From DEPARTMENT OF ONCOLOGY-PATHOLOGY
Karolinska Institutet, Stockholm, Sweden

PROTEOMICS OF INVASIVENESS OF HUMAN BREAST EPITHELIAL CELLS

Kah Wai Lin

Karolinska Institutet

Stockholm 2012

Published by Karolinska Institutet. Printed by Larserics Print Digital Print AB,
Landsvagen 65, Sundbyberg, Box 200, SE 17177 Stockholm, Sweden

ISBN 978-91-7457-634-4

ABSTRACT

Mechanisms of malignant transformation and cancer invasion and factors controlling them in response to various stimuli remain elusive. We used proteomics and systems biology to explore these mechanisms in human breast epithelial cells. The objective has been to identify a set of biomarkers for diagnostics and prognostics of breast cancer.

Acquiring of high proliferation by cells is a major hallmark of malignant transformation. Using global expression proteome profiling approach, we identified a set of proteins associated with high proliferation rate of human breast epithelial cells upon carcinogenic transformation (paper I). In this study, we described a proteome signature of cells with enhanced proliferation rate, and observed that deregulation of CDK4 and cyclin D3 may be among the early malignant transformation events.

Distal metastasis is the leading cause of death among breast cancer patients, and invasion of cancer cells is the first step in metastatic process. We established a highly invasive clone of MCF7 cells from non-invasive MCF7 cells (paper II). Using proteome profiling, we identified key regulators of invasiveness. Systemic analysis suggested that the invasive-specific network has features of a scale-free network, with TGFβ, EGFRB, TAF1, HNF4α, MYC and RB1 as key nodes. Analysis of TGFβ and EGF-centered network showed more than 30 key nodes which may define how TGFβ and EGF cooperate. Among these nodes were identified insulin, VEGF, HNF4α and NFκB. This result indicates that the insulin signaling disturbance may interfere with the invasiveness, thus explain the clinical observation of the increased risk of breast cancer metastasis in diabetes patients.

The correlation between protein translation and breast cancer is crucial in understanding of breast carcinogenesis. In paper III, we identified eukaryortic elongation factor 1 A1 (eEF1A1) as a direct substrate of type I transforming growth factor-β-receptor (TβRI). We showed that the phosphorylation of eEF1A1 at Ser300 by TβR-I mediates a direct inhibitory effect of TGFβ on protein synthesis, and contributes to effects on cell proliferation, anchorage-dependent and anchorage-independent cell growth. Furthermore, we showed that the phosphorylation of Ser300 is decreased in human breast tumors. In paper IV, we showed that eEF1A1 itself contributed to the increased proliferation of human breast epithelial cells by promoting transition of cells through the S- and G2/M-phases of the cell cycle. Therefore, our identification of eEF1A1 as a substrate of TβR-I unveiled novel translation-related regulatory pathway downstream of TβR-I, which is involved in breast tumorigenesis.

Breast cancer metastatic suppressor I (BRMS1) was identified by us as an invasiveness-related protein. In paper V, we showed that expression of BRMS1 resulted in a shift to epithelial morphology of otherwise mesenchymal morphology MDA-MB-231 cells. Our study concluded that TGFβ and EGF may modulate BRMS1-dependent breast cancer invasion by regulating focal adhesion and cytoskeletal rearrangement, and that Smad2 and Erk1/2 phosphorylation are involved in molecular mechanisms engaged by BRMS1.

Thus, presented here studies delivered a proteome signature of invasiveness and enhanced proliferation, and explored roles of eEF1A1 and BRMS1 in breast tumorigenesis. We described proteome signatures and proteins which may be considered as markers for diagnostics and prognostics of human breast cancer.

LIST OF PUBLICATIONS

I. Nimesh Bhaskaran*, Kah Wai Lin*, Aude Gautier*, Hanna Woksepp, Ulf Hellman, Serhiy Souchelnytskyi. **Comparative proteome profiling of MCF10A and 184A1 human breast epithelial cells emphasized involvement of cdk4 and cyclin D3 in cell proliferation**. *Proteomics - Clinical Applications*. 2009; 3(1): 68-77.

II. Kah Wai Lin, Serhiy Souchelnytskyi. **Proteome signature of invasiveness of human breast epithelial cells**. (Manuscript)

III. Kah Wai Lin*, Ihor Yakymovych*, Min Jia, Mariya Yakymovych, Serhiy Souchelnytskyi. **Phosphorylation of eukaryotic elongation factor eEF1A at Ser300 by type I transforming growth factor-β-receptor results in inhibition of mRNA translation**. *Current Biology*. 2010; 20(18): 1615-1625.

IV. #Kah Wai Lin, Serhiy Souchelnytskyi. **Eukaryotic elongation factor eEF1A1 promotes and Ser300 mutants of eEF1A1 inhibit transition through the S and G2/M phases of the cell cycle**. *Journal of Cell and Molecular Biology*. 2010; 8(2): 125-130.

V. #Kah Wai Lin, Serhiy Souchelnytskyi. **TGFβ and EGF coordinately modulate BRMS1-mediated breast cancer invasion by activation of focal adhesion and cytoskeleton rearrangement**. (Manuscript)

ADDITIONAL PUBLICATIONS

I. #Kah Wai Lin, Serhiy Souchelnytskyi. **Translational connection of TGFβ signaling: Phosphorylation of eEF1A1 by TβR-I inhibits protein synthesis**. *Small GTPase*. 2011; 2(2): 104-108.
(Invited Extra View)

II. #Kah Wai Lin, Min Jia, Serhiy Souchelnytskyi. **Application of bioinformatics tools in gel-based proteomics**. In *Proteomics / Book 2* (Eastwood Leung, ed, ISBN 979-953-307-693-4, InTech, 2011. (In Press)
(Invited Book Chapter)

III. Min Jia, Kah Wai Lin, Serhiy Souchelnytskyi. **Phosphoproteomics: detection, identification and importance of protein phosphorylation**. In *Proteomics / Book 2* (Eastwood Leung, ed), ISBN 979-953-307-693-4, InTech, 2011. (In Press)
(Invited Book Chapter)

* contributed equally
# corresponding author

CONTENTS

LIST OF ABBREVIATIONS

2D-DIGE	2 dimensional differential gel electrophoresis
2D-GE	2 dimensional gel electrophoresis
aa-tRNA	aminoacyl tRNA
BRCA1	breast cancer type I susceptibility protein
BRCA2	breast cancer type II susceptibility protein
BRMS1	breast cancer metastasis suppressor 1
cdc25a	cell division cycle 25 homolog A
cdk	cyclin-dependent kinase
CGAP	cancer genome anatomy project
c-myc	c-myelocytomatosis viral oncogene
Co-Smad	common-mediator Smad
DAG	directed acyclic graph
ECM	extracellular matrix
eEF1A	eukaryotic elongation factor 1 A
EGF	epidermal growth factor
EMT	epithelial to mesenchymal transition
erbB2	erythroblastic leukemia viral oncogene homolog 2
ERK	extracellular signal-regulated protein kinases
FACS	fluorescence activated cell sorting
GO	gene ontology
GOTM	gene ontology tree machine
GDP	guanosine diphosphate
GTP	guanosine triphosphate
GTPase	guanosine triphosphate hydrolase
HDAC	histone deacetylase
HGF	hepatocyte growth factor
HNF4α	hepatocyte nuclear factor 4 alpha
HPLC	high performance liquid chromatography
IGF	insulin-like growth factor
IHC	immunohistochemistry
LAP	latency-associated peptide
LC	liquid chromatography
LTBP	latent TGFβ-binding protein
L-TGFβ	latent-transforming growth factor beta
MALDI-TOF	matrix-assisted laser desorption/ionisation-time of flight
MAPK	mitogen-activated protein kinase
MMP-2	matrix metallopeptidase 2
MMP-9	matrix metallopeptidase 9
MS	mass spectrometry
Mw	molecular weight
NCBI	National Center for Biotechnology Information
NFκB	nuclear factor kappa B

OBO	open biomedical ontologies consortium
PMF	peptide mass fingerprinting
PRO	protein ontology
PTM	post-translational modification
PI3K	phosphatidylinositol 3-kinase
RPPA	reverse phase protein arrays
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
pI	isoelectric point
Rb	retinoblastoma
R-Smad	receptor-activated Smad
Smad	SMA-related, mothers against decapentaplegic protein
Src	sarcoma oncogene
TMA	tissue microarray
TGFβ	transforming growth factor beta
TIMP-1	tissue inhibitor of metalloproteinases 1
TIMP-2	tissue inhibitor of metalloproteinases 2
TNF	tumor necrosis factor
TβR-I	transforming growth factor receptor type I
VEGF	vascular endothelial growth factor
Wnt	wingless-type MMTV integration site family

1 INTRODUCTION

1.1 BREAST CANCER

Breast cancer is the most common female malignancy worldwide. In Europe, breast cancer is the most frequent cancer and the most frequent cause of cancer death among the women (Ferlay et al., 2007). Most deaths of the breast cancer patients, as a consequence of resistance to treatment, result from cancer metastasis to distal organs rather than from primary tumors.

Mechanisms of malignant transformation and invasion, and the intrinsic factors controlling them in response to various stimuli remain elusive. Unraveling the underlying mechanisms is a prerequisite step for further developing of therapeutics drugs. Over last several decades, cancer researchers pay great efforts in identifying underlying molecular pathways and biomarkers for diagnosis, treatment and prognosis purposes.

1.1.1 Malignant transformation in breast cancer

During malignant transformation, normal cells have to undergo multistep processes and acquire a series of changes that enable them to become cancerous. There are two main models of tumorigenesis, i.e. cancer stem cell/hierarchy model and clonal evolution/stochastic model of cancer growth (Fig. 1) (Dick, 2009; Shackleton *et al.*, 2009). The cancer stem cell/hierarchy model suggests that cancer is derived from population of highly specialized cancer stem cells, e.g. CD44+/CD24- cells represent the population of breast cancer stem cells. These cells are able to self-renew and proliferate extensively and eventually developed into tumor mass. However, the clonal evolution/stochastic model of cancer growth suggested that all cancer cells possess intrinsic potential that contribute to the unlimited growth (Dick, 2009; Shackleton *et al.*, 2009). To date, there is no definitive evident in favor of either model. This is probably depending on the type of cancer, stage of cancer development and environmental factor. Further studies are required to understand the origin and development of cancer.

In recent year, "Hallmarks of Cancer" have been proposed, which is an attempt to form organizing principles that provide a framework for understanding the complexity of cancer formation (Hanahan and Weinberg, 2000; Hanahan and Weinberg, 2011). The Hallmarks of Cancer compose of six essential alterations in cell physiology that lead to malignant transformation of normal cells. These included self-sufficiency of cells in growth signals, insensitivity to growth-inhibitory signals, evasion of apoptosis, limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis (Fig. 2) (Hanahan and Weinberg, 2000). The advance in cancer research over past decade has added two additional hallmarks, i.e. deregulation of cellular metabolism and avoiding of immune destruction; and two characteristics of cancer, i.e. genomic instability and mutation and tumor-promoting inflammation (Fig. 2) (Hanahan and Weinberg, 2011).

Each of these changes lead to the breaching of the anti-cancer defense mechanism of the body, followed by unlimited growth and proliferation of tumor cells, and eventually spread to the distal part of body.

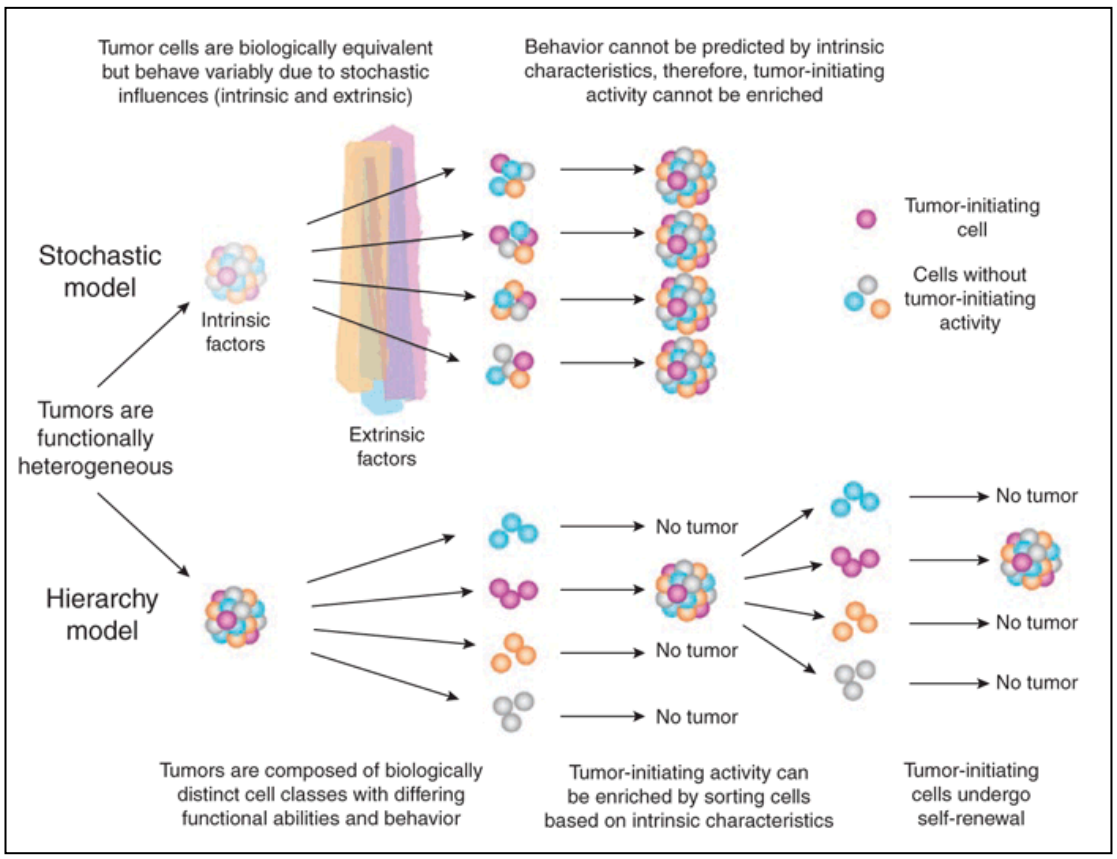

Fig 1. Models of tumorigenesis: cancer stem cell/hierarchy model and clonal evolution/stochastic model of cancer growth. The cancer stem cell/hierarchy model suggests that cancer is derived from population of highly specialized cancer stem cells. These cells are able to self-renew and proliferate extensively and eventually developed into tumor mass. Clonal evolution/stochastic model of cancer growth suggested that all cancer cells possess intrinsic potential that contribute to the unlimited growth. Reprinted from Dick, Nature Biotechnology 2009, with permission from Nature Publishing Group.

In addition to understand the biology of cancer, the Hallmarks of Cancer form a solid platform for the targeting therapy of cancer. Drugs that interfere with each of the acquired capabilities necessary for cancer formation have been developed, and are in clinical use for the treatment of various form of cancer (Fig. 2). For example, EGFR inhibitors can be used for targeting the proliferative signal of the cancer cells. Transtuzumab, or Herceptin, a humanized monoclonal antibody directed against HER-2, is used for the treatment for metastatic breast cancer in the clinic (Browne *et al.*, 2009; O'Donovan and Crown, 2007). Further understanding of the physiology of cancer will aid the development of novel therapeutic strategies.

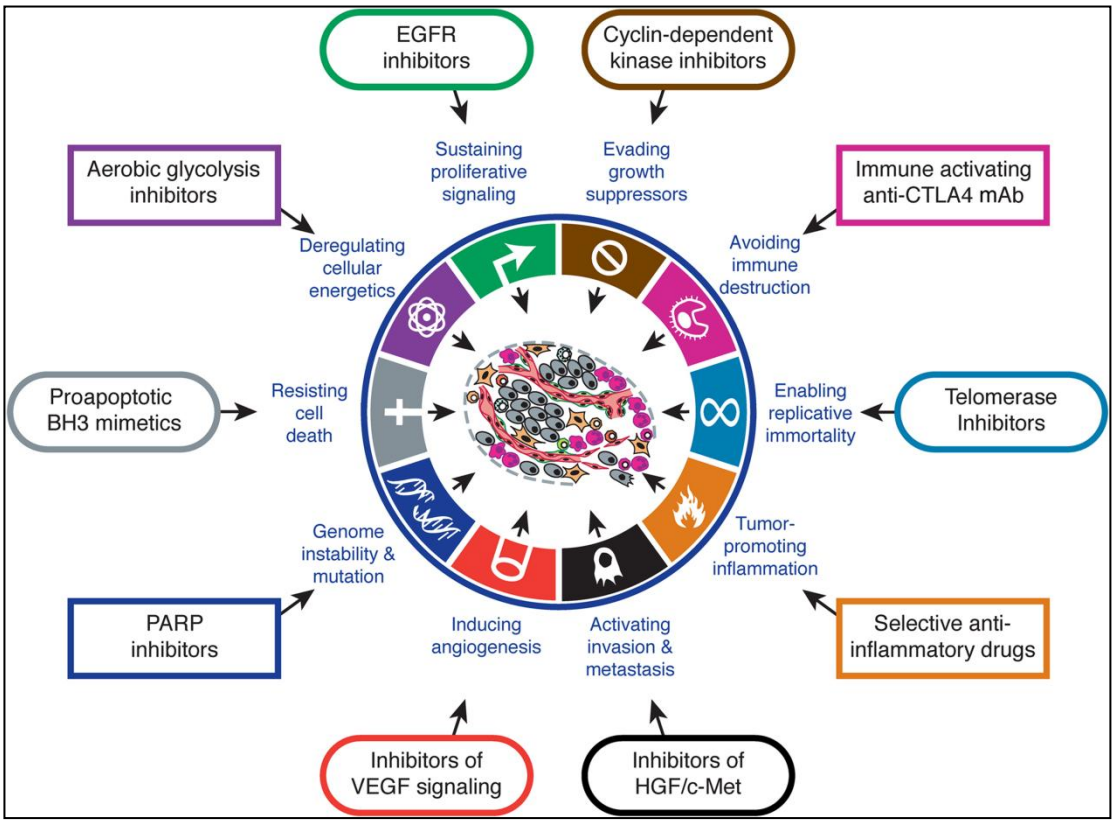

Fig. 2. Hallmark of cancer. Most cancers have acquired the same set of functional capabilities during their development through various mechanistic strategies. Acquiring of these capabilities lead to an unlimited growth and proliferation of tumor cells, eventually spread to the distal part of body. Various drugs that interfere with each of the acquired capabilities necessary for cancer formation have been developed and are in clinical use for the treatment of various form of cancer. Reprinted from Hanahan and Weinberg, 2011, with permission from Elsevier.

