MARMARA UNIVERSITY

INSTITUTE FOR GRADUATE STUDIES
IN PURE AND APPLIED SCIENCES

SYSTEMATIC AND INTEGRATIVE ANALYSIS OF BREAST CANCER AND OTHER ASSOCIATED DISEASES USING TRANSCRIPTOME DATA AND INTERACTOME NETWORKS

KÜBRA KARAGÖZ

Ph.D. THESIS

Department of Bioengineering

ADVISOR

Assoc. Prof. Kazım Yalçın ARĞA

ISTANBUL, 2015

MARMARA UNIVERSITY

INSTITUTE FOR GRADUATE STUDIES IN PURE AND APPLIED SCIENCES

SYSTEMATIC AND INTEGRATIVE ANALYSIS OF BREAST CANCER AND OTHER ASSOCIATED DISEASES USING TRANSCRIPTOME DATA AND INTERACTOME NETWORKS

KÜBRA KARAGÖZ

Ph.D. THESIS

Department of Bioengineering

ADVISOR

Assoc. Prof. Kazım Yalçın ARĞA

ISTANBUL, 2015

ACKNOWLEDGEMENT

First and foremost, I would like to represent my special appreciation and thanks to my advisor, Assoc. Prof. Kazım Yalçın ARĞA. It has been an honor to be his first PhD student. I would like to thank him for encouraging my research and for allowing me to grow as a research scientist. His advices on both research and on my career have been invaluable. I appreciate all his patience and contributions of ideas to make my PhD experience productive. I would also like to thank Prof. Betul KIRDAR and Prof Ahu ALTINKUT UNCUOGLU for serving as my committee members in my dissertation.

A special thanks to my precious family members. Words cannot express how grateful I am to my mother, Hüsniye Büyükyanbolu, father, Ragup Büyükyanbolu and also my brother, Burak Büyükyanbolu for all of the sacrifices. Your prayer for me was what sustained me thus far. I would also like to thank my cousin, Melek Arı, all of my friends from Systems Biomedicine Group in Marmara University, especially Tuba Sevimoğlu and Beste Çalımlıoğlu and colleagues from Istanbul Metropolitan Municipality who supported and believed me to achieve my goals. At the end I would like to express appreciation to my loving, supportive, encouraging, and patient husband, Turgay Karagöz whose faithful support during the all stages of this PhD and my beloved daughter, Zeynep Azra Karagöz who is the most directly affected through the way of my studies but always tries to be thoughtful and supporter although she is only five years old.

September, 2015 **Kübra KARAGÖZ**

TABLE OF CONTENTS

ÖZET

Tez Başlığı: Meme kanseri ve ilişkili hastalıkların transkriptom verileri ve interaktom ağları kullanılarak sistematik ve bütüncül analizi

Hastalıklar nadiren tek bir gendeki anormallikler sonucunda oluşurlar, genellikle karmaşik hücresel ağlardaki (protein-protein etkileşim, transkripsyonel düzenleyici ve metabolik ağlar gibi) bozulmalar sonucunda meydana gelirler. İnsan protein interaktomunun kalite ve kapsamının tıbbi biyoloji gibi çeşitli disiplinlerde kullanılabilmesi için geliştirilmesi gerekmektedir. Bu nedenle, yüksek güvenilirlikli protein etkileşim ağı kurmak için yeni bir skorlama yöntemi geliştirilmiştir. Metotun uygulanabilirliği mayada, *Saccharomyces cerevisiae*, gösterilmiştir ve sonrasında yüksek güvenilirlikli ve tamamlanmış bir insan interaktomu kurulmuştur.

Bu çalışmada meme kanseri, üçlü negatif meme kanseri ve yemek borusu kanserlerinin moleküler mekanizmaları hakkında detaylı bilgi sunan çoklu omik analizleri kullanılmıştır. Bütüncül analizler sonucu ise her hastalığın mekanizmasına ait önemli ve hedef olabilecek yolaklar belirlenmiş, potansiyel ilaç hedefleri ve biyo-işaretçiler tayin edilmiştir. RHOB, ARGHDIB, GSTM3, MLPH, ATF5, CSF2RA, ve CXCL10 genlerinin üçlü negatif meme kanseriyle, ACPP, C2orf54, DYNLT3, KANK1, ENDOU, FMO2, HOMER3, RFC4, COL10A1, FNDC3B ve MARCKSL1 genlerinin de yemek borusu kanserleriyle ilişkili oldukları ilk olarak bu çalışmayla sunulmuştur.

Hastalığa özel yeni ilaç hedefleri ve biyo-işaretçilerin belirlenmesini sağlayan önerdiğimiz çoklu-omik yaklaşımı meme kanseri, üçlü negatif meme kanseri ve yemek borusu kanserlerine yönelik yeni tanı ve kişiselleştirilmiş tıp tedavi yöntemlerinin tasarlanmasına olanak sağlayacaktır.

September, 2015 **Kübra KARAGÖZ**

ABSTRACT

Thesis Title: Systematic and Integrative Analysis of Breast Cancer and Other Associated Diseases Using Transcriptome Data and Interactome Networks

The diseases are rarely the outcomes of aberrance in a single gene but they are consequents of the disruptions of the complex cellular networks including protein-protein interaction (PPI), transcriptional regulatory and metabolic networks. The quality and coverage of human protein interactome needs to be improved to be used in various disciplines, especially in biomedicine. Therefore, novel scoring methodology was developed for reconstruction of highly reliable protein-protein interaction network. The applicability of method was shown in yeast, *Saccharomyces cerevisiae*, and then a consensus – highly reliable, almost complete – interactome of *Homo sapiens* was reconstructed.

A multi-omic discovery approach was developed which could offer detailed information about the molecular mechanisms of breast cancer, triple negative breast cancer (TNBC), and esophageal squamous cell carcinoma (ESCC). As a result of a comprehensive data assessment and processing, we identified disease specific signaling and metabolic pathways together with candidate novel drug targets and biomarkers. The associations of RHOB, ARGHDIB, GSTM3, MLPH, ATF5, CSF2RA, and CXCL10 with TNBC were reported for the first time. Moreover, ACPP, C2orf54, DYNLT3, KANK1, ENDOU, FMO2, HOMER3, RFC4, COL10A1, FNDC3B and MARCKSL1 were reported for the first time in ESCC.

The multi-omics molecular target and biomarker discovery approach presented here can offer ways forward on novel diagnostics and potentially help to design personalized therapeutics for breast cancer, triple negative breast cancer and esophageal squamous cell carcinoma in the future.

September, 2015 **Kübra KARAGÖZ**

CLAIM OF ORGINALITY

Breast cancer is the most common malignancy among women with an estimated 1.38 million new cases and 458000 deaths worldwide every year, and identifying reliable molecular biomarkers and drug targets of breast cancer is the major challenge for successful treatment. Moreover, triple negative breast cancer (TNBC), which constitutes approximately 15% of breast cancer cases, is typically distinguished by high metastatic and mortality rate among other types. The limited treatment options are available for patients with TNBC. Another cancer type focused on this study is esophageal squamous cell carcinoma (ESCC) with poor survival rate, i.e., fourth worst among all cancers. The proper biomarkers for clinical diagnosis and prognosis have yet not been defined for ESCC yet.

The aim of this thesis was to elucidate the molecular mechanisms of breast cancer, TNBC and ESCC through a network based approach. In this context, a novel scoring methodology was presented for reconstruction of highly reliable protein-protein interaction networks. We developed multi-omic discovery approach that offer detailed information about the molecular mechanisms of the diseases. This research as a significant intervention proposed novel candidate biomarkers as well as potential drug targets for further experimental analysis. The associations of RHOB, ARGHDIB, GSTM3, MLPH, ATF5, CSF2RA, and CXCL10 with TNBC were reported for the first time. Moreover, the down-regulated ACPP, C2orf54, DYNLT3, KANK1, ENDOU, FMO2 and up-regulated HOMER3, RFC4, COL10A1, FNDC3B and MARCKSL1 were reported for the first time in ESCC.

This study enables other researchers the standard multi-omic pipeline for data assessment and processing to enlighten the driving mechanisms of other complex diseases.

September, 2015 Assoc. Prof. Kazım Yalçın ARĞA Kübra KARAGÖZ

ABBREVIATIONS

AA: Arachidonic Acid

AUC: Area Under Curve

DEG: Differentially Expressed Genes

ECM: Extracellular Matrix

ER: Estrogen Receptor

ESCC: Esophageal Squamous Cell Carcinoma

GO: Gene Ontology

HCS: High Confidence Dataset

HER2: Human Epidermal Growth Factor 2

LCS: Low Confidence Dataset

MS: Mass Spectrometrometry

NMR: Nuclear Magnetic Resonance

PPI: Protein-protein Interaction

ROC: Receiver Operating Characteristic

TF: Transcriptional Factor

TLR: Tool like Receptor

TNBC: Triple Negative Breast Cancer

TSG: Tumor Suppressor Gene

Y2H: Yeast two-hybrid

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

The diseases are rarely the outcomes of aberrance in a single gene, yet they are consequents of the disruptions of the complex cellular networks. The network medicine as a systems biology platform provides to discover the molecular mechanisms of diseases and the molecular relationships among different phenotypes that offer to identify the disease specific molecules, modules and pathways. Therefore, improvements in the network medicine are vital to determine novel biomarkers as well as drug targets (Barabási et al., 2011).

Cancer, as a complex and heterogeneous disease, occurs as results of abnormalities in different cellular events. It was suggested that tumor phenotypes arised from particular interactions and subsequent activation of distinct subnetworks (Miller-Jensen et al., 2007). Therefore, targeted single gene/gene products therapies in which highly specific compounds were used and proposed to fail in case of not to perturb the associated network. The network-based approaches have arised in order to get through this difficulty (Pawson and Linding, 2008). The interaction of cellular molecules with each other is required to carry out their biological functions and complex networks are formed by these dynamic intracellular and intercellular interactions. The tumor phenotypes are determined by these interactions so cancer-associated networks could be promising targets for therapeutic discoveries. The signaling networks are dynamically altered in cancer cells with radio- and chemotherapies to protect themselves against external stimuli which induce resistance to therapeutics. Therefore, it is extremely important to identify which components have vital role in network and which subnetworks are activated or inactivated upon therapies to improve better strategies for cancer treatment (Seo et al., 2013).

In the next part, network based approaches which targets cellular networks to develop effective anticancer therapies are described briefly.

1.1 Network-based Approaches

Network-based systems biology, so called network biology, can study distinct biomolecular networks, and it includes multidisciplinary concepts such as mathematics, computer science and biology. The aim of the network biology is to explore how cellular mechanisms enable biological phenotypes that are resulted by the form of networks among genes, proteins and metabolites. The accelerated breakthroughs in network biology offer valuable pipelines to analyse high-throughput data that enhances our understanding of cellular mechanisms so that facilitates significant interventions in biomedicine (Barabasi and Oltvai, 2004 and Chen et al., 2009).

The "omics" tools including genomics, transcriptomics, proteomics and metabolomics and network-based approaches reforms researches in biomedicine that offer a chance to shift of complex diseases investigations from the traditional techniques to systems biology methods (Xiong et al., 2005 and Ideker and Sharan, 2008). Although the traditional biomedicine including experimental biology i.e. molecular biology and genetics has made significant advances to highlight molecular mechanisms of diseases, these experimental methods depend on investigating the role of individual genes and proteins which are not sufficient to elucidate mechanisms of complex diseases so the great challenges have emerged against the complex diseases for researches in biomedicine (Hornberg et al. 2006).

Network-based systems biology presents novel approaches to interpret the non-linear interactions among several components and to distinguish the phenotypic alterations through the analysis of dynamic nature of cellular networks. Moreover, it enlightens the complex diseases through the determination of the cellular pathways and networks and the impacts of environmental alterations to them (Chen et al., 2009).

Network-based approaches enable significant biological insights to multifactorial diseases. To understand the biological functions, the first step is the topological analysis of a network (Feldman et al., 2008). The topological attributes of a network including individual molecules (i.e. nodes), interacting pairs (i.e. edges), local structures (i.e. motifs and modules) and global structures of a network provide invaluable data to pinpoint disease associated genes/gene products and subnetworks as well as

dysregulated pathways. Furthermore, the network concepts also offer the candidate drug targets and drug combination for the treatments of complex diseases (Liu et al., 2012).

The therapies based on network biology for treatment of human cancers have many promising clinical interventions. Notably, hub molecules (highly connected nodes) could serve as biomarkers in a disease network and it was proposed that circumstances of cancer cells are governed by the dysregulation of hub molecules. Analysis of cellular networks and hub molecules will allow identifying new drug targets and diagnostic biomarkers to distinguish early stages of cancer. Single molecule targeted therapies have several limitations because biological networks includes many disease modules emerged in cancer cells. Additionally, multi-targeted therapies are not optimal cancer treatments yet since the structure of networks are dynamic and changed by a stimulus-dependent manner to keep homeostasis. Network-based integrative strategies could be the most auspicious methods for targeting particular mediators like novel hubs which alters the states of a network. Therefore, these methods could increase the effectiveness of classical therapies which particularly target hubs and perturb the main disease modules. A huge amount of bioinformatics data have been arised and obtained over time to explore the network structures and states. On the other hand, the modification of network state response to radio- and chemotherapies could be possible that utilize the discovery of novel drug combinations and personalized medicine strategies (Seo et al., 2013).

The molecular networks: *i) regulatory networks*, whose nodes represent transcriptional factors (TF) and genes and directed edges represent regulatory relationships between them. At genomic level, TFs as products of genes, regulate the activities of gene and the gene expression through gene regulatory network or a transcriptional regulatory network (Figure 1.1 A) (Carninci et al., 2005), *ii) protein interaction networks*, whose nodes are proteins and the edges express physical interactions among protein pairs. At proteomic level, proteins interact with each other to constitute protein complexes and biological pathways to take part in distinct biological processes. These associations between protein pairs are named as protein–protein interactions (PPIs), which form a PPI network (Figure 1.1 B) (Rual et al., 2005); *iii) metabolic networks*, whose nodes are metabolites and the edges express the connection if they take part in the same biochemical reactions (Figure 1.1 C) (Jeong et al., 2000). The metabolic networks are

constituted by the biochemical reactions and enzymes which catalyse the reactions, regulate fluxes (Chen et al., 2009). In the next parts, more detailed information was given about the molecular networks which were applied to elucidate the molecular mechanisms of breast cancer, triple negative breast cancer and esophageal squamous cell carcinoma.

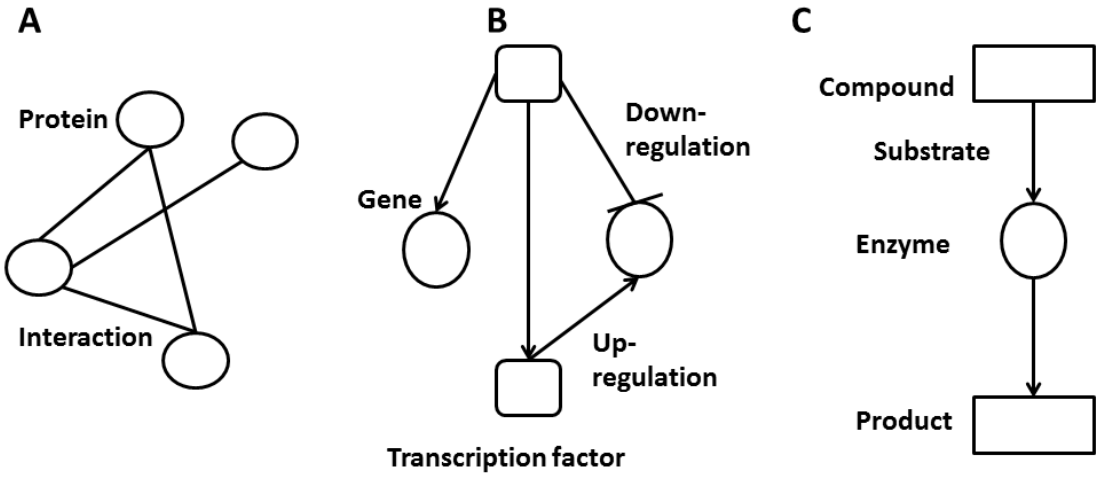


Figure 1.1 Graphical representations of biological networks, (**A)** Protein-protein interaction network, (B) Transcriptional regulatory network, (C) Metabolic networks.

1.1.1 Protein-protein interaction networks

Protein-protein interaction networks as undirected graph includes binary interactions between protein pairs. Nodes represent proteins and an edge among two nodes represents the interaction between the corresponding proteins (Wagner, 2003), as shown in Figure 1A. The analysis of PPI network offer several advantages including annotation of functionality since neighbor proteins in the network probably have common functionality and densely connected subnetworks are likely to constitute protein complexes as a unit of particular biological processes (Vittorio Cannistraci et al., 2013). The experimental techniques to identify PPIs include the *in vitro* and *in vivo* methods (i.e. affinity purification, yeast two hybrid, tandem affinity purification, coimmunoprecipitation, protein microarrays, X-ray crystallography, NMR spectroscopy) but these have their own disadvantages in cost and time and obtained data

are often noisy, incomplete and have more false positives. Therefore, in silico methods which include sequence-based approaches, structure-based approaches, chromosome proximity, gene fusion, *in silico* two hybrid, phylogenetic tree, phylogenetic profile, and gene expression-based approaches were developed to reduce false positive rates (Rao et al., 2014). Therefore, it is common to integrate different experimental and computational repositories to evaluate functional association among protein pairs.

The most biological networks are classified into scale-free networks which include highly connected nodes (i.e. hubs). In these networks, hub perturbation can induce a major loss of connectivity that results in network disturbance (Albert, 2005). Hubs are key players in several biological processes because of this property. The studies concerned with model organisms have presented that hub proteins are more often associated with essential genes. The genes encoding hub proteins also tend to be more conserved and evolve more slowly than ones encoding non-hub proteins (Jeong et al., 2001 and Fraser et al., 2002). Therefore, targeting hubs could result in considerable disturbances to an overall PPI network when it is compared to targeting non-hub proteins.

Cancer is a multifactorial disease resulting from dysregulation of multiple signaling networks. Thus, determining the underlying signaling network changes in cancer is not only required to uncover the molecular mechanism of carcinogenesis but also to classify cancer specific signature of signaling networks that can be applied in diagnosis, prognosis and treatment strategies (Nemunaitis et al., 2007 and Zhang et al., 2009). Cancer networks have been identified as scale-free networks and may contain a functionally important key player called a hub that is linked to a large number of interactors. The hubs of cancer-associated networks could offer the implementation to cancer diagnostic markers, prognostic markers as well as potential therapeutic targets. Thus, identifying hub proteins could be a promising strategy for drug discovery. The significance of hub proteins is also indicated through their potential as biomarkers so targeting the hub could be an effective strategy to increase the efficacy of cancer treatment (Seo et al., 2013).

The first step in exploring the association between PPI networks and cancer is to assess comprehensive and accurate interactome map of human. Although some studies

increase the coverage of human PPI networks and estimate the interactome size and correct for known biases (Venkatesan et al., 2008, Schwartz et al., 2009 and Barabási et al., 2011), human PPI maps still remain incomplete and noisy (Barabási et al., 2011). Therefore, the reconstruction of highly reliable, almost complete PPI network of *Homo sapiens* was one of the aim of this study. Hence, designed algorithm was applied firstly to model organism, *Saccharomyces cerevisiae* and then applied to *Homo sapiens*.

1.1.1.1 Protein-protein interaction network of *Saccharomyces cerevisiae*

The comprehensive identification and characterization of protein–protein interactions (PPIs) and their networks is essential in order to systematically define and understand the information transfer within and between biological processes. As a result of the accumulation of genome sequence information and development of different genomic and proteomic experimental technologies for identification of PPIs (Zhu and Snyder, 2003), enormous amounts of data have been generated and databases storing the experimentally achieved protein–protein interactions (Aranda et al., 2010; Ceol et al., 2010; Güldener et al., 2006; Salwinski et al., 2004; Stark et al., 2011 and Szklarczyk et al., 2011) have expanded significantly during the last decade. Unfortunately, the accuracy of these data is often criticized, since PPI data obtained from high-throughput experiments is thought to have a large number of false positives, i.e. the interactions that are spurious and do not occur in the cell (Mrowka et al., 2001 and von Mering et al., 2002). Especially the large amount of data produced by the high-throughput experimental methods need to be analyzed and verified before employment in biological research, otherwise these false positives will lead to spurious discoveries that can be potentially costly, e.g. wrong signaling routes, or wrong drug targets for diseases (Arga et al., 2007 and Tekir et al., 2009). Validation of PPI data is a difficult task.

A number of approaches have been proposed to assess the probability that an interaction is a biologically relevant true positive (von Mering et al., 2002; Iossifov et al., 2004; Chen et al., 2002; Deane et al., 2002; Deng et al., 2003; Edwards et al., 2002; Ewing et al., 2007; Giot et al., 2003; Jansen et al., 2002; Krogan et al., 2006; Kuchaiev et al., 2009; Parrish et al., 2007; Patil and Nakamura, 2005; Qi et al., 2005; Saito et al., 2003; Sambourg and Mieg, 2010; Scott and Barton, 2007; Sharan et al., 2005; Sprinzak et al., 2003; Titz et al., 2008; von Mering et al., 2005; Yamanishi et al., 2004; Yu and

Finley, 2009). First attempts were on designing methods taking into account an individual biological feature. The expression profile reliability method was based on the observation that interacting proteins are co-expressed (Deane et al., 2002). Two distributions of expression distances were predefined for non-interacting and reliably interacting protein sets. The distribution of expression distances for a protein dataset of interest was assumed to be a linear combination of two predefined distributions and the linearity coefficient was used in characterization of the accuracy of a given dataset. However, this method was useful for only datasets of proteins and is not applicable for individual interactions. A paralogues verification method by use of sequence homology information was also proposed. An interaction was judged to be true if the concerned proteins had paralogues that interact as well. But their results were limited by the number of proteins that have known paralogues. The structural information of proteins was also employed to assess the quality of interactions (Edwards et al., 2002). Although this method has a high reliability, it is strictly limited by the number of 3-D protein structures available in databanks. von Mering and co-workers (2002) identified the interactions that are observed in more than one high-throughput experiment and used them to estimate the fraction of true positives in the datasets. They hypothesized that if an interaction is supported by several methods, it should be more likely to be true. However, this idea results in ignoring true interactions that are only supported by one technique, e.g. transient and rare interactions. In addition, the overlaps between data-sets were surprisingly small to come with a conclusion. Protein localization method was also proposed (Sprinzak et al., 2003), which defined true positives as interacting proteins that are localized in the same cellular compartment and/or interacting proteins that are annotated to have a common cellular role. Their results showed that the accuracy of the data strongly depends on the experimental technique. In the following years, network topology was also used in identification of true interactions (Iossifov et al., 2004; Chen et al., 2002; Saito et al., 2003). Although these methods have low false positive rates, they also have low true positive rate, due to the scale-free topology of protein interaction networks, i.e. the number of proteins with more than one interaction partners is relatively few. Since the performance of studies integrating individual genomic features with protein interaction data were not satisfactory, it has been shown that the use of a combination of genomic features, such as sequence, structure and

functional annotation information, within a Bayesian network (Jansen et al., 2002; Patil and Nakamura, 2005) or logistic regression modeling framework (Yu and Finley, 2009) results in a more accurate prediction of the true interactions in high-throughput datasets. Moreover, MINT (Ceol et al., 2010), STRING (Szklarczyk et al., 2011), and HitPredict (Patil et al., 2011) databases provide scores that were calculated by using several biological features.

A disadvantage of most of the reported scoring schemes is that they work within a dataset obtained from a single type of experimentation or database. The scores designed for different datasets are mostly not comparable to each other, and for the interactions provided at other databases, a score assignment is not possible. Another disadvantage of these scoring approaches is that they use training data that consists of interactions that are only assumed to be true. Since these datasets are constructed based on simple assumptions, the approaches may produce training sets that are biased. A major disadvantage of all reported scoring approaches is that the high confidence datasets resulting from these scoring schemes are poor in terms of coverage and completeness, i.e. these datasets do not represent an interacting network, but represent numerous clusters of interacting proteins.

In spite of various studies on assessing the quality of protein–protein interaction data, a consensus – highly reliable, almost complete – PPI network of yeast is not presented yet. A dataset of high accuracy will not be very useful in biological research if its coverage is low. The main objective of this part of study was to fill up this gap by integrating multiple heterogeneous biological features and linking them onto a huge set of PPIs obtained from assembling of different publicly available databases.

1.1.1.2 Protein-protein interaction network of *Homo sapiens*

The biological function of a protein is usually determined by its physical interaction with other proteins. Understanding the disease pathogenesis of a biological system requires the knowledge of these proteins and their interacting partners. Even though the data for human protein interactions is incomplete, which presents restrictions to studies in research areas such as disease networks, the continuously increasing list of protein interactions has allowed scientists to better identify roles of these proteins and help in the global network analysis and identify protein biomarkers of a given disease which

leads us towards a better understanding of complex human diseases. Sevimoglu and Arga (2014) have compiled a list of studies regarding the role of protein interaction networks in systems biomedicine.

There are probably about 20500 genes identified by the Human Genome Project (www.genome.gov). The experimental methods developed to measure physical interactions between these identified proteins are also increasing in number. Two major techniques for mapping protein interactions are: yeast two-hybrid (Y2H) and affinity purification technique followed by mass spectrometry. In addition to these methods there are other methods such as immunoprecipitation and crosslinking (Berggard et al., 2007). There are publicly available PPI databases that collect and store interaction data from these experimental studies such as, HPRD (the Human Protein Reference Database) (Mathivanan et al., 2008), DIP (the Database of Interacting Proteins) (Salwinski et al., 2004) and IntAct (the IntAct molecular interaction database) (Hermjakob et al., 2004).

Even though the protein interactions mapped by these and other databases are increasing rapidly the noisiness of the data (false positives and negatives) cannot be overlooked. Techniques ranking the reliability of protein interactions are needed in order to construct a more reliable human protein interaction network (Karagoz and Arga, 2013). Chen et al. (2006) used network topological metrics to increase the confidence of protein interactions in yeast, fruit fly and worm. Patil et al. (2005) used a combination of three genomic features (domain domain ineractions, Gene Ontology annotations and sequence homology). Wu et al. (2010) proposed a two-step method to integrate biological and computational sources to achieve reliable PPIs. First they grouped PPIs to estimate the confidence scores for each biological or computational evidence source then classified these interactions to identify reliable interactions. HomoMINT is an online database for inferred network of protein interactions where interactions discovered in model organisms are mapped onto the corresponding human orthologs. InParanoid algorithm in conjunction with a string matching algorithm is used to construct the HomoMINT network and the inferred network is compared to an experimental interactome as well as a network formed by literature survey for validation (Persico et al., 2005). Another online database, HINT (High-quality INTeractomes) uses a filter system to eliminate the low quality interactions by considering the number of

experimental studies carried out to map these interactions as well as the number of interactions reported by each unique publication. They rank the publications in decreasing order of interactions reported per study (Das and Yu, 2012). BioGRID (the Biological General Repository for Interaction Datasets) includes interaction data from high-throughput experiments such as Y2H for binary interactions and MS analysis for purified complexes. To overcome the problem of false positives and negatives they have also included literature derived interactions from Pubmed called the "gold standart" (Stark et al., 2006). BIND (Biomolecular Interaction Network Database) stores information about interactions, molecular complexes and pathways and accepts interactions with as little as one publication reference to PubMed (Bader et al., 2001). HIPPIE (Human Integrated Protein-Protein Interaction Reference) is a human PPI dataset that uses a normalized scoring scheme integrating various experimental datasets (Schaefer et al., 2012). STRING is also an online database that uses confidence scores from three genomic associations (gene fusion, co-occurrence of proteins in a metabolic map and neighborhood evidence) to rank the reliability of protein interactome. This database maps interactions based on experimental studies as well as prediction algorithms.

Not all the interactome databases use a scoring system to map their interactions. Moreover the amount of data and content also differs from one database to another. Additionally the interactome data from different sources are usually non comparable. To get a more comprehensive reliable interactome one should combine the available databases but also use a scoring scheme that includes a wide variety of biological features. A more reliable human protein interactome is needed for a better understanding of the global human PPIs as well as disease networks such as different types of cancer and biological processes. The aim of this study is to achieve a high confidence human interactome by incorporating eight biological features with an extensive group of PPIs obtained by compiling data from various human protein interaction databases.

1.1.2 Transcriptional regulatory networks

Transcriptional regulatory networks are represented as directed graphs which describe the regulations of TFs on genes' expression by binding to target gene. The nodes

represent genes and a directed edge from one to the other demonstrates the regulatory effect of the first gene on the second. The edges are expressed by the mode of regulation such as up-, down-, or dual-regulation (Figure 1B). Some of the methods that identify the regulatory relationship between TFs and genes are ChIP-chip, genome-wide RNA interference and DNase I footprinting assay (Blais and Dynlacht, 2005). These experimental systems have limitations on mass sample. Therefore, forward engineering approaches have been applied to search TF binding sites in target sequences but this method has also disadvantage that has higher false positive rates. To decrease the false positives, integrative method was reported which was based on gene expression correlation (Yu et al., 2012; Li et al., 2014).

The research showed that human transcriptional networks are highly cell selective (Neph et al., 2012) and transcriptional regulatory networks are also defined as a scale-free network, including a few hubs and many other nodes with few connections, in which hubs have a key role to control cellular response to a distinct stimulus (Blais and Dynlacht, 2005). In addition, TF mutations have cascade impacts on many TF target genes which take roles in various biological processes (Li et al., 2014).

A proper regulation of gene expression is required for normal cellular growth and abnormal expression induces aberrant cell differentiation and proliferation that leads diseases including initiation and progression of carcinogenesis. The improper alterations in transcriptional regulations arise during tumor progression and metastasis through gain- or loss-of-function of transcription factors (Ell and Kang, 2014).

The dysregulation of oncogenes and tumor suppressors lead to increased cellular growth and survival during early stages of carcinogensis. TFs do not only play role in tumor formation but also regulate metastasis. For instance, MYC was reported as a dysregulated TF in 30% of all human cancers and MYC as an oncogene is involved in tumor formation and metastasis (Bernards and Weinberg, 2002; Dang, 2012). In contrast to oncogene TFs, there are also tumor suppressor TFs which act like metastatic suppressors including KLF4 as a tumor suppressor TF, inhibits cell growth (Chen et al., 2001). Thus, it is essential to elucidate the spatial and temporal regulation of TFs during tumor formation and progression due to their known roles in metastasis.

Metastasis associated TFs have been presented to govern resistance to chemotheraphy and apoptosis so targeted metastatic TFs could offer potential and many anti-metastasis therapeutics targeted metastatic transcriptional regulatory networks are still under development. To control the spread of cancer to other organs, downstream effectors of TFs are targeted but that kind of theraphy is unefficient to treat metastatic cancer. In addition, the treatment strategies that perturb the transcriptional regulatory network during cancer progression have potential to reduce the function of several TFs and downstream effectors simultaneously which increase the efficiency of therapeutics (Ell and Kang, 2013).

In this thesis, transcriptional control mechanisms including TFs, downstream effectors, upstream regulator and disease driven pathways of breast cancer, triple negative breast cancer and esophageal squamous cell carcinoma have been presented and further discussed.

1.1.3 Metabolic networks

Metabolic networks are represented as directed graph which presents chains of chemical reactions. The metabolites are represented by nodes and the connection between two metabolites represented by an edge. The direction of an edge shows whether the compound is a substrate or product of the reaction. The metabolic networks offer an integrated aspect of cellular metabolism through combining the data obtained from literature together with genomic, proteomic, metabolomic datasets (Feist et al., 2009). Nuclear magnetic resonance (NMR) and mass spectrometry (MS) are experimental techniques which are generally performed in metabolic studies (Malet-Martino and Holzgrabe, 2011).

Metabolic networks are comprised of interconnected biochemical reaction chains that are required to sustain the lives of organisms. These chains of reactions can synthesize larger compounds from simpler compounds so called anabolism or partition compounds into simpler ones to produce energy so called catabolism. Metabolic networks signify particular metabolic pathways, depending on the products of each group of reactions which govern a specific metabolic process.

Responses of cellular mechanisms to genetic and/or environmental distractions are generally regulated by metabolic alterations. These changes are occured through transcriptional changes caused by multifactorial control mechanisms regulating the activity of distinct metabolic pathways. Patil and Nielsen (2005) used these information to uncover signatures in the metabolic network and their algorithm identifed reporter metabolites around which the most significant transcriptional changes emerged and a group of connected genes that control cellular response to genetic or environmental disturbances. They reported that cells altered the expression patterns of many genes as responses to perturbations and these genes also played roles in the particular parts of metabolism in which a perturbation is presented. Then these alterations were spread out through the metabolic network due to the highly connected nature of cellular metabolism.

We identified the cancer specific reporter metabolites and their associated metabolic pathways.

1.2 Breast Cancer

Breast cancer is the most common malignancy among women with an estimated 1.38 million new cases and 458,000 deaths worldwide every year although the advances in our understanding of the cellular mechanisms of breast cancer and early detection methods together with progress in chemotherapies, surgery and radiotherapy provide considerably improved survival (Ferlay et al., 2010; Eccles et al., 2013). According to the American Cancer Society reports (2015), breast cancer is the most frequently diagnosed cancer and second leading cause of cancer death among women in United States and approximately 231,840 women will be diagnosed with BC and nearly 40,290 will die with this malignancy in 2015 (Siegel et al., 2015). Breast cancer is determined as a heterogeneous disease which has many subtypes and molecular characteristics so that identifying reliable molecular biomarkers and drug targets is the major challenge for successful treatment (Curtis et al., 2012). Surgery with/without radiation is the common treatment choice for localized breast cancer even though chemotherapy or endocrine therapy are applied to control tumor growth and enhance survival (Matsen and Neumayer, 2013). Many clinical circumstances such as tumor size, age, lymph node invasion are taken into consideration to decide the best therapeutic preference for a

breast cancer patient (Hammond et al., 2010). This heterogeneity causes various problems including accurate diagnosis of patients for optimal dosage and duration of treatment and it also makes difficult to understand the tumor biology including the genetic, epigenetic and other factors aberrant signaling pathways for development of new therapeutic strategies (Taherian-Fard et al., 2014).

One of the mostly applied techniques to determine novel biological markers is to identify alterations in gene expression levels of the diseased state that are not occured in healthy tissue through the analysis of high-throughput transcriptome data (Kashyap et al., 2009). There are several studies to investigate gene expression signatures and differentially expressed genes (DEGs) in breast cancer. Current gene expression datasets would be useful tool to elucidate the molecular mechanisms during tumorigenesis and to achieve hypotheses on accurate predictors. The comparison of gene expression profiles of stromal cells stemmed from invasive breast cancer was done by microarray techniques to highlight transcriptional alterations in stromal cells during tumor progression. The findings indicated that the tumor microenvironment showed particular properties according to type of tumor leading survival-predictive value. It was reported that extracellular matrix (ECM) encoding genes, COL11A1, COL10A1, COMP, MMP11, FN1, MFAP2, TNXB and MATN2 were associated with breast cancer stroma and they were suggested as a potential stromal markers of tumor prognosis (Planche et al., 2011). Clarke and coworkers used 13 microarray gene expression studies and applied "weighted gene co-expression network analysis" to identify co-expressed gene clusters from expression data sets and they demonstrated 11 co-regulated gene clusters and correlation of these clusters to tumor grade, survival and breast cancer molecular subtypes. Ubiquitin-conjugating enzyme E2S (UBE2S) was also suggested as a prognostic candidate in their study (Clarke et al., 2013). Another transcriptome analysis used in that study focused on metabolic profile of human breast tumors and discovered 2-hydroxyglutarate as a key metabolite which was determined at high concentration in human breast cancer cell lines. It was reported a relation between high 2-hydroxyglutarate levels and MYC pathway activation in breast cancer. Their results pointed out 2-hydroxyglutarate as a potential breast cancer specific metabolite related with MYC activation and poor prognosis (Terunuma et al., 2014).

Gene expression signatures which were reported for predicting prognosis in breast cancer patients, varied substantially between different data sets (van't Veer et al., 2002 and Staiger et al., 2012). Several reasons caused these differences including technical variations in different types of microarray platforms, diversities in samples or their treatment procedures in laboratories, or computational reasons due to the high dimensionality of the data integrated with small sample size (Fröhlich, 2011). As a consequence, obtained gene expression profiles were unlikely to decipher from a biological point of view. Hence, integrating secondary information such as PPI networks or biological pathways repositories to gene expression profiles was suggested to reduce variability and enhance the efficiency of prognostic signatures. Several rearch groups achieved higher accuracy in prediction of outcomes together with higher stability of the signatures through integration of transcriptome data with PPI network data (Chuang et al., 2007; Taylor et al., 2009; Fröhlich, 2011; Staiger et al., 2012). Chuang and colleagues (2007) applied a protein network- based approach to determine breast cancer biomarkers which were not assumed as individual genes but as subnetworks gained from interactome databases. Although the known breast cancer genes i.e. HER-2/neu, P53 and PIK3CA were not observed via gene expression analysis, these genes were shown to be a key part of subnetworks in which many DEGs were interconnected by these genes. Moreover, the constructed subnetworks indicated specific pathways involved in cancer development. They proposed that the reproducibility and accuracy of subnetwork markers was higher to classify metastatic and non-metastatic tumors. Taylor and coworkers (2009) investigated the dynamic modularity in PPI networks to predict the outcomes of breast cancer so they identified and then classified hub proteins according to their co-expression with their interactors in a tissue-specific manner. They also proposed targeting hubs in interactome for anticancer therapies. Consequently, alterations in modularity of interactome could be useful to characterize tumor phenotype which could ascertain patient prognosis. Sanz-Pamplona and coworkers (2012) implemented network-based approach to uncover organ-specific metastases of breast cancer by using transcriptome and interactome. They extracted most differentially expressed genes in different organs and obtained organ-specific functional proteins in brain (19 proteins), bone (23 proteins), liver (15 proteins) and lung (9 proteins). Then organ-specific PPI networks were reconstructed

that the proteins in brain, bone and liver metastasis associated networks involved in signal transduction, immune response and glucose metabolism, respectively. As a conclusion, they proposed selected significant hub proteins as metastatic markers or therapeutic targets (Sanz-Pamplona et al., 2012). The most recent study searched the PPIs between the proteins associated with breast cancer core and candidate genes. The authors demonstrated 599 novel high-confidence interactions among 44 core, 54 associated and 96 novel breast cancer proteins. The results displayed that their system-level method was a robust interpretation to highlight novel candidate proteins as well as uncover the mechanisms of those proteins whose previous roles were unknown in breast cancer. The effectiveness of their approach was shown as a validation of two breast cancer associated proteins involved in DNA damage response as an outcome particular rewiring interactions. The presented breast cancer related PPI network could be used as a template to incorporate clinical and molecular data and design novel treatment strategies (Arroyo et al., 2015).

Abnormalities in gene expression leads aberences in transcriptional regulatory networks that in turns result in human cancer including breast cancer and the studies concerned with transcriptional control mechanisms of breast cancer are still limited. Transcriptional control mechanisms lie at the center of many molecular associations driving the cellular pathways (Emmert-Streib et al., 2014). Transcription factors are crucial cellular compounds which control gene expression in turns their activities have impacts on cellular functions and responses to the environment. They also either activate or repress the activities of their targets (Venter et al., 2001). A few transcriptional factors form transcriptional cascades that affect several cellular processes. Transcriptional factors also control the entire transcriptional regulatory networks for particular phenotypes (Tovar et al., 2015).

Master regulators so called transcription factors were identified to classify breast cancer metastasis and applied approach presented the upstream transcriptional factors of the signature which were relevant with the network topology (Lim et al., 2009). Another study revealed the role of transcription factors MEF2C and MNDA as key master regulators in primary breast cancer and the potential connection between malignancy and metabolic dysfunction (Baca-López et al., 2012). Most recently, Tovar and coworkers (2015) ascertained the transcription factors (i.e. AGTR2, ZNF132, TFDP3

and others) as master regulators in this transcriptional regulatory network in breast cancer (Tavor et al., 2015).

Metabolomics allows designing novel strategies to identify predictive, diagnostic and prognostic markers of breast cancer. The metabolic signatures of breast tumors are distinct from healthy mammary epithelial tissues. The breast cancer cells which have resistance to therapeutics can reprogram their cellular metabolism. Higher prevalence and mortality rates of breast cancer indicate critical demands for biological markers which should be specific and sensitive as well as be measurable in high-throughput and also cost effective manner. In clinic, biomarkers can serve to determine individuals at risks, to follow disease progression and to evaluate the response to therapy (Mamas et al., 2011 and Vermeersch and Styczynski, 2013).

In breast cancer studies, metabolomics is used to ascertain alterations in breast cancer cells and classify breast cancer into subtypes (Curtis et al., 2012 and Cadoo et al., 2013). Moreover, several investigations have reported the importance of glycolytic flux in oncogenesis of breast cancer (Munoz-Pinedo et al., 2012 and Vander Heiden, 2013). Budczies and coworkers (2013) compared the metabolisms of estrogen receptor positive and negative breast cancer cells and reported changes in glutamine and beta alanine metabolism in estrogen receptor positive (Budczies et al., 2013). Another research group found a set of metabolites (i.e. acetoacetate, pyruvate, glycerol, histidine, glycoproteins) which could help to distinguish patients with metastatic breast cancer from the patients with localized breast cancer (Jobard et al., 2014). Metabolomics of non-invasive breast cancer was thoroughly in another study (Zhang et al., 2013). Moreover, Qiu and colleagues identified lipid panel to discriminate breast cancer patients from healthy controls (Qiu et al., 2013). The comparisons of breast cancer tissues to healthy ones indicated the alterations in central carbon metabolism (Budczies, 2012).

The identification of key metabolites offers determination of disease specific significant metabolic pathways. Several breast cancer studies in cells, tumors, serum or urine uncovered the metabolic pathways which underlied mechanisms of breast cancer progression under particular anti-cancer therapeutics (Asiago et al., 2010; Cao et al., 2013; Tsai et al., 2013; Bayet-Robert and Morvan, 2013; Jager et al., 2011). All of the

mentioned studies strengthened that metabolomics could be used as an effective tool to classify breast cancer subtypes as well as to provide the information concerned with biological pathways in response to anti-cancer therapeutics.

In this study, the network-based approaches were carried out to elucidate the mechanisms of breast cancer. Using transcriptome profiles of 428 samples from Gene Expression Omnibus, DEGs were obtained and common DEGs among datasets so called "core DEGs" among gene expression datasets were used as a template to reconstruct breast cancer specific PPI networks. The hub proteins were identified as results of topological analysis of interactome networks that resulted in 25 interactome-based protein signature of breast cancer. Transcriptional regulatory networks were also reconstructed through regulatory relations of core DEGs then driving molecular pathways were identified. Metabolic analysis also done by using transcriptome data to reveal reporter metabolites and associated metabolic pathways played role in breast cancer mechanisms. The applied approach was used to extract relevant biological information about deregulated functions and the relationship between them, as well as to select molecules that could be potential metastatic markers or therapeutic targets.

1.3 Triple Negative Breast Cancer

Breast cancer is intricate and heterogeneous in terms of its cellular source, disease evolution, mutations, metastatic capability, response to treatment, and clinical consequences (Coughlin and Ekwueme, 2009). It was the most commonly diagnosed types of cancer among women in 2013, ranked first as the leading cause of cancer-related deaths among women aged 20 to 59, and is alone expected to account for 29% of all new cancer cases among women in United States (Siegel et al., 2013). Four main molecular subtypes of breast cancer were characterized based upon gene expression analyses: luminal A, luminal B, human epidermal growth factor 2 (HER2) - enriched, and basal-like (Sorlie et al., 2001). The basal-like subtype has attracted more attention because of high frequency, and lack of effective targeted therapy. Triple negative breast cancer (TNBC) is more associated with basal-like breast cancer that is characterized by lack of expression of both estrogen receptor (ER) and progesterone receptor (PR), together with absence of human epidermal growth factor 2 (HER2) (Ismail-Khan and Bui, 2010). TNBC constitutes approximately 15% of breast cancer cases (Rastelli et al.,

2010). It is primarily encountered in younger women of African American or Hispanic descent. Furthermore, the tumors tend to be high grade and when it is recurrent, it spreads out to other organs such as brain, liver, and lungs so that it is typically distinguished by high metastatic and mortality rate among other types of breast cancer (Griffiths et al., 2012). Although important improvements in treatment of breast cancer have been made, such as the targeted therapies for ER + and HER2 + breast cancers (Baselga et al., 2012), limited treatment options are available for patients with TNBC (Turner and Reis-Filho, 2013). Thus, TNBC is still linked with poor prognosis due to its chemosensitivity. Ongoing targeted molecular therapies for TNBC have been focused around three types of agents: i) antiangiogenic agents, ii) epidermal growth factor receptor (EGFR) inhibitors, and iii) poly (ADP) ribose polymerase (PARP) inhibitors (Turner and Reis-Filho, 2013). A limited number of studies report findings on gene expression profiling of TNBC. Yang and co-workers (2011) performed a microarray study comparing TNBC with nontriple negative breast cancer (non-TNBC). Overexpression of several wingless-type signaling (WNT) pathway genes, such as frizzled class receptor 7 (FZD7), low density lipoprotein receptor related protein 6 (LRP6), and transcription factor 7 (TCF7), have been observed in TNBC. Their finding suggested that FZD7 involved canonical WNT signaling pathway was essential for tumorigenesis of TNBC, and thus, they proposed FZD7 as a promising biomarker and therapeutic target for TNBC. In another study, molecular markers defined by pathology immunohistochemical staining were assessed to identify potential therapeutic targets specific to the breast cancer subtypes (Abramovitz et al., 2011). A set of 30 genes were identified for distinguishing pathology-defined subtypes of breast cancer. Furthermore, to address whether gene expression in breast cancer-associated fibroblasts varies between breast cancer subtypes, Tchou and co-workers (2012) compared the gene expression profiles of early passage primary cancer-associated fibroblasts isolated from human breast cancer samples representing three main subtypes: ER+ , TNBC, and HER2 +. They observed significant expression differences between samples derived from HER2+ breast cancer and those from TNBC and ER+ cancers, particularly in pathways associated with cytoskeleton and integrin signaling. More recently, Dey and co-workers (2013) conducted an expression study of primary biopsies from two cohorts, including TNBC tumors, to identify molecular mechanisms of TNBC. They observed

over-representation of canonical WNT signaling components, and implicated transcriptional WNT signaling as a hallmark of TNBC disease associated with specific metastatic pathways. Though the genome expression studies represented significant findings in TNBC, conclusions were not sufficient to uncover the central molecular mechanisms behind the disease; since this requires integration of data from different 'omics' levels (Sevimoglu and Arga, 2014). In the present study, we recruited the high-throughput gene expression datasets within a holistic point of view. Statistical and comparative analyses of datasets led to identification and comparison of differentially expressed genes. Next, protein–protein interaction and transcriptional regulatory networks were reconstructed and analyzed to reveal TNBC-driven pathways, and further elucidate the molecular mechanisms in TNBC, which may identify specific therapeutic targets together with prognostic biomarkers that play significant roles in TNBC therapy.

1.4 Esophageal Squamous Cell Carcinoma

Esophageal cancer arises from the epithelial surface of the esophagus and comprises of two major histologic types: esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (EAC) (Enzinger and Mayer, 2003). ESCC stems from the squamous cells that cover the esophagus whereas EAC originates in granular cells found at the junction of esophagus and stomach, and involves gastro-esophageal reflux (Holmes and Vaughan, 2007). ESCC is the most common subtype of esophageal cancer, which is the sixth leading cause of cancer death worldwide (Kamangar et al., 2006). Moreover, ESCC is directly linked with the ethnicity and geography, and occurs at very high frequencies in China, Japan, Singapore and Puerto Rico (Ke, 2002). Low socio-economic and poor nutritional status, drinking beverages at high temperatures, eating pickled vegetables or mycotoxin-contaminated foods are reported factors leading to ESCC in China and Africa (Yang et al., 2005). On the other hand, heavy smoking and excessive alcohol consumption are the major predispositions to cause ESCC in North America and Europe (Morita et al., 2010). Taken together, ESCC is defined as a multifactorial disease since risks are related to both genetic and environmental factors (Abbaszadegan et al., 2005).

Although many significant advances are made in detection methods, prevention strategies, surgical techniques and chemotherapy, the survival rate is still poor, i.e., five-year survival rate of 19% is fourth worst among all cancers in USA (Siegel et al., 2013). The identification of efficient biomarkers for early detection and better understanding of the molecular mechanisms of ESCC may offer reduced mortality. Furthermore, the proper biomarkers for clinical diagnosis and prognosis have yet not been defined (Qi et al., 2012). Thus, it is imperative to investigate effective biomarkers for early diagnosis of ESCC.

One of the commonly used methods to identify new biomarkers is to find molecular changes of the diseased state that are not observed in normal tissue by analysis of high-throughput profiling data (Kashyap et al., 2009). Several efforts have been made to study gene expression profiles and differentially expressed genes (DEGs) in ESCC and presented gene expression datasets that would be useful to understand the molecular alterations during carcinogenesis and to generate hypotheses on accurate predictors. Hu and coworkers (2010) focused on DNA copy number neutral loss of heterozygosity (CNNLOH) and its relation to gene expression in ESCC (Hu et al., 2010). They observed that ninety percent of ESCCs showed CNNLOH, and expression of 168 genes (containing 1976 SNPs) differed significantly in tumors with loss of heterozygosity (LOH) versus tumors without LOH. As a result, they concluded that CNNLOH might have a profound impact on gene expression in ESCC, which in turn may affect tumor development (Hu et al., 2010). In another study, 159 dysregulated highly statistically significant genes were proposed as potential new targets for early detection and treatment of ESCC (Su et al., 2011). Recently, this group examined the deregulated PLCE1 mRNA expression in ESCC and validated its potential role in early detection and therapy (Li et al., 2014). Yan and coworkers (2012) used a tissue microdissection strategy and microarrays to measure gene expression profiles associated with cell differentiation versus tumorigenesis in twelve cases of patient-matched normal basal squamous (NB) epithelial cells, normal differentiated (ND) squamous epithelium, and ESCC. They identified 12 individual genes unique to the tumor/NB comparison as therapeutic targets, and 10 associated ESCC gene-drug pairs. More recently, a genome-wide screen of long noncoding RNAs (lncRNAs) and coding RNA from ESCC and matched adjacent non-neoplastic normal tissues was conducted (Cao et al., 2014). They

identified differentially expressed lncRNAs that are co-located and co-expressed with differentially expressed coding RNAs in ESCC, and demonstrated that lncRNAs may be involved in lipid metabolism in ESCC.

The information through transcriptome analyses may not be sufficient alone to figure out a complete understanding of the molecular mechanisms involved in the diseases. Nevertheless, proteins control main cellular behaviors and protein-protein interactions (PPI) take a central and crucial role in many biological processes (Sevimoglu and Arga, 2014). Therefore, a multi-omics discovery approach, in which transcriptome data is integrated with proteome and interactome data via reconstructed biological networks, provides the opportunity to highlight the molecular description of the mechanisms underlying diseases and to discover consistent biomarkers of the functional state of a cell. Recently, two independent groups analyzed a previously published transcriptome dataset (Hu et al., 2010) using this strategy. Yuan and coworkers (2014) reported a total of 538 DEGs, and constructed an interaction network around these DEGs. Via topological analysis of the constructed network, they identified highly-connected components (hubs) and proposed that the TPSO gene and its interactors may affect the cancer-specific gene expression by participating in the toll-like receptor signaling pathway in ESCC. In another study using the same experimental data, a total of 1063 DEGs were identified, a co-expression network around DEGs was constructed and topological properties were utilized to identify the candidate genes of ESCC (Shen et al., 2014). The top 24 genes, which were significantly enriched in bio-functions regarding cell differentiation, glucan biosynthetic process and immune response, were proposed as candidate genes of ESCC.

Though the genome expression studies represented significant findings in ESCC, conclusions were not sufficient to uncover the central molecular mechanisms involved in the disease due to (i) sampling from a narrow population, (ii) identification of divergent sets of genes associated with ESCC in each study, and (iii) lack of holistic point of view which requires integration of data from different ‘omics’ levels. Considering these drawbacks, we designed a comprehensive methodology in the present investigation that takes into account the gene expression datasets from five different studies to evaluate the molecular changes of the diseased state, which are not observed in normal tissue. Statistical and comparative analyses of datasets led to identification

and comparison of DEGs. To highlight the complex molecular mechanisms of ESCC, we focused on the receptors, transcription factors and tumor suppressor genes (TSGs) among the DEGs. Moreover, through the enrichment analysis based on DEGs, the major biological pathways and processes were delineated in order to determine novel factors contributing to onset and progression of ESCC. Next, we employed the common DEGs of ESCC as a template to construct PPI and transcriptional regulatory subnetworks. Topological analysis of the reconstructed networks revealed central proteins, ESCC-driven pathways and molecular mechanisms, which may provide specific therapeutic targets together with prognostic biomarkers that play significant roles in response to ESCC therapy. Finally, we also examined the metabolic alterations in ESCC, which often occur through transcriptional changes caused by complex regulatory mechanisms adjusting the activity of various metabolic pathways (Patil and Nielsen, 2005) and identified reporter metabolites around which the most significant transcriptional changes were exerted. Therefore, a comparative analysis between the active metabolic networks of diseased and healthy cases allowed the identification of ESCC-specific metabolic features that may constitute generic potential drug targets for cancer treatment.

1.5 Aim of the Study

The effectiveness of bioinformatics tools for gene prioritization, network analysis, gene ontology, pathway and disease enrichment analysis as well as all available data on gene expression during carcinogenesis bring a novel and worthful opportunity for an *in-silico* study of the diseases.

The proposed network based approach statistically analyses transcriptome data and integrates with protein-protein interaction, transcriptional regulatory and metabolic networks to elucidate the underlying molecular mechanisms of breast cancer, triple negative breast cancer and esophageal squamaous cell carcinoma. The objective of this study is to enlighten the driving mechanism of these diseases together with figuring out the possible novel biomarkers that can be considered in drug design and cancer therapy in future studies.

2 MATERIALS AND METHODS

2.1 Collection of *Saccharomyces cerevisiae* and *Homo sapiens* Protein-Protein Interaction Data

Five public databases, IntAct (Aranda et al., 2010), MINT (Ceol et al., 2010), MIPS (Güldener et al., 2006), DIP (Salwinski et al., 2004), and BioGRID (Stark et al., 2011), use the results of high-throughput experiments, and collect experimentally or computationally determined PPIs from the literature as well. We have assembled databases of physically interacting protein pairs for *Saccharomyces cerevisiae* from these public databases as available on April 5, 2011. To eliminate the uncertainties originated from the use of different identifiers in databases, yeast systematic names were used as reference. Totally 135 154 pairwise interactions between 6 191 proteins were acquired after elimination of the overlaps and self-interactions.

