REDEFINING THE MOLECULAR BASIS OF EPITHELIAL MESENCHYMAL TRANSITION IN BREAST CANCER METASTASIS

A Thesis
Submitted to
the Temple University Graduate Broad

In Partial Fulfillment
of the Requirements for the Degree
MASTER OF SCIENCE

---

by
Yubo Zhai
May 2013

Thesis Approvals:

Jose Russo, MD, FACP, Fox Chase Cancer Center

Charles Grubmeyer, Ph.D, Department of Biochemistry, Temple University

Dianne Soprano, Ph.D, Department of Biochemistry, Temple University

Barbara Hoffman, Ph.D, Department of Biochemistry, Temple University

ABSTRACT

Redefining the Molecular Basis of

Epithelial Mesenchymal Transition in Breast Cancer Metastasis

Yubo Zhai

Chair of the Supervisory Committee:

Jose Russo, MD, FACP

And

Charles Grubmeyer, Ph.D

Metastasis is a multi-step process that begins with cancer cells migrating and invading away from the primary tumor site and extravasating into distant organs to establish a secondary tumor. The loss of epithelial expression markers by neoplastic breast cancer cells in the primary tumor is believed to play a pivotal role during breast cancer metastasis. This phenomenon is the hallmark of the epithelial mesenchymal transition (EMT) process. Gene expression microarrays were performed to investigate key functional elements on an in vitro metastasis model derived from human breast epithelial cells (MCF-10F) treated with 17-beta estradiol. Functional profiling of dysregulated

genes revealed progressive changes in the integrin signaling pathway, and epithelial-mesenchymal transition. In tumorigenic cells, the levels of E-cadherin, desmoplakin and various keratins were low, whereas SLUG, integrinβ1 and fibronectin were high. SLUG, a zinc finger transcription factor acting as a transcriptional repressor, was defined as a promising target which led us establishing a SLUG-centered hypothetical pathway from the profile of dysregulated genes.

ACKNOWLEDGEMENTS

I would like to thank the members of my MS Supervisory committee Jose Russo, MD, FACP, Charles Grubmeyer, Ph.D, Dianne Soprano, Ph.D, and Barbara Hoffman, Ph.D for their support and guidance.

I would also like to thank all of those at the Breast cancer research lab at Fox Chase Cancer Center who helped me learn the laboratory and data analysis techniques that were required to complete this project: Jose Russo, MD, FACP, Irma H. Russo, MD, Julia Pereira, PhD, Ricardo Lopez, PhD, Yanrong Su, MD, Thomas Pogash, Maria Barton, Theresa D. Nguyen, Nathan Hopfinger.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Figure Number Page

LIST OF TABLES

Table Number Page

CHAPTER 1

INTRODUCTION

1.1 Breast Cancer and Metastasis

Metastatic progression of breast cancer is a complex and clinically daunting process lacking fully identified mechanisms [1]. However, it is known that cumulative and sustained exposure to estrogens increases the risk of developing breast cancer [2-4]. Of the five major breast cancer subtypes —basal-like, (ERBB2)-over expressing, normal breast tissue-like, subtype of luminal-like A and B [5-7] —basal-like breast cancer progresses toward malignancy with poor prognosis, as measured in time to development of distal metastasis, since these cells lose polarity and cell-to-cell junctions of epithelial differentiation and they acquire characteristics of mesenchymal cells lacking stable intercellular junctions [8-10].

1.2 *In-vitro* Model of Basal-like Breast Cancer Carcinogenesis

To mimic the process involved in the EMT, we used human breast epithelial cells (HBEC) derived from a previously developed *in vitro* model of carcinogenesis (Figure 1) [11-13]. The model uses an immortalized ERα-negative human breast epithelial cell line, MCF-10F. As previously described, MCF-10F cells were cultured and treated with 70 nmol/L 17β-estradiol (E2). Transformed cells were collected 24 h after the last treatment

and maintained for ten additional passages. These transformed "trMCF" cells, progressively express phenotypes of *in vitro* cell transformation, colony formation in agar methocel, decreased ductulogenesis in collagen assay, and increased invasiveness in a Matrigel invasion system. For the following step, "trMCF" cells were then seeded onto Matrigel invasion chambers, and cells that had degraded the reconstituted basal membrane and invaded were collected and identified as "bsMCF" cells. The tumorigenic ability of "bsMCF" cells was tested by injecting them into the mammary fat pad of 45-day-old female SCID mice. From the tumors formed by the "bsMCF" cells, the fourth cancer cell line "caMCF" was isolated (Figure 1). In addition, bsMCF has been shown to be highly metastatic when injected in the tail of SCID mice [12].

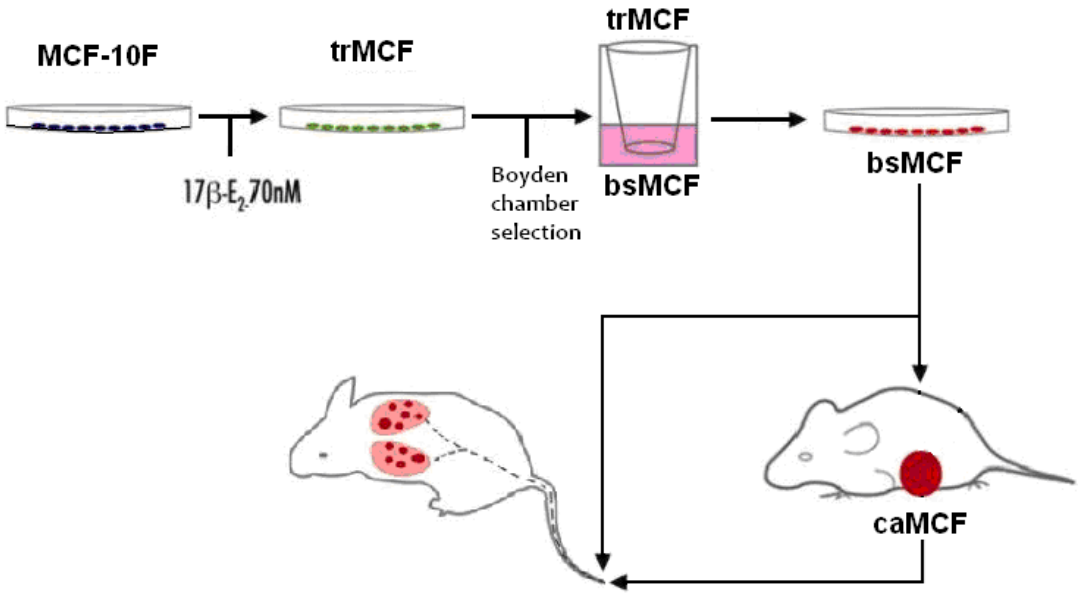


Fig.1 In-vitro Model of Basal-like Breast Cancer Carcinogenesis This figure depicts the different stages of *in vitro* model of basal breast cancer [11, 14].

1.3 Study Objectives

To better understand the molecular events within these estrogen-induced malignant progressive events, especially, the initiation phase of the cell transformation, gene expression analyses were performed using Affymetrix microarrays (HG-U133_Plus_2 arrays) to detect mRNA expression in different stages of *in vitro* model of basal breast cancer [15]. By integrating these data, we were able to identify several dysregulated pathways associated with progressive tumorigenic and invasive capacity.

CHAPTER 2

MATERIALS AND METHODS

2.1 Cell Culture

MCF-10F, trMCF, bsMCF and caMCF cell lines were maintained in growth medium comprising high calcium media (Fox Chase Cancer Center, Philadelphia, PA) supplemented with 5% horse serum. Cells were grown in a 37 ^{o}C incubator with 5% CO_2.

2.2 Affymetrix Gene Expression Microarray Assay

MCF-10F cells at three different passages, designated MCF-10F-1, MCF-10F-2, and MCF-10F-3 (passages 90, 91, and 93, respectively); trMCF cells at three different passages, designated trMCF-1, trMCF-2, and trMCF-3 (passages 55, 56, and 57, respectively); bsMCF 1, 2, and 3 (passages 72, 73, and 74, respectively); and caMCF 1, 2, and 3 (passages 36, 37, and 38, respectively) were used for gene expression microarrays. Total RNA was isolated by using QIAGEN RNeasy Mini Kit (QIAGEN, Inc.). Quality of every sample was assessed using Bioanalyzer 2100 (Agilent Inc.) and samples with good quality (RNA integrity number (RIN)>8) were prepared for hybridization to Affymetrix HG-U133_Plus_2 chips (Affymetrix) following the manufacture's protocol. The chips were scanned using GeneChip Scanner 3000.

2.3 Quality Control

The Simpleaffy package available by Bioconductor was used to examine the quality of the Affymetrix chips [16]. All the chips satisfied the quality control thresholds and were pre-processed using RMA (Robust Multi-array Average) for background adjustment and normalization.

2.4 Analyses of The Differentially Expressed Genes

Limma package [17] available by Bioconductor was used to identify differentially expressed probe sets in the different transformed cells compared to MCF-10F [18, 19]. False discovered rate (FDR) less than 0.01 and fold change (FC) of at least 2 were used as criteria to consider genes differentially expressed.

