaman et al. 1996, Brugarolas, Lei et al. 2004, Kim, Tchernyshyov et al. 2006), proliferation (Zundel, Schindler et al. 2000, Lal, Peters et al. 2001), apoptosis or necrosis (Bruick 2000, Velde, Cizeau et al. 2000, Sowter, Ratcliffe et al. 2001, Suzuki, Tomida et al. 2001), genetic instability (Bristow and Hill 2008) and migration (Imai, Horiuchi et al. 2003, Pennacchietti, Michieli et al. 2003, Semenza 2003).

The acid-bile reflux in Barrett's esophagus damages esophageal squamous epithelial cells and may cause deep ulceration. This generates a periodic hypoxic environment for the esophageal epithelial cells (Suchorolski, Paulson et al. 2013). A panel of metabolic related or hypoxia response genes, *Glut-1* (Younes, Ertan et al. 1997), pyruvate kinase isoform M2 (*PKM2*) (Koss, Harrison et al. 2004), *VEGF* (Couvelard, Paraf et al. 2000)and erythropoietin (*EPO*) (Griffiths, Pritchard et al. 2007) have been reported to be up-regulated in Barrett's esophagus tissue. A series of metabolic changes are also observed along the sequence of pre-malignant progression. Barrett's metaplastic cells generate ATP through normal mitochondrial oxidative phosphorylation. In the intermediate stages of Barrett's dysplastic cells, mitochondria are still active, electron

transport chain remains functional, but the cells boost up glycolysis in response to the substrate (Suchorolski, Paulson et al. 2013). However, the transcriptome and metabolic profile changes in Barrett's esophagus cells adapted to hypoxia are less clear.

In this study, eight hypoxia-adapted (HCPA, HCPB, HCPC, and HCPD) and their age-matched normoxic controls (CCPA, CCPB, CCPC, and CCPD) esophageal epithelial cell lines were used. These cell lines represent a sequence of pre-malignant progression in Barrett's esophagus from metaplasia to high-grade dysplasia. Whole transcriptome analysis of hypoxia-adapted cells and age-matched normoxic control cell lines after acute hypoxia treatment showed gene expression differences among the eight cell lines. Gene expressions differences between hypoxia and control cell lines indicated alterations in the metabolic process. Differential gene expression analysis revealed a drastic difference between HCPC and CCPC cells. The interplay between *TGFβ* and hypoxia induced responses were suppressed in HCPC cells. It was also found that, in contrast to hypoxia and control CPA, CPB and CPC cells, mitochondrial membrane potentials in HCPD cells were higher than CCPD cells. A better understanding of the resistance to the hypoxia mechanism will help physicians design new metabolic therapies for cancer.

6.3 Experiments

6.3.1 Cell lines

Barrett's Esophageal cell lines were derived from premalignant Barrett's esophagus tissue (Palanca-Wessels, Barrett et al. 1998), representing early (CP-A) and late (CP-B, CP-C and CP-D) stages in neoplastic progression.

Hypoxia-resistant Barrett's esophagus cells were selected following this procedure: cells were incubated under hypoxic conditions (<1% O_2) for 18 hours to reach 10% surviving rate. The surviving cells were cultured under normoxic conditions (22% O_2) and returned to the hypoxic condition after expansion. The cells underwent six rounds of selection and the surviving population was designated as hypoxia-selected cells (HCPA, HCPB, HCPC and HCPD). Their age-matched control cells are named as CCPA, CCPB, CCPC and CCPD.

Cells were cultured using Gibco keratinocyte cell growth medium (Invitrogen, Carlsbad, CA, USA), supplemented with hEGF (Peprotech, Rocky Hill, NJ, USA) at 2.5 μg/500 mL, BPE (bovine pituitary extract) at 25 mg/500 mL, 5% fetal bovine serum (FBS) (Invitrogen, Carlsbad, CA, USA) and penicillin–streptomycin solution (Invitrogen, Carlsbad, CA, USA) at 100 units/100 μg/mL. Cells were maintained at 37 °C in a humidified atmosphere containing 5% CO_2.

6.3.2 Acute hypoxia treatment

Cells cultured in 75 cm^2 flasks (Corning, Corning, NY) at 80% confluency were placed into a 37 °C incubator supplied with 5% CO_2 and 1% O_2. After 2 hours of incubation, cells were immediately lysed using 600 μL of Buffer RLT in an RNeasy mini kit (Qiagen, Valencia, CA, USA) according to the manufacture's protocol.

6.3.3 RNA extraction

Total RNA was extracted from the sorted cells using RNeasy mini kit per manufacture's recommendation. DNA was removed by the silica-gel membrane, the spin-column technology and DNase treatment. RNase-free DNase set (Qiagen, Valencia, CA, USA), which consists of 10 μL of DNase 1 stock solution in 70 μL of proprietary buffer RDD, was used to remove residual DNA for 15 min at 20–30 °C. RNA was eluted by adding 30 μL of RNase-free water into the silica-gel membrane and stored at -80 °C.

6.3.4 Whole transcriptome amplification

The whole transcriptome of the hypoxia-selected and age-matched normoxic control cells was amplified using Nugen Ovation RNA-Seq V2 kits (Nugen Technologies, San Carlos, CA, USA) per manufacturers' instructions on an Apollo 324 Library Preparation System (IntegenX, Pleasanton, CA, USA) as previously described in 5.3.4 Whole transcriptome amplification (Kurn, Chen et al. 2005).

6.3.5 Library preparation and Illumina sequencing

Illumina sequencing libraries were prepared on an Apollo 324 Library Preparation System using the PrepX™ ILM DNA Library Preparation kit (IntegenX, Pleasanton, CA, USA) with five different barcoded adapters for multiplexing. The adapter-ligated libraries were amplified by 10 cycles of PCR using KAPA HiFi Library Amplification Kits (Kapa Biosystems, Woburn, MA, USA). Amplified libraries were qualified using high sensitivity DNA assay on an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) and quantified using a KAPA Library Quantification Kit - Illumina/Universal (Kapa Biosystems, Woburn, MA, USA).

Clusters were generated using the cBot platform (Illumina, San Diego, CA, USA). Five samples were multiplexed per lane with two lane replicates. Single-end sequencing with 50 base reads was performed on an Illumina HiSeq 2000 (Illumina, San Diego, CA, USA) following the manufacturer's guidelines.

6.3.6 Next-generation sequencing alignment

The sequenced reads were parsed and analyzed using the GeneSifter® Analysis Edition pipeline (PerkinElmer, Inc., Seattle, WA, USA) as previously described in 5.3.6 Next-generation sequencing alignment. After quality assessment, raw reads were aligned to the *Homo sapiens* genome reference build 37.2 (GRCh37.p2) using Burrows-Wheeler Aligner (Li and Homer 2010) with Genome Analysis Toolkit (McKenna, Hanna et al. 2010) for variant calling.

6.3.7 Multiple condition gene expression comparison

In the GeneSifter project analysis pipeline, the raw read count was normalized by total mapped million reads (RPM) and reported as log2 values. A nonparametric Kruskal-Wallis test was run on the log transformed RPM between eight conditions to test whether transcript levels were significantly different. A Benjamini–Hochberg correction was performed for multiple testing corrections. Genes were considered as differentially expressed when the false discovery rate (FDR) was < 0.05.

A database for Annotation, Visualization, and Integrated Discovery (*http://david.abcc.ncifcrf.gov*) (Huang, Sherman et al. 2007) was used for functional classification and annotation of candidate genes in the glycolysis and oxidative phosphorylation pathways.

6.3.8 Differential gene expression

Differential gene expression was identified using three independent methods as previously described in Section 5.3.7 Differential gene expression.

Venn diagrams showing the overlaps of gene candidates from four different statistical tests were drawn using the Venn Diagram Plotter (PNNL; *http://omics.pnl.gov*).

6.3.9 Functional annotations and pathway enrichment

Functional and pathway analysis of differentially expressed genes were performed using the Ingenuity Pathway Analysis (Qiagen, Redwood City, CA, USA). Fisher's Exact Test was used to calculate p-values, which determined the probability of biological functions/pathways enriched in the candidate genes was due to random effects. A Benjamini–Hochberg correction was performed for multiple testing corrections. The activation Z-score was calculated by using information about the direction of gene regulation (Kramer, Green et al. 2014). Functions or pathways with p-values (Fisher's Exact Test) less than 0.05 were considered to be significantly relevant.

6.3.10 Mitochondrial membrane potential measurement using flow cytometry

Mitochondrial membrane potential was quantified with flow cytometry analysis using the potentiometric dye JC-1. Cells cultured in 25 cm^2 flasks (Corning, Corning, NY) at 80% confluency were incubated in Keratinocyte serum-free medium without serum containing JC-1 (100 ng/mL) for 15 min. Cells were washed twice with PBS, trypsinized and re-suspended in 300 µL of Keratinocyte serum-free medium.

Flow cytometry was performed using a Becton Dickinson FACS Calibur flow cytometer (Becton Dickinson, Franklin Lakes, NJ, USA). FL1 and FL2 corresponds to the 530/30 BP and 585/42 BP emission filters respectively. JC-1 fluorescence intensity

signals of 10,000 events (cells) were collected in each condition. The flow cytometry data was analyzed using custom-written MATLAB (MathWorks Inc., Natick, MA, USA) code.

6.3.11 Statistical analysis

Nonparametric Mann-Whitney test and Kruskal-Wallis tests were performed to determine the significance of differences. A p-value less than 0.05 was considered to be statistically significant.

6.4 Results and discussion

6.4.1 Mitochondrial membrane potential measurement of hypoxia-adapted and age-matched control cells

Mitochondrial membrane potential is an important parameter of mitochondrial function. It is generated and maintained by mitochondrial electron transport chain, mainly by oxidative phosphorylation. The collapse of mitochondrial membrane potential is an early event in apoptosis. High mitochondrial membrane potential is correlated with active oxidative phosphorylation. Several human cancers have high mitochondrial membrane potential and are resistant to apoptotic cell death (Vander Heiden, Chandel et al. 1997). JC-1 is a widely-used fluorescent probe for measuring mitochondrial membrane potential (Cossarizza, Baccaranicontri et al. 1993). Flow cytometry analysis of JC-1 stained cells revealed marked differences between hypoxia-adapted and age-matched control cells (Figure 6-1). Normalized against CCPA cells, the relative mitochondrial membrane potentials were significantly lower in hypoxia-adapted cells than control cells (CPA, CPB, and CPC); while it was significantly higher in hypoxia-adapted HCPD cells than control CCPD cells. This suggests that at later stages of progression, hypoxia adapted CPD cells have higher energy needs for proliferation, so they boost up their energy production.

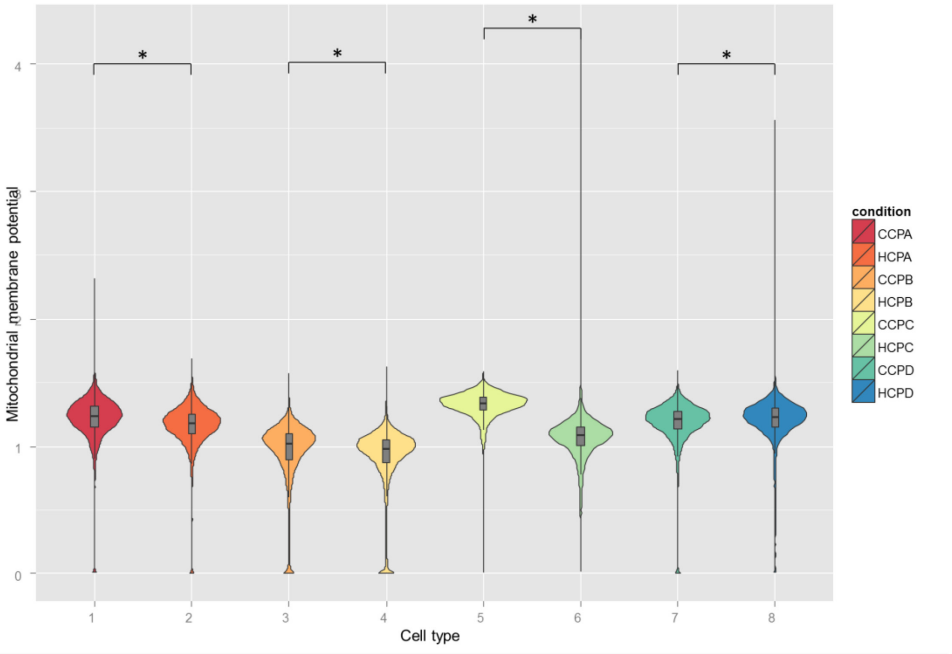


Figure 6-1 Violin plot of mitochondrial membrane potentials
The grey boxes inside the violin plot are the box plot of mitochondrial membrane potential distributions. Mann-Whitney test, *$p < 0.05$.

Mitochondrial heterogeneity within and between different types of cells has been reported before (Huang, Fowler et al. 2004, Wang, Shi et al. 2013). The descriptive statistics (Table 7) showed that the kurtosis values were comparable within the same hypoxia and control pairs (HCPA and CCPA, HCPB and CCPB, HCPC and CCPC). They were higher in HCPC and CCPC cells than the other two pairs, indicating higher heterogeneity in HCPC and CCPC cells. However, the hypoxia-adaptation did not significantly alter the heterogeneity. The kurtosis value in HCPD cells was the highest among all eight cell lines, suggesting a highly heterogeneous population in hypoxia-adapted CPD cells. The skewness values were negative and comparable in both hypoxia and control pairs of CPA and CPB cells. HCPC and HCPD cells have positive skewness values; while CCPC and CCPD cells have negative values. The skewness value of HCPD

cells was the highest, indicating that extremely high mitochondrial membrane potentials appeared in only a few cells. Taken together, HCPA, HCPB and HCPC cells, mitochondrial membrane potentials were lower than their age-matched controls, whereas membrane potentials in HCPD cells were higher than their controls. Hypoxia selection pressure may increase the heterogeneity in both CPC and CPD cells, evidenced by higher kurtosis values in HCPC and HCPC cells compared with their age-matched controls.

Table 7 Descriptive statistics of mitochondrial membrane potentials in eight cell lines at the single cell level

Cell Type	N total	Mean	S.D.	Kurtosis	Skewness
CCPA	9952	1.21	0.20	15.50	-3.12
HCPA	9965	1.16	0.19	15.71	-3.11
CCPB	8973	0.95	0.27	4.78	-2.16
HCPB	8652	0.90	0.27	4.49	-2.20
CCPC	9314	1.32	0.12	36.73	-4.19
HCPC	9853	1.06	0.19	33.21	0.55
CCPD	9859	1.17	0.21	15.68	-3.48
HCPD	8786	1.20	0.50	5849.36	69.04

6.4.2 RNA-Seq of hypoxia-adapted and age-matched control cells

In pre-malignant progression of esophageal adenocarcinoma, deep ulceration generates a periodic hypoxic environment for the esophageal epithelial cells (Suchorolski, Paulson et al. 2013). Esophageal cells may change their genotypes and phenotypes in response to hypoxia selection pressure. However, very few studies have been reported on transcriptome and metabolic changes in Barrett's esophagus cells under hypoxia selection.

To investigate how Barrett's esophagus cells adapt to selective pressure generated via acute repetitive hypoxia insult, hypoxia-adapted and age-matched normoxic control Barrett's esophagus cell lines were generated as described in Section 6.3.1 Cell lines. The cells underwent six rounds of selection and the surviving population was designated as hypoxia-selected cells (HCPA, HCPB, HCPC and HCPD). Their age-matched control cells are named as CCPA, CCPB, CCPC and CCPD. All eight cell lines were treated with acute hypoxia (1% O_2) for 2 hours in three biological replicates. The total RNA from each sample was extracted and subjected to RNA-Seq analysis. Five samples were multiplexed per lane with two lane replicates on an Illumina HiSeq 2000 sequencer. The raw reads were aligned to the *Homo sapiens* genome reference build 37.2 (GRCh37.p2) using Burrows-Wheeler Aligner (Li and Durbin 2009) with Genome Analysis Toolkit (McKenna, Hanna et al. 2010) for variant calling. From approximately 40 million single-end 50-bp sequencing reads, a median of 35 million mapped reads per sample were recovered. Most of the mapped reads fell within annotated gene regions (exon-intron), rRNA or snRNA, and intergenic regions.

6.4.3 Comparisons among hypoxia-adapted and control Barrett's esophagus cell lines

First, gene expressions among all eight hypoxia-adapted and control Barrett's esophagus cell lines (CCPA, CCPB, CCPC, CCPD, HCPA, HCPB, HCPC, and HCPD) were compared. A non-parametric Kruskal-Wallis test was performed on reads per million (RPM) of each gene. For multiple testing corrections, a Benjamini–Hochberg correction (Hochberg and Benjamini 1990) was applied to obtain the false discovery rate (FDR). After the FDR < 0.05 cutoff, it was found that the expressions of 21723 genes

were different among these eight cell lines. Sample cluster analysis (Figure 6-2) showed that the difference between the hypoxia and control pairs is bigger than that within the pairs. Also, the differences between HCPB and CCPB cells, HCPC and CCPC cells are bigger than the differences between hypoxia-adapted and control pairs in CPA and CPD cell lines.

Gene ontology analysis of these genes elucidated that metabolism, cellular signaling, transport and cell cycle related functions were different among these cell lines. Using CCPA cells as a control, most of the enriched metabolism-related functions were activated (Figure 6-3, threshold: z-score = 4). This indicated that metabolic reprogramming may be taking place during the pre-malignant progression in response to hypoxia selection pressure.