Interaction data of *Homo sapiens* were collected and assembled from six publicly available databases, BioGRID (Chatr-aryamontri et al., 2014), DIP (Salwinski et al., 2004), IntAct (Orchard et al., 2014), HIPPIE (Schaefer et al., 2012), HomoMINT (Persico et al., 2005) and HPRD (Mathivanan et al., 2008), which store data from several experimental systems. All the protein identifiers were converted to UniProt symbols to exclude the uncertainties arising from the use of distinct identifiers in different databases. After the removal of self-interactions and overlaps between datasets, a total of 295288 pairwise interactions between 15950 proteins were obtained for further analysis.

2.2 Biological Features

A total of eight heterogeneous genomic and proteomic features were used to quantify the probability that an interaction is a biologically relevant true positive: (1) The number of databases reporting the interaction, (2) The frequency of detection by different experimental systems, (3) The number of published papers reporting the interaction, (4) Functional coherence, (5) Biological process coherence, (6) Co-localization, (7) Structural conformity, and (8) Gene expression correlation between interacting proteins.

The information on experimental system used in the detection of PPIs and the publication data (i.e. the PubMed identifiers) were taken directly from its source, i.e. the databases (Aranda et al., 2010; Ceol et al., 2010; Güldener et al., 2006; Salwinski et al., 2004; Stark et al., 2011). Totally 6 230 PubMed identifiers and 75 different experimental systems were employed for *S. cerevisiae.* In case of uncertainty, the experimental technique used in detection of the PPI was manually checked from their publications by use of PubMed identifiers.

Functional coherence, biological process coherence and co-localization analyses were determined based on Gene Ontology (GO) terminology. The GO terms of yeast proteins were obtained from the consortium website as available on April 5, 2011.

To analyze structural conformity of interacting pairs, the protein–domain relationship for yeast proteins were obtained from PFAM (Bateman et al., 2000) and the domain–domain interaction data was directly taken from a previous study (Deng et al., 2002).

In determination of the correlation between the gene expression levels of interacting proteins, a large gene expression dataset was obtained from the SPELL tool (Hibbs et al., 2007) of *Saccharomyces* Genome Database (Cherry et al., 2012). The Pearson correlation coefficient for each protein pair, which ranges from −1.0 to 1.0, was used to enumerate the correlation.

Additionally, a total of 78 different experimental systems as well as 39460 PubMed identifiers were acquired for *Homo sapiens*. In order to avoid any possible uncertainty, unclear statements about the experimental methods were eliminated by manual check of related publications and the technique employed to detect the binary interactions was confirmed.

The Gene Ontology (GO) terms of human proteins were taken from the consortium website (as available on January 15, 2015) to analyze the functional and biological process coherence as well as co-localization. The protein–domain relationship for human proteins were obtained from UniProt (Uniprot Consortium, 2014) and the domain–domain interaction data was directly taken from Integrated Domain-Domain Interaction (IDDI) database (Kim et al., 2012) to evaluate structural conformity.

The correlations between the gene expression levels of pairwise human proteins were identified through gene expression datasets from the Gene Expression Omnibus (Barrett

et al., 2013). The expression values of healthy individuals from five datasets (GSE13355, GSE43346, GSE21510, GSE22255, and GSE21935) were normalized and the Pearson correlation coefficient for each protein pair (range between −1.0 to 1.0) was used to enumerate the co-expression level.

2.3 Statistical Analyses

To statistically describe and analyze confidence score distributions within biological features used in yeast, box-plot representation was employed. Each box represents the 25–75% range of the confidence score distribution. The medians are demonstrated by horizontal lines in the boxes. The lowest and highest data without outliers are shown by "Whiskers" and the stars express mild and extreme outliers. Since the score distributions do not follow the normality assumption, statistical analyses for the differences between the medians were performed by the Wilcoxon–Mann–Whitney test for pairwise comparisons and the Kruskal–Wallis ANOVA test for multiple comparisons.

2.4 Functional Enrichment Analyses

The functional enrichment analyses of the protein encoding gene sets were achieved through The Database for Annotation, Visualization and Integrated Discovery (DAVID) bioinformatics tool (Huang et al., 2009), which presents the enrichment results for molecular pathways, human diseases and Gene Ontology (GO) terms through investigating the well-known publicly available databases including pathway-related databases: KEGG (Kanehisa et al., 2012), Reactome (Croft et al., 2011), and disease related databases: Genetic Association Database (Becker et al., 2004) and Online Mendelian Inheritance in Man (OMIM). The outcomes with p-value < 0.05 were considered as statistically significant.

2.5 Reconstruction and Analysis of Protein-protein Interaction Subnetworks

The interactions associated with proteins corresponding to differentially expressed genes were identified and PPI subnetworks consisting of down- and upregulated genes were reconstructed by using high confidence PPI network of *Homo sapiens*. The PPI networks were converted into undirected graphs G= (V, E), where V is the collection of

nodes representing proteins and E is the edge set representing binary interactions between protein pairs. The topological analysis of the networks was performed via CytoHubba plugin of Cytoscape (Smoot et al., 2011). Two different topological metrics, degree and betweenness centrality were simultaneously employed to define hub nodes. 'Degree' of a protein is determined by the number of its adjacent proteins in network. In biological context, the degree indicates the regulatory relevance of the node that highly connected proteins in signaling network interact with many other signaling proteins. Moreover, shortest path indicates the efficiency of information flow in a network that plays a vital role in communication and transport. 'Betweenness centrality' distinguishes nodes by how often they exist on the shortest path between two other nodes in the network (Girvan et al., 2002). Therefore, nodes with highest betweenness centrality are key nodes which govern the flow of information in PPI network. The top 10 proteins with highest degree and betweenness centrality values were defined as hubs.

2.6 Transcriptome Datasets of Investigated Diseases

To characterize gene expression in breast cancer (GSE9574, GSE15852, GSE26910, GSE39004 and GSE42564), TNBC (GSE17650, GSE18539, GSE27447, and GSE37614) and ESCC (GSE17351, GSE29347, GSE23400, GSE29001 and GSE45350) independent high-throughput screening datasets were downloaded from the Gene Expression Omnibus (GEO) database (Barrett et al., 2013). A total of 428, 358 and 182 samples were investigated in breast cancer, TNBC and ESCC studies, respectively; including 138 normal and 290 breast tumors, 201 non-TNBC and 157 TNBC tumors, and 91 normal and 91 ESCC tumors (Table 2.1).

2.7 Identification of Differentially Expressed Genes

Each dataset was statistically analyzed independently to identify differentially expressed genes (DEGs). Raw data were normalized via quantile normalization using RMA (Bolstad et al., 2003) as implemented in the affy package (Gautier et al., 2004) of R/Bioconductor software (v.2.12) (Gentleman et al., 2004). DEGs were identified from the normalized log expression values using the multiple testing options of linear models for microarray data (LIMMA) method (Smyth, 2004). Benjamini Hochberg's (BH) method was used to control the false discovery rate (FDR) and the disease specific p-

value threshold was maintained to identify significantly expressed genes. The gene–protein associations were obtained from UniProt ID Mapping tool (Uniprot Consortium, 2014) and bioDBnet platform (Mudunuri et al., 2009).

Table 2.1 Microarray datasets employed in the present study.

GEO Series	# of Samples	Platform	Reference
GSE9574	15 Normal, 14 Breast Tumors	Affymetrix Human Genome U133A Array	Tripathi et al., 2008
GSE15852	43 Normal, 43 Breast Tumors	Affymetrix Human Genome U133A Array	Pau et al., 2010
GSE26910	6 Normal, 6 Breast Tumors	Affymetrix Human Genome U133 Plus 2.0 Array	Planche et al., 2011
GSE39004	57 Normal, 123 Breast Tumors	Affymetrix Human Gene 1.0 ST Array	Terunuma et al., 2014
GSE42568	17 Normal, 104 Breast Tumors	Affymetrix Human Genome U133 Plus 2.0 Array	Clarke et al., 2013
GSE17650	132 non-TNBC, 45 TNBC	Emory Univ. Custom BC DASL Panel	Abramovitz et al., 2011
GSE27447	14 non-TNBC, 5 TNBC	Affymetrix Human Gene 1.0 ST Array	Yang et al., 2011
GSE37614	13 non-TNBC, 6 TNBC	Illumina Human HT-12 v4.0 BeadChip	Tchou et al., 2012
GSE18539	42 non-TNBC, 101 TNBC	Emory Univ. Custom BC DASL Panel	Dey et al., 2013
GSE17351	5 Normal, 5 ESCC	Affymetrix U133 Plus 2.0 Array	Unpublished data
GSE20347	17 Normal, 17 ESCC	Affymetrix U133A 2.0 Array	Hu et al., 2010
GSE23400	53 Normal, 53 ESCC	Affymetrix U133B Array	Su et al., 2011
GSE29001	12 Normal, 12 ESCC	Affymetrix U133A 2.0 Array	Yan et al., 2012
GSE45350	4 Normal, 4 ESCC	Agilent-SurePrint G3 Human GE 8x60K Microarray	Cao et al., 2014

2.8 Transcriptional Regulatory Networks

Transcriptional interactions between transcription factors (TFs) and their downstream effectors in the regulatory network were gained from three different resources: TRED (Transcriptional Regulatory Element Database) (Jiang et al., 2007), GENOMATIX

(Genomatix Software Inc, Ann Arbor, MI, USA), and HTRIdb (Human Transcriptional Regulation Interaction Database) (Bovolenta et al., 2012). Moreover, upstream regulators of hub TFs were identified and their networks were constructed via IPA Bioinformatics tool (Ingenuity System Inc, USA) (Kramer et al., 2014).

2.9 Reporter Metabolites Analysis

The latest version of the genome-scale metabolic model (HMR2) of the generic human cell, composed of 3,765 genes associated with over 8,000 biochemical reactions and 3,000 unique metabolites (Mardinoglu et al., 2013), was integrated with transcriptomics datasets. The statistically significant metabolic alterations were identified using the reporter metabolites algorithm (Patil and Nielsen, 2005) implemented under the BioMet Toolbox (Cvijovic et al., 2010). The metabolites with p–value <0.05 were handled for further analyses. The resultant reporter metabolites from each dataset were compared and common metabolites and their metabolic pathways were investigated through MBROLE (Chagoyen and Pazos, 2011).

3 RESULTS AND DISCUSSION

3.1 A Novel Scoring Methodology for Reconstruction of Highly Reliable Protein-protein Interaction Networks

Considering the assembled dataset of PPIs, the association scores, S_i, of four features, i.e. the number of databases reporting the interaction, the frequency of detection by different experimental systems, the number of published papers reporting the interaction, and structural conformity, were calculated independently using a 4-step methodology. In Figure 2, this methodology is illustrated by calculation of the association scores for one of the biological features (i.e. the number of databases reporting the interaction) as an example. Initially, the frequency distribution that represents the number of feature value associated with the full set of interaction pairs was obtained (Figure 2.1A). For example, 95261 PPIs (70.48%) were reported in a single database, whereas 544 PPIs (0.40%) were reported in all of the databases considered in the study. In the next step, via nonlinear regression, an exponential model was fitted, which represent the probability that a randomly chosen interaction pair has

the corresponding feature value, i.e. the probability that a randomly chosen protein–protein interaction is reported in a certain number of databases (Figure 2.1B). Since the reliability increases with increasing feature value, the preliminary score distribution was considered as the complement of the probability distribution (Figure 2.1C), and the score distribution was normalized to ensure that the minimal value of the feature gives a zero score, i.e. no contribution to confidence (Figure 2.1D).

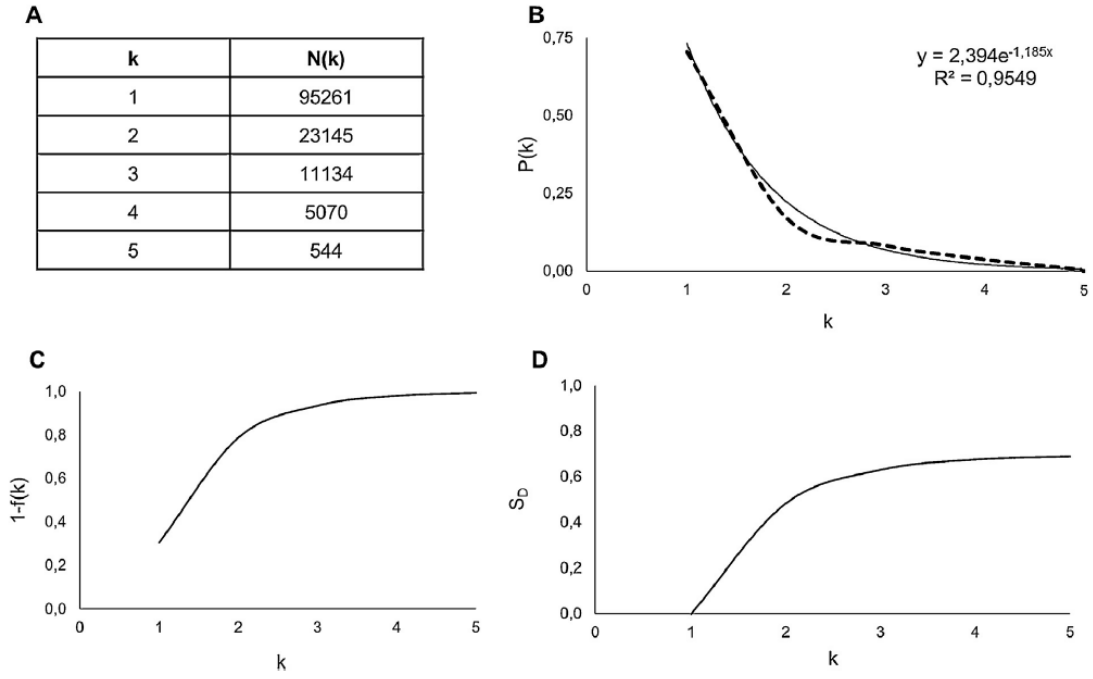


Figure 2.1 The illustration of the 4-step methodology used in the calculation of the association scores of the biological feature: the number of databases reporting the interaction. (A) The histogram representing the frequency (*N*) distribution of binary protein–protein interactions according to the number of databases reporting the interaction (*k*). (B) The exponential fit (straight line) to the data values (dashed line) representing the distribution of probabilities of a randomly chosen interaction pair being reported in a certain number of databases (*k*). (C) The preliminary score distribution obtained from the complement of the probability distribution. (D) The final association score (S_D) distribution.

The association scores for three yeast biological features, i.e. functional coherence, biological process coherence and co-localization, were determined based on GO

semantic similarity analysis (Yu et al., 2010) which computes the similarity of the GO terms annotated to interacting genes/proteins. This method takes into account the overall structure of the GO graph to assess the specificity of shared annotations and computes scores between 0 and 1. In addition, the absolute value of Pearson correlation coefficient was directly used as the score of gene expression correlation.

Naïve Bayesian approach is employed to calculate confidence score, CS, for any interacting pair of proteins so it is a function of the individual association scores:

$$CS = 1 - \prod(1 - Si) \quad (1)$$

Note that, the confidence scores obtained by Eq. (1) are always greater than the corresponding association scores, which signify increased confidence when an interaction is supported by various types of evidence.

3.1.1 Applicability of the method via *Saccharomyces cerevisiae* as a model system

The budding yeast, *S. cerevisiae*, is probably the most intensely studied model organism for functional interactions in eukaryotes, and therefore a huge number of experimental and computational studies have been conducted to detect PPIs in yeast. Five public databases, IntAct (Aranda et al., 2010), MINT (Ceol et al., 2010), MIPS (Güldener et al., 2006), DIP (Salwinski et al., 2004), and BioGRID (Stark et al., 2011), use the results of high-throughput experiments and collect experimentally/computationally determined PPIs from the literature as well. Since the entire PPI data reported in literature is not covered by a single database, in order to achieve the largest possible coverage, we have assembled databases of interacting protein pairs for *S. cerevisiae* from these public databases (as available on April 5, 2011) and obtained an extremely large PPI dataset which comprises of 135154 non-redundant physical interactions between 6191 yeast proteins.

The analysis of the sources of the PPIs in the dataset showed that less than 0.40% of all interactions (544 interactions) were reported in all of the databases considered. On the other hand, the majority of PPIs (70.5% of all) were stored by only one of these databases. These findings indicated a significant difference between the PPI databases, which arise mostly from the use of different resources, confidence sets and threshold

values (Reguly et al., 2006), and confirmed the need for integration of different repositories to increase the coverage.

Generally, it is assumed that the interactions that have been validated by multiple experimental techniques and/or more than one research groups have a reasonable confidence (Reguly et al., 2006). However, majority of the interactions (77%) in the dataset were explored by a single experimental system. In addition, 37 interactions (0.03% of all) were reported in 10 or more articles in literature, whereas the majority of the interactions (86.5%) were reported in a single publication. Analysis of the similarity between interacting protein pairs in terms of GO terms indicated significantly low semantic scores; for example, zero-scores were obtained for almost 66% of PPIs when GO molecular function terms were considered. Similarly, the semantic scores of 18.5% of all interactions were calculated as zero in terms of GO biological process and 12.3% of all interactions displayed a zero semantic score when GO cellular component terms were considered. Since proteins interact with each other through their specific domains, domain–domain interaction information also gives clues on the reliability of the PPIs (Chen et al., 2002; Edwards et al., 2002; Deng et al., 2002; Rain et al., 2001). Considering the 516 domains and 891 domain–domain interactions identified for *S. cerevisiae*, only 2.3% of all interactions were assigned by at least one domain–domain interaction. In terms of the correlation between gene expression profiles of interacting proteins, 4446 interacting pairs (3.3% of all) indicated a correlation coefficient higher than 0.60. All these preliminary analysis results on heterogeneous biological features indicated that each biological feature has its own pros and cons, but none of them is sufficient to define the accuracy of the PPI alone. Thus, the predictive power of the scoring system came from the integration of the multiple features.

Previous experiences on testing the quality of protein–protein interaction data indicated that as a result of the hidden physiochemical constraints inherent to each experimental method, conceptually different methods identify different type of interactions (Yu et al., 2008; Harrington et al., 2008). For example, interaction assays such as yeast two-hybrid system are better at identifying binary interactions; however, purification assays such as tandem affinity purification are better at identifying co-complex interactions, i.e. all proteins that are part of the same complex. Therefore, these experimental methods should not be compared, but should be considered as complementary (Harrington et al.,

2008). In addition, PPI data acquired from high-throughput approaches are reported to include large number of false positives (von Mering et al., 2002; Li et al., 2008). Thus, the interactions that were validated by multiple experimental methods are believed to have a reasonable confidence. Moreover, it is assumed that the interactions published in the literature are reliable since many of them have been validated by multiple experimental techniques and/or more than one research groups (Reguly et al., 2006). However, Cusick and co-workers (2009) suggested that the quality of literature-curated datasets may not be as high as commonly noticed. The extraction of correct information out of long publications is extremely hard and structured vocabularies have not been applied in publications.

The domain–domain interaction data provide significant information to predict unknown PPIs (Deng et al., 2002; Rain et al., 2001). Nevertheless, two proteins may not interact with each other although they have potentially interacting domains because their localization in the cell may be different or the expression of them during cell cycle may be different. Moreover, the incompleteness of current domain–domain interaction data also makes it difficult to predict PPI (Edwards et al., 2002; Deng et al., 2002).

Considering the assembled dataset of PPIs, the association scores, S_i, of eight biological features (Figure 3.1) were calculated independently, and integrated in a Bayesian approach to calculate confidence scores, CS, of all PPIs. The confidence scores indicated a broad distribution (Figure 3.2). Due to the nature of the scoring function, the range of the scores was varying between zero (non-significant interaction indicator) and 1 (interaction with high-confidence), and the confidence scores were always equal or greater than the corresponding association scores, which indicates increased confidence when an interaction is supported by various types of evidence. The confidence scores of 631 interactions (0.47% of all) were almost 1. These include binary interactions which were supported by almost all of the biological features. 29046 (21.50% of all) interactions were greater than 0.85, whereas scores of 86118 (63.72% of all) interactions were greater than 0.50.

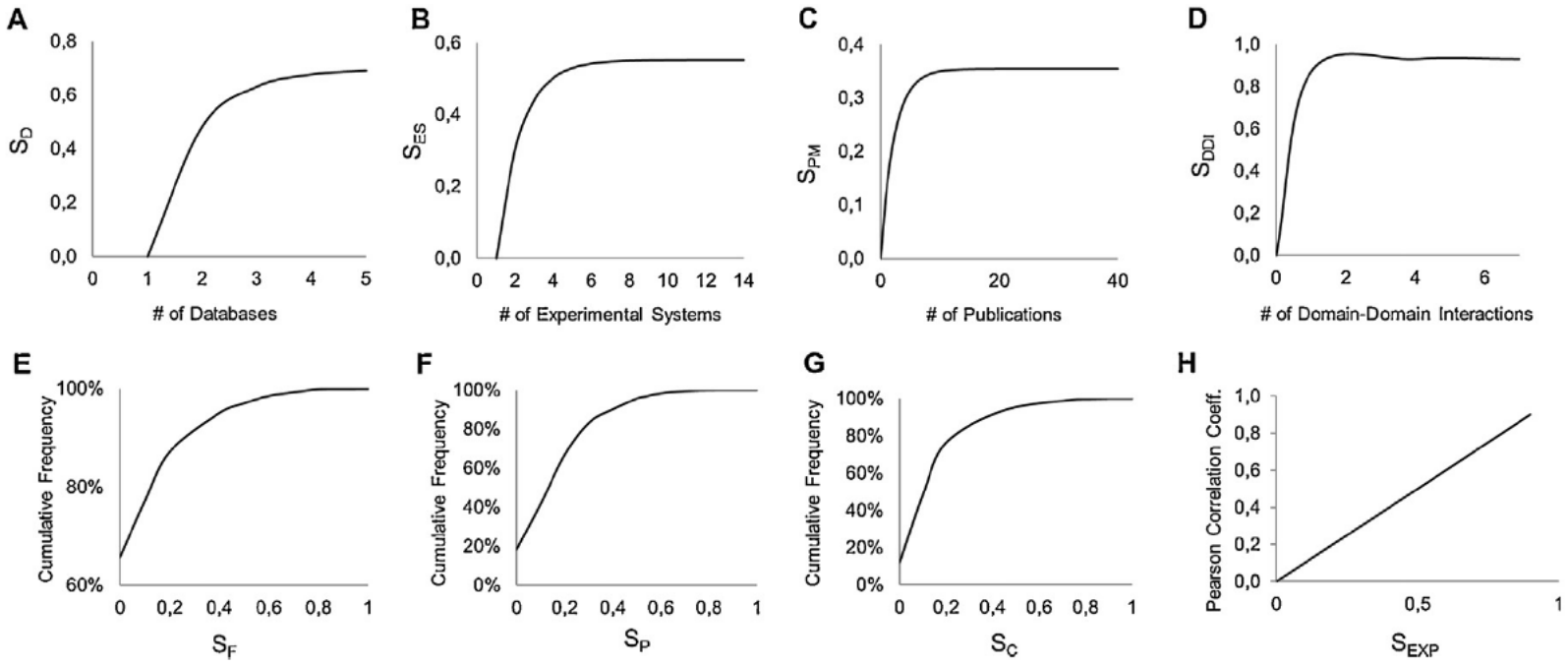


Figure 3.1 The score distributions of association scores for each heterogeneous feature. The association score (S_i) distributions, which were calculated using a 4-step methodology, for: (A) the number of databases reporting the interaction (S_D), (B) the frequency of detection by different experimental systems (S_{ES}), (C) the number of published papers reporting the interaction (S_{PM}), and (D) structural conformity (S_{DDI}). The association score distributions, which were determined based on GO semantic similarity analysis, are represented by cumulative frequencies for: (E) functional coherence (S_F), (F) biological process coherence (S_P), and (*G*) co-localization (S_C). (H) The association score distribution for gene expression correlation (S_{EXP}) as a linear function of Pearson correlation coefficients.

To assess the efficiency of our methodology in score calculation and to identify a high-confidence interaction dataset through the large PPI dataset, the resultant confidence scores were assessed in terms of a trade-off between the sensitivity (i.e. the ability to identify true-positives), the specificity (i.e. the ability to identify true-negatives) and the coverage (i.e. the fraction of all proteins covered by the dataset). The traditional method for determination of sensitivity and specificity is the employment of training data, i.e. the gold standard datasets, which consists of interactions that are only assumed to be true or false.

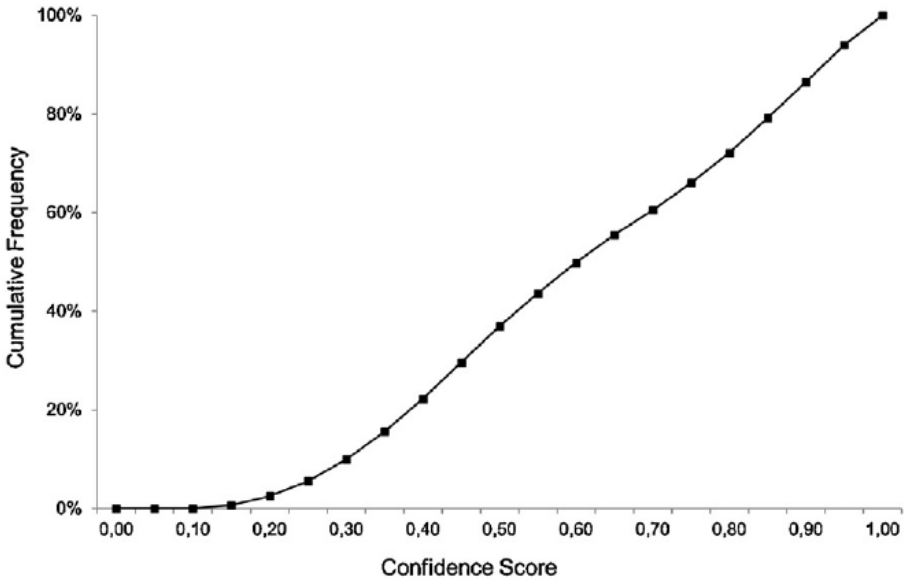


Figure 3.2 Cumulative distribution of confidence scores. Interactions with scores above 0.85 constitute only 21.5% of all and could be considered high-confidence interactions. According to the design of the scoring function, such score implies that the interaction is supported by multiple evidence.

On the other hand, construction of these datasets is a difficult task and requires attention. Since these datasets are constructed based on simple assumptions, the approaches may produce training sets that are biased toward particular type of interaction or highly conserved interactions or particularly well-studied pathways. Previously, the gold standard datasets of PPI were constructed to test the efficiency of

the signaling pathway reconstruction schemes (Arga et al., 2007). To determine sensitivity and specificity, the datasets of the previous study were updated and employed: The gold standard positive dataset consists of 2530 interactions and the gold standard negative dataset includes 2994 PPIs. The gold standard negative dataset consisted of interactions obtained from protein co-localization data in yeast (Jansen and Gerstein, 2004) under the assumption that if the proteins do not exist in the same sub-cellular compartment, they cannot be interacting (Arga et al., 2007; Jansen and Gerstein, 2004). The gold standard positive dataset was composed of physical interactions with very high confidence score (>0.999) in STRING (Szklarczyk et al., 2011), a set of interactions obtained by small-scale experimentation (Güldener et al., 2006), and interactions from the core network of DIP (Salwinski et al., 2004). Although each of the three subsets incorporated in construction of the positive dataset were considered as highly confident in literature, the intersection of all datasets constitute only 7.5% of all (Figure 3.3).

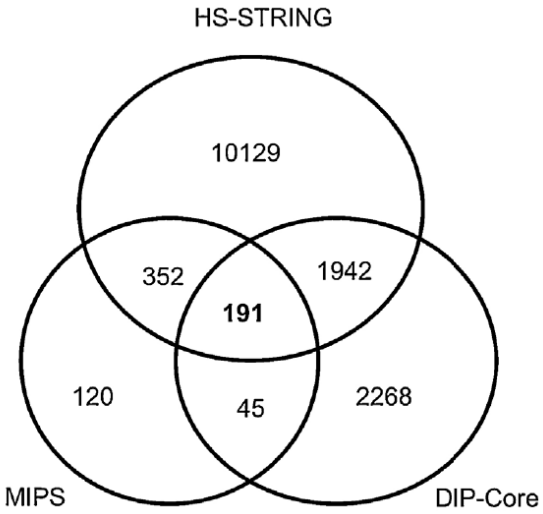


Figure 3.3 The distribution scheme of the interactions in the gold standard positive dataset to their sources. Although each of the three datasets incorporated in construction of the positive set were considered as highly confident in literature, the intersection of all datasets constitute only 7.5% of all.

As a result of the trade-off analysis for different score thresholds (Table 3.1), we determined the confidence score of 0.85 as an optimum threshold for identifying a high confidence dataset (HCS) of PPIs in yeast. This dataset, consisting of 29 046 binary

PPIs, constitute only 21.50% of all interactions in the large PPI dataset and indicates equilibrium between sensitivity, specificity and coverage (86%, 68% and 72%, respectively). When the association scores of these interactions in the HCS dataset was analyzed, it was found that these interactions were supported by multiple evidences, i.e. supported by at least five of the eight heterogeneous features considered in scoring scheme. In addition, a low confidence dataset (LCS) consisting of 49 036 (36.28%) binary interactions with confidence scores lower than 0.50 was also determined. These interactions were supported by weakly at most three of the eight biological features considered in the study.

Table 3.1. Trade-off analysis for different score thresholds.

Confidence score threshold	# of interactions in the network	Sensitivity (%)	Specificity (%)	Coverage (%)
0.99	2310	19.4	97.5	24.0
0.95	8928	45.4	90.3	47.0
0.90	19087	77.2	79.0	62.0
0.85	***29046***	***86.7***	***67.8***	***72.1***
0.80	38559	91.1	59.2	80.0
0.70	54049	94.5	46.5	89.9

To illustrate the predictive power of the scoring scheme, the Receiver Operating Characteristic (ROC) curve analysis was employed to give a more complete picture of the performance of the present scoring scheme when compared to some previous schemes (Figure 3.4).

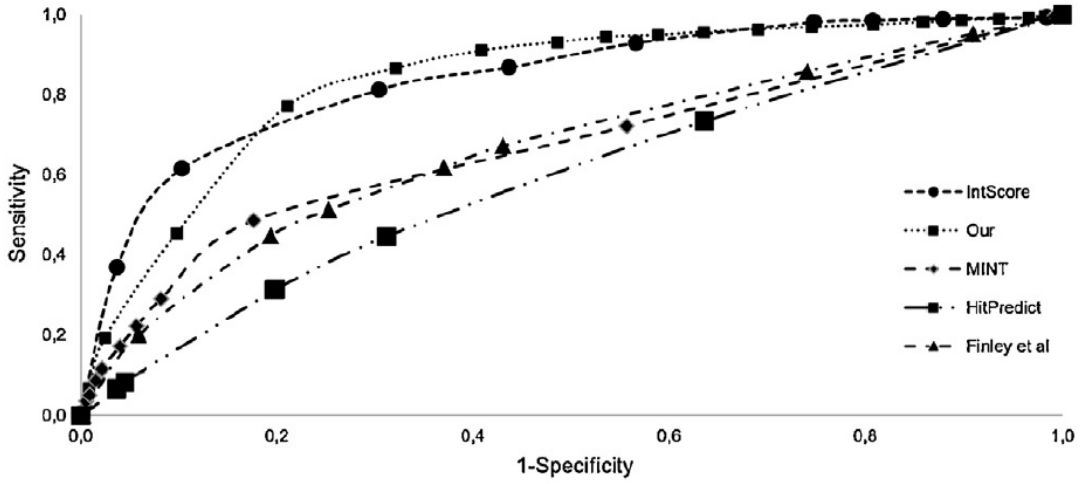


Figure 3.4 ROC curve analysis for comparison of the performance of the PPI scoring schemes under tradeoffs between sensitivity and specificity. The predictive power of the scoring methods was quantified by calculation of the area under curve (AUC) metric.

The present scoring scheme indicates an excellent discrimination in terms of the AUC value of 0.84 according to general rules used in interpretation of AUC values (Lin et al., 2004). The IntScore scoring scheme also indicates excellent discrimination with an AUC value of 0.84; other scoring schemes represents acceptable discrimination with AUC values between 0.59 (HitPredict) and 0.67 (Mint).

The extreme AUC value of the present scoring scheme is a result of many advantages of the present scoring scheme with respect to the reported scoring systems. First of all, it is more comprehensive than the others since an extremely large dataset of PPIs, which is the combination of widely used five PPI repertoires, was employed in the present scoring methodology. The other scoring schemes were based on a single dataset achieved from a single type of experimentation or database. Thus, the designed score for PPIs are mostly not analogous to each other, and for the interactions provided at other databases, a score assignment is not possible.

A major problem with most of the published scoring procedures was the use of training sets, which include only the interactions supposed to be true/false. For instance, Yu and Finley (Yu and Finley, 2009) synthesized four different gold positive training sets based

on interactions that (i) are associated with at least 10 PubMed identifiers in the interaction databases; (ii) are putative conserved interactions, which are those found in common between interaction sets for any two species (fly, human, worm and yeast); (iii) are high-throughput interactions reported to have high confidence by the original publications; and (iv) have expression correlation higher than 0.6. For each training set, a negative set of equal size was synthesized by drawing random samples from the list of all interactions, excluding those in that positive set. On the other hand, these training sets were doubtful since they might be biased toward highly conserved interactions or interactions from particularly well-known pathway or due to random sampling. In this point of view, a major advantage of the present scoring system is that the association scores were calculated considering the full set of PPIs in the large dataset, instead of reconstructing training sets.

Another major problem with other scoring schemes taking into account different types of biological features is the determination of weights assigned to the association scores of each feature in the calculation of the overall confidence score. However, in the present scoring system no artificial weights were assigned, i.e. weights were intrinsic to the scoring function. Application of the methodologies used in the calculation of association score of each feature resulted with different maxima for association scores of each feature (Figure 3.1), which can be considered as weights for each feature.

The present scoring scheme also allowed us to discuss the reliability of experimental systems used in detection of PPIs. Previously, it was claimed that the high-throughput methods, such as protein arrays and yeast-two-hybrid screens require the assessment of the reliability since 30–50% of the high-throughput interactions were known to be biologically relevant (Iossifov et al., 2004). On the other hand, low-throughput methods, such as X-ray crystallography and nuclear magnetic resonance, were believed to give the most detailed information about protein–protein interactions (Shoemaker and Panchenko, 2007). Schaefer and co-workers (Schaefer et al., 2012) assigned scores to various PPI detection methods via manual curation, and in vitro techniques like X-ray crystallography were assigned to the highest score that is 10. Complementation based assays and affinity based technologies were roughly equally scored with an average value of 5. The scores were slightly increased for those methods, such as FRET, that are used generally in homologous, more physiological setups. Methodologies that do not

directly provide evidence for interaction, such as co-localization or co-sedimentation, are scored with lower scores. These reliability scores assigned by manual curation necessarily reflect the individual opinion of researchers to these experimental techniques, therefore we evaluated our resultant confidence scores from this perspective and constructed a box-plot representing the distribution of confidence scores of all interactions characterized by the corresponding experimental system (Figure 3.5).

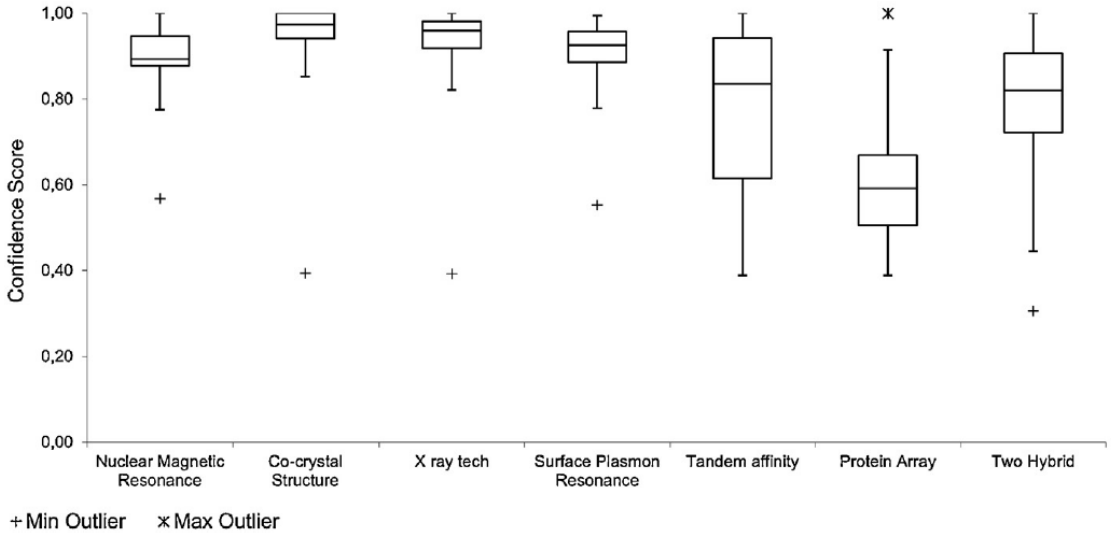


Figure 3.5 Box plot representing the confidence score distribution of experimental systems. The distribution for the high-throughput experimental systems, such as yeast-two-hybrid screen, tandem affinity purification and protein arrays, tend to possess lower median score values and broad score distribution when compared to those obtained by small-scale experimentation (i.e. low-throughput techniques) such as nuclear magnetic resonance, co-crystal structure, X-ray crystallography and surface plasmon resonance.

The score distributions of small-scale experimental systems (i.e. nuclear magnetic resonance, co-crystal structure, X-ray crystallography and surface plasmon resonance) possessed narrow distributions with significantly higher median values (p-value $<10^{-4}$), when compared to the score distributions of high-throughput experimentations (i.e. yeast-two-hybrid screen, tandem affinity purification and protein array), which possessed broad score distributions with significantly lower median scores with respect to the threshold score (p-value <0.01). These results were consistent with the previous discussions (Iossifov et al., 2004; Shoemaker and Panchenko, 2007) and with the

reliability scores of Schaefer and co-workers (2012), where nuclear magnetic resonance, co-crystal structure, X-ray crystallography and surface plasmon resonance methods were assigned to the highest score of 10; on the other hand, complementation-based assays and affinity based technologies, such as yeast-two-hybrid screen, tandem affinity purification and protein array, were roughly equally scored with an average value of 5.

3.2 Reconstruction of a Highly-Reliable Protein-Protein Interaction Network of *Homo sapiens*

Elucidating and analyzing the interacting partners of a protein highlights functional activity of proteins that is required to understand different biological processes ranging from cell cycle to apoptosis. Even though various experimentally and computationally derived protein interaction data of *Homo sapiens* databases exist, the coverage of a single database is not sufficient to demonstrate all reported PPI data in literature. We assembled physical interactions from 6 public databases including BioGRID (Chatr-aryamontri et al., 2014), DIP (Salwinski et al., 2004), IntAct (Orchard et al., 2013), HIPPIE (Schaefer et al., 2012), HomoMINT (Persico et al., 2005) and HPRD (Mathivanan et al., 2008) to investigate and to achieve a more reliable human interactome. 295 288 non-redundant physical interactions between 15 950 proteins were obtained that the majority of PPIs (83.13% of all) were covered by only one of these databases. This result confirmed the necessity to assemble repositories to enhance data coverage. The difference of each database comes from the use of different resources, confidence sets and threshold values (Reguly et al., 2006).

PPIs can be identified by numerous experimental techniques; therefore the reliability of the methods becomes crucial. In this regard, we used experimental scores which were assigned by HIPPIE database which presented the quality of the experimental systems (Schaefer et al., 2012). Reguly and coworkers (2006) reported that PPIs which have been confirmed by more than one research group and/or multiple experimental approaches have higher reliability. 78 different experimental systems with their individual scores were included in this study. Nevertheless, 245 585 (83.1%) PPI were examined by only a single experimental system and/or multiple lower confidence experimental techniques. On the other hand, 1 359 interactions (0.5%) were published in 10 or more articles in terms of literature search.

3 558 molecular function terms, 10 101 biological process terms and 1 256 cellular component terms were collected to analyze similarity between interacting protein pairs. The protein pairs of 60 273 (20.4%) PPI, 55 425 (18.8%) PPI and 73 148 (24.8%) had more than one common functional coherence, biological process coherence and co-localization terms, respectively. The PPIs occur through specific protein domains so domain–domain interaction data also enhances the confidence of the PPIs (Chen et al., 2002; Edwards et al., 2002; Deng et al., 2002; Rain et al., 2001). Domain-domain interaction database presented 7351 domains and 204 715 domain–domain interactions for *Homo sapiens* (Kim et al., 2012) that 94 520 (32.0%) interactions were reported by at least one domain–domain interaction in this research. With regards to the gene expression correlation of interacting proteins, 16 718 interacting pairs (5.7% of all) indicated a correlation coefficient higher than 0.80. These preliminary results of each heterogeneous attributes pinpointed pros and cons of each feature. Nevertheless, the accuracy of PPIs cannot be defined by only one of these features which reveal the importance of the predictive power of our scoring scheme which originated from the assembly of various attributes.

3.2.1 Evaluation of confidence scores

Each of the eight biological features has their own association scores, Si, which was computed independently in consideration of the combined data sources (Figure 3.6). The confidence scores, CS, of all PPIs which was computed by a Bayesian approach, displayed a wide range distribution (Figure 3.7). The scoring range varied between zero (non-significant interaction indicator) and 1 (interaction with high-confidence) by reason of characteristic of scoring function. The confidence scores were always greater than the corresponding association scores, which signified the increased confidence when an interaction is supported by different types of features. The CS of 2 440 interactions (0.82% of all) were almost 1 and 147 923 (50.09% of all) interactions were greater than 0.80, whereas CS of 220 281 (74.60% of all) interactions were greater than 0.50.

The confidence scores were evaluated with regards to a trade-off between the sensitivity, the specificity and the coverage to assess the reliability of our approach in scoring and to define a high-confidence PPI set through the PPI pool. The established

method for calculation of sensitivity (i.e. the ability to identify true-positives) and specificity (i.e. the ability to identify true-negatives) is the utilization of training data through the gold standard sets. The construction of these gold sets is a challenging task and required added attention. The gold standard positive dataset consists of 7 373 interactions and the gold standard negative dataset includes 14 479 PPIs. The gold standard negative dataset was constituted by applying the notion that protein which do not exist in the same sub-cellular compartment, cannot interact (Arga et al., 2007 and Jansen and Gerstein, 2004). The gold standard positive dataset was constituted by gathering physical interaction data of STRING with very high confidence PPI with score (>0.999) (Szklarczyk et al., 2011), co-complex interactomes of HINT database (Das et al., 2012), and high-quality binary human PPIs (Rolland et al., 2014). Although each of the three subsets incorporated in construction of the positive dataset were considered as highly confident in literature, the intersections of these datasets constitute only 14.95% of the whole data.

The trade analysis presented the confidence score of 0.80 as an optimum cut off to determine a high confidence dataset (HCS) in human (Table 3.2). This dataset, including 147 923 binary interactions, comprise only 59.09% of all in the pool of PPIs and revealed the equilibrium between sensitivity, specificity and coverage (94.5%, 80.9% and 82.8%, respectively). Furthermore, a low confidence dataset (LCS) including 75 007 (25.40%) binary PPIs with confidence scores lower than 0.50 was also determined.

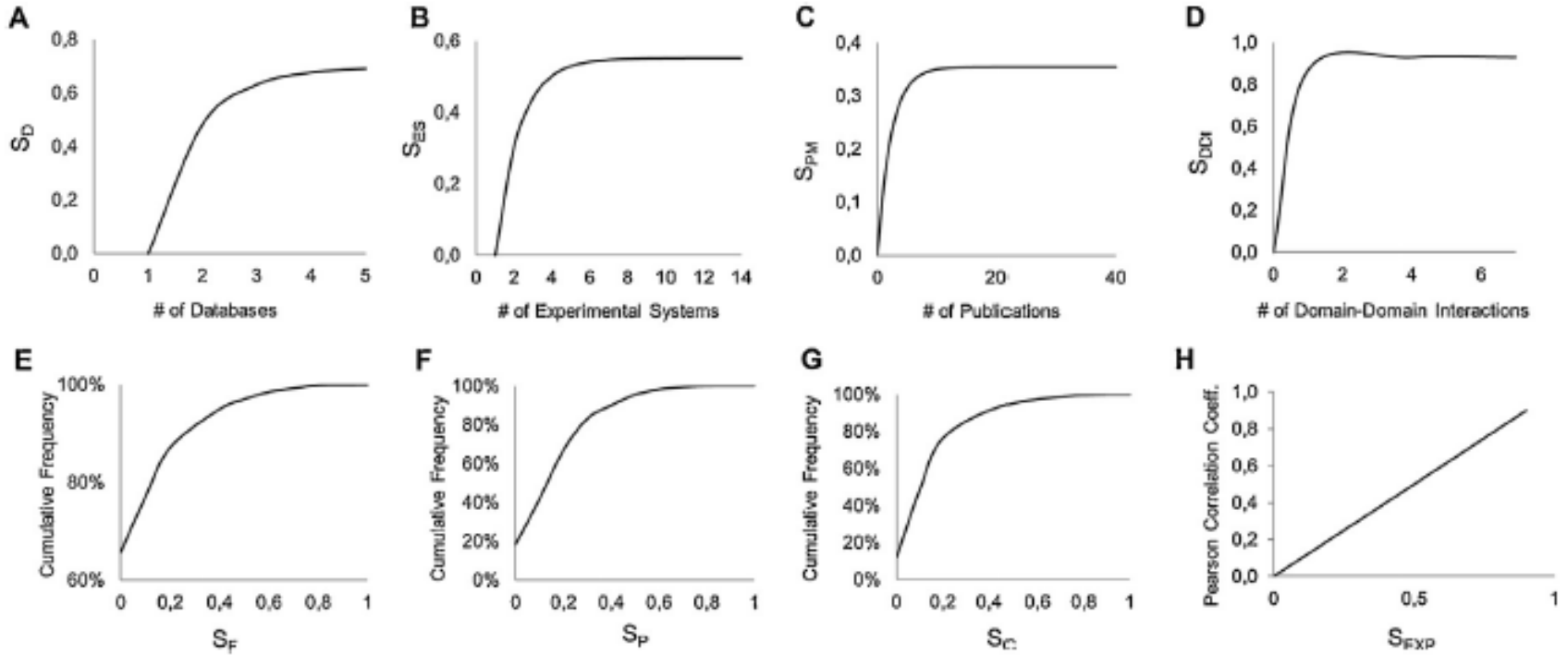


Figure 3.6 Distributions of association scores (S_i) for heterogeneous attributes. Scores are based on (A) the number of databases reporting the interaction (S_D), (B) the frequency of detection by different experimental systems (S_{ES}), (C) the number of published papers reporting the interaction (S_{PM}), and (D) structural conformity (S_{DDI}). The association score distributions, which were determined based on GO semantic similarity analysis, are represented by cumulative frequencies for: (E) functional coherence (S_F), (F) biological process coherence (S_P), and (G) co-localization (S_C). (H) The association score distribution for gene expression correlation (S_{EXP}) as a linear function of Pearson correlation coefficients.

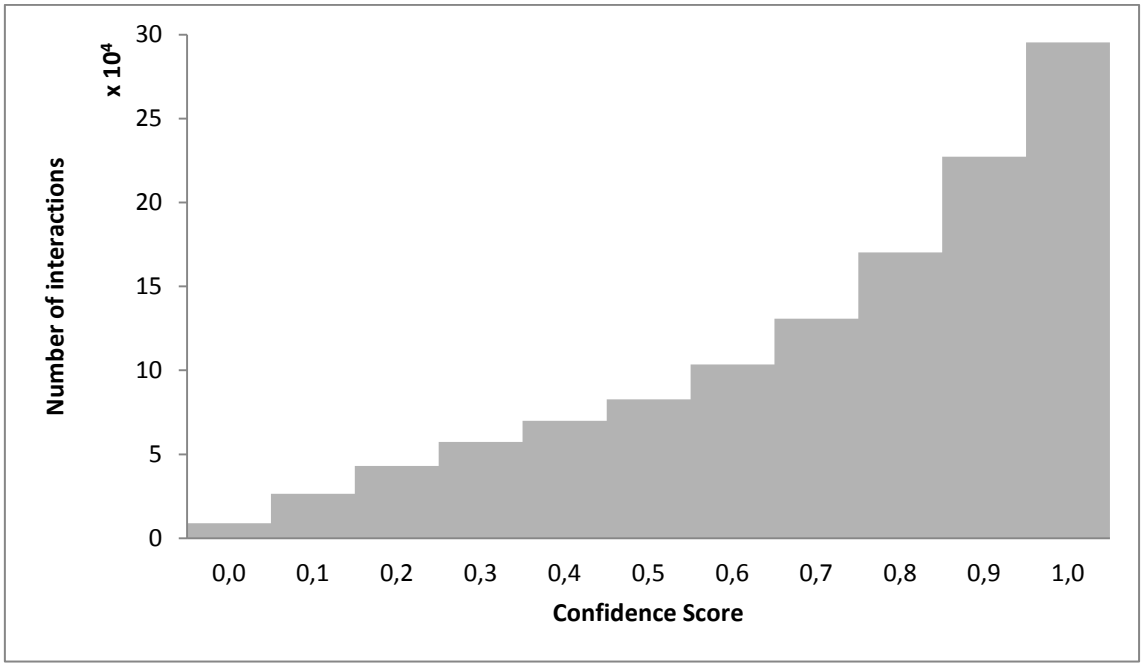


Figure 3.7. Cumulative distributions of confidence scores. Interactions with scores above 0.80 constitute only 21.5% of all and were considered as highly confident interactions. According to the design of the scoring function, such score implies that the interaction is supported significantly by multiple evidences.

Table 3.2 Trade off analysis of different confidence score cut offs.

Confidence Score	**# of PPIs**	**# of Proteins**	**Sensitivity (%)**	**Specificity (%)**	**Coverage (%)**
0.99	33475	6562	67.9	96.7	41.1
0.95	71920	9974	85.2	91.9	62.5
0.90	103437	11295	90.8	87.2	70.8
0.85	127272	12380	92.8	84.1	77.6
0.80	***147923***	***13213***	***94.5***	***80.9***	***82.8***
0.70	181839	14212	96.4	74.9	89.1
0.60	203787	14482	97.3	71.8	90.8
0.50	220281	14756	97.7	69.5	92.5

The power of the scoring methodology was assessed by the Receiver Operating Characteristic (ROC) curve analysis which provides the performance comparison of the presented scoring scheme with four previous schemes (Schaefer et al., 2012; Aranda et al., 2010; Persico et al., 2005; Kamburov et al., 2012). (Figure 3.8). Considering the general rules of AUC values analysis (Lin et al., 2004), the present scoring scheme expressed an excellent discrimination with regards to the AUC value (0.949), although scoring schemes of other databases indicated acceptable discrimination with AUC values which varied between 0.897 (HIPPIE) and 0.759 (INTSCORE). The highest AUC value belonged to the presented method and highlighted advantages of our scheme with respect to the reported scoring systems.

3.2.2 Characteristics of highly confident protein-protein interactome

Functional enrichment analysis of highly reliable human PPI network was done to define and analyze cancer and process subnetworks (Figure 3.9).

Disease enrichment analyses of this PPI network pinpointed various types of cancers and achieved cancer specific proteins which were used to reconstruct highly reliable cancer specific subnetworks. The power of PPI network in elucidating the cancer mechanism was appreciated through the identification of interacting partners that were dysregulated in tumorigenic cells. Dysregulated subnetworks enhanced diagnosis and prognosis of cancer over single-gene markers (Chowdhury et al., 2011). These highly informative subnetworks were explored in detail to evaluate their potential in highlighting the common and specific biological mechanisms. Sixteen cancer types (Bladder, breast, colorectal, ovarian, cervical, endometrial, esophageal, head and neck, lung, pancreatic, prostate and stomach cancers, leukemia and melanoma) were included in the subnetwork topological features and functional enrichment analysis, individually. Cell cycle, DNA repair and p53 signaling pathways were enriched in almost all cancer subnetworks. Furthermore, pathway enrichment analysis pinpointed significant signaling pathways including ErbB, FAS, insulin, MAPK, Neurotrophin and PI3 signaling pathways together with apoptosis, endocytosis, adherens junction and oocyte meiosis (Table 3.3).

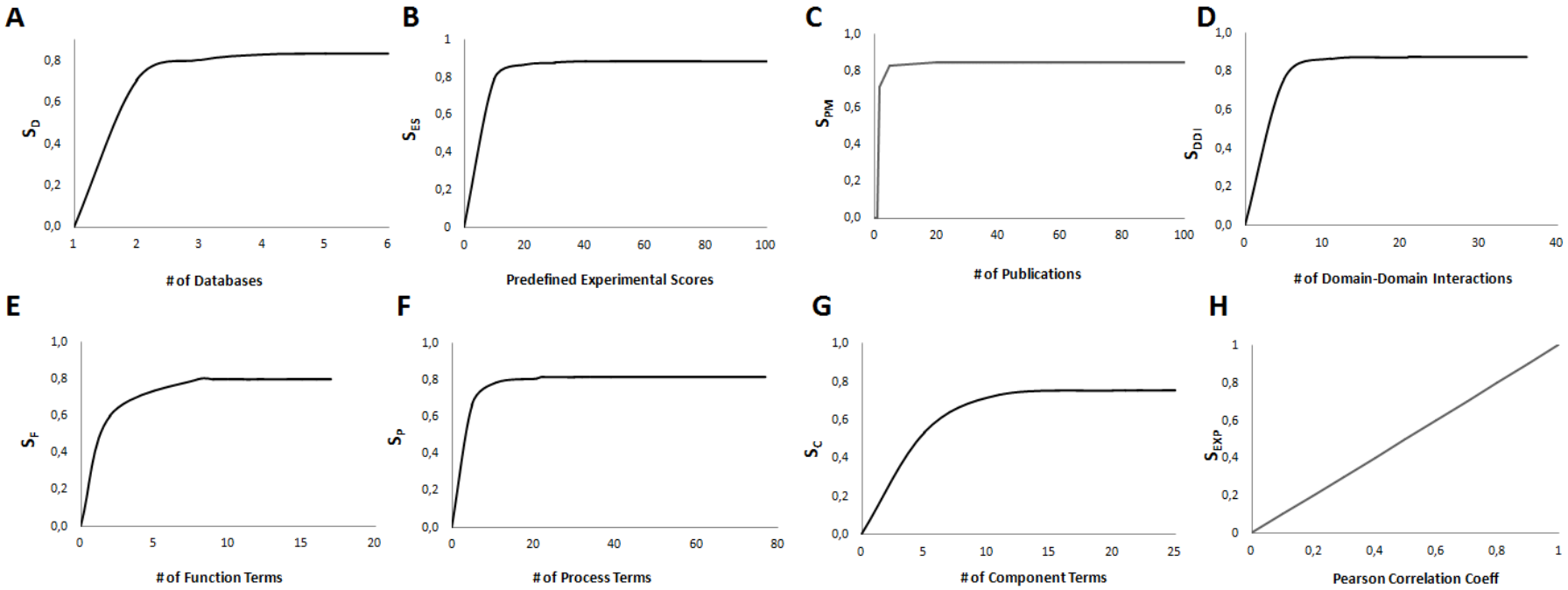
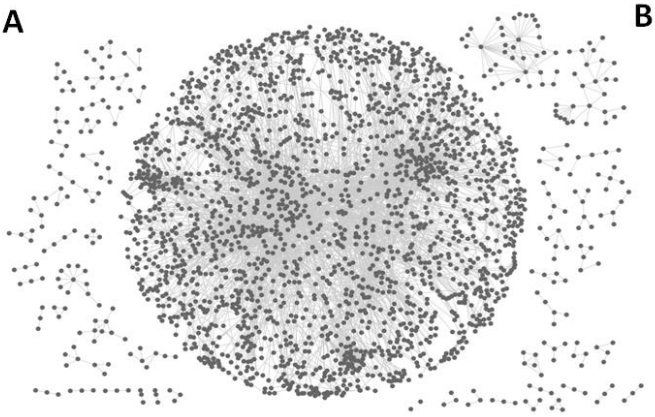


Figure 3.8 Performance analysis of the PPI scoring schemes via ROC curves. The predictive power of the scoring method was quantified by the area under curve (AUC) metric. The presented scoring scheme indicates an excellent discrimination (with AUC value of 0.949) whether other scoring schemes represents acceptable discrimination with AUC values between 0.897 (HIPPIE) and 0.759 (INTSCORE).