2.5 Data Mining

Genespring GX software (Agilent Inc.) was used to perform Gene Ontology (GO) enrichment analysis of the differentially expressed genes. GO terms with p-value smaller than 0.01 were considered over-represented among the genes dysregulated in the different groups. To identify pathways, networks and canonical pathways over-represented in the genes modulated during the transformation, we used Ingenuity Pathway Analysis software 5.5 (Ingenuity Systems, Redwood City, CA).

2.6 Real-Time RT-PCR

Real-time RT-PCR was performed to evaluate the mRNA expression levels of dysregulated genes identified in the *in vitro* model of basal breast cancer. 100 ng of RNA were used for each RT-PCR reaction using TaqMan® One-Step RT-PCR Master Mix (Applied Biosystems, Foster City, CA) in an ABI Prism 7000 Sequence Detection System (Applied Biosystems, Foster City, CA). For internal control, the housekeeping gene 18s was selected.

2.7 Immunoblot Analysis

The proteins in corresponding cells were collected, and the concentration of protein in supernatants was estimated using BCA assay (Pierce Chemical, Rockford, IL). Subsequently, 30 µg of protein was separated on NuPAGE® Bis-Tris Mini Gels and electroblotted on polyvinylidene fluoride membranes (Roche Applied Science, Germany). The blots were blocked overnight with Odyssey® Blocking Buffer (LI-COR Biosciences, Lincoln, NE) and incubated with primary antibodies at dilutions recommended by the suppliers. Immunoblots were detected by secondary antibodies using an Odyssey® Infrared Imaging System (LI-COR Biosciences, Lincoln, NE) and photographed.

2.8 Immunofluorescence Analysis

The localizations of target proteins were detected by immunofluorescence analysis. In brief, cells were grown on glass coverslips for three days. Then, cells were washed with PBS, fixed in 4% Paraformaldehyde in PBS at room temperature, and blocked in 5% non-immune goat serum for 1 h at room temperature. Next, the cells were incubated with indicated primary antibody with dilution of 1: 400 overnight at 4°C followed by incubation with Alexa Fluor 488 goat anti-mouse IgG or Alexa Fluor 555 goat anti-rabbit IgG (Cell Signaling, Danvers, MA) for 1 h at room temperature. The nuclei were stained with DAPI and the coverslips were mounted on a microscope slide with embedding medium. The cells were observed and photographed with a fluorescence microscope (Olympus, Japan).

CHAPTER 3

RESULTS

3.1 Gene Expression Altered During The EMT of *In-vitro* Model of Basal Breast Cancer

Table 1 The Number of Differentially Expressed Genes

	trMCF vs. MCF-10F	**bsMCF vs. MCF10-F**	**caMCF vs. MCF-10F**	**Common**
UP	1326	775	800	297
DOWN	**1265**	**953**	**1306**	**471**
Total	2579	1728	2106	

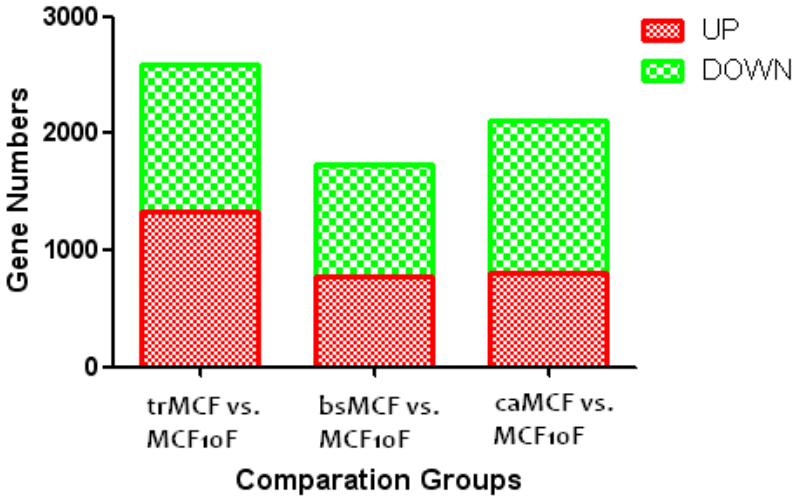


Fig.2 The Numbers of Differentially Expressed Genes within Comparisons This figure indicates the numbers of differentially expressed genes within comparisons trMCF vs MCF-10F, bsMCF vs MCF-10F and caMCF vs MCF-10F

The quality control results demonstrate that all chips showed a high level quality as measured by signal intensity across the chips, hybridization quality and reproducibility. Limma package compared the gene expression levels in trMCF, bsMCF and caMCF with genes in MCF-10F. The number of differentially expressed genes found in each comparison is shown in Table 1 and Fig. 2.

3.2 Biological Processes Enriched with The Dysregulated Genes

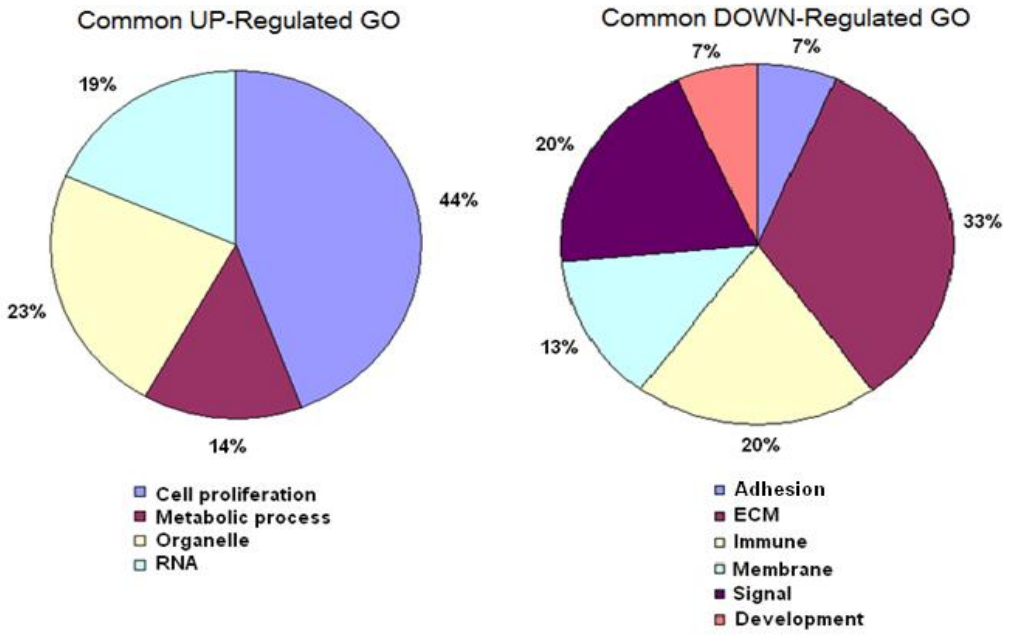


Fig.3 GO Analysis Enriched with Common Genes This figure indicates the GO analysis enriched with common genes in all three transformed cells

The differentially expressed genes were inputted into GeneSpringGX to perform Gene Ontology analysis. Enriched GO terms on genes found down-regulated in trMCF, bsMCF and caMCF were related to extra cellular matrix (ECM) and adhesion. For the up-regulated genes, cell proliferation takes the dominant position in GO terms (Fig. 3). Findings suggest that ECM transformation phenotype was involved in the metastasis initiation processes.

3.3 Identification of Significant Canonical Pathways

Using the Ingenuity Pathway Analysis, we compared the first 30 significant pathways with the highest –log(p-value) from trMCF vs. MCF-10F, bsMCF vs. MCF-10F and caMCF vs. MCF-10F groups. Ten common canonical pathways were revealed as Fig. 4 indicates. We found that the significantly changed elements in integrin pathway interact with several other pathways. This phenomenon indicates that the continuum changed integrin signaling can trigger the alterations in expression pattern throughout the progressive, malignant cell transformation of MCF-10F cells.