1.1.1.1 Regulation of cell proliferation in breast cancer

In a normal mammary gland, cell proliferation is tightly regulated by a complex interaction of hormones, growth factors and cytokines, and converges on activation of the cell cycle regulatory machinery, e.g. cyclins, cyclin-dependent kinases (CDK), CDK inhibitors, e.g. p16, and their regulators, e.g. c-myc, Ras/Raf/Erk. During the breast malignant transformation, expression and activities of these proteins is often deregulated, thereby the signaling pathways that control the normal proliferation are defective, causing the abnormal and unlimited growth of breast cancer cells (Butt *et al.*, 2008; Doisneau-Sixou *et al.*, 2003; Mester and Redeuilh, 2008). The expression of oncogenes, cell cycle regulators and mitogenic signaling pathways are associated with malignant and aggressive phenotype, of cancer.

In familial breast cancer, mutation of multiple breast cancer susceptibility genes, such as BRCA1, BRCA2 and p53 have been implicated. It has been shown that mutation of BRCA1, BRCA2, and p53 are involved in the deregulation of cell cycle control, cell proliferation, as well as genomic instability, subsequently leads to malignant

transformation of breast epithelial cells. The occurrence of these multiple mutations implicates the exertion of distinct selective pressure during malignant transformation in familial breast cancer (Gasco *et al.*, 2003; Gretarsdottir *et al.*, 1998; Holstege *et al.*, 2009).

To date, the complexities of regulation of cellular proliferation remain elusive. An in-depth understanding of regulation of cellular proliferation upon first steps of carcinogenic transformation of human breast epithelial cells may identify novel targets for diagnostic and therapeutic purposes.

1.1.1.2 Regulation of protein synthesis in breast cancer

Protein synthesis is the fundamental mechanism essential for any living cell, and is involved in normal physiology and disease development. Protein synthesis is a multiple step process that depends on the coordinated action of hundred of proteins and different RNAs. The protein synthesis consists of three phases: initiation, elongation and termination. Eukaryotic elongation factor 1A (eEF1A) is a GTP-binding protein that plays a crucial role in translational elongation process. Deregulation of translational elongation contributes to the development of cancer (Edmonds *et al.*, 1996; Liu *et al.*, 2010).

There are two eEF1A isoforms, eEF1A1 and eEF1A2, that are expressed in a tissue-specific manner (Knudsen *et al.*, 1993; Lee *et al.*, 1992). During protein translation, following the codon/anticodon matching, GTPase activity of eEF1A catalyzes the binding of aa-tRNA to the A site of ribosome. The formation of peptide bond between the amino acid and the growing peptide chain is catalyzed by ribosomal peptidyl-transferase. Following hydrolysis of GTP, eEF1A-GDP leaves the ribosome. In order for additional translocation events to occur, the GDP must be exchanged for GTP, which is carried out by eEF-1βγ (Moldave, 1985) (Fig. 3).

eEF1A is a key regulator of various physiological processes, such as embryogenesis, aging, proliferation, apoptosis, protein degradation and cytoskeletal rearrangement (Condeelis, 1995; Kato *et al.*, 1997; Lamberti *et al.*, 2004). It is believed that the role of eEF1A in protein synthesis is of the key importance in these activities. However, non-canonical functions of eEF1A, such as interaction with cytoskeleton, may also be considered. Deregulation of eEF1A in these processes may contribute to development of various diseases, including cancer (Edmonds *et al.*, 1996; Liu *et al.*, 2010).

Several studies suggested the enhanced expression of eEF1A in human breast cancer cells (Chen and Madura, 2005; Pecorari *et al.*, 2009). It was also reported that eEF1A increased metastatic potential of breast adenocarcinoma in rat (Edmonds et al., 1996). Therefore, eEF1A is involved in the breast tumorigenesis.

Several reports implicated that post-translational modifications (PTMs) of eEF1A are associated with its regulatory function. As an example, phosphorylation of eEF1A1 is

involved in stimulation of GDP/GTP-exchange rate in rabbit reticulocytes (Peters *et al.*, 1995), and reduces binding to F-actin in rat liver cells (Izawa et al., 2000). Methylation of eEF1A was associated with the SV40-dependent transformation of mouse 3T3B cells (Coppard *et al.*, 1983). As PTMs play a significant role in a wide array of cellular processes, uncovering novel PTMs in eEF1A1 enable us to understand how eEF1A1 functions, and their contribution in tumorigenesis (Lin and Souchelnytskyi, 2010; Lin and Souchelnytskyi, 2011; Lin *et al.*, 2010).

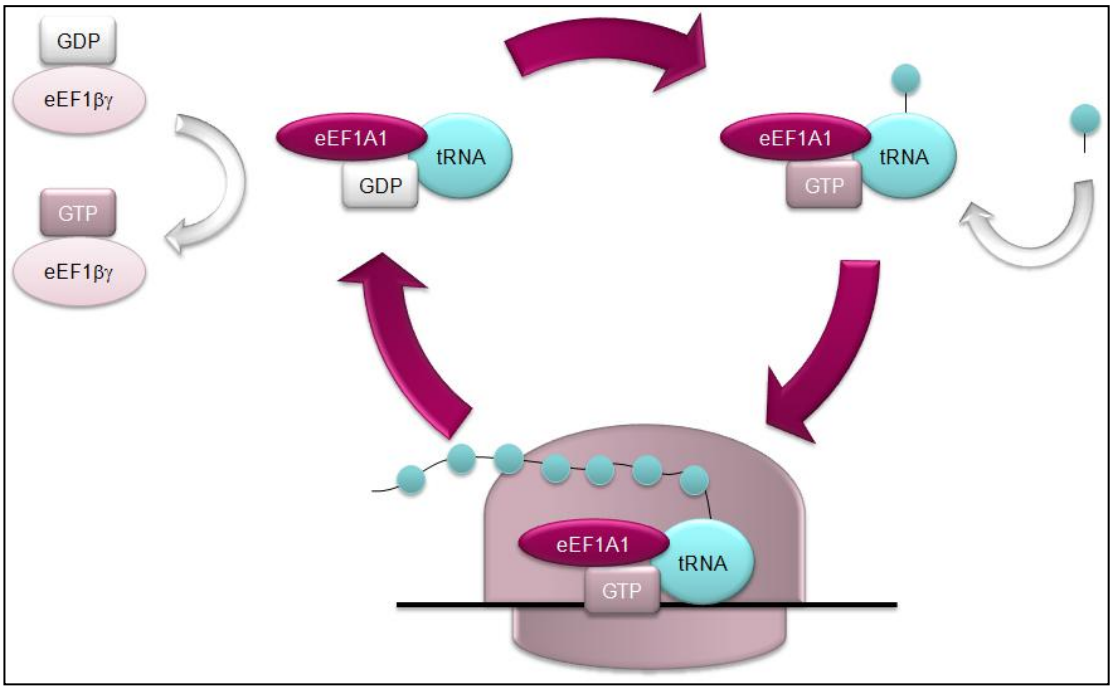

Figure 3. Elongation phase of protein synthesis. During protein translation, following the codon/anticodon matching, GTPase activity of eEF1A catalyzes the binding of aa-tRNA to the A site of ribosome. The formation of peptide bond between the amino acid and the growing peptide chain is catalyzed by ribosomal peptidyl-transferase. Following hydrolysis of GTP, eEF1A-GDP leaves the ribosome. In order for additional translocation events to occur, the GDP must be exchanged for GTP, which is carried out by eEF-1βγ.

1.1.2 Breast cancer invasion

Metastasis is the most common cause of death among the breast cancer patients. The understandings of mechanism of cancer metastasis enable us to develop the therapeutic strategies for targeting the metastatic cancer cells.

Metastasis is a complex biological process that involves a series of events, i.e. detachment of tumor cells from primary side, invading into surrounding stroma, enter the circulation (extravasation) directly or through lymphatic spread, arrested and exit from the circulation (intravasation) and eventually colonization in the distal organs (Fig. 4) (Gupta and Massague, 2006).

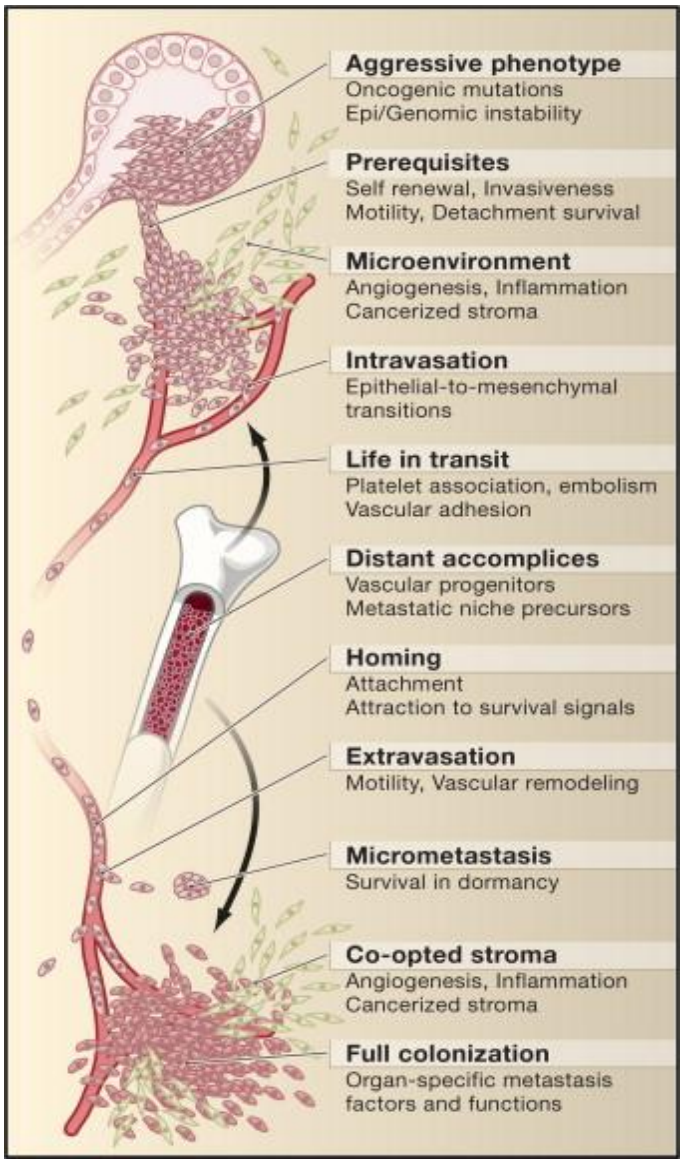

Figure 4. Model of cancer metastasis. Metastasis proceeds through the progressive acquisition of traits that allow malignant cells originating in one organ to disseminate and colonize a secondary site. This involves the detachment of primary tumor cells, invading into surrounding stroma, enter the circulation (extravasation), arrested and exit from the circulation (intravasation) and eventually colonization in the distal organs. Reprinted from Gupta and Massague, Cell 2006, with permission from Elsevier.

Most of the models of metastatic human breast epithelial cells have been developed by recovering cells from metastatic tumors formed in animals by originally non-metastatic cells. However, in these models there is no distinguishing between all the steps in cancer metastasis, e.g. no discrimination between invasiveness, extravasation, intravasation and colonization. Therefore, *in vitro* models reflecting specific steps of metastasis may provide crucial information on these specific steps, e.g. on detachment of cells from a primary tumor and invasion in stroma, passing in and from vessels through vessel walls, homing to specific organs and forming metastatic tumors in new tissues.

1.1.2.1 In vitro models of invasive breast cancer

Current in vitro models have been focused on use of highly invasive and metastatic cell lines, and genetically-modified cell lines. Among the most commonly studied cell lines, Hs578T, MDA-MB-231, BT-549, and MDA-MB-435S cells are classified as highly invasive cells, which are frequently used as *in vitro* model of breast cancer invasion (Zajchowski et al., 2001). Recently, an isogenic model of MCF10 has been developed for the study of malignant transformation and invasion. Invasive MCF-10CA1h cl2, MCF-10CA1a cl1, and MCF-10CA1d cl1 used as *in vitro* model of invasiveness, as compared to initial untransformed outgrowths cells (MCF10MS and MCF10A), hyperplastic cells (MCF10AT1, MCF10AT1kcl2), and cancerous cells (MCF10CA1h cl13) (Hurst *et al.*, 2009; Rhee *et al.*, 2008; Worsham *et al.*, 2006). However, these cells have also undergone passages in animals to select various clones, and therefore do not reflect strictly invasiveness but metastatic properties.

It was shown that many proteins play important roles in breast cancer invasion. Non-invasive cell lines transfected with these regulators also serve as *in vitro* model of breast cancer invasion. One of the examples is S100A4 model, in which the transfection with S100A4 breast cancer cell lines will lead to increased invasiveness (Jenkinson *et al.*, 2004).

More efforts are needed in order to understand the underlying mechanism of invasiveness, which may provide insight of both diagnostic and therapeutic importance. Special emphasis has to be on models which reflect well defined steps in metastasis.

1.1.2.2 Molecular mechanisms of cancer invasion

Breast cancer invasion is a complex biological process that involved a series of cellular events. These include an interaction of invasive cells with surrounding non-tumor stroma and cells, interaction between tumor cells, and changes in the tumor cell physiology.

Invasion of cancer cells into surrounding stroma represents a first barrier of cancer cells to metastasize. For decades, cancer invasion models have been centered on protease-driven mechanism (Liotta, 1986). A three-step model was developed to describe the biochemical events during cancer cell invasion of the extracellular matrix. The first step is adhesion, where cancer cell attach and bind to components of the extracellular matrix, such as laminin (for basement membrane) and fibronectin (for stroma). The second step is secretion of hydrolytic enzymes by malignant cells, or induction of host cells to secrete enzymes by anchored cancer cell, which can locally degrade matrix and attachment components. Lysis of the matrix most probably takes place in a highly localized region close to tumor cell surface. Some molecules, i.e. MMP-9, MMP-2, TIMP-1 and TIMP-2 are crucial in extracellular matrix and basement membrane degradation (Wolf et al., 2007). The third step is movement of cancer cell into region of

matrix modified by proteolysis. Continued invasion of extracellular matrix may take place by cyclic repetition of these steps. MMP regulate a broad spectrum of tumorigenic functions. Besides regulation of proteolytic activity in cancer invasion, MMP's also promote growth, angiogenesis, and migration of cancer cells (Fig. 5) (Rao, 2003).

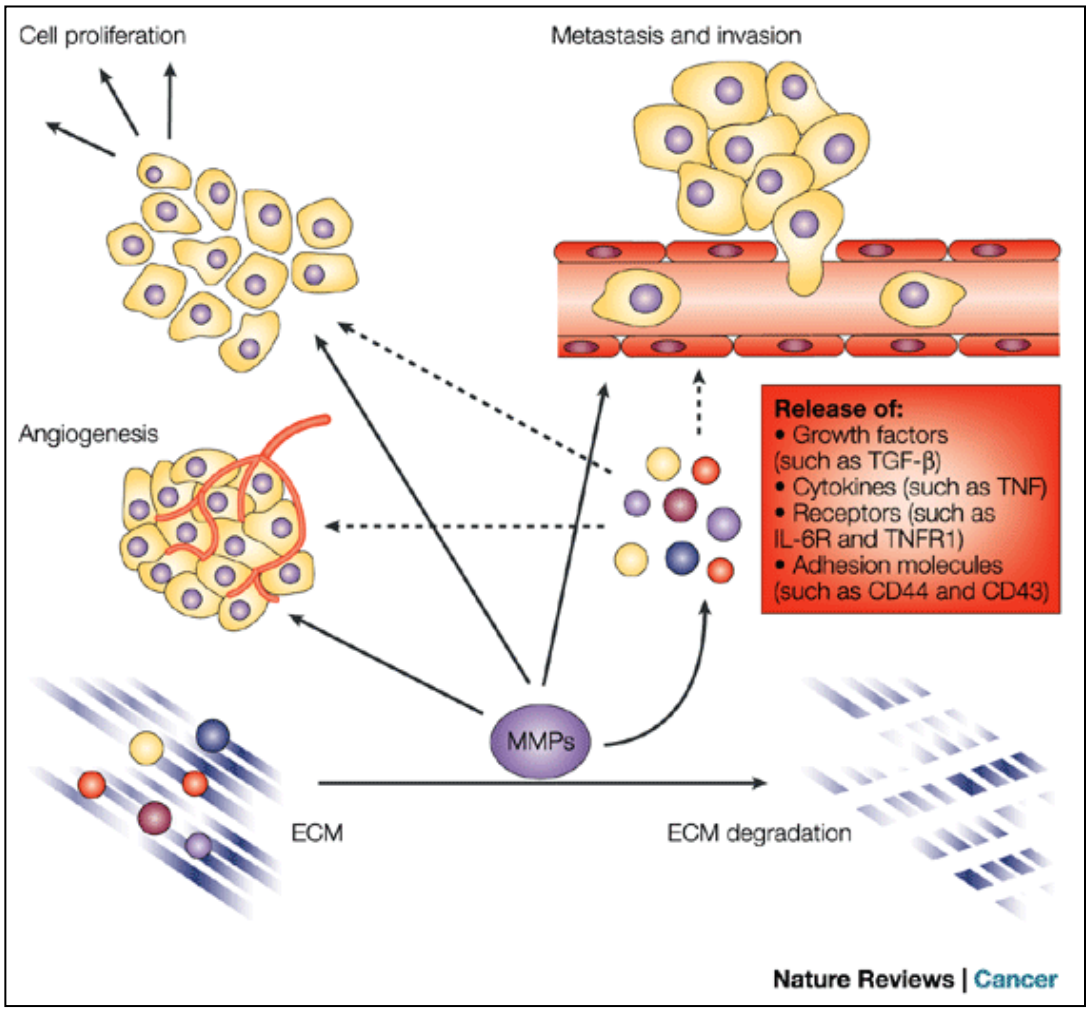

Figure 5. Matrix metalloproteinases are key regulators that promote tumor growth, angiogenesis, invasion and migration. Reprinted from Rao, Nature Reviews Cancer 2009, with permission from Nature Publishing Group.