A



B

Number of nodes (proteins)	2641
Number of edges (interactions)	6460
Clustering coefficient	0.147
Connected components	225
Network diameter	16
Network centralization	0.049
Characteristic path length	4.903
Avg number of neighbors	3.943
Network density	0.001
Network heterogenity	1.844

Figure 3.9 Highly reliable PPI network comprised of interactions with confidence score > 0.999 and its topological properties. (A) Graph representation of the network. (B) Values of selected topological metrics.

Topological analysis was also done for each subnetwork and hubs were identified. P53 appeared as a hub in almost all PPI network except cervical cancer. CCND1, CDN1A, EGFR, PARP1 were also the most common hubs among subnetworks. CDNA1 (bladder), EP300, SHC1 (breast), 1A01, 1B07, 1B08, 1B15, 1B44 (cervical), BLM, IRS1 (colorectal), PRGR (endometrial), DNLI3 (esophageal), ERCC2 (head and neck), STAT3 (kidney), ABL1, CDN1B, JAK2, PML, VDR, BCL2 (leukemia), XPF, MLH1 (melanoma), CDK4, RAD51 (ovarian), TSP1 (prostate), CCDN1, FBW1A, GSK3B, TGFR2 (stomach) were found as cancer specific hubs in this study (Figure 3.10).

Table 3.3 Molecular pathways enriched in cancer specific subnetworks (p-value < 0.05).

	Cell cycle	DNA Repair	Apoptosis	p53 signaling	ErbB signaling	FAS signaling	Insulin signaling	MAPK signaling	Neurotrophin signaling	PI3 kinase pathway	Endocytosis	Adherens junction	Oocyte meiosis
Breast Cancer	+	+	+	+	+	+	+	+	+	+		+	+
Colorectal Cancer	+	+	+	+			+		+	+		+	
Ovarian Cancer	+	+		+									
Melanoma	+	+	+	+	+			+			+		
Lung Cancer	+	+	+	+	+	+		+			+		
Stomach Cancer	+			+	+			+			+	+	
Prostate Cancer	+	+	+	+	+		+		+			+	+
Endometrial Cancer	+	+		+									
Leukemia	+	+	+	+	+				+				
Head and Neck Cancer	+	+		+	+						+		
Esophageal Cancer	+	+		+									
Cervical Cancer	+			+									
Pancreatic Cancer	+			+									
Bladder Cancer	+	+		+									

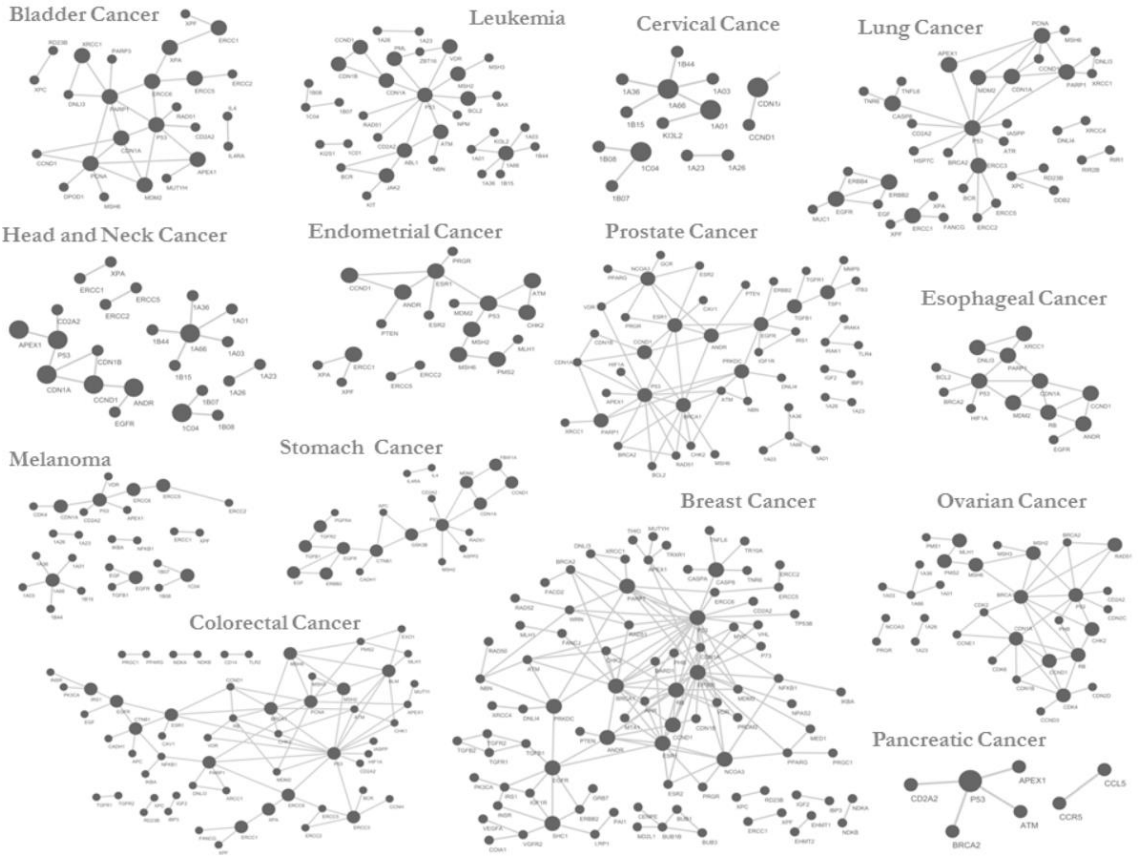


Figure 3.10 Cancer specific subnetworks comprising of highly reliable interactions (confidence score > 0.999) among proteins associated with the concerned disease according to Genetic Association Database (GAD).

The process enrichment analysis of the highest reliable network was done to examine the major biological processes related to this network. A total of twelve processes were selected for analysis. The genes involved in these processes have been extracted from DAVID and a PPI subnetwork was constructed for each biological process (Figure 3.11). The largest PPI subnetwork was of transcription process with 1318 binary interactions among 600 proteins and the smallest subnetwork was for nervous system development which included 147 binary interactions among 125 proteins. The hubs of these individual subnetworks were also identified. P53 was the mutual hub in ten of the sixteen processes (cell death, cell communication, cell cycle, cell differentiation, immune response, immune system process, localization, nervous system development, signal transduction and response to stress). LYN was the mutual hub in localization, nervous system development, phosphorylation, response to stress and signal

transduction. IKKB was specific to cell death, CDN1A, MDM2, MPIP1, PLK1, PRS8, PSA2, PSA3, PSA4, PSA5, PSB5, PSB7 were cell cycle specific hubs. STAT3 was cell differentiation specific and hubs specific to immune response were B2MG, VAV, TBK1, UBE2N. The hubs specific to immune system process were KPCT and TGFR1. Phosphorylation specific hubs were SHC1, JAK2, STAT1, ABL1 and MK03. Hubs specific to localization were CBL, A4, IMB1 and STXB1. Nervous system development specific hubs were ERBB2, PRKDC, NOTC1 and NCOA6. FAK2 was a specific hub for response to stress. The hubs specific to signal transduction were NCOR1, TRAF3, M3K7, VDR and RIPK2 and transcription specific hubs were EP300, HDAC2 and SIN3A.

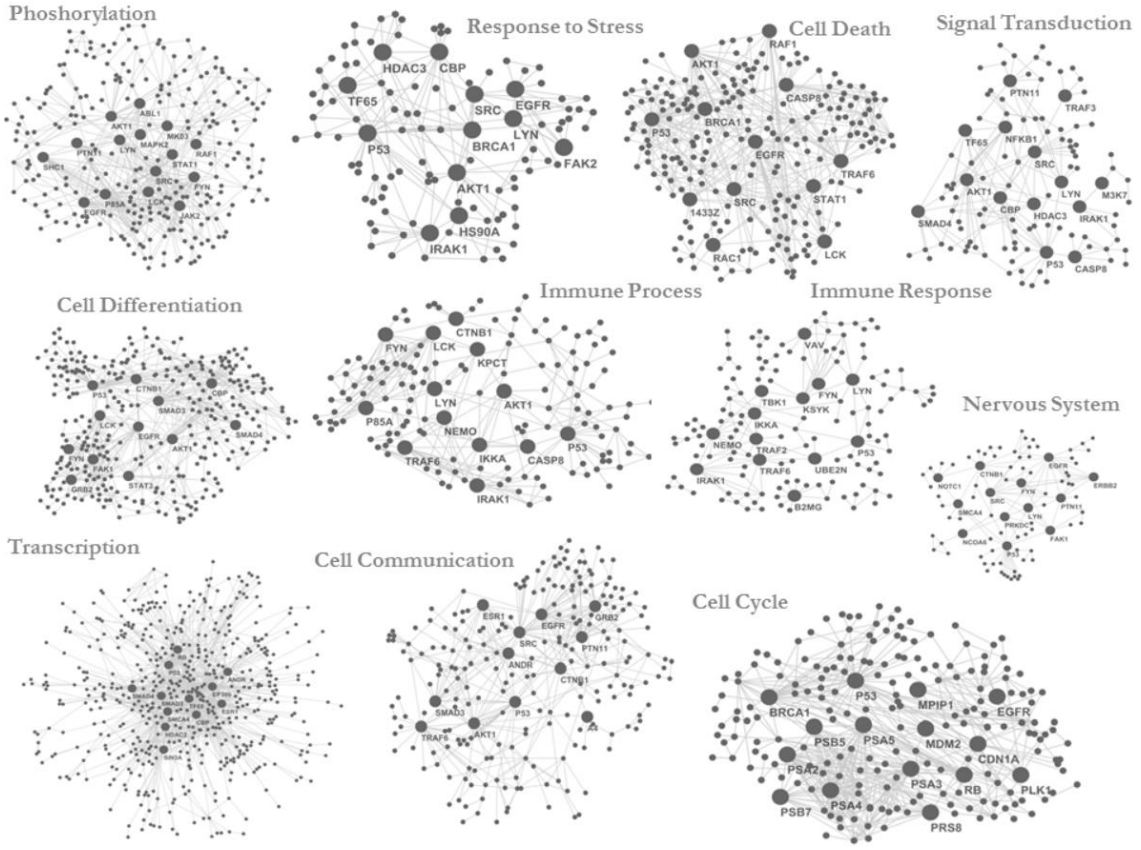


Figure 3.11 Process specific subnetworks comprising of highly reliable interactions (confidence score > 0.999) among proteins associated with the concerned biological process according to Gene Ontology (GO) annotations.

The signaling pathways related to each individual process specific subnetworks were also identified with neutrophin, TGF-beta, MAPK and NGF signaling pathways enriched in all of the twelve (Table 3.4). The diseases enriched in these subnetworks

include breast cancer, inflammatory bowel disease, leukemia, and type 1 diabetes (Table 3.5). For breast cancer subnetwork, the highest number of interaction for a given protein was 22, while there were 36 proteins that interact with only one protein.

Table 3.4 Signaling pathways enriched in process specific subnetworks (p-value < 0.05)

	Neurotrophin signaling	TGF-beta signaling	MAPK signaling	NGF signaling	FGF signaling	Jak-STAT signaling	Adipocytokine signaling	Apoptosis signaling	Chemokine signaling	Fc epsilon RI signaling	EGF receptor signaling	T cell receptor signaling	Toll-like receptor signaling	NOD-like receptor signaling	B cell receptor signaling
Programmed Cell Death/Apoptosis	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Cell Cycle	+	+	+	+							+				
Cell Differentiation	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Response to Stress	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Phosphorylation	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Signal Transduction	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Nervous System development	+	+	+	+	+						+				
Immune System Process	+	+	+	+		+	+	+	+	+		+	+	+	+
Cell Communication	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Immune Response	+	+	+	+		+	+	+	+	+		+	+	+	+
Localization	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Transcription	+	+	+	+		+	+	+				+	+	+	+

Table 3.5 Diseases enriched in process specific subnetworks (p-value < 0.05)

	Colorectal cancer	Breast cancer	Prostate cancer	Stomach cancer	Melanoma	Multiple sclerosis	Inflammatory bowel disease	Meningioma	Endometrial cancer	Head and neck cancer	Crohn's disease	Leukemia	Ovarian cancer	Lung cancer	Diabetes, type 1	Bladder cancer
Programmed Cell Death/Apoptosis	+	+	+	+	+	+	+	+	+	+	+	+	+	+		
Cell Cycle	+	+	+	+	+			+	+	+		+	+	+		
Cell Differentiation	+	+	+	+	+	+	+	+	+	+		+	+	+	+	
Response to Stress	+	+	+	+	+	+	+				+	+	+	+		
Phosphorylation	+	+	+	+	+	+	+	+								
Signal Transduction	+	+	+	+	+	+	+	+			+				+	
Nervous System Development	+	+	+	+	+			+	+			+		+	+	+
Immune System Process	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Cell Communication	+	+	+	+	+	+	+	+	+	+			+	+	+	
Immune Response	+	+	+	+	+	+	+			+	+	+	+	+	+	+
Localization	+	+	+	+	+	+	+	+				+			+	
Transcription	+	+	+	+	+								+		+	+

3.3 Systematic Analysis for Integration of Transcriptome Data with Biological Networks

High-throughput gene expression data from GEO was collected and analyzed to obtain differentially expressed genes which were used to reconstruct PPI, transcriptional regulatory and metabolic networks. Reconstructed biological networks were applied to identify potential biomarkers and therapeutic targets as well as to elucidate the underlying molecular mechanisms of breast cancer, triple negative breast cancer and esophageal squamous cell carcinoma.

3.3.1 Integrative analysis of breast cancer

The highly reliable protein-protein interaction network of *Homo sapiens* was reconstructed through previous analyses and this high confidence network was used to define the molecular mechanisms of breast cancer and other associated diseases.

3.3.1.1 Differentially expressed genes

In the present study, we utilized four publicly available gene expression datasets for breast cancer which included samples of healthy control and breast tumors. Statistical analyses of datasets led to identification of differentially expressed genes (DEGs) (Table 3.6). A total of 5993 and 7348 genes were identified as down- and upregulated, respectively.

Table 3.6 Numbers of differentially expressed genes in breast cancer datasets.

NCBI-GEO ID	# of Down-regulated Genes	# of Up-regulated Genes	Total # of DEGs
GSE9574	1148	1321	2469
GSE15852	1823	3089	4912
GSE26910	1971	2219	4190
GSE39004*	1938	1826	3764
GSE42568*	2821	2892	5713
Total	5993	7348	11690

P-value<0.05 except * p-value<0.0001

Among the DEGs, 41 genes were down-regulated and 17 genes were commonly up-regulated in all datasets. These genes are called ‘‘core DEGs’’ throughout the study. Through functional enrichment analysis, proteins associated with the down-regulated core genes were categorized into seven main groups: i) transcription factors (ATF3, CEBPD, EGR1, JUN, TCF7L2, KLF2, KLF4, KLF11, MAFF, NFIL3), ii) actin cytoskeleton organization related proteins (EPB41L2, EHBP1), iii) enzymes (ADAMTS1, DUSP1, FLT1, NLGN1, PPP2R1B), iv) signaling proteins (ADM, ANK2, ANXA1, CFLAR, EPS8, FOS, FOSB, GCOM1, IL6, ITSN1, SPRY1, TMEM100), v) nervous system related proteins (AHNAK, SEMA3A, SLIT3), vi) RNA processing related proteins (CELF2, EIF1, ZFP36), vii) extracellular matrix organization related proteins (FBLN5, ITIH5) as well as a receptor (THBD), sodium bicarbonate co-

transporter (SLC4A4), cell adhesion related protein (MPDZ) and epidermis development protein (EMP1). Furthermore, proteins of 17 upregulated core DEGs were classified into four groups: (i) actin cytoskeleton related proteins (EVL, PLEK2), ii) enzymes (AHCY, CANT1, PDIA4, SHMT2, SORD, SULF1, TRIM27), iii) protein transporters (NUP210, TMED3), and iv) signaling proteins (RAB15, STAT1) together with protease inhibitor (SPINT2), transmembrane protein (TMEM132A) and collagen related protein (COL11A1). Through the investigation of these core DEGs and their association with breast cancer, there is very limited number of studies on down-regulated (EHBP1, FBLN5, GCOM1, NLGN1 and TMEM100) and up-regulated (PDIA4, PLEK2, RAB15, SORD, TMED3, TMEM132A) core DEGs.

3.3.1.2 Reconstruction and analysis of active PPI subnetworks

Reconstructions of active PPI subnetworks in breast cancer were based upon proteins associated with 58 core DEGs and their physical interactors. As a result, the down-regulated interaction subnetwork consisted of 1477 binary interactions between 1228 proteins (Figure 3.12A), and the up-regulated interaction subnetwork included 473 interactions between 446 proteins (Figure 3.12D). Analyses of reconstructed networks identified a set of proteins as hubs (i.e., nodes exhibiting high topological centrality in the graph) (Figure 3.12B and 3.12C). These hub proteins deserve attention for further experimental studies, since they indicate significant potential for being candidate biomarkers for breast cancer. Down-regulated hub proteins were classified into four groups: i) transcription factors (ATF3, CEBPD, EGR1, FOS), ii) enzyme (PPP2R1B), iii) signaling proteins (ANXA1, CFLAR, EPS8, ITSN1), iv) RNA processing protein (ZFP36) (Table 3.7).

Table 3.7 Hub proteins of the reconstructed PPI networks in breast cancer.

Protein Symbol	Molecular Function	Direction of regulation	Novelty	References
EGR1	Transcriptional factor	Down	Known	Liu et al., 2000
FOS	Transcriptional factor	Down	Known	Milde-Langosch et al., 2004
CEBPD	Transcriptional factor	Down	Known	Palmiere et al., 2012
2AAB	Enzyme	Down	Known	Esplin et al., 2006
ITSN1	Signaling protein	Down	Known	Eom et al., 2008
CFLAR	Signaling protein	Down	Known	Day et al., 2008
EPS8	Signaling protein	Down	Known	Chen et al., 2007
ANXA1	Signaling protein	Down	Known	Cao et al., 2008
ATF3	Transcriptional factor	Down	Known	Wolford et al., 2013
TTP	RNA processing protein	Up	Known	Brennan et al., 2009
1B07	Immune system related protein	Up	Known	Torigoe et al., 2012
CANT1	Enzyme	Up	Known	Simpson et al., 2008
DHSO	Enzyme	Up	Known	Tarnow et al., 2013
EVL	Cytoskeletal protein	Up	Known	Hu et al., 2008
GLYM	Enzyme	Up	Known	Lee et al., 2014
HDAC1	Enzyme	Up	Known	Krusche et al., 2005
IMA7	Tranporter	Up	Known	Dahl et al., 2006
P53	Tumor suppressor	Up	Known	Dumay et al., 2012
P73	Tumor suppressor	Up	Known	Dominguez et al., 2001
PIAS1	Enzyme	Up	Known	Liu et al., 2014
PO210	Tranporter	Up	Known	Barros et al., 2010
SAHH	Enzyme	Up	Known	Cimino et al., 2008
SPIT2	Protease inhibitor	Up	Known	Jonsson et al., 2010
STAT1	Signaling protein	Up	Known	Soond et al.,2008
TRI27	Enzyme	Up	Known	Tezel et al., 2009

ATF3 (Activating transcription factor 3) is a basic leucine zipper transcription factor that involves in cell division, apoptosis, inflammation and homeostasis (Gilchrist et al., 2010; Hai et al., 2010). It was reported that ATF3 regulated the effect of TGF-β in turns increased cell motility (Yin et al., 2010). ATF3 was proposed as a regulator in myeloid cells and reported to be related with breast cancer metastasis (Wolford et al., 2013). The controversial functions of ATF3 as both oncogene and tumor suppressor was also investigated. As an oncogene, ATF3 overexpression was presented to be enhanced in patients with breast cancer (Yin et al., 2008). ATF3 expression has the potential to promote cancer progression, but the mechanisms and transcriptional targets are still largely undefined (Thompson et al., 2009). Other studies suggested that ATF3 could inhibit carcinogenesis by promoting apoptosis and decreasing the metastatic potential in cell lines. Moreover, ATF3 was proposed as a potential tumor suppressor through its role in TGFβ signaling pathway (Kang et al., 2003; Yan and Boyd et al., 2006). *CEBPD (CCAAT/Enhancer Binding Proteins)* are classified in a family of leucine zipper (bZip) proteins that play role cell growth and differentiation (Ramji et al., 2002; Vinson et al., 2002). Several studies reported down-regulation of CEBPD in carcinogen-induced mammary tumors and primary breast cancer (Kuramoto et al., 2002, Tang et al., 2006; Palmieri et al., 2012). CEBPD was included in 70-gene signature, which predicted better survival of patients with breast cancer (Naderi et al, 2007). Furthermore, it was proposed as a tumor suppressor gene of breast cancer (Tang et al., 2006, Thangaraju et al, 2005; Stein et al, 2009) as well as an epigenetic biomarker of primary breast cancer at enhanced risk of relapse and metastasis (Palmieri et al., 2012). *EGR1 (Early growth response 1)* is found in a group of proteins that are a family of zinc-finger transcription factors with various roles in proliferation, apoptosis and stress responses. EGR1 plays a key role in growth control and functions in many signaling cascades. EGR1 directly mediates several tumor breast suppressor genes including NDRG1, PTEN and P53 (de Belle et al., 1999; Redmond et al., 2010). EGR1 protein levels were reduced in a variety of human tumor cell lines including breast cancer (Liu et al., 2000; Huang et al., 1997). EGR1 was proposed as a tumor suppressor gene in breast cancer (Huang et al., 1997; Wenzel et al., 2007). *PPP2R1B (Protein Phosphatase 2, Regulatory Subunit A, Beta),* is a trimeric holoenzyme that has a key role in the regulation of many signaling pathways whose dysregulation can lead to cancer

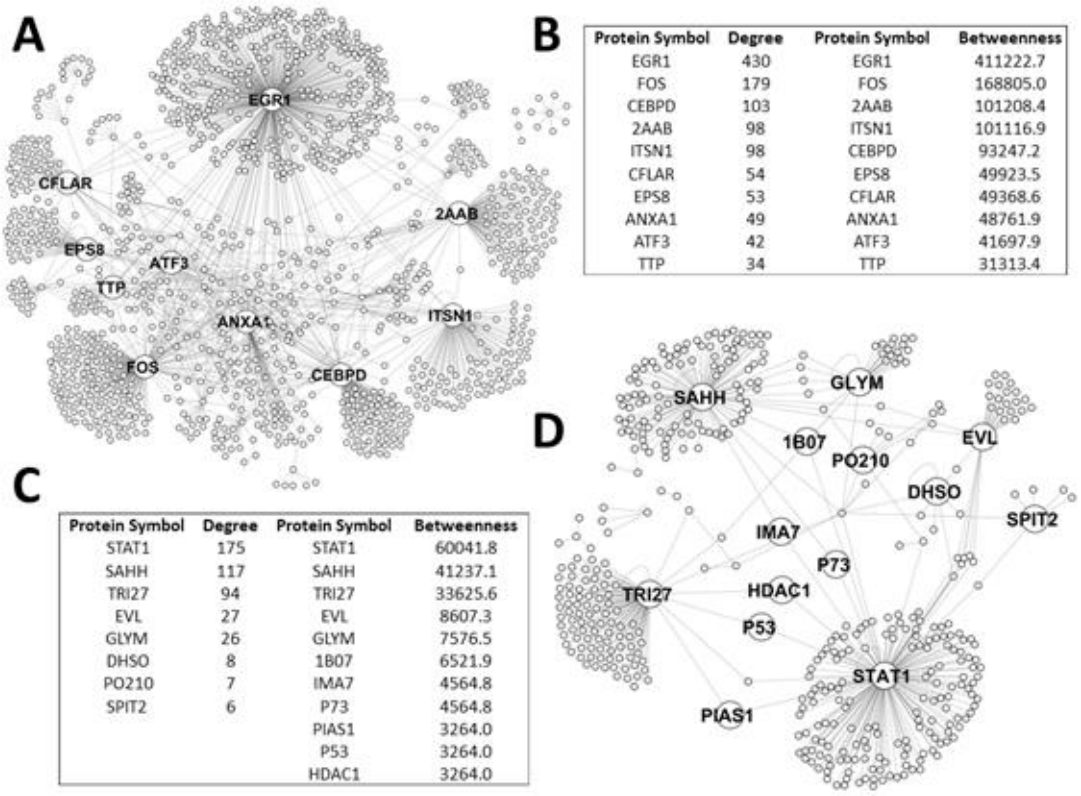
(Eichhorn et al., 2008). Inactivating somatic alterations in PPP2R1B was determined in breast cancer (Esplin et al., 2006; Westermarck and Hahn, 2008). Its function was also described as a tumor suppressor (Janssens et al., 2005). Wang and coworkers (2006) identified down-regulation of PP2R1B in breast cancer. *CFLAR (CASP8 and FADD-like Apoptosis Regulator)* is a member of the same family of CASP8 and CASP10, and serves as a negative regulator of apoptosis. CFLAR is taken a part in proliferation, cell cycle progression, and carcinogenesis. The knockdown of CFLAR initiated apoptosis and reduced the proliferation of breast cancer cells. Thus, CFLAR was proposed as a potential therapeutic target for breast cancer (Day et al., 2008; Bijangi-Vishehsaraei et al., 2010). *ITSN1 (Adaptor/scaffold proteins of the intersectin)* is involved in a family that forms important compounds of endocytic and signaling complexes. ITSN1 regulates cellular trafficking through actin cytoskeleton rearrangements and regulates the activity of several signaling pathways (Tsyba et al., 2013). There is a limited research on ITSN1 effects in breast cancer. Eom and coworkers expressed the reduced level of ITSN1 in fibroadenoma compared to ductal carcinoma (Eom et al., 2008). The recent study focused on the gene expression levels in response to efavirenz and reported that ITSN1 was involved in up-regulated cluster (Patnala et al., 2014). *EPS8 (Epidermal growth factor pathway substrate 8)* was determined as a substrate of EGF receptor that provides mitogenic signaling from receptor tyrosine kinases (Fazioli et al., 1993 and Gallo et al., 1997) and tyrosine phosphorylation of EPS8 was reported in several cancer cell lines (Matoskova et al., 1995). EPS8 was shown differentially expressed during breast cancer progression (Chen et al., 2007). The overexpression of EPS8 was reported in breast cancer cells and it was proposed among potential candidate breast cancer oncogenes (Yao et al., 2006; Li et al., 2013; Chen et al., 2015). *ANXA1 (Annexin A1),* so called lipocortin, is involved in an annexin superfamily of calcium- and phospholipid-binding proteins (Raynal and Pollard, 1994). ANXA1 plays role in hormone secretion, inflammatory response, EGFR degradation, apoptosis, membrane transport and cell differentiation (Sobral-Leite et al., 2015). The expression of ANXA1 was found to be concerned with the carcinogenesis process as well as metastasis (Yi et al., 2009; Guo et al., 2013; Boudhraa et al., 2014) including breast tumors (Cao et al., 2008; Bist et al., 2011; Swa et al., 2012). ANXA1 takes part in the initiation of breast cancer (Swa et al., 2012). Furthermore, the alteration of ANXA1 expression differs in subtypes of breast

cancer. For instance, overexpression of ANXA1 is related to the basal-like phenotype (Cao et al., 2008; de Graauw et al., 2010; Yom et al., 2011; Sobral-Leite et al., 2015) while the loss of ANXA1 expression is associated with invasive carcinoma (Shen et al., 2005; Yom et al., 2011). *ZFP36 (Zinc Finger Protein 36)* has a key role in the control of immune responses (Milke et al., 2013). It was shown to be involved in the regulation of proteins in invasion and metastasis (Al-Souhibani et al., 2010). Brennan and coworkers identified the reduced expression of ZFP36 as a negative prognostic marker in breast cancer (Brennan et al., 2009; Gebeshuber et al., 2009). Moreover, it was reported that knockdown of ZFP36 in breast cancer cell lines increased tumor growth (Milke et al., 2013). *FOS (BJ Murine Osteosarcoma Viral Oncogene Homolog)* involved in the processes of normal cell growth, cellular transformation and differentiation (Curran and Franza, 1988). It was shown that FOS was partly involved in tumor growth and progression and it was proposed as a possible tumor-suppressor (Durchdewald et al, 2009). The reduced expression level of c-FOS was reported in cancer studies (Jin et al, 2007). Moreover, FOS could have role in invasion of breast cancer cells (Milde-Langosch et al., 2004).

Up-regulated hub proteins were classified into three groups: (i) enzymes (CANT, SORD, SHMT2, HDAC1, PIAS1, AHCY, TRIM27), (ii) transporters (KPNA6, NUP210), (iii) tumor proteins (TP53, TP73) together with signaling protein (STAT1), protease inhibitor (SPINT2), immune system related protein (HLA-B), and cytoskeleton remodeling related protein (EVL).

AHCY (S-Adenosylhomocysteine hydrolase) is essential enzyme for cell growth and controlling gene expression. It was defined as a potential drug target for the design of antiviral and antiparasitic agents (Wang et al., 2014). AHCY was pinpointed as a direct c-MYC target and had roles in the tumorigenic activity of MYC (Fernandez-Sanchez et al., 2009). The up-regulation of AHCY was also reported in a previous breast cancer study (Cimino et al., 2008). *CANT1 (Calcium-activated nucleotidase)* is androgen mediated protein that commonly overexpressed in primary prostate cancer, and decreased CANT1 expression in prostate cancer cell lines lead to decrease in cell proliferation and migration (Gerhardt et al., 2011). The overexpression of CANT1 was identified in previous breast cancer researches (Tripathi et al., 2008; Simpson et al., 2008). *EVL (Ena/Vasodilator-Stimulated Phosphoprotein-Like protein)* plays role in

regulating the actin cytoskeleton. Higher EVL expression levels were determined in breast tumors and it was highlighted that overexpression of EVL was related to the clinical stages of breast cancer. Moreover, EVL was proposed to be associated with invasion and metastasis of human breast tumor (Hu et al., 2008). Another breast cancer investigation showed up-regulation of EVL which was linked to disease-related mortality (Andres et al., 2013).



Protein Symbol	Degree	Protein Symbol	Betweenness
EGR1	430	EGR1	411222.7
FOS	179	FOS	168805.0
CEBPD	103	2AAB	101208.4
2AAB	98	ITSN1	101116.9
ITSN1	98	CEBPD	93247.2
CFLAR	54	EPS8	49923.5
EPS8	53	CFLAR	49368.6
ANXA1	49	ANXA1	48761.9
ATF3	42	ATF3	41697.9
TTP	34	TTP	31313.4

Protein Symbol	Degree	Protein Symbol	Betweenness
STAT1	175	STAT1	60041.8
SAHH	117	SAHH	41237.1
TRI27	94	TRI27	33625.6
EVL	27	EVL	8607.3
GLYM	26	GLYM	7576.5
DHSO	8	1B07	6521.9
PO210	7	IMA7	4564.8
SPIT2	6	P73	4564.8
		PIAS1	3264.0
		P53	3264.0
		HDAC1	3264.0

Figure 3.12 Reconstructed PPI networks of breast cancer (A) Downregulated core network; (B) Topological properties of downregulated hubs; (C) Topological properties of upregulated hubs; (D) Upregulated core network.

HDAC1 (Histone deacetylases 1) ais associated with the development and progression of cancer. Substantial positive correlation between HDAC1 and hormone receptor expression was identified and this study propsoed a prognostic role of HDAC1 for breast cancer patients (Krusche et al., 2005). Several studies found overexpression of HDAC1 in hormone receptor positive tumors (Zhang et al., 2005; Muller et al., 2013). Moreover, Park and colleagues revealed the overexpression of HDAC1 and suggested

HDAC isoforms as potential biomarkers for breast cancer metastasis (Park et al., 2011). The results of another study demonstrated the cancer-promoting activity of HDAC1 in breast cancer, and they suggested HDAC1 as candidate targets of anti-cancer therapies (Wu et al., 2014). *HLA-B (HLA Class I Histocompatibility Antigen, B Alpha Chain)* is involved in the cell-mediated immune system, especially as antigen-presenting molecules for cytotoxic T lymphocytes (Torigoe et al., 2012). HLA-B was up-regulated in relapse-free patients compared to relapsed patients with breast cancer (Ascierto et al., 2011). HLA-B was proposed as one of the promising oncogenes for breast cancer (Xiang et al., 2014). *KNPA6 (Karyopherin α6)* took roles in the antioxidant response together with the cell redox homeostasis (Sánchez-Peña et al., 2013). KPNA6 was among the gene set that was predicted clinical response of stage IIB/III breast cancer patients (Dahl et al., 2006). *NUP210 (Nuclear pore complex glycoprotein Nucleoporin 210)* is an integral lipid-anchored membrane protein and is known to be involved in structural organization of the nuclear pore complex; in addition, it is directly related to gene regulation and mRNA transport into the cell cytoplasm (Prokhortchouk and Defossez, 2008). NUCP210 was defined as a predictive marker for breast cancer (Barros et al., 2010). *PIAS1 (Nuclear SUMO E3 ligase)* is a transcriptional repressor of STAT1 and NF-kappaB (Tahk et al., 2007), and PIAS1 is involved in T cell differentiation (Liu et al., 2014). Increased PIAS1 expression was reported in breast cancer that targeting the signaling pathway of PIAS1 was suggested as a novel therapeutic strategy to treat breast cancer (Liu et al., 2014). Deregulation of ER post-translational modifications by PIAS1 was involved in early stages of breast carcinogenesis (Karamouzis et al., 2007). *SHMT2 (hydroxymethyltransferase)* is required for cancer cell survival and it was overexpressed in breast cancer (Rhee et al., 2008; Lee et al., 2014). *STAT1 (Signal transducer and activator of transcription-1)* was involved in various cellular processes such as DNA damage, and activation of B and T cell immune responses. A tumor suppressor role of STAT1 was reported (Calo et al., 2003) as well as the loss of STAT1 expression have impact on the response to treatment of breast cancers (Bailey et al., 2012). Several groups showed the influence of STAT1 activity on the development of breast cancers (Soond et al., 2008; Klover et al., 2010; Khodarev et al., 2010; Bailey et al., 2012; Chan et al., 2012). Recently, Huang and coworkers presented that elevated expression and activation of STAT1 was increased in

endocrine-resistant breast cancer and they suggested that STAT1 inhibitors may be effective targets in resistant cells (Huang et al., 2014). *TRIM27 (Tripartite Motif-Containing Protein 27)* is a nuclear protein acting as a transcriptional repressor. The elevated expression level of TRIM27 was displayed in breast cancer cells (Cao et al., 1996; Tezel et al., 2009; Kato et al., 2009). Furthermore, TRIM27 was proposed as an oncogene (Hatakeyama, 2011). *SPINT2 (Serine Peptidase Inhibitor, Kunitz Type, 2*) is a transmembrane protein which was considered to take critical role in embryogenesis (Szabo et al., 2009). The lower expression level of SPINT2 was presented to be epigenetically silenced in cancer via DNA hypermethylation (Bergum and List, 2010; Dong et al., 2010; Nakamura et al., 2011). SPINT2 was considered as candidate biomarkers of breast cancer metastasis together with specific biomarkers for African American women (Yancy et al., 2007). Jonsson and coworkers demonstrated also the amplification of SPINT2 in breast cancer (Jonsson et al., 2010). *TP53 (Tumor Protein P53),* as a tumor suppressor, is mutated or inactivated in various types of cancers (Soussi and Lozano, 2005). *TP53* plays role in cell-cycle arrest, senescence, and apoptosis (Riley et al., 2008). Approximately, 30% of breast cancer patients have *TP53* mutation, but this frequency differs among breast cancer subgroups (Sorlie et al., 2001; Dumay et al., 2012; Walerych et al., 2012). *TP73 (Tumor Protein P73)* involved in the p53 family which had several p53-like characteristics including binding to p53 DNA target sites and induction of cell cycle arrest or apoptosis (Kaghad et al., 1997). p73 was activated like p53 by DNA damage and oncogenes (Zaika et al., 2001; Stiewe et al., 2004) .The elevated expression level of p73 was associated with an advanced tumor stage or poor prognostic parameters (Yang et al., 2000; Novak et al., 2001). Additionally, the results of breast cancer study suggested the up-regulation of that p73 which caused the tumorigenesis and poor clinical outcomes (Dominguez et al., 2001).

3.3.1.3 Functional enrichment analysis

According to pathway enrichment analyses of PPI networks, signaling pathways including JAK-STAT and Toll like receptor signaling were significantly up-regulated (p-value$<10^{-10}$), whereas Neurotrophin, T cell receptor, MAPK, apoptosis, FGF, EGFR, WNT, ERBB signaling pathways were significantly down-regulated (p-value$<10^{-10}$) in

breast cancer. On the other hand, chronic and acute myeloid leukemia, prostate and pancreatic cancer pathways were also significantly down-regulated (p-value<10^{-10}).

3.3.1.4 Transcriptional regulatory subnetworks

Transcriptional regulation data for twelve down-regulated (ATF3, CEPBD, EGR1, FOS, FOSB, JUN, NFIL3, TCF7L2, KLF2, KLF4, KLF11, MAFF) and one up-regulated (STAT1) TFs within the core DEGs were assembled from the literature and transcriptional regulatory subnetworks were reconstructed around these core-TFs and targeted core DEGs. The transcriptional regulatory subnetwork around down-regulated core-TFs consisted of 187 regulatory relations between these 12 down-regulated TFs and 25 down-regulated core DEGs (Figure 3.13A). Moreover, we explored upstream regulators of TFs into the pool of downregulated genes and identified 111 relations between 11 TFs (except CEPBD) and 82 genes (Figure 3.13B). Down-regulated TFs were involved in mainly MAPK (ATF3, FOS, FOSB, JUN), TNF (ATF3, FOS, JUN), TGFβ (ATF3, FOS, FOSB, JUN) and WNT (JUN, KLF4, TCF7L2) signaling pathways. We also explored the associated signaling pathways of the transcriptional regulatory networks as a whole, considering both downstream effectors and upstream regulators. The results of down-regulated genes associated regulatory networks also pointed out significant roles of MAPK (p-value: $1.7 \cdot 10^{-11}$), apoptosis ($4.0 \cdot 10^{-7}$), TGFβ (0.0011) and T cell receptor (0.015) signaling pathways.

The transcriptional regulatory subnetwork around the up-regulated core-TF, STAT1, consisted of five transcriptional regulatory relations between the TF and five core up-regulated DEGs (Figure 3.13C). STAT1 plays role in JAK-STAT, toll like receptor, chemokine, MAPK, and EGFR signaling pathways. Moreover, the upstream regulators of TFs were also explored into the pool of up-regulated genes and identified 14 relations between STAT1 and 14 up-regulated genes (Figure 3.13D). The signaling pathways associated with these transcriptional regulatory networks were toll receptor signaling pathway (p-value: 0.0096), Ras pathway (p-value: 0.019) and EGFR signaling pathway (p-value: 0.043).

The regulation of gene expression is required for normal cellular processes and the aberrant expression causes the abnormal cell growth as well as differentiation that lead to the development of cancer. The alteration of transcriptional regulation appeared

during tumor progression which has also impact on metastasis. Therefore, to maximize therapeutic efficiency, therapeutics which disturb the regulatory network of metastasis, can be chosen since they have potential to reduce the function of multiple effector genes simultaneously. Our approach allows identifying functional TFs that could be considered as potential therapeutic targets with promising effect on targeting the entire breast cancer gene network.

The transcriptional regulatory network around down-regulated core DEGs including 12 TFs and their upstream regulators together with downstream effectors was enriched with TGFB, MAPK and T cell receptor signaling pathways.

TGFB signaling pathway is involved in processes of immunity, inflammation, cancer, fibrosis and homeostasis (Massague, 2012). In the normal mammary gland, TGFB controls tissue homeostasis by inhibiting cell cycle progression, inducing differentiation and apoptosis, and maintaining genomic integrity (Barcellos-Hoff and Akhurst, 2009). TGFB signaling has a critical and dual role in breast tumorigenesis. In normal epithelium as well as in early tumorigenic lesions, TGFB has a tumor suppressive role through its ability to induce growth inhibition. As cancer progresses, however, tumor cells become resistant to TGFB-mediated growth inhibition, and instead TGFB promotes tumor progression and metastasis (Inman, 2011; Massague, 2012). The mechanism of this phenomenon is not well understood, but remains an important area of research (Smith et al., 2012). Several researchers proposed that the TGFB signaling pathway plays a critical role in progression of human breast cancer (Ivanovic et al., 2006; Grau et al., 2008). Moreover, genetic manipulation of the intrinsic TGFB signaling pathway in mammary cancer cells has provided direct evidence for its importance in driving the metastatic process (Tan et al., 2009). These studies have generated considerable enthusiasm for exploiting the TGFB pathway as a novel therapeutic target in breast cancer research (Tan et al., 2009; Ganapathy et al., 2010).

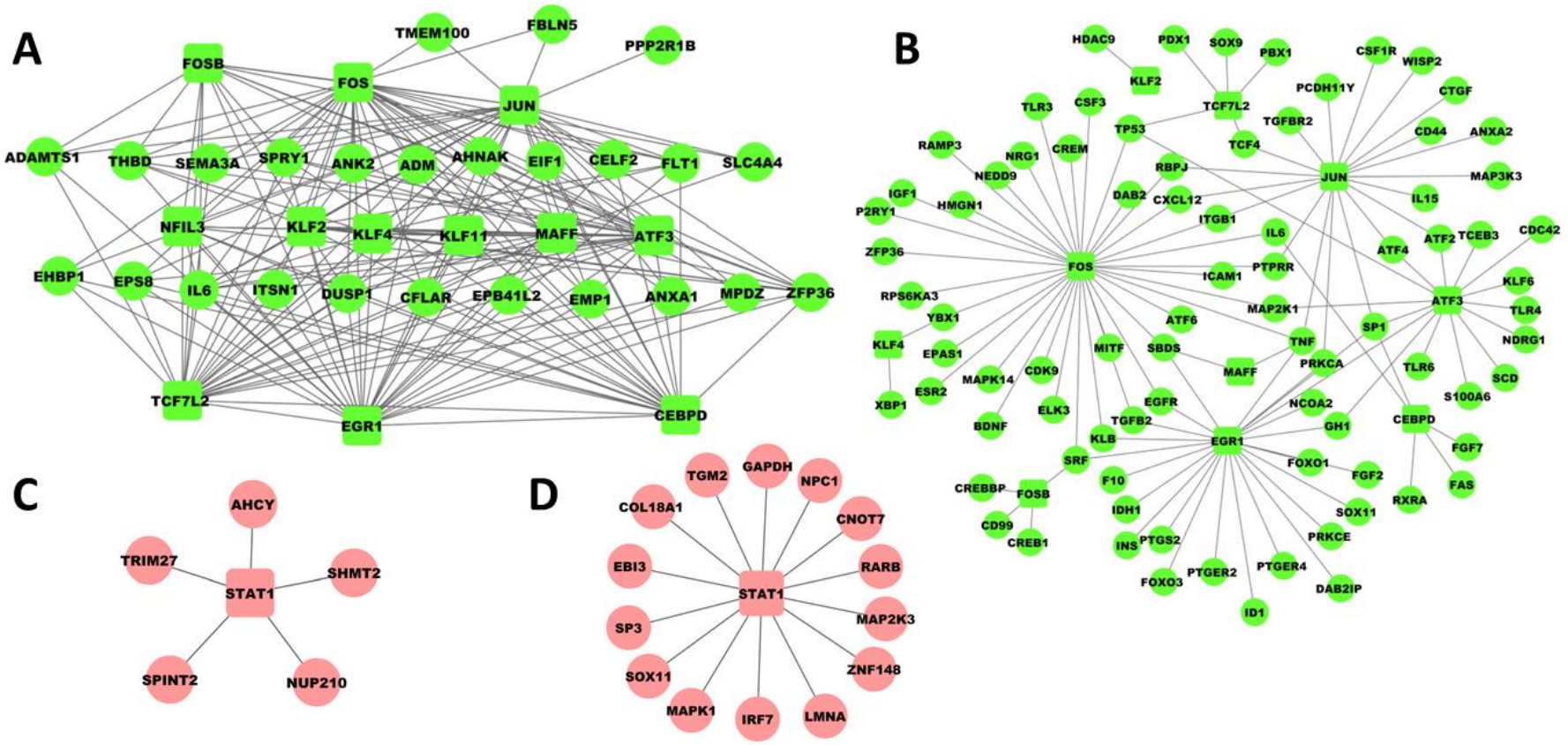


Figure 3.13. (A) Transcriptional regulatory subnetwork of downregulated core TFs. (B) Upstream regulators of downregulated core TFs. (C) Transcriptional regulatory subnetwork of upregulated core TF. (D) Upregulated upstream regulators of upregulated core TF.

MAPK signaling pathway is activated in cancer progression to promote cell proliferation, inhibit apoptosis, and regulate cell adhesion, migration and invasion capabilities (Reddy et al., 2003). It is also involved in progression of breast cancer, such as control of adhesion and migration (Zhu et al., 2011). MAPK (including P38, ERK1/2 and JNK1/2 members) signaling pathways play crucial roles in tumor cell proliferation, adhension and migration in breast cancer (Rosenthal et al., 2011; Meng et al., 2011). *T cell receptor signaling pathway* governs the maturation of T cells together with activation and proliferation of them. Activation of this pathway leads to activation of other signaling pathways that carry out cellular processes including mobilization of intracellular calcium stores, cytoskeletal reorganization and activation of transcription factors (Huang and Wange, 2004; Samelson, 2002). Moreover, T cell receptor signaling pathway is central to cell-mediated immunity, which is invariably activated by tumor associated antigens (Heemskerk, 2010). Bertucci and coworkers (2013) analyzed pathological response to primary chemotheraphy in patients with inflammatory breast cancer and they showed that various processes related to T-cell-dependent immunology were strongly enriched in their gene list, including the T-cell receptor signaling canonical pathway.

The only up-regulated TF was STAT1 involved in the STAT family of proteins that regulates cytokine and growth factor receptor signaling so they govern cell growth, survival, and differentiation (Yu and Jove, 2004; Stark and Darnell, 2012). STAT1 was reported to have a tumor suppressor role (Calo et al., 2003). Several research groups presented that STAT1 took significant role in the development of breast cancers (Klover et al., 2010; Raven et al., 2011; Chan et al., 2012; Bailey et al., 2012). The loss of STAT1 from tumor and surrounding mammary epithelium had impact on the progression and response to treatment of breast cancers (Bailey et al., 2012). Chan and colleagues (2012) also identified up-regulation of STAT1 expression is in ER-negative, stromal tissue than in ER-positive stromal tissue, and they also proposed that loss of STAT1 was important in the development of ER-positive breast tumors. Furthermore, another study showed that the overexpression of STAT1 was observed in endocrine-resistant breast cancer and STAT1 inhibitors were suggested to be effective in resistant cells (Huang et al., 2014). STAT1 dependent transcription regulation was suggested to

lead to therapeutic responses by regulating the efficacy of therapeutics and the progress of drug resistance (Koromilas and Sexl, 2013).

The transcriptional regulatory network around up-regulated core DEGs including STAT1 as a TF and its upstream regulators together with downstream effectors was enriched with Tool like receptor (TLR) signaling pathway, RAS pathway, and EGFR signaling pathway.

TLR binding activates several signaling pathways including nuclear factor κB (NFκB), mitogen-activated protein kinases (MAPKs) and IFN-regulatory factor 3, 5 and 7 (IRF3, IRF5 and IRF7) that are vital for innate and adaptive immune responses (Lee and Kim, 2007). Chronic inflammation was reported to have linked with tumor progression. In chronic inflammation, the TLRs activated NF-jB that caused proliferation of cancer cells out of control (Chen et al., 2007). Therefore, TLR signaling pathway in cancer progression was proposed as target for cancer therapy and prevention (Chen et al., 2006). The overexpression of TLRs was identified in breast cancer cell lines so it was suggested that TLR expression could be decreased to inhibit cell proliferation in breast cancer (Yang et al., 2010; Xie et al., 2009). Another study also reported the role of TLR4 in breast cancer progression and metastasis, i.e., the reduced metastasis was observed when TLR4 was targeted in breast cancer cell (Ahmed et al., 2013). Resler and coworkers also suggested strong the association between genetic variants in TLR pathway and breast cancer risk (Resler et al., 2013). *RAS pathway* is known to be one of the dysregulated pathways in cancer (Downward, 2003). It has role on transmitting the extracellular signals to intracellular targets that mediate cell metabolism, cell survival, cell cycle progression, and cell growth (Hollestelle et al., 2007). Clark and Der (1995) discovered the significance of down- and upstream regulators of RAS that was proposed to serve as drug targets to block the aberrant function of RAS signaling pathway in cancer. The hyperactivation of RAS/ERK pathway in 50% of breast cancers was ascertained and reported to be involved in tumor progression and recurrence of breast cancer (Sivaraman et al., 1997; von Lintig et al., 2000; Mueller et al., 2000). Furthermore, RAS signaling pathway was also reported in tumor formation that Hollestelle and coworkers identified mutations in RAS pathway genes in 40 human breast cancer cell lines (Hollestelle et al., 2007). A recent study also uncovered the activated RAS pathway in breast tumors (McLaughlin et al., 2013). *EGFR signaling*

pathway is associated with several cellular processes including apoptosis, cell proliferation, angiogenesis, and metastasis (Zhang et al., 2007). *EGFR* was reported to be up-regulated in almost 30% of tumors and associations with the stages of disease, prognosis and survival (Capdevila et al., 2009). The expression of *EGFR* was reported to be associated with BRCA1 status in breast cancer patients (van Diest et al., 2006). Pitteri and colleagues also demonstrated overexpression of *EGFR* in patients with breast cancer and proposed *EGFR* levels as an early biomarker (Pitteri et al., 2010). The EGFR-targeted therapies were also known in breast cancer treatments (Jonston et al., 2006). EGFR-Ras-ERK pathway was identified to govern the invasive characteristics of breast epithelial cells (Grass et al., 2013).

3.3.1.5 Reporter metabolic features

Reporter metabolite analyses highlighted three metabolites, chloral hydrate, EPA and NADPH, which were common metabolites of all datasets. In addition, 126 metabolites were shared at least by four datasets and 87 of them were lipid compounds. Furthermore, 21 reporter metabolites were having roles in biosynthesis of unsaturated fatty acids and 12 reporter metabolites were taking parts in fatty acid biosynthesis (Table 3.8).

Table 3.8 Metabolic pathway enrichment analysis of common reporter metabolites.

Metabolic Pathway	P-value	in background	in set
Biosynthesis of unsaturated fatty acids	$8.25 \cdot 10^{-12}$	54	21
Fatty acid biosynthesis	$3.65 \cdot 10^{-10}$	49	12
Pyruvate metabolism	$4.93 \cdot 10^{-6}$	32	7
Primary bile acid biosynthesis	$5.65 \cdot 10^{-4}$	47	6
Oxidative phosphorylation	$1.87 \cdot 10^{-5}$	16	5
Methane metabolism	$8.64 \cdot 10^{-4}$	34	5
Glycerolipid metabolism	$6.48 \cdot 10^{-4}$	32	5

Furthermore, we performed functional enrichment analysis on reporter metabolites of each dataset. These metabolic data indicated oxidative phosphorylation that was the only common term of each individual dataset. Other significant metabolic pathways

were lipid metabolic pathways including biosynthesis of unsaturated fatty acids, fatty acid biosynthesis, linoleic acid metabolism, glycerolipid and glycerophospholipid metabolism, amino acid metabolic pathways including beta-alanine and tyrosine metabolism, and carbohydrate metabolism including pyruvate metabolism (Figure 3.14).

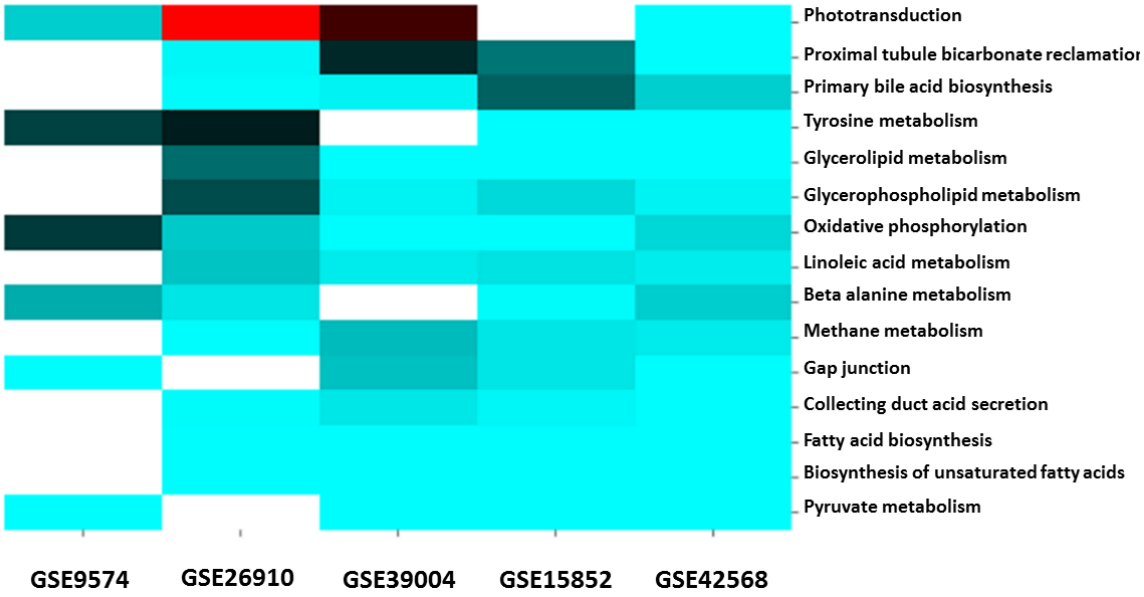


Figure 3.14. Metabolic pathways enrichment analysis through reporter metabolites of each dataset.