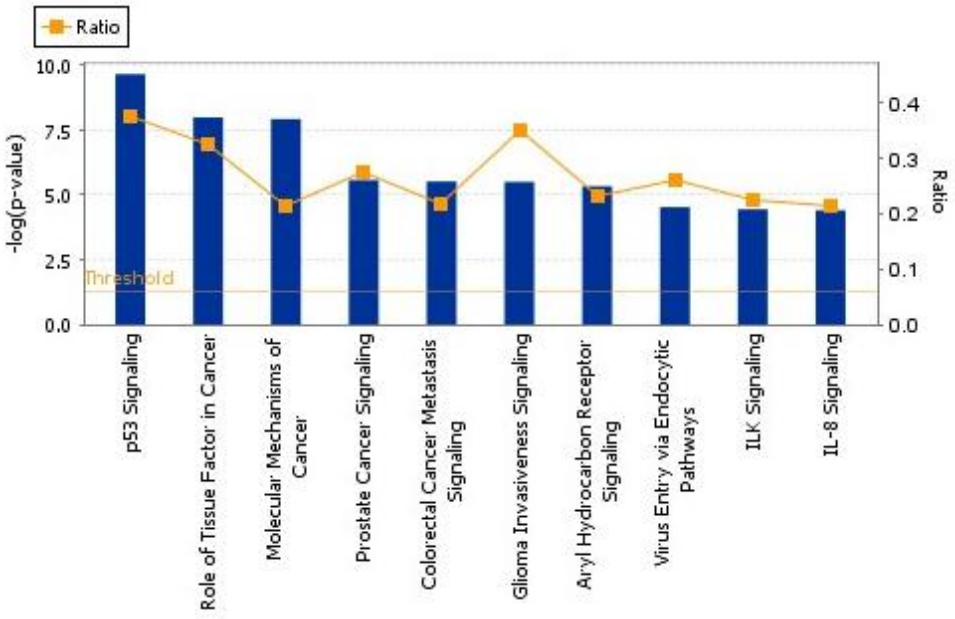


Fig.4 Canonical Pathways Altered in Progressive Malignant Cell Transformation

This figure indicates the canonical pathways altered in progressive malignant cell transformation. Using Ingenuity Pathway Analysis, the canonical pathways significantly enriched by the dysregulated genes were identified. The ratio between the number of differentially expressed genes found in our study and the number of genes that belong to each pathway is labeled above with yellow.

3.4 Composition of A SLUG-centered Hypothetical Pathway

We found 30 genes differentially expressed in our in vitro model that are involved in EMT (Table 2). Analyzing these genes, we have developed a SLUG-centered hypothetical pathway summarized in Fig. 5.

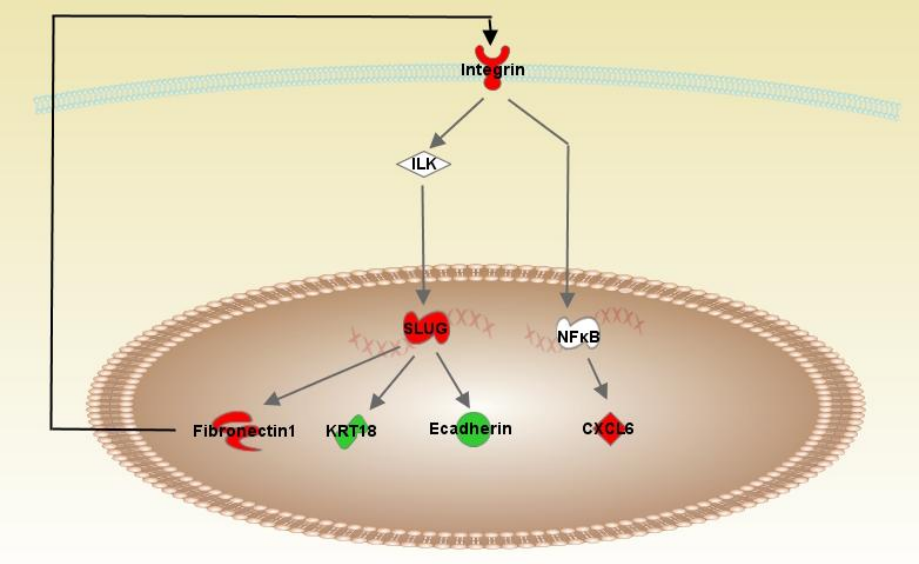


Fig.5 Hypothetical Pathway This figure indicates the hypothetical pathway for progressive molecular events in the 17β-estrodial mediated malignant transformation

In humans, the SNAI2 gene encodes SLUG, which is a member of the Snail family of C2H2-type zinc finger transcription factors. SLUG acts as a transcriptional repressor that binds to E-box motifs and is also likely to repress E-cadherin transcription in breast carcinoma [20-22]. Considering that MCF-10F is devoid of ERα and only expresses a very small amount of ERβ, 17β-estradiol could directly cause epigenetic alterations on SNAI2 gene, such as methylation and etc, without the mediation of ERα. As a zinc finger transcriptional repressor [23, 24], the down-regulation of SLUG tends to disrupt the integrin signaling pathway and therefore alter the ECM-cell relationship by enhancing transcription of Keratin 18 and E-cadherin. This disrupted cross-talk between the epithelium and stroma could in turn initiate the EMT [25-30]. Due to less integrin signal

input derived from altered ECM-cell attachment, the expression of SLUG is kept in a suppressed status through the regulation of integrin-linked kinase (ILK) [29]. Thus, the whole network forms a closed system under the control of SLUG expression. As the downstream of integrin pathways, NFκB is suggested to be down-regulated. As the response genes following NFκB signaling, components related to IL8, the CXC chemokine family and four TNF members all demonstrate suppressed status in trMCF [31-34].

Table 2 Dysregulated Genes Involved in Integrin Signaling Pathway

Symbol	Gene name	Log Fold Change		
		trMCF	**bsMCF**	**caMCF**
SNAI2(SLUG)	snail homolog 2 (Drosophila)	-2.61	-1.81	-2.15
FN1	fibronectin 1	-9.28	NS	-2.25
DSP	desmoplakin	2.22	NS	NS
KRT15	keratin 15	8.18	NS	-1.14
KRT6B	keratin 6B	6.79	-1.2	-1.36
KRT16	keratin 16	3.75	-1.04	-4.51
KRT6A	keratin 6A	3.38	NS	NS
VIM	vimentin	-3.11	-10.21	-10.12
CXCL5	chemokine (C-X-C motif) ligand 5	-3.3	-4.93	-5.27
CXCR7	chemokine (C-X-C motif) receptor 7	-5.41	-7.26	-7.08
CXCL6	chemokine (C-X-C motif) ligand 6 (granulocyte chemotactic protein 2)	-5.48	-2.45	-3.14
CXCL2	chemokine (C-X-C motif) ligand 2	-2.19	-2.18	-2.54
CXCL1	chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	-2.59	-2.78	-3.91

Table 2 (continued)

CXCL3	chemokine (C-X-C motif) ligand 3	-2.92	-2.45	-2.07
IL8	interleukin 8	-4.33	-1.59	-2.89
NFKBIZ	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta	-1.44	-1.09	-1.11
TNFAIP3	tumor necrosis factor, alpha-induced protein 3	-1.47	-1.24	-1.24
TNFRSF21	tumor necrosis factor receptor superfamily, member 21	-1.52	-1.88	-1.81
TNFRSF10D	tumor necrosis factor receptor superfamily, member 10d, decoy with truncated death domain	-1.74	-2.16	-3.93
TNFAIP6	tumor necrosis factor, alpha-induced protein 6	-2.92	-2.47	-2.61
CADM3	cell adhesion molecule 3	-2.46	NS	NS
CDH1	cadherin 1, type 1, E-cadherin (epithelial)	7.55	1.06	NS
CDH3	cadherin 3, type 1, P-cadherin (placental)	3.98	NS	-1.49
CDH16	cadherin 16, KSP-cadherin	1.59	-1.81	-2.15
ITGA6	integrin, alpha 6	-1.1	NS	-1.1
ITGB6	integrin, beta 6	6.73	-1.34	-1.59
ITGB8	integrin, beta 8	1.24	-1.61	NS
ITGB1	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	-1.77	-1.39	-1.63
ITGAV	integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51)	-1.24	NS	-1.43
ITGA6	integrin, alpha 6	-1.1	NS	-1.1

NS stands for “non-significant changed”

3.5 Characterization of Molecular Markers at Transcriptional Level

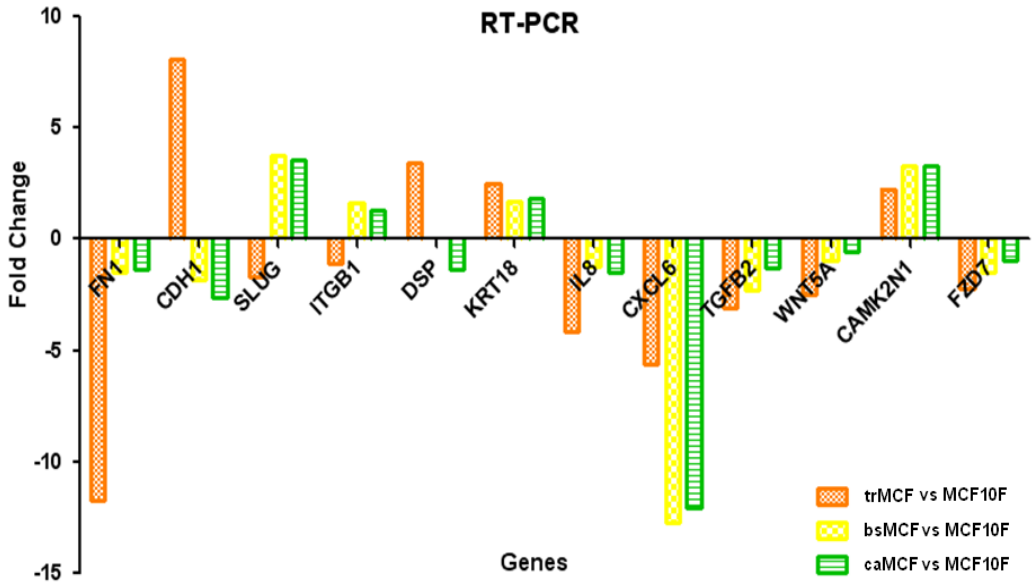


Fig.6 Results of Real Time RT-PCR This figure indicates the transcriptional characterization of the dysregulated genes conducted by real time RT-PCR. The fold change of differentially expressed genes is labeled with different colors.