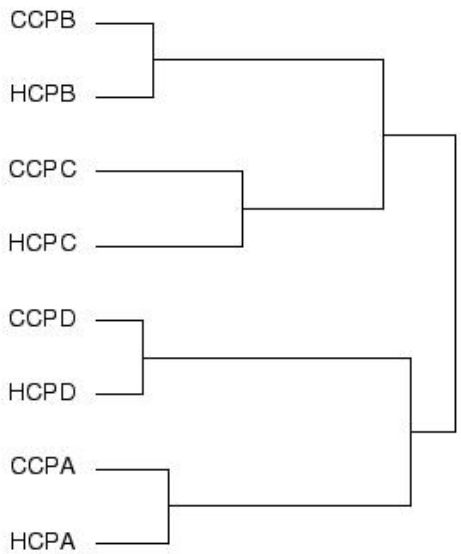


Figure 6-2 Sample cluster of differentially expressed genes
Distance: correlation; linkage: complete

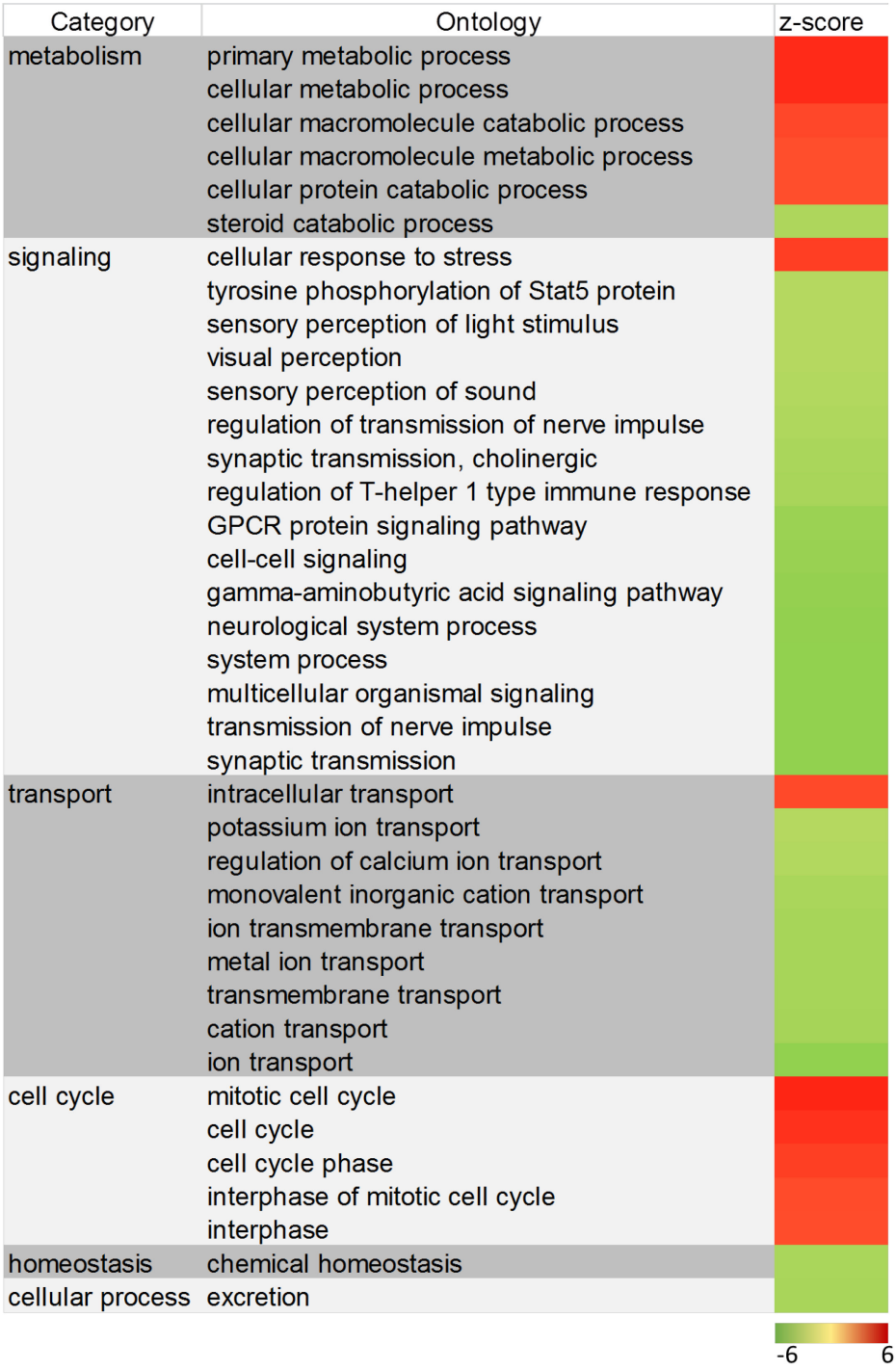


Figure 6-3 Gene ontology analysis of statistically significant genes

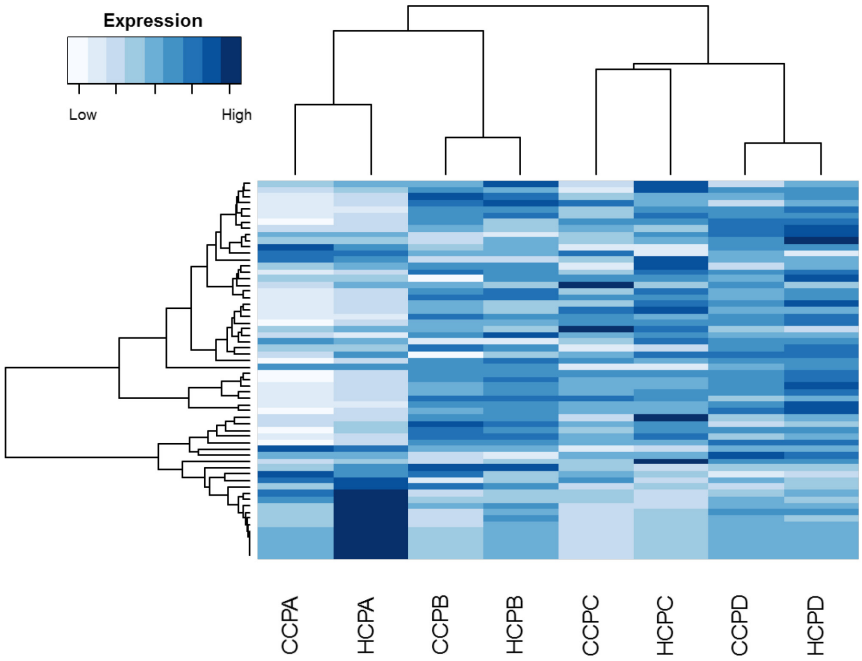


Figure 6-4 Hierarchical clustering of statistically significant genes in the glycolysis/gluconeogenesis pathway
Distance: correlation; linkage: complete

To gain further insights into the alterations in metabolism, these statistically significant genes were sorted based on their functions. Hierarchical clustering of genes and cell lines (Figure 6-4) clearly demonstrated that the expression levels of most genes in the glycolysis/gluconeogenesis pathway were higher in late-stage dysplastic cells (CCPB, HCPB, CCPC, HCPC, CCPD, and HCPD) than in metaplastic cells (CCPA and HCPA). The expressions of these genes in hypoxia-adapted cells were generally higher than age-matched control cells. The difference between hypoxia-adapted and age-matched control CPC cells was the largest.

Intriguingly, the last ten genes in the cluster showed a different pattern from the other genes (Figure 6-5): the expressions in hypoxia-adapted cells were higher than control cells in the pairs of CPA, CPB and CPC cells; while the expressions in hypoxia-

adapted HCPD cells were lower than control CCPD cells. Several alcohol dehydrogenases are among these genes. High expression levels of class I (Triano, Slusher et al. 2003) and II (Crabb, Matsumoto et al. 2004) alcohol dehydrogenase are associated with reduced risk of breast and esophageal cancer. This made for an interesting remark on the role of alcohol dehydrogenases in hypoxia-adaptation in esophageal cells.

Hierarchical clustering of genes involved in oxidative phosphorylation (Figure 6-5) indicated that expression levels of most genes were also higher in late-stage dysplastic cells than in metaplastic cells. In addition, their expressions in hypoxia-adapted cells were generally higher than in age-matched controls in the pairs of CPB, CPC and CPD cells. The last few genes in the cluster also manifested a different pattern, including cytochrome c oxidase subunit IV isoform 2 (*COX4I2*). In contrast, cytochrome c oxidase subunit IV isoform 1 (*COX4I1*) is in the upper part of the cluster. In hypoxic human cells, hypoxia-inducible factor 1 (*HIF-1*) can transactivate the *COX4I2* gene and mediate the degradation of *COX4I1* (Fukuda, Zhang et al. 2007). The COX subunit switching mechanism keeps homeostasis by maintaining optimal efficiency of mitochondrial respiration under hypoxia. The results suggested that as metaplastic (HCPA) and high-grade dysplastic (HCPB and HCPC) cells adapt to hypoxia, *COX4I2* is induced and acts as a physiological response in these cells. However, this was not observed in HCPD cells.

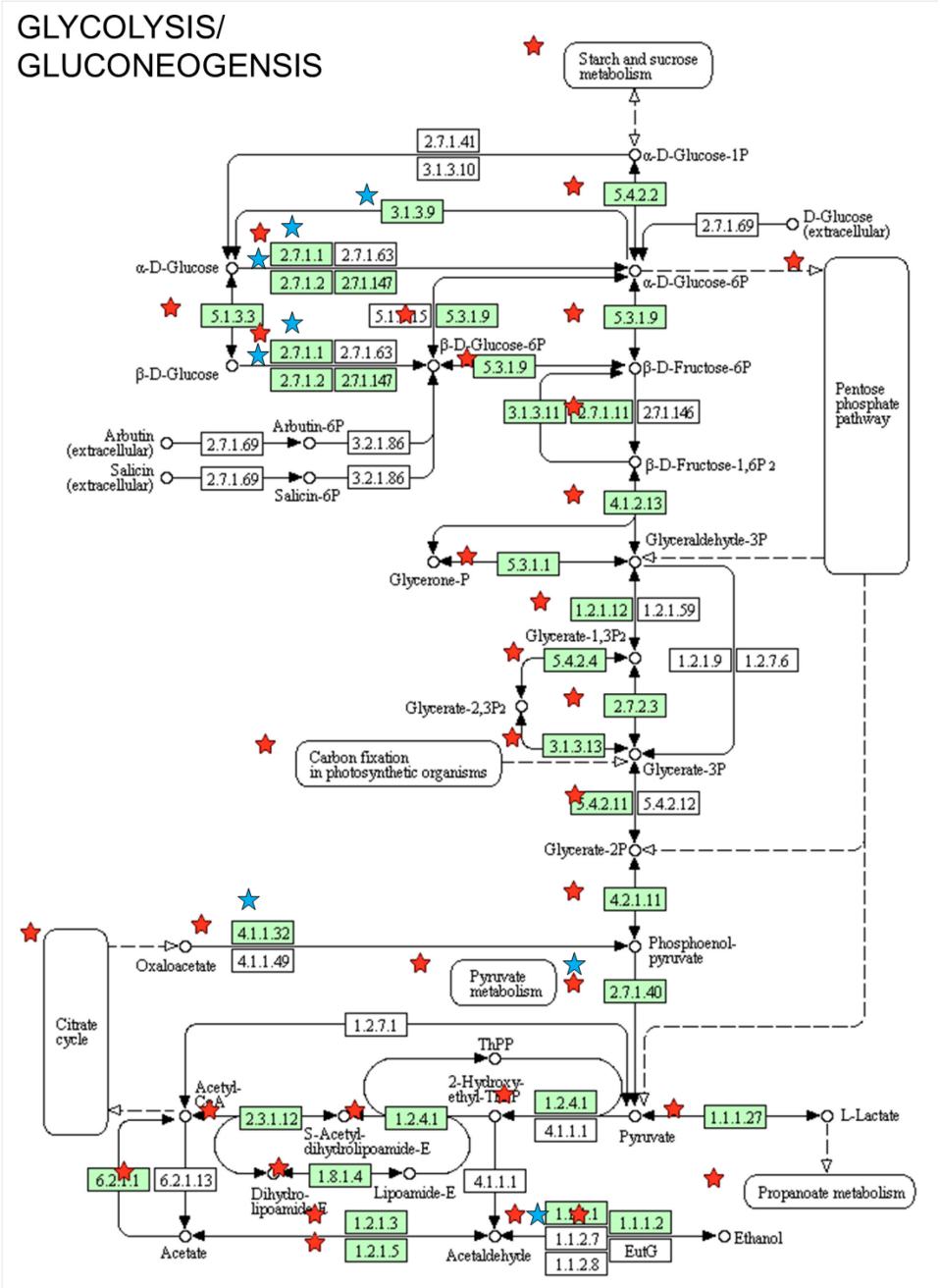


Figure 6-5 Statistically significant genes in the glycolysis/gluconeogenesis pathway

Red star: genes in row 1-51 of the cluster. *Blue star*: genes in row 52-60 of the cluster.

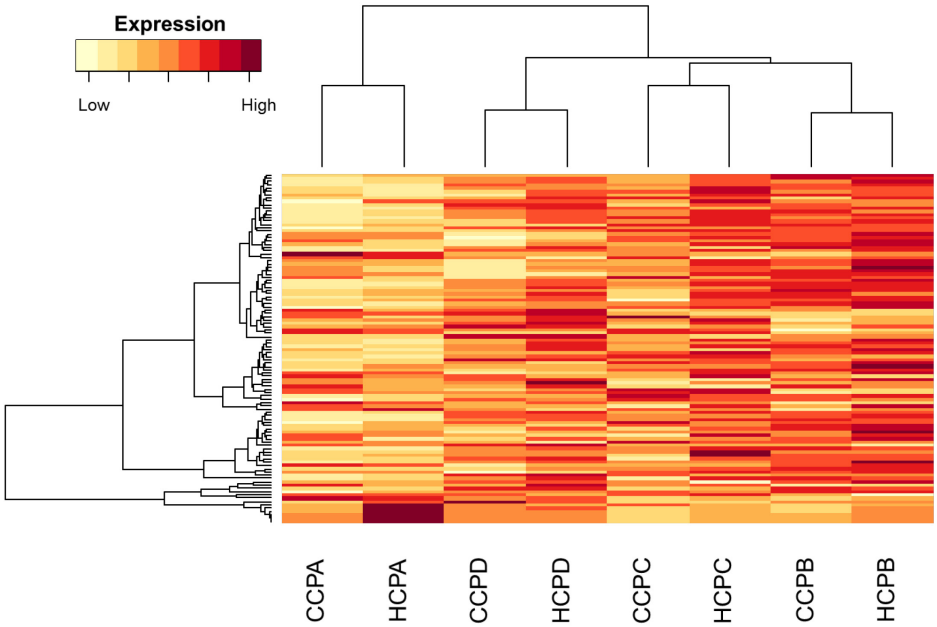


Figure 6-6 Hierarchical clustering of statistically significant genes in the oxidative phosphorylation pathway
Distance: correlation; linkage: complete

6.4.4 Differential gene expression analysis

Three methods were applied to identify differentially expressed genes: DESeq (Anders and Huber 2010), EdgeR (Robinson and Smyth 2008, Robinson, McCarthy et al. 2010) and welch's t-test (Dudoit, Yang et al. 2002). For multiple testing corrections, the Benjamini–Hochberg correction (Hochberg and Benjamini 1990) was performed to obtain the false discovery rate (FDR). The pairwise comparisons of hypoxia-adapted and age-matched control cells included four pairs: (1) HCPA vs. CCPA, (2) HCPB vs. CCPB, (3) HCPC vs. CCPC, and (4) HCPD vs. CCPD. The threshold of the Benjamini-Hochberg correction was set to FDR < 0.05 and a cutoff of logarithmic transformed fold change values ($\log_2$FC) was set at 2. With a FDR threshold at 0.05, the number of

statistically significant genes after $log_2FC = 2$ cutoff was much lower than that without

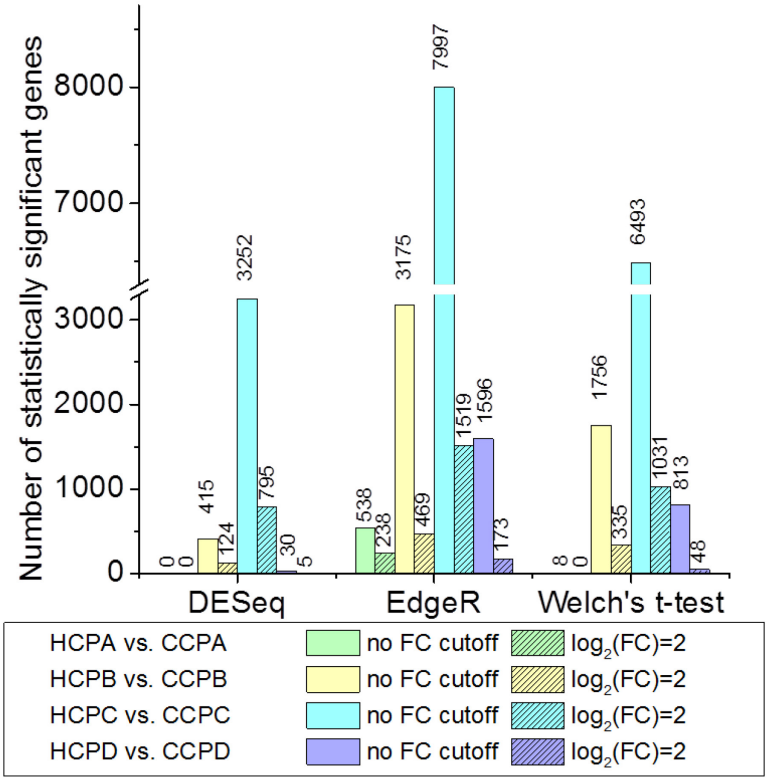


Figure 6-7 Number of statistically significant genes found by three differential gene expression tests

$log_2FC = 2$ cutoff (Figure 6-6).