Cell-cell interactions also play an important role in invasion of cancer cells. Homophilic epithelial cell-cell interaction could lead to activation of signaling to the cytoskeleton, E-cadherin and β-catenin. Interaction with fibroblast is shown to be essential in invasion of epithelial cells, as it is capable of extracellular matrix (ECM) remodeling (Gaggioli et al., 2007; Macpherson et al., 2007). Interaction with macrophages will activate Wnt 5a signaling critical for macrophage-induced invasion of MCF-7 (Pukrop et al., 2006).

Ability of epithelial-derived cancers to progress to an invasive, metastatic state correlates with a shift from adherent, epithelial shape to a motile, fibroblast-like

morphology, epithelial to mesenchymal transition (EMT) (Fig. 6) (Peinado *et al.*, 2007). In invasive cells, stable cell-cell and cell-ECM contacts are dissembled, actin cytoskeleton remodeled, and cells have an increased ability or migrate and degrade ECM (Thiery, 2002). Recent advances in DNA microarray technology have led to the isolation of several other metastasis-associated genes, many of the products of which are associated with the cellular cytoskeleton (Khanna et al., 2001) and its regulation (Clark *et al.*, 2000). Immunohistochemical studies reported correlation between E-cadherin loss and initiation and progression of tumors.

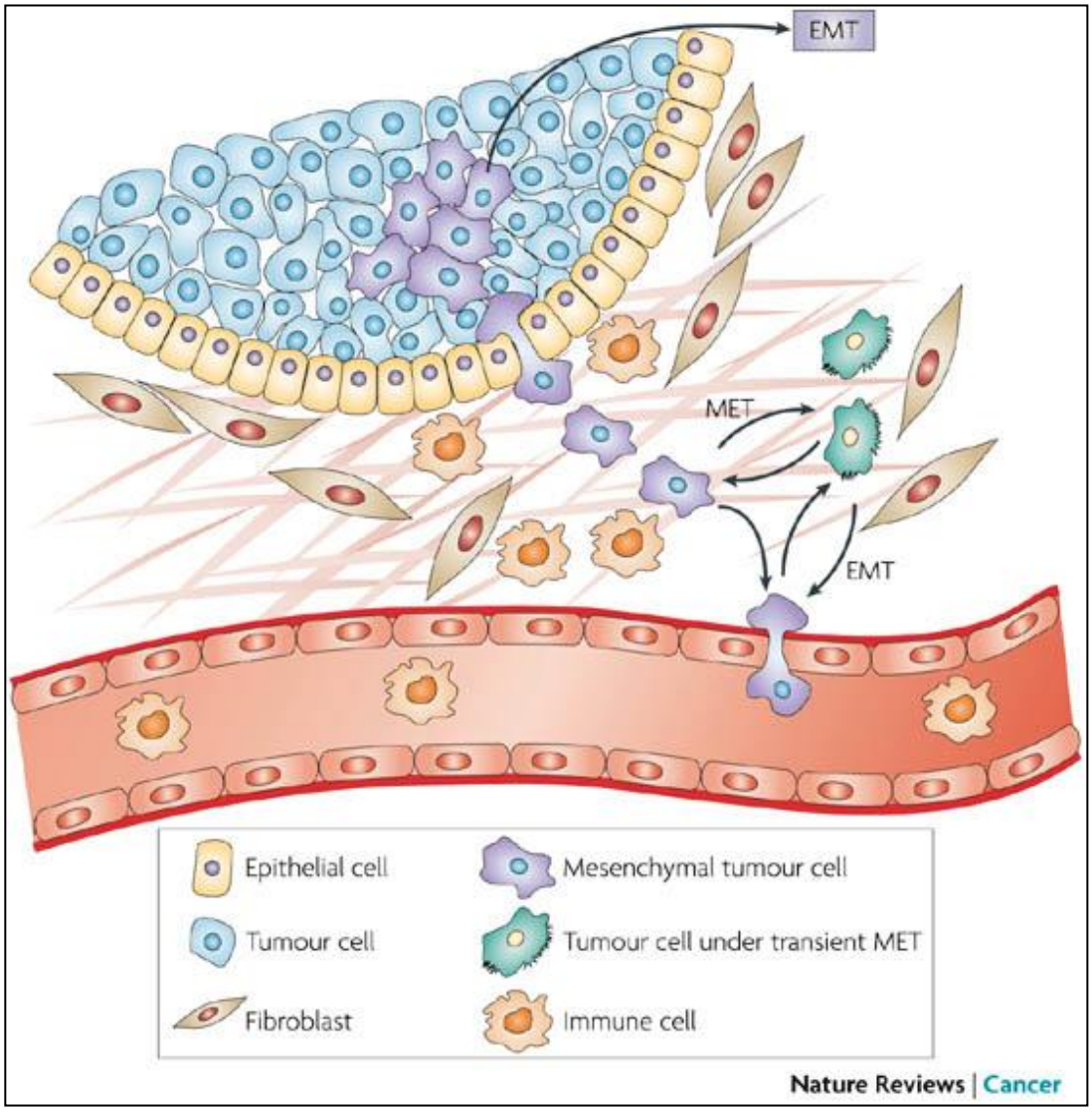

Figure 6. During EMT process, cancer cells are characterized by the loss of cell–cell adhesion and polarity accompanied by cytoskeleton rearrangements and increased cell motility. Reprinted from Peinado, Olmeda and Cano, Nature Reviews Cancer 2007, with permission from Nature Publishing Group.

Further studies are needed to reveal the detail mechanism of cancer invasion, its regulators are involved signaling pathways, which are crucial in the developing of drug targets to block the invasive-specific pathway. Transcriptomics studies have indicated that the number of metastasis-related regulators may be significantly higher than has

been reported biased one gene/one protein studies. As proteins are the main functional entities in cells, proteomics is expected to deliver a comprehensive description of all proteins involved in invasiveness mechanism.

1.1.2.3 BRMS1 in breast cancer invasion

Breast cancer metastasis suppressor 1 (BRMS1), a member of a growing metastasis suppressors family, has been reported to reduce breast and melanoma metastasis in vivo without affecting primary tumor growth (Hedley *et al.*, 2008; Hurst and Welch, 2011; Liu *et al.*, 2011). Recent progress in BRMS1 research leads us to the further insight in the molecular mechanism of BRMS1. It was suggested that BRMS1 can restore gap junctional intercellular communication of cells, as possible mechanism contributing to suppression of metastasis (Kapoor et al., 2004; Saunders et al., 2001). More recently, it was discovered that BRMS1 associates with large chromatin remodelling complexes such as SIN3:HDAC, which are the epigenetic regulators of gene expression. Besides, BRMS1 inhibits the activity of NF-κB (Cicek *et al.*, 2009; Li and Li, 2010). Despite the recent advance of our knowledge in BRMS1 role in cancer metastasis, the involvement of BRMS1 in other signalling pathway and regulation of invasiveness of breast cancer cells is relatively unexplored.

1.2 TGFβ SIGNALING

1.2.1 TGFβ signaling and breast cancer

Transforming growth factor-β (TGFβ) is a regulatory polypeptide which belongs to the TGFβ superfamily of secreted growth factors. Three isoforms of TGF-β, namely TGFβ1, TGFβ2, TGFβ3 were observed in mammalian cells (Massague *et al.*, 2000). TGFβ is secreted as biological inactive, latent form L-TGFβ, and formation of biological active TGFβ is the consequence of cleavage of the pro-form, and dissociation of active TGF-β from latency-conferring proteins, i.e. latency-associated peptide (LAP) and latent TGFβ-binding protein (LTBP).

The TGFβ signaling consists of Smad and non-Smad pathway. In Smad pathway, the TGFβ intracellular signalling is initiated by binding of the ligand to the heterotetrameric complex of TGFβ receptors type II and type I. Activated type II receptor kinase subsequently phosphorylated and activated type I receptor. The activated type I receptor phosphorylates receptor-activated Smads (R-Smads), i.e. Smad2 and Smad3. The activated R-Smads bind to a common-mediator Smad4 (Co-Smad) and this complex translocate into nucleus which leads to the activation or repression of transcription (Fig. 7) (Shi and Massague, 2003).

In addition to the canonical Smad-dependent TGF-β pathway, there are numbers of TGFβ signalling pathways that do not involve direct activation of Smad proteins (Derynck and Zhang, 2003; Mu *et al.*, 2011). These non-Smad pathways include

MAPK, Rho-like GTPase, and PI3K/Akt. These pathways are activated by receptor-interacting proteins, and also by non-Smad substrates of the receptor kinases. Among which ShcA and Par6 were shown to be directly phosphorylated by TGF-β receptor kinases (Lee *et al.*, 2007; Ozdamar *et al.*, 2005).

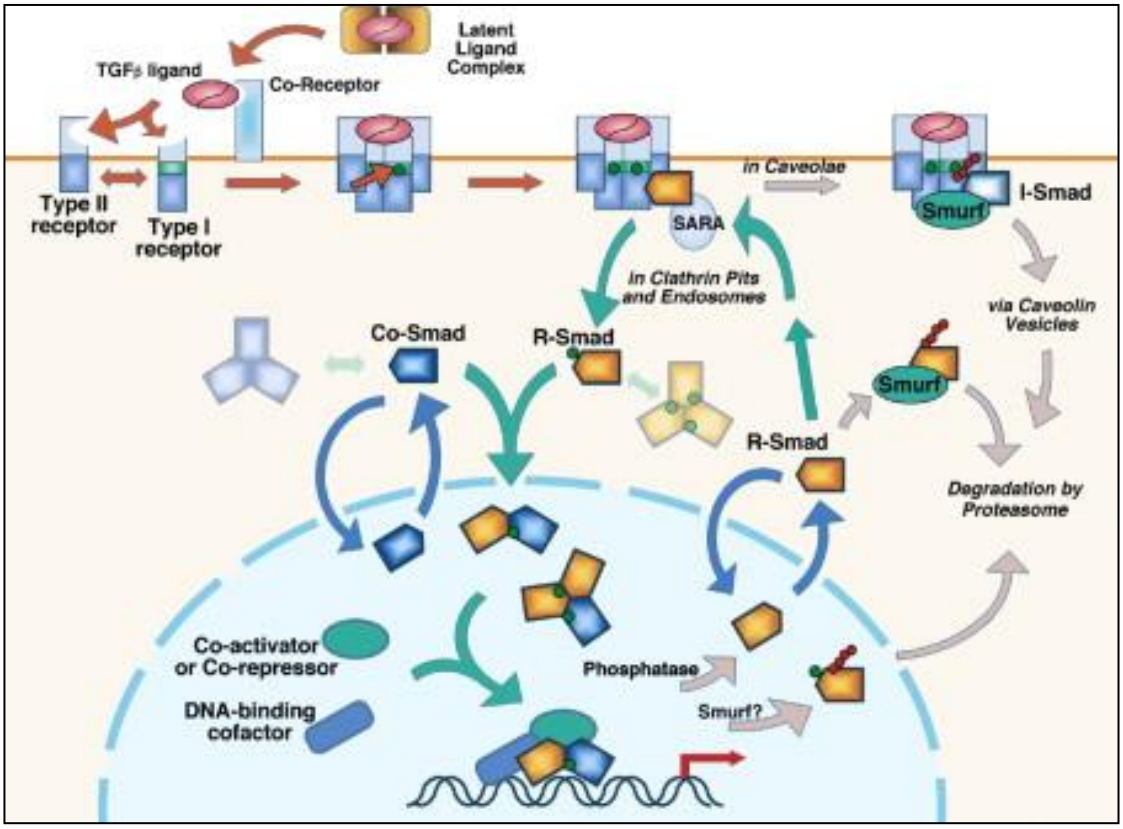

Figure 7. Smad-dependent TGF-β signaling. TGF-β intracellular signalling is initiated by binding of the ligand to the heterotetrameric complex of TGF-β receptors type II and type I. Activated type II receptor kinase subsequently phosphorylated and activated type I receptor. The activated type I receptor phosphorylates R-Smads. The activated R-Smads bind to a Co-Smad and this complex translocate into nucleus which leads to the activation or repression of transcription. Reprinted from Shi and Massague, Cell 2003, with permission from Elsevier.

TGFβ signaling plays important roles in a very broad range of cellular functions, i.e. proliferation, differentiation, migration, apoptosis, adhesion, angiogenesis, immune surveillance, survival, and ECM deposition. This diversity of TGFβ functions plays a role in TGFβ dual impact on tumor progression. In early stage of cancer, TGFβ plays a tumor suppressor role. In this stage, cancer cells are responsive to TGFβ, which will lead to growth inhibition, cell cycle arrest, and apoptosis. However, in advanced stage of cancer, TGFβ plays a tumor promotion role. The cancer cells may loss responsiveness to TGFβ and may acquire aberrant TGFβ signaling, followed by promotion of survival, proliferation, EMT and increased motility and invasiveness of the cells (Jakowlew, 2006).

In breast cancer progression, TGFβ have also been implicated in both tumor suppressor and tumor promotion roles. Overexpression of TGFβ1 in mouse breast tissue inhibited tumorigenesis, while interfering with it receptor promote tumorigenesis (Pierce et al., 1995). On the other hand, loss of TGFβ receptors are observed in advanced stage of human breast cancer (Gobbi et al., 2000). It was shown that expression of TGFβ associated with the progression of breast cancer. The increased expression of TGFβ1 was observed in late stage breast cancer, particularly associated with invasion and metastasis (Gorsch *et al.*, 1992; Walker and Dearing, 1992).

Due to the broad range of biological effects of TGFβ signaling in tumor progression, it has been extensively studied as a potential biomarker for diagnosis and prognosis and as a source of novel drug targets.

1.2.2 TGFβ in the regulation of cancer invasion

TGFβ play a pivotal role in invasive activity of tumor cells, either by affecting the stromal cells and microenvironment, or by affecting the tumor cells directly (Dumont and Arteaga, 2002). In the presence of TGFβ, ECM protein and protease secretion by fibroblast correlated with the invasiveness of breast cancer (Casey et al., 2008). On the other hand, direct affect of TGFβ on tumor cells can regulate their capacity of remodeling ECM by protease activities and EMT.

TGFβ regulates the protease activities of tumor cells through a complex interaction network of various signaling pathway. In human breast epithelial cells, MMP2 and MMP9 upregulated by TGFβ are dependent on p38-MAPK, but not ERK (Kim *et al.*, 2004). Upregulation of MMP expression, in turn leads to activation of L-TGFβ, and provides a positive feedback to the regulation of invasion and migration (Johansson *et al.*, 2000; Lin *et al.*, 2000). Smad7 inhibited invasion activity in breast tumor cell lines and mouse (Azuma et al., 2005). Inhibition of Smad signaling by Smad7 lead to inhibition of MMP1 and MMP13 production and subsequently invasion of the cells (Leivonen et al., 2006). HGF treatment increased cellular migration/invasion in response to TGFβ signaling through the JNK/pSmad3L pathway (Mori et al., 2004). The induction of TGFβ leads to upregulation of MMP and suppression of TIMP (Wick *et al.*, 2001).

TGFβ is a crucial regulator of EMT, which lead to invasive phenotype of the tumor cells. During carcinogenesis, tumor cells show EMT and become less sensitive to TGFβ-mediated growth inhibition while showing increased tumor invasion and metastasis (Oft *et al.*, 1996; Zavadil and Bottinger, 2005). Studies showed that TGFβ signaling interacts with other major protein hubs in regulation of invasion. Induction of β-integrin and Src expression by TGFβ is essential for induction of EMT and invasion in breast epithelial cells (Galliher and Schiemann, 2006). TGFβ activation of NF-κB lead to the induction of EMT and invasion of cancer cells (Neil and Schiemann, 2008). Earlier study showed that TGFβ activation of NF-κB mediated expression of EMT and

metastasis-related genes in response to TGFβ (Huber et al., 2004). Several studies showed that TGFβ and EGF crosstalk play a crucial role in invasion activity of tumor cells. Overexpression of erbB2 can alter cellular responses to TGF-β and lead to invasion and migration (Seton-Rogers et al., 2004; Ueda et al., 2004). TGFβ- and EGF-mediated invasion are probably linked to cytoskeleton rearrangement, however, further study is needed to elucidate the underlying mechanism of invasion.

Due to the complexity of biological effects of TGFβ in cancer invasion, it has been extensively studied as a potential biomarker for predicting the invasive and metastatic capacity of cancer.

1.3 PROTEOMICS AND SYSTEMS BIOLOGY

1.3.1 General concept

Personalized medicine is the most promising approaches in the treatment of various diseases, especially cancer. The use of appropriate biomarkers for personalized treatment has advantage over conventional therapeutics approach, as it confer maximum effectiveness with minimum side effect. Personalized treatment can be achieved by implementation of omic studies in clinical practices. Application of genomic, transcriptomic, proteomic and metabolomic studies deliver a vast amount of data that lead to the discovery of novel biomarkers for diagnostic, prognostic and therapeutic purposes. Therefore, further exploration in omic study could lead to the implementation of personalized medicine as a standard therapeutic scheme in the clinic (Lin *et al.*, 2011).