Based on the SLUG-centered hypothetical pathway, 12 molecular markers demonstrated in Fig. 6 were chosen as the targets for validation through real time RT-PCR. In trMCF cells, the results of RT-PCR agreed with results of microarray analysis. The mRNA of SLUG, ITGB1, FN1 and CXCL6 were suppressed, whereas CDH1 and DSP showed enhanced transcriptional levels. This transcriptional pattern supported the developed SLUG-centered pathway in trMCF. In bsMCF and caMCF cells, the results of RT-PCR conflicted with results of microarray analysis. We found that mRNA levels of SLUG and ITGB1 were up-regulated; meanwhile, CDH1 and DSP demonstrated suppressed

transcriptional status. However, this expression pattern can still fill into our SLUG-centered pathway since the up-regulation of SLUG can enhance the expression of ITGB1 and suppress CDH1 and DSP. RT-PCR results demonstrated that the expression pattern of SLUG-centered pathway might still exist in bsMCF and caMCF, but in the opposite direction.

3.6 Characterization of Molecular Markers at Translational Level

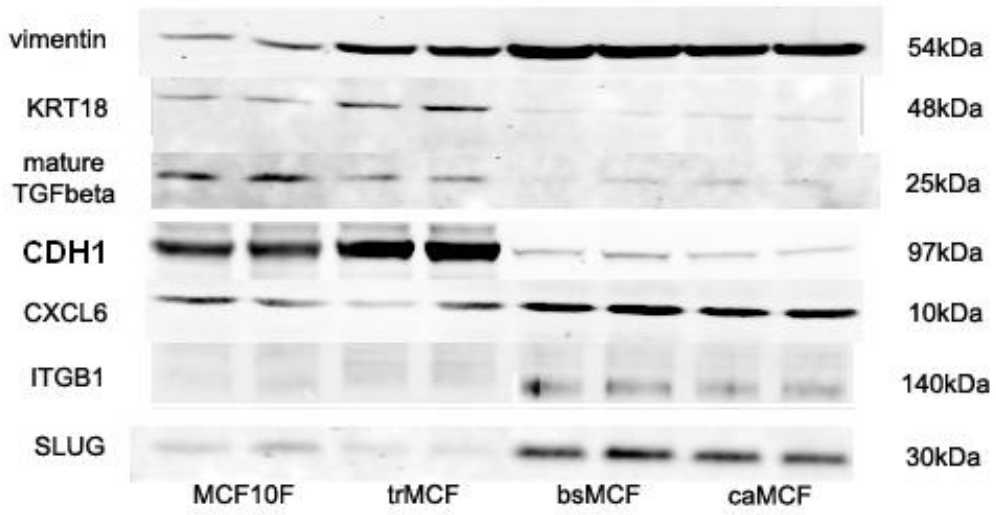


Fig.7 Results of Western Blot This figure indicates the translational characterization on dysregulated genes conducted by Western blot.

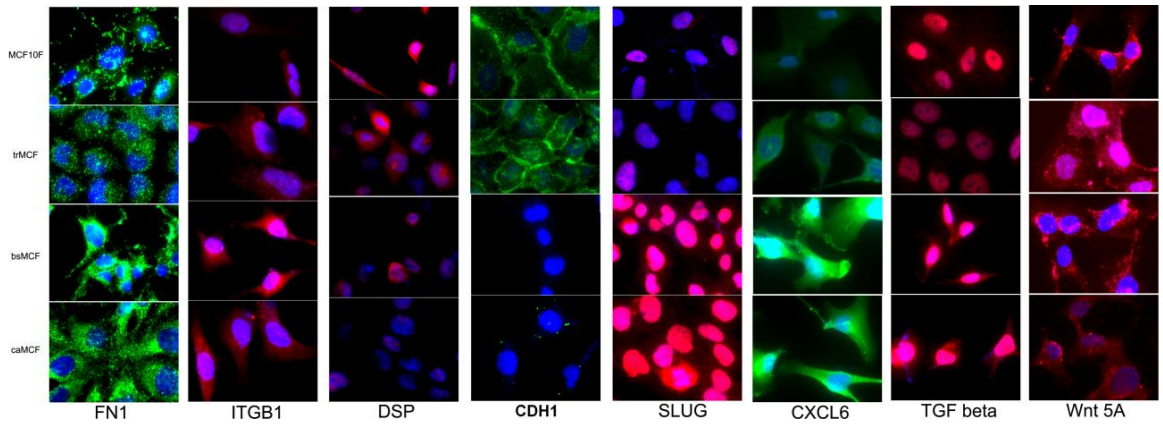


Fig.8 Results of Immunofluorescence Cell Culture Staining The localizations of dysregulated genes are characterized by immunocytochemistry. The blue dots represent the reaction with DAPI for nucleus staining. The green and red staining separately represent the reaction with specific antibodies as they are marked above.

We examined the translational expression pattern of the markers of SLUG-centered pathway in relation to the tumorigenic capacity of the cells in our model. The results of Western blot and Immunofluorescence are demonstrated in Fig. 7 and Fig. 8. In trMCF cells, the protein levels of SLUG and ITGB1 were suppressed, while CDH1 and DSP were enhanced. The protein expression of these four molecules agreed with their mRNA levels characterized in the RT-PCR supporting the SLUG-centered pathway in trMCF. In bsMCF and caMCF, the translational levels of SLUG, ITGB1, CDH1 and DSP demonstrated the same alterations as their transcriptional levels. This phenomenon indicates that the stem of SLUG-centered pathway is present in trMCF, bsMCF and caMCF. FN1 and CXCL6 possess a direct relationship with integrin signaling in our

SLUG-centered hypothetical pathway. In translational characterization, both FN1 (Fig 8) and CXCL6 (Fig 7 and 8) showed enhanced expressions in bsMCF and caMCF, following the same behavior as ITGB1. Therefore, the translational expressions of FN1 and CXCL6 support the function of SLUG-centered pathway in our model, although FN1 and CXCL6 demonstrated conflicts between their transcriptional and translational levels.

CHAPTER 4

DISCUSSION

This study integrates thanscriptomic data analyses to elucidate the progressive molecular events in the E2-mediated malignant transformation of ER-(-) HBECs. Expression patterns progressively accumulated as the cells expressed more aggressive phenotypes (i.e., in the tumorigenic bsMCF and caMCF) in comparison with the non-tumorigenic trMCF cells. These findings reveal an intrinsic and independent network within E2-induced gene expression alterations, and tumorigenesis in ERa (-) HBECs.

The hypothesis originates from the SLUG transcriptional change under the effects of E2 treatment, since SLUG can be regulated by estrogen through many pathways, such as the MTA gene and E-box [35, 36]. Snail family proteins are zinc-finger transcription factors and core EMT regulatory factors that are essential in both development and metastasis in carcinomas. Both SLUG and SNAI1 belong to the Snail family of proteins containing an NH_2-terminal repression domain and a COOH-terminal zinc-finger DNA-binding domain [35]. The down-regulation of SLUG exerts its role of transcriptional regulator through binding the E-box and alters the expression of many ECM components such as fibronectin1, E-cadherin, keratins and vimentin. In our model, the ECM component is represented by the expression variance of fibronectin1, which is involved in cell adhesion, growth, migration, differentiation and wound healing processes. The altered

mesenchymal gene expression is believed to be the hallmark of EMT, which involves differentiation of polarized epithelial cells to a migratory fibroblastoid phenotype, a phenomenon that is increasingly considered to be an important event during cancer progression and metastasis.

Of interest, SLUG was down-regulated in trMCF and the down-regulation of SLUG at the transformation phase was first examined by us. This sharp turn in expression may indicate a trigger mechanism of EMT in E2-induced basal-like breast cancer progression. The expression of SLUG is altered at a relative early phase during tumorigenesis; thus the mechanism can still exert its function through other oncogenes, like p53, which usually malfunctions at a later phase. The downstream components of SLUG are mainly ECM molecules, and many proven metastasis markers, such as vimentin and E-cadherin, are linked together by this hypothesis. Additionally, the ECM alteration triggered by SLUG can connect with the ERK system and control TGFß expression through integrin signaling, so that, the connections of both cell proliferation and invasiveness events are related to the hypothesis [37, 38].