EdgeR found the largest number of differentially expressed genes, followed by welch's t-test, and DESeq found the smallest number of genes. The number of differentially expressed genes determined by all three methods showed that the HCPC vs. CCPC group has the largest number, followed by the HCPB vs. CCPB and HCPD vs. CCPD groups; the HCPA vs. CCPA group has very few genes identified as differentially expressed.

The differentially expressed genes in three pairwise groups determined by DESeq, EdgeR and welch's t-test with FDR < 0.05 and $\log_2 FC \geq 2$ (Figure 6-7) were also compared. In all the groups except for HCPA vs. CCPA, a large portion of genes with differential expression found by DESeq were also identified by EdgeR. Previous evaluations showed that DESeq was often conservative, while EdgeR was too liberal and

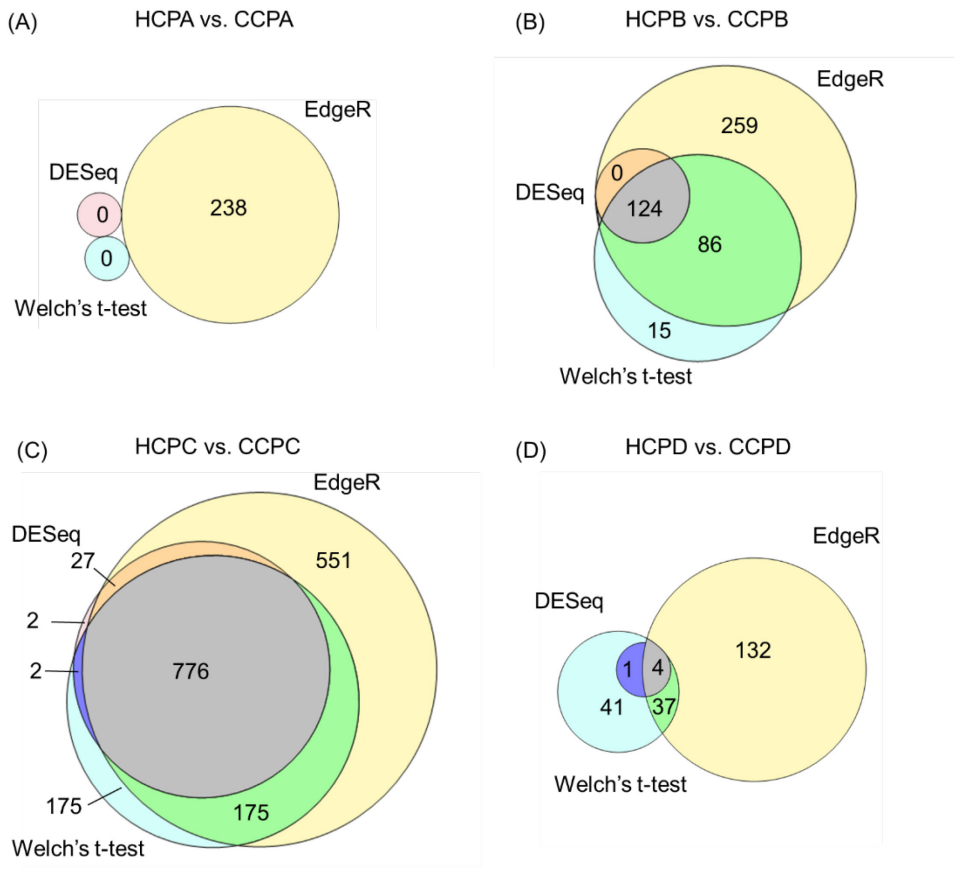


Figure 6-8 Venn diagrams of differentially expressed genes identified by three statistical methods

yields potential false positives (Soneson and Delorenzi 2013). To control false discovery rates, further functional annotations emphasized mainly the genes identified by DESeq.

6.4.5 Function enrichment and upstream analysis of differentially expressed genes between hypoxia-adapted and age-matched control cell lines

Gene expressions are significantly changed by hypoxia-adaptation in the pair of HCPC and CCPC cells based on the number of genes identified by statistical tests. To discover biological relevance of the transcriptome alterations, Ingenuity Pathway analysis (IPA) was performed on the list of differentially expressed genes identified independently by DESeq.

Seventy-two 72 significant bio-function terms enriched were in genes found by DESeq (Fisher's Exact Test and Benjamini-Hochberg correction, FDR < 0.05, activation z-score = 2, Figure 6-8 and Figure 6-9). Interestingly, most of these bio-functions were suppressed in HCPC cells. A majority of them are related to the vascular system: development of the vascular system, neovascularization, remodeling of blood vessel, activation of pericytes, and so forth. Other functional categories, including cellular movement, cell death and survival and cell cycle, are also down-regulated.

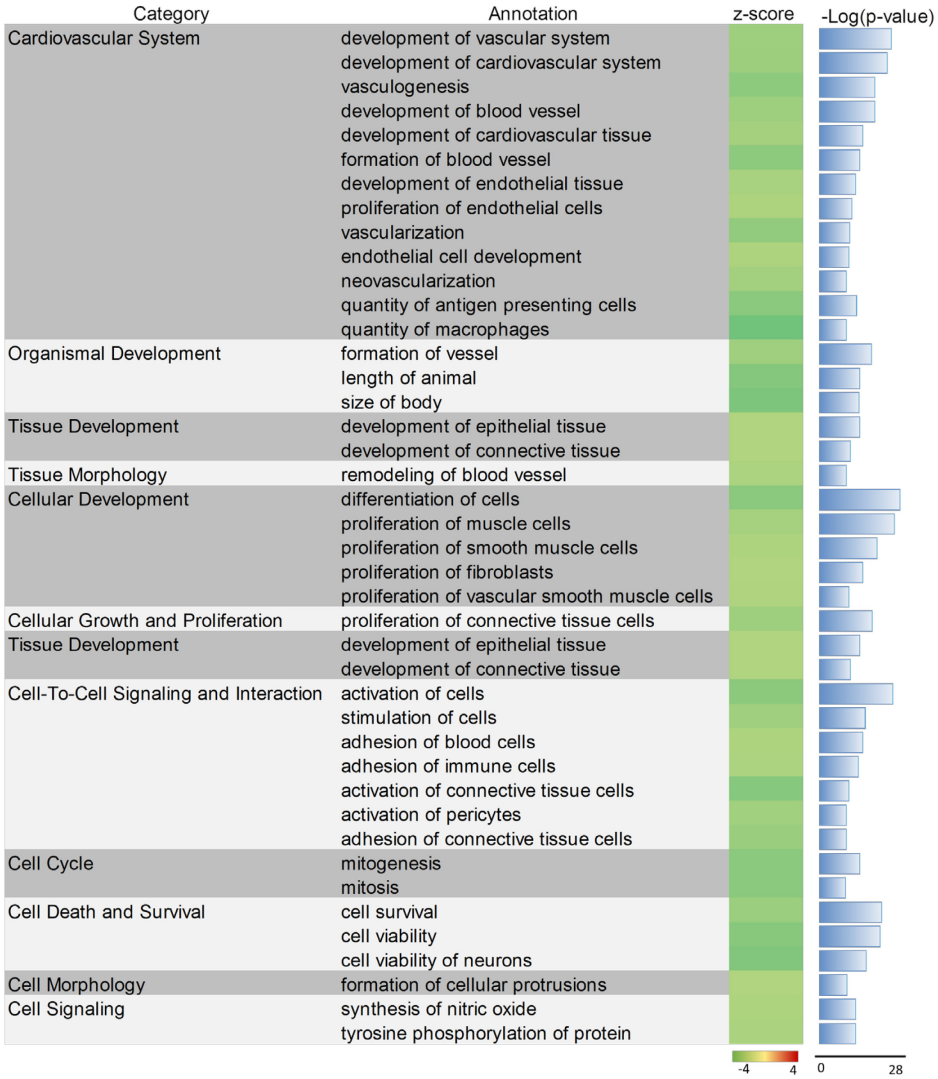


Figure 6-9 Vascular system related functions enriched in differentially expressed genes in the HCPC vs. CCPC group

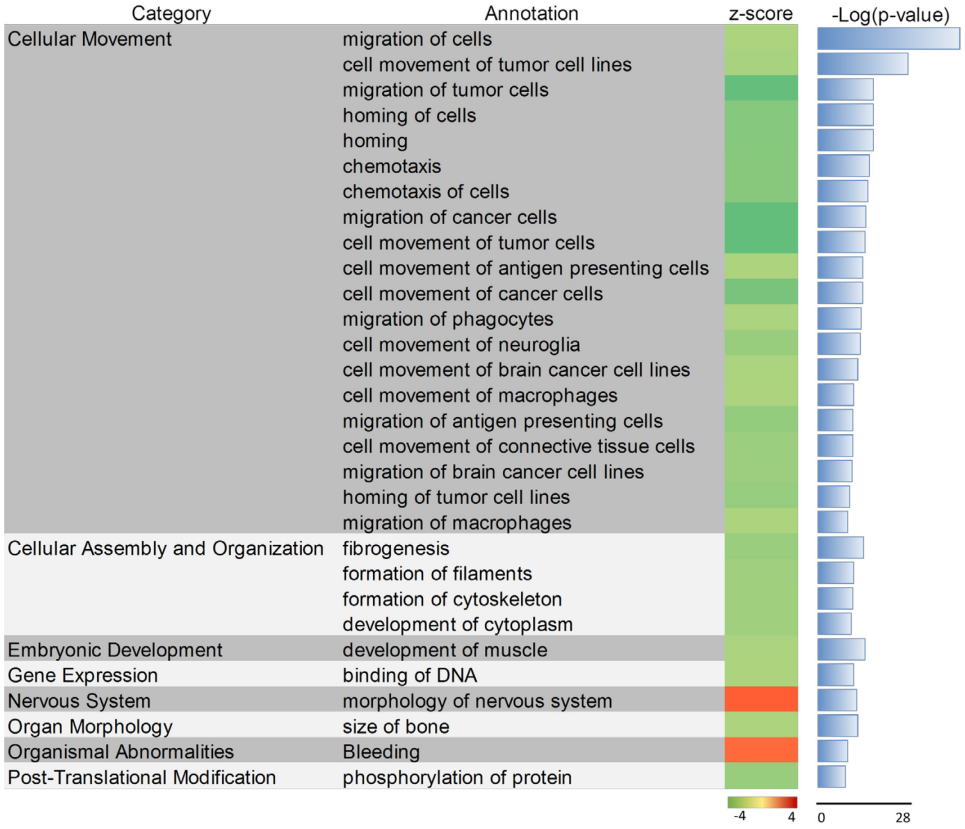


Figure 6-10 Other functions enriched in differentially expressed genes in the HCPC vs. CCPC group

To examine the upstream molecules that trigger the transcriptional changes, Ingenuity Pathway Analysis was used to predict the status of upstream regulators in comparing hypoxia-adapted and age-matched control cells.

Sixty-four upstream regulators were found to be activated or inhibited (threshold: activation z-score = 2, $p < 0.05$, Fisher's Exact Test) with expression level (log ratio) changes. These regulators include growth factors, cytokines, G-protein coupled receptors, transcription regulators, transmembrane receptors, enzymes, kinases, and other signal transducers (Figure 6-10 and Figure 6-11). The regulators were sorted based on their

enrichment p-value within each category and important regulators in adaptation to hypoxia selection pressure were identified.

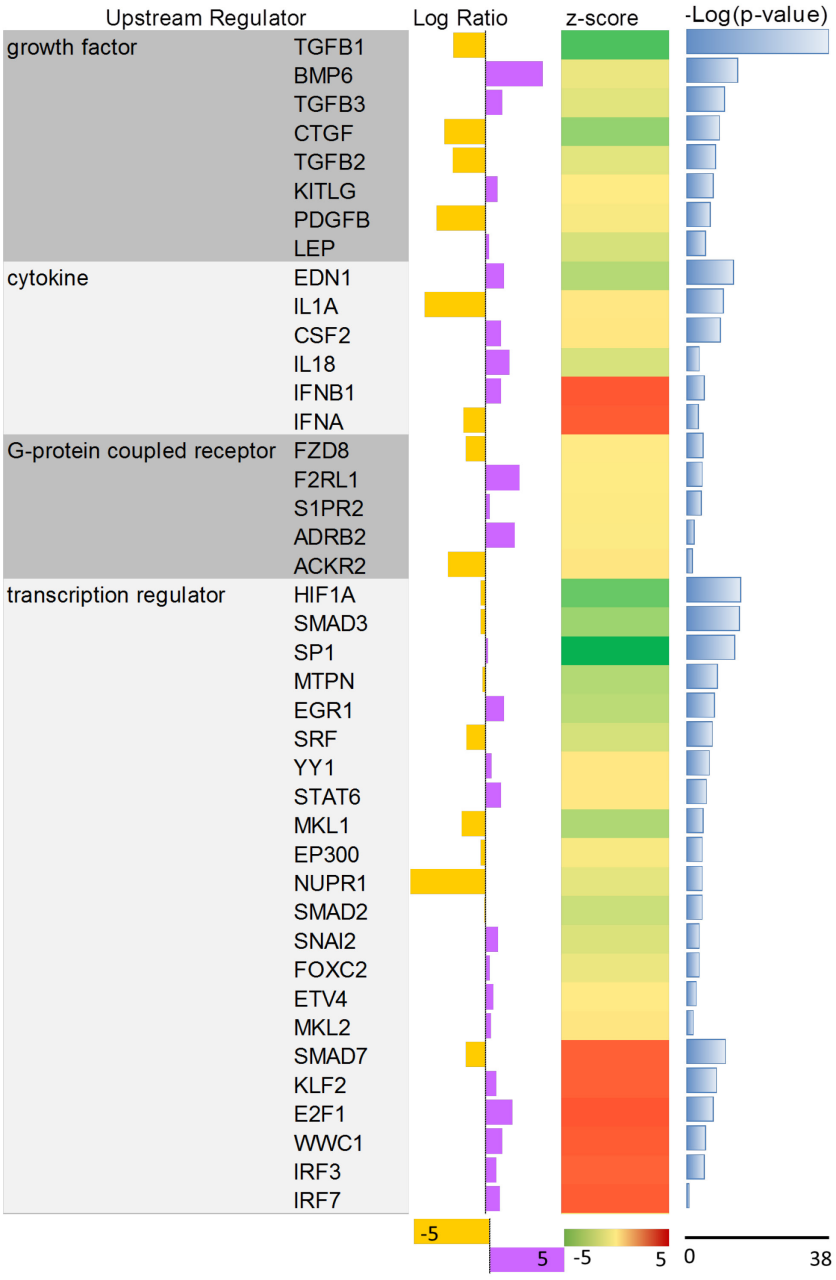


Figure 6-11 Upstream regulator analysis of differentially expressed genes (part 1)

The expression levels of several growth factors and transcription regulators are different between HCPC and CCPC cells: *TGFβ1* (transforming growth factor β), *TGFβ2* (transforming growth factor β2), *SMAD* family, *HIF-1α* and *SP1*. Noticeably, *SMAD2* and *SMAD3* were down-regulated, while *SMAD7* is activated in HCPC cells. Low oxygen

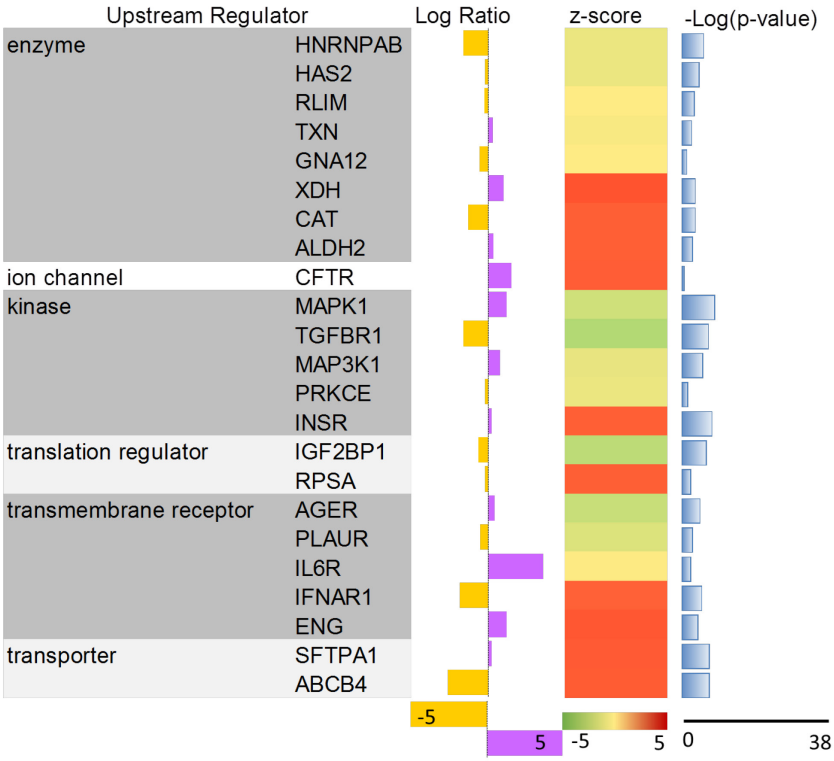


Figure 6-12 Upstream regulator analysis of differentially expressed genes (part 2)

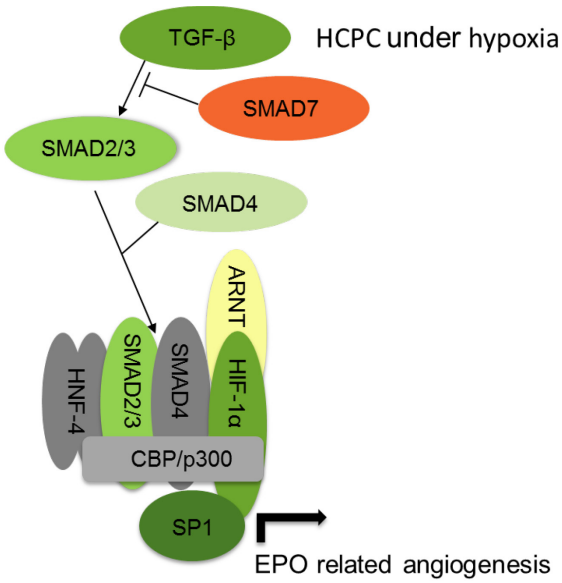


Figure 6-13 TGF-β mediated hypoxia response is suppressed in HCPC cells
Adapted from Sánchez-Elsner et al. 2004. *Green*: inhibition; *red*: activation.

levels can induce a cooperative interaction in which the *TGFβ* pathway interplays with promoter (*SP1*) and enhancer (*HIF-1α*) regions through *SMAD3* and enhances the transcription of genes associated with angiogenesis (Sánchez-Elsner, Ramírez et al. 2004, Basu, Hubchak et al. 2011). Under normoxic conditions, *SMAD7* inhibits the transforming growth factor-beta-activated signaling pathway and prevents carcinoma cell invasion. Hypoxia can convert *SMAD7* to a promoter of cancer invasion and progression (Heikkinen, Nummela et al. 2010). Given that the expressions of *TGFβ*, *SP1*, *HIF-1α*, *SMAD2* and *SMAD3* are inhibited, the crosstalk and functional consequences of these genes are suppressed in HCPC cells. These results suggested that hypoxia-adapted HCPC cells are more resistant to short-term acute hypoxia compared with their age-matched control CCPC cells.