The results of reporter metabolite analysis indicated the importance of oxidative phosphorylation and lipid metabolism in breast cancer. Pyruvate, fatty acids and amino acids are consumed to generate energy in form of ATP through oxidative phosphorylation in normally differentiated cells (Hinkle et al., 1991). The oxidative phosphorylation is also required for apoptosis (Dey and Moraes, 2000). It was reported that mitochondria controlled tumor growth through regulation of the TCA cycle and oxidative phosphorylation (Solaini et al., 2011). Thus, understanding the mitochondrial metabolic changes will provide new therapeutic approaches. The molecular mechanisms for the oxidative phosphorylation proteome network were dysregulated in breast cancer cell mitochondria leading mitochondria cytopathy (Putignani et al., 2008). Singh and coworkers (2009) examined the breast tumors and reported the mutations in the polymerase γ gene that caused serious mitochondrial DNA deficiency and oxidative phosphorylation disturbance (Singh et al., 2009). Mitochondrial DNA mutations were

suggested to activate the PI3K/AKT pathway leading to enhanced metastasis in breast cancer cells (Kulawiec et al., 2009). Another study of breast cancer by Sastre-Serra et al. (2010) showed that increased reactive oxygen species enhanced tumorigenicity. Therefore, a novel approaches in cancer treatment could target the energetic metabolism of the tumors. From this point of view, the bioenergetic signature of the breast tumor may provide to design future therapeutics that target the energy provision pathway.

The alterations in lipid metabolism have been defined as a common characteristic of tumor cells (Menendez and Lupu, 2007; Cairns et al., 2011). Lipid metabolism is also controlled by oncogenic signaling pathways in cancer cells, and was reported to be significant for the initiation and progression of tumorigenesis (Menendez and Lupu, 2007; Tennant et al., 2010). Cancer cells achieved membrane building blocks, signaling molecules, and energy supply to enhance cell proliferation via lipogenesis (Zhang and Du, 2012). Fatty acids are esterified to phospholipids that are required for signal transduction, intracellular trafficking, polarization, and migration in cancer cells (Menendez and Lupu, 2007; Swinnen et al., 2006). Fatty acid can be consumed through β-oxidation in response to glucose limitation to provide key substitute energy for cancer cell survival (Zhang and Du, 2012). Moreover, other lipids molecules including diacylglycerol, phosphatidic acid, and lysophosphatidic acid also regulate signaling pathways in cancer cells (Menendez and Lupu, 2007; Swinnen et al., 2006). The increased lipogenesis is an indicator of many cancers including breast cancer (Chajes et al., 1995). Hilvo and coworkers (2011) demonstrated that the regulation of lipid metabolism in breast cancer differed from that of normal physiology. Furthermore, they determined phospholipids that were diagnostic potential and the regulation of their metabolisms, and suggested to provide therapeutic opportunities in breast cancer treatment. Their results showed that lipid metabolism in tumor cells could be exploited as therapeutic targets. Thus, further understanding of dysregulated lipid metabolism and related signaling pathways is required to design efficient cancer therapeutics.

3.3.2 Integrative analysis of triple negative breast cancer

3.3.2.1 Differentially expressed genes

In the present study, we utilized four publicly available gene expression datasets for TNBC, which included samples of non-TNBC and TNBC tissues. Statistical analyses of datasets led to identification of DEGs (Table 3.9). A total of 1337 and 2088 genes were identified as down- and upregulated, respectively.

Table 3.9 The number of differentially expressed genes in TNBC datasets.

Dataset	**# of down-regulated DEGs**	**# of up-regulated DEGs**	**Total # of DEGs**
GSE17650	164	224	338
GSE18539	221	250	471
GSE27447	684	1245	1929
GSE37614	441	602	1043

Among the DEGs, only two genes, GATA3 and MLPH, were downregulated and a single gene, CHI3L2, was upregulated in all datasets (Figure 3.15). GATA3 gene encodes the GATA binding protein 3, which is a member of the GATA family of zinc-finger binding transcription factors. It plays a role in human growth and differentiation of breast tissues (Zheng and Blobel, 2010). The MLPH gene encodes a member of the exophilin subfamily of RAB effector proteins, and downregulation of MLPH genes in TNBC was reported by Fountzilas and co-workers (2008). In addition, the downregulation of MLPH gene expression was observed in lymph node positive breast cancer patients (Abba et al., 2007), and Thakkar and co-workers (2010) confirmed the lower expression levels of MLPH in ERa(-) breast cancer patients. The upregulated gene CHI3L2 encodes a kind of lectin that binds chito-oligosaccharides and other glycans with high affinity, and it was reported as one of the methylated genes in luminal B type breast cancer (Holm et al., 2010). GATA3 was recently reported as a biomarker for TNBC (Cimino-Mathews et al., 2013); however, the associations of MLPH and CHI3L2 genes with TNBC have not been studied yet.

Twenty-three downregulated (Figure 3.15A) and 17 upregulated (Figure 3.15B) genes were identified as common in at least three datasets. These genes were called ‘‘core DEGs’’ throughout the study. Through functional enrichment analysis, proteins associated with the downregulated core DEGs were categorized into six groups: i) transcription factors (GATA3, FOXA1, MYB), ii) receptors (ERBB4 and ESR1), iii) enzymes (GSTM3, ALDH6A1, ATP8B1, DUSP4), iv) signaling proteins (CACNG4, IL20, IGFBP2, RHOJ, STC2, ARGHDIB, RHOB, SEMA3F, TSPAN1, TSPAN13, HDGFRP3), v) transporters (MLPH, SLC39A6), and vi) keratin (KRT19). Furthermore, proteins of 17 upregulated core DEGs were classified into five groups: (i) transcription factors (ATF5, FOXC1, MYC and STAT5A), ii) receptors (CSF2RA, FZD7, CD274), iii) enzymes (CHEK2, PTGS2, CHI3L2, UHRF1), iv) metabolism-associated proteins (CRABP1, SLC2A3, SLC14A1), and v) signaling proteins (CXCL10, IL1b, SFRP1).

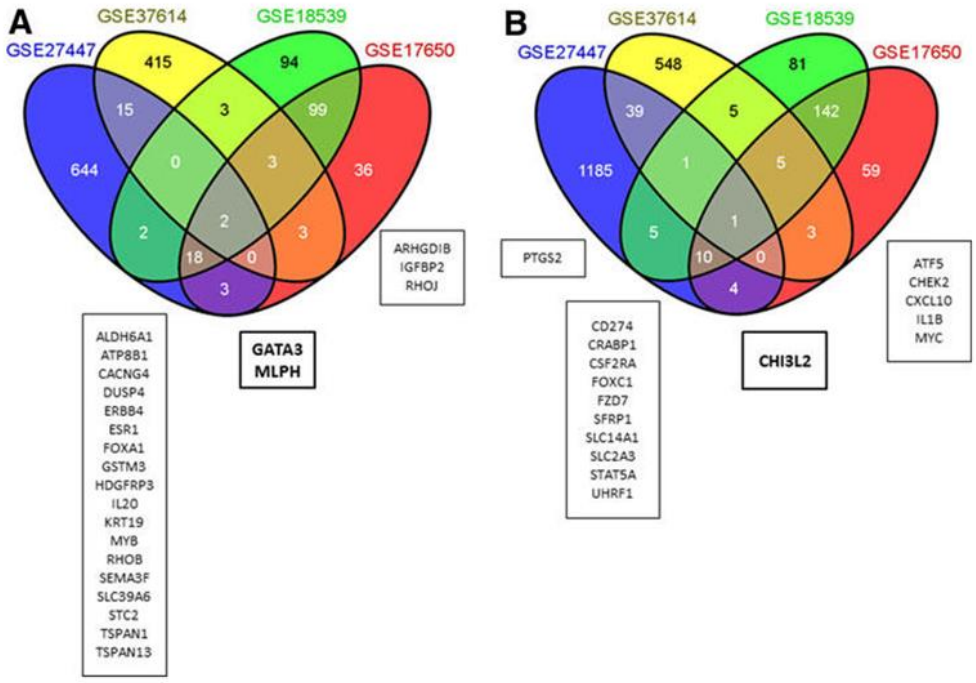


Figure 3.15 Comparison of the DEG sets in TNBC datasets (A) Downregulated DEGs, (B) Upregulated DEGs. Core DEGs are shown in boxes.

3.3.2.2 Reconstruction and analysis of active PPI subnetworks

Reconstructions of active PPI subnetworks in TNBC were based upon proteins associated with 40 core DEGs and their physical interactors. As a result, the downregulated interaction subnetwork consisted of 672 binary interactions between 615 proteins (Figure 3.16A), and the upregulated interaction subnetwork included 837 interactions between 755 proteins (Figure 3.16D). Analyses of reconstructed networks identified a set of proteins as hubs (i.e., nodes exhibiting high topological centrality in the graph (Figure 3.16B and 3.16C). These hub proteins deserve attention for further experimental studies, since they indicate significant potential for being candidate biomarkers for TNBC.

Downregulated hub proteins were classified into six groups: i) receptors (ESR1, ERBB4), (ii) transcription factors (MYB, GATA3, FOXA1), (iii) Rho related proteins (RHOB, GDIR2), (iv) enzymes (GSTM3), (v) cytokeratin (K1C19), and (vi) stanniocalcin (STC2). Upregulated hub proteins were classified into four groups: (i) receptors (CSF2RA, FZD7), (ii) transcription factors (MYC, CEBPB, TF65, ATF5, STAT5A), (iii) cytokines (IL1b, CXC10), and (iv) enzymes (CHIK2, UHRF1, PGH2) (Table 3.10). UBC was included in both lists since it has more than 10,000 interactions in protein interaction databases; since it came out in all hub lists, we excluded UBC in further analyses.

In the present study, we proposed 40 genes (23 downregulated, and 17 upregulated), which encode proteins with heterogeneous functionality, as core-DEGs, since these genes were commonly identified in at least three datasets (Figure 3.15). Furthermore, active PPI networks were reconstructed around core-DEGs, and via topological analyses 22 hub proteins were identified (Figure 3.16). These proteins mainly consisted of receptors, transcription factors, enzymes, and cytokines (Table 3.9). Seven downregulated hub proteins were previously reported with TNBC: ESR1, MYB, ERBB4, K1C19 (KRT19), GATA3, STC2, and FOXA1.

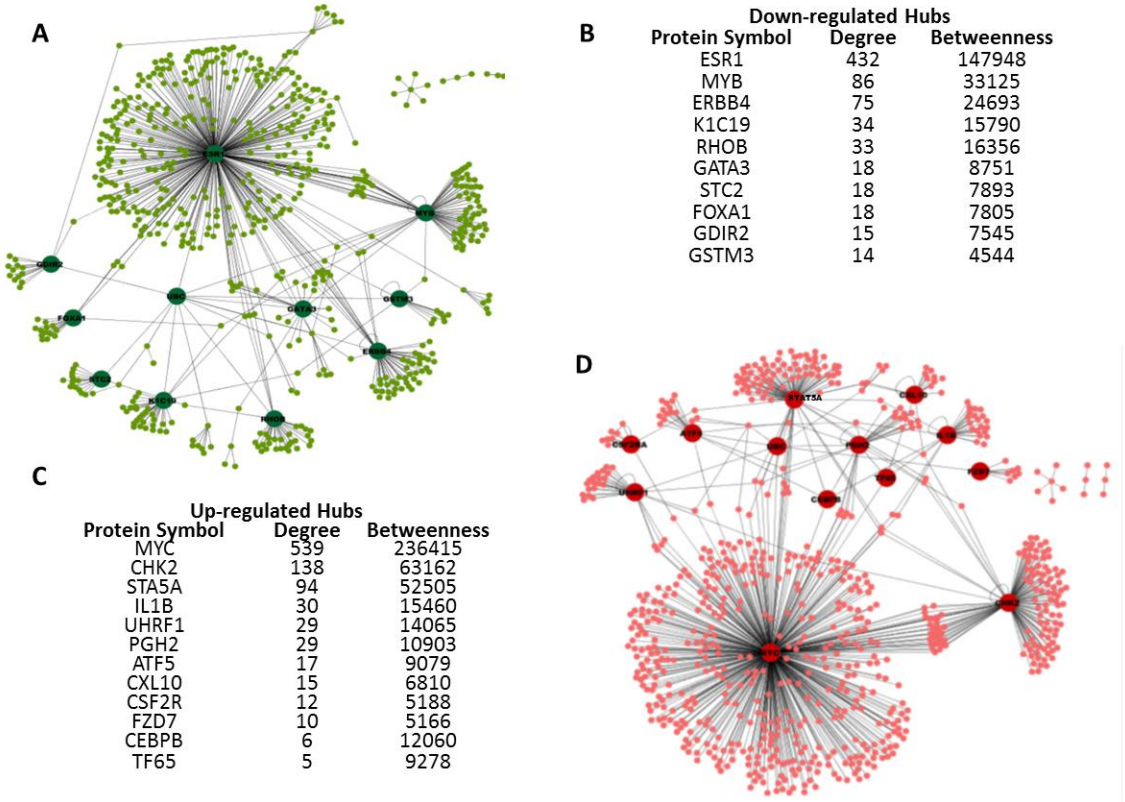


Figure 3.16 Reconstruction of PPI network of TNBC. (A) Downregulated core network; (B) Topological properties of downregulated hubs; (C) Topological properties of upregulated hubs; (D) Upregulated core network.

In addition, nine upregulated hub proteins were also previously associated with TNBC tumors: CHEK2, IL1b, MYC, STAT5A, UHRF1, TF65 (RELA), CEBPB, FZD7, and PGH2 (PTGS2). The present analyses proposed six novel proteins, which have not been reported in TNBC patients previously. These include three downregulated hub proteins, RHOB, GDIR2 (ARHGDIB), and GSTM3, and three upregulated hub proteins, ATF5, CSF2RA, and CXCL10. Among those, two downregulated proteins ESR1 (estrogen receptor 1) and ERBB4 (receptor tyrosine-protein kinase erbB-4) and two upregulated proteins, FZD7 (frizzled homolog 7) and CSF2RA (granulocyte-macrophage colony-stimulating factor receptor subunit alpha) were receptors that play a role in mainly four signaling pathways: estrogen, ERBB, WNT, and JAK-STAT signaling pathways.

Table 3.10 Hub proteins of the reconstructed PPI networks in TNBC.

Protein Symbol	Molecular Function	Direction of regulation	Novelty	References
ESR1	Receptor	Down	Known	Bender and Gabhann, 2013; Yu et al., 2013; Idowu et al., 2012
MYB	Transcription Factor	Down	Known	Bender and Gabhann, 2013
ERBB4	Enzyme	Down	Known	Hicks et al., 2013
K1C19	Keratin	Down	Known	Lehmann et al., 2011; Sparano et al., 2011
GATA3	Transcription Factor	Down	Known	Lee et al., 2013; Chu et al., 2012
STC2	Glycoprotein	Down	Known	Tate et al., 2012
FOXA1	Transcription Factor	Down	Known	Bender and Gabhann, 2013; Lee et al., 2013
CHEK2	Enzyme	Up	Known	Hicks et al., 2013
IL1β	Chemokine	Up	Known	Kuo et al., 2012
MYC	Transcription Factor	Up	Known	Lehmann et al., 2011; Liu et al., 2012; Horiuchi et al., 2012
STAT5A	Transcription Factor	Up	Known	Liu et al., 2012; Britschgi et al., 2012
UHRF1	Enzyme	Up	Known	Szarc vel Szic et al., 2014
TF65, RELA	Transcription Factor	Up	Known	Howe et al., 2012; Zhu et al., 2013
CEBPB	Transcription Factor	Up	Known	De Santi et al., 2011; Galluzzi et al., 2012
FZD7	Receptor	Up	Known	Yang et al., 2011; Dey et al., 2013
PGH2	Enzyme	Up	Known	Zhou et al., 2013; Hsia et al., 2013
RHOB	Signaling protein	Down	Novel	-
ARHGDIB	Signaling protein	Down	Novel	-
GSTM3	Enzyme	Down	Novel	-
ATF5	Transcription Factor	Up	Novel	-
CSF2RA	Receptor	Up	Novel	-
CXCL10	Chemokine	Up	Novel	-

ESR1 protein is an estrogen hormone receptor that affects cellular proliferation and differentiation in target tissues. Reduced level of ESR1 expression was observed in TNBC tumors (Bender and Gabhann, 2013; Idowu et al., 2012; Yu et al., 2013). ERBB4 (also named as HER4) is classified as a family member of the epidermal growth factor (EGF/ERBB) of receptor tyrosine kinases, which have significant roles in signaling pathways that are parts of cell division, migration, adhesion, differentiation, and apoptosis (Yarden and Sliwkowski, 2001). The ERBB family was proposed as a candidate for therapeutic studies (Saxena and Dwivedi, 2012). Downregulation of ERBB4 expression was reported in TNBC (Cascione et al., 2013; Hicks et al., 2013). FZD7 is a family member of the 'frizzled' gene family and encodes 7-transmembrane domain proteins that are receptors for WNT signaling proteins. Overexpression of several WNT pathway genes, including FZD7, was observed in TNBC patients (Dey et al., 2013; Yang et al., 2011). Moreover, FZD7 was reported as a crucial receptor for tumorigenesis of TNBC and it was proposed as a biomarker and a potential therapeutic target for TNBC (Yang et al., 2011). CSF2RA is a low affinity receptor that transduces signal resulting in proliferation, differentiation, and functional activation of hematopoietic cells. It further promotes JAK-STAT signaling pathway leading to activation of JAKs, and thereby regulating cell proliferation. Bailey and co-workers (2005) reported that CSF2RA was downregulated in MDA-MB-231 human breast cancer cells transfected with maspin. When genomic reprogramming in breast cancer tissues following treatment with 5-aza-2¢- deoxycytidine was analyzed, upregulation of CSF2RA expression was observed (El Baroudi et al., 2013). On the other hand, the association between CSF2RA receptor and TNBC is being reported for the first time in the present study. Three downregulated TFs, GATA3 (trans-acting T-cellspecific transcription factor GATA-3), MYB (transcriptional activator Myb), and FOXA1 (forkhead box protein A1), and five upregulated TFs, MYC (myc proto-oncogene protein), STAT5A (signal transducer and activator of transcription 5A), ATF5 (activating transcription factor 5), RELA (transcription factor p65), and CEBPB (CCAAT/enhancer-binding protein beta) were also reported as hub proteins. GATA3, a family member of zinc finger transcription factors, plays a role in human growth and differentiation. Dydensborg and colleagues (2009) addressed potential roles for GATA3 in breast tumor lung metastasis and progression using an aggressive breast cancer cell

line and concluded that GATA3 inhibits primary breast tumor outgrowth and reduces lung metastatic burden by regulating key genes involved in metastatic breast tumor progression. The expression of GATA3 in breast cancer is correlated with the ER status, and loss of GATA3 causes tumorigenesis in ER + breast cancers (Mehra et al., 2005; Voduc et al., 2008). GATA3 is involved in differentiation of breast luminal cell and progression of breast cancer (Cimino-Mathews et al., 2013; Voduc et al., 2008). Additionally, GATA3 was proposed as an effective biomarker for metastatic breast carcinoma, especially triplenegative and metaplastic carcinomas (Chu et al., 2012; Cimino-Mathews et al., 2013). MYB is a DNA-binding transcription factor involved in cell proliferation, differentiation, and apoptosis (Drabsch et al., 2010). MYB acts as an oncogene that suppresses the differentiation and apoptosis of human breast cancer cells (Drabsch et al., 2010; Thorner et al., 2010). MYB is a target of estrogen signaling and known to be associated with ER positivity (Drabsch et al., 2007). However, there is limited research on the control mechanism of breast carcinoma progression and dissemination by MYB (Knopfova et al., 2012). FOXA1, also known as hepatocyte nuclear factor 3-alpha (HNF-3A), is a member of the forkhead class of DNA binding proteins, and plays a role in apoptosis, establishment of tissue-specific gene expression, and regulation of gene expression in differentiated tissues. FOXA1 is required for estrogen-induced ER transcriptional activation and repression, and FOXA1 also regulates ER expression in breast cancer cells, especially luminal A type (Nakshatri and Badve, 2009; Perou et al., 2010). It was also confirmed that FOXA1 is a TNBC-related gene (Bender and Gabhann, 2013; Prat et al., 2013). In TNBC patients, the low expression levels of the ESR1 and PR, which were targets of these downregulated hub-TFs (GATA3, MYB, and FOXA1), were reported (Bender and Gabhann, 2013; Roll et al., 2013). MYC, one of the upregulated hub-TFs, is a proto-oncogene and plays roles in multiple signaling pathways, including those involved in cell growth, cell proliferation, metabolism, microRNA regulation, cell death, and cell survival (Dang, 1999; Eilers and Eisenman, 2008; Meyer and Penn, 2008). Furthermore, MYC signaling has been shown to be upregulated in high-grade mammary tumors with presumptive cancer stem cell properties (Ben-Porath et al., 2008; Wong et al., 2008). Several studies have identified a MYC transcriptional gene signature associated with the basal molecular subtype (Alles et al., 2009; Chandriani et al., 2009; Gatza et al., 2010). In addition, the upregulation of

MYC in TNBC and changing expression of MYC regulatory genes as a result of increasing activity of the MYC pathway were also reported (Horiuchi et al., 2012; Lehman et al., 2011; Liu et al., 2012b). STAT5A, a member of the STAT family of transcription factors, is also reported to be activated in breast cancer, though it is often associated with different tumor subtypes (Walker et al., 2009). STAT5A expression was upregulated by MYC, and altered MYC expression elevated antiapoptotic survival pathway *in vivo* involving STAT5A. It was suggested that MYC needed the activation of STAT5 as a mechanism to avoid apoptosis in the processes of tumor initiation and progression (Blakely et al., 2005). In addition, the upregulation of STAT5A by MYC was also reported in a subset of TNBC patients, indicating the promotion of metastasis (Liu et al., 2012a). RELA is classified in the NF-jB family that is involved in inflammation, immunity, differentiation, cell growth, tumorigenesis, and apoptosis. Howe and co-workers (2012) demonstrated that NF-jB activity was increased 6-fold in suspended TNBC cells, and identified RELA and NF-jB1 as the transcription factors responsible for suspension-induced upregulation of TRKB and NTF3 in anoikis-resistant breast cancer cells. In another report, upregulation ofRELAwas demonstrated in TNBC (Zhu et al., 2013). CEBPB is an important transcriptional activator in the regulation of genes, playing a role in immune and inflammatory responses (Hasko et al., 2008). Dysregulation of CEBPB was associated with the development and metastatic progression of breast cancer (Gomis et al., 2006) and upregulation of CEBPB in TNBC cell lines were reported (De Santi et al., 2011; Galluzzi et al., 2012). ATF5 is a member of the ATF/cAMP response element binding protein (CREB) family of basic zipper proteins that binds to several transcription regulatory DNA elements. Although several studies reported ATF5-induced promotion of gene regulation and cancer-specific cell survival in breast cancer cells (Dluzen et al., 2011; Li et al., 2009), its role in TNBC has not yet been investigated. Among the hub proteins, GSTM3 (glutathione S-transferase Mu 3) enzyme was downregulated, and CHEK2 (checkpoint kinase 2), UHRF1 (E3 ubiquitin-protein ligase), and PTGS2 (prostaglandin G/H synthase 2) enzymes were upregulated. GSTM3 functions by binding glutathione to electrophilic xenobiotics to make them more water soluble and easier to excrete. Previously, Bieche and co-workers (2004) displayed higher mRNA levels of GSTM3 in ERa-positive breast tumors than in normal breast tissue, and higher mRNA levels of GSTM3 in ERa-positive breast tumors

than in ERa-negative breast tumors. Li et al. (2009) listed GSTM3 among 24 genes that represent signature of breast cancer. In addition, Yu and co-workers (2010) indicated that the catalyzing ability of GSTM3 in normal breast tissue contribute to protect against breast cancer. More recently, GSTM3 was reported as a novel marker in colon cancer (Meding et al., 2012). However, the possible role of GSTM3 in TNBC has not yet been investigated. CHEK2 is serine/threonine-protein kinase that is involved in checkpoint-mediated cell cycle arrest, activation of DNA repair, and apoptosis. It was also defined as a candidate tumor suppressor in breast cancer (Bartek and Lucas, 2003; Bosch et al., 2010) and was recently proposed as a potential biomarker for TNBC (Hicks et al., 2013). CHEK2 directly modulates the activities of TP53 and BRCA1 through phosphorylation. Thus, targeting these pathways was proposed for therapeutics in TNBC (Tommiska et al., 2008). UHRF1 is a member of a subfamily of RING-finger type E3 ubiquitin ligases and is a key epigenetic regulator that impacts DNA methylation and chromatin modification. Previously, UHRF1 increased the proliferation of breast cancer cells by apoptosis inhibition, and therefore it was suggested that UHRF1 could be a new therapy target for breast cancer (Li et al., 2012). Komatsu and co-workers (2013) documented upregulation of UHRF1 in TNBC. Treatment of TNBC cells with withaferin A, which is an attractive phytochemical under investigation in various preclinical studies for treatment of different cancer types, lowered the expression of UHRF1 (Szarc vel Szic et al., 2014). PGH2 (PTGS2) encodes prostaglandin-endoperoxide synthase protein, also known as cyclooxygenase (COX), which is the crucial enzyme in prostaglandin biosynthesis, and acts both as a dioxygenase and as a peroxidase. Alteration of COX-2 expression was related with aggressive pathologic parameters and poor prognosis in many different carcinomas such as head and neck, colorectal, ovarian, and breast cancers (Ferrandina et al., 2002). It was also reported that COX-2 expression was positively correlated with poor prognosis in patients with TNBC, but not in patients with breast cancer and non-TNBC so it was proposed that COX-2 could be a useful prognostic factor of TNBC (Zhou et al., 2013). Another study demonstrated that the metastasis of TNBCs was promoted through the upregulation of EGFR and COX2 (Hsia et al., 2013). The chemokines IL1b (interleukin 1, beta) and CXCL10 (C-X-C motif chemokine 10) were upregulated hubs. IL1b is a family member of the interleukin 1 cytokine that is involved in various cellular

processes including proliferation, differentiation, and cell survival/apoptosis. IL1b was reported among 45-gene signature of TNBC, and was shown to regulate TGFb, which plays a role in cellular proliferation network (Kuo et al., 2012). IL1b is known to be expressed in 90% of invasive breast carcinomas, and high IL1b content in tumors correlated well with tumors aggressiveness. Previously, IL1b was suggested as an effective target for cancer therapy of TNBC (Roberti et al., 2012). CXCL10 is involved in a chemokine CXC subfamily and is the ligand for the receptor CXCR3. CXCL10 plays role in stimulation of monocytes, natural killer, and T-cell migration, and modulation of adhesion molecule expression. Datta and coworkers (2006) proposed that the Ras-induced regulation of CXCL10 and its receptor might be significantly involved in the development of human breast cancer through cancer cell proliferation. Mulligan et al. (2013) reported that CXCL10 affects the tumor microenvironment, and is involved in tumor invasiveness and progression as well. They proposed that the CXCL10–CXCR3 axis can be as a potential target in BRCA1 and basal breast cancers. On the other hand, there is TNBC. The other hub proteins include KRT19, STC2, ARHGDIB, and RHOB. Downregulated KRT19 is a member of the keratin family. It was previously reported that TNBC tumor subtypes present differential expression of both basal-like cytokeratins (KRT5, KRT6A, KRT6B, KRT14, KRT16, KRT17, KRT23, KRT81) and luminal cytokeratins (KRT7, KRT8, KRT18, KRT19) (Lehmann et al., 2011). Moreover, the low expression level of KRT19 was confirmed in some other TNBC studies (Lehmann et al., 2011; Sparano et al., 2011). STC2 (stanniocalcin 2), a secreted homodimeric glycoprotein, is expressed in a wide variety of tissues, and may have autocrine or paracrine functions. Charpentier and coworkers (2000) confirmed that STC2 expression was positively associated with the ER status of breast carcinomas. The association between STC2 expression with longer disease free survival was also reported (Esseghir et al., 2007). Additionally, the relation of STC2 with cell invasion and metastasis in cells was documented (Tate et al., 2012). Two Rho related genes, ARHGDIB (rho GDP dissociation inhibitor (GDI) beta) and RHOB (ras homolog family member B), were downregulated and their association with TNBC was reported for the first time in the present study. ARHGDIB, also known as D4-GDI, was originally identified in T- and B- cells (Scherle et al., 1993), but it was later documented as being expressed in nonhematopoietic neoplasms cells as well, including ovarian

(Tapper et al., 2001), colon (Ota et al., 2004), and gastric (Cho et al., 2009) carcinoma. Zhang and Zhang (2006) presented that D4-GDI is selectively expressed in invasive breast cancer cell lines as compared to normal mammary epithelial cells. Then they identified that D4-GDI protein is upregulated in the early stages of cancer progression (hyperplasia and benign), followed by a decrease in malignant tumors and metastatic lymph nodes compared with normal breast tissues, suggesting the dual roles of D4-GDI in regulation of breast cancer progression (Rivera Rosado et al., 2011). RhoB is a family member of small GTPases, which are significant compounds involved in growth factor signal transduction pathways (Guilluy et al., 2011). RhoB is assumed as a modulator of growth factor signals since RhoB maintains the cellular trafficking of some key signaling elements such as EGF receptor (EGFR), platelet-derived growth factor receptor (PDGFR), and functions as sensor of cellular stress to provide an early response to EGF, TGF-b, and Src activation (Huang et al., 2008; Liu et al., 2001). The expression level of RhoB is altered in various human cancers such as ovarian (Couderc et al., 2008), lung (Mazieres et al., 2004), and head and neck squamous cell carcinoma (Adnane et al., 2002). Moreover, RhoB acts as a tumor suppressor in cancer that restrain EGF receptor (EGFR) to bind the cell surface, as well as inhibit Akt signaling (Kazerounian et al., 2013; Prendergast et al., 2001). Although there has been no investigation on TNBC cells, RhoB functions to promote expression of ERa and PR in a manner correlated with cell proliferation in human breast cancer (Medale-Giamarchi et al., 2013).

3.3.2.3 Functional enrichment analyses

According to pathway enrichment analyses of DEG sets, signaling pathways including B-cell receptor signaling, NK signaling, T-cell receptor signaling, chemokine signaling, JAK-STAT signaling, and apoptosis were significantly upregulated, whereas adherens, tight and gap junction pathways, regulation of actin cytoskeletal, complement and coagulation cascades, TGF-beta signaling and endocytosis pathways were downregulated in TNBC (Figure 3.17A). On the other hand, DEG sets included both upregulated and downregulated members of MAPK signaling, ErbB signaling, p53 signaling, VEGF signaling, and mTOR signaling pathways (Figure 3.17B).

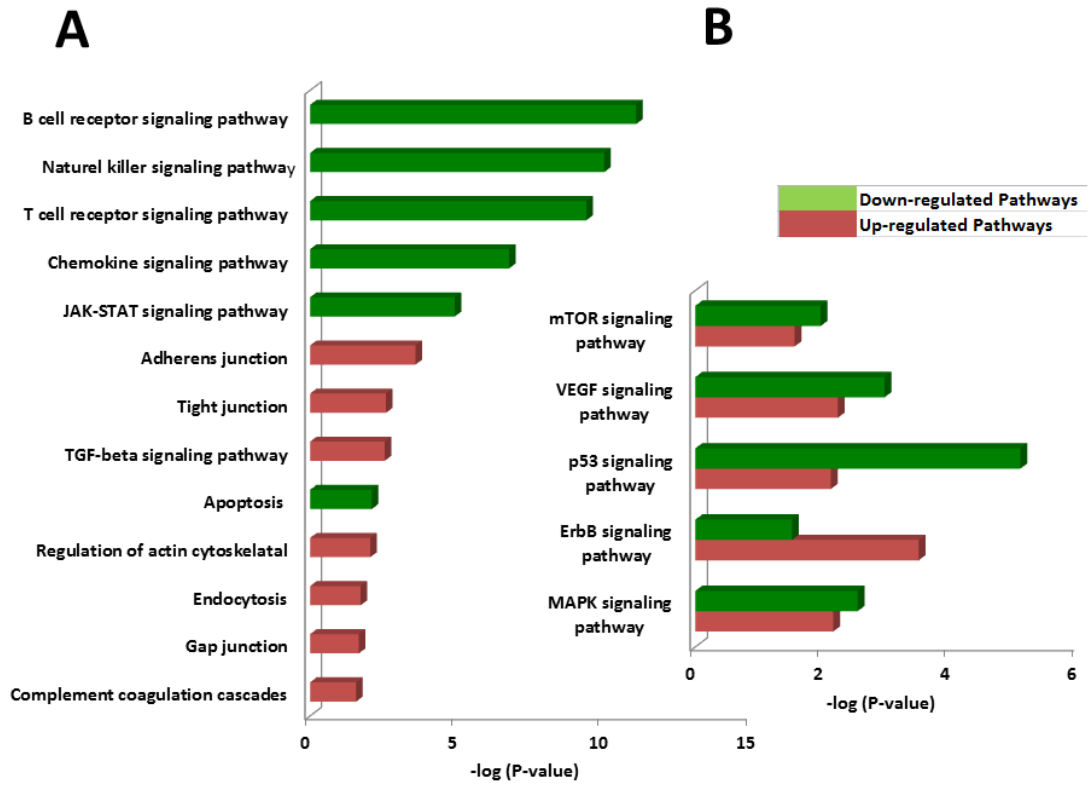


Figure 3.17. Enrichment results of molecular pathways. (A) Pathways either upregulated or downregulated. (B) Pathways including downregulated and upregulated members. Red: upregulation; green: downregulation.

3.3.2.4 Transcriptional regulatory subnetworks

Transcriptional regulation data for three downregulated (GATA3, FOXA1, MYB) and six upregulated (ATF5, CEBPB, FOXC1, MYC, STAT5A, RELA) hub TFs were assembled from the literature, and transcriptional regulatory subnetworks were reconstructed around these hub-TFs and their targeted DEGs. The transcriptional regulatory subnetwork around three downregulated hub-TFs consisted of 261 transcriptional regulatory relations between these TFs and 161 downregulated DEGs (Figure 3.18A). Proteins associated with the targeted genes were mainly playing a role in response to hormone stimulus (p-value: $2.6 \cdot 10^{-12}$), cell differentiation (p-value: $1.3 \cdot 10^{-8}$), response to estrogen stimulus (p-value: $1.4 \cdot 10^{-6}$), and apoptosis (p-value: $5.1 \cdot 10^{-3}$). Among the targets, four genes (ERBB2, RARA, BCL2L1, and BTG2) that participate in response to hormone stimulus (p-value: $6.5 \cdot 10^{-6}$) and regulation of anti-apoptosis (p-value: $1.1 \cdot 10^{-4}$) were regulated by all three TFs (Figure 3.18B). Another

remarkable result is that three genes (SEMA3F, KRT19, and STC2) among core DEGs were co-expressed (with Pearson correlation coefficients of 0.98 and higher) and coregulated by these downregulated hub-TFs (Figure 3.18C). The upstream regulators of hub-TFs, GATA3, MYB, and FOXA1, were also explored (Figure 3.19). ESR1, CTNNB1, CD28, TLR2, and EP300 were the only upstream regulators, which were differentially expressed in the TNBC datasets. These proteins play roles in various signaling pathways including glucocorticoid receptor, RhoGDI, p53, Wnt/b, ERK/MAPK, estrogen receptor, and NF-jB signaling. The transcriptional regulatory subnetwork around six upregulated hub-TFs consisted of 278 transcriptional regulatory relations between 5 TFs and 217 upregulated DEGs (Figure 3.20A). Proteins associated with the targeted genes were mainly involved in cell proliferation (p value: $1.3 \cdot 10^{-27}$) and regulation of apoptosis (p value: $5.4 \cdot 10^{-23}$) processes, and have a role in cell cycle (p value: $1.5 \cdot 10^{-8}$), p53 signaling (p value: $5.4 \cdot 10^{-8}$), and JAK-STAT signaling (p value: $4.4 \cdot 10^{-7}$) pathways. The highly connected module of the subnetwork consisted of 33 regulatory interactions between four upregulated hub-TFs and ten target genes regulated by more than three hub TFs (Figure 3.20B). The target genes, IFNG, IL6, and TNF, were regulated by all four TFs (RELA, CEBPB, STAT5A, and MYC). The module was enriched with the members of the JAK-STAT pathway (Gira et al., 2009; Vaclavicek et al., 2007). The transcriptional regulatory module including core-TFs and their targets among upregulated core DEGs (Figure 3.20C) included genes which were mainly involved in cell proliferation, immune system, apoptosis, and cell cycle. The roles of TFs, RELA, and CEBPB, and the targeted core DEGs, CXCL10, IL1b, and PTGS2, in TNF signaling were previously reported in several studies (Pan et al., 2010; Qi et al., 2009; Wang et al., 2011).

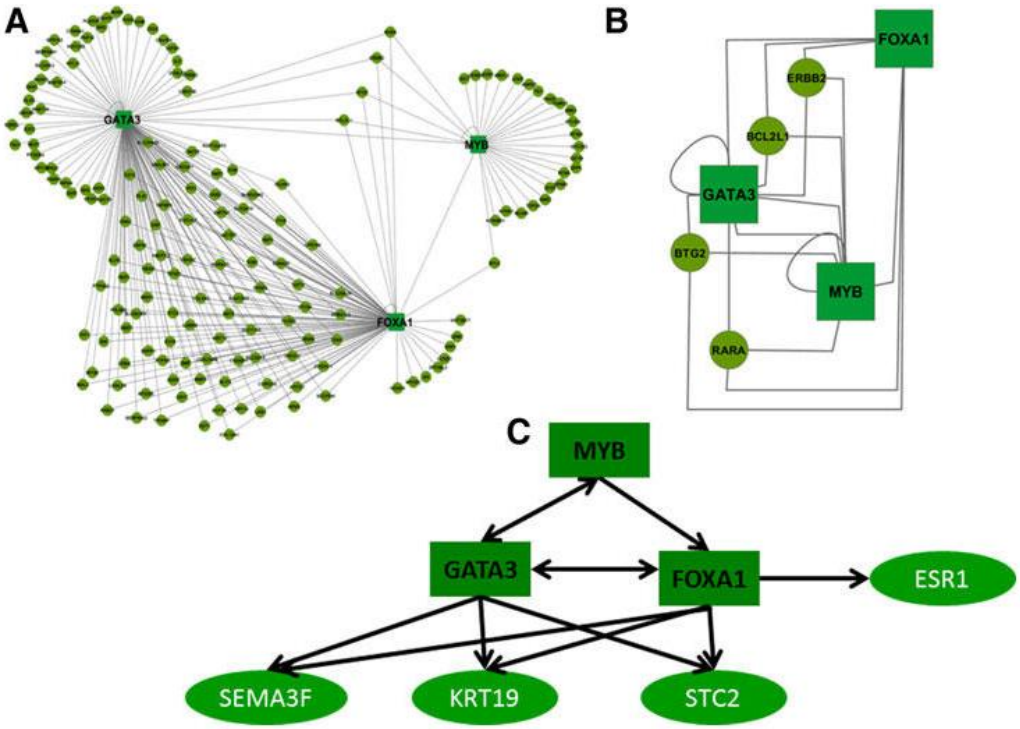


Figure 3.18 (A) Transcriptional regulatory subnetwork of downregulated hub-TFs. (B) Highly connected module in the regulatory subnetwork. (C) A transcriptional regulatory module including downregulated core-TFs and their targets among core DEGs.

Thirty-nine upstream regulators of MYC, RELA, and CEBPB were differentially expressed in TNBC datasets (Figure 3.21A). The pathway enrichment analysis of these genes indicated that 10 of these proteins were involved in MAPK signaling pathway (p-value: 1.9×10^{-5}). EGFR and TNF were the only regulators, which affect more than one TF. EGFR regulates CEBPB and MYC, and TNF regulates MYC and RELA. These regulations highlighted MAPK signaling, glucocorticoid receptor, colorectal cancer metastasis, IL6, ILK, and NF-jB signaling pathways (Figure 3.21B). Among these proteins, MYC, TNF, EGFR, and RELA were identified as biomarkers in a wide range of cancer types (Figure 3.21C). CEBPB was not defined as a specific biomarker previously but our analyses indicated CEBPB as a candidate biomarker for TNBC.

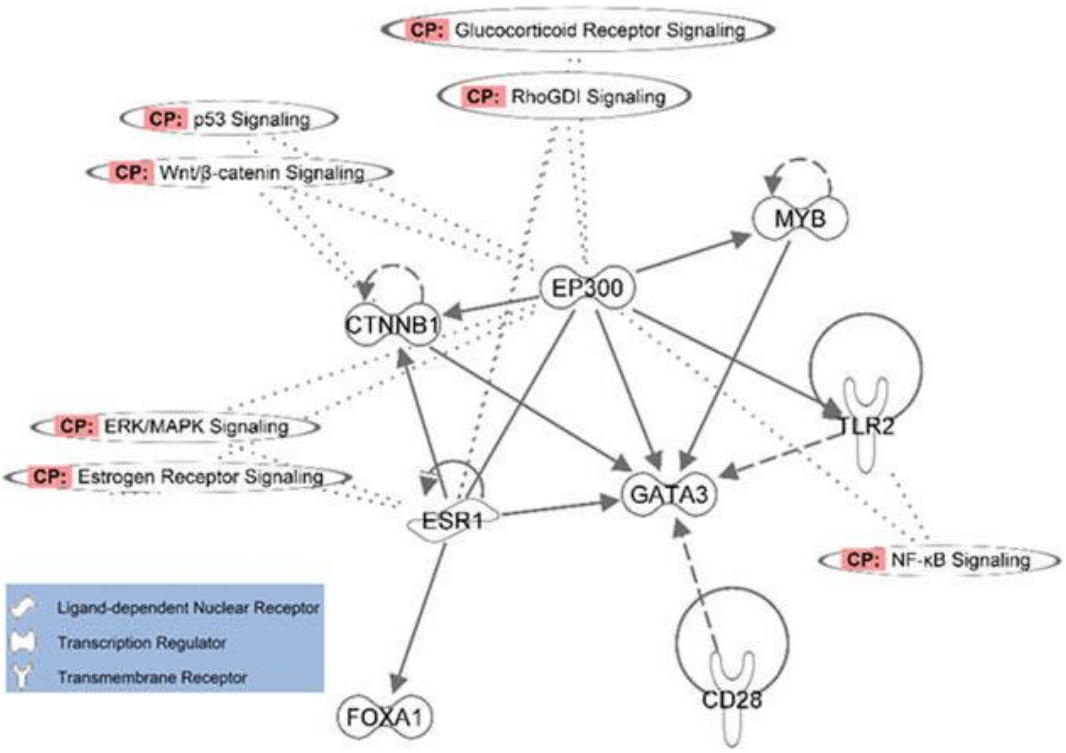


Figure 3.19 Upstream regulators of downregulated hub TFs.

The core down- and upregulated transcriptional regulatory networks (core TFs and their core targets) in which all genes were differentially expressed in at least three microarray datasets and more than 150 TNBC samples (Figure 2.18C and Figure 3.20C) were constructed individually. These networks specified the major regulatory mechanism of TNBC. The downstream targets, SEMA3F, KRT19, STC2, and ESR1, which were also ‘‘core DEGs,’’ and the upstream regulators CTNNB1, EP300, CD28, and TLR2 constructed the downregulated transcriptional regulatory network of TNBC. Although the role of SEMA3F was not completely enlightened in TNBC, Bender and Gabhann (2013) reported reduced expression level of SEMA3F in TNBC. SEMA3F was also explored as an inhibitor of tumor angiogenesis (Kessler et al., 2004) and characterized as a tumor suppressor gene (Xiang et al., 2002). Our study highlighted SEMA3F as a potential effective anti-angiogenic therapeutic target in TNBC.

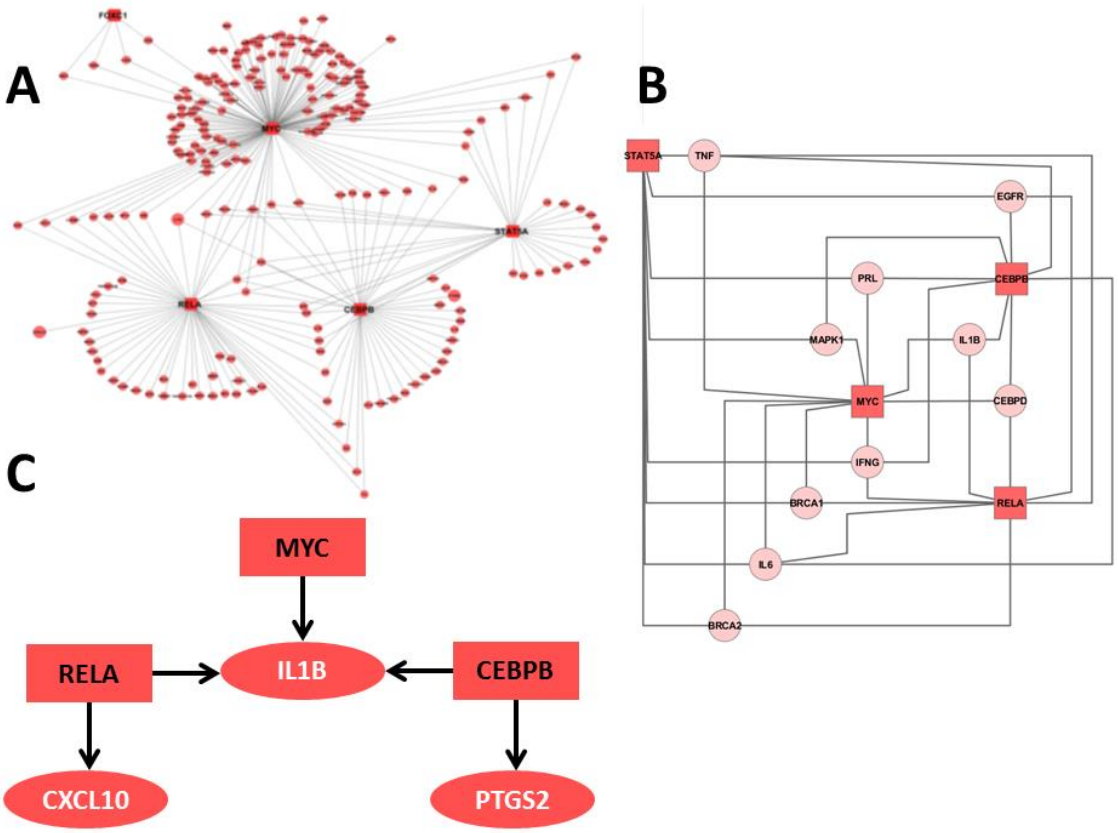
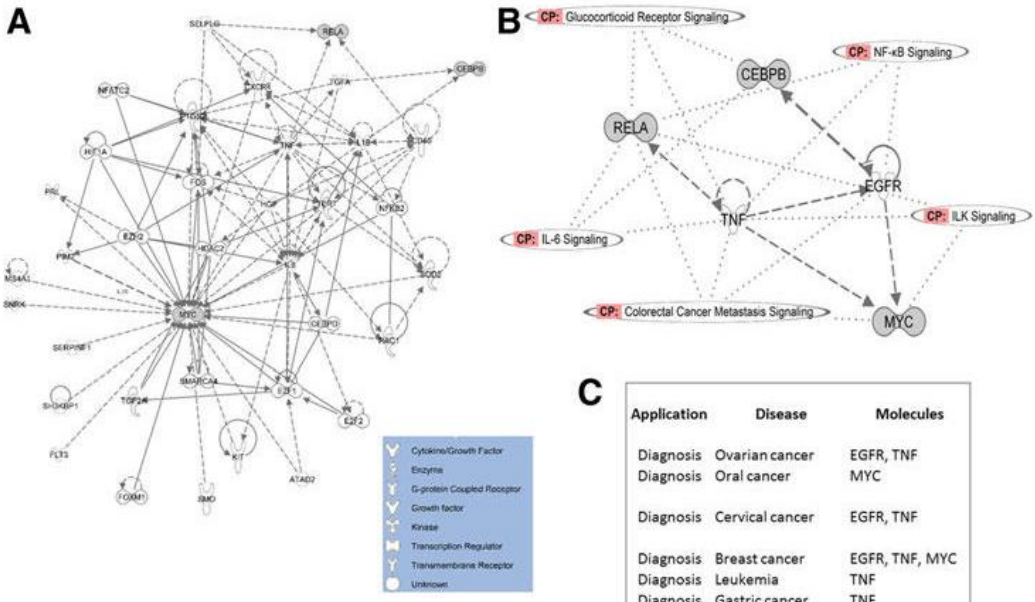


Figure 3.20 (A) Transcriptional regulatory subnetwork of upregulated hub-TFs. (B) The highly connected module in the regulatory subnetwork. (C) The transcriptional regulatory module including core-TFs and their targets among upregulated core DEGs.

Cytokeratins represent one of the most important elements of cytoskeleton in epithelial cells of humans and animals (Moll et al., 1982). Cytokeratins undergo the process of ubiquitination (taking place during proteasome-mediated protein degradation), which results in modification of structure and, indirectly, function of various proteins. Deformed cytokeratins disrupt shape of cytokeratin network in the cytoplasm, and destroy the natural cytoskeleton structure. Accumulation of the altered proteins increases aggressiveness and invasiveness character of the cells. Furthermore, processes of mitosis and apoptosis become disturbed, which promotes immortality of the altered cells (Iwaya and Mukai, 2005). Moreover, cytokeratins were applied in evaluation of the degree to which a tissue is abnormally differentiated (Galus and Wlodarski, 2007). Patients with CK-19 mRNA positive cells in their peripheral blood have a significantly increased probability for distant relapse and decreased overall survival (Stathopoulou et al., 2002).



C

Application	Disease	Molecules
Diagnosis	Ovarian cancer	EGFR, TNF
Diagnosis	Oral cancer	MYC
Diagnosis	Cervical cancer	EGFR, TNF
Diagnosis	Breast cancer	EGFR, TNF, MYC
Diagnosis	Leukemia	TNF
Diagnosis	Gastric cancer	TNF
Diagnosis	Colon cancer	EGFR
Diagnosis	Pancreatic cancer	RELA
Diagnosis	Bladder cancer	EGFR

Figure 3.21 (A) Upregulated upstream regulators and their hub TFs network. (B) Upstream regulators which affect more than one TFs. (C) Biomarkers in different cancer types.

The cytokeratins are proteins expressed in a majority of epithelial tumors and in most of the cells of these tumors (Pantel et al., 1999). While the administration of cytotoxic chemotherapy does not completely eliminate CK-positive tumor cells (Braun et al., 2000), possibly because of the dormant and noncycling nature of these cells, other treatments, such as monoclonal antibodies, may prove to be more effective (Braun et al., 1999). The upregulated transcriptional regulatory network included transcriptional factors, MYC, RELA, and CEBPB and the downstream targets, CXCL10, IL1b, and PTGS2, which were also ‘‘core DEGs’’ (Figure 22C). IL1b and CXCL10 are chemokines that play roles as regulators of cell proliferation, differentiation, migration, and invasion. IL1b were reported to be related with tumor aggressiveness and was proposed to be an effective target for TNBC therapy (Roberti et al., 2012). Furthermore, although another upstream regulator, CXCL10, was suggested as a potential target for

basal breast cancers (Mulligan et al., 2013), the role of CXCL10 in TNBC has not been investigated yet. PTGS2 is the crucial enzyme in prostaglandin biosynthesis. It was demonstrated that the metastasis of TNBCs was promoted through the upregulation of EGFR and COX-2 (Hsia et al., 2013). IL1b (as an inflammatory cytokine), PTGS2 (as a mediator of synthesis of inflammatory), and CXCL10 (as leukocyte recruitment), together with transcription factors RELA and CEBPB, indicated the significance of TNF signaling pathway. The role of TNF signaling pathway in TNBC has not been explored yet, although the roles of WNT signaling pathway and JAK-STAT signaling pathways in TNBC were investigated previously. Regulatory network analysis highlighted the importance of JAK-STAT and TNF signaling pathways and pinpoint specific targets (MYC, CEBPB, IL1b, CXCL10, and PTGS2) for therapeutic targeting since they were identified in more than 150 TNBC samples. The present study for the first time demonstrated an extensive characterization of the transcriptional regulatory network with both down- and upstream regulators of the TNBC that elucidate the candidate therapeutic targets. The effective therapeutics that disturb the transcriptional regulatory network of TNBC may also affect the function of multiple downstream genes and their associated pathways simultaneously so the efficiency of therapeutics increases within the treatment period.

3.3.3 Integrative analysis of esophageal squamous cell carcinoma

3.3.3.1 Differentially expressed genes

We obtained publicly available microarray datasets (Table 1) from the NCBI-GEO for esophageal squamous cell carcinoma (ESCC), which includes normal and ESCC tissue samples consisting of 182 paired-tissues from 91 patients.

As a result of statistical analyses of 5 datasets, a total of 1582 and 1078 genes were identified as down and up-regulated, respectively (Table 3.11), and analyzed further. Although difference between the number of down- and up-regulated genes were not significant in Affymetrix platforms, the difference was considerable in Agilent Platform which may indicate a possible effect of platform choice in detection of DEGs.

Table 3.11 Statistical results for differentially expressed genes of ESCC.

NCBI-GEO ID	# of down-regulated genes	# of up-regulated genes	Total # of genes
GSE17351	228	307	535
GSE20347	738	625	1363
GSE23400	283	271	554
GSE29001	321	312	633
GSE45350	1005	363	1368
Total (unique)	1582	1078	2660

30 down-regulated and 21 up-regulated genes, so-called core DEGs, were differentially expressed in all of the 91 ESCC patients (Figure 3.22).

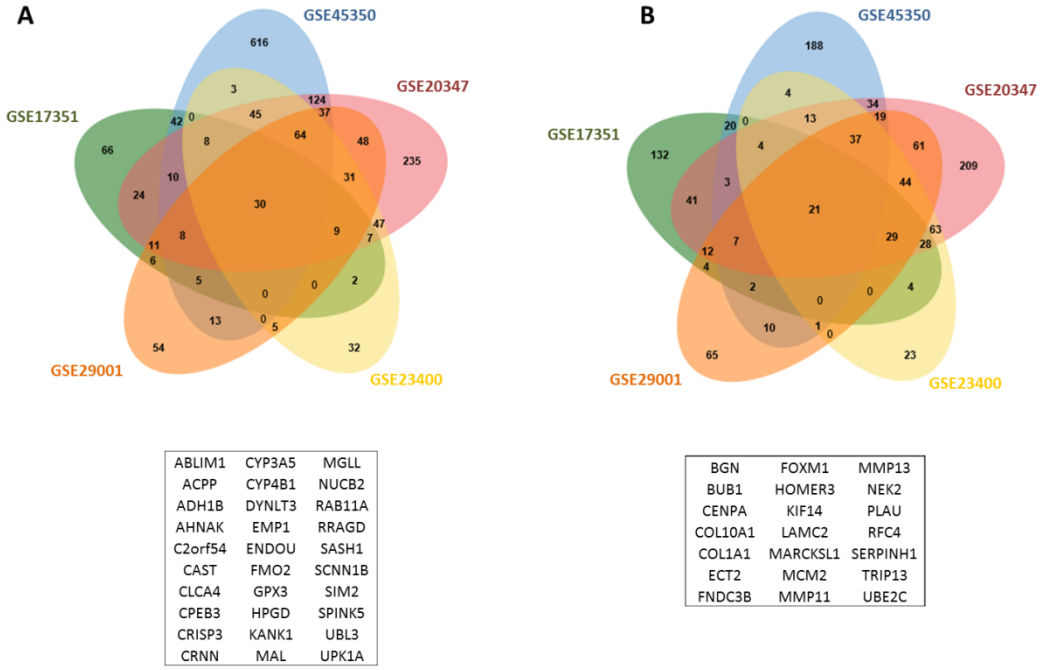


Figure 3.22 Comparison of the DEGs. (A) Down-regulated genes, (B) Up-regulated genes. Core DEGs are represented below the Venn schemes.