Integrins function as heterodimeric receptors for extracellular matrix proteins, mediating cell anchorage [39]. Due to the ECM destruction, the integrin signaling pathway was the most significantly altered pathway in the progression of neoplastic transformation. The integrin compositions recognizing fibronectin, collagen and laminin are altered to

cooperate with the change of ECM components; thus, the signal starting with integrin will be decreased. The integrin signaling pathway was enriched by the dysregulated genes in trMCF, indicating that this pathway was affected in early stages of cell transformation. In addition, GO analysis revealed enrichment of up-regulated genes in the cell proliferation process in tumorigenic cells, which can support the dysfunctional integrin signaling in same direction. Our results support the concept that E2-induced breast cancer is a polygenic disease with a large range of genomic instabilities. E2 and/or its metabolites can directly cause genomic aberrations and transcriptome alterations without the mediation of ERα. As a consequence, changes in gene expression result in disrupted integrin signaling and epithelial to mesenchymal transition.

REFERENCES

1. Thomassen, M., Q. Tan, and T.A. Kruse, *Gene expression meta-analysis identifies metastatic pathways and transcription factors in breast cancer.* BMC Cancer, 2008. **8**: p. 394.
2. Hoefnagel, L.D., et al., *Prognostic value of estrogen receptor alpha and progesterone receptor conversion in distant breast cancer metastases.* Cancer, 2012.
3. Tutt, A., et al., *Risk estimation of distant metastasis in node-negative, estrogen receptor-positive breast cancer patients using an RT-PCR based prognostic expression signature.* BMC Cancer, 2008. **8**: p. 339.
4. Chu, K.C., et al., *Frequency distributions of breast cancer characteristics classified by estrogen receptor and progesterone receptor status for eight racial/ethnic groups.* Cancer, 2001. **92**(1): p. 37-45.
5. Hattangadi-Gluth, J.A., et al., *Basal subtype of invasive breast cancer is associated with a higher risk of true recurrence after conventional breast-conserving therapy.* Int J Radiat Oncol Biol Phys, 2012. **82**(3): p. 1185-91.
6. Bertucci, F., P. Finetti, and D. Birnbaum, *Basal breast cancer: a complex and deadly molecular subtype.* Curr Mol Med, 2012. **12**(1): p. 96-110.
7. Chen, X.S., et al., *Molecular subtype can predict the response and outcome of Chinese locally advanced breast cancer patients treated with preoperative therapy.* Oncol Rep, 2010. **23**(5): p. 1213-20.

8. DiMeo, T.A., et al., *A novel lung metastasis signature links Wnt signaling with cancer cell self-renewal and epithelial-mesenchymal transition in basal-like breast cancer.* Cancer Res, 2009. **69**(13): p. 5364-73.

9. Sabatier, R., et al., *Kinome expression profiling and prognosis of basal breast cancers.* Mol Cancer, 2011. **10**: p. 86.

10. Marginean, F., et al., *Histological features of medullary carcinoma and prognosis in triple-negative basal-like carcinomas of the breast.* Mod Pathol, 2010. **23**(10): p. 1357-63.

11. Huang, Y., et al., *Epithelial to mesenchymal transition in human breast epithelial cells transformed by 17beta-estradiol.* Cancer Res, 2007. **67**(23): p. 11147-57.

12. Russo, J., et al., *17-Beta-estradiol induces transformation and tumorigenesis in human breast epithelial cells.* FASEB J, 2006. **20**(10): p. 1622-34.

13. Russo, J., et al., *17Beta-estradiol is carcinogenic in human breast epithelial cells.* J Steroid Biochem Mol Biol, 2002. **80**(2): p. 149-62.

14. Russo, J. and I.H. Russo, *The role of the basal stem cell of the human breast in normal development and cancer.* Adv Exp Med Biol, 2011. **720**: p. 121-34.

15. Kumar, R., A. Sharma, and R.K. Tiwari, *Application of microarray in breast cancer: An overview.* J Pharm Bioallied Sci, 2012. **4**(1): p. 21-6.

16. Gautier, L., et al., *affy - analysis of Affymetrix GeneChip data at the probe level.* Bioinformatics, 2004. **20**(3): p. 307-315.

17. Gentleman, R., *Bioinformatics and computational biology solutions using R and Bioconductor*. Statistics for biology and health. 2005, New York: Springer Science+Business Media. xix, 473 p.

18. Diboun, I., et al., *Microarray analysis after RNA amplification can detect pronounced differences in gene expression using limma.* BMC Genomics, 2006. **7**.

19. Kerr, M.K., *Linear models for microarray data analysis: hidden similarities and differences.* Journal of Computational Biology, 2003. **10**(6): p. 891-901.

20. Maruyama, K. and S. Sugano, *Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides.* Gene, 1994. **138**(1-2): p. 171-4.

21. Savagner, P., K.M. Yamada, and J.P. Thiery, *The zinc-finger protein slug causes desmosome dissociation, an initial and necessary step for growth factor-induced epithelial-mesenchymal transition.* J Cell Biol, 1997. **137**(6): p. 1403-19.

22. Suzuki, Y., et al., *Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library.* Gene, 1997. **200**(1-2): p. 149-56.

23. Shih, J.Y., et al., *Transcription repressor slug promotes carcinoma invasion and predicts outcome of patients with lung adenocarcinoma.* Clin Cancer Res, 2005. **11**(22): p. 8070-8.

24. Ye, Y., et al., *ERalpha suppresses slug expression directly by transcriptional repression.* Biochem J, 2008. **416**(2): p. 179-87.

25. Schmidt, C.R., et al., *E-cadherin is regulated by the transcriptional repressor SLUG during Ras-mediated transformation of intestinal epithelial cells.* Surgery, 2005. **138**(2): p. 306-12.

26. Zhang, K., et al., *Slug regulates proliferation and invasiveness of esophageal adenocarcinoma cells in vitro and in vivo.* Med Oncol, 2011. **28**(4): p. 1089-100.

27. Turner, F.E., et al., *Slug regulates integrin expression and cell proliferation in human epidermal keratinocytes.* J Biol Chem, 2006. **281**(30): p. 21321-31.

28. Emadi Baygi, M., et al., *Slug/SNAI2 regulates cell proliferation and invasiveness of metastatic prostate cancer cell lines.* Tumour Biol, 2010. **31**(4): p. 297-307.

29. Shields, M.A., et al., *Interplay between beta1-integrin and Rho signaling regulates differential scattering and motility of pancreatic cancer cells by snail and Slug proteins.* J Biol Chem, 2012. **287**(9): p. 6218-29.

30. Vuoriluoto, K., et al., *Vimentin regulates EMT induction by Slug and oncogenic H-Ras and migration by governing Axl expression in breast cancer.* Oncogene, 2011. **30**(12): p. 1436-48.

31. Hirsch, J., et al., *PEDF inhibits IL8 production in prostate cancer cells through PEDF receptor/phospholipase A2 and regulation of NFkappaB and PPARgamma.* Cytokine, 2011. **55**(2): p. 202-10.

32. Natoli, G., et al., *Tumor necrosis factor (TNF) receptor 1 signaling downstream of TNF receptor-associated factor 2. Nuclear factor kappaB (NFkappaB)-inducing kinase requirement for activation of activating protein 1 and NFkappaB but not of c-Jun N-terminal kinase/stress-activated protein kinase.* J Biol Chem, 1997. **272**(42): p. 26079-82.

33. Relaix, F., et al., *Peg3/Pw1 is an imprinted gene involved in the TNF-NFkappaB signal transduction pathway.* Nat Genet, 1998. **18**(3): p. 287-91.

34. Morel, J.C., et al., *Interleukin-18 induces rheumatoid arthritis synovial fibroblast CXC chemokine production through NFkappaB activation.* Lab Invest, 2001. **81**(10): p. 1371-83.

35. Come, C., et al., *Roles of the transcription factors snail and slug during mammary morphogenesis and breast carcinoma progression.* J Mammary Gland Biol Neoplasia, 2004. **9**(2): p. 183-93.

36. Shih, J.Y. and P.C. Yang, *The EMT regulator slug and lung carcinogenesis.* Carcinogenesis, 2011. **32**(9): p. 1299-304.

37. Karam, M., et al., *Protein kinase D1 stimulates proliferation and enhances tumorigenesis of MCF-7 human breast cancer cells through a MEK/ERK-dependent signaling pathway.* Exp Cell Res, 2012. **318**(5): p. 558-69.