Proteomics is a global study of entire proteome of cell, tissue and organism in a particular condition and time point (Graves and Haystead, 2002). Proteomics is a very comprehensive discipline that includes the study of expression, function, localization, structure, modification of proteins, as well as protein-protein interaction (Graves and Haystead, 2002; Lim and Elenitoba-Johnson, 2004). A proteomics experiment generates vast amount of data that require further analysis, and systems biology is the main approach. Systems biology is an integrative science that studies the complex behavior of biological entities at the systems level (Kitano, 2002a; Kitano, 2002b). Integrating the proteomics data into systems biology language is an important approach in understanding the behavior of the complex organisms at various levels (Souchelnytskyi, 2005). In recent years, our knowledge of proteomics and systems biology is growing rapidly and create an excitement in scientific community because of its potential in novel biomarker and drug discovery (Duncan and Hunsucker, 2005).

Proteomics studies are highly dependent on the technology for protein separation and identification, and bioinformatics for data analysis. By protein separation techniques, gel-based and liquid chromatography (LC)-based approaches represent the primary stream in proteomics. In gel-based approach, that is, conventional 2D gel electrophoresis (2D-GE) and 2D differential gel electrophoresis (2D-DIGE), the

proteins are separated by their isoelectric point (pI) and molecular weight (Mw). In LC-based approach, the proteins or peptides are separated by using high performance liquid chromatography (HPLC) (Aebersold and Mann, 2003; Cravatt *et al.*, 2007). The identification and characterization of proteins or peptides by mass spectrometry are followed after separation (Kolker *et al.*, 2006). In recent years, antibody-based methods emerging as important approaches in proteomics. These approaches included the use of immunohistochemistry (IHC) on tissue microarrays (TMAs), pathway analysis using reverse phase protein arrays (RPPAs) and serum-based diagnostic assays using antibody arrays (Borrebaeck and Wingren, 2007; Brennan *et al.*, 2010; Wingren and Borrebaeck, 2004).

Ontological Classification		**Systems and Network Analysis**
Query Tools	**Visualization Tools**	Osprey (Breitkreutz et al., 2002)
GO-TermFinder (Boyle et al., 2004) AmiGO (Carbon et al., 2009) MatchMiner (Bussey et al., 2003)	GoMiner (Bussey et al., 2003) FatiGO (Al-Shahrour et al., 2004, 2007) Onto-Express (Draghici et al., 2003; Khatri et al., 2002) GOSurfer (Zhong et al., 2004) GOTM (Zhang et al., 2004)	BioLayout (Enright & Ouzounis, 2001) CellDesigner (Funahashi et al., 2007) Cytoscape (Kohl et al. 2011)

Table 1. List of bioinformatics tools that are commonly used for gel-based proteomics.

In the subsequent section, we focus our discussion on the various ways of translating gel-based proteomics data into systems biology using different bioinformatics approaches. Firstly, we will discuss the dataset from gel-based proteomics, including the acquisition of primary data and type of data for bioinformatics analysis. In the subsequent section, we will discuss the several way of analyzing the data acquired from gel-based proteomics, which included the ontological-based classification, systems and network analysis (Table 1). We will focus our discussion on the general concepts of the analysis, type of datasets used and bioinformatics software. Some examples of studies and future directions are presented for each approach (Lin *et al.*, 2011).

1.3.2 Dataset in gel-based proteomics

The general workflow of bioinformatics analysis of gel-based proteomics is shown in Figure 8. In gel-based proteomics, various types of datasets can be generated. There can be an annotated 2D gel, mass spectra, and list of identified proteins (Taylor et al., 2003). These dataset can be qualitative or quantitative. In this section, we focus on the analysis of 2 type of datasets generated from annotated 2D gel, i.e. global expression profile and differential expression profile.

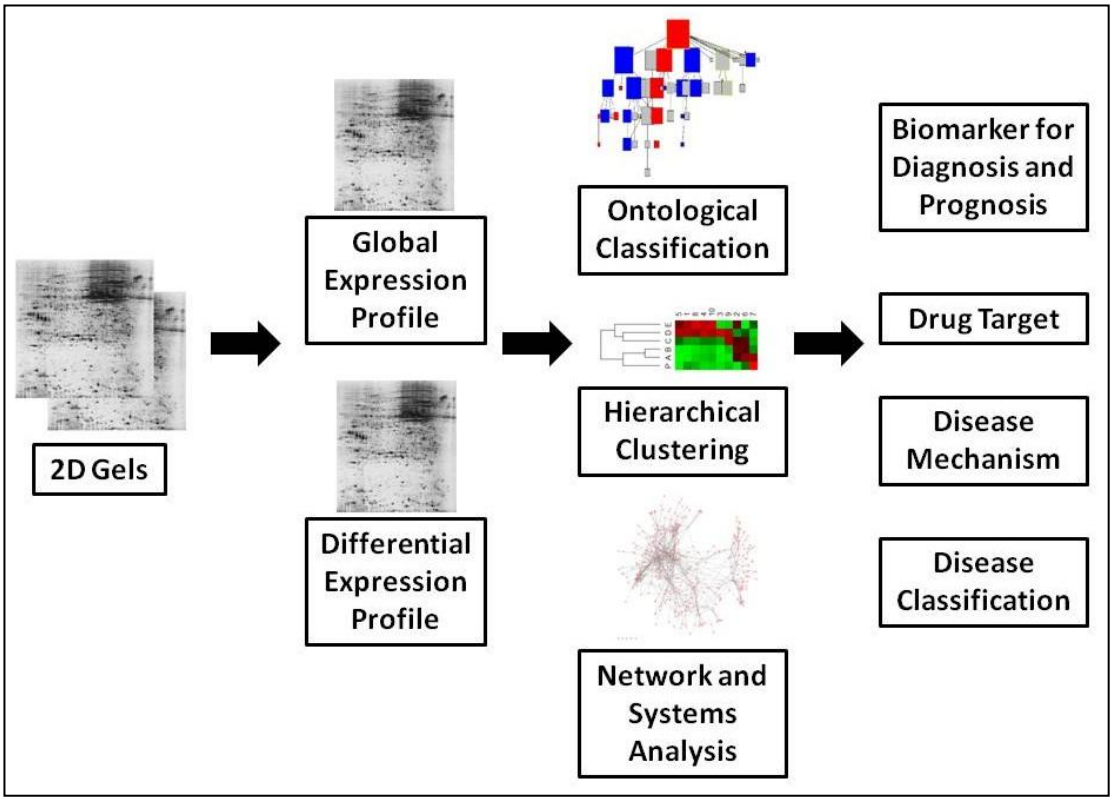

Figure 8. General workflow of bioinformatics analysis of gel-based proteomics. Once the 2D gels are generated, 2 type of dataset can be acquired from annotated gel, i.e. global expression and differential expression profiles. These datasets can be used for further analysis by various approaches, such as ontological classification, hierarchical clustering and systems/network analysis. These analysis approaches can improve our insight into particular biological questions, such as discovery of novel disease biomarkers for diagnosis and prognosis, drug target, study of disease mechanism and disease classification.

By identifying the protein spots on a 2D gel, a comprehensive, global protein expression profile can be generated. This approach can deliver a list of proteins expressed in a cell or tissue in a particular condition, which is exceptionally useful in understanding their biological characteristic. An example is a recent study on proteome profiling of breast epithelial cells with various proliferation potential. This study generate the most comprehensive 2D protein expression map with 183 proteins identified in 184A1 cells and 318 proteins identified in MCF10A cells, which lead to

the understanding of their biological properties and delivered a list of potential biomarkers of early event of tumorigenesis (Bhaskaran et al., 2009).

By identifying the protein spots in 2D gels that are different in their staining intensity in different conditions, a differential expression profile can be generated. Various biological questions can be addressed by differential expression analysis. The proteome changes upon drugs application can be studied by comparing the 2D gel of a particular cell treated or not with drugs. For example, cellular response to histone deacetylase inhibitor in colon cancer cells was evaluated by such approach (Milli et al., 2008). Besides, various disease stages can also be compared, for example, a list of proteins were identified to be differentially regulated between normal liver tissue and hepatocellular carcinoma (Corona et al., 2010). Furthermore, the dynamic changes of proteome can also be studied. Comparison of the differentially expressed proteins in the neuroblastoma grown in mice in different time interval revealed the proteome changes of the disease progression and effect of host-tumor interaction (Turner et al., 2009). Therefore, differential expression analysis of 2D gels often called comparative proteomics.

By applying various systems biology analysis tools, these proteomics dataset can further improve our insight into particular biological questions. The first objective of gel-based proteomics data mining is to search for protein of biological importance, such as diagnostic biomarker and potential drug target. By comparing two or more predefined biological conditions, we can precisely define the proteins of interest among thousands of spots in the 2D gel (Meunier et al., 2007). This can be achieved by using differential expression proteome profile, or by comparative analysis of two or more global protein expression profiles. The second objective of gel-based proteomics data mining is to use clustering approach to group or classify the proteins. This is important for understanding the complex biological systems, such as classification of tumor according to the expression of proteins, for the diagnostics and therapeutics purposes (Meunier et al., 2007). This approach can be realized applying the bioinformatics tools on both differential expression and global expression profile. In the subsequent section, we will discuss the analysis of gel-based proteomics dataset by using various approaches, and their biological significance.

1.3.3 Ontological classification

The postgenomic era has brought an exponential growth of biological databases. In recent years, researchers have begun to use unique identifiers to describe components of a database, and the relationship between them. The concept of unique identifiers forms the basis of ontology. Ontology can be described by a set of representative, unambiguous and non-redundant vocabulary or identifiers, which define classes, relations, functions, objects and theories (Gruber, 1993). It is not only represents an individual component but also its related components (Dimmer et al., 2008).

The Open Biomedical Ontology (OBO) consortium (http://www.obofoundry.org/) provides a resource where biomedical ontologies are presented in a standard format. Ontology-based approaches for data integration provide a platform of communication between researchers. It also allowed the retrieval/query of information across multiple resources and more efficient data mining and exploration. To gain the functional insight in a large-scale proteomics study, the traditional "literature mining" method is laborious and inefficient. Therefore, ontology-based approach is an effective solution.

In gel-based proteomics, the large dataset can be annotated and explored by application of Gene Ontology (GO) (http://www.geneontology.org/). Gene Ontology is a part of the OBO, which is the most widely used ontology in biomedical research community (Smith et al., 2007). The main objective of GO is to produce a controlled and unified vocabulary for genes and gene products, such as proteins, that can be applied to all organisms. Furthermore, classification of these components in defined groups or classes allowed us to gain the functional insight in the large-scale proteomics data.

GO annotation organizes genes or gene products into hierarchical order based on 3 categories: cellular component, biological process and molecular function (The Gene Ontology Consortium, 2000). Cellular component describe the localization of particular active gene products in the cell or its extracellular environment. It may be particular cellular structure, e.g. mitochondrion, Golgi apparatus; or gene products groups, e.g. proteosome, ribosome. Biological process describes the biochemical reaction of gene products in the cells. Examples of higher order categories are cell death, signal transduction. Examples of lower order categories are lipid metabolism, purine metabolism. Molecular function describes the elemental activities of gene products at molecular levels. Examples of higher order categories are enzyme, cytoskeletal regulator. Examples of lower order categories are glycine dehydrogenase, apoptosis activator. Since March 2007, 25,000 unique GO identifiers have been created, these provide researchers a broad set of descriptors for cellular component, biological process and molecular function for genes and their products (Dimmer et al., 2008).

There are various GO tools available for this task (Table 1). The complete list of tools can be found in http://www.geneontology.org/. These tools belong to either query tools or visualization tools. Prior to analysis, the genes or proteins have to be converted from generic or common name into the unique identifier, i.e. GO term, by using query tools. The most commonly used query tools are GO-TermFinder (Boyle et al., 2004), AmiGO (Carbon et al., 2009), and MatchMiner (Bussey et al., 2003). For example, the GO identifier for cyclin D3 is CCND3.

Once the list of GO identifiers are generated, visualization the data are carried out, using the tools such as GoMiner (Zeeberg et al., 2003), FatiGO (Al-Shahrour *et al.*, 2004; Al-Shahrour *et al.*, 2007), Onto-Express (Draghici *et al.*, 2003; Khatri *et al.*, 2002), GOSurfer (Zhong et al., 2004), and GOTM (Zhang *et al.*, 2004). These tools provide visualization of data in the form of either AmiGo view or Direct Acyclic Graph (DAG) view (Fig. 9). AmiGO view is in the form of expandable tree structures, and it

is linked to external databases, such as NCBI and CGAP. DAG is similar to hierarchies but differ in that a more specialized and narrower term or "child" can be related to more than one less specialized and broader term or "parent". Each term are represented by a node and they connected by path in hierarchical order. Each node can often be reached from multiple paths, which allow the comparison of genes/gene products involved in more than one molecular function or biological processes.

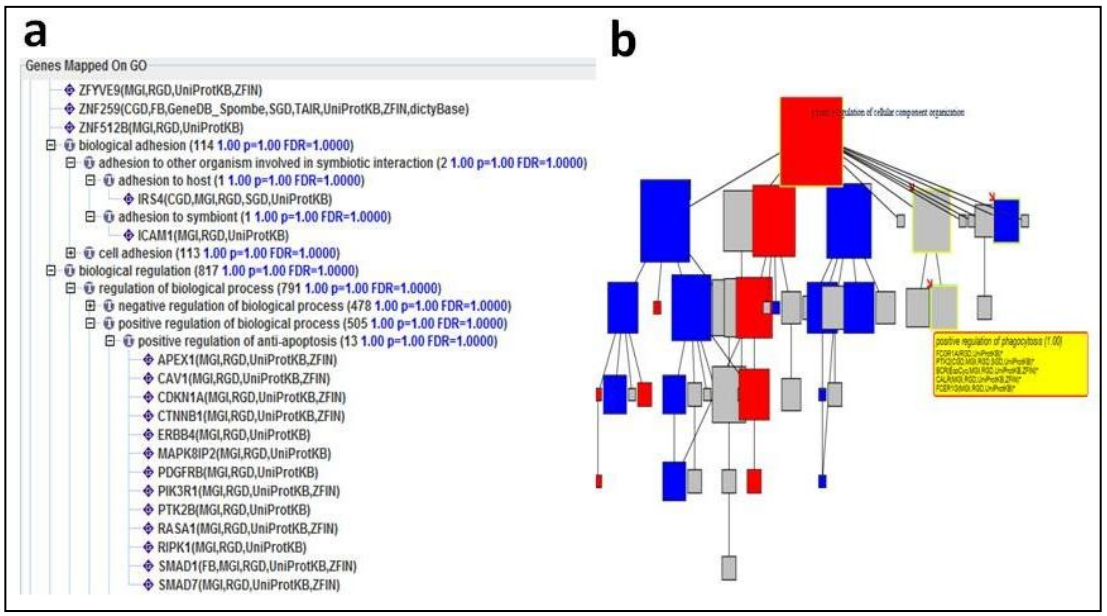

Figure 9. Data visualization of ontological-based classification. Gene Ontology tools, such as GoMiner showed in this figure, provides visualization of data in the form of either AmiGo view or Direct Acyclic Graph (DAG) view. (a) AmiGO view is in the form of expandable tree structures, and it is linked to external databases. (b) In DAG view, each GO term are represented by a node and they connected by path in hierarchical order. Each node can often be reached from multiple paths, which allow the comparison of genes/gene products involved in more than one category.

In gel-based proteomics, data generated from global expression and differential expression profiles can be used for ontological-based classification. Many studies suggested that ontological classification is a powerful tool in functional characterization of the cells in gel-based proteomics studies. For instance, a study from Alfonso et al. showed the use of ontological classification in a gel-based proteomics study to provide a functional insight of the colorectal cancer. In this study, 41 out of 52 analyzed proteins were unambiguously identified as being differentially expressed in colorectal cancer (Alfonso et al., 2005). An ontology analysis of these proteins revealed that they were mainly involved in regulation of transcription, cellular reorganization and cytoskeleton, cell communication and signal transduction, and protein synthesis and folding (Alfonso et al., 2005).

Although current proteomics study benefit from using Gene Ontology, the major drawback is that Gene Ontology does not describe and annotate the multiple forms of a

gene, such as alternative slicing, proteolytic cleavage and post-translational modification. Therefore, Gene Ontology cannot describe the functional stage of the gene products. In recent year, Protein Ontology (PRO) database has been created, which provide a formal classification of proteins (Natale et al., 2007; Natale et al., 2011; Reeves et al., 2008). The PRO included the classification of proteins based on the basis of evolutionary relationships and the structured representation of multiple protein forms of a gene. An initial attempt in applying PRO for the annotation of TGFβ signaling proteins showed that PRO provide a more accurate annotation and also facilitate various analysis, such as cross-species analysis, pathway analysis and disease modeling (Arighi et al., 2009). Despite of that, implementation of PRO in proteomics study is still in the infancy stage and there is no tools developed for the analysis of large-scale proteomics data. This implicates that further refinement and development of tools for PRO is needed in order to fill the gap.

1.3.4 Systems and network analysis

The behavior of a biological system, such as cells, is the consequence of complex interaction between their individual components, such as DNAs, proteins, metabolites, and other biological active molecules. In the past decades, signaling pathway has been the only approach to understand the interaction between these components. However, it is impossible to predict the behavior of biological systems solely from understanding of their individual component or single signaling pathway. Integration of signaling pathways into a higher order biological network is a very crucial approach for studying the complex behavior of a biological system. These can be achieved by implementation of systems and network analysis tools.

Over the past few years, application of system and network analysis in genomics and proteomics study had showed a great promise in understanding of complex behavior of biological systems. Global mapping of the cells or organelles using these tools enable us to discover, visualize and explore the behavior of the biological systems relevant to our experimental design. In addition, by studying the topological, functional, and dynamic properties of biological networks, the regulatory and control mechanism of the cells underlying the changes of environment can be explored. Examples are studies of the overexpression of certain signaling pathway of the tumor cells under the challenge with chemotherapeutics drug (Barabasi and Oltvai, 2004; Kwoh and Ng, 2007).