6.5 Conclusion

Whole transcriptome profiles and mitochondrial functions have revealed changes due to hypoxia-selection in pre-malignant Barrett's esophagus. Under normal conditions, mitochondrial membrane potentials are lower in HCPA, HCPB and HCPC cells compared with CCPA, CCPB and CCPC cells. However, HCPD cells have higher mitochondrial membrane potential and a higher heterogeneity than CCPD cells. This is probably because high grade dysplastic cells have higher energy needs for proliferation and other cellular processes. Under hypoxic conditions, transcriptome profiling using RNA-Seq showed that more differentially expressed genes are found between the HCPB vs CCPB group and the HCPC vs CCPC group. The *TGFβ* mediated hypoxia response network is less active in HCPC than CCPC cells. Metabolism and signal transduction related genes are also differentially expressed among all eight cell lines. Taken all these together, hypoxia adaptation makes cells more resistant to acute hypoxia by changing their metabolism and hypoxia-response signaling pathways.

From these results, an evolutionary process can be proposed (Figure 6-14): as the hypoxic selection pressure is applied to these cells, CPA, metaplastic cells, change their phenotypes in adaptation to hypoxia. However, when they are treated by another short-term acute hypoxia, CPA cells did not retain these changes and the adaptation was reversed back. CPB and CPC are the dysplastic cells at transitional states of progression. They made a lot of transcriptional changes and retained accumulative effects of hypoxia adaptation in this process of pre-malignant progression. CPD cells are at the latest stage of esophageal adenocarcinoma. They probably have already gone through several rounds of hypoxia selection. Therefore, hypoxia-adaptation did not change too much on CPD

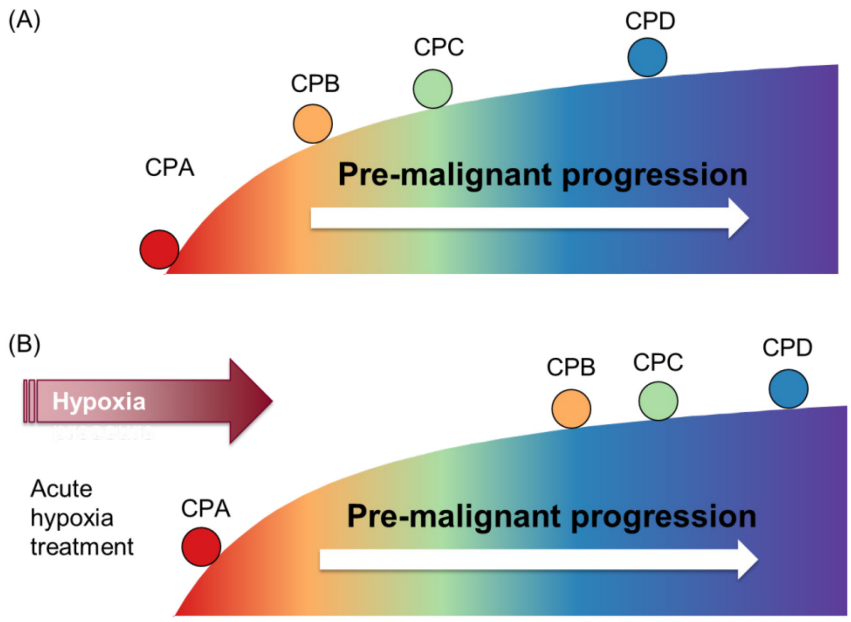


Figure 6-14 The proposed evolutionary process of Barrett's esophagus cells adapted to the hypoxic selection pressure
(A) Before acute hypoxia treatment. (B) After acute hypoxia treatment.

cells. Compared with CPA and CPD cells, dysplastic CPB and CPC cells are at a transitional state and made lots of gene expression changes to adapt to hypoxia. This model suggested that hypoxia adaptation can be used as a risk stratification marker in Barrett's esophagus. Also, because of the functional plasticity in dysplastic cells, physicians need to target multiple metabolic pathways such as oxidative phosphorylation and glycolysis to treat premalignant conditions and cancer.

6.6 Acknowledgments

I thank Tong Fu for help with fluorescence-activated cell sorting, Jason Steel and Kristina Buss from the Center for Personalized Diagnostic for sequencing services, Patti Senechal-Willis and Nanna Hansen for the assistance with cell culture, and Dr. Brian Reid's group from the Fred Hutchinson Cancer Research Center for providing cell lines. This work was supported by the following grant to Dr. Deirdre R. Meldrum from the NIH National Human Genome Research Institute, Centers of Excellence in Genomic Sciences, Grant Number 5P50HG002360.

CHAPTER 7

CONCLUSION AND FUTURE WORK

7.1 Conclusion

The major findings and contributions of my dissertation include:

(i) Developed and optimized a new method for retrieving single adherent cells of different types with minimal perturbation from growth cultures. The method combines mechanical (shear flow) and biochemical (enzymatic digestion) treatments and is relatively simple to use and replicate by other research groups in the field. Analyzed expression levels of stress-related genes in individual cells to a range of different combinations of shear flow and enzymatic digestion. Identified optimal conditions of shear force and trypsinization time for retrieving single cells with minimal perturbation. The study encompasses two important aspects that are broadly applicable in the field of single-cell analysis: (a) a new method for single-cell harvesting of different types of cells from growth cultures and co-cultures with minimal perturbation; (b) a detailed study of the cell response to a range of mechanical and biochemical stress levels in individual cells. This study will be of interest to a broad cell biology research community and especially to the researchers working in the field of single-cell analysis.

(ii) Developed a single cell RT-qPCR method for analyzing expression levels of multiple genes in individual mammalian cells with high sensitivity and reproducibility. It allows for reliable detection of up to ten genes of interest in a single cell with multiple technical replicates. Using this method, hypoxia response genes in 36 single cells exposed to hypoxia and 36 grown under normal physiological conditions were analyzed. This method can also detect cell-to-cell differences of gene expression levels in single

cells under hypoxia treatment. The compatibility of the method with most of commercially available RT-qPCR instrumentation and its relatively low cost makes it amenable to many applications focused on gene expression analysis in single cells, such as high throughput, chip-based techniques, which will enable further insights into cellular mechanisms involved in disease genesis and progression at the single-cell level.

(iii) Characterized the transcriptome wide gene expression changes due to cell-cell interactions in neoplastic progression of Barrett's esophagus. The results indicated that co-culturing of dysplastic cells and normal esophageal epithelial cells significantly changed the expression of hundreds of genes in dysplastic cells. These genes are related to cellular movement, tissue morphology and cancer functions. Identified upstream regulators of differentially expressed genes. Signaling networks regulated by *TGFβ* and *EGF* were significantly inhibited in dysplastic cells by heterotypic interactions. Functional validations showed that proliferation and cellular motility are also changed in co-culture of dysplastic and normal cells. The results suggest that heterotypic interactions in Barrett's esophagus are complicated and dynamic processes involve myriads of transcriptional and phenotypic changes. This study provides deep insights into the role of cell-cell interactions and the tumor microenvironment in the neoplastic progression of Barrett's esophagus. The transcriptional changes due to the presence of cell-cell interactions will shed light on esophageal adenocarcinoma progression, and are likely to reveal the mechanisms of neoplastic progressions in cancer as a general disease.

(iv) Analyzed whole-transcriptome profiles and mitochondrial functions in pre-malignant Barrett's esophagus cell lines adapted to hypoxia. Using RNA-Seq and differential gene expression analysis, observed that gene expression differences between

hypoxia and control cell lines are related to alterations in metabolic processes. The gene expression differences were larger in the HCPB vs. CCPB group and the HCPC vs. CCPC group than in the HCPA vs. CCPA group and the HCPD vs. CCPD group. From upstream analysis, discovered that the interplay between *TGFβ* and hypoxia induced responses were suppressed in hypoxia-adapted high grade dysplastic HCPC cells compared with CCPC cells. Mitochondrial membrane potentials in hypoxia-selected and age-matched control cell lines are significantly different: potentials are lower in hypoxia-selected HCPA, HCPB and HCPC cells than their age-matched controls; potentials are higher in hypoxia-selected HCPD cells than the age-matched control. Results from this study will lead to a greater understanding of mitochondrial functions in neoplasia and tumor progression. It will also help clinicians develop better adjuvant therapeutic strategies targeting cancer metabolism.

With these results, one paper was published, one manuscript was submitted and two more manuscripts are in preparation for publication. All of them are listed below.

1. Zeng, J.*, Wang, J.*, Gao, W., Mohammadreza, A., Kelbauskas, L., Zhang, W., Johnson, R. H., and Meldrum, D. R. (2011). Quantitative single-cell gene expression measurements of multiple genes in response to hypoxia treatment. Analytical and Bioanalytical Chemistry 401, 3-13. (*Co-first authorship) PMID:21614642. *Published in a special, accelerated section, "Paper in Forefront."*

2. Zeng, J., Mohammadreza, A., Gao, W., Merza, S., Smith D., Kelbauskas, L., and Meldrum, D. R. A minimally invasive method for retrieving single adherent cells of different types from cultures. *Under review*.

3. Whole transcriptome and metabolic profiling of intercellular interactions between normal and pre-malignant esophageal cells. *In preparation*.

4. Alterations in gene expression levels and metabolic phenotype in response to hypoxic selection in pre-malignant Barrett's esophagus cells. *In preparation*.

7.2 Future work

In this dissertation research, single cell harvesting and RT-qPCR methods were developed for implementing multi-parameter single-cell analysis and correlating genotype to phenotype in various diseases. The transcriptome-wide differential gene expression analysis of cell-cell interactions and hypoxia responses in Barrett's esophagus cells allows novel inquiries into neoplastic progression.

However, more work will be required in the future to apply these single-cell technologies in an integrated framework to link genotype to phenotype in neoplastic progression. First, the throughput of single-cell harvesting needs to be increased. Currently it takes about 1 min to retrieve a single cell from a microwell. A micro-pipette array can be designed and tested for collecting single cell or single cell lysate, which may enable high-throughput harvesting of single cells from an ultra-high density micro-well array. Second, the single cell RT-qPCR method should be combined with Fluidigm's BioMark platform. Even though the single cell RT-qPCR method developed here can detect up to ten genes without pre-amplification, a larger number of genes is still desired. It will be helpful if an interface or pipeline can be built in the future, bridging the extraction-reverse-transcription step of the RT-qPCR method developed here and Fluidigm BioMark's 96 cells × 96 genes assays. Third, a 3D culture system of intercellular communications in the tumor microenvironment should be created. Cell-cell

interactions in Barrett's esophagus experiments were performed in 2D cell culture. This model is still far away from the native state of the cells. Creating a 3D culture system and further an organotypic model for cell-cell interactions in the tumor microenvironment will allow further inquiries into pre-malignant progression. Fourth, there is a need to track cell-cell interactions for longer periods of time. The process of cell-cell interactions is dynamic. RNA-Seq analysis after 24 hours of co-culturing dysplastic and normal cells only took a snapshot of the ever-changing microenvironment. The landscape of the transcriptome might be drastically different after longer culture times, when both cell lines have adapted to the environment. Therefore, time series experiments on a multi-dimensional co-culture system holds the potential to uncover more events and underlying mechanisms in neoplastic progression of Barrett's esophagus.

REFERENCES

Alison, M., Hunt, T. and Forbes, S. (2002). Minichromosome maintenance (MCM) proteins may be pre-cancer markers. *Gut* **50**(3): 290-291.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215**(3): 403-410.

Anders, S. and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biology* **11**(10): R106.

Anderson, A. R., Weaver, A. M., Cummings, P. T. and Quaranta, V. (2006). Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. *Cell* **127**(5): 905-915.

Anis, Y. H., Holl, M. R. and Meldrum, D. R. (2010). Automated selection and placement of single cells using vision-based feedback control. *IEEE Transactions on Automation Science and Engineering* **7**(3): 598-606.

Anis, Y. H., Houkal, J., Holl, M., Johnson, R. H. and Meldrum, D. R. (2011). Diaphragm pico-liter pump for single-cell manipulation. *Biomedical Microdevices* **13**(4): 651-659.

Arai, F., Ng, C., Maruyama, H., Ichikawa, A., El-Shimy, H. and Fukuda, T. (2005). On chip single-cell separation and immobilization using optical tweezers and thermosensitive hydrogel. *Lab on a Chip* **5**(12): 1399-1403.

Arany, Z., Huang, L. E., Eckner, R., Bhattacharya, S., Jiang, C., Goldberg, M. A., Bunn, H. F. and Livingston, D. M. (1996). An essential role for p300/CBP in the cellular response to hypoxia. *Proceedings of the National Academy of Sciences of the United States of America* **93**(23): 12969-12973.

Ausprunk, D. H. and Folkman, J. (1977). Migration and proliferation of endothelial cells in preformed and newly formed blood vessels during tumor angiogenesis. *Microvascular Research* **14**(1): 53-65.

Bacon, A. and Harris, A. (2004). Hypoxia-inducible factors and hypoxic cell death in tumour physiology. *Annals of Medicine* **36**(7): 530-539.

Bailey, T., Biddlestone, L., Shepherd, N., Barr, H., Warner, P. and Jankowski, J. (1998). Altered cadherin and catenin complexes in the Barrett's esophagus-dysplasia-adenocarcinoma sequence: correlation with disease progression and dedifferentiation. *The American Journal of Pathology* **152**(1): 135.

Balda, M. S. and Matter, K. (2003). Epithelial cell adhesion and the regulation of gene expression. *Trends in Cell Biology* **13**(6): 310-318.

Bani-Hani, K., Martin, I. G., Hardie, L. J., Mapstone, N., Briggs, J. A., Forman, D. and Wild, C. P. (2000). Prospective study of cyclin D1 overexpression in Barrett's esophagus: association with increased risk of adenocarcinoma. *Journal of the National Cancer Institute* **92**(16): 1316-1321.

Barba, M., Czosnek, H. and Hadidi, A. (2014). Historical perspective, development and applications of next-generation sequencing in plant virology. *Viruses* **6**(1): 106-136.

Barcellos-Hoff, M. H., Lyden, D. and Wang, T. C. (2013). The evolution of the cancer niche during multistage carcinogenesis. *Nature Reviews Cancer* **13**(7): 511-518.

Barer, G. R., Howard, P. and Shaw, J. (1970). Stimulus—response curves for the pulmonary vascular bed to hypoxia and hypercapnia. *The Journal of Physiology* **211**(1): 139-155.

Barrett, M. T., Sanchez, C. A., Prevo, L. J., Wong, D. J., Galipeau, P. C., Paulson, T. G., Rabinovitch, P. S. and Reid, B. J. (1999). Evolution of neoplastic cell lineages in Barrett oesophagus. *Nature Genetics* **22**(1): 106-109.

Barrett, M. T., Yeung, K. Y., Ruzzo, W. L., Hsu, L., Blount, P. L., Sullivan, R., Zarbl, H., Delrow, J., Rabinovitch, P. S. and Reid, B. J. (2002). Transcriptional analyses of Barrett's metaplasia and normal upper GI mucosae. *Neoplasia* **4**(2): 121.

Bartfai, T., Buckley, P. T. and Eberwine, J. (2012). Drug targets: single-cell transcriptomics hastens unbiased discovery. *Trends in Pharmacological Sciences* **33**(1): 9-16.