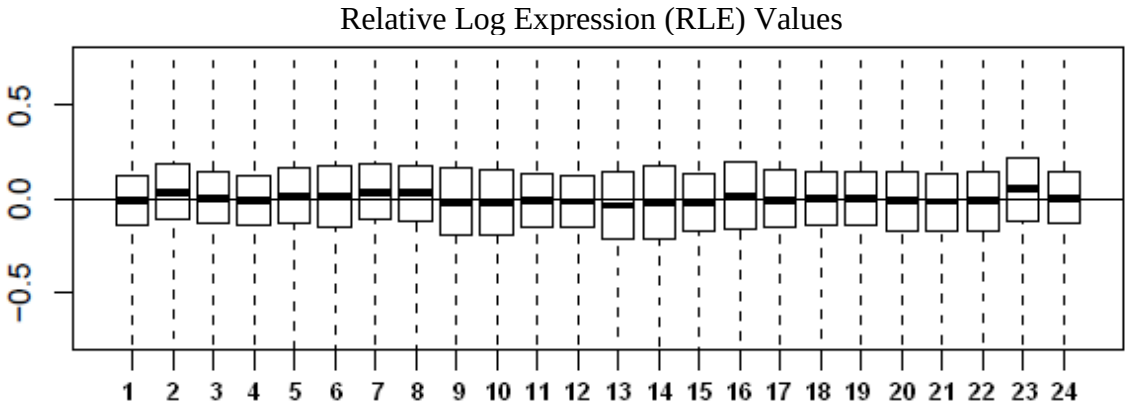
38. Xu, Z., et al., *TGFbeta and EGF synergistically induce a more invasive phenotype of epithelial ovarian cancer cells.* Biochem Biophys Res Commun, 2010. **401**(3): p. 376-81.

39. Li, D.M. and Y.M. Feng, *Signaling mechanism of cell adhesion molecules in breast cancer metastasis: potential therapeutic targets.* Breast Cancer Res Treat, 2011. **128**(1): p. 7-21.

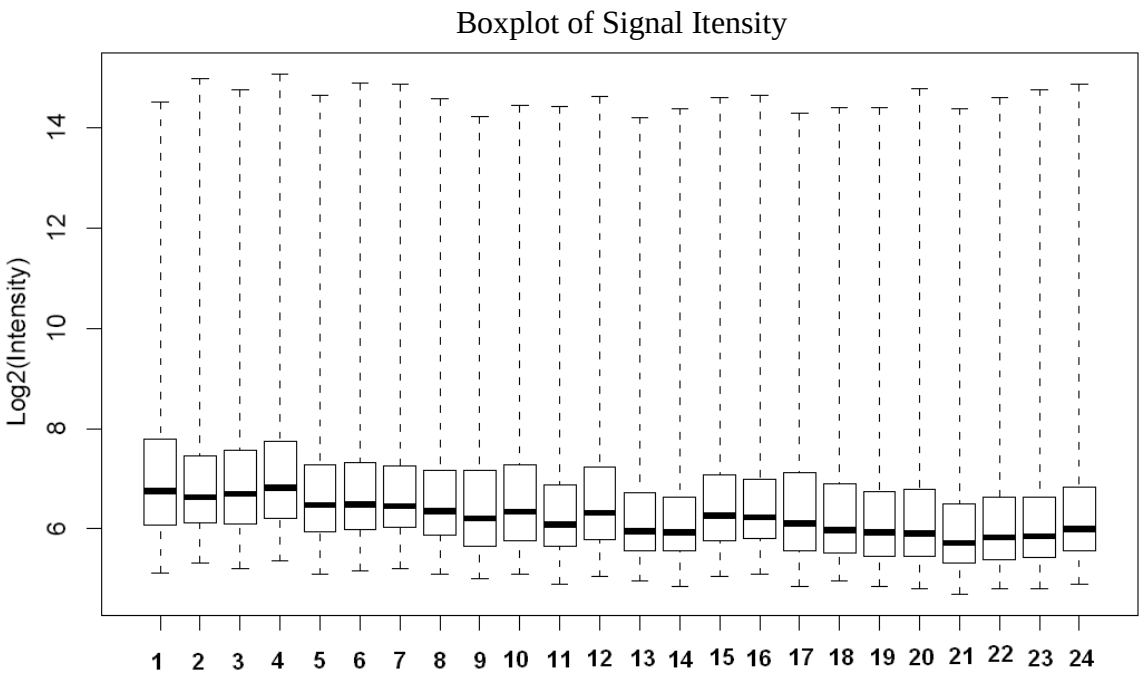
APPENDIX: SUPPLEMENTAL FIGURES AND TABLES

Supplemental Table 1 RNA Concentration and RNA Integrity Number (RIN)

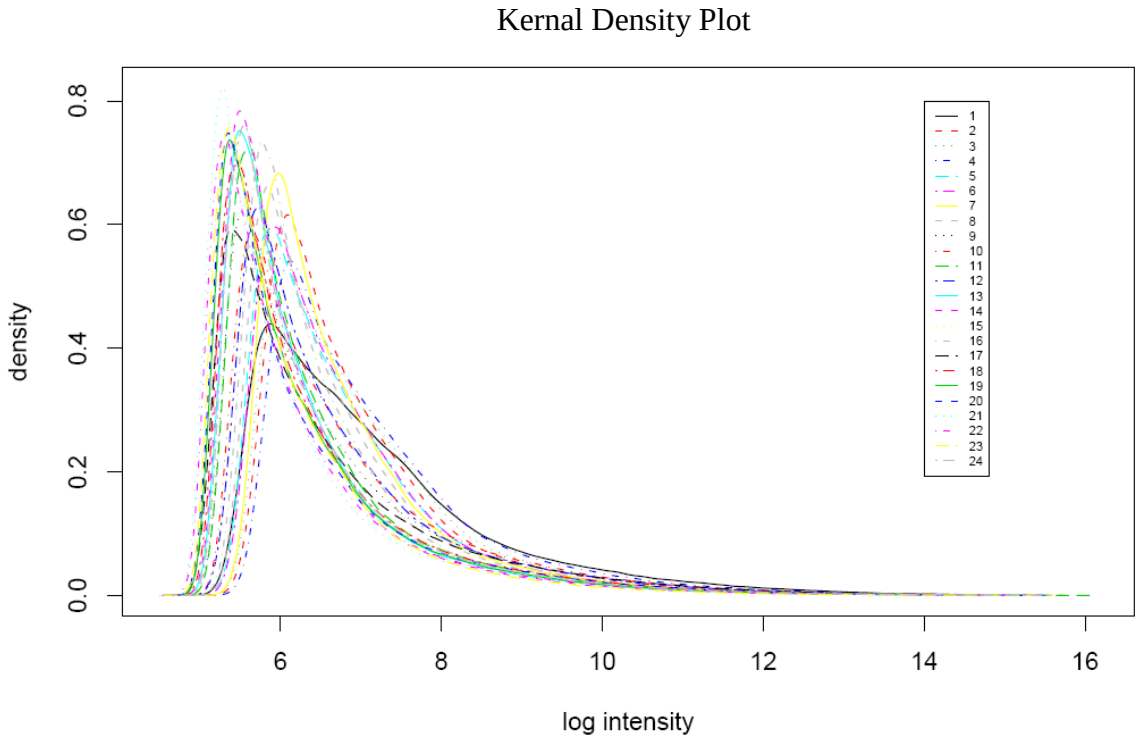
Cell line	Passage	Sample mark	RNA Concentration(ng/uL)	RIN
MCF-10F	p93	MCF-10F -1	19.43	8.7
		MCF-10F -2	45.69	9.1
	p90	MCF-10F -3	45.2	8.5
		MCF-10F -4	30.4	9.9
	p91	MCF-10F -5	69.2	9
		MCF-10F -6	19.1	8.5
trMCF	p55	trMCF-1	174.2	9.2
		trMCF-2	213.7	8.8
	p56	trMCF-3	269	8.7
		trMCF-4	144.4	9
	p57	trMCF-5	72.4	8.9
		trMCF-6	115.2	8.9
bsMCF	p72	bsMCF-1	460.5	8.1
		bsMCF-2	308.9	8.4
	p73	bsMCF -3	325	8.3
		bsMCF -4	152.2	8.7
	p74	bsMCF -5	343.4	8.5
		bsMCF -6	379.4	8.3
caMCF	p36	caMCF -1	170	8.1
		caMCF -2	89.5	8.5
	p37	caMCF -3	239.4	9.1
		caMCF -4	172.2	9
	p38	caMCF -5	250.5	8.7
		caMCF -6	348.4	8.1