Among down-regulated core DEGs, 9 genes (ACPP, ADH1B, CYP3A5, CYP4B1, ENDOU, FMO2, GPX3, HPGD, and MGLL) encode enzymes, 6 DEGs (ABLIM1, AHNAK, KANK1, MAL, RAB11A and SIM2) were playing role in nervous system, 4

DEGs (CAST, CLCA4, CRNN and NUCB2) were related to calcium metabolism, SCNNB1 was related to sodium metabolism, 4 DEGs (KANK1, MGLL, RRAGD and SASH1) were signaling proteins, 2 DEGs (CRISP3 and SPINK5) were involved in immune response, UPK1A and SPINK5 were playing role in cell differentiation, CREB3, DYNLT3 and EMP1 were involved in cell proliferation, cell cycle and translation, respectively. Additionally, the biological processes of C2orf54 and UBL3 were unknown. 8 (CRNN, EMP1, GPX3, HPGD, KANK1, MAL, SASH1 and UPK1A) of 30 down-regulated genes were identified as TSGs in TSGene Database.

Among up-regulated core DEGs, 7 DEGs (BUB1, MMP11, MMP13, MCM2, NEK2, PLAU, and UBE2C) were enzymes, 5 DEGs (BGN, COL1A1, COL10A1, MMP13 and SERPINH1) were related with collagen, 9 DEGs (BUB1, CENPA, FOXM1, LAMC2, MCM2, NEK2, UBE2C, ECT2, RFC4) were associated with cell cycle, cell division and mitosis, 2 DEGs (HOMER3 and KIF4) were signaling proteins, FNDC3B and MARCKSL1 were involved in fat cell differentiation and cell proliferation, respectively and TRIP13 was a hormone dependent transcription factor. Furthermore, MMP11 was identified as TSG in the TSGene Database.

The associations of 6 down-regulated core DEGs (ACPP, C2orf54, DYNLT3, ENDOU, FMO2, and KANK1) and 5 up-regulated core genes (COL10A1, FNDC3B, HOMER3, MARCKSL1, and RFC4) with ESCC were reported for the first time in the present study.

The key molecules involved in ESCC: Receptors, Transcriptional Factors and Tumor Suppressor Genes: We explored down- and up-regulated receptors by curating DEGs to highlight the signaling mechanisms associated with ESCC. 61 and 56 receptors were down and up-regulated, respectively. Interleukin receptors, lysophosphatidic acid receptors, nuclear receptors and olfactory receptors were mainly down-regulated. Whereas up-regulated receptors included high affinity immunoglobin epsilon receptors, low density lipoprotein receptors and tumor necrosis factor receptors. Several members of the G-protein coupled receptors were either down- or up-regulated. Four down-regulated (CXCR2, EPHA1, EPHA2, and TLR3) and seven up-regulated receptors (F2R, HMMR, IGF1R, MET, PLAUR, TFRC, and TNFRSF10B) have previously been reported to be associated with ESCC.

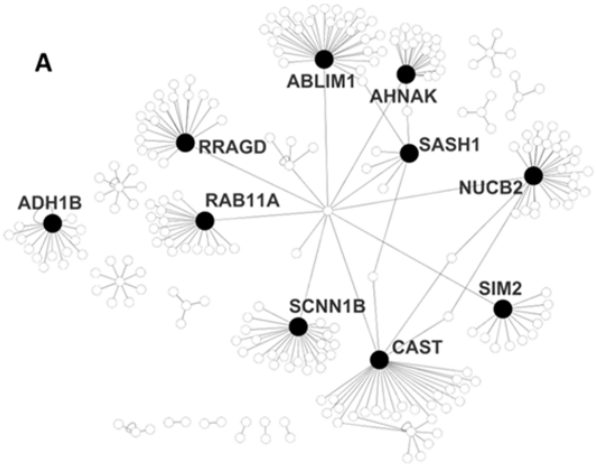
In addition, 15 and 17 transcriptional factors (TFs) were down- and up-regulated, respectively. Three main families were remarkable in up-regulated TFs; basic leucine zipper family, E2F family and runt-related TFs, whereas down-regulated TFs were not enriched in any superfamily. Two down-regulated (CLOCK and SSBP2) and six up-regulated TFs (E2F1, HLTF, RUNX3, STAT1, TCF3 and TWIST) were previously shown to be associated with ESCC.

Tumor suppressor genes (TSGs) are a group of guardian genes that play critical roles in controlling cell proliferation processes such as cell cycle checkpoints and apoptosis. They may influence cancer development by loss-of-function. Identification of these genes and understanding their function is critical for further investigation of tumorigenesis. TSGs were explored through TSGene Database (Zhao et al., 2013) and manual curation in PubMed.

Among the 716 TSGs present in TSGene Database, 68 and 58 TSGs were down and upregulated, respectively. In addition, manual curation of literature (for >300 publications) resulted in 78 TSGs associated with ESCC. Among those 12 TSGs CACNA2D3, CDKN2A, CRNN, EMP1, FBXW7, KLF5, PLCD1, PTK6, RAB25, RBMS3, S100A14, UPK1A) were down-regulated and 4 TSGs (COL4A1, IGFBP3, RUNX3, TP63) were up-regulated, in the present study.

3.3.3.2 Reconstruction and analysis of active PPI subnetworks

The interactions associated with proteins corresponding to core 30 down-regulated and 21 up-regulated genes were collected that the networks related with down-regulated and up-regulated genes comprised of 243 interactions among 245 proteins and 775 interactions among 728 proteins, respectively (Figure 3.23). Analyses of reconstructed networks identified a set of proteins as hubs which deserve attention for further experimental studies, since they indicate significant potential for being candidate biomarkers for ESCC.



B

Node Name	Degree	Node Name	Betweenness
ABLIM1	29	CAST	471231
NUCB2	27	ABLIM1	460213
CAST	26	NUCB2	419376
SCNN1B	23	AHNAK	301046
AHNAK	19	SASH1	78781
RRAGD	18	SRC	29166
RAB11A	17	PIK3R1	29166
ADH1B	15	CASP7	23958
SIM2	11	CASP3	23958

C

Node Name	Degree	Node Name	Betweenness
MCM2	132	MCM2	106747
TRIP13	100	COL1A1	85656
CENPA	100	TRIP13	63350
RFC4	69	CENPA	60400
COL1A1	65	RFC4	48922
NEK2	54	NEK2	32905
SERPINH1	36	SERPINH1	22401
UBE2C	36	FOXM1	19404
FOXM1	35	RUVBL2	11564
HOMER3	30	LAMC2	7460

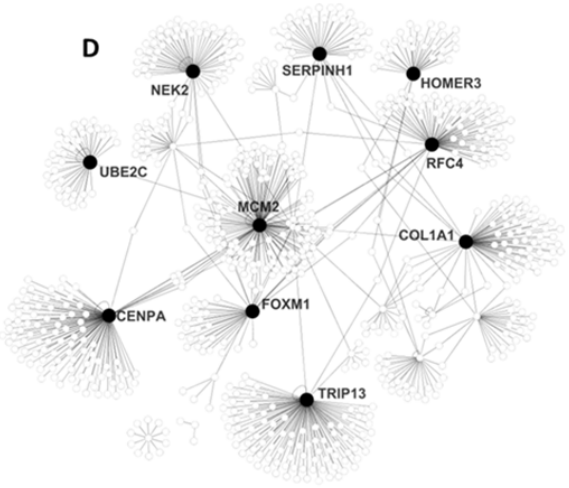


Figure 3.23 Reconstructed protein interaction networks around core DEGs (hubs are emphasized). (A) PPI network of down-regulated genes, (B) Topological properties of down-regulated hubs, (C) Topological properties of up-regulated hubs, (D) PPI network of up-regulated genes.

Down-regulated hub proteins were clustered into 4 groups: i) enzyme (ADH1B), ii) calcium-sodium associated proteins (AHNAK, CAST, NUCB2 and SCNNB1), iii) proteins in signaling pathways (SASH1, RRAGD), and iv) proteins in nervous system (ABLIM1, AHNAK, CAST, NUCB2, SIM2, RAB11A). On the other hand, up-regulated hubs were classified into 4 groups: i) enzymes (MCM2, NEK2 and UBE2C), ii) collagen related proteins (COL1A1 and SERPINH1), iii) proteins in cell cycle and division (CENPA, FOXM1, MCM2, NEK2, RFC4, TRIP13), and iv) signaling protein (HOMER3) (Table 3.12).

Most cellular components interact with each other to accomplish biological functions and biological processes were carried out by interacting proteins and metabolites in cells and tissues. Here, we presented a network reconstruction and analysis approach to

uncover the ESCC driven pathways and biomarkers. In this context, biological networks were represented by a graph theoretical modeling, and topological properties of graphs, such as degree and betweenness connectivity were explored to identify hub proteins.

Table 3.12 Hub proteins of the reconstructed PPI networks in ESCC.

Protein Symbol	Molecular Function	Direction of Regulation	Novelty	References
ADH1B	Enzyme	Down	Known	Wu et al., 2013
ICAL	Proteins in nervous system	Down	Known	Liu et al., 2006
ABLM1	Proteins in nervous system	Down	Known	Ling et al., 2010
NUCB2	Proteins in nervous system	Down	Known	Sato et al., 2006
SCNNB	Calcium-sodium associated protein	Down	Known	Nakamura et al., 2011
SIM2	Proteins in nervous system	Down	Known	Uchikado et al., 2006
RB11A	Proteins in nervous system	Down	Known	Kazemi-Noureini et al., 2004
SASH1	Signaling protein	Down	Known	Su et al., 2011
RRAGD	Signaling protein	Down	Known	Nakamura et al., 2011
AHNK	Calcium-sodium associated protein	Down	Known	Fan et al., 2012
SERPH	Collagen related protein	Up	Known	Kwon et al., 2009
CENPA	Proteins in cell cycle and division	Up	Known	Tao et al., 2012
CO1A1	Collagen related protein	Up	Known	Wong et al., 2012
FOXM1	Proteins in cell cycle and division	Up	Known	Hui et al., 2012
LAMC2	Heparin binding	Up	Known	Hansen et al., 2008
MCM2	Enzyme	Up	Known	Kato et al., 2003
NEK2	Proteins in cell cycle and division	Up	Known	Li et al., 2014
RFC4	Proteins in cell cycle and division	Up	Novel	This study
PCH2	Proteins in cell cycle and division	Up	Known	Zhang et al., 2010
HOME3	Signaling protein	Up	Novel	This study
UBE2C	Enzyme	Up	Known	Kashyap et al., 2009

Remarkably, 6 of 10 down-regulated hub proteins indicated the importance of nervous system in ESCC. The nervous system has a key role in cell migration so that supports metastasis at least two ways. Firstly, cell migration is known to be induced by neurotransmitters that also regulate some part of metastasic multi-step process. Secondly, nerve fibers can be used by tumor cells as routes for invasion and emigration from the primary tumors (Voss et al., 2010). Additionally, neuropeptides were shown to be involved in regulation of tumor growth and differentiation that is why neural factors could affect the tumorgenesis and tumor development (Tallet et al., 1996; Lu et al., 2003). Lu and coworkers (2003) suggested that esophageal and cardiac carcinomas were innervated by peptidergic nerve fibers, which are important for the extention of neuronal processes. A few studies focused on nervous system related genes such as CD171 (Rawnaq et al., 2009) and NGF (Tsunoda et al., 2006) in ESCC and these were proposed as potential therapeutic targets. However, the connection between ESCC and nervous system has not been clarified yet. Another significant aspect was the calcium related 4 of 10 hub proteins. Calcium regulates the switch between proliferation and differentiation (Dotto, 1999). Intracellular calcium concentration is transduced by the 'activation' of Ca^{2+}-binding proteins and calcium- and phospholipid-binding proteins (Kerkhoff et al., 1999). Luo and coworkers (2004) found that expression levels of several calcium associated genes were reduced in ESCC. Moreover, they proposed that calcium was involved in esophageal epithelium proliferation and differentiation and the progress of ESCC but there is limited information about calcium mechanisms in ESCC in literature.

ABLIM1 (Actin Binding LIM Protein 1) was reported to be localized in a genomic region often deleted in human cancers and was proposed to play a role in axon guidance (Krupp et al., 2006). It was also suggested to provide connection between the actin-based cytoskeleton and signaling pathways (Roof et al., 1997). ABLIM1 was reported to be a candidate tumor suppressor gene (Kim et al., 1997) and was regulated in ESCC (Ling et al., 2010). *ADH1B (Alcohol Dehydrogenase 1B, Beta Polypeptide)* is the key enzyme that metabolizes ethanol consumed in alcohol beverages to acetaldehyde, and then to acetate (Yokoyama et al., 2005). ADH1B plays a significant role in determining the the risk of developing ESCC (Ye et al., 2014; Wu et al., 2013; Zhang et al., 2010). *AHNAK (Desmoyokin)* has been recently linked to several biological processes

including calcium regulation, actin cytoskeleton, cellular migration and invasion (Haase et al., 1999; Gentil et al., 2001). Moreover, it was investigated in head and neck cancer (Dumitru et al., 2013), breast cancer (Leong et al., 2012) and acute myeloid leukemia (Mougeot et al., 2011). AHNAK was also reported to be dysregulated in rat esophagus (Seguin et al., 2011) and has been implicated in human ESCC (Fan et al., 2012). Further investigation is needed to study the role of AHNAK in ESCC. *CAST (Calpastatin)* is an endogenous inhibitor of calpain (calcium-dependent cysteine protease) (Hanna et al., 2008). Its depletion induced calpain dysregulation and resulted in cytoskeleton disruption and neurodegeneration (Rao et al., 2008). Calpain/calpastatin expression levels are effective markers for predicting the sensitivity of human esophageal carcinoma cells to treatment (Liu et al., 2006). *NUCB2 (Nucleobindin 2)* is a calcium binding protein involved in signaling adaptive responses and controls the body functions associated with energy reserves. It also behaves as an integral regulator of energy homeostasis and closely related neuroendocrine functions (Garcia-Galiano et al., 2010). NUCB2 interacts with a G protein-coupled receptor, and regulates calcium inflow through the activation of protein kinase A in neurons and is altered in many diseases including epilepsy, diabetes, and anxiety (Aydin, 2013). NUCB2 was reported to be down-regulated in ESCC (Sato et al., 2006; Uchikado et al., 2006). *RAB11A (Ras-Related Protein Rab-11A)* plays part in secretory pathways, vesicle trafficking and apical recycling system in epithelial cells (Rodman and Wandinger-Ness, 2000). RAB11A also facilitates cell migration by internalizing integrins at the rear of cell and transporting them forward at the leading edge to form new contacts with the extra cellular matrix (Kamei et al., 1999). It transported molecules in neurons, where it governed the traffic of AMPA receptors in dendrites during long-term potentiation (Park et al., 2004; Brown et al., 2007; Correia et al., 2008) and axonal trafficking of Trk receptors (Ascano et al., 2009), recycled endosomes lead to the trafficking of membrane proteins into axons (Kamiguchi et al., 2003; Winckler et al., 2004). The alterations of RAB11A expression were reported in ESCC (Kazemi-Noureini et al., 2004). *RRAGD (Ras-Related GTP Binding D*) plays significant role in the activation of the TOR signaling cascade by amino acids (Sancak et al., 2010). RRAGD was also defined as down-regulated genes (Nakamura et al., 2011). *SASH1 (SAM and SH3 Domain Containing 1)* was suggested to have a role in signaling pathways (Zeller et al., 2003).

It was reported as tumor suppressor gene in breast and colon cancer (Zeller et al., 2003; Rimkus et al., 2006). Su and coworkers have reported SASH1 in the list of down-regulated genes in ESCC (Su et al., 2011) but did not characterize its role. *SCNN1B (Sodium Channel, Non Voltage Gated 1 Beta Subunit)* is one of three homologous subunits forming the amiloride-sensitive epithelial sodium channel (Voilley et al., 1995). SCNN1B coding CACNA1G and CALCA1I were shown to be found in human tumors (Jacinto et al., 2007). SCNN1B was proposed to have role in sodium, calcium, potassium, and chloride channels that regulated cell proliferation, migration, and invasion (Jacinto et al., 2007). SCNN1B was included in the list of down-regulated genes in ESCC (Nakamura et al., 2011). *SIM2 (Single-Minded Family BHLH Transcription Factor 2)* encodes transcription factors involved in brain development (Rachidi et al., 2005) and neuronal differentiation (Goshu et al., 2004). It might be playing a role in pathogenesis of solid tumors, and was suggested as a molecular target for cancer therapy (DeYoung et al., 2003). The overexpression of SIM2-s was related with tumors of the colon, pancreas and prostate (Halvorsen et al., 2007). It was only listed in a table of suppressed genes that correlated with lymph node metastasis in ESCC (Uchikado et al., 2006).

Apparently, the importance of cell cycle was pointed out through the up-regulated hub proteins. It was reported that deregulation of cell cycle proteins induced the cell proliferation that lead to cancer. The cell cycle proteins that provided integrated information transduced through upstream signaling networks was suggested to be an alternative target for diagnostic and therapeutic interventions. Analysis of several mitotic proteins in human tissues contributed to identification of novel biomarkers for cancer detection and yielded target validation for cell cycle-directed therapies (Williams and Stoeber, 2012).

CENPA (Centromere Protein A) recruits kinetochore components to centromeres (Palmer et al., 1991) and is situated in the inner kinetochore plate throughout the cell cycle (Guo et al., 2008). It is overexpressed in colorectal cancer cells (Tomonaga et al., 2003) and up-reguated in ESCC (Tao et al., 2012). *COL1A1 (Collagen, Type 1, Alpha 1)* was reported to induce disruption of cell cell contacts and promote proliferation of pancreatic carcinoma cells (Koenig et al., 2006). The up-regulation of collagen genes in primary tumors with metastatic potential is consistent with recent observations that

epithelial-mesenchymal interactions are critical determinants of tumor cell behavior (Olumi et al., 1999). High levels of type 1 collagen have been reported in metastatic lesions as well as in the serum of individuals with metastatic disease (Ramaswamy et al., 2003). COL1A1 was documented as an up-regulated hub in ESCC (Wong et al., 2009; Sakamoto et al., 2010). *FOXM1 (Forkhead Box M1)* is a member of the Forkhead transcription factor family that is involved in metabolism and development, cell proliferation, apoptosis and invasion, longevity, and cancer (Myatt and Lam, 2007). It controls the cell cycle by modulating the transcription of various cell cycle-related genes (The et al., 2002). FOXM1 over expression has been repored in basal cell carcinomas, glioblastomas, non-small-cell lung cancers, and breast cancers. These results proposed a role for FOXM1 in tumor initiation and progression (Wierstra and Alves, 2007). Cytoplasmic FOXM1 was also suggested to be a biomarker for ESCC (Hui et al., 2012). *HOMER3 (Homer Scaffolding Protein 3*) has a key role in the differentiation and development of some tissues (e.g., muscle and nervous system) (Ishiguro et al., 2004; Shiraishi et al., 2004; Bortoloso et al., 2006). The up-regulation of HOMER3 is implicated in in the occurrence and development of acute myeloid leukemia (Li et al., 2013). In this study, HOMER3 was reported for the first time in ESCC. *MCM2 (Minichromosome Maintenance Complex Component 2)* is involved in regulating eukaryotic DNA replication (Bell and Dutta, 2002) and has been investigated in bladder, cervical, and colon cancer. MCM proteins are considered useful biomarkers that reflect the cell cycle state (Huang et al., 2011). MCM2 was shown to be necessary for ESCC development (Tao et al., 2012) and could prove to be a useful proliferative biomarker of esophageal epithelium (Huang et al., 2011). *NEK2 (NIMA-Related Kinase 2)* regulates cell cycle (Fry, 2002). NEK2 protein was up-regulated in several cancers including cervical, ovarian, breast, prostate, and leukemia (Hayward et al., 2004). Yan and coworkers (2012) demonstrated up-regulation of NEK2 in ESCC. *RFC4 (Replication Factor C4)* seems to be overexpressed in cancer cells including the lung, prostate, colon, stomach, and skin (Jung et al., 2009). Upon DNA damage, the checkpoint kinases Chk1 and Chk2 phosphorylate and activate FOXM1, which in turn activated the expression of genes important for homologous recombination, including RFC4 (Zhao et al., 2012). This study reports RFC4 for the first time as being overexpressed in ESCC. *SERPINH1 (Serpin Peptidase Inhibitor)* binds specifically to

collagen, which is a major component of ECM in tissues such as skin and bone and also takes part in the biosynthetic pathway of collagen (Widmer et al., 2012). SERPINH1 was reported to be differentially expressed in ESCC (Wong et al., 2013; Sugimoto et al., 2007) and was proposed as a potential biomarker for squamous cell carcinoma of the head and neck (Schaaij-Visser et al., 2010). *TRIP13 (Thyroid Hormone Receptor Interactor 13)* is a thyroid receptor interacting protein (Lee et al., 1995) and TRIP13 (by its interaction with CMT2) was involved in cancer progression as spindle checkpoint defects were shown to contribute to loss of chromosome integrity, which aneuploidy, a common feature of cancer cells (Bharadwaj and Yu, 2004). Moreover, it was reported as up-regulated genes in ESCC (Tao et al., 2012; Zhang et al., 2010). *UBE2C (Ubiquitin-Conjugating Enzyme E2C)* has a critical role in M-phase cell-cycle progression by inactivating the M-phase check point (Reddy et al., 2007). UBE2C was overexpressed in various types of solid tumors including breast cancer, colon cancer, gastro-esophageal cancer, lung cancer, ovary cancer, thyroid cancer and late-stage prostate cancer (Wagner et al., 2004; Lin et al., 2006; Fujita et al., 2009; Wang et al., 2009). UBE2C was listed as up-regulated gene in ESCC (Sugimoto et al., 2007; Kashyap et al., 2009).

Literature search was performed by curated ESCC genes using three extensive sources: i) Agrawal and coworkers (2012) conducted exomic sequencing on 12 ESCC cases from the United States to explore its genetic origins and identified high confidence mutations on 893 genes in ESCC patients, ii) Lin and co-workers (2009) performed a detailed literature survey, reviewed protein alterations, and presented 221 altered proteins in ESCC, iii) 101 genes were associated with ESCC in GAD, which is a database of genetic association data from complex diseases and disorders. The comparative analysis of ESCC-associated gene lists from these three independent studies demonstrated unequal distribution of genes, i.e., only two genes (TP53 and CDKN2A) were common in all sets. CDKN2A, a known TSG was also reported in our down-regulated core genes list (Figure 3.24).

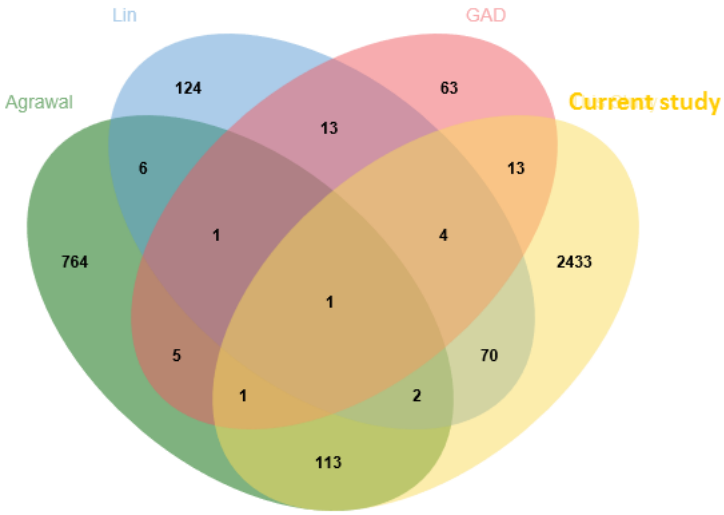


Figure 3.24 The literature-curated ESCC gene pool from three different sources.

Down-regulated hub protein, ADH1B is stated as ESCC related gene in GAD but there is no ESCC specification for other hubs. However, physical interacting partners (1st neighbors) of these hubs were searched for the relation with known ESCC genes. 7 of 10 (ADH1B, AHNAK, CAST, NUCB2, RAB11A, SASH1, and SCNN1B) have interacting partners, which are ESCC associated genes (based on literature search). Furthermore, AHNAK interacts physically with TSGs (CDKN2A), which is known gene associated with ESCC. These results indicated the power of interactome to identify disease specific genes. On the other hand, ABLIM1, RRAGD and SIM2 do not have any known interacting partner related with ESCC in these literature sources.

We also explored the interacting partners of 20 core up-regulated genes. 9 hub proteins (CENPA, COL1A1, FOXM1, HOMER3, MCM2, NEK2, RFC4, SERPINH1, and TRIP13) had interacting partners except UBE2C that were known to be associated with ESCC. Additionally, the interacting partners of COL1A1, FOXM1, MCM2, and NEK2 were TSGs of ESCC. COL1A1 interacts with IGFBP3 and MCM2 interacts with CDKN2A. IGFBP3 and CDKN2A were also in our DEGs list. FOXM1 interacts with CDH1 which was known ESCC related TSG.

3.3.3.3 Functional enrichment analysis

Through disease enrichment analysis, we found associations of ESCC with other cancer forms such as prostate cancer (p-value: 0.021) and colorectal cancer (p-value: 0.040). The gene-disease associations revealed that down-regulated genes were mainly involved in esophageal cancer, prostate cancer, and hepatitis; while up-regulated genes were predominantly associated with abdominal aortic aneurysm, rheumatoid arthritis, brain aneurysm, colorectal cancer, periodontitis, systemic sclerosis, and gastric ulcer.

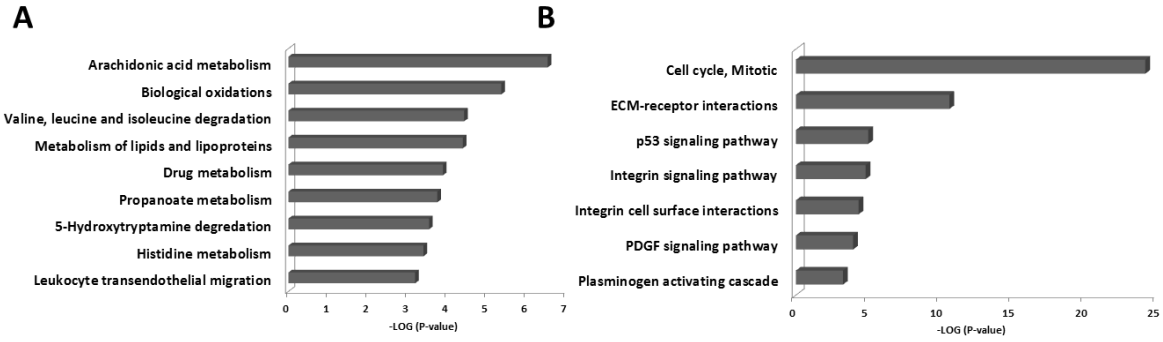


Figure 3.25 Pathway enrichment analysis results for (A) Down-regulated genes and (B) Up-regulated genes.

The pathway enrichment analysis mapped significant pathways via, KEGG, REACTOME and PANTHER databases. We applied a rigid p-value cutoff of 10^{-4} in order to focus only on the highest quality predictions. The down-regulated genes were associated with diverse metabolic pathways including arachidonic acid metabolism, valine, leucine and isoleucine degradation, histidine metabolism, drug metabolism, propanoate metabolism, leukocyte transendothelial migration, biological oxidations and metabolism of lipids and lipoproteins (Figure 3.25A). On the other hand, up-regulated genes highlighted cell cycle and several signaling pathways such as p53 signaling, integrin signaling, platelet-derived growth factor (PDGF) signaling and plasminogen activating cascade (Figure 3.25B).

GO process enrichment analysis indicated significant down-regulation of epidermis and ectoderm development, oxidation-reduction, lipid metabolism (lipid metabolic process, monocarboxylic acid metabolic process, icosanoid metabolic process, organic acid, carboxylic acid, oxoacid metabolic processes) and protein-DNA complex assembly

process under ESCC (Figure 3.26). The GO biological process enrichment analysis revealed up-regulation of cell cycle, cell proliferation, collagen metabolic process, and extracellular matrix organization.

The pathway enrichment analyses highlighted the molecular mechanisms of ESCC that the down-regulated core genes were associated with diverse metabolic pathways, while up-regulated core genes involved in cell cycle and several signaling pathways.

The most significant down-regulated pathway in this study involves arachidonic acid (AA) metabolism (p-value < 10^{-6}), which is related to many different biologically active lipids playing role in apoptosis and cell proliferation. The down-regulation of AA metabolic pathway in esophageal cancers was also reported by previous studies (Zhi et al., 2003). Moreover, Zhi and coworkers (2003) proposed that better diagnostic and treatment approaches could be developed by further investigation of AA metabolism involved in ESCC. Other down-regulated vital pathways included 5-Hydroxytryptamine degradation, biological oxidations and metabolism of lipids and lipoproteins. There is not much known about 5-Hydroxytryptamine degradation in ESCC, however, Wang and coworkers demonstrated that low-density lipoproteins, high-density lipoproteins and total cholesterol were reduced in ESCC patients. They also found decreased levels of antibodies against biologically oxidized low-density lipoproteins in ESCC patients (Wang et al., 2008).

The up-regulated pathways: i) *cell cycle:* the majority of cell cycle regulators were either disabled or activated in ESCC thus they probably played significant roles in this disease (Smeds et al., 2002; Lin et al., 2009), ii) *ECM-receptor signaling:* the migration of tumor cells is related with the degradation of the extracellular matrix (ECM) (Gu et al., 2005). Transformation of ECM is an essential factor for various physiological and pathological processes, including tissue formation and development, cell proliferation and differentiation, as well as invasion and metastasis of neoplastic cells. ECM remodeling is also a regulator of stromal-epithelial interactions during carcinogenesis (Groblewska et al., 2012). Several differentially expressed ECM genes, such as MMP13, MMP11 and EMP1 which were also listed as DEGs in the present study, are reported to be involved in ESCC (Kashyap et al., 2009), iii) *p53 signaling pathway:* deactivation of the p53 pathway was caused by the p53 mutation itself, overexpression

of the MDM2 protein or p14ARF hypermethylation, resulting in disabling of esophageal cell cycle arrest and apoptosis (Lin et al., 2009), iv) *Integrin signaling pathway:* important cellular processes critical to tumor development and progression, including cell proliferation, cell survival, angiogenesis, cell motility, and invasiveness that was known to play a critical role in driving ESCC were influenced (Tong et al., 2012), v) *PDGF signaling pathway:* most significant cancer associated pathway and targeted by tyrosine kinase inhibitors in some cancers, but not much explored in ESCC. For the first time, Yan and coworkers (2012) identified PDGF signaling pathway as therapeutic target in ESCC, vi) *plasminogen activating cascade:* the urokinase plasminogen activator (uPA) system (uPAS) consists of the uPA, its cognate receptor (uPAR) and two specific inhibitors, the plasminogen activator inhibitor 1 (PAI-1) and 2 (PAI-2). The uPA converts the proenzyme - plasminogen in the serine protease plasmin, involved in a number of physiopathological processes requiring basement membrane and/or ECM remodelling, including tumor progression and metastasis (Ulisse et al., 2009). PAI-1 was reported as a new prognostic marker for ESCC, because significant increases in expression of PAI-1 were observed in advanced tumor stages and metastasis-positive ESCCs, compared with metastasis-negative ones (Sakakibara et al., 2006).

Biological process enrichment analyses indicated significant down-regulation of epidermis and ectoderm development, oxidation-reduction, lipid metabolism (lipid metabolic process, monocarboxylic acid metabolic process, icosanoid metabolic process, organic acid, carboxylic acid, oxoacid metabolic processes), keratinocyte differentiation, and vitamin metabolic process in ESCC (Figure 28). Gao and coworkers (2012) combined meta-analysis and graph clustering to find markers of ESCC and identified gene clusters, which were involved in epidermis and ectoderm development. Moreover, Essack and coworkers (2009) formulated a gene list involving epidermis development and unique to esophageal cancer. It can be assumed that lipid and its constituents (oxLDL-ab, oxLDL-lgG and oxLDL-lgM) may also be associated with different stages in esophageal carcinogenesis proceeding (Diao et al., 2009). Vitamins play a significant role as coenzymes or enzymes in many processes for normal functioning of the body. Currently, it is known that vitamins can have an important role in the prevention and treatment of cancer (Mamede et al., 2011). Factors that may

protect against the development of esophageal cancer include foods containing dietary fiber, folate, pyridoxine, carotenoids, and vitamins E and C. (Vermeulen et al., 2013). Moreover, low vitamin E and selenium nutritional status is known to be associated with increased esophageal carcinogenesis in rat (Yang et al., 2011).

Enrichment analysis figured out up-regulation of cell cycle, cell proliferation, collagen metabolic process, immune system process, extracellular matrix organization and neurogenesis (Figure 3.26). In ESCC, the majority of cell cycle regulators, which probably played important roles in this disease, were either activated or disabled (Smeds et al., 2002 and Lin et al., 2009). The migration of tumor cells is associated with the degradation of the extracellular matrix (ECM). The ECM is a dynamic structure composed of various macromolecules such as collagen, fibronectin, laminin and proteoglycans of connective tissue.

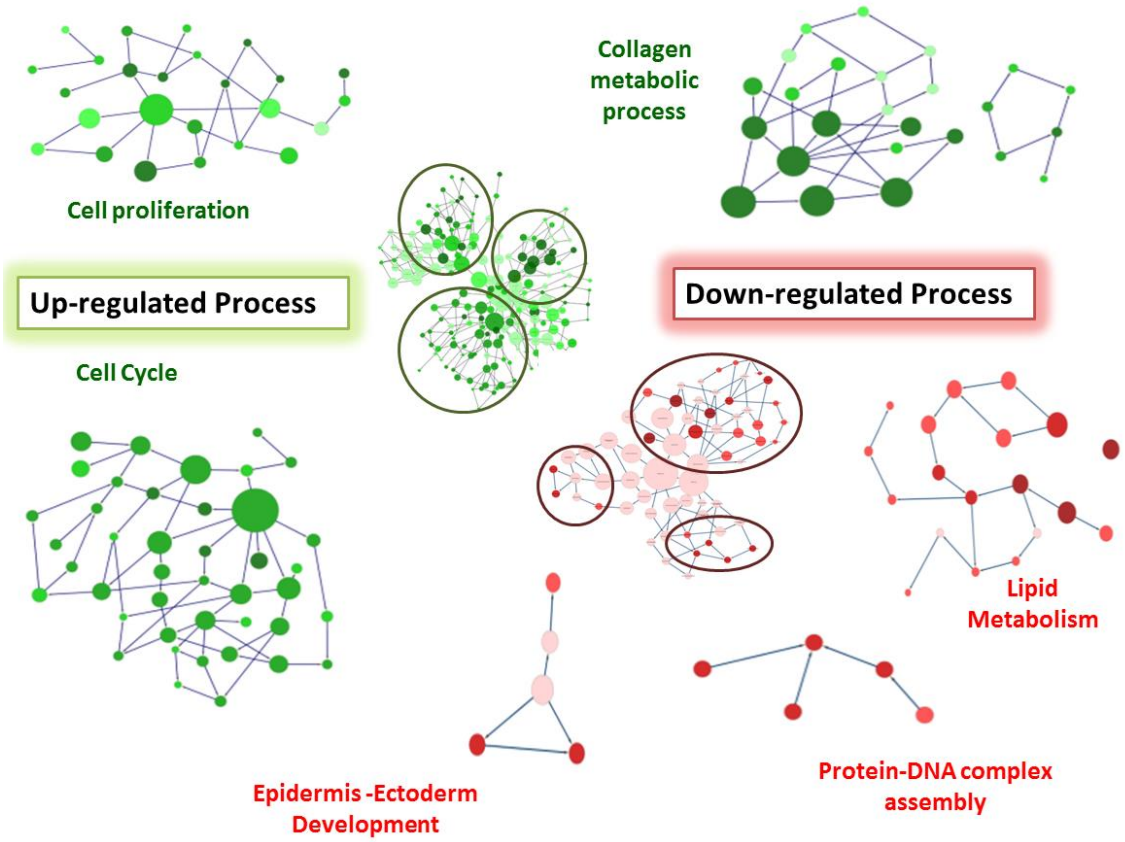


Figure 3.26 GO biological process enrichment analysis of down- and up-regulated genes. Green: Up-regulated genes associated processes and Red: down-regulated genes associated processes.

Diseases in general are not consistently considered independent of one another. Indeed, different disease modules can overlap, so that perturbations caused by one disease can affect other disease modules. Uncovering such links between diseases not only helps us understand how different phenotypes, often addressed by different medical sub-disciplines, are linked at the molecular level, but can also help us to comprehend why certain groups of diseases arise together. The "diseasome" offers insights that may yield new approaches to diagnosis, disease prevention, and treatment. Diseasome-based approaches could also aid drug discovery, in particular when it comes to the use of approved drugs to treat molecularly linked diseases. Here, we determined through analyzing DEGs that in ESCC and other related diseases that down-regulated core genes were mainly involved in esophageal cancer, prostate cancer, and hepatitis; while up-regulated core genes were predominantly found in abdominal aortic aneurysm,

rheumatoid arthritis, brain aneurysm, colorectal cancer, periodontitis, systemic sclerosis, and gastric ulcer.

The association of ESCC and abdominal aortic aneurysm through significant enzymes, matrix metalloproteinases (MMPs) and their tissue inhibitors (TIMPs), that abnormal matrix turnover were reported to play important role in the pathogenesis of abdominal aortic aneurysm (Morris et al., 2014). It was also reported that they played also a key role in esophageal carcinogenesis (Groblewska et al., 2012). Moreover, the main roles of MMPs in pathology included tissue destruction that was observed in cancer metastasis and rheumatoid arthritis. It was proved that these enzymes were involved in the development of many malignant neoplasms: gastric, colorectal and breast cancer (Groblewska et al., 2012). On the other hand, scleroderma is skin thickening and hardening associated with many different medical conditions, and it is called systemic sclerosis when other organs are also involved. In the esophagus, it predominantly causes distal esophageal hypomotility and weak lower esophageal sphincter tone. Clinically, it can present as heartburn and dysphagia, and it is associated with ESCC. A more recent review of seven studies showed that the link between scleroderma and cancer was not overwhelming with probably a modest increase in lung cancer. This cancer risk might be much lower among localized scleroderma patients (Kenneth et al., 2012).

3.3.3.4 Transcriptional regulatory subnetwork

A transcriptional regulatory network consisting of 103 regulatory relationships between 18 TFs and their targeted 28 down-regulated core genes was reconstructed (Figure 3.27A). Enrichment analysis of the reconstructed network indicated significant enrichment for estrogen mediated S phase entry, aryl hydrocarbon receptor signaling, nicotine degradation, bupropion degradation and telomerase signaling pathways (Figure 3.27B).

The reconstructed transcriptional regulatory network around up-regulated core DEGs was consisting of 82 regulatory relationships between 25 TFs and their 20 up-regulated core target genes (Figure 3.27D). Functional enrichment analysis indicated estrogen mediated S phase entry, aryl hydrocarbon receptor signaling, chronic myeloid leukemia signaling, telomerase signaling and atherosclerosis signaling pathways as significant biological processes (Figure 3.27C).

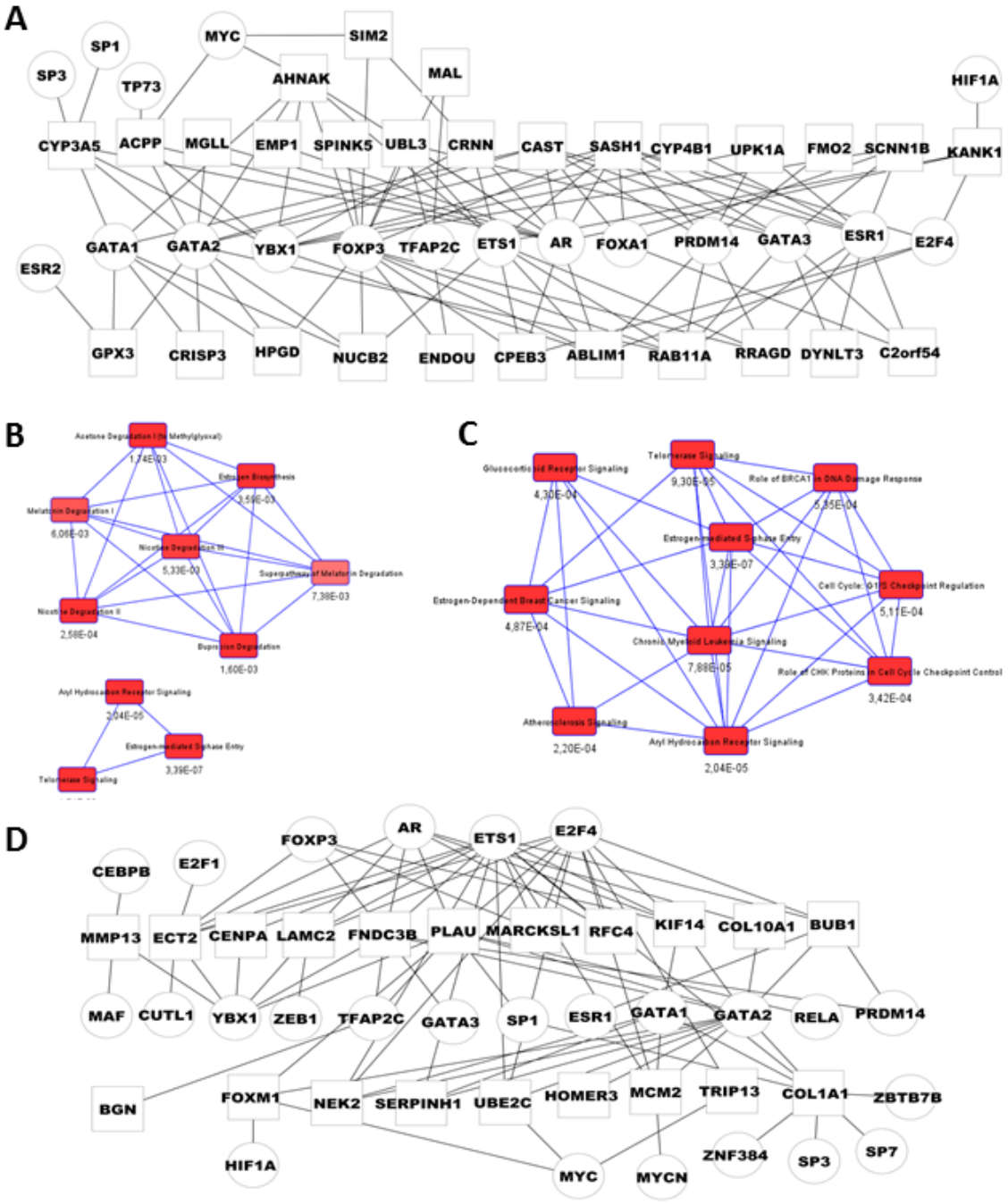


Figure 3.27 Transcriptional regulatory networks and associated pathways. A) Network around down-regulated core genes. B) Associated pathways of down-regulated network. C) Associated pathways of up-regulated network. D) Network around up-regulated core genes. (Circle: TF, square: gene).

TFs have roles as regulatory endpoints in signaling pathways, and their altered activities lead to many human diseases, including cancer. To uncover novel TFs that would govern clinical outcomes in ESCC, we analyzed core sets of DEGs. TFs serve as oncogenes or tumor suppressors in various human cancers. The normal balance between tumor-suppressive and oncogenic gene networks is altered by dysregulation of gene expression in cancer cells so that changing the differentiation state of cells leads to proliferation, apoptosis, and metastasis. For that reason, a very effective therapeutic strategy can be planned through the activation of tumor-suppressive TFs and/or the inhibition of oncogenic TFs for cancer treatment (Grivas et al., 2011; Park et al., 2012).

AR was the only TF that was shown to be down-regulated in this study. AR influenced 13 core down-regulated genes out of which ABLIM1, AHNAK, CAST, SIM2, SASH1 were hubs (Figure 3.27A). Down-regulated TFs can be drug target since the target genes of these TFs were also candidate tumor suppressors of ESCC.

Additionally, transcription factors, AR, CEBPB and E2F1 were identified to be up-regulated in this study (Figure 3.27D). AR targets 7 core up-regulated genes among which RFC4 was reported as a hub. The inhibition of up-regulated TFs as drug targets can have significant impact on these up-regulated oncogenic core genes (Figure 3.27C-D). CEBPB is involved in TNF signaling pathway (Nanes, 2003), which can promote the process of tumor metastasis (Mocellin et al., 2005). Moreover, E2F1 and AR promoted cell cycle and cell proliferation, respectively thus playing a role in prostate cancer pathway (Lee et al., 2009; Heinlein and Chang, 2004). Although the relation between prostate cancer and ESCC was investigated in particular for role of maspin (Cai et al., 2009) and N-cahderin (Li et al., 2010) the prostate cancer and ESCC connection remains unclear.

3.3.3.5 Reporter metabolic features

Reporter metabolite analyses revealed three metabolites, formate, proton (H+) and 10-Hydroxyeicosatetraenoate (10-HETE), as common reporter in all five datasets. In addition, 35 metabolites were common in four datasets (excluding GSE17351) (Figure 3.28A). Among the 35 common reporter metabolites, 33 of them were lipids, more specifically eicosanoids (23 metabolites), streoids (9 metabolites) and a retinoid. In addition, 13 reporter metabolites were playing roles in arachidonic acid (AA)

metabolism (Figure 3.28B), 8 reporter metabolites were involved in steroid hormone biosynthesis (Figure 3.28C), and 7 reporter metabolites were having role in ovarian steroidogenesis (Table 3.13).

Table 3.13 Reporter metabolites of ESCC.

Arachidonic acid metabolism	**KEGG ID**	**Molecular Class**
11,14,15-Trihydroxyicosatrienoic acid (11,14,15-THETA)	C14814	LIPIDS
11-Hydroxy-14,15-epoxyeicosatrienoic acid (11h-14,15-EETA)	C14813	LIPIDS
11,12,15-Trihydroxyicosatrienoic acid (11,12,15-THETA)	C14782	LIPIDS
15-Hydroxy-11,12-epoxyeicosatrienoic acid (15H-11,12-EETA)	C14781	LIPIDS
11-Hydroxyeicosatetraenoate (11-HETE)	C14780	LIPIDS
16R-Hydroxyeicosatetraenoate (16(R)-HETE)	C14778	LIPIDS
8-Hydroxyeicosatetraenoate (8-HETE)	C14776	LIPIDS
14,15 Epoxyeicosatrienoic acid (14,15-EET)	C14771	LIPIDS
11,12 Epoxyeicosatrienoic acid (11,12-EET)	C14770	LIPIDS
8,9 Epoxyeicosatrienoic acid (8,9-EET)	C14769	LIPIDS
5,6 Epoxyeicosatrienoic acid (5,6-EET)	C14768	LIPIDS
19(S)-Hydroxyeicosatetraenoate (19(S)-HETE)	C14749	LIPIDS
15(S)-Hydroperoxyeicosatetraenoate (15(S)-HPETE)	C05966	LIPIDS
Steroid hormone biosynthesis	**KEGG ID**	**Molecular Class**
2-hydroxyestradiol-17beta	C05301	LIPIDS
estradiol-17beta	C00951	LIPIDS
16alpha-hydroxyestrone	C05300	LIPIDS
19-oxoandrostenedione	C05297	LIPIDS
19-oxo-testosterone	C05295	LIPIDS
19-hydroxytestosterone	C05294	LIPIDS
19-hydroxyandrostenedione	C05290	LIPIDS
4-androstene-3,17-dione	C00280	LIPIDS
Ovarian steroidogenesis	**KEGG ID**	**Molecular Class**
2-hydroxyestradiol-17beta	C05301	LIPIDS
estradiol-17beta	C00951	LIPIDS
4-androstene-3,17-dione	C00280	LIPIDS
11,12 Epoxyeicosatrienoic acid (11,12-EET)	C14770	LIPIDS
8,9 Epoxyeicosatrienoic acid (8,9-EET)	C14769	LIPIDS
5,6 Epoxyeicosatrienoic acid (5,6-EET)	C14768	LIPIDS
15(S)-Hydroperoxyeicosatetraenoate (15(S)-HPETE)	C05966	LIPIDS

Table 3.13 Continued.

Serotonergic synapse	**KEGG ID**	**Molecular Class**
14,15 Epoxyeicosatrienoic acid (14,15-EET)	C14771	LIPIDS
11,12 Epoxyeicosatrienoic acid (11,12-EET)	C14770	LIPIDS
8,9 Epoxyeicosatrienoic acid (8,9-EET)	C14769	LIPIDS
5,6 Epoxyeicosatrienoic acid (5,6-EET)	C14768	LIPIDS
15(S)-Hydroperoxyeicosatetraenoate (15(S)-HPETE)	C05966	LIPIDS
Vascular smooth muscle contraction	**KEGG ID**	**Molecular Class**
14,15 Epoxyeicosatrienoic acid (14,15-EET)	C14771	LIPIDS
11,12 Epoxyeicosatrienoic acid (11,12-EET)	C14770	LIPIDS
8,9 Epoxyeicosatrienoic acid (8,9-EET)	C14769	LIPIDS
5,6 Epoxyeicosatrienoic acid (5,6-EET)	C14768	LIPIDS
Inflammatory mediator regulation of TRP channels	**KEGG ID**	**Molecular Class**
8,9 Epoxyeicosatrienoic acid (8,9-EET)	C14769	LIPIDS
5,6 Epoxyeicosatrienoic acid (5,6-EET)	C14768	LIPIDS
15(S)-Hydroperoxyeicosatetraenoate (15(S)-HPETE)	C05966	LIPIDS
Linoleic acid metabolism	**KEGG ID**	**Molecular Class**
12,13-epoxyoctadecenoate (12(13)-EpOME)	C14826	LIPIDS
9,10-epoxyoctadecenoate (9(10)-EpOME)	C14825	LIPIDS
Prolactin signaling pathway	**KEGG ID**	**Molecular Class**
estradiol-17beta	C00951	LIPIDS
4-androstene-3,17-dione	C00280	LIPIDS
Prostate cancer	**KEGG ID**	**Molecular Class**
estradiol-17beta	C00951	LIPIDS
4-androstene-3,17-dione	C00280	LIPIDS
Riboflavin metabolism	**KEGG ID**	**Molecular Class**
Adenine-riboflavine dinucleotide (FAD)	C00016	Nucleosides, Nucleotides, and Analogue
Dihydroflavine-adenine dinucleotide (FADH2)	C01352	Nucleosides, Nucleotides, and Analogues

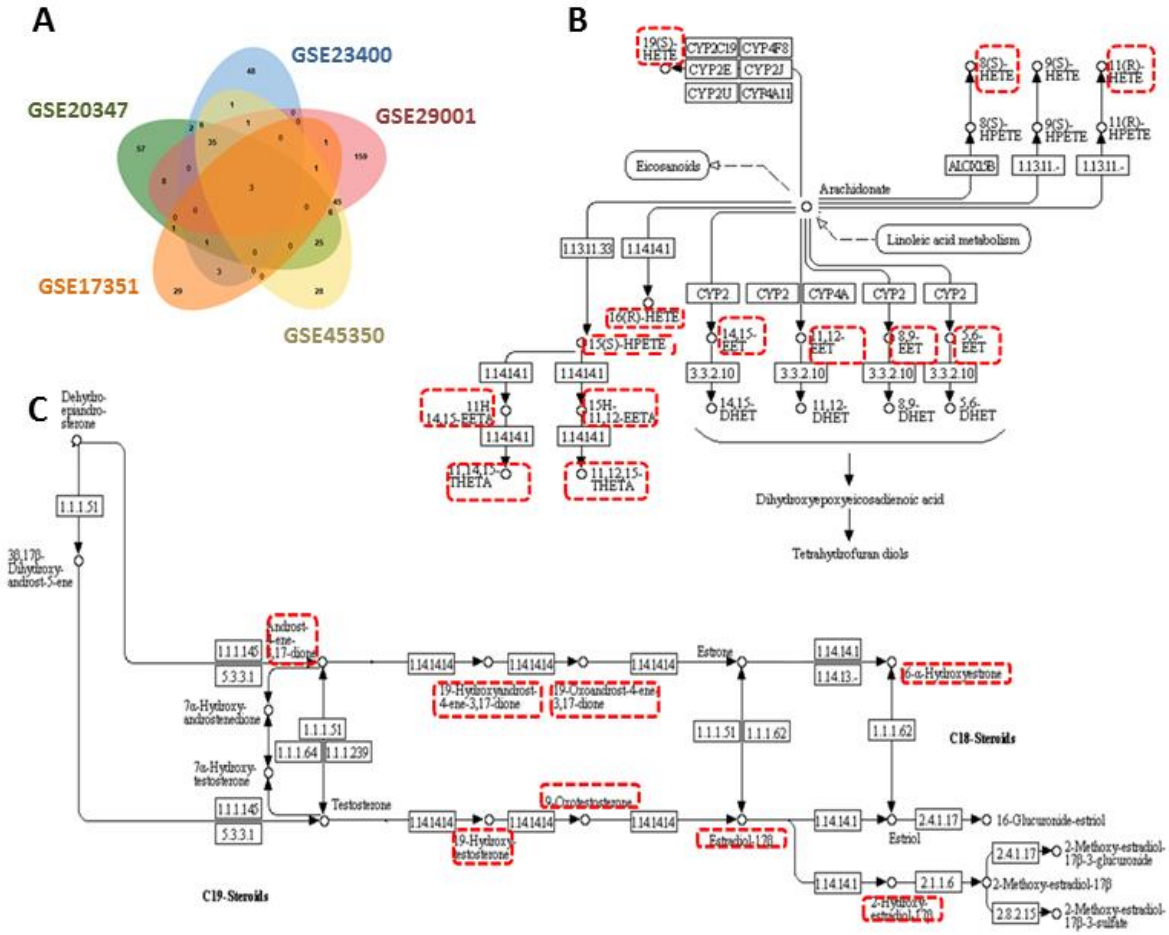


Figure 3.28 (A) The number of reporter metabolites in datasets analysed. (B) Arachidonic acid metabolic pathway. (C) Steroid hormone biosynthesis. Red dash-lined rectangles indicate the reporter metabolites included in our list.