Networks are displayed as graphs, which represented by nodes and edges/links. Nodes are displayed in various shapes, which represent various types of molecules, such as genes, proteins, and metabolites. The nodes are connected with each other by the edges or links. Edges or links represent the biological relationships between the nodes, such as induction, activation, inhibition, post-translational modification, enzymatic-substrate reaction, and physical binding. Most of the biological networks are scale-free, in which most of the nodes have only a few links, while a few nodes with a very large number of links, which are called hubs (Barabasi and Oltvai, 2004).

The general principle of network construction is based on the known interaction pair of gene or protein. In brief, Swiss-Prot and GeneBank accession numbers from the experimental dataset are used to search against the external databases that contain information about the interaction between the genes or proteins. Subsequently, the genes or proteins from the experiment data were integrated and merged with their known interacting partners and pathways. This process is continued until all proteins of interest are included into the network.

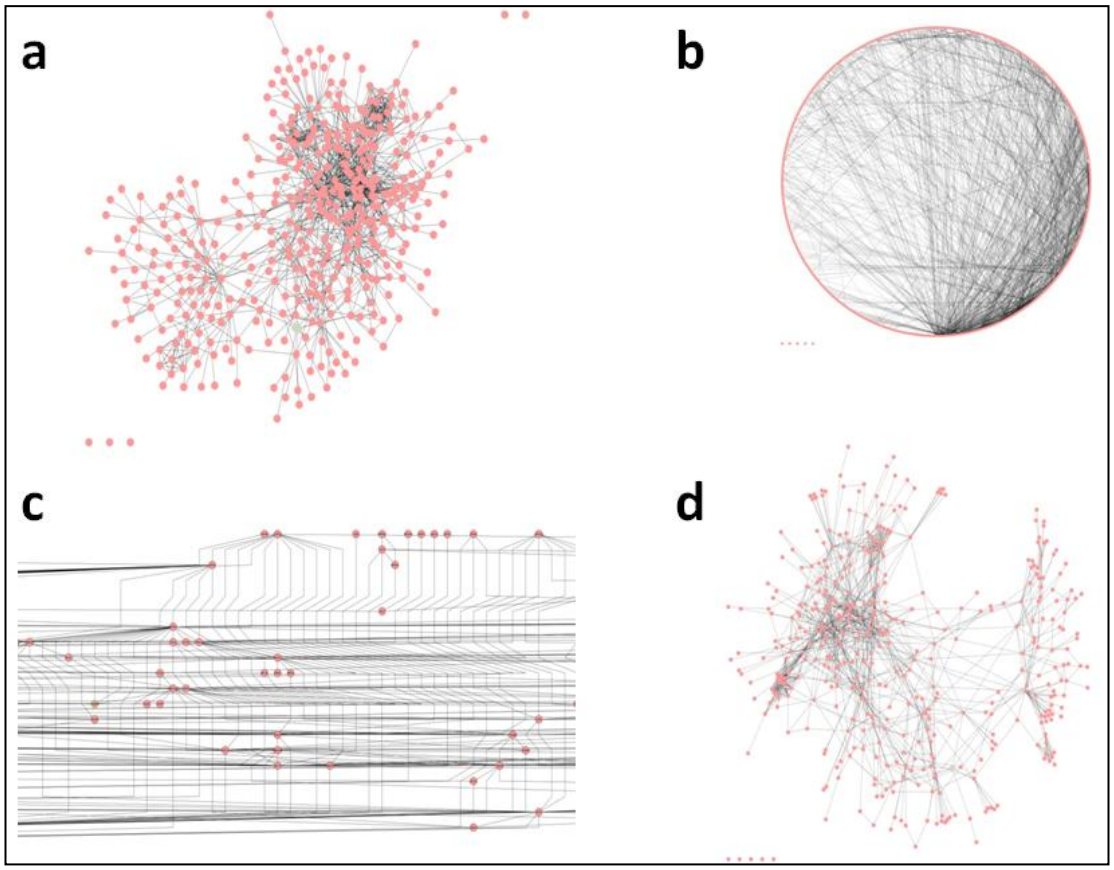

Figure 10. Visualization of network structure using Cytoscape. Networks are displayed as graphs, which represented by nodes and edges. For visualization of network structure, Cytoscape supports a variety of network layout algorithms, such as (a) force-directed layout, (b) circular layout, (c) hierarchical layout, and (d) spring-embedded layout.

There are a number of available tools for construction and analysis of networks (Thomas and Bonchev, 2010), such as Osprey (Breitkreutz *et al.*, 2002; Breitkreutz *et al.*, 2003), BioLayout (Enright and Ouzounis, 2001), CellDesigner (Funahashi *et al.*, 2007), and Cytoscape (Kohl *et al.*, 2011; Smoot *et al.*, 2011). Each tool has distinct functional features. Although most of these tools were initially designed for genomics data analysis, most of them are well adapted for proteomics data analysis. For gel-based proteomics, both global expression profile and differential expression profile can be used to construct the network, depending on the experimental design and question to be answered.

Here we show an example of workflow of network analysis in gel-based proteomics, by using Cytoscape. Cytoscape is open source software that provides basic functionality for integrating proteomics data on the network, editing and visualization of network, and also implementation of external plug-ins for network analysis. Data generated from gel-based proteomics, i.e. the list of the proteins, are integrated with the graph using tools for network construction, such as MiMi (Gao et al., 2009), cPath (Cerami *et al.*, 2006) and BioNetBuilder (Avila-Campillo *et al.*, 2007). Subsequently, using the annotation tools, the node and edge can be annotated with attribute and expression data, such as expression ratio obtained from 2D gel analysis. For visualization of network structure, Cytoscape supports a variety of network layout algorithms, such as spring-embedded layout, circular layout and hierarchical layout (Fig. 10). In order to reduce the complexity of a large network, user can selectively display the set of nodes and edges in the graph, using graph selection and filtering tools. Nodes and edges can be selected according to a wide variety of criteria, including selection by name or by the property of the attribute (Fig. 11) (Shannon et al., 2003).

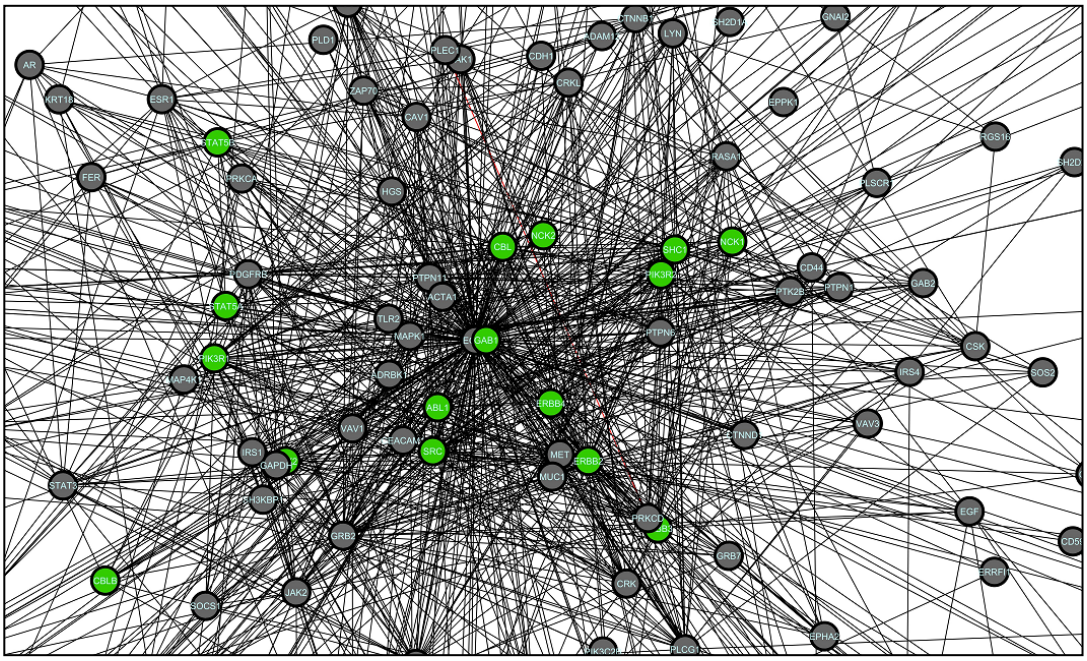

Figure 11. Graph selection tool in Cytoscape. User can use graph selection tool to reduce the complexity of the graph. In this example, the components of ERBB pathway were selected and colored (green) using selection tools.

When the network construction is complete, user can implement various external plug-ins for analysis of the network. This is one of the most powerful functionality of Cytoscape for solving biological questions by mean of network exploration. There is a variety of plug-ins which is commonly used in network analysis. Several examples of Cytoscape plug-ins for network analysis, such as MCODE (Bader and Hogue, 2003), NetworkAnalyzer (Assenov *et al.*, 2008) and Centiscape (Scardoni *et al.*, 2009), are discussed here. MCODE is a plug-in that search for clusters or highly interconnected regions in the network (Bader and Hogue, 2003). NetworkAnalyzer is a Java plug-in

that analyses and visualizes the molecular interaction networks (Assenov et al., 2008). These enable us to understanding the property of biological network, such as protein signaling network, protein-protein interaction network, that are of biological importance. Centiscape is another plug-in for analysis of complex topology of biological network (Scardoni et al., 2009). Centiscape computes centrality indexes of each node in the network, and relationship between the nodes. This may enable us to identify the critical nodes and regulatory circuits in the protein network.

In gel-based proteomics, network construction and pathway analysis are very useful in identifying novel regulatory mechanism of diseases and drug target discovery (Dudley and Butte, 2009). This was showed by a recent study that network analysis of proteomics data from clear cell renal cell carcinoma patient revealed the role of TNFα in clear cell renal cell carcinoma pathogenesis. In addition, it was suggested that clinically available TNFα inhibitors, such as thalidomide and etanercept can be used for the treatment of renal cell carcinoma (Perroud et al., 2006).

Network and pathway analysis is a robust approach in analyzing large proteomics dataset. However, network analysis is unbiased and hypothesis-free because the built of network are based on known interaction sets that recruited from published data. As a consequent, network analysis is not able to uncover the new or unknown pathway and interaction. Nevertheless, network analysis remains a powerful tool in understanding the gel-based proteomics data, and it can serve as a good starting point for a further exploration of the dataset.

1.4 PROTEOMICS AND SYSTEMS BIOLOGY IN BREAST CANCER RESEARCH

Tremendous efforts have been made during past decade in understanding the biology of normal and diseased cells at systemic level. In recent year, proteomics is one of the most common approaches in the study of breast cancer. Proteome profiling of a number of breast cancer cell lines has been reported, such as MCF7, HMEC, MCF10A, HB4a, HB2, 8701-BC, EM-G3, and 184A1 cells (Hamler et al., 2004; Jacobs et al., 2004; Page et al., 1999; Selicharova et al., 2007). Most of the breast cancer cell lines were used to study specific treatments, such as chemotherapeutic drugs (Souchelnytskyi, 2002). Besides, proteome profiling of various biological materials, including serum, plasma, tissue, nipple aspirate fluid, ductal lavage fluid and saliva has also been reported (Gast *et al.*, 2009).

Despite many proteins from proteome profiling study have been reported to possess significant diagnostic and prognostic value, only few of them have been characterized. Moreover, none of the identified biomarkers has been validated and investigated for their utility as breast cancer markers in clinical practices. The importance of proteomics analysis at systems level has been implicated (Souchelnytskyi, 2005). Combination of proteomics and systems biology is one of the most comprehensive approaches in

identifying the novel function and signaling pathway of the identified markers, which lead to the further development of biomarker for clinical application.

2 AIMS OF STUDY

The specific aims are:

- To identify set of proteins associated with acquiring of the high proliferation rate upon carcinogenic transformation of human breast epithelial cells by global expression proteome profiling approach.

- To establish an invasiveness model of isogenic human breast epithelial cells and to identify key regulators and prognostic signature of breast cancer invasion by differential expression proteome profiling approach.

- To identify and explore eEF1A1 as a novel substrate of TβR-I and the role of serine 300 phosphorylation of eEF1A by TβR-I in the regulation of protein synthesis, cell proliferation, and promotion of carcinogenesis.

- To study the role of serine 300 phosphorylation of eEF1A by TβR-I in the regulation of cell cycle regulation.

- To study the mechanism of BRMS1 in the invasiveness of breast cancer cells and the underlying signaling pathway.

3 RESULTS AND DISCUSSION

3.1 PAPER I

Comparative proteome profiling of MCF10A and 184A1 human breast epithelial cells emphasized involvement of cdk4 and cyclin D3 in cell proliferation

To gain insights into molecular mechanisms which may explain how proliferation rates of cells are defined, we performed comprehensive proteome profiling of 184A1 and MCF10A cells. 184A1 and MCF10A cells are often used in breast cancer research (Ethier, 1996; Hondermarck *et al.*, 2001; Lacroix and Leclercq, 2004; Souchelnytskyi, 2002). These two cell lines were established from normal human breast epithelial cells and non-disease tissue, and they have showed significant similarities. Notably, 184A1 and MCF10A cells express wild-type p53 and pRb, show enhanced telomerase activity, have near-diploid karyotypes with minimal rearrangements, negative ER status, and express keratins 14, 8 and 18 (Ethier, 1996; Lacroix and Leclercq, 2004; Stampfer and Yaswen, 2000). The important feature of 184A1 and MCF10A cells is that they both are immortalized, but not tumorigenic. These characteristics put these cells at the step of early changes on the scale of tumorigenic transformation. However, evaluation of proliferation rates confirmed that MCF10A cells proliferate faster, as compared to 184A1 cells, and as measured by [^{3}H]thymidine incorporation assay. This implicates that MCF10A cells are more advanced in transformation process, as compared to 184A1 cells.

In this paper, we described the most comprehensive proteome expression map of 184A1 and MCF10A cells. We performed proteome profiling of 184A1 and MCF10A cells. By systemic identification of all proteins detected in 2D gels using MALDI-TOF mass spectrometry, we identified 183 proteins in 184A1 cells and 318 proteins in MCF10A cells. In combination with other reports, our data contribute to annotation of proteomes of human breast epithelial cells. As an example, the overall protein expression maps of the normal primary luminal and myoepithelial cells were similar to those of 184A1 and MCF10A (Page *et al.*, 1999).

Functional clusters were defined by GoMiner, which is a tool used to identify biological processes and functions represented by proteins and genes of interest (http://discover.nci.nih.gov/gominer/). As it was expected, analysis of functional domains represented by identified proteins confirmed significant similarities of 184A1 and MCF10A cells, despite differences in protocols for establishing of 184A1 and MCF10A cells.

To explore relations between identified proteins in 184A1 and MCF10A cells, we built large-scale networks in which identified proteins formed connections between

themselves. The networks allowed incorporation of other proteins which are in direct physical or functional relations with the proteins identified by us. Such approach partially compensated for the limitations of identification of only a part of proteome, and introduced into the analysis proteins and genes which have not been identified or detected in 2-D gels. Networks which were built with all identified proteins showed complex character for both 184A1 and MCF10A cells. Both networks had features of a scale-free network, although highly connected components were observed. These networks showed similarities in regulatory mechanisms unveiled by identified proteins and their analysis. Similarities included involvement of TNF, AKT, F2, and IGF hubs. The second important observation is the higher expression in MCF10A cells proteins directly involved in cell cycle regulation and mitogenic signaling, as compared to 184A1 cells. Such enhanced expressions of proliferation-dependent and growth-promoting proteins are likely to explain why MCF10A cells have higher proliferation rate, as compared to 184A1 cells.

Among proteins of interest, we focused on cdk4 and cyclin D3 which are known as potent direct regulators of the cell cycle (Jiang *et al.*, 2007; Yu *et al.*, 2006). Cyclin D3 can bind cdk4, and the resulting complex can phosphorylate pRb. The level of D-cyclins, including D3, is regulated by extracellular stimuli, and therefore cyclin D3 can be considered as a mediator of signals to the core cell cycle machinery. Expression of cyclin D3 is required for progression of the cell cycle in human breast cancer cells (Jiang *et al.*, 2007). In human breast tumors high levels of expression of cdk4 and cyclin D3 were observed, as compared to normal breast tissues (Ito *et al.*, 1996; Montanaro *et al.*, 2007). Cdk4 kinase activity was also found to be required for breast tumorigenesis (Yu *et al.*, 2006), and involvement of cdk4 and cyclin D3 in a number of other cancers was reported, e.g. in melanoma, prostate cancer and acute myeloid leukemia. These reports of cdk4 and cyclin D3 role in regulation of the cell proliferation and their expression in breast tumors and normal tissues suggested that our data are of relevance to the observations with clinical samples. Our data showed that higher expression of cdk4 and cyclin D3 in MCF10A cells may promote higher proliferation rate, as compared to 184A1 cells, and as was validated by using siRNA approach. Thus, our results suggest that enhanced expression of cdk4 and cyclin D3 may be among the early events in the breast cancer progression.

A comprehensive proteome profiling of human breast epithelial cells is required for understanding of carcinogenesis. Reported here proteome datasets of 184A1 and MCF10A cells will be of importance not only for cancer studies in vitro, but also for an extrapolation of results obtained with cultured cells to proteome changes in tumors from patients.

3.2 PAPER II

Proteome signature of invasiveness of human breast epithelial cells

Metastases are the main cause of lethality in breast cancer. Invasiveness of malignant cells from the site of a primary tumor is the first step of metastasis. Knowledge of mechanisms of invasiveness is essential for development of novel treatments and novel prediction markers (Mego *et al.*, 2010; Valastyan and Weinberg, 2011). However, current in vitro models have been focused on use of highly invasive and metastatic cell lines, and genetically-modified cell lines, these cells do not reflect strictly invasiveness properties. In this study, we report a development of an isogenic model of invasiveness of human breast epithelial cells and its proteome profiling. This allows us to identify key regulators of breast tumour invasiveness and may be explored as a prognostic signature of invasiveness.