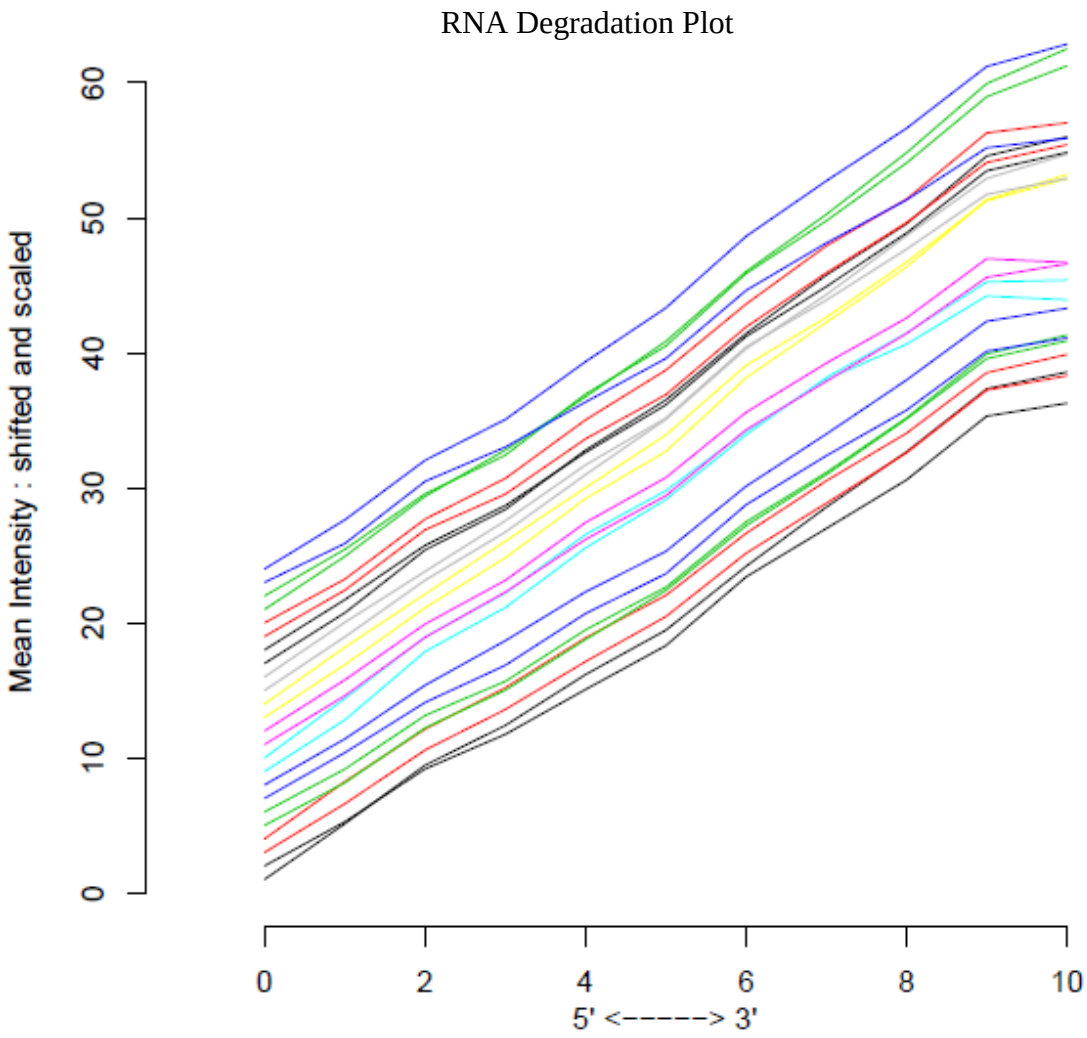
Supplemental Fig.1 Relative Log Expression (RLE) Values RLE values are computed for each probe by comparing the expression value on each array against the median expression value for that probe across all arrays. Assuming that most genes are not changing in expression, the ideally values will be near 0. For all these samples, the RLE values are in the vicinity of 0, which means all the RNA have the same expression level.



Supplemental Fig.2 Boxplot of Signal Intensity Box plot of signal intensity indicates the intensity of each sample in $\log_2$ form across all the chips.

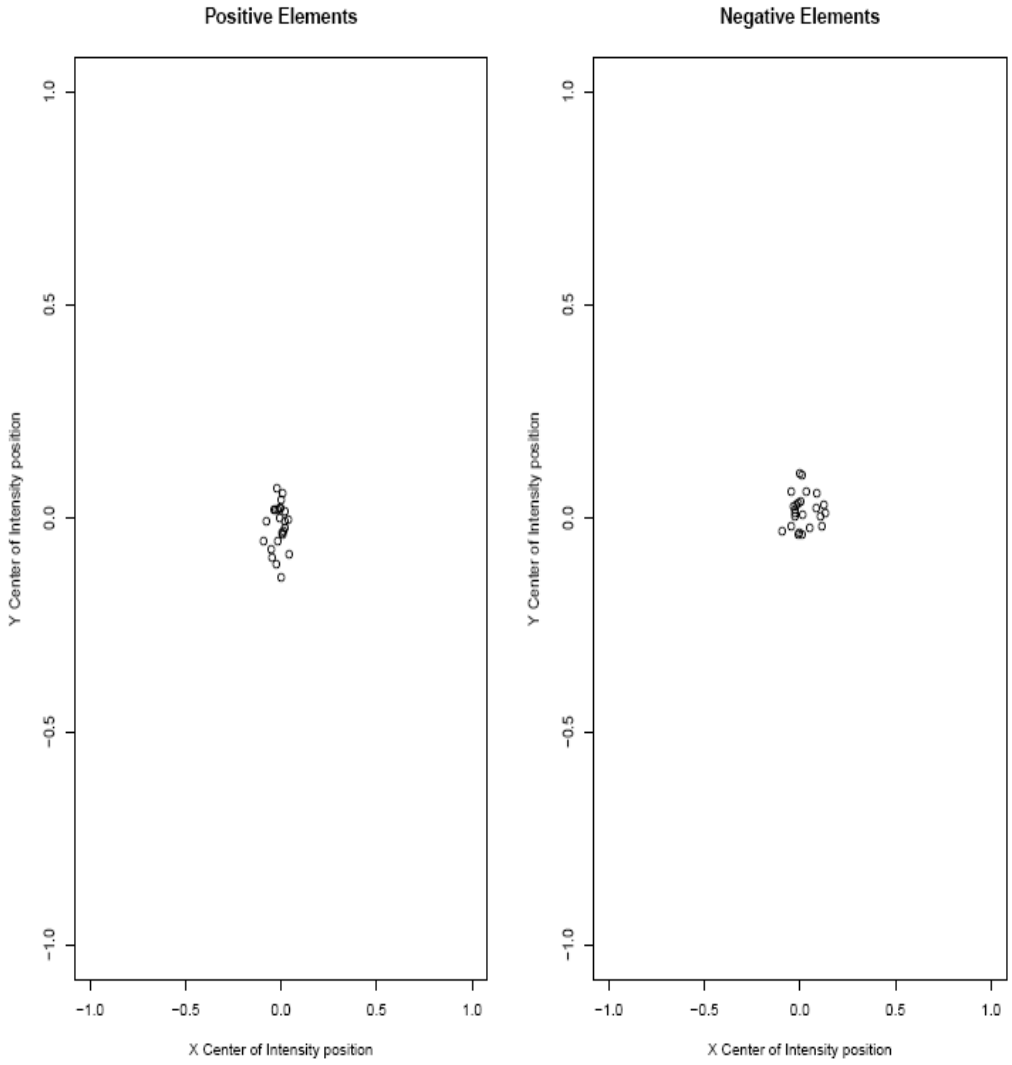


Supplemental Fig.3 Kernal Density Plot The Kernal density plot assumes that in the whole array, most of the probes are not differentially expressed. Thus, it is expected the signal intensities across all the chips are similar. Any array with a low average intensity or a significantly different shaped density would be suspected. For our samples, all of them form the density peaks around log intensity 6 and flat trend outside peak areas. This points out that the intensity distribution satisfied the requirement of affymetrix.

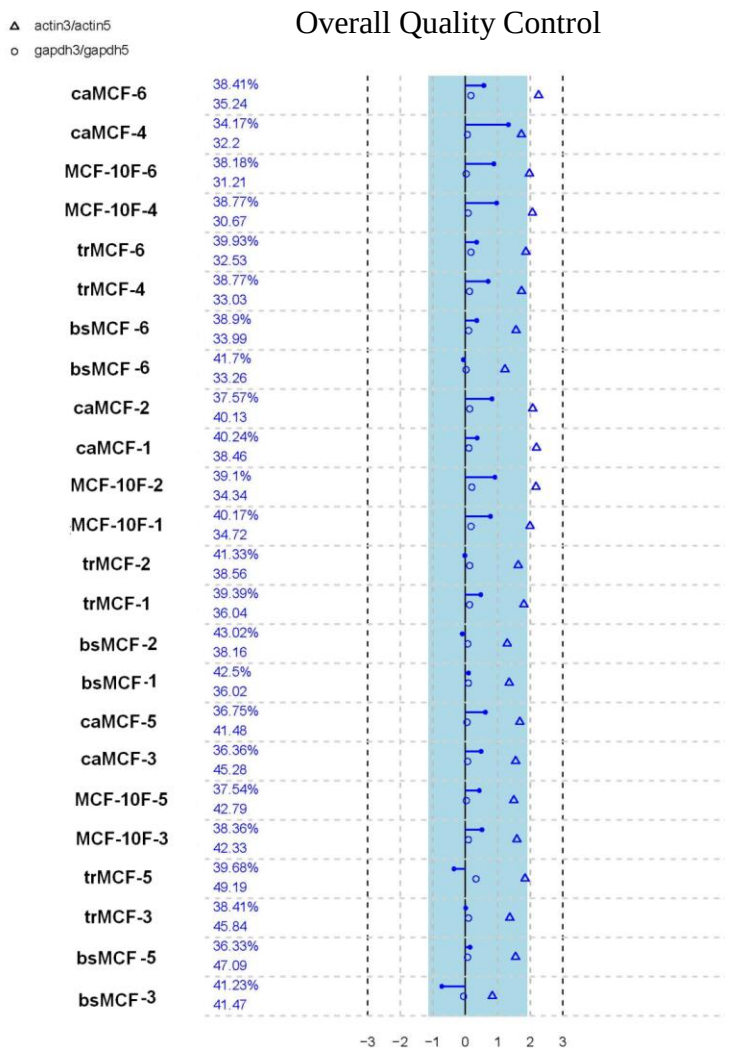


Supplemental Fig.4 RNA Degradation Plot RNA degradation plot demonstrates the RNA integrity degree affected by degradation. Since RNA degradation typically starts from the 5’ end of the molecule, it is expected that probe intensities to be systematically lowered at 5’ end of a probe when compared to the 3’ end. This result shows that none of the samples had unexpected RNA degradation.