One of three common metabolites among five datasets was formate and its metabolism was shown at the center of many activities that were controlling tumor growth. Cancer tissue was known to maintain a very low redox state which is maintained with increased NADPH production. This NADPH production was in part fueled by formate dehydrogenation (Jellinek, 2008). Moreover, reporter metabolite analysis showed that 10 of 38 metabolites were HETEs and 10-HETE was revealed in all five datasets. It was reported that HETEs lead to increase in cell proliferation and tumor angiogenesis via activating numerous signaling pathways involved in cell growth and by stimulating the transcription of growth factors (Moreno, 2009). Wu and coworkers (2003) displayed that some HETEs and related metabolites such as hydroxyoctadecadienoic acids

functioned as a tumor-suppressor although other studies suggest that the up-regulated HETE synthesis was related to enhanced tumor growth and progression (Gupta et al., 2001; Chen et al., 2002). For instance, elevated levels of certain eicosanoids such as 5-HETE and 12-HETE together with PGE_2 and LTB_4 might promote tumor progression, whereas others such as 15-HETE might interfere with the progression to malignancy (Shureiqi and Lippman, 2008; Moreno, 2009). Thus, it is important to identify which receptors, transcription factors and signaling pathways are playing a role in HETE synthesis, to design new therapeutic strategies for modulating cell proliferation and cancer progression (Moreno, 2009).

Progression and metastasis are defined as multi-step as well as multi-factorial processes that take many years to achieve a clinically apparent phenotype. To identify critical factors for the malignant progression of cancers it is important to develop a better treatment or secondary prevention but, this is a big challenge. Among many factors underpinning tumor progression, cyclooxygenase (COX), lipoxygenase (LOX) and P450 epoxygenase pathways take key roles in cancer-related inflammation and progression. Arachidonic acid (AA) can be converted to bioactive eicosanoids through these pathways in response to different growth factors, hormones or cytokines. The reporter metabolite analysis highlighted the importance of eicosanoids hence AA metabolism in ESCC. The specified eicosanoids can function as the second messenger or through their cognate receptors evoke various cellular responses (Wang et al., 2007). Among them, it is well known that cytochrome P450-derived eicosanoids take parts in inflammation and angiogenesis rather than cancer pathways (Panigrahy et al., 2010). However, the important function of CYP enzymes is the detoxification of compounds i.e., anti-cancer drugs so the efficacy of cancer drug delivery is improved by blocking these enzymes (Nebert and Dalton, 2006). Moreover, cytochrome P450 epoxygenases was reported as a potential tumor promoting enzyme and its metabolites epoxyeicosatrienoic acids (EETs) were suggested to be involved in cancer metastasis (Jiang et al., 2007). On the other hand, activated cytochrome P450 enzymes by prodrugs are applied to inhibit tumor growth by targeting the tumor cells (Swanson et al., 2010). EETs take significant roles in cellular proliferation, migration and inflammation that their major target is blood vessels (Spector and Norris, 2007; Fleming, 2007). It was shown that EET antagonists inhibited prostate cancer cell motility and suggested a

novel mechanism of EET antagonists acting directly on tumor cells (Nithipatikom et al., 2010). Another study found the linkage between eicosanoids to stroma inflammation in epithelial ovarian cancer (Freedman et al., 2007).

Zhi and colleagues (2003) identified a member of the AA metabolizing enzymes (COX-2) and many genes activated as part of the AA metabolism pathway were regulated in ESCC. Therefore, they suggested that AA metabolism pathway through both the COX and LOX metabolism pathways is modulated during tumorigenesis of ESCC. Furthermore, AA is converted to leukotrienes by LOX that Xu and coworkers reported the loss of 15-LOX-2 in esophageal cancers and the induction of 15-LOX-2 inhibited cancer cell proliferation (Xu et al., 2003). The results another study showed that GRP was up-regulated in ESCC and GRP was involved at early stage of ESCC and may be responsible for promoting AA metabolism (Fang et al., 2004). Kashyap and colleagues (2009) identified down-regulation of several genes involved in AA metabolism in ESCC that were not reported earlier by other studies. The genes encoding COX-2 and 12-LOX were figured out as genetically polymorphic that influenced in the expression of these enzymes. Thus, these results demonstrated that aberrant AA metabolism took a place in the development of esophageal cancer (Guo et al., 2006). Consequently, our results of reporter metabolite analysis suggest that intervention through AA metabolic pathway might be useful for arresting cancer progression. Further investigations of AA metabolism in ESCC could be essential for the improvement of better treatment methods and diagnostic approaches.

In the present study, the reporter feature analysis also indicated 8 steroids that are involved in steroid hormone biosynthesis. They could be categorized in two groups: estrogens and androgens. It was reported that hormone-related cancers such as breast, endometrium, ovary, prostate, testis, thyroid and osteosarcoma, had a unique mechanism of carcinogenesis in common since endogenous and exogenous hormones caused cell proliferation (Henderson and Feigelson, 2000).

Paliakov and coworkers (2014) analyzed various proteins including metabolic enzymes across different cancers and presented a list of genes that were up-regulated in various types of cancer. They further delineated major molecular pathways that were abundant in cancer protein machinery. Glycolysis/gluconeogenesis and oxidative phosphorylation

were emphasized along with steroid hormone synthesis as the most abundant pathways. Meanwhile, the impact of estrogens and androgens in esophageal cancer development and treatment should be defined to uncover the alterations in ESCC metabolism. The metabolic feature results also indicated two significant core genes, CYP3A5 and CYP4B1. These were encoded for aryl hydrocarbon hydroxylase which was involved in both AA metabolism and steroid hormone biosynthesis. Lastly, another outcome of the reporter analysis pinpointed the steroid hormone biosynthesis that provides insight into highlighted transcriptional factor AR in this study.

4 CONCLUSION

Cancer is a multifactorial disease that emerges from dysregulation of normal biological networks. Uncovering the underlying protein signaling network changes aids in understanding the molecular mechanism of carcinogenesis and identifies the characteristic network signatures specific for different cancers and distinct cancer subtypes. The identified signatures can be used for cancer diagnosis, prognosis, and personalized treatment.

Comprehensive identification and characterization of PPIs and their networks is essential in order to systematically define and understand the information transfer within and between biological processes. We realized that mapping techniques of PPI are rapidly increasing and generating an extensive amount of noisy PPI data. While there are various databases that collect and store this data their coverage and quality is debatable and their overlap is interestingly low. A more comprehensive and reliable global human network is essential for progress in the areas such as medicine, pharmacology and systems biology.

Initially, in presented yeast study, a scoring scheme considering eight heterogeneous biological features was proposed and a high-confidence protein–protein interactome of yeast *S. cerevisiae* was reconstructed. The proposed methodology is expected to increase the efficiency of the methodologies in biological research which make use of protein interaction networks in several aspects that the proposed scoring scheme then was adapted to the interactome of *Homo sapiens* and a highly reliable human interactome is achieved by compiling various datasets and ranking PPIs using the proposed scoring scheme. Comparison of the present scoring system to other methods

revealed that the current scheme provided greater coverage as well as higher reliability. Cancer and process subnetworks resulting from the enriched highly reliable human PPI data were also presented which could enable a better understanding of disease and process networks as well as signaling pathways associated with them. The reconstructed high-confidence interactome will prevent from spurious discoveries that can be potentially costly, e.g. wrong signaling routes, or wrong drug targets for diseases.

The characterization of efficient molecular targeted therapy is significantly challenging process since the various cellular pathways and multiple crosstalk mechanisms may develop resistance in cancer cells. Thus, multiple targets are required to reduce the development of resistance together with enhancing the clinical effects; our approach pinpoints many cellular targets and driving pathways for breast cancer, triple negative breast cancer and esophageal squamous cell carcinoma that could be explored.

Herein, we suggested a gene signature including 41 down-regulated and 16 up-regulated DEGs as candidate biological markers for breast cancer. The limited studies on down-regulated EHBP1, FBLN5, GCOM1, NLGN1 and TMEM100 and up-regulated PDIA4, PLEK2, RAB15, SORD, TMED3, TMEM132A with breast cancer indicated the necessity of further validations on these genes. In addition, we assessed the association of 57 core-DEGs and 25 hub proteins with breast cancer, which could also be assumed as potential biomarkers and therapeutic targets. The analyses of transcriptional regulatory mechanism signified the driving pathways of breast cancer. As down-regulated TFs were involved in mainly MAPK (ATF3, FOS, FOSB, JUN), TNF (ATF3, FOS, JUN), TGFβ (ATF3, FOS, FOSB, JUN) and WNT (JUN, KLF4, TCF7L2) signaling pathways, the only upregulated TF, STAT1 was involved in JAK-STAT, toll like receptor, chemokine, MAPK, EGFR signaling pathways. Within the framework of network approach, transcriptional control networks included both downstream effectors and upstream regulators were driven by significantly downregulated MAPK, apoptosis, TGFβ and T cell receptor signaling pathways, and upregulated JAK-STAT, the toll receptor, Ras and EGFR signaling pathways. The presented TFs could be selected as therapeutic targets to maximize efficiency of treatment strategies, since they have potential to reduce the function of multiple effector genes simultaneously by targeting the entire breast cancer gene network. Through the metabolic point of view, functional enrichment analysis on reporter metabolites demonstrated the alterations in significant

metabolic pathways including oxidative phosphorylation, lipid metabolic pathways (i.e. biosynthesis of unsaturated fatty acids, fatty acid biosynthesis, linoleic acid metabolism, glycerolipid metabolism and glycerophospholipid metabolism), amino acid metabolic pathways (i.e. eta-Alanine and tyrosine metabolism) and carbohydrate metabolism (i.e. pyruvate metabolism). Hence, further understanding of dysregulated metabolic pathways together with their related signaling pathways is required to design efficient cancer therapeutics. Our network-based approach elucidates the molecular mechanisms of breast cancer through the analyses of genes, gene products, reporter metabolites and associated biological pathways and offers invaluable information to diagnosis and design novel therapeutics for breast cancer.

In the TNBC study, the downregulated CHI3L2, together with upregulated MLPH, were proposed as potential biomarkers for TNBC. In addition, we proposed the association of 40 genes, so-called core-DEGs, and 22 hub proteins with TNBC, which could also be considered as candidate biomarkers and drug targets. The associations of RHOB, ARGHDIB, GSTM3, MLPH, ATF5, CSF2RA, and CXCL10 with TNBC were reported for the first time. Further reliable, low-throughput experimentation (e.g., quantitative RT-PCR) is required to validate the differential expression of these genes encoding potential biomarkers and/or drug targets in breast biopsies from TNBC patients compared to non-TNBC patients. In addition, the protein–protein interaction networks need to confirmed as potential targets for therapeutics using known TNBC cell lines *in vitro*. This study also represents that TNBC is characterized by complex changes in various molecular pathways. Substantial alterations in biochemical pathways of immune system, the cell cycle, apoptosis, proliferation, and cell survival were predicted. The analyses illustrate the part of transcriptional control mechanism that indicates the driving pathways of TNBC, including JAK-STAT and TNF signaling pathways, and possible drug target TFs including MYB, GATA3, FOXA1, MYC, RELA, and CEBPB. To sum up, these results implement an exhaustive approach to figure out the molecular framework for better therapeutic studies in the treatment of TNBC.

In ESCC study, we proposed 30 down-regulated as well as 21 up-regulated DEGs as potential biomarkers for ESCC. Additionally, we asserted the association of 51 genes, so-called core-DEGs, and 20 hub proteins with ESCC, which could also be assumed as promising candidate biomarkers and potential drug targets. The down-regulated ACPP,

C2orf54, DYNLT3, KANK1, ENDOU, FMO2 and up-regulated HOMER3, RFC4, COL10A1, FNDC3B and MARCKSL1 were reported for the first time in ESCC. Further, the validation of these genes encoding potential biomarkers and/or drug targets by RT-qPCR is required in esophageal biopsies from ESCC patients compared to healthy controls. Moreover, the PPI networks could be studied as potential targets for therapeutics using known ESCC cell lines *in vitro.* Another significant direction for our systems approach would be to validate human DEGs from ESCC in the carcinogen-induced rodent models of ESCC. ESCC is characterized by complex changes in different molecular pathways that distinct metabolic pathways such as AA metabolism, biological oxidations and metabolism of lipids and lipoproteins were down-regulated whereas cell cycle and several signaling pathways such as p53 signaling, integrin signaling, PDGF signaling and plasminogen activating cascade were up-regulated in ESCC. The analyses highlights the part of transcriptional control mechanism which signify the driving pathways of ESCC such as estrogen-mediated S phase entry, aryl hydrocarbon receptor signaling, nicotine degradation, bupropion degradation and telomerase signaling pathways for down-regulated core genes together with estrogen mediated S phase entry, aryl hydrocarbon receptor signaling, chronic myeloid leukemia signaling, telomerase signaling and atherosclerosis signaling for up-regulated core genes and possible drug target TFs (AR, CEBPB and E2F1). Additionally, the present metabolomics work identifies the reporter metabolites and related pathways including AA metabolism and steroid hormone biosynthesis.

To sum up, our multi-omics analyses reveal insight into genes/proteins, reporter metabolites and molecular pathways mediating breast cancer, triple-negative breast cancer and ESCC development and provide potentially valuable information for diagnosis and designing novel therapeutic interventions for these diseases.

5 REFERENCES

Abba, M.C., Sun, H., Hawkins, K.A., Drake, J.A., Hu, Y., Nunez, M.I., Gaddis, S., Shi, T., Horvath, S., Sahin, A., Aldaz, C.M. (2007) Breast cancer molecular signatures as determined by SAGE: correlation with lymph node status. Molecular Cancer Research, 5, 881-890.

Abbaszadegan, M.R., Raziee, H.R., Ghafarzadegan, K., Shakeri, M.T., Afsharnezhad, S., Ghavamnasiry, M.R. (2005) Aberrant p16 methylation, a possible epigenetic risk factor in familial esophageal squamous cell carcinoma. International Journal of Gastrointestinal Cancer, 36, 47-54.

Abramovitz, M., Barwick, B.G., Willis, S., Young, B., Catzavelos, C., Li, Z., Kodani, M., Tang, W., Bouzyk, M., Moreno, C.S., Leyland-Jones, B. (2011) Molecular characterisation of formalin-fixed paraffin-embedded (FFPE) breast tumour specimens using a custom 512-gene breast cancer bead array-based platform. British Journal of Cancer, 105, 1574-1581.

Adnane, J., Muro-Cacho, C., Mathews, L., Sebti, S.M., Munoz-Antonia T. (2002) Suppression of rho B expression in invasive carcinoma from head and neck cancer patients. Clinical Cancer Research, 8, 2225-2232.

Agrawal, N., Jiao, Y., Bettegowda, C., Hutfless, S.M., Wang, Y., David, S., Cheng, Y., Twaddell,W.S., Latt, N.L., Shin, E.J., Wang, L.D., Wang, L., Yang, W., Velculescu, V.E., Vogelstein, B., Papadopoulos, N., Kinzler, K.W., Meltzer, S.J. (2012) Comparative genomic analysis of esophageal adenocarcinoma and squamous cell carcinoma. Cancer Discovery, 2(10), 899-905.

Ahmed, A., Wang, J.H., Redmond, H.P. (2013). Silencing of TLR4 increases tumor progression and lung metastasis in a murine model of breast cancer. Annals of Surgical Oncology, 20 Suppl 3, S389-396.

Albert, R. (2005) Scale-free networks in cell biology. Journal of Cell Science, 118(Pt 21), 4947-57.

Alles, M.C., Gardiner-Garden, M., Nott, D.J., Wang, Y., Foekens, J.A., Sutherland, R.L., Musgrove, E.A., Ormandy, C.J. (2009) Meta-analysis and gene set enrichment relative to er status reveal elevated activity of MYC and E2F in the "basal" breast cancer subgroup. PLoS One, 4, e4710.

Al-Souhibani, N., Al-Ahmadi, W., Hesketh, J.E., Blackshear, P.J., Khabar, K.S. (2010). The RNA-binding zinc-finger protein tristetraprolin regulates AU-rich mRNAs involved in breast cancer-related processes. Oncogene, 29(29), 4205-4215.

Amberger, J., Bocchini, C., Hamosh, A. (2011) A new face and new challenges for Online Mendelian Inheritance in Man (OMIM(R)). Human Mutation, 32, 564-567.

Andres, S.A., Brock, G.N., Wittliff, J.L. (2013). Interrogating differences in expression of targeted gene sets to predict breast cancer outcome. BMC Cancer, 13, 326.

Aranda, B., Achuthan, P., Alam-Faruque, Y., Armean, I., Bridge, A., Derow, C., Feuermann, M., Ghanbarian, A.T., Kerrien, S., Khadake, J., Kerssemakers, J., Leroy, C., Menden, M., Michaut, M., Montecchi-Palazzi, L., Neuhauser, S.N., Orchard, S., Perreau, V., Roechert, B., van Eijk, K., Hermjakob, H. (2010) The IntAct molecular interaction database in 2010. Nucleic Acids Research, 38, 525–531.

Arga, K.Y., Onsan, Z.I., Kırdar, B., Ulgen, K.O., Nielsen, J. (2007) Understanding signaling in yeast: insights from network analysis. Biotechnology and Bioengineering, 97, 1246–1258.

Arroyo, R., Sune, G., Zanzoni, A., Duran-Frigola, M., Alcalde, V., Stracker, T.H., Soler-Lopez, M., Aloy, P. (2015). Systematic identification of molecular links between core and candidate genes in breast cancer. Journal of Molecular Biology, 427(6 Pt B), 1436-1450.

Ascano, M., Richmond, A., Borden, P., Kuruvilla, R. (2009) Axonal targeting of Trk receptors via transcytosis regulates sensitivity to neurotrophin responses. Journal of Neuroscience, 29(37), 11674-11685.

Ascierto, M.L., Kmieciak, M., Idowu, M.O., Manjili, R., Zhao, Y., Grimes, M., Dumur, C., Wang, E., Ramakrishnan, V., Wang, X.Y., Bear, H.D., MArincola, F.M., Manjili, M.H. (2012). A signature of immune function genes associated with recurrence-free survival in breast cancer patients. Breast Cancer Research and Treatment, 131(3), 871-880.

Asiago, V.M., Alvarado, L.Z., Shanaiah, N., Gowda, G.A., Owusu-Sarfo, K., Ballas, R.A., Raftery, D. (2010). Early detection of recurrent breast cancer using metabolite profiling. Cancer Research, 70(21), 8309-8318.

Aydin, S. (2013) Presence of adropin, nesfatin-1, apelin-12, ghrelins and salusins peptides in the milk, cheese whey and plasma of dairy cows. Peptides, 43, 83-87.

Baca-Lopez, K., Mayorga, M., Hidalgo-Miranda, A., Gutierrez-Najera, N., Hernandez-Lemus, E. (2012). The role of master regulators in the metabolic/transcriptional coupling in breast carcinomas. PLoS One, 7(8), e42678.

Bailey, C.M., Khalkhali-Ellis, Z., Kondo, S., Margaryan, N.V., Seftor, R.E., Wheaton, W.W., Amir, S., Pins, M.R., Schutte, B.C., Hendrix, M.J. (2005) Mammary serine protease inhibitor (Maspin) binds directly to interferon regulatory factor 6: identification of a novel serpin partnership. Journal of Biological Chemistry, 280, 34210-34217.

Bailey, S.G., Cragg, M.S., Townsend, P.A. (2012). Role of STAT1 in the breast. JAKSTAT, 1(3), 197-199.

Barabasi, A.L., Gulbahce, N., Loscalzo, J. (2011) Network medicine: a network-based approach to human disease. Nature Reviews Genetics, 12(1), 56-68.

Barabasi, A.L., Oltvai, Z.N. (2004) Network biology: understanding the cell's functional organization. Nature Reviews Genetics, 5(2), 101-13.

Barcellos-Hoff, M.H., Akhurst, R.J. (2009). Transforming growth factor-beta in breast cancer: too much, too late. Breast Cancer Research, 11(1), 202.

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C.L., Serova, N., Davis, S., Soboleva, A. (2013) NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Research, 41(Database issue), D991-995.

Barros, F.F., Powe, D.G., Ellis, I.O., Green, A.R. (2010). Understanding the HER family in breast cancer: interaction with ligands, dimerization and treatments. Histopathology, 56(5), 560-572.

Bartek, J., Lukas, J. (2003) Chk1 and Chk2 kinases in checkpoint control and cancer. Cancer Cell, 3, 421-429.

Baselga, J., Segalla, J.G., Roche, H., Del, Giglio, A., Pinczowski, H., Ciruelos, E.M., Filho, S.C., Gómez, P., Van, Eyll, B., Bermejo, B., Llombart, A., Garicochea, B., Durán M.Á., Hoff, P.M., Espié, M., de Moraes, A.A., Ribeiro, R.A., Mathias, C., Gil Gil, M., Ojeda, B., Morales, J., Kwon Ro, S., Li, S., Costa, F. (2012) Sorafenib in combination with capecitabine: an oral regimen for patients with HER2-negative locally advanced or metastatic breast cancer. Journal of Clinical Oncology, 30, 1484-1491.

Bateman, A., Birney, E., Durbin, R., Eddy, S.R., Howe, K.L., Sonhammer, E.L.L. (2000) The Pfam protein families database. Nucleic Acids Research, 28, 263–266.

Bayet-Robert, M., Morvan, D. (2013). Metabolomics reveals metabolic targets and biphasic responses in breast cancer cells treated by curcumin alone and in association with docetaxel. PLoS One, 8(3), e57971.

Becker, K.G., Barnes, K.C., Bright, T.J., Wang, S.A. (2004) The genetic association database. Nature Genetics, 36, 431-432.

Bell, S.P., Dutta, A. (2002) DNA replication in eukaryotic cells. Annual Review of Biochemistry, 71, 333-374.

Bender, R.J., Mac Gabhann, F. (2013) Expression of VEGF and semaphorin genes define subgroups of triple negative breast cancer. PLoS One, 8, e61788.

Ben-Porath, I., Thomson, M.W., Carey, V.J., Ge, R., Bell, G.W., Regev, A., Weinberg, R.A. (2008) An embryonic stem cell–like gene expression signature in poorly differentiated aggressive human tumors. Nature Genetics, 40, 499-507.

Bergum, C., List, K. (2010). Loss of the matriptase inhibitor HAI-2 during prostate cancer progression. Prostate, 70(13), 1422-1428.

Bernards, R., Weinberg, R.A. (2002) A progression puzzle. Nature, 418(6900), 823.

Bertucci, F., Ueno, N.T., Finetti, P., Vermeulen, P., Lucci, A., Robertson, F.M., Marsan, M., Iwamoto, T., Krishnamurthy, S., Masuda, H., Van Dam, P., Woodward, W.A., Cristofanilli, M., Reuben, J.M., Dirix, L., Viens, P., Symmans, W.F., Birnbaum, D., Van Laere, S.J. (2014). Gene expression profiles of inflammatory breast cancer: correlation with response to neoadjuvant chemotherapy and metastasis-free survival. Annals of Oncology, 25(2), 358-365.

Bharadwaj, R., Yu, H. (2004) The spindle checkpoint, aneuploidy, and cancer. Oncogene, 23(11), 2016-2027.

Bieche, I., Girault, I., Urbain, E., Tozlu, S., Lidereau, R. (2004) Relationship between intratumoral expression of genes coding for xenobiotic-metabolizing enzymes and benefit from adjuvant tamoxifen in estrogen receptor alpha-positive postmenopausal breast carcinoma. Breast Cancer Research, 6, R252-263.

Bijangi-Vishehsaraei, K., Saadatzadeh, M.R., Huang, S., Murphy, M.P., Safa, A.R. (2010). 4-(4-Chloro-2-methylphenoxy)-N-hydroxybutanamide (CMH) targets mRNA of the c-FLIP variants and induces apoptosis in MCF-7 human breast cancer cells. Molecular and Cellular Biochemistry, 342(1-2), 133-142.

Bist, P., Leow, S.C., Phua, Q.H., Shu, S., Zhuang, Q., Loh, W.T., Nguyen, T.H., Zhou, J.B., Hooi, S.C., Lim, L.H. (2011). Annexin-1 interacts with NEMO and RIP1 to constitutively activate IKK complex and NF-kappaB: implication in breast cancer metastasis. Oncogene, 30(28), 3174-3185.

Blais, A., Dynlacht, B.D. (2005) Constructing transcriptional regulatory networks. Genes and Development, 19(13), 1499-511.

Blakely, C.M., Sintasath, L., D'Cruz, C.M., Hahn, K.T., Dugan, K.D., Belka, G.K., Chodosh, L.A. (2005) Developmental stage determines the effects of MYC in the mammary epithelium. Development, 132, 1147-1160.

Bolstad, B.M., Irizarry, R.A., Astrand, M., Speed, T.P. (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics, 19,185-193.

Bortoloso, E., Pilati, N., Megighian, A., Tibaldo, E., Sandona, D., Volpe, P. (2006) Transition of Homer isoforms during skeletal muscle regeneration. American Journal of Physiology Cell Physiology, 290(3), C711-718.

Bosch, A., Eroles, P., Zaragoza, R., Vina, J.R., Lluch, A. (2010) Triple-negative breast cancer: molecular features, pathogenesis, treatment and current lines of research. Cancer Treat Review, 36, 206-215.

Boudhraa, Z., Rondepierre, F., Ouchchane, L., Kintossou, R., Trzeciakiewicz, A., Franck, F., Kanitakis, J., Labeille, B., Joubert-Zakeyh, J., Bouchon, B., Perrot, J.L., Mansard, S., Papon, J., Dechelotte, P., Chezal, J.M., Miot-Noirault, E., Bonnet, M., D'Incan, M., Degoul, F. (2014). Annexin A1 in primary tumors promotes melanoma dissemination. Clinical and Experimental Metastasis, 31(7), 749-760.

Bovolenta, L.A., Acencio, M.L., Lemke, N. (2012) HTRIdb: an open-access database for experimentally verified human transcriptional regulation interactions. BMC Genomics, 13, 405.

Boyle, E.I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J.M., Sherlock, G. (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, 3710-3715.

Braun, S., Hepp, F., Kentenich, C.R., Janni, W., Pantel, K., Riethmuller, G., Willgeroth, F., Sommer, H.L. (1999) Monoclonal antibody therapy with edrecolomab in breast cancer patients: monitoring of elimination of disseminated cytokeratin-positive tumor cells in bone marrow. Clinical Cancer Research, 5, 3999-4004.

Braun, S., Pantel, K., Muller, P., Janni, W., Hepp, F., Kentenich, C.R., Gastroph, S., Wischnik, A., Dimpfl, T., Kindermann, G., Riethmüller, G., Schlimok, G. (2000) Cytokeratin-positive cells in the bone marrow and survival of patients with stage I, II, or III breast cancer. New England Journal of Medicine 342, 525-533.

Brennan, M.E., Houssami, N., Lord, S., Macaskill, P., Irwig, L., Dixon, J.M., Warren, R.M., Ciatto, S. (2009). Magnetic resonance imaging screening of the contralateral breast in women with newly diagnosed breast cancer: systematic review and meta-analysis of incremental cancer detection and impact on surgical management. Journal of Clinical Oncology, 27(33), 5640-5649.

Britschgi, A., Andraos, R., Brinkhaus, H., Klebba, I., Romanet, V., Muller, U., Murakami, M., Radimerski, T., Bentires-Alj, M. (2012) JAK2/STAT5 inhibition circumvents resistance to PI3K/mTOR blockade: a rationale for cotargeting these pathways in metastatic breast cancer. Cancer Cell, 22, 796-811.

Brown, T.C., Correia, S.S., Petrok, C.N., Esteban, J.A. (2007) Functional compartmentalization of endosomal trafficking for the synaptic delivery of AMPA receptors during long-term potentiation. Journal of Neuroscience, 27(48), 13311-13315.

Budczies, J., Brockmoller, S.F., Muller, B.M., Barupal, D.K., Richter-Ehrenstein, C., Kleine-Tebbe, A., Griffin, J.L., Orešič, M., Dietel, M., Denkert, C., Fiehn, O. (2013). Comparative

metabolomics of estrogen receptor positive and estrogen receptor negative breast cancer: alterations in glutamine and beta-alanine metabolism. Journal of Proteomics, 94, 279-288.

Cadoo, K. A., Fornier, M. N., Morris, P. G. (2013). Biological subtypes of breast cancer: current concepts and implications for recurrence patterns. The Quarterly Journal of Nuclear Medicine and Molecular Imaging, 57(4), 312-321.

Cai, Z., Zhou, Y., Lei, T., Chiu, J.F., He, Q.Y. (2009) Mammary serine protease inhibitor inhibits epithelial growth factor-induced epithelial-mesenchymal transition of esophageal carcinoma cells. Cancer, 115(1), 36-48.

Cairns, R.A., Harris, I., McCracken, S., Mak, T.W. (2011). Cancer cell metabolism. Cold Spring Harbor Symposia on Quantitative Biology, 76, 299-311.

Calo, V., Migliavacca, M., Bazan, V., Macaluso, M., Buscemi, M., Gebbia, N., Russo, A. (2003). STAT proteins: from normal control of cellular events to tumorigenesis. Journal of Cell Physiology, 197(2), 157-168.

Cannistraci, C.V., Alanis-Lobato, G., Ravasi, T. (2013) Minimum curvilinearity to enhance topological prediction of protein interactions by network embedding. Bioinformatics, 29(13), i199-209.

Cao, B., Li, M., Zha, W., Zhao, Q., Gu, R., Liu, L., Shi, J., Zhou, J., Zhou, F., Wu, X., Wu, Z., Wang, G., Aa, J. (2013). Metabolomic approach to evaluating adriamycin pharmacodynamics and resistance in breast cancer cells. Metabolomics, 9(5), 960-973.

Cao, T., Duprez, E., Borden, K.L., Freemont, P.S., Etkin, L.D. (1998). Ret finger protein is a normal component of PML nuclear bodies and interacts directly with PML. Journal of Cell Science, 111 (Pt 10), 1319-1329.

Cao, W., Shi, F., Wu, L., Yang, K., Tian, F., Chen, G., Wang, W., Wu, W. (2014) Genome wide

screening and characterization of long non-coding RNAs in esophageal cancer. Zhonghua Yi Xue Yi Chuan Xue Za Zhi, 31(5), 587-90.

Cao, Y., Li, Y., Edelweiss, M., Arun, B., Rosen, D., Resetkova, E., Wu, Y., Liu, J., Sahin, A., Albarracin, C.T. (2008). Loss of annexin A1 expression in breast cancer progression. Applied Immunohistochemistry and Molecular Morphology, 16(6), 530-534.

Capdevila, J., Elez, E., Macarulla, T., Ramos, F.J., Ruiz-Echarri, M., Tabernero, J. (2009). Anti-epidermal growth factor receptor monoclonal antibodies in cancer treatment. Cancer Treatmemt Review, 35(4), 354-363.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Impiombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Bansal, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner,G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T.,Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Huminiecki, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H.,Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavesi, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Semple, C.A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugiura, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi,

K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessy, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J., Hayashizaki, Y., Fantom Consortium, Riken Genome Exploration Research Group and Group Genome Science (2005) The Transcriptional Landscape of the Mammalian Genome. Science, 309 (5740), 1559-63.

Cascione, L., Gasparini, P., Lovat, F., Carasi, S., Pulvirenti, A., Ferro, A., Alder, H., He, G., Vecchione, A., Croce, C.M., Shapiro, C.L., Huebner, K. (2013) Integrated microRNA and mRNA signatures associated with survival in triple negative breast cancer. PLoS One, 8, e55910.

Ceol, A., Aryamontri, A.C., Licata, L., Peluso, D., Briganti, L., Perfetto, L., Castagnoli, L., Cesareni, G. (2010) MINT, the molecular interaction database: 2009 uptake. Nucleic Acids Research, 38, 532–539.

Chagoyen, M., Pazos, F. (2011) MBRole: enrichment analysis of metabolomic data. Bioinformatics, 27, 730-731.

Chajes, V., Lanson, M., Fetissof, F., Lhuillery, C., Bougnoux, P. (1995). Membrane fatty acids of breast carcinoma: contribution of host fatty acids and tumor properties. International Journal of Cancer, 63(2), 169-175.

Chan, S.R., Vermi, W., Luo, J., Lucini, L., Rickert, C., Fowler, A.M., Lonardi, S., Arthur, C., Young, L.J., Levy, D.E., Welch, M.J., Cardiff, R.D., Schreiber, R.D. (2012). STAT1-deficient mice spontaneously develop estrogen receptor alpha-positive luminal mammary carcinomas. Breast Cancer Research, 14(1), R16.

Chandriani, S., Frengen, E., Cowling, V.H., Pendergrass, S.A., Perou, C.M., Whitfield, M.L., Cole, M.D. (2009) A core MYC gene expression signature is prominent in basal-like breast cancer but only partially overlaps the core serum response. Plos One, 4, e6693.

Charpentier, A.H., Bednarek, A.K., Daniel, R.L., Hawkins, K.A., Laflin, K.J., Gaddis, S., MacLeod, M.C., Aldaz, C.M. (2000) Effects of estrogen on global gene expression: identification of novel targets of estrogen action. Cancer Research, 60, 5977-5983.

Chen, C., Liang, Z., Huang, W., Li, X., Zhou, F., Hu, X., Han, M., Ding, X., Xiang, S. (2015). Eps8 regulates cellular proliferation and migration of breast cancer. International Journal of Oncology, 46(1), 205-214.

Chen, C., Zhao, J.F., Huang, Q., Wang, R.S., Zhang, X.S. (2002) Inferring domain–domain interactions from protein–protein interactions. Genome Research, 12, 1540–1548.

Chen, J. J., Tsai, C. A., Tzeng, S., Chen, C.H. (2007) Gene selection with multiple ordering criteria. BMC Bioinformics, 8, 74.

Chen, L., Wang, R.S., Zhang, X.S. (2009) Biomolecular networks: methods and applications in systems biology, Wiley series in bioinformatics, A John Wiley and Sons Inc., New Jersey, USA.

Chen, R., Alvero, A.B., Silasi, D.A., Mor, G. (2007) Inflammation, cancer and chemoresistance: taking advantage of the toll like receptor signaling pathway. American Journal of Reproductive Immunology, 57(2), 93-107.

Chen, X., Johns, D.C., Geiman, D.E., Marban, E., Dang, D.T., Hamlin, G., Sun., R., Yang, V.W. (2001) Kruppel-like factor 4 (gut-enriched Kruppel-like factor) inhibits cell proliferation by blocking G1/S progression of the cell cycle. The Journal of Biological Chemistry, 276(32), 30423-8.

Chen, X., Li, N., Wang, S., Hong, J., Fang, M., Yousselfson, J., Yang, P., Newman, R.A., Lubet, R.A., Yang, C.S. (2002) Aberrant arachidonic acid metabolism in esophageal adenocarcinogenesis, and the effects of sulindac, nordihydroguaiaretic acid, and alpha-difluoromethylornithine on tumorigenesis in a rat surgical model. Carcinogenesis, 23(12), 2095-2102.

Chen, Y., Choong, L.Y., Lin, Q., Philp, R., Wong, C.H., Ang, B.K., Tan, Y.L., Loh, M. C., Hew, C. L., Shah, N., Druker, B.J., Chong, P.K., Lim, Y. P. (2007). Differential expression of novel tyrosine kinase substrates during breast cancer development. Molecular and Cellular Proteomics, 6(12), 2072-2087.

Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R., Fisk, D.G., Hirschman, J.E., Hitz, B.C., Karra, K., Krieger, C.J., Miyasato, S.R., Nash, R.S., Park, J., Skrzypek, M.S., Simison, M., Weng, S., Wong, E.D. (2012) Saccharomyces genome database: the genomics resource of budding yeast. Nucleic Acids Research, 40, 700–705.

Cho, H.J., Baek, K.E., Park, S.M., Kim, I.K., Choi, Y.L., Cho, H.J., Nam, I.K., Hwang, E.M., Park, J.Y., Han, J.Y., Kang, S.S., Kim, D.C., Lee, W.S., Lee, M.N., Oh, G.T., Kim, J.W., Lee, C.W., Yoo, J. (2009) RhoGDI2 expression is associated with tumor growth and malignant progression of gastric cancer. Clinical Cancer Research, 15, 2612-2619.

Chu, I.M., Michalowski, A.M., Hoenerhoff, M., Szauter, K.M., Luger, D., Sato, M., Flanders, K., Oshima, A., Csiszar, K., Green, J.E. (2012) GATA3 inhibits lysyl oxidase-mediated metastases of human basal triple-negative breast cancer cells. Oncogene, 31, 2017-2027.

Chuang, H.Y., Lee, E., Liu, Y.T., Lee, D., Ideker, T. (2007). Network-based classification of breast cancer metastasis. Molecular System Biology, 3, 140.

Cimino, D., Fuso, L., Sfiligoi, C., Biglia, N., Ponzone, R., Maggiorotto, F., Russo, G., Cicatiello, L., Weisz, A., Taverna, D., Sismondi P., De Bortoli, M. (2008). Identification of new genes associated with breast cancer progression by gene expression analysis of predefined

sets of neoplastic tissues. International Journal of Cancer, 123(6), 1327-1338.

Cimino-Mathews, A., Subhawong, A.P., Illei, P.B., Sharma, R., Halushka, M.K., Vang, R., Fetting, J.H., Park, B.H., Argani, P. (2013) GATA3 expression in breast carcinoma: utility in triple-negative, sarcomatoid, and metastatic carcinomas. Human Pathology, 44, 1341-1349.

Clark, G.J., Der, C.J. (1995). Aberrant function of the Ras signal transduction pathway in human breast cancer. Breast Cancer Research and Treatment, 35(1), 133-144.

Clarke, C., Madden, S.F., Doolan, P., Aherne, S.T., Joyce, H., O'Driscoll, L., Gallagher, W. M., Hennessy, B.T., Moriarty, M., Crown, J., Kennedy, S., Clynes, M. (2013). Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. Carcinogenesis, 34(10), 2300-2308.

Correia, S.S., Bassani, S., Brown, T.C., Lise, M.F., Backos, D.S., El-Husseini, A., Passafora, M., Esteban, J.A. (2008) Motor protein-dependent transport of AMPA receptors into spines during long-term potentiation. Nature Neuroscience, 11(4), 457-466.

Couderc, B., Pradines, A., Rafii, A., Golzio, M., Deviers, A., Allal, C., Berg, D., Penary, M., Teissie, J., Favre, G. (2008) In vivo restoration of RhoB expression leads to ovarian tumor regression. Cancer Gene Therapy, 15, 456-464.

Coughlin, S.S., Ekwueme, D.U. (2009) Breast cancer as a global health concern. Cancer Epidemiology, 33, 315-318.

Croft, D., O'Kelly, G., Wu, G., Haw, R., Gillespie, M., Matthews, L. Caudy, M., Garapati, P., Gopinath, G., Jassal, B., Jupe, S., Kalatskaya, I., Mahajan, S., May, B., Ndegwa, N., Schmidt, E., Shamovsky, V,. Yung, C., Birney, E., Hermjakob, H., D'Eustachio, P., Stein, L. (2011) Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Research, 39(Database issue), D691-697.

Curran, T., & Franza, B. R., Jr. (1988). Fos and Jun: the AP-1 connection. Cell, 55(3), 395-397.

Curtis, C., Shah, S.P., Chin, S.F., Turashvili, G., Rueda, O. M., Dunning, M.J., Speed, D.,

Lynch, A.G., Samarajiwa, S., Yuan, Y., Gräf, S., Ha, G., Haffari, G., Bashashati, A., Russell, R., McKinney, S., METABRIC Group, Langerød, A., Green, A., Provenzano, E., Wishart, G., Pinder, S., Watson, P., Markowetz, F., Murphy, L., Ellis, I., Purushotham, A., Børresen-Dale, A.L., Brenton, J.D., Tavaré, S., Caldas, C., Aparicio, S. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature, 486(7403), 346-352.

Curtis, C., Shah, S.P., Chin, S.F., Turashvili, G., Rueda, O.M., Dunning, M.J., Speed, D., Lynch, A.G., Samarajiwa, S., Yuan, Y., Graf, S., Ha, G., Haffari, G., Bashashati, A., Russell, R., McKinney, S., Metabric Group, Langerod, A., Green, A., Provenzano, E., Wishart, G., Pinder, S., Watson, P., Markowetz, F., Murphy, L., Ellis, I., Purushotham, A., Borresen-Dale, A.L., Brenton, J.D., Tavare, S., Caldas C., Aparicio, S. (2012) The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature, 486(7403), 346-52.

Cusick, M.E., Yu, H., Smolyar, A., Venkatesan, K., Carvunis, A.R., Simonis, N., Rual, J.F., Borick, H., Braun, P., Dreze, M., Vandenhaute, J., Galli, M., Yazaki, J., Hill, D.E., Ecker, J.R., Roth, F.P., Vidal, M. (2009) Addendum: literature-curated protein interaction datasets. Nature Methods, 6, 934–935.

Cvijovic, M., Olivares-Hernandez, R., Agren, R., Dahr, N., Vongsangnak, W., Nookaew, I., Patil, K.R, Nielsen, J. (2010) BioMet Toolbox: genome-wide analysis of metabolism. Nucleic Acids Research, 38(Web Server issue), W144-149.

Dang, C.V. (1999) c-MYC target genes involved in cell growth, apoptosis and metabolism. Molecular and Cellular Biology, 19, 1-11.

Dang, C.V. (2012) MYC on the path to cancer. Cell, 149(1), 22-35.

Datta, D., Flaxenburg, J.A., Laxmanan, S., Geehan, C., Grimm, M., Waaga-Gasser, A.M., Briscoe, D.M., Pal, S. (2006) Ras-induced modulation of CXCL10 and its receptor splice variant CXCR3-B in MDA-MB-435 and MCF-7 cells: relevance for the development of human breast cancer. Cancer Research, 66, 9509-9518.

Day, T.W., Huang, S., Safa, A.R. (2008). c-FLIP knockdown induces ligand-independent DR5-, FADD-, caspase-8-, and caspase-9-dependent apoptosis in breast cancer cells. Biochemical Pharmacology, 76(12), 1694-1704.

de Graauw, M., van Miltenburg, M.H., Schmidt, M.K., Pont, C., Lalai, R., Kartopawiro, J., Pardali, E., Le Dévédec, S.E., Smit, V.T., van der Wal, A., Van't Veer, L.J., Cleton-Jansen, A.M., ten Dijke, P., van de Water, B. (2010). Annexin A1 regulates TGF-beta signaling and promotes metastasis formation of basal-like breast cancer cells. PNAS, Proceedings of the National Academy of Sciences U S A, 107(14), 6340-6345.

De Santi, M., Galluzzi, L., Lucarini, S., Paoletti, M.F., Fraternale, A., Duranti, A., De Marco, C., Fanelli, M., Zaffaroni, N., Brandi, G., Magnani, M. (2011) The indole-3-carbinol cyclic tetrameric derivative CTet inhibits cell proliferation via overexpression of p21/CDKN1A in both estrogen receptor-positive and triple-negative breast cancer cell lines. Breast Cancer Research, 13, R33.

Deane, C.M., Salwinski, L., Xenarios, I., Eisenberg, D. (2002) Protein interactions: two methods for assessment of the reliability of high throughput observations. Molecular and Cellular Proteomics, 1, 349–356.

Deng, M., Mehta, S., Sun, F., Chen, T. (2002) Inferring domain–domain interactions from protein–protein interactions. Genome Research, 12, pp. 1540–1548.

Deng, M., Zhang, K., Mehta, S., Chen, T., Sun, F. (2003) Prediction of protein function using protein–protein interaction data. Journal of Computational Biology, 10, pp. 947–960.

Dey, N., Barwick, B.G., Moreno, C.S., Ordanic-Kodani, M., Chen, Z., Oprea-Ilies, G., Tang, W., Catzavelos, C., Kerstann, K.F., Sledge, G.W.Jr, Abramovitz, M., Bouzyk, M., De, P., Leyland-Jones, B.R. (2013) Wnt signaling in triple negative breast cancer is associated with metastasis. BMC Cancer, 13, 537.

Dey, R., Moraes, C.T. (2000). Lack of oxidative phosphorylation and low mitochondrial membrane potential decrease susceptibility to apoptosis and do not modulate the protective effect of Bcl-x(L) in osteosarcoma cells. Journal of Biological Chemistry, 275(10), 7087-7094.

DeYoung, M.P., Tress, M., Narayanan, R. (2003) Down's syndrome-associated Single Minded 2 gene as a pancreatic cancer drug therapy target. Cancer Letters, 200(1), 25-31.

Diao, Y., Li, H., Li, H., Zhou, Y., Ma, Q., Wang, Y., Li, D. (2009) Association of serum levels of lipid and its novel constituents with the different stages of esophageal carcinoma. Lipids in Health and Disease, 8, 48.

Dluzen, D., Li, G., Tacelosky, D., Moreau, M., Liu, D.X. (2011) BCL-2 is a downstream target of ATF5 that mediates the prosurvival function of ATF5 in a cell type-dependent manner. Journal of Biological Chemistry, 286, 7705-7713.

Dominguez, G., Silva, J.M., Silva, J., Garcia, J.M., Sanchez, A., Navarro, A., Gallego, I., Provencio, M., España, P., Bonilla, F. (2001). Wild type p73 overexpression and high-grade malignancy in breast cancer. Breast Cancer Research and Treatment, 66(3), 183-190.

Dong, W., Chen, X., Xie, J., Sun, P., Wu, Y. (2010). Epigenetic inactivation and tumor suppressor activity of HAI-2/SPINT2 in gastric cancer. International Journal of Cancer, 127(7), 1526-1534.

Dotto, G.P. (1999) Signal transduction pathways controlling the switch between keratinocyte growth and differentiation. Critical Reviews in Oral Biology and Medicine, 10(4), 442-457.

Downward, J. (2003). Targeting RAS signalling pathways in cancer therapy. Nature Reviews Cancer, 3(1), 11-22.

Drabsch, Y., Hugo, H., Zhang, R., Dowhan, D.H., Miao, Y.R., Gewirtz, A.M., Barry, S.C., Ramsay, RG., Gonda, T.J. (2007) Mechanism of and requirement for estrogen-regulated MYB

expression in estrogen-receptor-positive breast cancer cells. Proceedings of the National Academy Sciences USA, 104, 13762-13767.

Drabsch, Y., Robert, R.G., Gonda, T.J. (2010) MYB suppresses differentiation and apoptosis of human breast cancer cells. Breast Cancer Research, 12, R55.

Dressman, H.K., Hans, C., Bild, A., Olson, J.A., Rosen, E., Marcom, P.K., Liotcheva, V.B., Jones, E.L., Vujaskovic, Z., Marks, J., Dewhirst, M.W., West, M., Nevins, J.R., Blackwell, K. (2006). Gene expression profiles of multiple breast cancer phenotypes and response to neoadjuvant chemotherapy. Clinical Cancer Research, 12(3 Pt 1), 819-826.

Dumay, A., Feugeas, J. P., Wittmer, E., Lehmann-Che, J., Bertheau, P., Espie, M., Plassa, L.F., Cottu, P., Marty, M., André, F., Sotiriou, C., Pusztai, L., de The, H. (2013). Distinct tumor protein p53 mutants in breast cancer subgroups. International Journal of Cancer, 132(5), 1227-1231.

Dumitru, C.A., Bankfalvi, A., Gu, X., Zeidler, R., Brandau, S., Lang, S. (2013) AHNAK and inflammatory markers predict poor survival in laryngeal carcinoma. PLoS One, 8(2), e56420.

Durchdewald, M., Angel, P., Hess, J. (2009). The transcription factor Fos: a Janus-type regulator in health and disease. Histology and Histopathology, 24(11), 1451-1461.

Dydensborg, A.B., Rose, A.A., Wilson, B.J., Grote, D., Paquet, M., Giguere, V., Siegel, P.M., Bouchard M. (2009) GATA3 inhibits breast cancer growth and pulmonary breast cancer metastasis. Oncogene, 28, 2634-2642.

Eccles, S.A., Aboagye, E.O., Ali, S., Anderson, A.S., Armes, J., Berditchevski, F., Blaydes, J.P., Brennan, K., Brown, N.J., Bryant, H.E., Bundred, N.J., Burchell, J.M., Campbell, A.M., Carroll, J.S., Clarke, R.B., Coles, C.E., Cook, G.J., Cox, A., Curtin, N.J., Dekker, L.V., Silva Idos, S., Duffy, S.W., Easton, D.F., Eccles, D.M., Edwards, D.R., Edwards, J., Evans, D., Fenlon, D.F., Flanagan, J.M., Foster, C., Gallagher, W.M., Garcia-Closas, M., Gee, J.M., Gescher, A.J., Goh, V., Groves, A.M., Harvey, A.J., Harvie, M., Hennessy, B.T., Hiscox, S.,

Holen, I., Howell, S.J., Howell, A., Hubbard, G., Hulbert-Williams, N., Hunter, M.S., Jasani, B., Jones, L.J., Key, T.J., Kirwan, C.C., Kong, A., Kunkler, I.H., Langdon, S.P., Leach, M.O., Mann, D.J., Marshall, J.F., Martin, L., Martin, S.G., Macdougall, J.E., Miles, D.W., Miller, W.R., Morris, J.R., Moss, S.M., Mullan, P., Natrajan, R., O'Connor, J.P., O'Connor, R., Palmieri, C., Pharoah, P.D., Rakha, E.A., Reed, E., Robinson, S.P., Sahai, E., Saxton, J.M., Schmid, P., Smalley, M.J., Speirs, V., Stein, R., Stingl, J., Streuli, C.H., Tutt, A.N., Velikova, G., Walker, R.A., Watson, C.J., Williams, K.J., Young L.S., Thompson, A.M. (2013) Critical research gaps and translational priorities for the successful prevention and treatment of breast cancer. Breast Cancer Research, 15(5), R92.

Edwards, A.M., Kus, B., Jansen, R., Greenbaum, D., Greenblatt, J., Gerstein, M. (2002) Bridging structural biology and genomics: assessing protein interaction data with known complexes. Trends in Genetics, 18, pp. 529–536.

Eichhorn, P.J., Creyghton, M.P., Bernards, R. (2009). Protein phosphatase 2A regulatory subunits and cancer. Biochimica et Biophysica Acta, 1795(1), 1-15.

Eilers, M., Eisenman, R.N. (2008) Myc's broad reach. Genes and Development 22, 2755-2766.
El Baroudi, M., La Sala, D., Cinti, C., Capobianco, E. (2013) DAC-driven integrative network regulation and pathway coordination in breast cancer. IWBBIO Proceedings, 413-425.

Ell, B., Kang, Y. (2013) Transcriptional control of cancer metastasis. Trends Cell Biology, 23, 603-611.

Ell, B., Kang, Y. (2014) MicroRNAs as regulators of bone homeostasis and bone metastasis. BoneKEy Reports. 2014;3:549.

Emmert-Streib, F., de Matos Simoes, R., Mullan, P., Haibe-Kains, B., Dehmer, M. (2014). The gene regulatory network for breast cancer: integrated regulatory landscape of cancer hallmarks. Frontiers in Genetics, 5, 15.

Enzinger, P.C., Mayer, R.J. (2003) Esophageal cancer. The New England Journal of Medicine,

349(23), 2241-2252.

Eom, M., Han, A., Yi, S.Y., Shin, J.J., Cui, Y., Park, K.H. (2008). RHEB expression in fibroadenomas of the breast. Pathology International, 58(4), 226-232.

Esplin, E D., Ramos, P., Martinez, B., Tomlinson, G.E., Mumby, M.C., Evans, G. A. (2006). The glycine 90 to aspartate alteration in the Abeta subunit of PP2A (PPP2R1B) associates with breast cancer and causes a deficit in protein function. Genes Chromosomes Cancer, 45(2), 182-190.

Essack, M., Radovanovic, A., Schaefer, U., Schmeier, S., Seshadri, S.V., Christoffels, A., Akur, M., Bajic, V.B. (2009) DDEC: Dragon database of genes implicated in esophageal cancer. BMC Cancer, 9, 219.

Esseghir, S., Kennedy, A., Seedhar, P., Nerurkar, A., Poulsom, R., Reis-Filho, J.S., Isacke, C.M. (2007) Identification of NTN4, TRA1, and STC2 as prognostic markers in breast cancer in a screen for signal sequence encoding proteins. Clinical Cancer Research, 13, 3164-3173.

Ewing, R.M., Chu, P., Elisman, F., Li, H., Taylor, P., Climie, S., McBroom-Cerajewski, L., Robinson, M.D., O'Connor, L., Li, M., Taylor, R., Dharsee, M., Ho, Y., Heilbut, A., Moore, L., Zhang, S., Ornatsky, O., Bukhman, Y.V., Ethier, M., Sheng, Y., Vasilescu, J., Abu-Farha, M., Lambert J.P., Duewel, H.S., Stewart, I.I., Kuehl, B., Hogue, K., Colwill, K., Gladwish, K., Muskat, B., Kinach, R., Adams, S.L., Moran, M.F., Morin, G.B., Topaloglou, T., Figeys, D. (2007) Large-scale mapping of human protein–protein interactions by mass spectrometry. Molecular Systems Biology, 3, 1–17.

Fan, N.J., Gao, C.F., Wang, C.S., Zhao, G., Lv, J.J., Wang, X.L., Chun, G.H., Yin, J., Li, D.H., Chen, X., Yuan, X.T., Meng, N.L. (2012) Identification of the up-regulation of TP-alpha, collagen alpha-1(VI) chain, and S100A9 in esophageal squamous cell carcinoma by a proteomic method. Journal of Proteomics., 75(13), 3977-3986.

Fang, M.Z., Liu, C., Song, Y., Yang, G.Y., Nie, Y., Liao, J., Zhao, X., Shimada, Y., Wang, L.D., Yang, C.S. (2004) Over-expression of gastrin-releasing peptide in human esophageal

squamous cell carcinomas. Carcinogenesis, 25(6), 865-871.

Fazioli, F., Minichiello, L., Matoska, V., Castagnino, P., Miki, T., Wong, W.T., Di Fiore, P. P. (1993). Eps8, a substrate for the epidermal growth factor receptor kinase, enhances EGF-dependent mitogenic signals. The EMBO Journal, 12(10), 3799-3808.

Feist, A.M., Herrgard, M.J., Thiele, I., Reed, J.L., Palsson, B.O. (2009) Reconstruction of biochemical networks in microorganisms. Nature Reviews Microbiology. 7(2), 129-43.

Feldman, I., Rzhetsky, A., Vitkup, D. (2008) Network properties of genes harboring inherited disease mutations. Proceedings of the National Academy of Sciences of the United States of America. 105(11), 4323-8.

Ferlay, J., Parkin, D.M., Steliarova-Foucher, E. (2010) Estimates of cancer incidence and mortality in Europe in 2008. European Journal of Cancer. 46(4), 765-81.

Ferrandina, G., Lauriola, L., Zannoni, G.F., Fagotti, A., Fanfani, F., Legge, F., Maggiano, N., Gessi, M., Mancuso, S., Ranelletti, F.O., Scambia, G. (2002) Increased cyclooxygenase-2 (COX-2) expression is associated with chemotherapy resistance and outcome in ovarian cancer patients. Annals of Oncology, 13, 1205-1211.

Fleming, I. (2007) Epoxyeicosatrienoic acids, cell signaling and angiogenesis. Prostaglandins Other Lipid Mediators, 82(1-4), 60-67.