To generate isogenic model of invasiveness of human breast epithelial cells, we use non-invasive MCF7 cells. This was performed using collagen invasion assay. Collagen invasion assay was used to collect more than 50 clones of MCF7 cells which showed enhanced invasiveness. Two cycles of selection were performed, with each cycle consisting of collection of the collagen-invading cells and their expansion. Selected invasive cell clones were further subjected to a rigorous validation using membrane invasiveness assay to confirm their invasive phenotype. For evaluation of invasiveness, membranes were coated with matrigel. For proteomics study, we selected a clone MCF7c46, as this clone was among the most invasive and stable in maintaining of its invasive phenotype over the long-term culturing.

2D gels were generated for MCFc46, MCF7 and MDA-MB-231 cells. For each cell line, 2000 protein spots were detected in a 2D gel in average. The overall patterns of protein separations in 2D gels of tested cells were similar, which is in line with similarity of origin of the cells, i.e. from breast epithelium. Gel image analysis was performed to detect protein spots differentially expressed between MCF7c46 and MCF7, MCF7 and MDA-MB-231, and MCF7c46 and MDA-MB-231. These three combinations allowed extraction of proteins which are invasiveness-specific, and exclude proteins which changed their expression due to differences between the cell lines and due to invasiveness-unrelated changes upon selection of the MCF7c46 clone.

We identified 150 proteins which change their expression upon acquisition of invasive phenotype by MCF7 cells (parental MCF7 vs invasive MCF7c46). We identified also 302 proteins with different expression in invasive MCF7c46 and metastatic MDA-MB-231, and 279 proteins with different expression in non-invasive parental MCF7 and MDA-MB-231 cells.

We validated expression of HNF4α and BRMS1 in MCF7, MCF7c46 and MDA-MB-231 cells. For BRMS1, proteome profiling data showed that expression is enhanced in MCF7, as compared to MCF7c46 and MDA-MB-231. To validate this observation, we

perform immunoblotting of total cell extract with specific antibodies. We observed that BRMS1 was expressed at higher levels in MCF7 cells, as compared to MCF7c46 and MDA-MB-231 cells. Correlations of proteomics data and immunoblotting results were also observed for HNF4α protein. These results support observed changes in the proteome profiling.

The analysis of identified proteins was performed to extract only proteins which would be related to acquisition of invasive phenotype. Notably, difference between MCF7c46 and MDA-MB-231 was considered as cell type-related. Difference between MCF7 and MDA-MB-231 was considered as a combination of the cell type-, invasiveness- and metastasis-related. Difference between MCF7 and MCF7c46 were considered as invasiveness-related if they were present in MCF7 vs MDA-MB-231 and absent in MCF7c46 vs MDA-MB-231 comparisons. These combinations showed that the invasiveness signature contains 84 proteins. This analysis was further elaborated by analysis of networks formed by the identified proteins, as described below.

Functional clustering of identified proteins showed that overall representation of domains in all 3 datasets was similar, with cellular metabolism being the largest affected functional domain. However, a portion of differentially expressed proteins involved in ECM organization and biogenesis were found to represent non-invasive versus invasive/metastatic subsets.

To extract invasive-specific network, dependencies common for the MCF7/MCF7c46 and MCF7/MDA-MB-231 were extracted from these two networks. This common network was then compared with the network formed by MCF7c46/MDA-MB-231 differentially expressed proteins. The MCF7c46/MDA-MB-231 network represents proteins different due to the cell line origin and metastasis, and probably not related to acquiring invasiveness only. Dependencies related to the cell line differences were subtracted from the invasive-specific network. The final network represents invasive-specific dependencies with 924 nodes of the whole network, as evaluated by the Cytoscape-built network. Top functions represented by the invasiveness-related network are regulation of metabolism, cell cycle, proliferation, cell death and cell motion. Top sub-networks represented these cellular functions, and defined sets of specific nodes for each functions. As an example, cyclin-dependent kinases and estrogen receptor-α were among nodes involved in regulation of cell proliferation. TGFβ, EGF and FGF were also strongly represented as regulators of cell functions affected by acquisition of invasiveness phenotype.

Topology analysis showed that the invasive-specific network has features of a scale-free network, with close to rectangular hyperbolic distribution of frequency to the number of shared neighbors. This suggests that the regulation of invasiveness is rather a robust system. The nodes of high importance for robustness of a scale-free network are often the nodes of highest weight in the network. These nodes represent key regulators of the studied process, as they have strongest impact on the system stability and response to perturbation. In the context of invasiveness, we expect that the key nodes of

the invasiveness network would be crucial regulators of invasiveness. Analysis of the degree of connectivity of nodes identified EGFRB, TAF1, HNF4α, MYC and RB1 as potential key regulators. An impact of other nodes, e.g. TGFβ, was estimated as significant due to higher betweenness as expected to a node with such connectivity.

TGFβ and EGF were two growth factors that appeared as key nodes and in the top of functional domains. TGFβ and EGF are also known to affect breast tumorigenesis, and their signaling mechanisms are targeted in treatment of breast cancer. Therefore we explored how TGFβ and EGF may cooperate in regulation of invasiveness. The TGFβ and EGF-centered network showed 2-levels with more than 30 highly connected nodes which may define how TGFβ and EGF may cooperate. Among these nodes were identified insulin, VEGF, HNF4α and NFκB. Studies of diabetes impact on breast cancer output have shown that diabetes, and therefore deregulated insulin signaling, correlated with worse prognosis and aggressive development of the disease. The rate of death from breast cancer was enhanced from 20 % to 50 % for patients who suffered also from diabetes. Our study indicates that the insulin signaling disturbance may interfere with the invasiveness-related network signaling.

In conclusion, we generated MCF7c46, an isogenic model of breast cancer invasion, which allowed the study of cancer invasion activity as a distinct step in metastasis. Proteome profiling and systemic analysis showed relations of TGFβ, EGF, insulin, VEGF, HNF4α and NFκB in regulation of invasiveness. This indicates that insulin signaling disturbance may interfere with the invasive network, thus explained the increased risk of breast cancer metastasis in diabetes patients.

3.3 PAPER III

Phosphorylation of eukaryotic elongation factor eEF1A at Ser300 by type I transforming growth factor-β-receptor (TβR-I) results in inhibition of protein translation

More than 100 proteins associated with TGFβ receptor complex were identified, and many of these interactions are dynamically dependent on the activation of the receptor complex. These suggested that, in addition Smad2 and Smad3, and Smad-independent signalling substrate (i.e. ShcA, Par6), TGFβ receptors might have more substrates to be uncovered. Since the kinetics of phosphorylation is very rapid by its nature, and the interaction between a kinase and its substrate is transient, identification of substrates of TGFβ receptors remains a daunting task. Undoubtedly, characterization of the novel substrates might reveal novel molecular mechanisms of TGFβ signalling.

To identify novel substrate of TβR-I, we screened a human fetal lung λgt11 phage library with purified TβR-I containing constitutively active kinase. The screening identified eEF1A1 as a novel substrate that is directly phosphorylated by TβR-I containing constitutively active kinase. Phosphopeptide mapping and phosphoamino acid analysis identified Ser300 as the site of phosphorylation. We observed TGF-β1-dependent phosphorylation of eEF1A1 upon transfection of wild-type TβR-I, strong phosphorylation upon transfection of constitutively active TβR-I, and no phosphorylation upon transfection of kinase-inactive TβR-I. SB431542 strongly inhibited Ser300 phosphorylation, further confirm the involvement of TβR-I in phosphorylation of eEF1A1.

We showed that the phosphorylation of eEF1A1 at Ser300 by TβR-I is a mechanism of a direct inhibitory effect of TGFβ on protein synthesis, cell proliferation, anchorage-dependent cell growth and anchorage-independent cell growth. We found that the phosphorylation of Ser300 is also altered in human breast cancer. Enhanced expression of eEF1A was observed in human cancers, without discriminating its isoforms1 and 2, and eEF1A was found to promote carcinogenic transformation of epithelial cells and fibroblasts (Anand *et al.*, 2002; Cans *et al.*, 2003; Chen and Madura, 2005; Joseph *et al.*, 2002; Taniguchi *et al.*, 1991; Tatsuka *et al.*, 1992). eEF1A is highly expressed in metastatic cell lines (Taniguchi *et al.*, 1991; Tatsuka *et al.*, 1992), tumorigenesis. TGFβ signaling also has a strong impact on tumorigenesis (Feng and Derynck, 2005; Massague, 2008; Miyazono *et al.*, 2003). Our results show that eEF1A1 is a convergence point of TGFβ signaling and regulation of protein synthesis, with the direct phosphorylation of eEF1A1 by TβR-I as a triggering mechanism of this crosstalk.

eEF1A1 has a number of functions in mammalian cells (Liu *et al.*, 2002). Location of Ser300 in domain II of eEF1A pointed to the possible interference of Ser300

phosphorylation with aa-tRNA binding, which was confirmed experimentally. This is in agreement with the potent growth-inhibitory activity of TGFβ and may be the mechanism that prepares cells for the lower rate of proliferation by inhibiting protein synthesis. Our observations indicate that TGFβ has an impact on protein synthesis via targeting eEF1A1. This pathway does not require transcriptional activation of genes and may function in cooperation with the other pathways of TGFβ, e.g., the Smad-dependent pathway (Feng and Derynck, 2005; Massague, 2008; Miyazono *et al.*, 2003). An indirect effect of TGFβ on the cell cycle via activation of dephosphorylation of p70S6K in EpH4 cells has been reported (Petritsch *et al.*, 2000), although in smooth muscle cells, TGFβ had no effect on p70S6K activity (Krymskaya *et al.*, 1997). An involvement of the mTOR pathway in TGFβ-dependent regulation of cell size, and therefore protein synthesis, via S6K1 and eIF4E binding protein 1 has also been reported (Lamouille and Derynck, 2007). Two proteins known to be involved in protein synthesis, TRIP-1 (a component of the eIF3 complex) and eIF2a, have also been reported as modulators of TGF-β signaling, but not in the context of regulation of mRNA translation (Chen *et al.*, 1995; Choy and Derynck, 1998; McGonigle *et al.*, 2002). Multiplicity of TGF-β initiated pathways is a feature that ensures robustness of TGFβ action. Smad2 and Smad3 phosphorylation by TβR-I is observed as early as after 20 min of treatment with TGFβ, followed by initiation of transcriptional responses (Feng and Derynck, 2005; Massague, 2008; Miyazono *et al.*, 2003). eEF1A1 phosphorylation at Ser300 contributes to the inhibition of cell proliferation predominantly at the late stage of TGFβ action, after 4 to 6 hr of cell treatment with TGFβ.

Furthermore, we also observed decreased Ser300 phosphorylation in human breast cancer, as compare to normal breast tissue, indicates that this mechanism may be implicated in tumorigenesis, because decreased phosphorylation at Ser300 correlates with the higher proliferation rate of cells. Thus, identification of eEF1A1 as a substrate of TβR-I unveils a novel translation-related regulatory pathway downstream of TβR-I.

3.4 PAPER IV

Eukaryotic elongation factor eEF1A1 promotes and Ser300 mutants of eEF1A1 inhibit transition through the S and G2/M phases of the cell cycle

In this paper, we described that the enhanced expression of the wild-type (WT) eEF1A1 promoted proliferation of MCF-7 cells. Mutations of the Ser300 residue in eEF1A1 result in inhibition of aa-tRNA loading onto eEF1A1, and subsequently in inhibition of protein synthesis. Expression of the Ser300Ala or Ser300Glu mutants of eEF1A1 decreased proliferation of MCF-7 cells.

TGFβ1 is known to have a direct effect on the cell cycle by regulating the activity of CDKs, CDK inhibitors and cyclins (Feng and Derynck, 2005; Massague, 2008; Miyazono *et al.*, 2003). When the MCF-7 cells were treated with TGFβ1, the [^{3}H]thymidine incorporation was dose-dependently inhibited in the WT eEF1A1 expressing cells and in the empty vector expressing control cells. No TGFβ1 responsiveness was observed in cells expressing Ser300 mutants of eEF1A1. This is expected, as the mutants cannot be phosphorylated by TβR-I (Lin *et al.*, 2010). [^{3}H]thymidine incorporation test measures synthesis of the genomic DNA and provides a readout of how fast the cell cycle is.

To examine the effect of eEF1A1 on the regulation of the cell cycle, we analyzed progression of the cells transfected with WT or various mutants of eEF1A1 through G0/G1, S, and G2/M phases. Using FACS analysis, we monitored distribution of cells in the various phases of the cell cycle. FACS results showed that the WT eEF1A1 promoted transition of MCF-7 cells through the S- and G2/M-phases and accumulation in G0/G1 phase. Abrogation of the binding of aa-tRNA, and therefore inhibition of protein synthesis, by mutating Ser300 in eEF1A1, resulted in slower transition of the S-phase, as compared to the WT eEF1A1. Upon treatment with TGFβ1, eEF1A1-WT decreased the accumulation of cells in G0/G1 phase. Abrogation of Ser300 (S300A) and mimic phosphorylation at Ser300 (S300E) of eEF1A1 inhibited TGFβ-dependent accumulation of cells in G0/G1 phase.

In conclusion, our study suggested that overexpression of eEF1A1contributed to the increased proliferation of cells by promoting transition of cells through the S- and G2/M-phases of the cell cycle. aa-tRNA binding-impaired mutants of eEF1A1, especially Ser300Glu, showed the opposite effect. This indicates that the main contribution of eEF1A1 to the cell cycle regulation is by the promotion of the transition through the cell cycle. Future study will focus on investigating the detail signaling pathway of eEF1A1 and the Ser300 mutation in regulation of cell cycle.

3.5 PAPER V

TGFβ and EGF coordinately Modulate BRMS1-mediated Breast Cancer Invasion by Activation of Focal Adhesion and Cytoskeletal Rearrangement

Breast cancer metastasis suppressor 1 (BRMS1) has been reported to reduce breast and melanoma metastasis in vivo without affecting primary tumour growth (Hedley *et al.*, 2008; Hurst and Welch, 2011; Liu *et al.*, 2011). Recent progress in BRMS1 research leads us to the further insight in the molecular mechanism of BRMS1. It was suggested that BRMS1 can restore gap junctional intercellular communication of cells, a possibly mechanism contribute to metastasis suppression (Kapoor *et al.*, 2004; Saunders *et al.*, 2001). Despite the recent advance of our knowledge in BRMS1 role in cancer metastasis, the involvement of BRMS1 in other signalling pathway and regulation of invasiveness of breast cancer cells is relatively unexplored. This study aim to investigate the role of TGFβ and EGF signalling in BRMS1-regulated invasiveness of breast cancer cells.

In this study, we generated BRMS1 stable transfected MDA-MB-231. The expression of BRMS1 was demonstrated by immunoblotted with anti-BRMS1 antibody. BRMS1-transfected MDA-MB-231 cells showed more of the epithelial morphology, as compared to predominantly mesenchymal morphology of wild-type MDA-MB-231 cells. Thus, we generated cells stably transfected with BRMS1, and observed that these cells were viable and changed morphology.

In order to study the cellular activity of BRMS1, we performed proliferation, migration and invasion assays. These 3 assays were performed with the treatment of cells with TGFβ, EGF, SB431542, Iressa, alone and in combinations. We showed that TGFβ and EGF coordinately regulate the proliferation, migration and invasion activity of BRMS1-transfected MDA-MB-231 breast cancer cells.

In proliferation assay, our data suggested that TGFβ enhanced proliferation of MDA-MB-231 cells but decreased the proliferation of BRMS1-transfected cells. The combinational treatment of TGFβ and Iressa significantly enhanced the proliferation of BRMS1-transfected MDA-MB-231 cells, as compared to the treated parental cells.

To investigate the migration activity, we performed wound healing assay. We observed that TGFβ had a significantly enhanced effect on would closure upon expression of BRMS1. Simultaneous treatment with EGF prevented this effect. No other significant effects of BRMS1 expression or treatments were observed.

The most pronounced effect of BRMS1 on the cross-talk between TGFβ and EGF was observed on regulation of cell invasiveness. We observed that TGF-β enhanced invasiveness of control MDA-MB-231 cells but decreased the invasiveness of BRMS1-

transfected cells. EGF treatment led also to enhanced invasiveness, but no inhibitory effect of BRMS1 was observed. Combined treatment of cells with TGFβ and EGF led to even higher invasiveness of parental cells and re-appearance of inhibitory effect of BRMS1 expression. Surprisingly, the strongest stimulatory effect was observed when control cells were treated with TGFβ and inhibitor of TβR-I kinase. Even more striking is that this stimulation was strongly hampered by BRMS1 expression. Thus, BRMS1 expression had most pronounce impact on TGFβ and EGF-dependent regulation of cell invasiveness.

The effect of BRMS1 in invasiveness of MDA-MB-231 cells is likely due to cytoskeletal rearrangement and expression of molecules involved in focal adhesion. Expression of BRMs1 resulted in more epithelial morphology of otherwise more mesenchymal morphology MDA-MB-231 cells. We observed also that the vinculin staining was more diffuse in cytoplasm of BRMS1-expressing cells, as compared to defined patched staining in parental cells. Treatment of BRMS1-expressing cells with TGF-β and Iressa reverted vinculin staining pattern to patched, and led to more mesenchymal type of staining for actin (phalloidin staining).

Phosphorylation of Smad2 and Erk1/2 reflect activation of signaling down-stream of TGFβ and EGF. We observed that expression of BRMS1 decreased intensity of Smad2 phosphorylation upon treatments with TGFβ and EGF, and TGFβ and Iressa. At the same time, phosphorylation of Smad2 upon treatment with TGFβ only was similar in control and BRMS1 expressing cells. Phosphorylation of Erk1/2 was found being enhanced in control MDA-MB-231 cells, and significantly decreased upon expression of BRMS1. Induction of Erk1/2 phosphorylation upon different treatments was also modulated by BRMS1 expression. As an example, simultaneous treatment of cells with TGFβ and EGF enhanced Erk1/2 phosphorylation in parental cells, but this enhanced was not observed upon expression of BRMS1. Observed differences in phosphorylation of Smad2 and Erk1/2 indicate that BRMS1 may affect these signaling pathways in modulation of TGFβ and EGF effects. However, partial correlations of results of immunoblotting study with functional assays suggest that it may be also additional mechanisms employed by BRMS1 in modulation of TGFβ and EGF signaling.