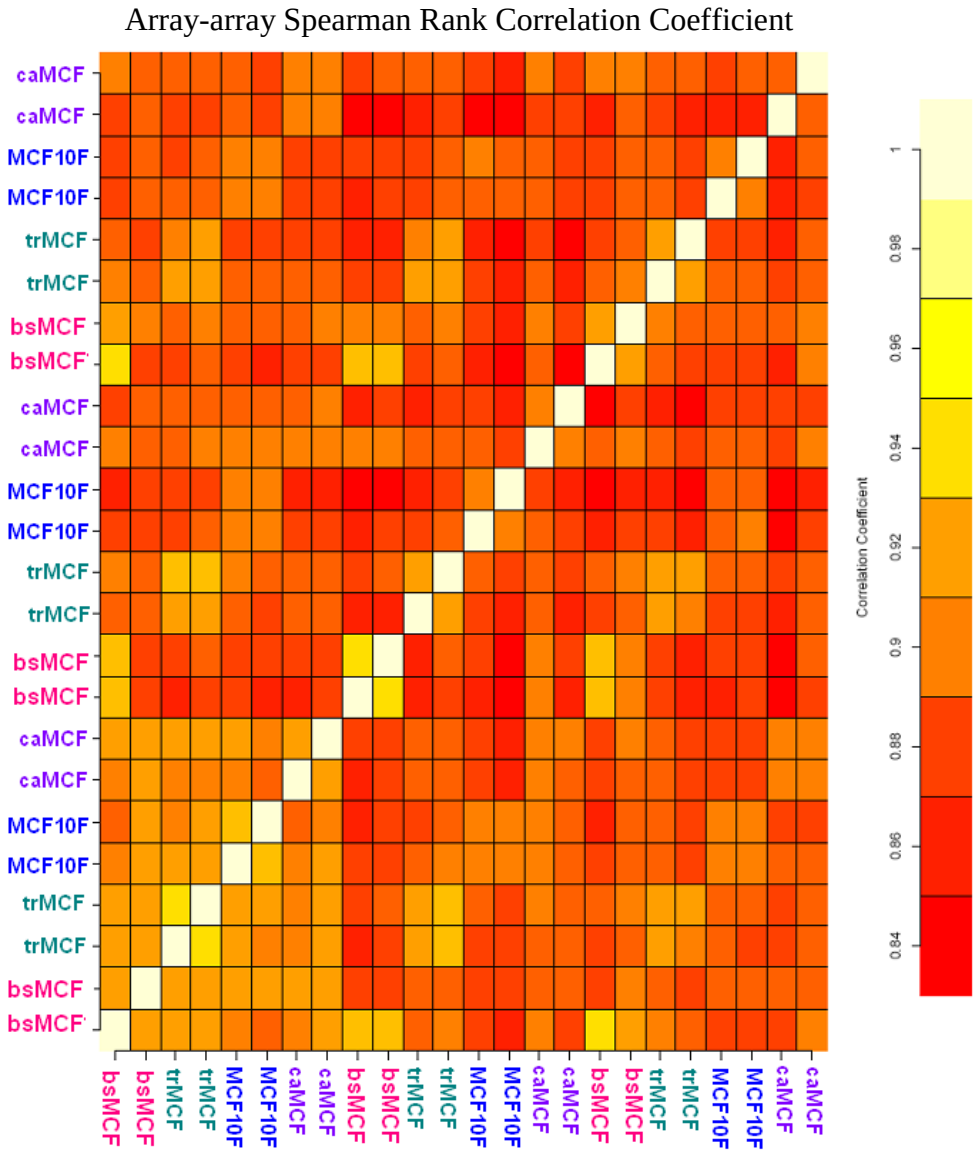
Center of Intensity (COI)



Supplemental Fig.5 Center of Intensity "Center of intensity" (COI) is calculated from the control elements on the outer edges of the Affymetrix arrays and based on which edge of the array they are located the. If the hybridization is uniform across the array, the COI will be located at the physical center of the array. Any spatial variations in the hybridization will cause the COI to move from center. The uniformity of the back-ground can be evaluated through both positive and negative controls.

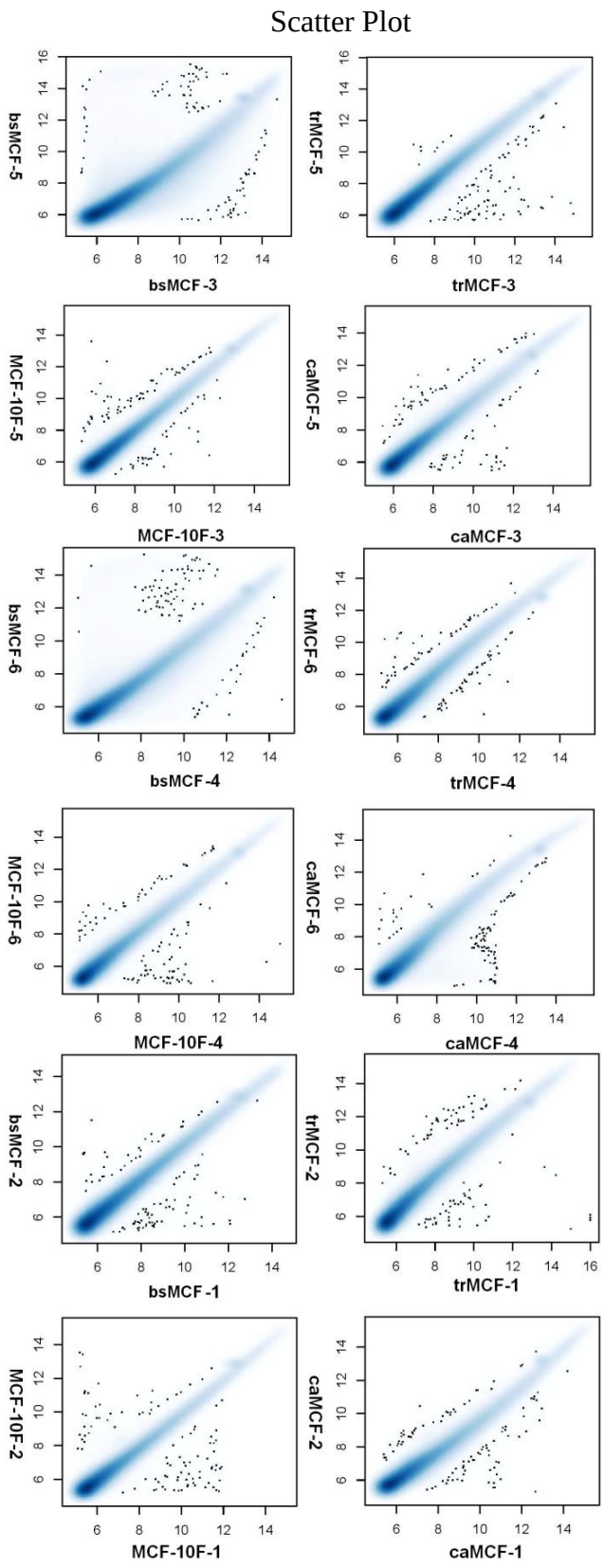


Supplemental Fig.6 Overall Quality Control This chart is generate by using the Internal control parameters: GAPDH 3’ : 5’ andβ-Actin 3’ : 5’. Theoretically, if there is no degradation in RNA, the ratio should be 1, however, since the degradation on the 5’end of the gene is higher. For GAPDH, the ratios up to 1.25 are considered acceptable; for β-Actin 3’ : 5’ , a longer gene, the recommendation is to be below 3 values The blue stripe in the image represents the range where scale factors are within 3-fold of the mean for all chips. Present calls and average background are also returned with this chart.



Supplemental Fig.7 Array-array Spearman Rank Correlation Coefficient

The arrays are arranged by using the phenotypic data in order to place arrays with similar samples adjacent to each other. Self-self correlations are on the diagonal and by definition have a correlation coefficient of 1.0. Data from similar tissues or treatments will tend to have higher coefficients. These results show that the corresponding samples are good replications for each others.



Supplemental Fig.8 Scatter Plot
Scatter plot is the comparisons conducted between the samples from same cell line in same batch. The concentrated dots distribution and straight shape indicate the good quality in reproducibility of these two samples.