Fountzilas, G., Dafni, U., Gogas, H., Linardou, H., Kalofonos, H.P., Briasoulis, E., Pectasides D, Samantas E, Bafaloukos D, Stathopoulos, G.P., Karina, M., Papadimitriou, C., Skarlos, D., Pisanidis, N., Papakostas, P., Markopoulos, C., Tzorakoeleftherakis, E., Dimitrakakis, K., Makrantonakis, P., Xiros, N., Polichronis, A., Varthalitis, I., Karanikiotis, C., Dimopoulos, A.M., Hellenic Cooperative Oncology Group. (2008) Postoperative dose-dense sequential chemotherapy with epirubicin, paclitaxel and CMF in patients with high-risk breast cancer: safety analysis of the Hellenic Cooperative Oncology Group randomized phase III trial HE 10/00. Annals of Oncology, 19, 853-860.

Fraser, H.B., Hirsh, A.E., Steinmetz, L.M., Scharfe, C., Feldman, M.W. (2002) Evolutionary rate in the protein interaction network. Science. 296(5568), 750-2.

Freedman, R.S., Wang, E., Voiculescu, S., Patenia, R., Bassett, R.L., Deavers, M., Marincola, F.M, Yang, P., Newman, R.A. (2007) Comparative analysis of peritoneum and tumor eicosanoids and pathways in advanced ovarian cancer. Clinical Cancer Research, 13(19), 5736-5744.

Frohlich, H. (2011). Network based consensus gene signatures for biomarker discovery in breast cancer. PLoS One, 6(10), e25364.

Fry, A.M. (2002) The Nek2 protein kinase: a novel regulator of centrosome structure. Oncogene, 21(40), 6184-6194.

Fujita, T., Ikeda, H., Taira, N., Hatoh, S., Naito, M., Doihara, H. (2009) Overexpression of UbcH10 alternates the cell cycle profile and accelerate the tumor proliferation in colon cancer. BMC Cancer, 9, 87.

Gallo, R., Provenzano, C., Carbone, R., Di Fiore, P.P., Castellani, L., Falcone, G., Alema, S. (1997). Regulation of the tyrosine kinase substrate Eps8 expression by growth factors, v-Src and terminal differentiation. Oncogene, 15(16), 1929-1936.

Galluzzi, L., De Santi, M., Crinelli, R., De Marco, C., Zaffaroni, N., Duranti, A., Brandi, G., Magnani, M. (2012) Induction of endoplasmic reticulum stress response by the indole-3-carbinol cyclic tetrameric derivative CTet in human breast cancer cell lines. PLoS One, 7, e43249.

Galus, R., Wlodarski, K. (2007) Cytokeratin detection as a diagnostic tool in oncology. Pol Merkur Lekarski, 23, 209-211.

Ganapathy, V., Ge, R., Grazioli, A., Xie, W., Banach-Petrosky, W., Kang, Y., Lonning, S., McPherson, J., Yingling, J.M., Biswas, S., Mundy, G.R., Reiss, M. (2010). Targeting the Transforming Growth Factor-beta pathway inhibits human basal-like breast cancer metastasis.

Molecular Cancer, 9, 122.

Gao, Z., Cox, J.L., Gilmore, J.M., Ormsbee, B.D., Mallanna, S.K., Washburn, M.P., Rizzino, A. (2012) Determination of protein interactome of transcription factor Sox2 in embryonic stem cells engineered for inducible expression of four reprogramming factors. Journal of Biological Chemistry, 287(14), 11384-11397.

Garcia-Galiano, D., Navarro, V.M., Gaytan, F., Tena-Sempere, M. (2010) Expanding roles of NUCB2/nesfatin-1 in neuroendocrine regulation. Journal of Molecular Endocrinology, 45(5), 281-290.

Gatza, M.L., Lucas, J.E., Barry, W.T., Kim, J.W., Wang, Q., Crawford, M.D., Datto, M.B., Kelley, M., Mathey-Prevot, B., Potti, A., Nevins, J.R. (2010) A pathway-based classification of human breast cancer. Proceedings National Academy Sciences, 107, 6994-6999.

Gautier, L., Cope, L., Bolstad, B.M., Irizarry, R.A. (2004) Affy- analysis of Affymetrix GeneChip data at the probe level. Bioinformatics, 20,307-315.

Gebeshuber, C.A., Zatloukal, K., Martinez, J. (2009). miR-29a suppresses tristetraprolin, which is a regulator of epithelial polarity and metastasis. EMBO Reports, 10(4), 400-405.

Gentil, B.J., Delphin, C., Mbele, G.O., Deloulme, J.C., Ferro, M., Garin, J., Baudier, J. (2001) The giant protein AHNAK is a specific target for the calcium- and zinc-binding S100B protein: potential implications for Ca2+ homeostasis regulation by S100B. Journal of Biological Chemistry, 276(26), 23253-23261.

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

Gerhardt, J., Steinbrech, C., Buchi, O., Behnke, S., Bohnert, A., Fritzsche, F., Liewen, H., Stenner, F., Wild, P., Hermanns, T., Müntener, M., Dietel, M., Jung, K., Stephan, C., Kristiansen, G. (2011). The androgen-regulated Calcium-Activated Nucleotidase 1 (CANT1) is commonly overexpressed in prostate cancer and is tumor-biologically relevant in vitro. American Journal of Pathology, 178(4), 1847-1860.

Gilchrist, M., Henderson, W.R., Jr., Morotti, A., Johnson, C.D., Nachman, A., Schmitz, F., Schmitz, F., Smith, K.D., Aderem, A. (2010). A key role for ATF3 in regulating mast cell survival and mediator release. Blood, 115(23), 4734-4741.

Giot, L., Bader, J.S., Brouwer, C., Chaudhuri, C., Kuang, B., Li, Y., Hao, Y.L., Ooi, C.E., Godwin, B., Vitols, E., Vijayadamodar, G., Pochart, P., Machineni, H., Welsh, M., Kong, Y., Zerhusen, B., Malcolm, R., Varrone, Z., Colis, A., Minton, M., Burgess, S., McDaniel, L., Stimpson, E., Spriggs, F., Williams, J., Neurath, K., Ioime, N., Agee, M., Voss, E., Furtak, K., Renzulli, R., Aanensen, N., Carrolla, S., Bickelhaupt, E., Lazovatsky, Y., DaSilya, A., Zhong, J., Stanyon Jr., C.A., Finley, R.L., White, K.P., Braverman, M., Jarvie, T., Gold, S., Leach, M., Knight, J., Shimkets, R.A., McKenna, M.P., Chant, J., Rothberg, J.M. (2003) A protein interaction map of Drosophila melanogaster. Science, 302, 1727–1736.

Gira, A.K., Kowalczyk, A.P., Feng, Y., Swerlick RA (2009) Iron chelators and hypoxia mimetics inhibit IFNgamma-mediated Jak-STAT signaling. Journal of Investigative Dermatology, 129, 723-729.

Girvan, M., Newman, M.E. (2002) Community structure in social and biological networks. Proceedings of National Academy of Sciences U S A., 99(12), 7821-7826.

Gomis, R.R., Alarcon, C., Nadal, C., Van Poznak, C., Massague, J. (2006) C/EBPbeta at the core of the TGFbeta cytostatic response and its evasion in metastatic breast cancer cells. Cancer Cell, 10, 203-214.

Goshu, E., Jin, H., Lovejoy, J., Marion, J.F., Michaud, J.L., Fan, C.M. (2004) Sim2 contributes to neuroendocrine hormone gene expression in the anterior hypothalamus. Molecular

Endocrinology, 2004, 18(5), 1251-1262.

Grass, G.D., Tolliver, L.B., Bratoeva, M., Toole, B.P. (2013). CD147, CD44, and the epidermal growth factor receptor (EGFR) signaling pathway cooperate to regulate breast epithelial cell invasiveness. Journal of Biological Chemistry, 288(36), 26089-26104.

Grau, A.M., Wen, W., Ramroopsingh, D.S., Gao, Y.T., Zi, J., Cai, Q., Shu, X.O., Zheng, W. (2008). Circulating transforming growth factor-beta-1 and breast cancer prognosis: results from the Shanghai Breast Cancer Study. Breast Cancer Research and Treatment, 112(2), 335-341.

Griffiths, C.L., Olin, J.L. (2012) Triple negative breast cancer: a brief review of its characteristics and treatment options. Journal of Research in Pharmacy Practice, 25, 319-323.

Grivas, P.D., Kiaris, H., Papavassiliou, A.G. (2011) Tackling transcription factors: challenges in antitumor therapy. Trends in Molecular Medicine, 17(10), 537-538.

Groblewska, M., Siewko, M., Mroczko, B., Szmitkowski, M. (2012) The role of matrix metalloproteinases (MMPs) and their inhibitors (TIMPs) in the development of esophageal cancer. Folia Histochem et Cytobiologica, 50(1), 12-19.

Gu, Z.D., Li, J.Y., Li, M., Gu, J., Shi, X.T., Ke, Y., Chen, K.N. (2005) Matrix metalloproteinases expression correlates with survival in patients with esophageal squamous cell carcinoma. The American Journal of Gastroenterology, 100(8), 1835-1843.

Guilluy, C., Garcia-Mata, R., Burridge, K. (2011) Rho protein crosstalk: another social network? Trends Cell Biology, 21, 718-726.
Guo, C., Liu, S., Sun, M.Z. (2013). Potential role of Anxa1 in cancer. Future Oncology, 9(11), 1773-1793.

Guo, M.; Roman, R. J.; Fenstermacher, J. D.; Brown, S. L.; Falck, J. R.; Arbab, A. S.; Edwards, P.A.; Scicli, A. G. (2006) 9L gliosarcoma cell proliferation and tumor growth in rats are suppressed by N-hydroxy-N'-(4-butyl-2-methylphenol) formamidine (HET0016), a selective

inhibitor of CYP4A. Journal of Pharmacology and Experimental Therapeutics, 317(1), 97-108.

Guo, X.Z.; Zhang, G.; Wang, J.Y.; Liu, W.L.; Wang, F.; Dong, J.Q.; Xu, L.H; Cao, J.Y.; Song, L.B.;Zeng, M.S. (2008)Prognostic relevance of Centromere protein H expression in esophageal carcinoma. BMC Cancer, 8, 233.

Gupta, S., Srivastava, M., Ahmad, N., Sakamoto, K., Bostwick, D.G., Mukhtar, H. (2001) Lipoxygenase-5 is overexpressed in prostate adenocarcinoma. Cancer, 91(4), 737-743.

Güldener, U., Münsterkötter, M., Oesterheld, M., Pagel, P., Ruepp, A., Mewes, H.W., Stümpflen, V. (2006) MPact: the MIPS protein interaction resource on yeast. Nucleic Acids Research, 34, 436–441.

Haase, H., Podzuweit, T., Lutsch, G., Hohaus, A., Kostka, S., Lindschau, C., Kott, M., Kraft, R., Morano, I. (1999) Signaling from beta-adrenoceptor to L-type calcium channel: identification of a novel cardiac protein kinase A target possessing similarities to AHNAK. Faseb Journal, 13(15), 2161-2172.

Hai, T., Wolford, C.C., Chang, Y.S. (2010). ATF3, a hub of the cellular adaptive-response network, in the pathogenesis of diseases: is modulation of inflammation a unifying component? Gene Expression, 15(1), 1-11.

Halvorsen, O.J., Rostad, K., Oyan, A.M., Puntervoll, H., Bo, T.H., Stordrange, L., Olsen, S., Haukaas, S.A., Hood, L., Jonassen, I., Akslen, L.A. (2007) Increased expression of SIM2-s protein is a novel marker of aggressive prostate cancer. Clinical Cancer Research, 13(3), 892-897.

Hammond, M.E., Hayes, D.F., Dowsett, M., Allred, D.C., Hagerty, K.L., Badve, S., Fitzgibbons, P.L., Francis, G., Goldstein, N.S., Hayes, M., Hicks, D.G., Lester, S., Love, R., Mangu, P.B., McShane, L., Miller, K., Osborne, C.K., Paik, S., Perlmutter, J., Rhodes, A., Sasano, H., Schwartz, J.N., Sweep, F.C., Taube, S., Torlakovic, E.E., Valenstein, P., Viale, G.,Visscher, D., Wheeler, T., Williams, R.B., Wittliff, J.L., Wolff, A.C. (2010) American

Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. Journal of Clinical Oncology, 28(16), 2784-2795.

Hanna, R.A., Campbell, R.L., Davies, P.L. (2008) Calcium-bound structure of calpain and its mechanism of inhibition by calpastatin. Nature, 456(7220), 409-412.

Hansen, L.V., Laerum, O.D., Illemann, M., Nilesen, B.Y., Ploug, M. (2008) Altered expression of the urokinase receptor homologue, C4.4A, in invasive areas of human esophageal squamous cell carcinoma. Cancer Cell Biology, 122(4), 734-741.

Harrington, E.D., Jensen, L.,J., Bork, P. (2008) Predicting biological networks from genomic data. FEBS Letters, 582, 1251–1258.

Hasko, G., Linden, J., Cronstein, B., Pacher, P. (2008) Adenosine receptors: therapeutic aspects for inflammatory and immune diseases. Nature Reviews Drug Discovery, 7, 759-770.

Hatakeyama, S. (2011). TRIM proteins and cancer. Nature Reviews Cancer, 11(11), 792-804.
Hayward, D.G., Clarke, R.B., Faragher, A.J., Pillai, M.R., Hagan, I.M., Fry, A.M. (2004) The centrosomal kinase Nek2 displays elevated levels of protein expression in human breast cancer. Cancer Research, 64(20), 7370-7376.

Heemskerk, M.H. (2010). T-cell receptor gene transfer for the treatment of leukemia and other tumors. Haematologica, 95(1), 15-19.

Heinlein, C.A., Chang, C. (2004) Androgen receptor in prostate cancer. Endocrine Reviews, 25(2), 276-308.

Henderson, B.E., Feigelson, H.S. (2000) Hormonal carcinogenesis. Carcinogenesis, 21(3), 427-433.

Hibbs, M.A., Hess, D.C., Myers, C.L., Huttenhower, C., Li, K., Hess, D.C., Myers, C.L., C.

Huttenhower, C., Li, K., Troyanskaya, O.G. (2007) Exploring the functional landscape of gene expression: directed search of large microarray compendia. Nucleic Acids Research, 23, 2692–2699.

Hicks, C., Kumar, R., Pannuti, A., Backus, K., Brown, A., Monico, J., Miele, L. (2013) An Integrative Genomics Approach for Associating GWAS Information with Triple-Negative Breast Cancer. Cancer Informatics, 12, 1-20.

Hilvo, M., Denkert, C., Lehtinen, L., Muller, B., Brockmoller, S., Seppanen-Laakso, T., Budczies, J., Bucher, E., Yetukuri, L., Castillo, S., Berg, E., Nygren, H., Sysi-Aho, M., Griffin, J.L., Fiehn, O., Loibl, S., Richter-Ehrenstein, C., Radke, C., Hyötyläinen, T., Kallioniemi, O., Iljin, K., Oresic, M. (2011). Novel theranostic opportunities offered by characterization of altered membrane lipid metabolism in breast cancer progression. Cancer Research, 71(9), 3236-3245.

Hinkle, P.C., Kumar, M.A., Resetar, A., Harris, D.L. (1991). Mechanistic stoichiometry of mitochondrial oxidative phosphorylation. Biochemistry, 30(14), 3576-3582.

Hollestelle, A., Elstrodt, F., Nagel, J.H., Kallemeijn, W.W., Schutte, M. (2007). Phosphatidylinositol-3-OH kinase or RAS pathway mutations in human breast cancer cell lines. Molecular Cancer Research, 5(2), 195-201.

Holm, K., Hegardt, C., Staaf, J., Vallon-Christersson, J., Jonsson, G., Olsson, H., Borg, A., Ringnér, M. (2010) Molecular subtypes of breast cancer are associated with characteristic DNA methylation patterns. Breast Cancer Research, 12, R36.

Holmes, R.S., Vaughan, T.L. (2007) Epidemiology and pathogenesis of esophageal cancer. Seminars in Radiation Oncology, 17(1), 2-9.

Horiuchi, D., Kusdra, L., Huskey, N.E., Chandriani, S., Lenburg, M.E., Gonzalez-Angulo, A.M., Creasman, K.J., Bazarov, A.V., Smyth, J.W., Davis, S.E., Yaswen, P., Mills, G.B., Esserman, L.J., Goga, A. (2012) MYC pathway activation in triple-negative breast cancer is

synthetic lethal with CDK inhibition. Journal of Experimental Medicine, 209, 679-696.

Hornberg, J.J., Bruggeman, F.J., Westerhoff, H.V., Lankelma J. (2006) Cancer: a Systems Biology disease. Bio Systems. 83(2-3), 81-90.

Howe, E.N., Cochrane, D.R., Cittelly, D.M., Richer, J.K. (2012) miR-200c targets a NF-kappaB up-regulated TrkB/NTF3 autocrine signaling loop to enhance anoikis sensitivity in triple negative breast cancer. PLoS One, 7, e49987.

Hsia, T.C., Tu, C.Y., Chen, Y.J., Wei, Y.L., Yu, M.C., Hsu, S.C., Tsai, S.L., Chen, W.S., Yeh, M.H., Yen, C.J., Yu, Y.L., Huang, T.C., Huang, C.Y., Hung, M.C., Huang, W.C. (2013) Lapatinib-mediated cyclooxygenase-2 expression via epidermal growth factor receptor/HuR interaction enhances the aggressiveness of triple-negative breast cancer cells. Molecular Pharmacology, 83, 857-869.

Hu, L.D., Zou, H.F., Zhan, S.X., Cao, K.M. (2008). EVL (Ena/VASP-like) expression is up-regulated in human breast cancer and its relative expression level is correlated with clinical stages. Oncology Reports, 19(4), 1015-1020.

Hu, N., Clifford, R.J., Yang, H.H., Wang, C., Goldstein, A.M., Ding, T., Taylor, P.R., Lee, M.P. (2010) Genome wide analysis of DNA copy number neutral loss of heterozygosity (CNNLOH) and its relation to gene expression in esophageal squamous cell carcinoma. BMC Genomics, 11, 576.

Huang da, W., Sherman, B.T., Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols, 4(1), 44-57.

Huang, B., Hu, B., Su, M., Tian, D., Guo, Y., Lian, S., Wu, X., Li, Q., Zheng, R., Gao, Y. (2011) Potential role of minichromosome maintenance protein 2 as a screening biomarker in esophageal cancer high-risk population in China. Human Pathology, 42(6), 808-816.

Huang, L., He, Z., Guo, L., Wang, H. (2008) Improvement of cognitive deficit and neuronal

damage in rats with chronic cerebral ischemia via relative long-term inhibition of rho-kinase. Cellular and Molecular Neurobiology, 28, 757-768.

Huang, R., Faratian, D., Sims, A.H., Wilson, D., Thomas, J.S., Harrison, D.J., Langdon, S.P. (2014). Increased STAT1 signaling in endocrine-resistant breast cancer. PLoS One, 9(4), e94226.

Huang, R.P., Fan, Y., de Belle, I., Niemeyer, C., Gottardis, M.M., Mercola, D., Adamson, E. D. (1997). Decreased Egr-1 expression in human, mouse and rat mammary cells and tissues correlates with tumor formation. International Journal of Cancer, 72(1), 102-109.

Huang, Y., Wange, R.L. (2004). T cell receptor signaling: beyond complex complexes. The Journal of the Biological Chemistry, 279(28), 28827-28830.

Hui, M.K., Chan, K.W., Luk, J.M., Lee, N.P., Chung, Y., Cheung, L.C., Srivastava, G., Tsao, S.W., Tang, J.C., Law, S. (2012) Cytoplasmic Forkhead box M1 (FoxM1) in esophageal squamous cell carcinoma significantly correlates with pathological disease stage. World Journal of Surgery., 36(1), 90-97.

Idowu, M.O., Kmieciak, M., Dumur, C., Burton, R.S., Grimes, M.M., Powers, C.N., Manjili, M.H. (2012) CD44(+)/CD24(-/low) cancer stem/progenitor cells are more abundant in triple-negative invasive breast carcinoma phenotype and are associated with poor outcome. Human Pathology, 43, 364-373.

Inman, G.J. (2011). Switching TGFbeta from a tumor suppressor to a tumor promoter. Current Opinion in Genetics and Development, 21(1), 93-99.

Iossifov, I., Krauthammer, M., Friedman, C., Hatzivassiloglou, V., Bader, J.S., White, K.P., Rzhetsky, A. (2004) Probabilistic inference of molecular networks from noisy data sources. Bioinformatics, 20, 1205–1213.

Ishiguro, K., Xavier, R. (2004) Homer-3 regulates activation of serum response element in T

cells via its EVH1 domain. Blood, 103(6), 2248-2256.

Ismail-Khan, R., Bui, M.M. (2010) A review of triple-negative breast cancer. Cancer Control, 17, 173-176.

Ivanovic, V., Demajo, M., Krtolica, K., Krajnovic, M., Konstantinovic, M., Baltic, V., Prtenjak, G., Stojiljković, B., Breberina, M., Nesković-Konstantinović, Z., Nikolić-Vukosavljević, D., Dimitrijevic, B. (2006). Elevated plasma TGF-beta1 levels correlate with decreased survival of metastatic breast cancer patients. Clinica Chimica Acta, 371(1-2), 191-193.

Iwaya, K., Mukai, K. (2005) Accumulation of ubiquitin-conjugated cytokeratin fragments in tumor cells. Seminars in Cancer Biology, 15, 309-318.

Jacinto, F.V., Ballestar, E., Ropero, S., Esteller, M. (2007) Discovery of epigenetically silenced genes by methylated DNA immunoprecipitation in colon cancer cells. Cancer Research, 67(24), 11481-11486.

Jager, W., Gruber, A., Giessrigl, B., Krupitza, G., Szekeres, T., Sonntag, D. (2011). Metabolomic analysis of resveratrol-induced effects in the human breast cancer cell lines MCF-7 and MDA-MB-231. OMICS, 15(1-2), 9-14.

Jansen, R., Gerstein, M. (2004) Analyzing protein function on a genomic scale: the importance of gold-standard positives and negatives for network prediction. Current Opinion in Microbiology, 7, 535–545.

Jansen, R., Greenbaum, D., Gerstein, M. (2002) Relating whole-genome expression data with protein–protein interactions. Genome Research, 12, 37–46.

Janssens, V., Goris, J., Van Hoof, C. (2005). PP2A: the expected tumor suppressor. Current Opinion in Genetics and Development, 15(1), 34-41.

Jellinek, M. (2008) The need for a multi-level biochemical approach to defeat cancer that will

also support the host. Medical Hypotheses, 71(4), 515-526.

Jeong, H., Mason, S.P., Barabasi, A.L., Oltvai, Z.N. (2001) Lethality and centrality in protein networks. Nature. 411(6833), 41-2.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., Barabasi, A.L. (2000) The large-scale organization of metabolic networks. Nature. 407(6804), 651-4.

Jiang, C., Xuan, Z., Zhao, F., Zhang, M.Q. (2007) TRED: a transcriptional regulatory element database, new entries and other development. Nucleic Acids Research, 35(Database issue), D137-140.

Jiang, J.G., Ning, Y.G., Chen, C., Ma, D., Liu, Z.J., Yang, S., Zhou, J., Xiao, X., Zhang, X.A., Edin, M.L., Card, J.W., Wang, J., Zeldin, D.C., Wang, D.W. (2007) Cytochrome p450 epoxygenase promotes human cancer metastasis. Cancer Research, 67(14), 6665-6674.

Jin, S.P., Kim, J.H., Kim, M.A., Yang, H.K., Lee, H.E., Lee, H.S., Kim, W.H. (2007). Prognostic significance of loss of c-fos protein in gastric carcinoma. Pathology and Oncology Research, 13(4), 284-289.

Jobard, E., Pontoizeau, C., Blaise, B.J., Bachelot, T., Elena-Herrmann, B., Tredan, O. (2014). A serum nuclear magnetic resonance-based metabolomic signature of advanced metastatic human breast cancer. Cancer Letters, 343(1), 33-41.

Johnston, J.B., Navaratnam, S., Pitz, M.W., Maniate, J.M., Wiechec, E., Baust, H., Gingerich, J., Skliris, G.P., Murphy, L.C., Los, M. (2006). Targeting the EGFR pathway for cancer therapy. Current Medicinal Chemistry, 13(29), 3483-3492.

Jonsson, G., Staaf, J., Vallon-Christersson, J., Ringner, M., Holm, K., Hegardt, C., Gunnarsson, H., Fagerholm, R., Strand, C., Agnarsson, B.A., Kilpivaara, O., Luts, L., Heikkilä, P., Aittomäki, K., Blomqvist, C., Loman, N., Malmström, P., Olsson, H., Johannsson, O.T., Arason, A., Nevanlinna, H, Barkardottir R.B., Borg, A. (2010). Genomic subtypes of breast cancer identified by array-comparative genomic hybridization display distinct molecular and clinical

characteristics. Breast Cancer Research, 12(3), R42.

Jung, H.M., Choi, S.J., Kim, J.K. (2009) Expression profiles of SV40-immortalization-associated genes upregulated in various human cancers. Journal of Cell Biochemistry, 106(4), 703-713.

Kaghad, M., Bonnet, H., Yang, A., Creancier, L., Biscan, J.C., Valent, A., Minty, A., Chalon, P., Lelias, J.M., Dumont, X., Ferrara, P., McKeon, F., Caput, D. (1997). Monoallelically expressed gene related to p53 at 1p36, a region frequently deleted in neuroblastoma and other human cancers. Cell, 90(4), 809-819.

Kamangar, F., Dores, G.M., Anderson, W. F. (2006) Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. Journal of Clinical Oncology, 24, 2137–2150.

Kamburov, A., Stelzi, U., Herwig, R. (2012) IntScore: a web tool for confidence scoring of biological interactions. Nucleic Acids Research,

Kamei, T.; Matozaki, T.; Takai, Y. (1999) [Mechanisms of cell adhesion and migration]. Gan To Kagaku Ryoho., 26(9), 1359-1366.

Kamiguchi, H. (2003) The mechanism of axon growth: what we have learned from the cell adhesion molecule L1. Molecular Neurobiology, 28(3), 219-228.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., Tanabe, M. (2012) KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Research, 40(Database issue), D109-114.

Kang, Y., Chen, C.R., Massague, J. (2003). A self-enabling TGFbeta response coupled to stress signaling: Smad engages stress response factor ATF3 for Id1 repression in epithelial cells. Molecular Cell, 11(4), 915-926.

Karagoz, K., Sinha, R., Arga, K.Y. (2015) Triple negative breast cancer: a multi-omics network discovery strategy for candidate targets and driving pathways. OMICS, 19(2):115-30.
Karamouzis, M.V., Badra, F.A., Papavassiliou, A.G. (2007). Breast cancer: the upgraded role of HER-3 and HER-4. International Journal of Biochemistry and Cell Biology, 39(5), 851-856.

Kashyap, M.K., Marimuthu, A., Kishore, C.J., Peri, S., Keerthikumar, S., Prasad, T.S, Mahmood, R., Rao, S., Ranganathan, P., Sanjeeviah, R.C., Vijayakumar, M., Kumar, K.V., Pandey, A. (2009) Genomewide mRNA profiling of esophageal squamous cell carcinoma for identification of cancer biomarkers. Cancer Biology and Theraphy, 8(1), 36-46.

Kato, T., Shimono, Y., Hasegawa, M., Jijiwa, M., Enomoto, A., Asai, N., Murakumo, Y., Takahashi, M., Takahashi, M. (2009). Characterization of the HDAC1 complex that regulates the sensitivity of cancer cells to oxidative stress. Cancer Research, 69(8), 3597-3604.

Kazemi-Noureini, S., Colonna-Romano, S., Ziaee, A.A., Malboobi, M.A., Yazdanbod, M., Setayeshgar, P., Maresca, B. (2004) Differential gene expression between squamous cell carcinoma of esophageus and its normal epithelium; altered pattern of mal, akr1c2, and rab11a expression. World Journal of Gastroenterology, 10(12), 1716-1721.

Kazerounian, S., Gerald, D., Huang, M., Chin, Y.R., Udayakumar, D., Zheng, N., O'Donnell, R.K., Perruzzi, C., Mangiante, L., Pourat, J., Phung, T.L., Bravo-Nuevo, A., Shechter, S., McNamara, S., Duhadaway ,J.B., Kocher, O.N., Brown, L.F., Toker, A., Prendergast, G.C., Benjamin, L.E. (2013) RhoB differentially controls Akt function in tumor cells and stromal endothelial cells during breast tumorigenesis. Cancer Research, 73, 50-61.

Ke, L. (2002) Mortality and incidence trends from esophagus cancer in selected geographic areas of China circa 1970-90.International Journal of Cancer, 102(3), 271-274.

Kenneth, K., Wang, M.D. (2012) Benign and Malignant Tumors of the Esophagus. Principles of Deglutition, Springer, 579-598.

Kerkhoff, C., Klempt, M., Kaever, V., Sorg, C. (1999) The two calcium-binding proteins,

S100A8 and S100A9, are involved in the metabolism of arachidonic acid in human neutrophils. Journal of Biological Chemistry, 274(46), 32672-32679.

Kessler, O., Shraga-Heled, N., Lange, T., Gutmann-Raviv, N., Sabo, E., Baruch, L., Machluf, M., Neufeld, G. (2004) Semaphorin-3F is an inhibitor of tumor angiogenesis. Cancer Research, 64, 1008-1015.

Khodarev, N., Ahmad, R., Rajabi, H., Pitroda, S., Kufe, T., McClary, C., Joshi, M.D., MacDermed, D., Weichselbaum, R., Kufe, D. (2010). Cooperativity of the MUC1 oncoprotein and STAT1 pathway in poor prognosis human breast cancer. Oncogene, 29(6), 920-929.

Kim, A.C., Peters, L.L., Knoll, J.H., Van Huffel, C., Ciciotte, S.L., Kleyn, P.W., Chishti, A.H. (1997) Limatin (LIMAB1), an actin-binding LIM protein, maps to mouse chromosome 19 and human chromosome 10q25, a region frequently deleted in human cancers. Genomics., 46(2), 291-293.

Klover, P.J., Muller, W.J., Robinson, G.W., Pfeiffer, R.M., Yamaji, D., Hennighausen, L. (2010). Loss of STAT1 from mouse mammary epithelium results in an increased Neu-induced tumor burden. Neoplasia, 12(11), 899-905.

Knopfova, L., Benes, P., Pekarcikova, L., Hermanova, M., Masarik, M., Pernicova, Z., Souček, K., Smarda, J. (2012) c-Myb regulates matrix metalloproteinases 1/9, and cathepsin D: implications for matrix-dependent breast cancer cell invasion and metastasis. Molecular Cancer, 11, 1-15.

Koenig, A., Mueller, C., Hasel, C., Adler, G., Menke, A. (2006) Collagen type I induces disruption of E-cadherin-mediated cell-cell contacts and promotes proliferation of pancreatic carcinoma cells. Cancer Research, 66(9), 4662-4671.

Komatsu, M., Yoshimaru, T., Matsuo, T., Kiyotani, K., Miyoshi, Y., Tanahashi, T., Rokutan, K., Yamaguchi, R., Saito, A., Imoto, S., Miyano, S., Nakamura, Y., Sasa, M., Shimada, M., Katagiri, T. (2013) Molecular features of triple negative breast cancer cells by genome-wide gene expression profiling analysis. International Journal of Oncology, 42, 478-506.

Koromilas, A.E., Sexl, V. (2013). The tumor suppressor function of STAT1 in breast cancer. JAKSTAT, 2(2), e23353.

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

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Joyce, L., Shuy, P., Datta, N., Tikuisis, A.P., Punna, T., Peregrín-Alvarez, J.M., Shales, M., Zhang, X., Davey, M., Robinson, M.D., Paccanaro, A., Bray, J.E., Sheung, A., Beattie, B., Richards, D.P., Canadien, V., Lalev, A., Mena, F., Wong, P., Starostine, A., Canete, M.M., Vlasblom, J., Wu, S., Orsi, C., Collins, S.R., Chandran, S., Haw, R., Rilstone, J.J., Gandi, K., Thompson, N.J., Musso, G., St Onge, P., Ghanny, S., Lam, M.H.Y., Butland, G., Altaf-Ul, A.,M., Kanaya, S., Shilatifard, A., O'Shea, E., Weissman, J.S., Ingles, C.J., Hughes, T.R., Parkinson, J., Gerstein, M., Wodak, S.J., Emili, A., Greenblatt, J.F. (2006) Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature, 440, 637–643.

Krupp, M., Weinmann, A., Galle, P.R., Teufel, A. (2006) Actin binding LIM protein 3 (abLIM3). International Journal of Molecular Medicine, 17(1), 129-133.

Krusche, C.A., Wulfing, P., Kersting, C., Vloet, A., Bocker, W., Kiesel, L., Beier, H.M., Alfer, J. (2005). Histone deacetylase-1 and -3 protein expression in human breast cancer: a tissue microarray analysis. Breast Cancer Research and Treatment, 90(1), 15-23.

Kuchaiev, O., Rasajski, M., Higham, D.J., Przulj, N. (2009) Geometric de-noising of protein–protein interaction networks. PLoS Computational Biology, 5, 1–10.

Kulawiec, M., Owens, K.M., Singh, K.K. (2009). mtDNA G10398A variant in African-American women with breast cancer provides resistance to apoptosis and promotes metastasis in mice. Journal of Human Genetics, 54(11), 647-654.

Kuo, W.H., Chang, Y.Y., Lai, L.C., Tsai, M.H., Hsiao, C.K., Chang, K.J., Chuang, E Y. (2012) Molecular characteristics and metastasis predictor genes of triple-negative breast cancer: a clinical study of triple-negative breast carcinomas. PLoS One, 7, e45831.

Kuramoto, T., Morimura, K., Yamashita, S., Okochi, E., Watanabe, N., Ohta, T., Ohki, M., Fukushima, S., Sugimura, T., Ushijima, T. (2002). Etiology-specific gene expression profiles in rat mammary carcinomas. Cancer Research, 62(13), 3592-3597.

Lee, G.Y., Haverty, P. M., Li, L., Kljavin, N.M., Bourgon, R., Lee, J., Stern, H, Modrusan, Z, Seshagiri, S., Zhang, Z., Davis, D., Stokoe, D., Settleman, J., de Sauvage, F.J., Neve, R. M. (2014). Comparative oncogenomics identifies PSMB4 and SHMT2 as potential cancer driver genes. Cancer Research, 74(11), 3114-3126.

Lee, J.W., Ryan, F., Swaffield, J.C., Johnston, S.A., Moore, D.D. (1995) Interaction of thyroid-hormone receptor with a conserved transcriptional mediator. Nature, 374(6517), 91-94.

Lee, K.H., Chen, Y.L., Yeh, S.D., Hsiao, M., Lin, J.T., Goan, Y.G., Lu, P.J. (2009) MicroRNA-330 acts as tumor suppressor and induces apoptosis of prostate cancer cells through E2F1-mediated suppression of Akt phosphorylation. Oncogene, 28(38), 3360-3370.

Lee, M.S., Kim, Y.J. (2007) Signaling pathways downstream of pattern-recognition receptors and their cross talk. Annual Review of Biochemistry, 76,447-80.

Lee, S.T., Feng, M., Wei, Y., Li, Z., Qiao, Y., Guan, P., Jiang, X., Wong, C.H., Huynh, K., Wang, J., Li, J., Karuturi, K.M., Tan, E.Y., Hoon, D.S.B., Kang, Y., Yu, Q. (2013) Protein tyrosine phosphatase UBASH3B is overexpressed in triple-negative breast cancer and promotes invasion and metastasis. Proceedings National Academy Sciences USA, 110, 11121-11126.

Lee, U., Frankenberger, C., Yun, J., Bevilacqua, E., Caldas, C., Chin, S.F., Rueda, O.M., Reinitz, J., Rosner, M.R. (2013) A prognostic gene signature for metastasis-free survival of

triple negative breast cancer patients. PLoS One, 8, e82125.,

Lehmann, B.D., Bauer, J.A., Chen, X., Sanders, M.E., Chakravarthy, A.B., Shyr, Y., Pietenpol, J.A. (2011) Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. Journal of Clinical Investigation, 121, 2750-2767.

Leong, S., Nunez, A.C., Lin, M.Z., Crossett, B., Christopherson, R.I., Baxter, R.C. (2012) iTRAQ-based proteomic profiling of breast cancer cell response to doxorubicin and TRAIL. J Proteome Research, 11(7), 3561-3572.

Li, D., Liu, W., Liu, Z., Wang, J., Liu, Q., Zhu, Y., He, F. (2008) PRINCESS, a protein interaction confidence evaluation system with multiple data sources. Molecular and Cellular Proteomics, 7, 1043–1052.

Li, K., He, W., Lin, N., Wang, X., Fan, Q.X. (2010) Downregulation of N-cadherin expression inhibits invasiveness, arrests cell cycle and induces cell apoptosis in esophageal squamous cell carcinoma. Cancer Investigation, 28(5), 479-486.

Li, L.F., Xu, X.J., Zhao, Y., Liu, Z.B., Shen, Z.Z., Jin, W.R., Shao, Z.M., (2009) Integrated gene expression profile predicts prognosis of breast cancer patients. Breast Cancer Research and Treatment, 113, 231-237.

Li, L., Lian, B., Li, C., Li, W., Li, J., Zhang, Y., He, X., Li, Y., Xie, L. (2014) Integrative analysis of transcriptional regulatory network and copy number variation in intrahepatic cholangiocarcinoma. PloS one, 9(6), e98653.

Li, W.Q., Hu, N., Burton, V.H., Yang, H.H., Su, H., Conway, C. M., Wang, L., Wang, C., Ding, T., Xu, Y., Giffen, C., Abnet, C.C., Goldstein, A.M., Hewitt, S.M., Taylor, P.R. (2014) PLCE1 mRNA and protein expression and survival of patients with esophageal squamous cell carcinoma and gastric adenocarcinoma. Cancer Epidemiology, Biomarkers and Prevention, 23(8), 1579-1588.

Li, X.L., Xu, J.H., Nie, J.H., Fan, S,J. (2012) Exogenous expression of UHRF1 promotes proliferation and metastasis of breast cancer cells. Oncology Reports, 28, 375-383.

Li, Y.H., Xue, T.Y., He, Y. Z., Du, J.W. (2013). Novel oncoprotein EPS8: a new target for anticancer therapy. Future Oncology, 9(10), 1587-1594.

Li, Z., Qiu, H.Y., Jiao, Y., Cen, J.N., Fu, C.M., Hu, S.Y., Zhu, M.Q., Qi, X.F. (2013) Growth and differentiation effects of Homer3 on a leukemia cell line. Asian Pacfic Journal of Cancer Prevention, 14(4), 2525-2528.

Lim, W.K., Lyashenko, E., Califano, A. (2009). Master regulators used as breast cancer metastasis classifier. Pacific Symposium on Biocomputing, 504-515.

Lin, C.Y., Chin, C.H., Wu, H.H., Chen, S.H., Ho, C.W., Ko, M.T. (2008) Hubba: hub objects analyzer--a framework of interactome hubs identification for network biology. Nucleic Acids Research, 36(Web Server issue), W438-443.

Lin, D.C., Du, X.L., Wang, M.R. (2009) Protein alterations in ESCC and clinical implications: a review. Disases of Esophagus., 22(1), 9-20.

Lin, J., Raoof, D.A., Wang, Z., Lin, M.Y., Thomas, D.G., Greenson, J.K., Giordano, T.J., Orringer, M.B., Chang, A.C., Beer, D.G., Lin, L. (2006) Expression and effect of inhibition of the ubiquitin-conjugating enzyme E2C on esophageal adenocarcinoma. Neoplasia, 8(12), 1062-1071.

Lin, N., Wu, B., Jansen, R., Gerstein, M., Zhao, H. (2004) Information assessment on predicting protein–protein interactions. BMC Bioinformatics, 5, 154.

Ling, Z.Q., Mukaisho, K., Yamamoto, H., Chen, K.H., Asano, S., Araki, Y., Sugihara, H., Mao, W.M., Hattori, T. (2010) Initiation of malignancy by duodenal contents reflux and the role of ezrin in developing esophageal squamous cell carcinoma. Cancer Science, 101(3), 624-630.

Liu, B., Yee, K.M., Tahk, S., Mackie, R., Hsu, C., Shuai, K. (2014). PIAS1 SUMO ligase regulates the self-renewal and differentiation of hematopoietic stem cells. The EMBO Journal, 33(2), 101-113.

Liu, B.P., Strittmatter, S.M. (2001) Semaphorin-mediated axonal guidance via Rho-related G proteins. Current Opinion in Cell Biology, 13, 619-626.

Liu, C., Yao, J., Mercola, D., Adamson, E. (2000). The transcription factor EGR-1 directly transactivates the fibronectin gene and enhances attachment of human glioblastoma cell line U251. Journal of Biological Chemistry, 275(27), 20315-20323.

Liu, D., He, J., Yuan, Z., Wang, S., Peng, R., Shi, Y., Teng, X., Qin, T. (2012) EGFR expression correlates with decreased disease-free survival in triple-negative breast cancer: a retrospective analysis based on a tissue microarray. Medical Oncology, 29, 401-405.

Liu, F., Verin, A.D., Wang, P., Day, R., Wersto, R.P., Chrest, F.J., English, D.K., Garcia, J.G.. (2001) Differential regulation of sphingosine-1-phosphate- and VEGF-induced endothelial cell chemotaxis. Involvement of G(ialpha2)-linked Rho kinase activity. American Journal of Respiratory Cell Molecular Biology, 24, 711-719.

Liu, H., Radisky, D.C., Yang, D., Xu, R., Radisky, E.S., Bissell, M.J., Bishop, J.M. (2012) MYC suppresses cancer metastasis by direct transcriptional silencing of alphav and beta3 integrin subunits. Nature Cell Biology, 14, 567-574.

Liu, L.Y., Chang, L.Y., Kuo, W.H., Hwa, H.L., Lin, Y.S., Huang, S.F., Chen, C.N., Chang, K.J., Hsieh, F.J. (2012) Major Functional Transcriptome of an Inferred Center Regulator of an ER(-) Breast Cancer Model System. Cancer Informatics, 11, 87-111.

Liu, T.L., Shimada, H., Ochiai, T., Shiratori, T., Lin, S.E., Kitagawa, M., Harigaya, K., Maki, M., Oka, M., Abe, T., Takiguchi, M., Hiwasa, T. (2006) Enhancement of chemosensitivity toward peplomycin by calpastatin-stabilized NF-kappaB p65 in esophageal carcinoma cells:

possible involvement of Fas/Fas-L synergism. Apoptosis, 11(6), 1025-1037.

Liu, Z.P., Wang, Y., Zhang, X.S., Chen, L. (2012) Network-based analysis of complex diseases. IET Systems Biology. 6(1), 22-33.

Lu, S.H., Zhou, Y., Que, H.P., Liu, S.J. (2003) Peptidergic innervation of human esophageal and cardiac carcinoma. World Journal of Gastroenterology, 9(3), 399-403.

Luo, A., Kong, J., Hu, G., Liew, C.C., Xiong, M., Wang, X., Ji, J., Wang, T., Zhi, H., Wu, M., Liu, Z. (2004) Discovery of Ca2+-relevant and differentiation-associated genes downregulated in esophageal squamous cell carcinoma using cDNA microarray. Oncogene, 23(6), 1291-1299.

Ma, S., Bao, J.Y., Kwan, P.S., Chan, Y.P., Tong, C. M., Fu, L., Zhang, N., Tong, A.H., Quin, Y.R., Tsao, S.W., Chan, K.W., Lok, S., Guan, X.Y. (2012) Identification of PTK6, via RNA sequencing analysis, as a suppressor of esophageal squamous cell carcinoma. Gastroenterology, 143(3), 675-686 e671-612.

Maere, S., Heymans, K., Kuiper, M. (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics, 21(16), 3448-3449.

Malet-Martino, M., Holzgrabe, U. (2011) NMR techniques in biomedical and pharmaceutical analysis. Journal of Pharmaceutical and Biomedical Analysis. 55(1), 1-15.

Mamas, M., Dunn, W.B., Neyses, L., Goodacre, R. (2011). The role of metabolites and metabolomics in clinically applicable biomarkers of disease. Archives of Toxicology, 85(1), 5-17.

Mamede, A.C., Tavares, S.D., Abrantes, A.M., Trindade, J., Maia, J.M., Botelho, M.F. (2011) The role of vitamins in cancer: a review. Nutrition and Cancer., 63(4), 479-494.

Mardinoglu, A., Gatto, F., Nielsen, J. (2013) Genome-scale modeling of human metabolism - a

systems biology approach. Biotechnology Journal, 8(9), 985-996.

Massague, J. (2012). TGF-beta signaling in development and disease. FEBS Letters, 586(14), 1833.

Matoskova, B., Wong, W.T., Salcini, A.E., Pelicci, P.G., Di Fiore, P.P. (1995). Constitutive phosphorylation of eps8 in tumor cell lines: relevance to malignant transformation. Molecular and Cellular Biology, 15(7), 3805-3812.

Matsen, C.B., Neumayer, L.A. (2013). Breast cancer: a review for the general surgeon. JAMA Surgery, 148(10), 971-979.

Mazieres, J., Antonia,.T., Daste, G., Muro-Cacho, C., Berchery, D., Tillement, V., Pradines, A., Sebti, S., Favre, G. (2004) Loss of RhoB expression in human lung cancer progression. Clinical Cancer Research, 10, 2742-2750.

McLaughlin, S.K., Olsen, S.N., Dake, B., De Raedt, T., Lim, E., Bronson, R.T., Beroukhim, R., Polyak, K., Brown, M., Kuperwasser, C., Cichowski, K. (2013). The RasGAP gene, RASAL2, is a tumor and metastasis suppressor. Cancer Cell, 24(3), 365-378.

Medale-Giamarchi, C., Lajoie-Mazenc, I., Malissein, E., Meunier, E., Couderc, B., Berge, Y., Filleron, T., Keller, L., Marty, C., Lacroix-Triki, M., Dalenc, F., Doisneau-Sixou, S.F., Favre, G. (2013) RhoB modifies estrogen responses in breast cancer cells by influencing expression of the estrogen receptor. Breast Cancer Research, 15, R6.

Meding, S., Balluff, B., Elsner, M., Schone, C., Rauser, S., Nitsche, U., Maak, M., Schäfer, A., Hauck, S.M., Ueffing, M., Langer, R., Höfler, H., Friess, H., Rosenberg, R., Walch, A. (2012) Tissue-based proteomics reveals FXYD3, S100A11 and GSTM3 as novel markers for regional lymph node metastasis in colon cancer. Journal of Pathology, 228, 459-470.

Mehra, R., Varambally, S., Ding, L., Shen, R., Sabel, M.S., Ghosh, D., Chinnaiyan, A.M., Kleer, C.G. (2005) Identification of GATA3 as a breast cancer prognostic marker by global gene expression meta-analysis. Cancer Research, 65, 11259-11264.

Menendez, J.A., Lupu, R. (2007). Transphosphorylation of kinase-dead HER3 and breast cancer progression: a new standpoint or an old concept revisited? Breast Cancer Research, 9(5), 111.

Meng, F., Zhang, H., Liu, G., Kreike, B., Chen, W., Sethi, S., Miller, F.R., Wu, G. (2011). p38gamma mitogen-activated protein kinase contributes to oncogenic properties maintenance and resistance to poly (ADP-ribose)-polymerase-1 inhibition in breast cancer. Neoplasia, 13(5), 472-482.

Meyer, N., Penn, L.Z. (2008) Reflecting on 25 years with MYC. Nature Reviews Cancer, 8, 976-990.

Mi, H., Dong, Q., Muruganujan, A., Gaudet, P., Lewis, S., Thomas, P.D. (2010) PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium. Nucleic Acids Research, 38(Database issue), D204-210.

Milde-Langosch, K., Roder, H., Andritzky, B., Aslan, B., Hemminger, G., Brinkmann, A., Bamberger, C.M., Löning, T., Bamberger, A. M. (2004). The role of the AP-1 transcription factors c-Fos, FosB, Fra-1 and Fra-2 in the invasion process of mammary carcinomas. Breast Cancer Research and Treatment, 86(2), 139-152.

Milke, L., Schulz, K., Weigert, A., Sha, W., Schmid, T., Brune, B. (2013). Depletion of tristetraprolin in breast cancer cells increases interleukin-16 expression and promotes tumor infiltration with monocytes/macrophages. Carcinogenesis, 34(4), 850-857.

Miller-Jensen, K., Janes, K.A., Brugge, J.S., Lauffenburger, D.A. (2007) Common effector processing mediates cell-specific responses to stimuli. Nature. 448(7153), 604-8.

Mocellin, S., Rossi, C.R., Pilati, P., Nitti, D. (2005) Tumor necrosis factor, cancer and anticancer therapy. Cytokine Growth Factor Reviews, 16(1), 35-53.

Moll, R., Franke, W.W., Schiller, D.L., Geiger, B., Krepler, R. (1982) The catalog of human

cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. Cell, 31, 11-24.

Morel, N., Bon, S., Greenblatt, H.M., Van Belle, D., Wodak, S.J., Sussman, J.L., Sussman, J.L., Massoulié, J., Silman, I. (1999). Effect of mutations within the peripheral anionic site on the stability of acetylcholinesterase. Molecular Pharmacology, 55(6), 982-992.

Moreno, J.J. (2009). New aspects of the role of hydroxyeicosatetraenoic acids in cell growth and cancer development. Biochemical Pharmacology, 77(1), 1-10.

Morita, M., Kumashiro, R., Kubo, N., Nakashima, Y., Yoshida, R., Yoshinaga, K., Saeki, H., Emi, Y., Kakeji, Y., Sakaguchi, Y., Toh, Y., Maehara, Y. (2010) Alcohol drinking, cigarette smoking, and the development of squamous cell carcinoma of the esophagus: epidemiology, clinical findings, and prevention. International Journal of Clinical Oncology, 15(2), 126-134.

Morris, D.R., Biros, E., Cronin, O., Kuivaniemi, H., Golledge, J. (2014) The association of genetic variants of matrix metalloproteinases with abdominal aortic aneurysm: a systematic review and meta-analysis. Heart, 100(4), 295-302.

Mougeot, J.L., Bahrani-Mougeot, F.K., Lockhart, P.B., Brennan, M.T. (2011) Microarray analyses of oral punch biopsies from acute myeloid leukemia (AML) patients treated with chemotherapy. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology 112(4), 446-452.

Mrowka, R., Patzak, A., Herzel, H. (2001) Is there a bias in proteome research? Genome Research, 11, pp. v1971–v1973.

Mudunuri, U., Che, A., Yi, M., Stephens, R.M., (2009) bioDBnet: the biological database network. Bioinformatics, 25, 555-556.

Mueller, H., Flury, N., Eppenberger-Castori, S., Kueng, W., David, F., Eppenberger, U. (2000). Potential prognostic value of mitogen-activated protein kinase activity for disease-free survival

of primary breast cancer patients. International Journal of Cancer, 89(4), 384-388.

Muller, B.M., Jana, L., Kasajima, A., Lehmann, A., Prinzler, J., Budczies, J., Winzer, K.J., Dietel, M., Weichert, W., Denkert, C. (2013). Differential expression of histone deacetylases HDAC1, 2 and 3 in human breast cancer--overexpression of HDAC2 and HDAC3 is associated with clinicopathological indicators of disease progression. BMC Cancer, 13, 215.

Mulligan, A.M., Raitman, I., Feeley, L., Pinnaduwage, D., Nguyen, L.T., O'Malley, F.P., Ohashi, P.S., Andrulis, I.L. (2013) Tumoral lymphocytic infiltration and expression of the chemokine CXCL10 in breast cancers from the Ontario Familial Breast Cancer Registry. Clinical Cancer Research, 19, 336-346.

Munoz-Pinedo, C., El Mjiyad, N., Ricci, J.E. (2012). Cancer metabolism: current perspectives and future directions. Cell Death and Disease, 3, e248.

Myatt, S.S., Lam, E.W. (2007) The emerging roles of forkhead box (Fox) proteins in cancer. Nature Reviews Cancer, 7(11), 847-859.

Naderi, A., Teschendorff, A.E., Barbosa-Morais, N.L., Pinder, S.E., Green, A.R., Powe, D. G., Robertson, J.F., Aparicio, S., Ellis, I.O., Brenton, J.D., Caldas, C. (2007). A gene-expression signature to predict survival in breast cancer across independent data sets. Oncogene, 26(10), 1507-1516.

Nakamura, K., Hongo, A., Kodama, J., Hiramatsu, Y. (2011). The role of hepatocyte growth factor activator inhibitor (HAI)-1 and HAI-2 in endometrial cancer. International Journal of Cancer, 128(11), 2613-2624.

Nakamura, T., Ota, M., Sato, T., Yamamoto, M. (2011) [Pre-treatment tumor staging of esophageal cancer: difficulty for accurate diagnosis]. Nihon Rinsho, 69 Suppl 6, 129-132.

Nakshatri, H., Badve, S. (2009) FOXA1 in breast cancer. Expert Reviews in Molecular Medicine, 11, e8.

Nanes, M.S. (2003) Tumor necrosis factor-alpha: molecular and cellular mechanisms in skeletal pathology. Gene, 321, 1-15.

Nebert, D.W., Dalton, T.P. (2006) The role of cytochrome P450 enzymes in endogenous signalling pathways and environmental carcinogenesis. Nature Reviews Cancer, 6(12), 947-960.

Nemunaitis, J., Senzer, N., Khalil, I., Shen, Y., Kumar, P., Tong, A., Kuhn, J., Lamont, J., Nemunaitis, M., Rao, D., Zhang, Y.A., Zhou, Y., Vorhies, J., Maples, P., Hill C., Shanahan, D. (2007) Proof concept for clinical justification of network mapping for personalized cancer therapeutics. Cancer Gene Therapy, 14(8), 686-95.

Neph, S., Stergachis, A.B., Reynolds, A., Sandstrom, R., Borenstein, E., Stamatoyannopoulos, J.A. (2012) Circuitry and dynamics of human transcription factor regulatory networks. Cell. 150(6), 1274-86.

Nithipatikom, K., Brody, D.M., Tang, A.T., Manthati, V.L., Falck, J.R., Williams, C.L., Campbell, W.B. (2010) Inhibition of carcinoma cell motility by epoxyeicosatrienoic acid (EET) antagonists. Cancer Science, 101(12), 2629-2636.

Novak, U., Grob, T.J., Baskaynak, G., Peters, U.R., Aebi, S., Zwahlen, D., Tschan, M.P., Kreuzer, K.A., Leibundgut, E.O., Cajot, J.F., Tobler, A., Fey, M. F. (2001). Overexpression of the p73 gene is a novel finding in high-risk B-cell chronic lymphocytic leukemia. Annals of Oncology, 12(7), 981-986.

Olumi, A.F., Grossfeld, G.D., Hayward, S.W., Carroll, P.R., Tlsty, T.D., Cunha, G.R. (1999) Carcinoma-associated fibroblasts direct tumor progression of initiated human prostatic epithelium. Cancer Research, 59(19), 5002-5011.

Ota, M., Shimada, H., Masui, H., Tanaka, K., Yamaguchi, S., Ichikawa, Y., Togo, S., Ike, H., Oki, S. (2004) Adjuvant hepatic arterial infusion chemotherapy after curative resection for

Dukes C colorectal cancer: a pilot study. Hepatogastroenterology, 51, 124-127.