Basu, R. K., Hubchak, S., Hayashida, T., Runyan, C. E., Schumacker, P. T. and Schnaper, H. W. (2011). Interdependence of HIF-1α and TGF-β/Smad3 signaling in normoxic and hypoxic renal epithelial cell collagen expression. *American Journal of Physiology-Renal Physiology* **300**(4): F898.

Bendall, S. C., Simonds, E. F., Qiu, P., El-ad, D. A., Krutzik, P. O., Finck, R., Bruggner, R. V., Melamed, R., Trejo, A. and Ornatsky, O. I. (2011). Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**(6030): 687-696.

Bengtsson, M., Hemberg, M., Rorsman, P. and Stahlberg, A. (2008). Quantification of mRNA in single cells and modelling of RT-qPCR induced noise. *BMC Molecular Biology* **9**: 63-74.

Bertout, J. A., Patel, S. A. and Simon, M. C. (2008). The impact of O2 availability on human cancer. *Nature Reviews Cancer* **8**(12): 967-975.

Bissell, M. J. and Hines, W. C. (2011). Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. *Nature Medicine* **17**(3): 320-329.

Boonstra, J. J., van Marion, R., Beer, D. G., Lin, L., Chaves, P., Ribeiro, C., Pereira, A. D., Roque, L., Darnton, S. J. and Altorki, N. K. (2010). Verification and unmasking of widely used human esophageal adenocarcinoma cell lines. *Journal of the National Cancer Institute* **102**(4): 271-274.

Bristow, R. G. and Hill, R. P. (2008). Hypoxia and metabolism: hypoxia, DNA repair and genetic instability. *Nature Reviews Cancer* **8**(3): 180-192.

Brugarolas, J., Lei, K., Hurley, R. L., Manning, B. D., Reiling, J. H., Hafen, E., Witters, L. A., Ellisen, L. W. and Kaelin, W. G. (2004). Regulation of mTOR function in response to hypoxia by REDD1 and the TSC1/TSC2 tumor suppressor complex. *Genes & Development* **18**(23): 2893-2904.

Bruick, R. K. (2000). Expression of the gene encoding the proapoptotic Nip3 protein is induced by hypoxia. *Proceedings of the National Academy of Sciences of the United States of America* **97**(16): 9082-9087.

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W. and Shipley, G. L. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry* **55**(4): 611-622.

Cameron, A. J., Ott, B. J. and Payne, W. S. (1985). The incidence of adenocarcinoma in columnar-lined (Barrett's) esophagus. *New England Journal of Medicine* **313**(14): 857-859.

Cazes, A., Galaup, A., Chomel, C., Bignon, M., Brechot, N., Le Jan, S., Weber, H., Corvol, P., Muller, L., Germain, S. and Monnot, C. (2006). Extracellular matrix-bound angiopoietin-like 4 inhibits endothelial cell adhesion, migration, and sprouting and alters actin cytoskeleton. *Circulation Research* **99**(11): 1207-1215.

Chaudhry, M. A. (2008). Induction of gene expression alterations by culture medium from trypsinized cells. *Journal of Biological Sciences* **8**(1): 81-87.

Chen, K. D., Li, Y. S., Kim, M., Li, S., Yuan, S., Chien, S. and Shyy, J. Y. (1999). Mechanotransduction in response to shear stress. Roles of receptor tyrosine kinases, integrins, and Shc. *The Journal of Biological Chemistry* **274**(26): 18393-18400.

Chen, Q., Zhang, X. H.-F. and Massagué, J. (2011). Macrophage binding to receptor VCAM-1 transmits survival signals in breast cancer cells that invade the lungs. *Cancer Cell* **20**(4): 538-549.

Chiu, D. T. and Lorenz, R. M. (2009). Chemistry and biology in femtoliter and picoliter volume droplets. *Accounts of Chemical Research* **42**(5): 649-658.

Clemons, N. J., McColl, K. E. and Fitzgerald, R. C. (2007). Nitric oxide and acid induce double-strand DNA breaks in Barrett's esophagus carcinogenesis via distinct mechanisms. *Gastroenterology* **133**(4): 1198-1209.

Cocquet, J., Chong, A., Zhang, G. and Veitia, R. A. (2006). Reverse transcriptase template switching and false alternative transcripts. *Genomics* **88**(1): 127-131.

Cohen, A. A., Geva-Zatorsky, N., Eden, E., Frenkel-Morgenstern, M., Issaeva, I., Sigal, A., Milo, R., Cohen-Saidon, C., Liron, Y., Kam, Z., Cohen, L., Danon, T., Perzov, N. and Alon, U. (2008). Dynamic proteomics of individual cancer cells in response to a drug. *Science* **322**(5907): 1511-1516.

Cossarizza, A., Baccaranicontri, M., Kalashnikova, G. and Franceschi, C. (1993). A new method for the cytofluorometric analysis of mitochondrial membrane potential using the J-aggregate forming lipophilic cation 5, 5′, 6, 6′-tetrachloro-1, 1′, 3, 3′-tetraethylbenzimidazolcarbocyanine iodide (JC-1). *Biochemical and Biophysical Research Communications* **197**(1): 40-45.

Couvelard, A., Paraf, F., Gratio, V., Scoazec, J. Y., Henin, D., Degott, C. and Fléjou, J. F. (2000). Angiogenesis in the neoplastic sequence of Barrett's oesophagus. Correlation with VEGF expression. *The Journal of Pathology* **192**(1): 14-18.

Crabb, D. W., Matsumoto, M., Chang, D. and You, M. (2004). Overview of the role of alcohol dehydrogenase and aldehyde dehydrogenase and their variants in the genesis of alcohol-related pathology. *Proceedings of the Nutrition Society* **63**(01): 49-63.

Dalerba, P., Kalisky, T., Sahoo, D., Rajendran, P. S., Rothenberg, M. E., Leyrat, A. A., Sim, S., Okamoto, J., Johnston, D. M. and Qian, D. (2011). Single-cell dissection of

transcriptional heterogeneity in human colon tumors. *Nature Biotechnology* **29**(12): 1120-1127.

Di Carlo, D., Wu, L. Y. and Lee, L. P. (2006). Dynamic single cell culture array. *Lab on a Chip* **6**(11): 1445-1449.

di Pietro, M., Lao-Sirieix, P., Boyle, S., Cassidy, A., Castillo, D., Saadi, A., Eskeland, R. and Fitzgerald, R. C. (2012). Evidence for a functional role of epigenetically regulated midcluster HOXB genes in the development of Barrett esophagus. *Proceedings of the National Academy of Sciences of the United States of America* **109**(23): 9077-9082.

Diercks, A., Kostner, H. and Ozinsky, A. (2009). Resolving cell population heterogeneity: real-time PCR for simultaneous multiplexed gene detection in multiple single-cell samples. *PLoS One* **4**(7): e6326.

Drewitz, D., Sampliner, R. and Garewal, H. (1997). The incidence of adenocarcinoma in Barrett's esophagus: a prospective study of 170 patients followed 4.8 years. *American Journal of Gastroenterology* **92**(2).

Dubash, A. D., Menold, M. M., Samson, T., Boulter, E., García-Mata, R., Doughman, R. and Burridge, K. (2009). Focal adhesions: new angles on an old structure. *International Review of Cell and Molecular Biology* **277**: 1-65.

Dudoit, S., Yang, Y. H., Callow, M. J. and Speed, T. P. (2002). Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Statistica Sinica* **12**(1): 111-140.

Dulak, A. M., Stojanov, P., Peng, S., Lawrence, M. S., Fox, C., Stewart, C., Bandla, S., Imamura, Y., Schumacher, S. E. and Shefler, E. (2013). Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational complexity. *Nature Genetics* **45**(5): 478-486.

Eberwine, J., Yeh, H., Miyashiro, K., Cao, Y. X., Nair, S., Finnell, R., Zettel, M. and Coleman, P. (1992). Analysis of gene-expression in single live neurons. *Proceedings of the National Academy of Sciences of the United States of America* **89**(7): 3010-3014.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P. and Bettman, B. (2009). Real-time DNA sequencing from single polymerase molecules. *Science* **323**(5910): 133-138.

El-Ali, J., Sorger, P. K. and Jensen, K. F. (2006). Cells on chips. *Nature* **442**(7101): 403-411.

Elowitz, M. B., Levine, A. J., Siggia, E. D. and Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science* **297**(5584): 1183-1186.

Emmert-Buck, M. R., Bonner, R. F., Smith, P. D., Chuaqui, R. F., Zhuang, Z., Goldstein, S. R., Weiss, R. A. and Liotta, L. A. (1996). Laser capture microdissection. *Science* **274**(5289): 998-1001.

Fan, H. C., Wang, J., Potanina, A. and Quake, S. R. (2011). Whole-genome molecular haplotyping of single cells. *Nature Biotechnology* **29**(1): 51-57.

Femino, A. M., Fay, F. S., Fogarty, K. and Singer, R. H. (1998). Visualization of single RNA transcripts in situ. *Science* **280**(5363): 585-590.

Fitzgerald, R., Abdalla, S., Onwuegbusi, B., Sirieix, P., Saeed, I., Burnham, W. and Farthing, M. (2002). Inflammatory gradient in Barrett's oesophagus: implications for disease complications. *Gut* **51**(3): 316-322.

Flaberg, E., Markasz, L., Petranyi, G., Stuber, G., Dicső, F., Alchihabi, N., Oláh, E., Csízy, I., Józsa, T. and Andrén, O. (2011). High‐throughput live‐cell imaging reveals differential inhibition of tumor cell proliferation by human fibroblasts. *International Journal of Cancer* **128**(12): 2793-2802.

Flatz, L., Roychoudhuri, R., Honda, M., Filali-Mouhim, A., Goulet, J.-P., Kettaf, N., Lin, M., Roederer, M., Haddad, E. K. and Sékaly, R. P. (2011). Single-cell gene-expression profiling reveals qualitatively distinct CD8 T cells elicited by different gene-based vaccines. *Proceedings of the National Academy of Sciences of the United States of America* **108**(14): 5724-5729.

Fleige, S. and Pfaffl, M. W. (2006). RNA integrity and the effect on the real-time qRT-PCR performance. *Molecular Aspects of Medicine* **27**(2-3): 126-139.

Flicek, P. and Birney, E. (2009). Sense from sequence reads: methods for alignment and assembly. *Nature Methods* **6**: S6-S12.

Flores, D., Battini, L., Gusella, G. L. and Rohatgi, R. (2010). Fluid shear stress induces renal epithelial gene expression through polycystin-2-dependent trafficking of extracellular regulated kinase. *Nephron Physiology* **117**(4): p27-p36.

Foldager, C. B., Munir, S., Ulrik-Vinther, M., Soballe, K., Bunger, C. and Lind, M. (2009). Validation of suitable house keeping genes for hypoxia-cultured human chondrocytes. *BMC Molecular Biology* **10**: 94-102.

Forman, S. and Jia, X. (2012). Nucleic acid purification, EP Patent 2,479,274.

Fraser, D. and Kaern, M. (2009). A chance at survival: gene expression noise and phenotypic diversification strategies. *Molecular Microbiology* **71**(6): 1333-1340.

Fujita, J. (1999). Cold shock response in mammalian cells. *Journal of Molecular Microbiology and Biotechnology* **1**(2): 243-255.

Fukuda, R., Zhang, H., Kim, J.-w., Shimoda, L., Dang, C. V. and Semenza, G. L. (2007). HIF-1 regulates cytochrome oxidase subunits to optimize efficiency of respiration in hypoxic cells. *Cell* **129**(1): 111-122.

Furusawa, C. and Kaneko, K. (2009). Chaotic expression dynamics implies pluripotency: when theory and experiment meet. *Biology Direct* **4**: 17-28.

Galaup, A., Cazes, A., Le Jan, S., Philippe, J., Connault, E., Le Coz, E., Mekid, H., Mir, L. M., Opolon, P., Corvol, P., Monnot, C. and Germain, S. (2006). Angiopoietin-like 4 prevents metastasis through inhibition of vascular permeability and tumor cell motility and invasiveness. *Proceedings of the National Academy of Sciences of the United States of America* **103**(49): 18721-18726.

Galipeau, P. C., Cowan, D. S., Sanchez, C. A., Barrett, M. T., Emond, M. J., Levine, D. S., Rabinovitch, P. S. and Reid, B. J. (1996). 17p (p53) allelic losses, 4N (G2/tetraploid) populations, and progression to aneuploidy in Barrett's esophagus. *Proceedings of the National Academy of Sciences of the United States of America* **93**(14): 7081-7084.

Galipeau, P. C., Li, X., Blount, P. L., Maley, C. C., Sanchez, C. A., Odze, R. D., Ayub, K., Rabinovitch, P. S., Vaughan, T. L. and Reid, B. J. (2007). NSAIDs modulate CDKN2A, TP53, and DNA content risk for progression to esophageal adenocarcinoma. *PLoS Medicine* **4**(2): e67.

Gao, W., Zhang, W. and Meldrum, D. R. (2011). Quantitative RT-qPCR analysis of gene expression in single bacterial cells. . *Journal of Microbiological Methods* **85**(3): 221-227.

Gatenby, R. A. and Gillies, R. J. (2008). A microenvironmental model of carcinogenesis. *Nature Reviews Cancer* **8**(1): 56-61.

Gentil, C., Le Jan, S., Philippe, J., Leibowitch, J., Sonigo, P., Germain, S. and Pietri-Rouxel, F. (2006). Is oxygen a key factor in the lipodystrophy phenotype? *Lipids in Health and Disease* **5**: 27-38.

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A. J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J. Y. and Zhang, J. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology* **5**(10): R80.

Goetz, J. J. and Trimarchi, J. M. (2012). Transcriptome sequencing of single cells with Smart-Seq. *Nature Biotechnology* **30**(8): 763-765.

Gogvadze, V., Orrenius, S. and Zhivotovsky, B. (2008). Mitochondria in cancer cells: what is so special about them? *Trends in Cell Biology* **18**(4): 165-173.

Gogvadze, V., Zhivotovsky, B. and Orrenius, S. (2010). The Warburg effect and mitochondrial stability in cancer cells. *Molecular Aspects of Medicine* **31**(1): 60-74.

Golding, I., Paulsson, J., Zawilski, S. M. and Cox, E. C. (2005). Real-time kinetics of gene activity in individual bacteria. *Cell* **123**(6): 1025-1036.

Golubeva, Y., Salcedo, R., Mueller, C., Liotta, L. A. and Espina, V. (2013). Laser capture microdissection for protein and NanoString RNA analysis. *Methods in Molecular Biology* **931**: 213-257.

Gong, Y. A., Ogunniyi, A. O. and Love, J. C. (2010). Massively parallel detection of gene expression in single cells using subnanolitre wells. *Lab on a Chip* **10**(18): 2334-2337.

Gonzalez, M., Artimez, M., Rodrigo, L., Lopez-Larrea, C., Menendez, M., Alvarez, V., Perez, R., Fresno, M., Perez, M. and Sampedro, A. (1997). Mutation analysis of the p53, APC, and p16 genes in the Barrett's oesophagus, dysplasia, and adenocarcinoma. *Journal of Clinical Pathology* **50**(3): 212-217.

Greaves, M. and Maley, C. C. (2012). Clonal evolution in cancer. *Nature* **481**(7381): 306-313.

Griffiths, E. A., Pritchard, S., McGrath, S., Valentine, H. R., Price, P. M., Welch, I. and West, C. M. (2007). Increasing expression of hypoxia-inducible proteins in the Barrett's metaplasia–dysplasia–adenocarcinoma sequence. *British Journal of Cancer* **96**(9): 1377-1383.

Gustavsson, M., Mallard, C., Vannucci, S. J., Wilson, M. A., Johnston, M. V. and Hagberg, H. (2007). Vascular response to hypoxic preconditioning in the immature brain. *Journal of Cerebral Blood Flow & Metabolism* **27**(5): 928-938.

Guzhova, I. and Margulis, B. (2006). Hsp70 chaperone as a survival factor in cell pathology. *International Review of Cytology* **254**: 101-149.

Halm, U., Tannapfel, A., Breitung, B., Breidert, M., Wittekind, C. W. and Mössner, J. (1999). Apoptosis and cell proliferation in the metaplasia-dysplasia-carcinoma-sequence of Barrett's esophagus. *Hepato-gastroenterology* **47**(34): 962-966.

Hanahan, D. and Coussens, L. M. (2012). Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* **21**(3): 309-322.

Hanahan, D. and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell* **100**(1): 57-70.

Hanahan, D. and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* **144**(5): 646-674.

Harada, H., Nakagawa, H., Oyama, K., Takaoka, M., Andl, C. D., Jacobmeier, B., von Werder, A., Enders, G. H., Opitz, O. G. and Rustgi, A. K. (2003). Telomerase Induces Immortalization of Human Esophageal Keratinocytes Without p16INK4a inactivation. *Molecular Cancer Research* **1**(10): 729-738.

Harris, A. L. (2002). Hypoxia—a key regulatory factor in tumour growth. *Nature Reviews Cancer* **2**(1): 38-47.

Hartshorn, C., Anshelevich, A. and Wangh, L. J. (2005). Rapid, single-tube method for quantitative preparation and analysis of RNA and DNA in samples as small as one cell. *BMC Biotechnology* **5**: 2-15.

Hartshorn, C., Eckert, J. J., Hartung, O. and Wangh, L. J. (2007). Single-cell duplex RT-LATE-PCR reveals Oct4 and Xist RNA gradients in 8-cell embryos. *BMC Biotechnology* **7**: 87-101.