In conclusion, our study showed that TGFβ and EGF may modulate BRMS1-mediated breast cancer invasion by regulating focal adhesion and cytoskeletal rearrangement, and that Smad2 and Erk1/2 phosphorylation are involved in modulatory molecular mechanisms engaged by BRMS1.

4 GENERAL CONCLUSION

Results of these studies contribute to the better understanding of the mechanism of human breast tumorigenesis, development of novel targets of anti-cancer treatment, and signatures for diagnostics and monitoring of breast cancer:

- Proteome profiling of human breast epithelial cells MCF10A and 184A1 cells showed that enhanced expression of cdk4 and cyclin D3 represent the early proliferation-specific event in the malignant transformation of breast epithelial cells.

- MCF7c46, an isogenic model of breast cancer invasion has been generated, which allowed the study of cancer invasion activity as a distinct step in metastasis. Systemics analysis showed relations of TGFβ, EGF, insulin, VEGF, HNF4α and NFκB in regulation of invasiveness. This indicates that insulin signaling disturbance may interfere with the invasive network, thus explained the increased risk of breast cancer metastasis in diabetes patients.

- Direct phosphorylation on serine 300 of eEF1A1 by TβRI provides novel mechanism of TGFβ-dependent regulation of protein synthesis, cell proliferation, and promotion of carcinogenic phenotype.

- eEF1A1 contribute to the increased proliferation of cells by promoting transition of cells through the S- and G2/M-phases of the cell cycle. Phosphorylation of eEF1A1 at Ser300 that impaired aa-tRNA binding inhibits this process.

- TGFβ and EGF may modulate BRMS1-mediated breast cancer invasion by regulating focal adhesion and cytoskeletal rearrangement, and that Smad2 and Erk1/2 phosphorylation are involved in molecular mechanisms engaged by BRMS1.

5 ACKNOWLEDGEMENTS

I would like to thanks my supervisor, Serhiy Souchelnytskyi, for the opportunity to pursuit my Ph.D., as well as the unlimited support and guidance; my co-supervisor, Rolf Lewensohn and Roman Zubarev, and external mentor, Ulf Hellman, for the guidance during my study; my friends and colleagues in Sweden and Malaysia for the support and encouragement; my family, my parents, grandmother and my wife, Mei Fong, for the never-ending support and love.

6 REFERENCES

Aebersold R, Mann M (2003). Mass spectrometry-based proteomics. *Nature* **422:** 198-207.

Al-Shahrour F, Diaz-Uriarte R, Dopazo J (2004). FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. *Bioinformatics* **20:** 578-80.

Al-Shahrour F, Minguez P, Tarraga J, Medina I, Alloza E, Montaner D *et al* (2007). FatiGO +: a functional profiling tool for genomic data. Integration of functional annotation, regulatory motifs and interaction data with microarray experiments. *Nucleic Acids Res* **35:** W91-6.

Alfonso P, Nunez A, Madoz-Gurpide J, Lombardia L, Sanchez L, Casal JI (2005). Proteomic expression analysis of colorectal cancer by two-dimensional differential gel electrophoresis. *Proteomics* **5:** 2602-11.

Anand N, Murthy S, Amann G, Wernick M, Porter LA, Cukier IH *et al* (2002). Protein elongation factor EEF1A2 is a putative oncogene in ovarian cancer. *Nat Genet* **31:** 301-5.

Arighi CN, Liu H, Natale DA, Barker WC, Drabkin H, Blake JA *et al* (2009). TGF-beta signaling proteins and the Protein Ontology. *BMC Bioinformatics* **10 Suppl 5:** S3.

Assenov Y, Ramirez F, Schelhorn SE, Lengauer T, Albrecht M (2008). Computing topological parameters of biological networks. *Bioinformatics* **24:** 282-4.

Avila-Campillo I, Drew K, Lin J, Reiss DJ, Bonneau R (2007). BioNetBuilder: automatic integration of biological networks. *Bioinformatics* **23:** 392-3.

Azuma H, Ehata S, Miyazaki H, Watabe T, Maruyama O, Imamura T *et al* (2005). Effect of Smad7 expression on metastasis of mouse mammary carcinoma JygMC(A) cells. *J Natl Cancer Inst* **97:** 1734-46.

Bader GD, Hogue CW (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* **4:** 2.

Barabasi AL, Oltvai ZN (2004). Network biology: understanding the cell's functional organization. *Nat Rev Genet* **5:** 101-13.

Bhaskaran N, Lin KW, Gautier A, Woksepp H, Hellman U, Souchelnytskyi S (2009). Comparative proteome profiling of MCF10A and 184A1 human breast epithelial cells emphasized involvement of CDK4 and cyclin D3 in cell proliferation. *Proteomics Clin Appl* **3:** 68-77.

Borrebaeck CA, Wingren C (2007). High-throughput proteomics using antibody microarrays: an update. *Expert Rev Mol Diagn* **7:** 673-86.

Boyle EI, Weng S, Gollub J, Jin H, Botstein D, Cherry JM *et al* (2004). GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* **20:** 3710-5.

Breitkreutz BJ, Stark C, Tyers M (2002). Osprey: a network visualization system. *Genome Biol* **3:** PREPRINT0012.

Breitkreutz BJ, Stark C, Tyers M (2003). Osprey: a network visualization system. *Genome Biol* **4:** R22.

Brennan DJ, O'Connor DP, Rexhepaj E, Ponten F, Gallagher WM (2010). Antibody-based proteomics: fast-tracking molecular diagnostics in oncology. *Nat Rev Cancer* **10:** 605-17.

Browne BC, O'Brien N, Duffy MJ, Crown J, O'Donovan N (2009). HER-2 signaling and inhibition in breast cancer. *Curr Cancer Drug Targets* **9:** 419-38.

Bussey KJ, Kane D, Sunshine M, Narasimhan S, Nishizuka S, Reinhold WC *et al* (2003). MatchMiner: a tool for batch navigation among gene and gene product identifiers. *Genome Biol* **4:** R27.

Butt AJ, Caldon CE, McNeil CM, Swarbrick A, Musgrove EA, Sutherland RL (2008). Cell cycle machinery: links with genesis and treatment of breast cancer. *Adv Exp Med Biol* **630:** 189-205.

Cans C, Passer BJ, Shalak V, Nancy-Portebois V, Crible V, Amzallag N *et al* (2003). Translationally controlled tumor protein acts as a guanine nucleotide dissociation inhibitor on the translation elongation factor eEF1A. *Proc Natl Acad Sci U S A* **100:** 13892-7.

Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S (2009). AmiGO: online access to ontology and annotation data. *Bioinformatics* **25:** 288-9.

Casey TM, Eneman J, Crocker A, White J, Tessitore J, Stanley M *et al* (2008). Cancer associated fibroblasts stimulated by transforming growth factor beta1 (TGF-beta 1) increase invasion rate of tumor cells: a population study. *Breast Cancer Res Treat* **110:** 39-49.

Cerami EG, Bader GD, Gross BE, Sander C (2006). cPath: open source software for collecting, storing, and querying biological pathways. *BMC Bioinformatics* **7:** 497.

Chen L, Madura K (2005). Increased proteasome activity, ubiquitin-conjugating enzymes, and eEF1A translation factor detected in breast cancer tissue. *Cancer Res* **65:** 5599-606.

Chen RH, Miettinen PJ, Maruoka EM, Choy L, Derynck R (1995). A WD-domain protein that is associated with and phosphorylated by the type II TGF-beta receptor. *Nature* **377:** 548-52.

Choy L, Derynck R (1998). The type II transforming growth factor (TGF)-beta receptor-interacting protein TRIP-1 acts as a modulator of the TGF-beta response. *J Biol Chem* **273:** 31455-62.

Cicek M, Fukuyama R, Cicek MS, Sizemore S, Welch DR, Sizemore N *et al* (2009). BRMS1 contributes to the negative regulation of uPA gene expression through recruitment of HDAC1 to the NF-kappaB binding site of the uPA promoter. *Clin Exp Metastasis* **26:** 229-37.

Clark EA, Golub TR, Lander ES, Hynes RO (2000). Genomic analysis of metastasis reveals an essential role for RhoC. *Nature* **406:** 532-5.

Condeelis J (1995). Elongation factor 1 alpha, translation and the cytoskeleton. *Trends Biochem Sci* **20:** 169-70.

Coppard NJ, Clark BF, Cramer F (1983). Methylation of elongation factor 1 alpha in mouse 3T3B and 3T3B/SV40 cells. *FEBS Lett* **164:** 330-4.

Corona G, De Lorenzo E, Elia C, Simula MP, Avellini C, Baccarani U *et al* (2010). Differential proteomic analysis of hepatocellular carcinoma. *Int J Oncol* **36:** 93-9.

Cravatt BF, Simon GM, Yates JR, 3rd (2007). The biological impact of mass-spectrometry-based proteomics. *Nature* **450:** 991-1000.

Derynck R, Zhang YE (2003). Smad-dependent and Smad-independent pathways in TGF-beta family signalling. *Nature* **425:** 577-84.

Dick JE (2009). Looking ahead in cancer stem cell research. *Nat Biotechnol* **27:** 44-6.

Dimmer EC, Huntley RP, Barrell DG, Binns D, Draghici S, Camon EB *et al* (2008). The Gene Ontology - Providing a Functional Role in Proteomic Studies. *Proteomics*.

Doisneau-Sixou SF, Sergio CM, Carroll JS, Hui R, Musgrove EA, Sutherland RL (2003). Estrogen and antiestrogen regulation of cell cycle progression in breast cancer cells. *Endocr Relat Cancer* **10:** 179-86.

Draghici S, Khatri P, Bhavsar P, Shah A, Krawetz SA, Tainsky MA (2003). Onto-Tools, the toolkit of the modern biologist: Onto-Express, Onto-Compare, Onto-Design and Onto-Translate. *Nucleic Acids Res* **31:** 3775-81.

Dudley JT, Butte AJ (2009). Identification of discriminating biomarkers for human disease using integrative network biology. *Pac Symp Biocomput***:** 27-38.

Dumont N, Arteaga CL (2002). The tumor microenvironment: a potential arbitrator of the tumor suppressive and promoting actions of TGFbeta. *Differentiation* **70:** 574-82.

Duncan MW, Hunsucker SW (2005). Proteomics as a tool for clinically relevant biomarker discovery and validation. *Exp Biol Med (Maywood)* **230:** 808-17.

Edmonds BT, Wyckoff J, Yeung YG, Wang Y, Stanley ER, Jones J *et al* (1996). Elongation factor-1 alpha is an overexpressed actin binding protein in metastatic rat mammary adenocarcinoma. *J Cell Sci* **109 (Pt 11):** 2705-14.

Enright AJ, Ouzounis CA (2001). BioLayout--an automatic graph layout algorithm for similarity visualization. *Bioinformatics* **17:** 853-4.

Ethier SP (1996). Human breast cancer cell lines as models of growth regulation and disease progression. *J Mammary Gland Biol Neoplasia* **1:** 111-21.

Feng XH, Derynck R (2005). Specificity and versatility in tgf-beta signaling through Smads. *Annu Rev Cell Dev Biol* **21:** 659-93.

Ferlay J, Autier P, Boniol M, Heanue M, Colombet M, Boyle P (2007). Estimates of the cancer incidence and mortality in Europe in 2006. *Ann Oncol* **18:** 581-92.

Funahashi A, Jouraku A, Matsuoka Y, Kitano H (2007). Integration of CellDesigner and SABIO-RK. *In Silico Biol* **7:** S81-90.

Gaggioli C, Hooper S, Hidalgo-Carcedo C, Grosse R, Marshall JF, Harrington K *et al* (2007). Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells. *Nat Cell Biol* **9:** 1392-400.

Galliher AJ, Schiemann WP (2006). Beta3 integrin and Src facilitate transforming growth factor-beta mediated induction of epithelial-mesenchymal transition in mammary epithelial cells. *Breast Cancer Res* **8:** R42.

Gao J, Ade AS, Tarcea VG, Weymouth TE, Mirel BR, Jagadish HV *et al* (2009). Integrating and annotating the interactome using the MiMI plugin for cytoscape. *Bioinformatics* **25:** 137-8.

Gasco M, Yulug IG, Crook T (2003). TP53 mutations in familial breast cancer: functional aspects. *Hum Mutat* **21:** 301-6.

Gast MC, Schellens JH, Beijnen JH (2009). Clinical proteomics in breast cancer: a review. *Breast Cancer Res Treat* **116:** 17-29.

Gobbi H, Arteaga CL, Jensen RA, Simpson JF, Dupont WD, Olson SJ *et al* (2000). Loss of expression of transforming growth factor beta type II receptor correlates with high tumour grade in human breast in-situ and invasive carcinomas. *Histopathology* **36:** 168-77.

Gorsch SM, Memoli VA, Stukel TA, Gold LI, Arrick BA (1992). Immunohistochemical staining for transforming growth factor beta 1 associates with disease progression in human breast cancer. *Cancer Res* **52:** 6949-52.

Graves PR, Haystead TA (2002). Molecular biologist's guide to proteomics. *Microbiol Mol Biol Rev* **66:** 39-63; table of contents.

Gretarsdottir S, Thorlacius S, Valgardsdottir R, Gudlaugsdottir S, Sigurdsson S, Steinarsdottir M *et al* (1998). BRCA2 and p53 mutations in primary breast cancer in relation to genetic instability. *Cancer Res* **58:** 859-62.

Gruber TR (1993). A translation approach to portable ontologies. *Knowledge Acquisition* **5:** 199-220.

Gupta GP, Massague J (2006). Cancer metastasis: building a framework. *Cell* **127:** 679-95.

Hamler RL, Zhu K, Buchanan NS, Kreunin P, Kachman MT, Miller FR *et al* (2004). A two-dimensional liquid-phase separation method coupled with mass spectrometry for proteomic studies of breast cancer and biomarker identification. *Proteomics* **4:** 562-77.

Hanahan D, Weinberg RA (2000). The hallmarks of cancer. *Cell* **100:** 57-70.

Hanahan D, Weinberg RA (2011). Hallmarks of cancer: the next generation. *Cell* **144:** 646-74.

Hedley BD, Vaidya KS, Phadke P, MacKenzie L, Dales DW, Postenka CO *et al* (2008). BRMS1 suppresses breast cancer metastasis in multiple experimental models of metastasis by reducing solitary cell survival and inhibiting growth initiation. *Clin Exp Metastasis* **25:** 727-40.

Holstege H, Joosse SA, van Oostrom CT, Nederlof PM, de Vries A, Jonkers J (2009). High incidence of protein-truncating TP53 mutations in BRCA1-related breast cancer. *Cancer Res* **69:** 3625-33.

Hondermarck H, Vercoutter-Edouart AS, Revillion F, Lemoine J, el-Yazidi-Belkoura I, Nurcombe V *et al* (2001). Proteomics of breast cancer for marker discovery and signal pathway profiling. *Proteomics* **1:** 1216-32.

Huber MA, Azoitei N, Baumann B, Grunert S, Sommer A, Pehamberger H *et al* (2004). NF-kappaB is essential for epithelial-mesenchymal transition and metastasis in a model of breast cancer progression. *J Clin Invest* **114:** 569-81.

Hurst DR, Welch DR (2011). Unraveling the enigmatic complexities of BRMS1-mediated metastasis suppression. *FEBS Lett* **585:** 3185-90.

Hurst DR, Xie Y, Edmonds MD, Welch DR (2009). Multiple forms of BRMS1 are differentially expressed in the MCF10 isogenic breast cancer progression model. *Clin Exp Metastasis* **26:** 89-96.

Ito Y, Kobayashi T, Takeda T, Komoike Y, Wakasugi E, Tamaki Y *et al* (1996). Immunohistochemical study of Cell Cycle Modulators in G(1)-S Transition in Clinical Breast Cancer Tissue. *Breast Cancer* **3:** 93-104.

Izawa T, Fukata Y, Kimura T, Iwamatsu A, Dohi K, Kaibuchi K (2000). Elongation factor-1 alpha is a novel substrate of rho-associated kinase. *Biochem Biophys Res Commun* **278:** 72-8.

Jacobs JM, Mottaz HM, Yu LR, Anderson DJ, Moore RJ, Chen WN *et al* (2004). Multidimensional proteome analysis of human mammary epithelial cells. *J Proteome Res* **3:** 68-75.

Jakowlew SB (2006). Transforming growth factor-beta in cancer and metastasis. *Cancer Metastasis Rev* **25:** 435-57.

Jenkinson SR, Barraclough R, West CR, Rudland PS (2004). S100A4 regulates cell motility and invasion in an in vitro model for breast cancer metastasis. *Br J Cancer* **90:** 253-62.

Jiang J, Wei Y, Liu D, Zhou J, Shen J, Chen X *et al* (2007). E1AF promotes breast cancer cell cycle progression via upregulation of Cyclin D3 transcription. *Biochem Biophys Res Commun* **358:** 53-8.

Johansson N, Ala-aho R, Uitto V, Grenman R, Fusenig NE, Lopez-Otin C *et al* (2000). Expression of collagenase-3 (MMP-13) and collagenase-1 (MMP-1) by transformed keratinocytes is dependent on the activity of p38 mitogen-activated protein kinase. *J Cell Sci* **113 Pt 2:** 227-35.

Joseph P, Lei YX, Whong WZ, Ong TM (2002). Oncogenic potential of mouse translation elongation factor-1 delta, a novel cadmium-responsive proto-oncogene. *J Biol Chem* **277:** 6131-6.

Kapoor P, Saunders MM, Li Z, Zhou Z, Sheaffer N, Kunze EL *et al* (2004). Breast cancer metastatic potential: correlation with increased heterotypic gap junctional intercellular communication between breast cancer cells and osteoblastic cells. *Int J Cancer* **111:** 693-7.