Hashimshony, T., Wagner, F., Sher, N. and Yanai, I. (2012). CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Reports* **2**(3): 666-673.

Hebenstreit, D. (2012). Methods, challenges and potentials of single cell RNA-seq. *Biology* **1**(3): 658-667.

Heikkinen, P. T., Nummela, M., Jokilehto, T., Grenman, R., Kähäri, V.-M. and Jaakkola, P. M. (2010). Hypoxic conversion of SMAD7 function from an inhibitor into a promoter of cell invasion. *Cancer Research* **70**(14): 5984-5993.

Helm, J., Enkemann, S. A., Coppola, D., Barthel, J. S., Kelley, S. T. and Yeatman, T. J. (2005). Dedifferentiation precedes invasion in the progression from Barrett's metaplasia to esophageal adenocarcinoma. *Clinical Cancer Research* **11**(7): 2478-2485.

Hochberg, Y. and Benjamini, Y. (1990). More powerful procedures for multiple significance testing. *Statistics in Medicine* **9**(7): 811-818.

Hu, X., Bessette, P. H., Qian, J., Meinhart, C. D., Daugherty, P. S. and Soh, H. T. (2005). Marker-specific sorting of rare cells using dielectrophoresis. *Proceedings of the National Academy of Sciences of the United States of America* **102**(44): 15757-15761.

Hu, Z. Y., Fan, C., Livasy, C., He, X. P., Oh, D. S., Ewend, M. G., Carey, L. A., Subramanian, S., West, R., Ikpatt, F., Olopade, O. I., van de Rijn, M. and Perou, C. M. (2009). A compact VEGF signature associated with distant metastases and poor outcomes. *BMC Medicine* **7**: 9-23.

Huang, D. W., Sherman, B. T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M. W. and Lane, H. C. (2007). DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Research* **35**(suppl 2): W169-W175.

Huang, H.-M., Fowler, C., Zhang, H. and Gibson, G. E. (2004). Mitochondrial heterogeneity within and between different cell types. *Neurochemical Research* **29**(3): 651-658.

Hunt, T. and Westervelt, R. (2006). Dielectrophoresis tweezers for single cell manipulation. *Biomedical Microdevices* **8**(3): 227-230.

Hynes, R. O. (2009). The extracellular matrix: not just pretty fibrils. *Science* **326**(5957): 1216-1219.

Imai, T., Horiuchi, A., Wang, C., Oka, K., Ohira, S., Nikaido, T. and Konishi, I. (2003). Hypoxia attenuates the expression of E-cadherin via up-regulation of SNAIL in ovarian carcinoma cells. *The American Journal of Pathology* **163**(4): 1437-1447.

Irish, J. M., Hovland, R., Krutzik, P. O., Perez, O. D., Bruserud, O., Gjertsen, B. T. and Nolan, G. P. (2004). Single cell profiling of potentiated phospho-protein networks in cancer cells. *Cell* **118**(2): 217-228.

Iscove, N. N., Barbara, M., Gu, M., Gibson, M., Modi, C. and Winegarden, N. (2002). Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA. *Nature Biotechnology* **20**(9): 940-943.

Iyer, N. V., Kotch, L. E., Agani, F., Leung, S. W., Laughner, E., Wenger, R. H., Gassmann, M., Gearhart, J. D., Lawler, A. M., Yu, A. Y. and Semenza, G. L. (1998). Cellular and developmental control of O2 homeostasis by hypoxia-inducible factor 1 alpha. *Genes & Development* **12**(2): 149-162.

Jaiswal, K., Morales, C., Feagins, L., Gandia, K., Zhang, X., Zhang, H. Y., Hormi‐Carver, K., Shen, Y., Elder, F. and Ramirez, R. (2007). Characterization of telomerase‐immortalized, non‐neoplastic, human Barrett's cell line (BAR‐T). *Diseases of the Esophagus* **20**(3): 256-264.

Jankowski, J. (1993). Gene expression in Barrett's mucosa: modulation of acute and chronic adaptive respones. *Gut* **34**: 1012-1014.

Jankowski, J. A., Harrison, R. F., Perry, I., Balkwill, F. and Tselepis, C. (2000). Barrett's metaplasia. *The Lancet* **356**(9247): 2079-2085.

Jankowski, J. A., Wright, N. A., Meltzer, S. J., Triadafilopoulos, G., Geboes, K., Casson, A. G., Kerr, D. and Young, L. S. (1999). Molecular evolution of the metaplasia-dysplasia-adenocarcinoma sequence in the esophagus. *The American Journal of Pathology* **154**(4): 965-973.

Jean, C., Gravelle, P., Fournie, J. and Laurent, G. (2011). Influence of stress on extracellular matrix and integrin biology. *Oncogene* **30**(24): 2697-2706.

Jiang, G., Huang, A. H., Cai, Y., Tanase, M. and Sheetz, M. P. (2006). Rigidity sensing at the leading edge through $\alpha v\beta 3$ Integrins and RPTPα. *Biophysical Journal* **90**(5): 1804-1809.

Jin, Z., Cheng, Y., Gu, W., Zheng, Y., Sato, F., Mori, Y., Olaru, A. V., Paun, B. C., Yang, J. and Kan, T. (2009). A multicenter, double-blinded validation study of methylation biomarkers for progression prediction in Barrett's esophagus. *Cancer Research* **69**(10): 4112-4115.

Jing, S., Wang, Y., Chen, L. Q., Sang, M., Zheng, M., Sun, G., Liu, Q., Cheng, Y. and Yang, C. (2013). Hypoxia suppresses E‐cadherin and enhances matrix metalloproteinase‐2 expression favoring esophageal carcinoma migration and invasion via hypoxia inducible factor‐1 alpha activation. *Diseases of the Esophagus* **26**(1): 75-83.

Joglekar, M. V., Wei, C. and Hardikar, A. A. (2010). Quantitative estimation of multiple miRNAs and mRNAs from a single cell. *Cold Spring Harbor Protocols* **2010**(8): pdb prot5478.

Katada, N., Hinder, R. A., Smyrk, T. C., Hirabayashi, N., Perdikis, G., Lund, R. J., Woodward, T. and Klingler, P. J. (1997). Apoptosis is inhibited early in the dysplasia-carcinoma sequence of Barrett esophagus. *Archives of Surgery* **132**(7).

Kelbauskas, L., Ashili, S. P., Houkal, J., Smith, D., Mohammadreza, A., Lee, K. B., Forrester, J., Kumar, A., Anis, Y. H., Paulson, T. G., Youngbull, C. A., Tian, Y. Q., Holl, M. R., Johnson, R. H. and Meldrum, D. R. (2012). Method for physiologic phenotype characterization at the single-cell level in non-interacting and interacting cells. *Journal of Biomedical Optics* **17**(3): 0370081.

Kelly, R. T. and Woolley, A. T. (2005). Microfluidic systems for integrated, high-throughput DNA analysis. *Analytical Chemistry* **77**(5): 96a-102a.

Kim, J.-w., Tchernyshyov, I., Semenza, G. L. and Dang, C. V. (2006). HIF-1-mediated expression of pyruvate dehydrogenase kinase: a metabolic switch required for cellular adaptation to hypoxia. *Cell Metabolism* **3**(3): 177-185.

Kim, R., Weissfeld, J. L., Reynolds, J. C. and Kuller, L. H. (1997). Etiology of Barrett's metaplasia and esophageal adenocarcinoma. *Cancer Epidemiology Biomarkers & Prevention* **6**(5): 369-377.

Kimchi, E. T., Posner, M. C., Park, J. O., Darga, T. E., Kocherginsky, M., Karrison, T., Hart, J., Smith, K. D., Mezhir, J. J. and Weichselbaum, R. R. (2005). Progression of Barrett's metaplasia to adenocarcinoma is associated with the suppression of the transcriptional programs of epidermal differentiation. *Cancer Research* **65**(8): 3146-3154.

Klump, B., Hsieh, C.-J., Holzmann, K., Gregor, M. and Porschen, R. (1998). Hypermethylation of the CDKN2/p16 promoter during neoplastic progression in Barrett's esophagus. *Gastroenterology* **115**(6): 1381-1386.

Kong, J., Crissey, M. A., Funakoshi, S., Kreindler, J. L. and Lynch, J. P. (2011). Ectopic Cdx2 expression in murine esophagus models an intermediate stage in the emergence of Barrett's esophagus. *PloS One* **6**(4): e18280.

Kosoff, R. E., Gardiner, K. L., Merlo, L. M., Pavlov, K., Rustgi, A. K. and Maley, C. C. (2012). Development and characterization of an organotypic model of Barrett's esophagus. *Journal of Cellular Physiology* **227**(6): 2654-2659.

Koss, K., Harrison, R., Gregory, J., Darnton, S., Anderson, M. and Jankowski, J. (2004). The metabolic marker tumour pyruvate kinase type M2 (tumour M2-PK) shows increased expression along the metaplasia–dysplasia–adenocarcinoma sequence in Barrett's esophagus. *Journal of Clinical Pathology* **57**(11): 1156-1159.

Kramer, A., Green, J., Pollard, J., Jr. and Tugendreich, S. (2014). Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30**(4): 523-530.

Krishnan, M. and Erickson, D. (2012). Optically induced microfluidic reconfiguration. *Lab on a Chip* **12**(3): 613-621.

Kubista, M., Andrade, J. M., Bengtsson, M., Forootan, A., Jonák, J., Lind, K., Sindelka, R., Sjöback, R., Sjögreen, B. and Strömbom, L. (2006). The real-time polymerase chain reaction. *Molecular Aspects of Medicine* **27**(2): 95-125.

Kurn, N., Chen, P., Heath, J. D., Kopf-Sill, A., Stephens, K. M. and Wang, S. (2005). Novel isothermal, linear nucleic acid amplification systems for highly multiplexed applications. *Clinical Chemistry* **51**(10): 1973-1981.

Lader, E. S. (2001). Methods and reagents for preserving RNA in cell and tissue samples, US Patent 6,204,375.

Lai, L. A., Paulson, T. G., Li, X., Sanchez, C. A., Maley, C., Odze, R. D., Reid, B. J. and Rabinovitch, P. S. (2007). Increasing genomic instability during premalignant neoplastic progression revealed through high resolution array‐CGH. *Genes, Chromosomes and Cancer* **46**(6): 532-542.

Lal, A., Peters, H., Croix, B. S., Haroon, Z. A., Dewhirst, M. W., Strausberg, R. L., Kaanders, J. H., van der Kogel, A. J. and Riggins, G. J. (2001). Transcriptional response to hypoxia in human tumors. *Journal of the National Cancer Institute* **93**(17): 1337-1343.

Lao-Sirieix, P. and Fitzgerald, R. (2010). Role of the micro-environment in Barrett's carcinogenesis. *Biochemical Society Transactions* **38**(2): 327.

Lao-Sirieix, P., Lovat, L. and Fitzgerald, R. C. (2007). Cyclin A immunocytology as a risk stratification tool for Barrett's esophagus surveillance. *Clinical Cancer Research* **13**(2): 659-665.

Lasken, R. (2013). Single-cell sequencing in its prime. *Nature Biotechnology* **31**(3): 211.

Le, T. T. and Cheng, J. X. (2009). Single-cell profiling reveals the origin of phenotypic variability in adipogenesis. *PLoS One* **4**(4): e5189.

Lee, J. J., Natsuizaka, M., Ohashi, S., Wong, G. S., Takaoka, M., Michaylira, C. Z., Budo, D., Tobias, J. W., Kanai, M., Shirakawa, Y., Naomoto, Y., Klein-Szanto, A. J. P., Haase, V. H. and Nakagawa, H. (2010). Hypoxia activates the cyclooxygenase-2-prostaglandin E synthase axis. *Carcinogenesis* **31**(3): 427-434.

Lee, P. J., Hung, P. J., Shaw, R., Jan, L. and Lee, L. P. (2005). Microfluidic application-specific integrated device for monitoring direct cell-cell communication via gap junctions between individual cell pairs. *Applied Physics Letters* **86**(22): 223902.

Levine, D. M., Ek, W. E., Zhang, R., Liu, X., Onstad, L., Sather, C., Lao-Sirieix, P., Gammon, M. D., Corley, D. A. and Shaheen, N. J. (2013). A genome-wide association study identifies new susceptibility loci for esophageal adenocarcinoma and Barrett's esophagus. *Nature Genetics* **45**(12): 1487-1493.

Levsky, J. M., Shenoy, S. M., Pezo, R. C. and Singer, R. H. (2002). Single-cell gene expression profiling. *Science* **297**(5582): 836-840.

Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**(14): 1754-1760.

Li, H. and Homer, N. (2010). A survey of sequence alignment algorithms for next-generation sequencing. *Briefings in Bioinformatics* **11**(5): 473-483.

Lidstrom, M. E. and Konopka, M. C. (2010). The role of physiological heterogeneity in microbial population behavior. *Nature Chemical Biology* **6**(10): 705-712.

Lidstrom, M. E. and Meldrum, D. R. (2003). Life-on-a-chip. *Nature Reviews Microbiology* **1**(2): 158-164.

Liu, C.-J., Lien, K.-Y., Weng, C.-Y., Shin, J.-W., Chang, T.-Y. and Lee, G.-B. (2009). Magnetic-bead-based microfluidic system for ribonucleic acid extraction and reverse transcription processes. *Biomedical Microdevices* **11**(2): 339-350.

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., Lin, D., Lu, L. and Law, M. (2012). Comparison of next-generation sequencing systems. *BioMed Research International* **2012**.

Liu, Y. X., Cox, S. R., Morita, T. and Kourembanas, S. (1995). Hypoxia Regulates Vascular Endothelial Growth-Factor Gene-Expression in Endothelial-Cells - Identification of a 5'-Enhancer. *Circulation Research* **77**(3): 638-643.

Lopez-Lazaro, M. (2007). Why do tumors metastasize? *Cancer Biology & Therapy* **6**(2): 141-144.

Lord, R. V., Park, J. M., Wickramasinghe, K., DeMeester, S. R., Oberg, S., Salonga, D., Singer, J., Peters, J. H., Danenberg, K. D. and DeMeester, T. R. (2003). Vascular endothelial growth factor and basic fibroblast growth factor expression in esophageal

adenocarcinoma and Barrett esophagus. *The Journal of Thoracic and Cardiovascular Surgery* **125**(2): 246-253.

Lord, R. V., Salonga, D., Danenberg, K. D., Peters, J. H., DeMeester, T. R., Park, J. M., Johansson, J., Skinner, K. A., Chandrasoma, P. and DeMeester, S. R. (2000). Telomerase reverse transcriptase expression is increased early in the Barrett's metaplasia, dysplasia, adenocarcinoma sequence. *Journal of Gastrointestinal Surgery* **4**(2): 135-142.

Losick, R. and Desplan, C. (2008). Stochasticity and cell fate. *Science* **320**(5872): 65-68.

Lu, P., Weaver, V. M. and Werb, Z. (2012). The extracellular matrix: a dynamic niche in cancer progression. *The Journal of Cell Biology* **196**(4): 395-406.

Luthra, R., Wu, T.-T., Luthra, M. G., Izzo, J., Lopez-Alvarez, E., Zhang, L., Bailey, J., Lee, J. H., Bresalier, R. and Rashid, A. (2006). Gene expression profiling of localized esophageal carcinomas: association with pathologic response to preoperative chemoradiation. *Journal of Clinical Oncology* **24**(2): 259-267.

Maley, C. C., Galipeau, P. C., Finley, J. C., Wongsurawat, V. J., Li, X., Sanchez, C. A., Paulson, T. G., Blount, P. L., Risques, R.-A. and Rabinovitch, P. S. (2006). Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nature Genetics* **38**(4): 468-473.

Mammoto, A., Mammoto, T. and Ingber, D. E. (2012). Mechanosensitive mechanisms in transcriptional regulation. *Journal of Cell Science* **125**(13): 3061-3073.

Marcus, J. S., Anderson, W. F. and Quake, S. R. (2006). Microfluidic single-cell mRNA isolation and analysis. *Analytical Chemistry* **78**(9): 3084-3089.

Mardis, E. R. (2011). A decade's perspective on DNA sequencing technology. *Nature* **470**(7333): 198-203.

Mardis, E. R. (2013). Next-generation sequencing platforms. *Annual Review of Analytical Chemistry* **6**: 287-303.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J. and Chen, Z. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**(7057): 376-380.

Martins, R. P., Finan, J. D., Guilak, F. and Lee, D. A. (2012). Mechanical regulation of nuclear structure and function. *Annual Review of Biomedical Engineering* **14**: 431.

Masson, N., Willam, C., Maxwell, P. H., Pugh, C. W. and Ratcliffe, P. J. (2001). Independent function of two destruction domains in hypoxia‐inducible factor‐α chains activated by prolyl hydroxylation. *The EMBO journal* **20**(18): 5197-5206.

Mayer, M. P. and Bukau, B. (2005). Hsp70 chaperones: Cellular functions and molecular mechanism. *Cellular and Molecular Life Sciences* **62**(6): 670-684.

Mazumdar, J., Dondeti, V. and Simon, M. C. (2009). Hypoxia-inducible factors in stem cells and cancer. *Journal of Cellular and Molecular Medicine* **13**(11-12): 4319-4328.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. and DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**(9): 1297-1303.

Merlo, L. M., Kosoff, R. E., Gardiner, K. L. and Maley, C. C. (2011). An in vitro co-culture model of esophageal cells identifies ascorbic acid as a modulator of cell competition. *BMC Cancer* **11**(1): 461.

Merlo, L. M., Pepper, J. W., Reid, B. J. and Maley, C. C. (2006). Cancer as an evolutionary and ecological process. *Nature Reviews Cancer* **6**(12): 924-935.

Metzker, M. L. (2010). Sequencing technologies—the next generation. *Nature Reviews Genetics* **11**(1): 31-46.

Miller, C. T., Moy, J. R., Lin, L., Schipper, M., Normolle, D., Brenner, D. E., Iannettoni, M. D., Orringer, M. B. and Beer, D. G. (2003). Gene amplification in esophageal adenocarcinomas and Barrett's with high-grade dysplasia. *Clinical Cancer Research* **9**(13): 4819-4825.

Miller, J. R., Koren, S. and Sutton, G. (2010). Assembly algorithms for next-generation sequencing data. *Genomics* **95**(6): 315-327.

Morales, C. P., Lee, E. L. and Shay, J. W. (1998). In situ hybridization for the detection of telomerase RNA in the progression from Barrett's esophagus to esophageal adenocarcinoma. *Cancer* **83**(4): 652-659.

Morales, C. P., Souza, R. F. and Spechler, S. J. (2002). Hallmarks of cancer progression in Barrett's esophagus. *The Lancet* **360**(9345): 1587-1589.

Morris, C. D., Armstrong, G. R., Bigley, G., Green, H. and Attwood, S. E. (2001). Cyclooxygenase-2 expression in the Barrett's metaplasia–dysplasia–adenocarcinoma sequence. *The American Journal of Gastroenterology* **96**(4): 990-996.

Mu, J., Brozinick, J. T., Valladares, O., Bucan, M. and Birnbaum, M. J. (2001). A role for AMP-activated protein kinase in contraction- and hypoxia-regulated glucose transport in skeletal muscle. *Molecular Cell* **7**(5): 1085-1094.

Murata, M., Yudo, K., Nakamura, H., Chiha, J., Okamoto, K., Suematsu, N., Nishioka, K., Beppu, M., Inoue, K., Kato, T. and Masuko, K. (2009). Hypoxia Upregulates the Expression of Angiopoietin-Like-4 in Human Articular Chondrocytes: Role of Angiopoietin-Like-4 in the Expression of Matrix Metalloproteinases and Cartilage Degradation. *Journal of Orthopaedic Research* **27**(1): 50-57.

Murphy, M. E. (2013). The HSP70 family and cancer. *Carcinogenesis* **34**(6): 1181-1188.

Murray, L., Sedo, A., Scott, M., McManus, D., Sloan, J. M., Hardie, L. J., Forman, D. and Wild, C. P. (2006). TP53 and progression from Barrett's metaplasia to oesophageal adenocarcinoma in a UK population cohort. *Gut* **55**(10): 1390-1397.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M. and Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**(5881): 1344-1349.

Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D. and Esposito, D. (2011). Tumour evolution inferred by single-cell sequencing. *Nature* **472**(7341): 90-94.

Naya, M., Pereboom, D., Ortego, J., Alda, J. and Lanas, A. (1997). Superoxide anions produced by inflammatory cells play an important part in the pathogenesis of acid and pepsin induced oesophagitis in rabbits. *Gut* **40**(2): 175-181.

Neshat, K., Sanchez, C., Galipeau, P., Blount, P., Levine, D., Joslyn, G. and Reid, B. (1994). p53 mutations in Barrett's adenocarcinoma and high-grade dysplasia. *Gastroenterology* **106**(6): 1589.

Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science* **194**(4260): 23-28.

Oh, H., Takagi, H., Suzuma, K., Otani, A., Matsumura, M. and Honda, Y. (1999). Hypoxia and vascular endothelial growth factor selectively up-regulate angiopoietin-2 in

bovine microvascular endothelial cells. *Journal of Biological Chemistry* **274**(22): 15732-15739.

Okawa, T., Michaylira, C. Z., Kalabis, J., Stairs, D. B., Nakagawa, H., Andl, C., Johnstone, C. N., Klein-Szanto, A. J., El-Deiry, W. S. and Cukierman, E. (2007). The functional interplay between EGFR overexpression, hTERT activation, and p53 mutation in esophageal epithelial cells with activation of stromal fibroblasts induces tumor development, invasion, and differentiation. *Genes & Development* **21**(21): 2788-2803.

Ong, C.-A. J., Lao-Sirieix, P. and Fitzgerald, R. C. (2010). Biomarkers in Barrett's esophagus and esophageal adenocarcinoma: predictors of progression and prognosis. *World Journal of Gastroenterology* **16**(45): 5669.

Onwuegbusi, B. A., Rees, J. R., Lao-Sirieix, P. and Fitzgerald, R. C. (2007). Selective Loss of TGFβ Smad-Dependent Signalling Prevents Cell Cycle Arrest and Promotes Invasion in Oesophageal Adenocarcinoma Cell Lines. *PLoS One* **2**(1): e177.

Overholt, B. F., Lightdale, C. J., Wang, K. K., Canto, M. I., Burdick, S., Haggitt, R. C., Bronner, M. P., Taylor, S. L., Grace, M. G. and Depot, M. (2005). Photodynamic therapy with porfimer sodium for ablation of high-grade dysplasia in Barrett's esophagus: international, partially blinded, randomized phase III trial. *Gastrointestinal Endoscopy* **62**(4): 488-498.

Palanca-Wessels, M. C., Barrett, M. T., Galipeau, P. C., Rohrer, K. L., Reid, B. J. and Rabinovitch, P. S. (1998). Genetic analysis of long-term Barrett's esophagus epithelial cultures exhibiting cytogenetic and ploidy abnormalities. *Gastroenterology* **114**(2): 295-304.

Pan, X., Durrett, R. E., Zhu, H., Tanaka, Y., Li, Y., Zi, X., Marjani, S. L., Euskirchen, G., Ma, C. and LaMotte, R. H. (2013). Two methods for full-length RNA sequencing for low quantities of cells and single cells. *Proceedings of the National Academy of Sciences of the United States of America* **110**(2): 594-599.

Pande, A., Iyer, R., Rani, A., Maddipatla, S., Yang, G., Nwogu, C., Black, J., Levea, C. and Javle, M. (2008). Epidermal growth factor receptor-directed therapy in esophageal cancer. *Oncology* **73**(5-6): 281-289.

Parashar, D., Chauhan, D., Sharma, V. and Katoch, V. (2006). Applications of real-time PCR technology to mycobacterial research. *Indian Journal of Medical Research* **124**(4): 385.

Pareek, C. S., Smoczynski, R. and Tretyn, A. (2011). Sequencing technologies and genome sequencing. *Journal of Applied Genetics* **52**(4): 413-435.

Park, N. J., Yu, T., Nabili, V., Brinkman, B. M., Henry, S., Wang, J. and Wong, D. T. (2006). RNAprotect saliva: An optimal room-temperature stabilization reagent for the salivary transcriptome. *Clinical Chemistry* **52**(12): 2303-2304.

Paulson, T. G., Maley, C. C., Li, X., Li, H., Sanchez, C. A., Chao, D. L., Odze, R. D., Vaughan, T. L., Blount, P. L. and Reid, B. J. (2009). Chromosomal instability and copy number alterations in Barrett's esophagus and esophageal adenocarcinoma. *Clinical Cancer Research* **15**(10): 3305-3314.

Paulson, T. G. and Reid, B. J. (2004). Focus on Barrett's esophagus and esophageal adenocarcinoma. *Cancer Cell* **6**(1): 11-16.

Peitz, I. and van Leeuwen, R. (2010). Single-cell bacteria growth monitoring by automated DEP-facilitated image analysis. *Lab on a Chip* **10**(21): 2944-2951.

Pennacchietti, S., Michieli, P., Galluzzo, M., Mazzone, M., Giordano, S. and Comoglio, P. M. (2003). Hypoxia promotes invasive growth by transcriptional activation of the *met* protooncogene. *Cancer Cell* **3**(4): 347-361.

Pepe, M. S., Etzioni, R., Feng, Z., Potter, J. D., Thompson, M. L., Thornquist, M., Winget, M. and Yasui, Y. (2001). Phases of biomarker development for early detection of cancer. *Journal of the National Cancer Institute* **93**(14): 1054-1061.

Phelan, M. C. (2007). Basic techniques in mammalian cell tissue culture. *Current Protocols in Cell Biology*: 1.1.1-1.1.18.

Picelli, S., Björklund, Å. K., Faridani, O. R., Sagasser, S., Winberg, G. and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nature Methods*.

Pohl, H. A. and Pohl, H. (1978). Dielectrophoresis: the behavior of neutral matter in nonuniform electric fields, Cambridge university press, Cambridge.

Portale, G., Hagen, J. A., Peters, J. H., Chan, L. S., DeMeester, S. R., Gandamihardja, T. A. and DeMeester, T. R. (2006). Modern 5-year survival of resectable esophageal adenocarcinoma: single institution experience with 263 patients. *Journal of the American College of Surgeons* **202**(4): 588-596.

Prinz, C., Tegenfeldt, J. O., Austin, R. H., Cox, E. C. and Sturm, J. C. (2002). Bacterial chromosome extraction and isolation. *Lab on a Chip* **2**(4): 207-212.

Rabinovitch, P., Reid, B., Haggitt, R., Norwood, T. and Rubin, C. (1989). Progression to cancer in Barrett's esophagus is associated with genomic instability. *Laboratory Investigation* **60**(1): 65-71.

Raj, A. and van Oudenaarden, A. (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**(2): 216-226.

Ramsköld, D., Luo, S., Wang, Y.-C., Li, R., Deng, Q., Faridani, O. R., Daniels, G. A., Khrebtukova, I., Loring, J. F. and Laurent, L. C. (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nature Biotechnology* **30**(8): 777-782.

Rasmussen, J. G., Frøbert, O., Pilgaard, L., Kastrup, J., Simonsen, U., Zachar, V. and Fink, T. (2011). Prolonged hypoxic culture and trypsinization increase the pro-angiogenic potential of human adipose tissue-derived stem cells. *Cytotherapy* **13**(3): 318-328.

Rattigan, Y. I., Patel, B. B., Ackerstaff, E., Sukenick, G., Koutcher, J. A., Glod, J. W. and Banerjee, D. (2012). Lactate is a mediator of metabolic cooperation between stromal carcinoma associated fibroblasts and glycolytic tumor cells in the tumor microenvironment. *Experimental Cell Research* **318**(4): 326-335.

Rees, J. R., Onwuegbusi, B. A., Save, V. E., Alderson, D. and Fitzgerald, R. C. (2006). In vivo and in vitro evidence for transforming growth factor-β1-mediated epithelial to mesenchymal transition in esophageal adenocarcinoma. *Cancer Research* **66**(19): 9583-9590.

Reid, B. J., Levine, D. S., Longton, G., Blount, P. L. and Rabinovitch, P. S. (2000). Predictors of progression to cancer in Barrett's esophagus: baseline histology and flow cytometry identify low-and high-risk patient subsets. *The American Journal of Gastroenterology* **95**(7): 1669-1676.

Reid, B. J., Prevo, L. J., Galipeau, P. C., Sanchez, C. A., Longton, G., Levine, D. S., Blount, P. L. and Rabinovitch, P. S. (2001). Predictors of progression in Barrett's esophagus II: baseline 17p (p53) loss of heterozygosity identifies a patient subset at increased risk for neoplastic progression. *The American Journal of Gastroenterology* **96**(10): 2839-2848.

Reis-Filho, J. S. (2009). Next-generation sequencing. *Breast Cancer Research* **11**(Suppl 3): S12.

Riegman, P. H., Vissers, K. J., Alers, J. C., Geelen, E., Hop, W. C., Tilanus, H. W. and van Dekken, H. (2001). Genomic alterations in malignant transformation of Barrett's esophagus. *Cancer Research* **61**(7): 3164-3170.

Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**(1): 139-140.

Robinson, M. D. and Smyth, G. K. (2008). Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics* **9**(2): 321-332.

Roca-Cusachs, P., Iskratsch, T. and Sheetz, M. P. (2012). Finding the weakest link–exploring integrin-mediated mechanical molecular pathways. *Journal of Cell Science* **125**(13): 3025-3038.

Roman, B. L. and Pekkan, K. (2012). Mechanotransduction in embryonic vascular development. *Biomechanics and Modeling in Mechanobiology* **11**(8): 1149-1168.

Rothberg, J. M., Hinz, W., Rearick, T. M., Schultz, J., Mileski, W., Davey, M., Leamon, J. H., Johnson, K., Milgrew, M. J. and Edwards, M. (2011). An integrated semiconductor device enabling non-optical genome sequencing. *Nature* **475**(7356): 348-352.

Russnes, H. G., Navin, N., Hicks, J. and Borresen-Dale, A.-L. (2011). Insight into the heterogeneity of breast cancer through next-generation sequencing. *The Journal of Clinical Investigation* **121**(10): 3810-3818.

Sánchez-Elsner, T., Ramírez, J. R., Rodriguez-Sanz, F., Varela, E., Bernabéu, C. and Botella, L. M. (2004). A cross-talk between hypoxia and TGF-β orchestrates erythropoietin gene regulation through SP1 and Smads. *Journal of Molecular Biology* **336**(1): 9-24.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**(12): 5463-5467.

Schadt, E. E., Turner, S. and Kasarskis, A. (2010). A window into third-generation sequencing. *Human Molecular Genetics* **19**(R2): R227-R240.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. and Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**(7): 676-682.

Schlecht, N. F., Platt, R. W., Duarte-Franco, E., Costa, M. C., Sobrinho, J. P., Prado, J. C., Ferenczy, A., Rohan, T. E., Villa, L. L. and Franco, E. L. (2003). Human papillomavirus infection and time to progression and regression of cervical intraepithelial neoplasia. *Journal of the National Cancer Institute* **95**(17): 1336-1343.

Schmitz, B., Radbruch, A., Kümmel, T., Wickenhauser, C., Korb, H., Hansmann, M., Thiele, J. and Fischer, R. (1994). Magnetic activated cell sorting (MACS)—a new immunomagnetic method for megakaryocytic cell isolation: comparison of different separation techniques. *European Journal of Haematology* **52**(5): 267-275.

Schnell, T. G., Sontag, S. J., Chejfec, G., Aranha, G., Metz, A., O'connell, S., Seidel, U. J. and Sonnenberg, A. (2001). Long-term nonsurgical management of Barrett's esophagus with high-grade dysplasia. *Gastroenterology* **120**(7): 1607-1619.

Schober, M., Raghavan, S., Nikolova, M., Polak, L., Pasolli, H. A., Beggs, H. E., Reichardt, L. F. and Fuchs, E. (2007). Focal adhesion kinase modulates tension signaling to control actin and focal adhesion dynamics. *The Journal of Cell Biology* **176**(5): 667-680.

Schubert, C. (2011). Single-cell analysis: The deepest differences. *Nature* **480**(7375): 133-137.

Schulmann, K., Sterian, A., Berki, A., Yin, J., Sato, F., Xu, Y., Olaru, A., Wang, S., Mori, Y. and Deacu, E. (2005). Inactivation of p16, RUNX3, and HPP1 occurs early in Barrett's-associated neoplastic progression and predicts progression risk. *Oncogene* **24**(25): 4138-4148.

Semenza, G. L. (2002). HIF-1 and tumor progression: pathophysiology and therapeutics. *Trends in Molecular Medicine* **8**(4): S62-S67.

Semenza, G. L. (2003). Targeting HIF-1 for cancer therapy. *Nature Reviews Cancer* **3**(10): 721-732.

Shaheen, N. and Ransohoff, D. F. (2002). Gastroesophageal reflux, Barrett esophagus, and esophageal cancer: scientific review. *The Journal of the American Medical Association* **287**(15): 1972-1981.

Shaheen, N. J., Crosby, M. A., Bozymski, E. M. and Sandler, R. S. (2000). Is there publication bias in the reporting of cancer risk in Barrett's esophagus? *Gastroenterology* **119**(2): 333-338.

Shay, J. and Bacchetti, S. (1997). A survey of telomerase activity in human cancer. *European Journal of Cancer* **33**(5): 787-791.

Shendure, J. and Ji, H. (2008). Next-generation DNA sequencing. *Nature Biotechnology* **26**(10): 1135-1145.

Shi, X., Gao, W., Chao, S.-h., Zhang, W. and Meldrum, D. R. (2013). Monitoring the Single-Cell Stress Response of the Diatom Thalassiosira pseudonana by Quantitative Real-Time Reverse Transcription-PCR. *Applied and Environmental Microbiology* **79**(6): 1850-1858.

Shimizu, D., Vallböhmer, D., Kuramochi, H., Uchida, K., Schneider, S., Chandrasoma, P. T., Shimada, H., DeMeester, T. R., Danenberg, K. D. and Peters, J. H. (2006). Increasing cyclooxygenase‐2 (cox‐2) gene expression in the progression of Barrett's esophagus to adenocarcinoma correlates with that of Bcl‐2. *International Journal of Cancer* **119**(4): 765-770.

Shree, T., Olson, O. C., Elie, B. T., Kester, J. C., Garfall, A. L., Simpson, K., Bell-McGuinn, K. M., Zabor, E. C., Brogi, E. and Joyce, J. A. (2011). Macrophages and cathepsin proteases blunt chemotherapeutic response in breast cancer. *Genes & Development* **25**(23): 2465-2479.

Siegal-Gaskins, D. and Crosson, S. (2008). Tightly regulated and heritable division control in single bacterial cells. *Biophysical Journal* **95**(4): 2063-2072.

Siegel, R., Ma, J., Zou, Z. and Jemal, A. (2014). Cancer statistics, 2014. *CA: A Cancer Journal for Clinicians* **64**(1): 9-29.