Kato MV, Sato H, Nagayoshi M, Ikawa Y (1997). Upregulation of the elongation factor-1alpha gene by p53 in association with death of an erythroleukemic cell line. *Blood* **90:** 1373-8.

Khanna C, Khan J, Nguyen P, Prehn J, Caylor J, Yeung C *et al* (2001). Metastasis-associated differences in gene expression in a murine model of osteosarcoma. *Cancer Res* **61:** 3750-9.

Khatri P, Draghici S, Ostermeier GC, Krawetz SA (2002). Profiling gene expression using onto-express. *Genomics* **79:** 266-70.

Kim ES, Kim MS, Moon A (2004). TGF-beta-induced upregulation of MMP-2 and MMP-9 depends on p38 MAPK, but not ERK signaling in MCF10A human breast epithelial cells. *Int J Oncol* **25:** 1375-82.

Kitano H (2002a). Computational systems biology. *Nature* **420:** 206-10.

Kitano H (2002b). Systems biology: a brief overview. *Science* **295:** 1662-4.

Knudsen SM, Frydenberg J, Clark BF, Leffers H (1993). Tissue-dependent variation in the expression of elongation factor-1 alpha isoforms: isolation and characterisation of a cDNA encoding a novel variant of human elongation-factor 1 alpha. *Eur J Biochem* **215:** 549-54.

Kohl M, Wiese S, Warscheid B (2011). Cytoscape: software for visualization and analysis of biological networks. *Methods Mol Biol* **696:** 291-303.

Kolker E, Higdon R, Hogan JM (2006). Protein identification and expression analysis using mass spectrometry. *Trends Microbiol* **14:** 229-35.

Krymskaya VP, Hoffman R, Eszterhas A, Ciocca V, Panettieri RA, Jr. (1997). TGF-beta 1 modulates EGF-stimulated phosphatidylinositol 3-kinase activity in human airway smooth muscle cells. *Am J Physiol* **273:** L1220-7.

Kwoh CK, Ng PY (2007). Network analysis approach for biology. *Cell Mol Life Sci* **64:** 1739-51.

Lacroix M, Leclercq G (2004). Relevance of breast cancer cell lines as models for breast tumours: an update. *Breast Cancer Res Treat* **83:** 249-89.

Lamberti A, Caraglia M, Longo O, Marra M, Abbruzzese A, Arcari P (2004). The translation elongation factor 1A in tumorigenesis, signal transduction and apoptosis: review article. *Amino Acids* **26:** 443-8.

Lamouille S, Derynck R (2007). Cell size and invasion in TGF-beta-induced epithelial to mesenchymal transition is regulated by activation of the mTOR pathway. *J Cell Biol* **178:** 437-51.

Lee MK, Pardoux C, Hall MC, Lee PS, Warburton D, Qing J *et al* (2007). TGF-beta activates Erk MAP kinase signalling through direct phosphorylation of ShcA. *EMBO J* **26:** 3957-67.

Lee S, Francoeur AM, Liu S, Wang E (1992). Tissue-specific expression in mammalian brain, heart, and muscle of S1, a member of the elongation factor-1 alpha gene family. *J Biol Chem* **267:** 24064-8.

Leivonen SK, Ala-Aho R, Koli K, Grenman R, Peltonen J, Kahari VM (2006). Activation of Smad signaling enhances collagenase-3 (MMP-13) expression and invasion of head and neck squamous carcinoma cells. *Oncogene* **25:** 2588-600.

Li J, Li G (2010). Cell cycle regulator ING4 is a suppressor of melanoma angiogenesis that is regulated by the metastasis suppressor BRMS1. *Cancer Res* **70:** 10445-53.

Lim MS, Elenitoba-Johnson KS (2004). Proteomics in pathology research. *Lab Invest* **84:** 1227-44.

Lin KW, Jia M, Souchelnytskyi S (2011). Application of bioinformatics tools in gel-based proteomics. *Proteomics / Book 2 (Eastwood Leung, ed), ISBN 979-953-307-693-4, InTech. (In Press).*

Lin KW, Souchelnytskyi S (2010). Eukaryotic elongation factor eEF1A1 promotes and Ser300 mutants of eEF1A1 inhibit transition through the S and G2/M phases of the cell cycle. *J Cell Mol Biol* **8:** 125-130.

Lin KW, Souchelnytskyi S (2011). Translational connection of TGFbeta signaling: Phosphorylation of eEF1A1 by TbetaR-I inhibits protein synthesis. *Small Gtpases* **2:** 104-108.

Lin KW, Yakymovych I, Jia M, Yakymovych M, Souchelnytskyi S (2010). Phosphorylation of eEF1A1 at Ser300 by TbetaR-I results in inhibition of mRNA translation. *Curr Biol* **20:** 1615-25.

Lin SW, Lee MT, Ke FC, Lee PP, Huang CJ, Ip MM *et al* (2000). TGFbeta1 stimulates the secretion of matrix metalloproteinase 2 (MMP2) and the invasive behavior in human ovarian cancer cells, which is suppressed by MMP inhibitor BB3103. *Clin Exp Metastasis* **18:** 493-9.

Liotta LA (1986). Tumor invasion and metastases--role of the extracellular matrix: Rhoads Memorial Award lecture. *Cancer Res* **46:** 1-7.

Liu G, Grant WM, Persky D, Latham VM, Jr., Singer RH, Condeelis J (2002). Interactions of elongation factor 1alpha with F-actin and beta-actin mRNA: implications for anchoring mRNA in cell protrusions. *Mol Biol Cell* **13:** 579-92.

Liu H, Ding J, Chen F, Fan B, Gao N, Yang Z *et al* (2010). Increased expression of elongation factor-1alpha is significantly correlated with poor prognosis of human prostate cancer. *Scand J Urol Nephrol* **44:** 277-83.

Liu Y, Mayo MW, Nagji AS, Smith PW, Ramsey CS, Li D *et al* (2011). Phosphorylation of RelA/p65 promotes DNMT-1 recruitment to chromatin and represses transcription of the tumor metastasis suppressor gene BRMS1. *Oncogene*.

Macpherson IR, Hooper S, Serrels A, McGarry L, Ozanne BW, Harrington K *et al* (2007). p120-catenin is required for the collective invasion of squamous cell carcinoma cells via a phosphorylation-independent mechanism. *Oncogene* **26:** 5214-28.

Massague J (2008). TGFbeta in Cancer. *Cell* **134:** 215-30.

Massague J, Blain SW, Lo RS (2000). TGFbeta signaling in growth control, cancer, and heritable disorders. *Cell* **103:** 295-309.

McGonigle S, Beall MJ, Pearce EJ (2002). Eukaryotic initiation factor 2 alpha subunit associates with TGF beta receptors and 14-3-3 epsilon and acts as a modulator of the TGF beta response. *Biochemistry* **41:** 579-87.

Mego M, Mani SA, Cristofanilli M (2010). Molecular mechanisms of metastasis in breast cancer--clinical applications. *Nat Rev Clin Oncol* **7:** 693-701.

Mester J, Redeuilh G (2008). Proliferation of breast cancer cells: regulation, mediators, targets for therapy. *Anticancer Agents Med Chem* **8:** 872-85.

Meunier B, Dumas E, Piec I, Bechet D, Hebraud M, Hocquette JF (2007). Assessment of hierarchical clustering methodologies for proteomic data mining. *J Proteome Res* **6:** 358-66.

Milli A, Cecconi D, Campostrini N, Timperio AM, Zolla L, Righetti SC *et al* (2008). A proteomic approach for evaluating the cell response to a novel histone deacetylase inhibitor in colon cancer cells. *Biochim Biophys Acta* **1784:** 1702-10.

Miyazono K, Suzuki H, Imamura T (2003). Regulation of TGF-beta signaling and its roles in progression of tumors. *Cancer Sci* **94:** 230-4.

Moldave K (1985). Eukaryotic protein synthesis. *Annu Rev Biochem* **54:** 1109-49.

Montanaro L, Vici M, Donati G, Ceccarelli C, Santini D, Trere D *et al* (2007). Controversial relationship between the expression of the RB pathway components and RB protein phosphorylation in human breast cancer. *Histol Histopathol* **22:** 769-75.

Mori S, Matsuzaki K, Yoshida K, Furukawa F, Tahashi Y, Yamagata H *et al* (2004). TGF-beta and HGF transmit the signals through JNK-dependent Smad2/3 phosphorylation at the linker regions. *Oncogene* **23:** 7416-29.

Mu Y, Gudey SK, Landstrom M (2011). Non-Smad signaling pathways. *Cell Tissue Res*.

Natale DA, Arighi CN, Barker WC, Blake J, Chang TC, Hu Z *et al* (2007). Framework for a protein ontology. *BMC Bioinformatics* **8 Suppl 9:** S1.

Natale DA, Arighi CN, Barker WC, Blake JA, Bult CJ, Caudy M *et al* (2011). The Protein Ontology: a structured representation of protein forms and complexes. *Nucleic Acids Res* **39:** D539-45.

Neil JR, Schiemann WP (2008). Altered TAB1:I kappaB kinase interaction promotes transforming growth factor beta-mediated nuclear factor-kappaB activation during breast cancer progression. *Cancer Res* **68:** 1462-70.

O'Donovan N, Crown J (2007). EGFR and HER-2 antagonists in breast cancer. *Anticancer Res* **27:** 1285-94.

Oft M, Peli J, Rudaz C, Schwarz H, Beug H, Reichmann E (1996). TGF-beta1 and Ha-Ras collaborate in modulating the phenotypic plasticity and invasiveness of epithelial tumor cells. *Genes Dev* **10:** 2462-77.

Ozdamar B, Bose R, Barrios-Rodiles M, Wang HR, Zhang Y, Wrana JL (2005). Regulation of the polarity protein Par6 by TGFbeta receptors controls epithelial cell plasticity. *Science* **307:** 1603-9.

Page MJ, Amess B, Townsend RR, Parekh R, Herath A, Brusten L *et al* (1999). Proteomic definition of normal human luminal and myoepithelial breast cells purified from reduction mammoplasties. *Proc Natl Acad Sci U S A* **96:** 12589-94.

Pecorari L, Marin O, Silvestri C, Candini O, Rossi E, Guerzoni C *et al* (2009). Elongation Factor 1 alpha interacts with phospho-Akt in breast cancer cells and regulates their proliferation, survival and motility. *Mol Cancer* **8:** 58.

Peinado H, Olmeda D, Cano A (2007). Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat Rev Cancer* **7:** 415-28.

Perroud B, Lee J, Valkova N, Dhirapong A, Lin PY, Fiehn O *et al* (2006). Pathway analysis of kidney cancer using proteomics and metabolic profiling. *Mol Cancer* **5:** 64.

Peters HI, Chang YW, Traugh JA (1995). Phosphorylation of elongation factor 1 (EF-1) by protein kinase C stimulates GDP/GTP-exchange activity. *Eur J Biochem* **234:** 550-6.

Petritsch C, Beug H, Balmain A, Oft M (2000). TGF-beta inhibits p70 S6 kinase via protein phosphatase 2A to induce G(1) arrest. *Genes Dev* **14:** 3093-101.

Pierce DF, Jr., Gorska AE, Chytil A, Meise KS, Page DL, Coffey RJ, Jr. *et al* (1995). Mammary tumor suppression by transforming growth factor beta 1 transgene expression. *Proc Natl Acad Sci U S A* **92:** 4254-8.

Pukrop T, Klemm F, Hagemann T, Gradl D, Schulz M, Siemes S *et al* (2006). Wnt 5a signaling is critical for macrophage-induced invasion of breast cancer cell lines. *Proc Natl Acad Sci U S A* **103:** 5454-9.

Rao JS (2003). Molecular mechanisms of glioma invasiveness: the role of proteases. *Nat Rev Cancer* **3:** 489-501.

Reeves GA, Eilbeck K, Magrane M, O'Donovan C, Montecchi-Palazzi L, Harris MA *et al* (2008). The Protein Feature Ontology: a tool for the unification of protein feature annotations. *Bioinformatics* **24:** 2767-72.

Rhee DK, Park SH, Jang YK (2008). Molecular signatures associated with transformation and progression to breast cancer in the isogenic MCF10 model. *Genomics* **92:** 419-28.

Saunders MM, Seraj MJ, Li Z, Zhou Z, Winter CR, Welch DR *et al* (2001). Breast cancer metastatic potential correlates with a breakdown in homospecific and heterospecific gap junctional intercellular communication. *Cancer Res* **61:** 1765-7.

Scardoni G, Petterlini M, Laudanna C (2009). Analyzing biological network parameters with CentiScaPe. *Bioinformatics* **25:** 2857-9.

Selicharova I, Smutna K, Sanda M, Ubik K, Matouskova E, Bursikova E *et al* (2007). 2-DE analysis of a new human cell line EM-G3 derived from breast cancer progenitor cells and comparison with normal mammary epithelial cells. *Proteomics* **7:** 1549-59.

Seton-Rogers SE, Lu Y, Hines LM, Koundinya M, LaBaer J, Muthuswamy SK *et al* (2004). Cooperation of the ErbB2 receptor and transforming growth factor beta in induction of migration and invasion in mammary epithelial cells. *Proc Natl Acad Sci U S A* **101:** 1257-62.

Shackleton M, Quintana E, Fearon ER, Morrison SJ (2009). Heterogeneity in cancer: cancer stem cells versus clonal evolution. *Cell* **138:** 822-9.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D *et al* (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13:** 2498-504.

Shi Y, Massague J (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell* **113:** 685-700.

Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W *et al* (2007). The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nat Biotechnol* **25:** 1251-5.

Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27:** 431-2.

Souchelnytskyi S (2002). Proteomics in studies of signal transduction in epithelial cells. *J Mammary Gland Biol Neoplasia* **7:** 359-71.

Souchelnytskyi S (2005). Bridging proteomics and systems biology: what are the roads to be traveled? *Proteomics* **5:** 4123-37.

Stampfer MR, Yaswen P (2000). Culture models of human mammary epithelial cell transformation. *J Mammary Gland Biol Neoplasia* **5:** 365-78.

Taniguchi S, Miyamoto S, Sadano H, Kobayashi H (1991). Rat elongation factor 1 alpha: sequence of cDNA from a highly metastatic fos-transferred cell line. *Nucleic Acids Res* **19:** 6949.

Tatsuka M, Mitsui H, Wada M, Nagata A, Nojima H, Okayama H (1992). Elongation factor-1 alpha gene determines susceptibility to transformation. *Nature* **359:** 333-6.

Taylor CF, Paton NW, Garwood KL, Kirby PD, Stead DA, Yin Z *et al* (2003). A systematic approach to modeling, capturing, and disseminating proteomics experimental data. *Nat Biotechnol* **21:** 247-54.

Thiery JP (2002). Epithelial-mesenchymal transitions in tumour progression. *Nat Rev Cancer* **2:** 442-54.

Thomas S, Bonchev D (2010). A survey of current software for network analysis in molecular biology. *Hum Genomics* **4:** 353-60.

Turner KE, Kumar HR, Hoelz DJ, Zhong X, Rescorla FJ, Hickey RJ *et al* (2009). Proteomic analysis of neuroblastoma microenvironment: effect of the host-tumor interaction on disease progression. *J Surg Res* **156:** 116-22.

Ueda Y, Wang S, Dumont N, Yi JY, Koh Y, Arteaga CL (2004). Overexpression of HER2 (erbB2) in human breast epithelial cells unmasks transforming growth factor beta-induced cell motility. *J Biol Chem* **279:** 24505-13.

Valastyan S, Weinberg RA (2011). Tumor metastasis: molecular insights and evolving paradigms. *Cell* **147:** 275-92.

Walker RA, Dearing SJ (1992). Transforming growth factor beta 1 in ductal carcinoma in situ and invasive carcinomas of the breast. *Eur J Cancer* **28:** 641-4.

Wick W, Platten M, Weller M (2001). Glioma cell invasion: regulation of metalloproteinase activity by TGF-beta. *J Neurooncol* **53:** 177-85.

Wingren C, Borrebaeck CA (2004). High-throughput proteomics using antibody microarrays. *Expert Rev Proteomics* **1:** 355-64.

Wolf K, Wu YI, Liu Y, Geiger J, Tam E, Overall C *et al* (2007). Multi-step pericellular proteolysis controls the transition from individual to collective cancer cell invasion. *Nat Cell Biol* **9:** 893-904.

Worsham MJ, Pals G, Schouten JP, Miller F, Tiwari N, van Spaendonk R *et al* (2006). High-resolution mapping of molecular events associated with immortalization, transformation, and progression to breast cancer in the MCF10 model. *Breast Cancer Res Treat* **96:** 177-86.

Yu Q, Sicinska E, Geng Y, Ahnstrom M, Zagozdzon A, Kong Y *et al* (2006). Requirement for CDK4 kinase function in breast cancer. *Cancer Cell* **9:** 23-32.

Zajchowski DA, Bartholdi MF, Gong Y, Webster L, Liu HL, Munishkin A *et al* (2001). Identification of gene expression profiles that predict the aggressive behavior of breast cancer cells. *Cancer Res* **61:** 5168-78.

Zavadil J, Bottinger EP (2005). TGF-beta and epithelial-to-mesenchymal transitions. *Oncogene* **24:** 5764-74.

Zeeberg BR, Feng W, Wang G, Wang MD, Fojo AT, Sunshine M *et al* (2003). GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biol* **4:** R28.

Zhang B, Schmoyer D, Kirov S, Snoddy J (2004). GOTree Machine (GOTM): a web-based platform for interpreting sets of interesting genes using Gene Ontology hierarchies. *BMC Bioinformatics* **5:** 16.

Zhong S, Storch KF, Lipan O, Kao MC, Weitz CJ, Wong WH (2004). GoSurfer: a graphical interactive tool for comparative analysis of large gene sets in Gene Ontology space. *Appl Bioinformatics* **3:** 261-4.