Sihvo, E. I., Ruohtula, T., Auvinen, M. I., Koivistoinen, A., Harjula, A. L. and Salo, J. A. (2003). Simultaneous progression of oxidative stress and angiogenesis in malignant transformation of Barrett esophagus. *The Journal of Thoracic and Cardiovascular Surgery* **126**(6): 1952-1957.

Sirieix, P. S., O'Donovan, M., Brown, J., Save, V., Coleman, N. and Fitzgerald, R. C. (2003). Surface expression of minichromosome maintenance proteins provides a novel method for detecting patients at risk for developing adenocarcinoma in Barrett's esophagus. *Clinical Cancer Research* **9**(7): 2560-2566.

Snijder, B. and Pelkmans, L. (2011). Origins of regulated cell-to-cell variability. *Nature Reviews Molecular Cell Biology* **12**(2): 119-125.

Soneson, C. and Delorenzi, M. (2013). A comparison of methods for differential expression analysis of RNA-seq data. *BMC Bioinformatics* **14**(1): 91.

Sotgia, F., Martinez-Outschoorn, U. E., Howell, A., Pestell, R. G., Pavlides, S. and Lisanti, M. P. (2012). Caveolin-1 and cancer metabolism in the tumor microenvironment: markers, models, and mechanisms. *Annual Review of Pathology: Mechanisms of Disease* **7**: 423-467.

Souza, R. F., Shewmake, K., Pearson, S., Sarosi, G. A., Feagins, L. A., Ramirez, R. D., Terada, L. S. and Spechler, S. J. (2004). Acid increases proliferation via ERK and p38 MAPK-mediated increases in cyclooxygenase-2 in Barrett's adenocarcinoma cells. *American Journal of Physiology-Gastrointestinal and Liver Physiology* **287**(4): G743-G748.

Sowter, H. M., Ratcliffe, P. J., Watson, P., Greenberg, A. H. and Harris, A. L. (2001). HIF-1-dependent regulation of hypoxic induction of the cell death factors BNIP3 and NIX in human tumors. *Cancer Research* **61**(18): 6669-6673.

Spechler, S. J. (2013). Barrett esophagus and risk of esophageal cancer: a clinical review. *The Journal of the American Medical Association* **310**(6): 627-636.

Stahlberg, A. and Bengtsson, M. (2010). Single-cell gene expression profiling using reverse transcription quantitative real-time PCR. *Methods* **50**(4): 282-288.

Stairs, D. B., Nakagawa, H., Klein-Szanto, A., Mitchell, S. D., Silberg, D. G., Tobias, J. W., Lynch, J. P. and Rustgi, A. K. (2008). Cdx1 and c-Myc foster the initiation of transdifferentiation of the normal esophageal squamous epithelium toward Barrett's esophagus. *PLoS One* **3**(10): e3534.

Su, Z., Gay, L. J., Strange, A., Palles, C., Band, G., Whiteman, D. C., Lescai, F., Langford, C., Nanji, M. and Edkins, S. (2012). Common variants at the MHC locus and at chromosome 16q24. 1 predispose to Barrett's esophagus. *Nature Genetics* **44**(10): 1131-1136.

Suarez-Quian, C., Goldstein, S., Pohida, T., Smith, P., Peterson, J., Wellner, E., Ghany, M. and Bonner, R. (1999). Laser capture microdissection of single cells from complex tissues. *Biotechniques* **26**(2): 328-335.

Suchorolski, M. T., Paulson, T. G., Sanchez, C. A., Hockenbery, D. and Reid, B. J. (2013). Warburg and Crabtree effects in premalignant Barrett's esophagus cell lines with active mitochondria. *PloS One* **8**(2): e56884.

Suzuki, H., Tomida, A. and Tsuruo, T. (2001). Dephosphorylated hypoxia-inducible factor 1α as a mediator of p53-dependent apoptosis during hypoxia. *Oncogene* **20**(41).

Tang, F., Barbacioru, C., Nordman, E., Li, B., Xu, N., Bashkirov, V. I., Lao, K. and Surani, M. A. (2010). RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nature Protocols* **5**(3): 516-535.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B. and Siddiqui, A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods* **6**(5): 377-382.

Taniguchi, K., Kajiyama, T. and Kambara, H. (2009). Quantitative analysis of gene expression in a single cell by qPCR. *Nature Methods* **6**(7): 503-506.

Taniguchi, Y., Choi, P. J., Li, G.-W., Chen, H., Babu, M., Hearn, J., Emili, A. and Xie, X. S. (2010). Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**(5991): 533-538.

Tay, S., Hughey, J. J., Lee, T. K., Lipniacki, T., Quake, S. R. and Covert, M. W. (2010). Single-cell NF-[kgr] B dynamics reveal digital activation and analogue information processing. *Nature* **466**(7303): 267-271.

Taylor, M. D., Smith, P. W., Brix, W. K., Wick, M. R., Theodosakis, N., Swenson, B. R., Kozower, B. D. and Jones, D. R. (2009). Correlations between selected tumor markers and fluorodeoxyglucose maximal standardized uptake values in esophageal cancer. *European Journal of Cardio-Thoracic Surgery* **35**(4): 699-705.

Tian, Y., Shumway, B. R. and Meldrum, D. R. (2010). A new cross-linkable oxygen sensor covalently bonded into poly (2-hydroxyethyl methacrylate)-co-polyacrylamide thin film for dissolved oxygen sensing. *Chemistry of Materials* **22**(6): 2069-2078.

Tietjen, I., Rihel, J. M., Cao, Y., Koentges, G., Zakhary, L. and Dulac, C. (2003). Single-cell transcriptional analysis of neuronal progenitors. *Neuron* **38**(2): 161-175.

Tischler, J. and Surani, M. A. (2013). Investigating transcriptional states at single-cell-resolution. *Current Opinion in Biotechnology* **24**(1): 69-78.

Triano, E. A., Slusher, L. B., Atkins, T. A., Beneski, J. T., Gestl, S. A., Zolfaghari, R., Polavarapu, R., Frauenhoffer, E. and Weisz, J. (2003). Class I alcohol dehydrogenase is highly expressed in normal human mammary epithelium but not in invasive breast cancer implications for breast carcinogenesis. *Cancer Research* **63**(12): 3092-3100.

Tselepis, C., Perry, I., Dawson, C., Hardy, R., Darnton, S. J., McConkey, C., Stuart, R. C., Wright, N., Harrison, R. and Jankowski, J. (2002). Tumour necrosis factor-a in Barrett's oesophagus: a potential novel mechanism of action. *Oncogene* **21**: 6071-6081.

Tseng, P., Judy, J. W. and Di Carlo, D. (2012). Magnetic nanoparticle-mediated massively parallel mechanical modulation of single-cell behavior. *Nature Methods* **9**(11): 1113-1119.

Ungefroren, H., Sebens, S., Seidl, D., Lehnert, H. and Hass, R. (2011). Interaction of tumor cells with the microenvironment. *Cell Communication and Signaling* **9**(1): 18.

Valencia-Burton, M., Shah, A., Sutin, J., Borogovac, A., McCullough, R. M., Cantor, C. R., Meller, A. and Broude, N. E. (2009). Spatiotemporal patterns and transcription kinetics of induced RNA in single bacterial cells. *Proceedings of the National Academy of Sciences of the United States of America* **106**(38): 16399-16404.

Van Gelder, R. N., von Zastrow, M. E., Yool, A., Dement, W. C., Barchas, J. D. and Eberwine, J. H. (1990). Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proceedings of the National Academy of Sciences of the United States of America* **87**(5): 1663-1667.

Vander Heiden, M. G., Chandel, N. S., Williamson, E. K., Schumacker, P. T. and Thompson, C. B. (1997). Bcl-xL regulates the membrane potential and volume homeostasis of mitochondria. *Cell* **91**(5): 627-637.

Vannucci, S. J., Seaman, L. B. and Vannucci, R. C. (1996). Effects of hypoxia-ischemia on GLUT1 and GLUT3 glucose transporters in immature rat brain. *Journal of Cerebral Blood Flow & Metabolism* **16**(1): 77-81.

Velde, C. V., Cizeau, J., Dubik, D., Alimonti, J., Brown, T., Israels, S., Hakem, R. and Greenberg, A. (2000). BNIP3 and genetic control of necrosis-like cell death through the mitochondrial permeability transition pore. *Molecular and Cellular Biology* **20**(15): 5454-5468.

von Wichert, G., Jiang, G., Kostic, A., De Vos, K., Sap, J. and Sheetz, M. P. (2003). RPTP-α acts as a transducer of mechanical force on αv/β3-integrin–cytoskeleton linkages. *The Journal of Cell Biology* **161**(1): 143-153.

Wang, B., Wood, I. S. and Trayhurn, P. (2008). PCR arrays identify metallothionein-3 as a highly hypoxia-inducible gene in human adipocytes. *Biochemical and Biophysical Research Communications* **368**(1): 88-93.

Wang, D. and Bodovitz, S. (2010). Single cell analysis: the new frontier in 'omics'. *Trends in Biotechnology* **28**(6): 281-290.

Wang, H.-B., Dembo, M., Hanks, S. K. and Wang, Y.-l. (2001). Focal adhesion kinase is involved in mechanosensing during fibroblast migration. *Proceedings of the National Academy of Sciences of the United States of America* **98**(20): 11295-11300.

Wang, J., Qin, R., Ma, Y., Wu, H., Peters, H., Tyska, M., Shaheen, N. J. and Chen, X. (2009). Differential gene expression in normal esophagus and Barrett's esophagus. *Journal of Gastroenterology* **44**(9): 897-911.

Wang, J., Shi, X., Johnson, R. H., Kelbauskas, L., Zhang, W. and Meldrum, D. R. (2013). Single-cell analysis reveals early manifestation of cancerous phenotype in pre-malignant esophageal cells. *PloS One* **8**(10): e75365.

Wang, J. S., Guo, M., Montgomery, E. A., Thompson, R. E., Cosby, H., Hicks, L., Wang, S., Herman, J. G. and Canto, M. I. (2009). DNA promoter hypermethylation of p16 and APC predicts neoplastic progression in Barrett's esophagus. *The American Journal of Gastroenterology* **104**(9): 2153-2160.

Wang, X., Ouyang, H., Yamamoto, Y., Kumar, P. A., Wei, T. S., Dagher, R., Vincent, M., Lu, X., Bellizzi, A. M. and Ho, K. Y. (2011). Residual embryonic cells as precursors of a Barrett's-like metaplasia. *Cell* **145**(7): 1023-1035.

Wang, X. and Seed, B. (2003). A PCR primer bank for quantitative gene expression analysis. *Nucleic Acids Res* **31**(24): e154.

Wang, Z., Gerstein, M. and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **10**(1): 57-63.

Warburg, O., Wind, F. and Negelein, E. (1927). The metabolism of tumors in the body. *The Journal of General Physiology* **8**(6): 519-530.

Wassarman, P. M. and Keller, G. M. (2003). Differentiation of embryonic stem cells, Academic Press.

Weinhouse, S. (1976). The Warburg hypothesis fifty years later. *Zeitschrift Für Krebsforschung Und Klinische Onkologie* **87**(2): 115-126.

Weston, A. P., Banerjee, S. K., Sharma, P., Tran, T. M., Richards, R. and Cherian, R. (2001). p53 protein overexpression in low grade dysplasia (LGD) in Barrett's esophagus:

immunohistochemical marker predictive of progression. *The American Journal of Gastroenterology* **96**(5): 1355-1362.

Weston, A. P., Sharma, P., Topalovski, M., Richards, R., Cherian, R. and Dixon, A. (2000). Long-term follow-up of Barrett's high-grade dysplasia. *The American Journal of Gastroenterology* **95**(8): 1888-1893.

Wheeler, A. R., Throndset, W. R., Whelan, R. J., Leach, A. M., Zare, R. N., Liao, Y. H., Farrell, K., Manger, I. D. and Daridon, A. (2003). Microfluidic device for single-cell analysis. *Analytical Chemistry* **75**(14): 3581-3586.

White, A. K., VanInsberghe, M., Petriv, I., Hamidi, M., Sikorski, D., Marra, M. A., Piret, J., Aparicio, S. and Hansen, C. L. (2011). High-throughput microfluidic single-cell RT-qPCR. *Proceedings of the National Academy of Sciences of the United States of America* **108**(34): 13999-14004.

Wilson, K. T., Fu, S., Ramanujam, K. S. and Meltzer, S. J. (1998). Increased expression of inducible nitric oxide synthase and cyclooxygenase-2 in Barrett's esophagus and associated adenocarcinomas. *Cancer Research* **58**(14): 2929-2934.

Xu, X., Hou, Y., Yin, X., Bao, L., Tang, A., Song, L., Li, F., Tsang, S., Wu, K. and Wu, H. (2012). Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell* **148**(5): 886-895.

Yang, M., Li, C.-W. and Yang, J. (2002). Cell docking and on-chip monitoring of cellular reactions with a controlled concentration gradient on a microfluidic device. *Analytical Chemistry* **74**(16): 3991-4001.

Yoshiji, H., Gomez, D. E., Shibuya, M. and Thorgeirsson, U. P. (1996). Expression of vascular endothelial growth factor, its receptor, and other angiogenic factors in human breast cancer. *Cancer Research* **56**(9): 2013-2016.

Younes, M., Ertan, A., Lechago, L. V., Somoano, J. and Lechago, J. (1997). Human erythrocyte glucose transporter (Glut1) is immunohistochemically detected as a late event during malignant progression in Barrett's metaplasia. *Cancer Epidemiology Biomarkers & Prevention* **6**(5): 303-305.

Younes, M., Lechago, J., Ertan, A., Finnie, D. and Younes, A. (2000). Decreased expression of fas (CD95/AP01) associated with goblet cell metaplasia in barrett's esophagus. *Human Pathology* **31**(4): 434-438.

Younes, M., Schwartz, M. R., Finnie, D. and Younes, A. (1999). Overexpression of Fas ligand (FasL) during malignant transformation in the large bowel and in Barrett's metaplasia of the esophagus. *Human Pathology* **30**(11): 1309-1313.

Yu, C.-h., Law, J. B. K., Suryana, M., Low, H. Y. and Sheetz, M. P. (2011). Early integrin binding to Arg-Gly-Asp peptide activates actin polymerization and contractile movement that stimulates outward translocation. *Proceedings of the National Academy of Sciences of the United States of America* **108**(51): 20585-20590.

Yun, H., Kim, K. and Lee, W. G. (2013). Cell manipulation in microfluidics. *Biofabrication* **5**(2): 022001.

Zare, R. N. and Kim, S. (2010). Microfluidic platforms for single-cell analysis. *Annual Review of Biomedical Engineering* **12**: 187-201.

Zeng, J., Wang, J., Gao, W., Mohammadreza, A., Kelbauskas, L., Zhang, W., Johnson, R. H. and Meldrum, D. R. (2011). Quantitative single-cell gene expression measurements of multiple genes in response to hypoxia treatment. *Analytical and Bioanalytical Chemistry* **401**(1): 3-13.

Zhang, H. and Liu, K.-K. (2008). Optical tweezers for single cells. *Journal of The Royal Society Interface* **5**(24): 671-690.

Zhang, H. Y., Zhang, Q., Zhang, X., Yu, C., Huo, X., Cheng, E., Wang, D. H., Spechler, S. J. and Souza, R. F. (2011). Cancer-related inflammation and Barrett's carcinogenesis: interleukin-6 and STAT3 mediate apoptotic resistance in transformed Barrett's cells. *American Journal of Physiology-Gastrointestinal and Liver Physiology* **300**(3): G454.

Zhang, X., Jones, P. and Haswell, S. J. (2008). Attachment and detachment of living cells on modified microchannel surfaces in a microfluidic-based lab-on-a-chip system. *Chemical Engineering Journal* **135**: S82-S88.

Zhong, H., Agani, F., Baccala, A. A., Laughner, E., Rioseco-Camacho, N., Isaacs, W. B., Simons, J. W. and Semenza, G. L. (1998). Increased expression of hypoxia inducible factor-1 alpha in rat and human prostate cancer. *Cancer Research* **58**(23): 5280-5284.

Zhong, H. and Simons, J. W. (1999). Direct comparison of GAPDH, beta-actin, cyclophilin, and 28S rRNA as internal standards for quantifying RNA levels under hypoxia. *Biochemical and Biophysical Research Communications* **259**(3): 523-526.

Zhu, H., Holl, M., Ray, T., Bhushan, S. and Meldrum, D. R. (2009). Characterization of deep wet etching of fused silica glass for single cell and optical sensor deposition. *Journal of Micromechanics and Microengineering* **19**(6): 065013.

Zong, C., Lu, S., Chapman, A. R. and Xie, X. S. (2012). Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* **338**(6114): 1622-1626.

Zundel, W., Schindler, C., Haas-Kogan, D., Koong, A., Kaper, F., Chen, E., Gottschalk, A. R., Ryan, H. E., Johnson, R. S. and Jefferson, A. B. (2000). Loss of PTEN facilitates HIF-1-mediated gene expression. *Genes & Development* **14**(4): 391-396.

APPENDIX A

PERMISSIONS

This dissertation contains previously published work in which I am listed as the co-first author as well as a figure adapted from a published work

Permissions have been obtained to reprint copyrighted material.

“Quantitative single-cell gene expression measurements of multiple genes in response to hypoxia treatment” has been licensed from Springer under the license number 3352171310300.

“A Cross-Talk Between Hypoxia and TGF-β Orchestrates Erythropoietin Gene Regulation Through SP1 and Smads” has been licensed from Elsevier under the license number 3375051353913.