Ota, T., Maeda, M., Suto, S., Tatsuka, M. (2004) LyGDI functions in cancer metastasis by anchoring Rho proteins to the cell membrane. Molecular Carcinogenesis, 39, 206-220.

Palmer, D.K., O'Day, K., Trong, H.L., Charbonneau, H., Margolis, R.L. (1991) Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. Proceedings National Academy of Sciences U S A., 88(9), 3734-3738.

Palmieri, C., Monteverde, M., Lattanzio, L., Gojis, O., Rudraraju, B., Fortunato, M., Syed, N., Thompson, A., Garrone, O., Merlano, M., Lo Nigro, C., Crook, T. (2012). Site-specific CpG methylation in the CCAAT/enhancer binding protein delta (CEBPdelta) CpG island in breast cancer is associated with metastatic relapse. British Journal of Cancer, 107(4), 732-738.

Pan, W.W., Li, J.D., Huang, S., Papadimos, T.J., Pan, Z.K., Chen, L.Y. (2010) Synergistic activation of NF-{kappa}B by bacterial chemoattractant and TNF{alpha} is mediated by p38 MAPK-dependent RelA acetylation. Journal of Biological Chemistry, 285, 34348-34354.

Panigrahy, D., Kaipainen, A., Greene, E.R., Huang, S. (2010) Cytochrome P450-derived eicosanoids: the neglected pathway in cancer. Cancer Metastasis Reviews, 29(4), 723-735.

Pantel, K., Cote, R.J., Fodstad, O. (1999) Detection and clinical importance of micrometastatic disease. Journal of the National Cancer Institute, 91, 1113-1124.

Park, M., Penick, E.C., Edwards, J.G., Kauer, J.A., Ehlers, M.D. (2004) Recycling endosomes supply AMPA receptors for LTP. Science, 305(5692), 1972-1975.

Park, S.Y., Jun, J.A., Jeong, K.J., Heo, H.J., Sohn, J.S., Lee, H.Y., Park, C.G., Kang, J. (2011). Histone deacetylases 1, 6 and 8 are critical for invasion in breast cancer. Oncology Reports, 25(6), 1677-1681.

Park, Y.Y., Jung, S.Y., Jennings, N.B., Rodriguez-Aguayo, C., Peng, G., Lee, S.R., Kim, S.B., Kim, K., Leem, S.H., Lin, S.Y., Lopez-Berestein, G., Sood, A.K., Lee, J.S. (2012) FOXM1

mediates Dox resistance in breast cancer by enhancing DNA repair. Carcinogenesis, 33(10), 1843-1853.

Parrish, J.R., Yu, J., Liu, G., Hines, J.A., Chan, J.E., Mangiola, B.A., Zhang, H., Pacifico, S., Fotouhi, F., DiRita, V.J., Ideker, T., Andrews Jr., P., Finley, R.L. (2007) A proteome-wide protein interaction map for Campylobacter jejuni. Genome Biology, 8, R130.

Patil, A., Nakai, K., Nakamura, H. (2011) HitPredict: a database of quality assessed protein–protein interactions in nine species. Nucleic Acids Research, 39, pp. 744–749.

Patil, A., Nakamura, H. (2005) Filtering high-throughput protein–protein interaction data using a combination of genomic features. BMC Bioinformatics, 6, 1–13.

Patil, K. R., Nielsen, J. (2005) Uncovering transcriptional regulation of metabolism by using metabolic network topology. Proceedings of National Academy of Sciences U S A., 102(8), 2685-2689.

Patnala, R., Lee, S.H., Dahlstrom, J.E., Ohms, S., Chen, L., Dheen, S.T., Rangasamy, D. (2014). Inhibition of LINE-1 retrotransposon-encoded reverse transcriptase modulates the expression of cell differentiation genes in breast cancer cells. Breast Cancer Research and Treatment, 143(2), 239-253.

Perou, C.M., Parker, J.S., Prat, A., Ellis, M.J., Bernard, P.S. (2010) Clinical implementation of the intrinsic subtypes of breast cancer. Lancet Oncology, 11, 718-719; 720-711.

Pitteri, S.J., Amon, L.M., Busald Buson, T., Zhang, Y., Johnson, M.M., Chin, A., Kennedy, J., Wong, C.H., Zhang, Q., Wang, H., Lampe, P.D., Prentice, R.L., McIntosh, M.W., Hanash, S.M., Li, C. I. (2010). Detection of elevated plasma levels of epidermal growth factor receptor before breast cancer diagnosis among hormone therapy users. Cancer Research, 70(21), 8598-8606.

Planche, A., Bacac, M., Provero, P., Fusco, C., Delorenzi, M., Stehle, J.C., Stamenkovic, I.

(2011). Identification of prognostic molecular features in the reactive stroma of human breast and prostate cancer. PLoS One, 6(5), e18640.

Poliakov, E., Managadze, D., Rogozin, I.B. (2014) Generalized portrait of cancer metabolic pathways inferred from a list of genes overexpressed in cancer. Genetics Research International, 646193.

Prat, A., Adamo, B., Cheang, M.C., Anders, C.K., Carey, L.A., Perou, C.M. (2013) Molecular characterization of basal-like and non-basal-like triple-negative breast cancer. Oncologist, 18, 123-133.

Prendergast, G.C. (2001) Actin' up: RhoB in cancer and apoptosis. Nature Reviews Cancer 1, 162-168.

Prokhortchouk, E., Defossez, P.A. (2008). The cell biology of DNA methylation in mammals. Biochimica Biophysica Acta, 1783(11), 2167-2173.

Putignani, L., Raffa, S., Pescosolido, R., Aimati, L., Signore, F., Torrisi, M.R., Grammatico, P. (2008). Alteration of expression levels of the oxidative phosphorylation system (OXPHOS) in breast cancer cell mitochondria. Breast Cancer Research and Treatment, 110(3), 439-452.

Qi, X.F., Kim, D.H., Yoon, Y.S., Jin, D., Huang, X.Z., Li, J.H., Deung, Y.K., Lee, K.J. (2009) Essential involvement of cross-talk between IFN-gamma and TNF-alpha in CXCL10 production in human THP-1 monocytes. Journal of Cellular Physiology, 220, 690-697.

Qi, Y., Klein-Seetharaman, J., Bar-Joseph, Z. (2005) Random forest similarity for protein–protein interaction prediction from multiple sources. Pacific Symposium on Biocomputing, 531–542.

Qi, Y.J., Chao, W.X., Chiu, J.F. (2012) An overview of esophageal squamous cell carcinoma proteomics. Journal of Proteomics, 75(11), 3129-3137.

Qiu, Y., Zhou, B., Su, M., Baxter, S., Zheng, X., Zhao, X., Yen, Y., Jia, W. (2013). Mass spectrometry-based quantitative metabolomics revealed a distinct lipid profile in breast cancer patients. International Journal of Molecular Sciences, 14(4), 8047-8061.

Rachidi, M., Lopes, C., Charron, G., Delezoide, A.L., Paly, E., Bloch, B., Delabar, J.M. (2005) Spatial and temporal localization during embryonic and fetal human development of the transcription factor SIM2 in brain regions altered in Down syndrome. International Journal of Developmental Neuroscience, 23(5), 475-484.

Rain, J.C., Selig, L., De Reuse, H., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J., Schachter, V., Chemama, Y., Labigne, A., Legrain, P., (2001) The protein–protein interaction map of Helicobacter pylori. Nature, 409, 211–215.

Ramaswamy, S., Ross, K.N., Lander, E.S., Golub, T.R. (2003) A molecular signature of metastasis in primary solid tumors. Nature Genetics, 33(1), 49-54.

Ramji, D.P., Foka, P. (2002). CCAAT/enhancer-binding proteins: structure, function and regulation. Biochemical Journal, 365(Pt 3), 561-575.

Rao, M.V., Mohan, P.S., Peterhoff, C.M., Yang, D.S., Schmidt, S.D., Stavrides, P.H., Campbell, J., Chen, Y., Jiang, Y., Paskevich, P.A., Cataldo, A.M., Haroutunian, V., Nixon, R.A. (2008) Marked calpastatin (CAST) depletion in Alzheimer's disease accelerates cytoskeleton disruption and neurodegeneration: neuroprotection by CAST overexpression. Journal of Neuroscience, 28(47), 12241-12254.

Rao, V.S., Srinivas, K., Sujini, G.N., Kumar, G.N. (2014) Protein-protein interaction detection: methods and analysis. International Journal of Proteomics. 2014:147648.

Rastelli, F., Biancanelli, S., Falzetta, A., Martignetti, A., Casi, C., Bascioni, R., Giustini, L., Crispino, S. et al.(2010) Triple-negative breast cancer: current state of the art. Tumori, 96,875–88.

Rawnaq, T., Kleinhans, H., Uto, M., Schurr, P.G., Reichelt, U., Cataldegirmen, G., Gawad, K.A., Yekebas, E.F., Schachner, M., Izbicki, J.R., Kaifi, J.T. (2009) Subset of esophageal adenocarcinoma expresses adhesion molecule l1 in contrast to squamous cell carcinoma. Anticancer Research, 29(4), 1195-1199.

Raynal, P., Pollard, H.B. (1994). Annexins: the problem of assessing the biological role for a gene family of multifunctional calcium- and phospholipid-binding proteins. Biochimica Biophysica Acta, 1197(1), 63-93.

Razick, S., Magklaras, G., Donaldson, I.M. (2008) iRefIndex: a consolidated protein interaction database with provenance. BMC Bioinformics, 9, 405.

Reddy, K.B., Nabha, S.M., Atanaskova, N. (2003). Role of MAP kinase in tumor progression and invasion. Cancer Metastasis Reviews, 22(4), 395-403.

Reddy, S.K., Rape, M., Margansky, W.A., Kirschner, M.W. (2007) Ubiquitination by the anaphase-promoting complex drives spindle checkpoint inactivation. Nature, 446(7138), 921-925.

Redmond, K.L., Crawford, N.T., Farmer, H., D'Costa, Z.C., O'Brien, G.J., Buckley, N.E., Kennedy, R.D., Johnston, P.G., Harkin, D.P., Mullan, P. B. (2010). T-box 2 represses NDRG1 through an EGR1-dependent mechanism to drive the proliferation of breast cancer cells. Oncogene, 29(22), 3252-3262.

Reguly, R., Breitkreutz, A., Boucher, L., Breitkreutz, B.J., Hon, G.C., Myers, C.L., Parsons, A., Friesen, H., Oughtred, R., Tong, A., Stark, C., Ho, Y., Botstein, D., Andrews, B., Boone, C., Troyanskya, O.G., Ideker, Dolinski, K., Batada, N.N., Tyers, M. (2006) Comprehensive curation and analysis of global interaction networks in Saccharomyces cerevisiae. Journal of Biology, 5, p. 1.

Resler, A.J., Malone, K.E., Johnson, L.G., Malkki, M., Petersdorf, E.W., McKnight, B., Madeleine, M.M. (2013). Genetic variation in TLR or NFkappaB pathways and the risk of breast cancer: a case-control study. BMC Cancer, 13, 219.

Rhee, D.K., Park, S.H., Jang, Y.K. (2008). Molecular signatures associated with transformation and progression to breast cancer in the isogenic MCF10 model. Genomics, 92(6), 419-428.

Riley, T., Sontag, E., Chen, P., Levine, A. (2008). Transcriptional control of human p53-regulated genes. Nature Reviews Molecular Cell Biology, 9(5), 402-412.

Rimkus, C., Friederichs, J., Rosenberg, R., Doll, D., Siewert, J.R., Holzmann, B., Janssen, K. P. (2006) Prognostic significance of downregulated expression of the candidate tumour suppressor gene SASH1 in colon cancer. British Journal of Cancer., 95(10), 1419-1423.

Rivera Rosado, L.A., Rodriguez-Canales, J., Zhang, B. (2011) Association of D4-GDI expression with breast cancer progression. Cancer Biomarkers, 10, 163-173.

Roberti, M.P., Arriaga, J.M., Bianchini, M., Quinta, H.R., Bravo, A.I., Levy, E.M., Mordoh, J., Barrio, M.M. (2012) Protein expression changes during human triple negative breast cancer cell line progression to lymph node metastasis in a xenografted model in nude mice. Cancer Biology and Therapy, 13, 1123-1140.

Roll, J.D., Rivenbark, A.G., Sandhu, R., Parker, J.S., Jones, W.D., Carey, L.A., Livasy, C.A., Coleman, W.B. (2013) Dysregulation of the epigenome in triple-negative breast cancers: basal-like and claudin-low breast cancers express aberrant DNA hypermethylation. Experimental and Molecular Pathology, 95, 276-287.

Roof, D.J., Hayes, A., Adamian, M., Chishti, A.H., Li, T. (1997) Molecular characterization of abLIM, a novel actin-binding and double zinc finger protein. Journal of Cell Biology, 138(3), 575-588.

Rosenthal, D.T., Iyer, H., Escudero, S., Bao, L., Wu, Z., Ventura, A.C., Kleer, C.G., Arruda, E.M., Garikipati, K., Merajver, S. D. (2011). p38gamma promotes breast cancer cell motility and metastasis through regulation of RhoC GTPase, cytoskeletal architecture, and a novel leading edge behavior. Cancer Research, 71(20), 6338-6349.

Rual, J.F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G.F., Gibbons, F.D., Dreze, M., Ayivi-Guedehoussou, N., Klitgord, N., Simon, C., Boxem, M., Milstein, S., Rosenberg, J., Goldberg, D.S., Zhang, L.V., Wong, S.L., Franklin, G., Li, S., Albala, J.S., Lim, J., Fraughton, C., Llamosas, E., Cevik, S., Bex, C., Lamesch, P., Sikorski, R.S., Vandenhaute, J., Zoghbi, H.Y., Smolyar, A., Bosak, S., Sequerra, R., Doucette-Stamm, L., Cusick, M.E., Hill, D.E., Roth F.P., Vidal, M. (2005) Towards a proteome-scale map of the human protein-protein interaction network. Nature, 437(7062), 1173-8.

Saito, R., Suzuki, H., Hayashizaki, Y. (2003) Construction of reliable protein–protein interaction networks with a new interaction generality measure. Bioinformatics, 19, 756–763.

Sakakibara, T., Hibi, K., Koike, M., Fujiwara, M., Kodera, Y., Ito, K., Nakao, A. (2006) PAI-1 expression levels in esophageal and colorectal cancers are closely correlated to those in corresponding normal tissues. Anticancer Research, 26(6B), 4343-4347.

Sakamoto, N., Oue, N., Noguchi, T., Sentani, K., Anami, K., Sanada, Y., Yoshida, K., Yasui, W. (2010) Serial analysis of gene expression of esophageal squamous cell carcinoma: ADAMTS16 is upregulated in esophageal squamous cell carcinoma. Cancer Science, 101(4), 1038-1044.

Salwinski, L., Miller, C.S., Smith, A.J., Pettit, F.K., Bowie, J.U., Eisenberg, D. (2004) The database of interacting proteins: 2004 update. Nucleic Acids Research, 32, 449–451.

Sambourg, L., Mieg, N.T.(2010) New insights into protein–protein interaction data lead to increased estimates of the S. cerevisiae interactome size. BMC Bioinformatics, 11, 605–615.

Sancak, Y., Bar-Peled, L., Zoncu, R., Markhard, A.L., Nada, S., Sabatini, D.M. (2010) Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. Cell, 141(2), 290-303.

Sanchez-Pena, M.L., Isaza, C.E., Perez-Morales, J., Rodriguez-Padilla, C., Castro, J. M., Cabrera-Rios, M. (2013). Identification of potential biomarkers from microarray experiments using multiple criteria optimization. Cancer Medicine, 2(2), 253-265.

Sanz-Pamplona, R., Garcia-Garcia, J., Franco, S., Messeguer, X., Driouch, K., Oliva, B., Sierra, A. (2012). A taxonomy of organ-specific breast cancer metastases based on a protein-protein interaction network. Molecular BioSystems, 8(8), 2085-2096.

Sastre-Serra, J., Valle, A., Company, M.M., Garau, I., Oliver, J., Roca, P. (2010). Estrogen down-regulates uncoupling proteins and increases oxidative stress in breast cancer. Free Radical Biology and Medicine, 48(4), 506-512.

Sato, T., Iizuka, N., Hamamoto, Y., Yoshino, S., Abe, T., Takeda, S., Uchimure, S., Miyamoto, T., Sei, F., Hamada, K., Yamada-Okabe, H., Oka, M. (2006) Esophageal squamous cell carcinomas with distinct invasive depth show different gene expression profiles associated with lymph node metastasis. International Journal of Oncology, 28(5), 1043-1055.

Saxena, R., Dwivedi, A. (2012) ErbB family receptor inhibitors as therapeutic agents in breast cancer: current status and future clinical perspective. Medicinal Research Reviews, 32, 166-215.

Schaaij-Visser, T.B., Brakenhoff, R.H., Leemans, C.R., Heck, A.J., Slijper, M. (2010) Protein biomarker discovery for head and neck cancer. Journal of Proteomics., 73(10), 1790-1803.

Schaefer, M.H., Fontaine, J.F., Vinayagam, A., Porras, P., Wanker, E.E., Andrade-Navarro, M.A. (2012) HIPPIE: integrating protein interaction networks with experiment based quality scores. PLoS ONE, 7, p. 2.

Scherle, P., Behrens, T., Staudt, L.M. (1993) Ly-GDI, a GDP-dissociation inhibitor of the RhoA GTP-binding protein, is expressed preferentially in lymphocytes. Proceedings National Academy Sciences USA, 90, 7568-7572.

Schwartz, A.S., Yu, J., Gardenour, K.R., Finley, R.L., Jr., Ideker, T. (2009) Cost-effective strategies for completing the interactome. Nature Methods. 6(1), 55-61.

Scott, M.S., Barton, G.J. (2007) Probabilistic prediction and ranking of human protein–protein interactions. BMC Bioinformatics, 8 (239).

Seguin, C.M., Wang, L.S., Stoner, G.D. (2011) Chemopreventive Effects of Berries and Berry Components in the Rodent Esophagus. Berries and Cancer Prevention, 143-161.

Seo, H., Kim, W., Lee, J., Youn, B. (2013) Network-based approaches for anticancer therapy (Review). International Journal of Oncology. 43(6), 1737-44.

Sevimoglu, T., Arga, K.Y. (2014) The role of protein interaction networks in system medicine. Computional and Structural Biotechnology Journal, 11, 22-24.

Sharan, R., Suthram, S., Kelley, R.M., Kuhn, T., McCuine, S., Uetz, P., Sittler, T., Karp, R.M., Ideker, T. (2005) Conserved patterns of protein interaction in multiple species. Proceedings of the National Academy of Science, 102, 1974–1979.

Shen, D., Nooraie, F., Elshimali, Y., Lonsberry, V., He, J., Bose, S., Chia, D., Seligson, D., Chang, H.R., Goodglick, L. (2006). Decreased expression of annexin A1 is correlated with breast cancer development and progression as determined by a tissue microarray analysis. Human Pathology, 37(12), 1583-1591.

Shen, Y., Tantai, J., Zhao, H. (2014) Ranking candidate genes of esophageal squamous cell carcinomas based on differentially expressed genes and the topological properties of the co-expression network. European Journal of Medical Research, 19:52.

Shiraishi, Y., Mizutani, A., Yuasa, S., Mikoshiba, K., Furuichi, T. (2004) Differential expression of Homer family proteins in the developing mouse brain. Journal of Comparative Neurology, 473(4), 582-599.

Shoemaker, B.A., Panchenko, A.R. (2007) Deciphering protein–protein interactions. Part I. Experimental techniques and databases. PLoS Computational Biology, 3, p. 3.

Shureiqi, I., Lippman, S.M. (2001) Lipoxygenase modulation to reverse carcinogenesis. Cancer Research, 61(17), 6307-6312.

Siegel, R., Naishadham, D., Jemal, A. (2013) Cancer statistics, 2013. CA Cancer Journal for Clinicians, 63, 11-30.

Siegel, R.L., Miller, K.D., Jemal, A. (2015) Cancer statistics, 2015. CA: A Cancer Journal for Clinicians. 65(1), 5-29.

Simpson, P.T., Reis-Filho, J.S., Lambros, M.B., Jones, C., Steele, D., Mackay, A., Iravani, M., Fenwick, K., Dexter, T., Jones, A., Reid, L., Da Silva, L., Shin, S.J., Hardisson, D., Ashworth, A., Schmitt, F.C., Palacios, J., Lakhani, S. R. (2008). Molecular profiling pleomorphic lobular carcinomas of the breast: evidence for a common molecular genetic pathway with classic lobular carcinomas. Journal of Pathology, 215(3), 231-244.

Singh,K.K., Ayyasamy, V., Owens, K.M., Koul, M.S., Vujcic, M. (2009) Mutations in mitochondrial DNA polymerase-gamma promote breast tumorigenesis. Journal of Human Genetics, 54(9), 516-24.

Sivaraman, V.S., Wang, H., Nuovo, G.J., Malbon, C.C. (1997). Hyperexpression of mitogen-

activated protein kinase in human breast cancer. Journal of Clinical Investigation, 99(7), 1478-1483.

Smeds, J., Berggren, P., Ma, X., Xu, Z., Hemminki, K., Kumar, R. (2002) Genetic status of cell cycle regulators in squamous cell carcinoma of the oesophagus: the CDKN2A (p16(INK4a) and p14(ARF)) and p53 genes are major targets for inactivation. Carcinogenesis, 23(4), 645-655.

Smith, A.L., Iwanaga, R., Drasin, D.J., Micalizzi, D.S., Vartuli, R.L., Tan, A.C., Ford, H.L. (2012). The miR-106b-25 cluster targets Smad7, activates TGF-beta signaling, and induces EMT and tumor initiating cell characteristics downstream of Six1 in human breast cancer. Oncogene, 31(50), 5162-5171.

Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P.L., Ideker, T. (2011) Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics, 27, 431-432.

Smyth G.K. (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Statistical Applications in Genetics and Molecular Biology, 3 (article 3).
Smyth, G.K. (2005) Limma: linear models for microarray data in Bioinformatics and computational biology solutions using R and Bioconductor. Springer, 397-420.

Sobral-Leite, M., Wesseling, J., Smit, V. T., Nevanlinna, H., van Miltenburg, M.H., Sanders, J., Hofland, I., Blows, F.M., Coulson, P., Patrycja, G., Schellens, J.H., Fagerholm, R., Heikkilä, P., Aittomäki, K., Blomqvist, C., Provenzano, E., Ali, H.R., Figueroa, J., Sherman, M., Lissowska, J., Mannermaa, A., Kataja, V., Kosma, V.M., Hartikainen, J.M., Phillips, K.A., kConFab/AOCS Investigators, Couch, F.J., Olson, J.E., Vachon, C., Visscher, D., Brenner, H., Butterbach, K., Arndt, V., Holleczek, B., Hooning, M.J., Hollestelle, A., Martens, J.W., van Deurzen, C.H.,van de Water, B., Broeks, A., Chang-Claude, J., Chenevix-Trench, G., Easton, D.F., Pharoah, P.D., García-Closas, M., de Graauw, M., Schmidt, M.K. (2015). Annexin A1 expression in a pooled breast cancer series: association with tumor subtypes and prognosis. BMC Medicine, 13, 156.

Solaini, G., Sgarbi, G., Baracca, A. (2011). Oxidative phosphorylation in cancer cells. Biochimica Biophysica Acta, 1807(6), 534-542.

Sommers, C.L., Park, C.S., Lee, J., Feng, C., Fuller, C.L., Grinberg, A., Hildebrand, J.A., Lacaná, E., Menon, R.K., Shores, E.W., Samelson, L.E., Love, P.E. (2002). A LAT mutation that inhibits T cell development yet induces lymphoproliferation. Science, 296(5575), 2040-2043.

Somsel Rodman, J., Wandinger-Ness, A. (2000) Rab GTPases coordinate endocytosis. Journal of Cell Science, 113 Pt 2, 183-192.

Soond, S. M., Townsend, P. A., Barry, S. P., Knight, R. A., Latchman, D. S., Stephanou, A. (2008). ERK and the F-box protein betaTRCP target STAT1 for degradation. Journal of Biological Chemistry, 283(23), 16077-16083.

Sorlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Thorsen, T., Quist, H., Matese, J.C., Brown, P.O., Botstein, D., Lønning, P.E., Børresen-Dale, A.L. (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings National Academy Sciences USA, 98, 10869-10874.

Soussi, T., Lozano, G. (2005). p53 mutation heterogeneity in cancer. Biochemical and Biophysical Research Communications, 331(3), 834-842.

Sparano, J.A., Goldstein, L.J., Childs, B.H., Shak, S., Brassard, D., Badve, S., Baehner, F.L., Bugarini, R., Rowley, S., Perez, E.A., Shulman, L.N., Martino, S., Davidson, N.E., Kenny, P.A., Sledge, G.W. Jr., Gray R. (2011) Relationship between quantitative GRB7 RNA expression and recurrence after adjuvant anthracycline chemotherapy in triple-negative breast cancer. Clinical Cancer Research, 17, 7194-7203.

Spector, A.A., Norris, A.W. (2007) Action of epoxyeicosatrienoic acids on cellular function. American Journal Physiology Cell Physiology, 292(3), C996-1012.

Sprinzak, E., Sattath, S., Margalit, H. (2003) How reliable are experimental protein–protein interaction data? Journal of Molecular Biology, 327, 919–923.

Staiger, C., Cadot, S., Kooter, R., Dittrich, M., Muller, T., Klau, G.W., Wessels, L.F. (2012). A critical evaluation of network and pathway-based classifiers for outcome prediction in breast cancer. PLoS One, 7(4), e34796.

Stark, C., Breitkreutz, B.J., Chatr-aryamontri, A., Boucher, L., Oughther, R., Liystone, M.S., Nixon, J., van Auken, K., Wang, X., Shi, X., Reguly, T., Rust, J.M., Winter, A., Dolinski, K., Tyers, M. (2011) The BioGRID interaction database: 2011 update. Nucleic Acids Research, 39, 698–704.

Stark, G.R., Darnell, J.E., Jr. (2012). The JAK-STAT pathway at twenty. Immunity, 36(4), 503-514.

Stathopoulou, A., Vlachonikolis, I., Mavroudis, D., Perraki, M., Kouroussis, C., Apostolaki, S., Malamos, N., Kakolyris, S., Kotsakis, A., Xenidis, N., Reppa, D., Georgoulias, V. (2002) Molecular detection of cytokeratin-19-positive cells in the peripheral blood of patients with operable breast cancer: evaluation of their prognostic significance. Journal of Clinical Oncology, 20, 3404-3412.

Stein, R.A., Gaillard, S., McDonnell, D.P. (2009). Estrogen-related receptor alpha induces the expression of vascular endothelial growth factor in breast cancer cells. The Journal of Steroid Biochemistry and Molecular Biology, 114(1-2), 106-112.

Stiewe, T., Tuve, S., Peter, M., Tannapfel, A., Elmaagacli, A. H., Putzer, B. M. (2004). Quantitative TP73 transcript analysis in hepatocellular carcinomas. Clinical Cancer Research, 10(2), 626-633.

Su, H., Hu, N., Yang, H.H., Wang, C., Takikita, M., Wang, Q.H., Giffen, C., Clifford, R., Hewitt, S.M., Shou, J.Z., Goldstein, A.M., Lee, M.P., Taylor, P.R. (2011) Global gene

expression profiling and validation in esophageal squamous cell carcinoma and its association with clinical phenotypes. Clinical Cancer Research, 17(9), 2955-2966.

Sugimoto, T., Arai, M., Shimada, H., Hata, A., Seki, N. (2007) Integrated analysis of expression and genome alteration reveals putative amplified target genes in esophageal cancer. Oncology Reports, 18(2), 465-472.

Swa, H.L., Shaik, A.A., Lim, L.H., Gunaratne, J. (2015). Mass spectrometry based quantitative proteomics and integrative network analysis accentuates modulating roles of annexin-1 in mammary tumorigenesis. Proteomics, 15(2-3), 408-418.

Swanson, H.I., Njar, V.C., Yu, Z., Castro, D.J., Gonzalez, F.J., Williams, D.E., Huang, Y., Kong, A.N., Doloff, J.C., Ma, J., Waxman, D.J., Scott, E.E. (2010) Targeting drug-metabolizing enzymes for effective chemoprevention and chemotherapy. Drug Metabolism adn Disposition, 38(4), 539-544.

Swinnen, J.V., Brusselmans, K., Verhoeven, G. (2006). Increased lipogenesis in cancer cells: new players, novel targets. Current Opinion in Clinical Nutrition and Metabolic Care, 9(4), 358-365.

Szabo, R., Hobson, J.P., Christoph, K., Kosa, P., List, K., Bugge, T. H. (2009). Regulation of cell surface protease matriptase by HAI2 is essential for placental development, neural tube closure and embryonic survival in mice. Development, 136(15), 2653-2663.

Szarc vel Szic, K., Op de Beeck, K., Ratman, D., Wouters, A., Beck, I.M., Declerck, K. Heyninck, K., Fransen, E., Bracke, M., De Bosscher, K., Lardon, F., Van Camp, G., Vanden Berghe, W. (2014) Pharmacological levels of Withaferin A (Withania somnifera) trigger clinically relevant anticancer effects specific to triple negative breast cancer cells. PLoS One, 9, e87850.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P., Jensen, L.J., von Mering, C. (2011) The STRING database in

2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Research, 39, 561–568.

Taherian-Fard, A., Srihari, S., Ragan, M.A. (2015). Breast cancer classification: linking molecular mechanisms to disease prognosis. Briefings in Bioinformics, 16(3), 461-474.

Tahk, S., Liu, B., Chernishof, V., Wong, K.A., Wu, H., Shuai, K. (2007). Control of specificity and magnitude of NF-kappa B and STAT1-mediated gene activation through PIASy and PIAS1 cooperation. Proceedings of the National Academy of Sciences U S A, 104(28), 11643-11648.

Tallett, A., Chilvers, E.R., Hannah, S., Dransfield, I., Lawson, M.F., Haslett, C., Sethi, T. (1996) Inhibition of neuropeptide-stimulated tyrosine phosphorylation and tyrosine kinase activity stimulates apoptosis in small cell lung cancer cells. Cancer Research, 56(18), 4255-4263.

Tan, A.R., Alexe, G., Reiss, M. (2009). Transforming growth factor-beta signaling: emerging stem cell target in metastatic breast cancer? Breast Cancer Research and Treatment, 115(3), 453-495.

Tang, D., Sivko, G.S., DeWille, J.W. (2006). Promoter methylation reduces C/EBPdelta (CEBPD) gene expression in the SUM-52PE human breast cancer cell line and in primary breast tumors. Breast Cancer Research and Treatment, 95(2), 161-170.

Tao, Y., Chai, D., Ma, L., Zhang, T., Feng, Z., Cheng, Z., Wu, S., Quin, Y., Lai, M. (2012) Identification of distinct gene expression profiles between esophageal squamous cell carcinoma and adjacent normal epithelial tissues. The Tohoku Journal of Experimantal Medicine, 226(4), 301-311.

Tapper, J., Kettunen, E., El-Rifai, W., Seppala, M., Andersson, L.C., Knuutila, S. (2001) Changes in gene expression during progression of ovarian carcinoma. Cancer Genetics and Cytogenetics, 128, 1-6.

Tarnow, P., Tralau, T., Hunecke, D., Luch, A. (2013) Effects of Effects of triclocarban on the

transcription of estrogen, androgen and aryl hydrocarbon receptor responsive genes in human breast cancer cells. Toxicology in Vitro, 27, 1467-1475.

Tate, C.R., Rhodes, L.V., Segar, H.C., Driver, J.L., Pounder, F.N., Burow, M.E., Collins-Burow, B.M. (2012) Targeting triple-negative breast cancer cells with the histone deacetylase inhibitor panobinostat. Breast Cancer Research, 14, R79.

Taylor, I.W., Linding, R., Warde-Farley, D., Liu, Y., Pesquita, C., Faria, D., Bull, S., Pawson, T, Morris, Q., Wrana, J.L. (2009). Dynamic modularity in protein interaction networks predicts breast cancer outcome. Nature Biotechnology, 27(2), 199-204.

Tchou, J., Kossenkov, A.V., Chang, L., Satija, C., Herlyn, M., Showe, L.C., Ellen, P. (2012) Human breast cancer associated fibroblasts exhibit subtype specific gene expression profiles. BMC Medical Genomics, 5, 39.

Teh, M.T., Wong, S.T., Neill, G.W., Ghali, L.R., Philpott, M.P., Quinn, A.G. (2002) FOXM1 is a downstream target of Gli1 in basal cell carcinomas. Cancer Research, 62(16), 4773-4780.

Tekir, S.D., Arga, K.Y., Ulgen, K.O. (2009) Drug targets for tumorigenesis: insights from structural analysis of EGFR signaling network. Journal of Biomedical Informatics, 42, 228–236.

Tennant, D.A., Duran, R.V., Gottlieb, E. (2010). Targeting metabolic transformation for cancer therapy. Nature Reviews Cancer, 10(4), 267-277.

Terunuma, A., Putluri, N., Mishra, P., Mathe, E.A., Dorsey, T.H., Yi, M., Wallace, T.A., Issaq, H.J., Zhou, M., Killian, J.K., Stevenson, H.S., Karoly, E.D., Chan, K., Samanta, S., Prieto, D., Hsu, T.Y., Kurley, S.J., Putluri, V., Sonavane, R., Edelman, D.C., Wulff, J., Starks, A.M., Yang, Y., Kittles, R.A., Yfantis, H.G., Lee, D.H., Ioffe, O.B., Schiff, R., Stephens, R.M., Meltzer, P.S., Veenstra, T.D., Westbrook, T.F., Sreekumar, A., Ambs, S. (2014). MYC-driven accumulation of 2-hydroxyglutarate is associated with breast cancer prognosis. J Clin Invest, 124(1), 398-412.

Tezel, G.G., Uner, A., Yildiz, I., Guler, G., Takahashi, M. (2009). RET finger protein

expression in invasive breast carcinoma: relationship between RFP and ErbB2 expression. Pathology – Research and Practice, 205(6), 403-408.

Thakkar, A.D., Raj, H., Chakrabarti, D., Ravishankar Saravanan, N., Muthuvelan, B., Balakrishnan, A., Padigaru, M. (2010) Identification of gene expression signature in estrogen receptor positive breast carcinoma. Biomarkers in Cancer, 2, 1-15.

Thangaraju, M., Rudelius, M., Bierie, B., Raffeld, M., Sharan, S., Hennighausen, L., Huang A.M., Sterneck, E. (2005). C/EBPdelta is a crucial regulator of pro-apoptotic gene expression during mammary gland involution. Development, 132(21), 4675-4685.

Thompson, M.R., Xu, D., Williams, B.R. (2009). ATF3 transcription factor and its emerging roles in immunity and cancer. Journal of Molecular Medicine (Berlin), 87(11), 1053-1060.
Thorner, A.R., Parker, J.S., Hoadley, K.A., Perou C.M. (2010) Potential tumor suppressor role for the c-Myb oncogene in luminal breast cancer. PLoS One, 5, e13073.
Tian, S., Krueger, J.G., Li, K., Jabbari, A., Brodmerkel, C., Lowes, M.A., Suarez-Farinas, M. (2012) Meta-analysis derived (MAD) transcriptome of psoriasis defines the "core" pathogenesis of disease. PLoS One, 7(9), e44274.

Titz, B., Rajagopala, S.V., Goll, J., Hauser, R., McKevitt, M.T., Palzkill, T., Uetz, P. (2008) The binary protein interactome of Treponema pallidum – the syphilis spirochete. PLoS ONE, 3 p.

Tommiska, J., Bartkova, J., Heinonen, M., Hautala, L., Kilpivaara, O., Eerola, H., Aittomäki, K., Hofstetter, B., Lukas, J., von Smitten, K., Blomqvist, C., Ristimäki, A., Heikkilä P., Bartek, J., Nevanlinna, H. (2008) The DNA damage signalling kinase ATM is aberrantly reduced or lost in BRCA1/BRCA2-deficient and ER/PR/ERBB2-triple-negative breast cancer. Oncogene, 27, 2501-2506.

Tomonaga, T., Matsushita, K., Yamaguchi, S., Oohashi, T., Shimada, H., Ochiai, T., Yoda, K., Nomura, F. (2003) Overexpression and mistargeting of centromere protein-A in human primary colorectal cancer. Cancer Research, 63(13), 3511-3516.

Tong, M., Chan, K.W., Bao, J.Y., Wong, K.Y., Chen, J.N., Kwan, P.S., Tang, K.H., Fu, L., Qin, Y.R., Lok, S., Guan, X.Y., Ma, S. (2012) Rab25 is a tumor suppressor gene with antiangiogenic and anti-invasive activities in esophageal squamous cell carcinoma. Cancer Research, 72(22), 6024-6035.

Torigoe, T., Asanuma, H., Nakazawa, E., Tamura, Y., Hirohashi, Y., Yamamoto, E., Kanaseki, T., Hasegawa, T., Sato, N. (2012). Establishment of a monoclonal anti-pan HLA class I antibody suitable for immunostaining of formalin-fixed tissue: unusually high frequency of down-regulation in breast cancer tissues. Pathology International, 62(5), 303-308.

Tripathi, A., King, C., de la Morenas, A., Perry, V.K., Burke, B., Antoine, G.A., Hirsch, E.F., Kavanah, M., Mendez, J., Stone, M., Gerry, N.P., Lenburg, M.E., Rosenberg, C. L. (2008). Gene expression abnormalities in histologically normal breast epithelium of breast cancer patients. International Journal of Cancer, 122(7), 1557-1566.

Tsai, C.H., Muo, C.H., Tzeng, H.E., Tang, C.H., Hsu, H.C., Sung, F.C. (2013). Fracture in asian women with breast cancer occurs at younger age. PLoS One, 8(9), e75109.

Tsunoda, S., Okumura, T., Ito, T., Mori, Y., Soma, T., Watanabe, G., Kaganoi, J., Itami, A., Sakai, Y., Shimada, Y. (2006) Significance of nerve growth factor overexpression and its autocrine loop in oesophageal squamous cell carcinoma. British Journal of Cancer., 95(3), 322-330.

Turner, N.C., Reis-Filho, J.S. (2013) Tackling the diversity of triple-negative breast cancer. Clinical Cancer Research, 19, 6380-6388.

Uchikado, Y., Inoue, H., Haraguchi, N., Mimori, K., Natsugoe, S., Okumura, H., Aikou, T., Mori, M. (2006) Gene expression profiling of lymph node metastasis by oligomicroarray analysis using laser microdissection in esophageal squamous cell carcinoma. International Journal of Oncology, 29(6), 1337-1347.

Ulisse, S., Baldini, E., Sorrenti, S., D'Armiento, M. (2009) The urokinase plasminogen activator

system: a target for anti-cancer therapy. Current Cancer Drug Targets., 9(1), 32-71.

UniProt C. (2014) Activities at the Universal Protein Resource (UniProt). Nucleic Acids Research, 42, D191-D198.

Vaclavicek, A., Bermejo, J.L., Schmutzler, R.K., Sutter, C., Wappenschmidt B., Meindl A., Kiechle, M., Arnold, N., Weber, B.H., Niederacher, D., Burwinkel, B., Bartram, C.R.,Hemminki, K., Försti, A. (2007) Polymorphisms in the Janus kinase 2 (JAK)/signal transducer and activator of transcription (STAT) genes: putative association of the STAT gene region with familial breast cancer. Endocrine Related Cancer, 14, 267-277.

van Diest, P.J., van der Groep, P., van der Wall, E. (2006). EGFR expression predicts BRCA1 status in patients with breast cancer. Clinical Cancer Research, 12(2), 670; author reply 671.

van 't Veer, L.J., Dai, H., van de Vijver, M.J., He, Y.D., Hart, A.A., Mao, M., Peterse, H.L., van der Kooy, K., Marton, M.J., Witteveen, A.T., Schreiber, G.J., Kerkhoven, R.M., Roberts, C., Linsley, P.S., Bernards, R., Friend, S. H. (2002). Gene expression profiling predicts clinical outcome of breast cancer. Nature, 415(6871), 530-536.

Vander Heiden, M.G. (2013). Exploiting tumor metabolism: challenges for clinical translation. The Journal of Clinical Investigation, 123(9), 3648-3651.

Venkatesan, K., Rual, J.F., Vazquez, A., Stelzl, U., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Zenkner, M., Xin, X., Goh, K.I., Yildirim, M.A., Simonis, N., Heinzmann, K., Gebreab, F., Sahalie, J.M., Cevik, S., Simon, C., de Smet, A.S., Dann, E., Smolyar, A., Vinayagam, A., Yu, H., Szeto, D., Borick, H., Dricot, A., Klitgord, N., Murray, R.R., Lin, C., Lalowski, M., Timm, J., Rau, K., Boone, C., Braun, P., Cusick, M.E., Roth, F.P., Hill, D.E., Tavernier, J., Wanker, E.E., Barabasi A.L., Vidal, M. (2009) An empirical framework for binary interactome mapping. Nature Methods, 6(1), 83-90.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng, X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor Miklos, G.L., Nelson C, Broder, S., Clark, A.G., Nadeau, J., McKusick, V.A., Zinder, N., Levine, A.J., Roberts, R.J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A.E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T.J., Higgins, M.E., Ji, R.R., Ke, Z., Ketchum, K.A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G.V., Milshina, N., Moore, H.M., Naik, A.K., Narayan, V.A., Neelam, B., Nusskern, D., Rusch, D.B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M.L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y.H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N.N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J.F., Guigó, R., Campbell, M.J., Sjolander, K.V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y.H., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C.,

Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D.,Wu, M., Xia, A., Zandieh, A., Zhu, X. (2001). The sequence of the human genome. Science, 291(5507), 1304-1351.

Vermeersch, K.A., Styczynski, M.P. (2013). Applications of metabolomics in cancer research. Journal of Carcinogenesis, 12, 9.

Vermeulen, E., Zamora-Ros, R., Duell, E.J., Lujan-Barroso, L., Boeing, H., Aleksandrova, K., Bueno-de-Mesquita, H.B., Scalbert, A., Romieu, I., Fedirko, V., Touillaud, M., Fagherazzi, G., Perquier, F., Molina-Montes, E., Chirlaque, M.D., Vicente Argüelles, M., Amiano, P., Barricarte, A., Pala, V., Mattiello, A., Saieva, C., Tumino, R., Ricceri, F., Trichopoulou, A., Vasilopoulou, E., Ziara, G., Crowe, F.L., Khaw, K.T., Wareham, N.J., Lukanova, A., Grote, V.A., Tjønneland, A., Halkjær, J., Bredsdorff, L., Overvad, K., Siersema, P.D., Peeters, P.H., May, A.M., Weiderpass, E., Skeie, G., Hjartåker, A., Landberg, R., Johansson, I., Sonestedt, E., Ericson, U., Riboli, E., Gonzalez, C.A. (2013) Dietary flavonoid intake and esophageal cancer risk in the European prospective investigation into cancer and nutrition cohort. American Journal of Epidemiology, 178(4), 570-581.

Vinson, C., Myakishev, M., Acharya, A., Mir, A.A., Moll, J.R., Bonovich, M. (2002). Classification of human B-ZIP proteins based on dimerization properties. Molecular and Cellular Biology, 22(18), 6321-6335.

Voduc, D., Cheang, M., Nielsen, T. (2008) GATA-3 expression in breast cancer has a strong association with estrogen receptor but lacks independent prognostic value. Cancer Epidemiology Biomarkers and Prevention, 17, 365-373.

Voilley, N., Bassilana, F., Mignon, C., Merscher, S., Mattei, M.G., Carle, G.F., Lazdunski, M., Barbry, P. (1995) Cloning, chromosomal localization, and physical linkage of the beta and gamma subunits (SCNN1B and SCNN1G) of the human epithelial amiloride-sensitive sodium channel. Genomics, 28(3), 560-565.

von Lintig, F.C., Dreilinger, A.D., Varki, N.M., Wallace, A.M., Casteel, D.E., Boss, G.R. (2000). Ras activation in human breast cancer. Breast Cancer Research and Treatment, 62(1), 51-62.

von Mering, C., Jensen, L.J., Snel, B., Hooper, S.D., Krupp, M., Foglierini, M., Jouffre, N., Huynen, M.A., Bork, P. (2005) STRING: known and predicted protein–protein associations, integrated and transferred across organisms. Nucleic Acids Research, 33, 433–437.

von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S.G., Fields, S., Bork, P. (2002) Comparative assessment of large-scale datasets of protein–protein interactions. Nature, 417, 399–403.

Voss, M. J., Entschladen, F. (2010) Tumor interactions with soluble factors and the nervous system. Cell Communication and Signaling, 8, 21.

Wagner, A. (2003) How the global structure of protein interaction networks evolves. Proceedings Biological sciences / The Royal Society, 270(1514), 457-66.

Wagner, K.W., Sapinoso, L.M., El-Rifai, W., Frierson, H.F., Butz, N., Mestan, J., Hofmann, F., Deveraux, Q.L., Hampton, G.M. (2004) Overexpression, genomic amplification and therapeutic potential of inhibiting the UbcH10 ubiquitin conjugase in human carcinomas of diverse anatomic origin. Oncogene, 23(39), 6621-6629.

Walerych, D., Napoli, M., Collavin, L., Del Sal, G. (2012). The rebel angel: mutant p53 as the driving oncogene in breast cancer. Carcinogenesis, 33(11), 2007-2017.

Walker, S.R., Nelson, E.A., Zou, L., Chaudhury, M., Signoretti, S., Richardson, A., Frank, D.A. (2009) Reciprocal effects of STAT5 and STAT3 in breast cancer. Molecular Cancer Research, 7, 966-976.

Walker, S.R., Xiang, M., Frank, D.A. (2014) Distinct roles of STAT3 and STAT5 in the pathogenesis and targeted therapy of breast cancer. Molecular Cell Endocrinology, 382, 616-621.

Wang, M.T., Honn, K.V., Nie, D. (2007) Cyclooxygenases, prostanoids, and tumor progression. Cancer Metastasis Reviews, 26(3-4), 525-534.

Wang, Q., Li, W., Zhang, Y., Yuan, X., Xu, K., Yu, J., Beroukhim, R., Wang, H., Lupien, M., Wu, T., Regan, M.M., Meyer, C.A., Carroll, J.S., Manrai, A.K., Jänne, O.A., Balk, S.P., Mehra, R., Han, B., Chinnaiyan, A.M., Rubin, M.A., True, L., Fiorentino, M., Fiore, C., Loda, M., Kantoff, P.W., Liu, X.S., Brown, M. (2009) Androgen receptor regulates a distinct transcription program in androgen-independent prostate cancer. Cell, 138(2), 245-256.

Wang, X., Yuan, S., Wang, J., Lin, P., Liu, G., Lu, Y., Zhang, J., Wang, W., Wei, Y. (2006). Anticancer activity of litchi fruit pericarp extract against human breast cancer in vitro and in vivo. Toxicology and Applied Pharmacology, 215(2), 168-178.

Wang, Y., Kavran, J.M., Chen, Z., Karukurichi, K. R., Leahy, D. J., Cole, P. A. (2014). Regulation of S-adenosylhomocysteine hydrolase by lysine acetylation. The Journal of Biological Chemistry, 289(45), 31361-31372.

Wang, Y., Li, H., Diao, Y., Li, H., Zhang, Y., Yin, C., Cui, Y., Ma, Q., Fang, X., Zhou, Y., Yang, Y. (2008) Relationship between oxidized LDL antibodies and different stages of esophageal carcinoma. Archives of Medical Research, 39(8), 760-767.

Wang, Y., Tang, Z., Xue, R., Singh, G.K., Shi, K., Lv, Y., Yang, L. (2011) Combined effects of TNF-alpha, IL-1beta, and HIF-1alpha on MMP-2 production in ACL fibroblasts under mechanical stretch: an in vitro study. Journal of Orthopaedic Research, 29, 1008-1014.

Wenzel, K., Daskalow, K., Herse, F., Seitz, S., Zacharias, U., Schenk, J.A., Schulz, H., Hubner, N., Micheel, B., Schlag, P.M., Osterziel, K.J., Ozcelik, C., Scherneck, S., Jandrig, B. (2007). Expression of the protein phosphatase 1 inhibitor KEPI is downregulated in breast cancer cell lines and tissues and involved in the regulation of the tumor suppressor EGR1 via the MEK-

ERK pathway. Biological Chemistry, 388(5), 489-495.

Westermarck, J., Hahn, W.C. (2008). Multiple pathways regulated by the tumor suppressor PP2A in transformation. Trends in Molecular Medicine, 14(4), 152-160.

Widmer, C., Gebauer, J.M., Brunstein, E., Rosenbaum, S., Zaucke, F., Drogemuller, C., Leeb, T., Baumann, U. (2012) Molecular basis for the action of the collagen-specific chaperone Hsp47/SERPINH1 and its structure-specific client recognition. Proceedings Natural Academy of Sciences U S A., 109(33), 13243-13247.

Wierstra, I., Alves, J. (2007) FOXM1, a typical proliferation-associated transcription factor. Biological Chemistry, 388(12), 1257-1274.

Williams, G.H., Stoeber, K. (2012) The cell cycle and cancer. Journal of Pathology, 226(2), 352-364.

Winckler, B. (2004) Scientiae forum / models and speculations pathways for axonal targeting of membrane proteins. Biology of the Cell., 96(9), 669-674.

Wolford, C.C., McConoughey, S.J., Jalgaonkar, S.P., Leon, M., Merchant, A.S., Dominick, J.L., Yin, X., Chang, Y., Zmuda, E.J., O'Toole, S.A., Millar, E.K.A., Roller, S.L., Shapiro, C.L., Ostrowski, M.C., Sutherland, R.S., Hai, T. (2013). Transcription factor ATF3 links host adaptive response to breast cancer metastasis. The Journal of Clinical Investigation, 123(7), 2893-2906.

Wong, D.J., Liu, H., Ridky, T.W., Cassarino, D., Segal, E., Chang, H.Y. (2008) Module map of stem cell genes guides creation if epithelial cancer stam cells. Cell Stem Cell, 2, 333-344.

Wong, F.H., Huang, C.Y., Su, L.J., Wu, Y.C., Lin, Y.S., Hsia, J.Y., Tsai, H.T., Lee, S.A., Lin, C.H., Tzeng, C.H., Chen, P.M., Chen, J., Liang, S.C., Lai, J.M., Yen, C.C. (2009) Combination of microarray profiling and protein-protein interaction databases delineates the minimal discriminators as a metastasis network for esophageal squamous cell carcinoma. International Journal of Oncology, 34(1), 117-128.

Wong, G.S., Lee, J.S., Park, Y.Y., Klein-Szanto, A.J., Waldron, T.J., Cukierman, E., Herlyn, M., Gimotty, P., Nakagawa, H., Rustgi, A.K. (2013) Periostin cooperates with mutant p53 to mediate invasion through the induction of STAT1 signaling in the esophageal tumor microenvironment. Oncogenesis, 2, e59.

Wu, J.; Xia, H. H.; Tu, S. P.; Fan, D. M.; Lin, M. C.; Kung, H. F.; Lam, S.K.; Wong, B. C. (2003) 15-Lipoxygenase-1 mediates cyclooxygenase-2 inhibitor-induced apoptosis in gastric cancer. Carcinogenesis, 24(2), 243-247.

Wu, M. Y., Fu, J., Xiao, X., Wu, J., Wu, R. C. (2014). MiR-34a regulates therapy resistance by targeting HDAC1 and HDAC7 in breast cancer. Cancer Letters, 354(2), 311-319.

Wu, M., Chang, S.C., Kampman, E., Yang, J., Wang, X.S., Gu, X.P., Han, R.Q., Liu, A.M., Wallar, G., Zhou, J.Y., Kok, F.J., Zhao, J.K., Zhang, Z.F. (2013) Single nucleotide polymorphisms of ADH1B, ADH1C and ALDH2 genes and esophageal cancer: a population-based case-control study in China. International Journal of Cancer, 132(8), 1868-1877.

Xiang, R., Davalos, A.R., Hensel, C.H., Zhou, X.J., Tse, C., Naylor, S.L. (2002) Semaphorin 3F gene from human 3p21.3 suppresses tumor formation in nude mice. Cancer Research, 62, 2637-2643.

Xiang, Y. J., Guo, M. M., Zhou, C. J., Liu, L., Han, B., Kong, L. Y., Gao, Z.C., Ma, Z.B., Wang, L., Feng, M., Chen, H.Y., Jia, G.T., Gao, D.Z., Zhang, Q., Li, L., Li, Y.Y., Yu, Z. G. (2014). Absence of gamma-interferon-inducible lysosomal thiol reductase (GILT) is associated with poor disease-free survival in breast cancer patients. PLoS One, 9(10), e109449.

Xie, W., Wang, Y., Huang, Y., Yang, H., Wang, J., Hu, Z. (2009). Toll-like receptor 2 mediates invasion via activating NF-kappaB in MDA-MB-231 breast cancer cells. Biochemical and Biophysical Research Communications, 379(4), 1027-1032.

Xu, X.C., Shappell, S.B., Liang, Z., Song, S., Menter, D., Subbarayan, V., Iyengar, S., Tang, D.G., Lippman, S.M. (2003) Reduced 15S-lipoxygenase-2 expression in esophageal cancer specimens and cells and upregulation in vitro by the cyclooxygenase-2 inhibitor, NS398. Neoplasia, 5(2), 121-127.

Yamanishi, Y., Vert, J.P., Kanehisa, M. (2004) Protein network inference from multiple genomic data: a supervised approach. Bioinformatics, 20, 363–370.

Yan, C., Boyd, D.D. (2006). ATF3 regulates the stability of p53: a link to cancer. Cell Cycle, 5(9), 926-929.

Yan, W., Shih, J.H., Rodriguez-Canales, J., Tangrea, M.A., Ylaya, K., Hipp, J., Player, A., Hu, N., Goldstein, A.M., Taylor, P.R., Emmert-Buck, M.R., Erickson, H.S. (2012) Identification of unique expression signatures and therapeutic targets in esophageal squamous cell carcinoma. BMC Research Notes, 5, 73.

Yancy, H.F., Mason, J.A., Peters, S., Thompson, C.E., 3rd, Littleton, G.K., Jett, M., Day, A. A. (2007). Metastatic progression and gene expression between breast cancer cell lines from African American and Caucasian women. Journal of Carcinogenesis, 6, 8.

Yang, A., McKeon, F. (2000). P63 and P73: P53 mimics, menaces and more. Nature Reviews Molecular Cell Biology, 1(3), 199-207.

Yang, C.X., Matsuo, K., Wang, Z.M., Tajima, K. (2005) Phase I/II enzyme gene polymorphisms and esophageal cancer risk: a meta-analysis of the literature. World Journal of Gastroenterology, 11(17), 2531-2538.

Yang, H., Fang, J., Jia, X., Han, C., Chen, X., Yang, C.S., Li, N. (2011) Chemopreventive effects of early-stage and late-stage supplementation of vitamin E and selenium on esophageal carcinogenesis in rats maintained on a low vitamin E/selenium diet. Carcinogenesis, 32(3), 381-388.

Yang, H., Zhou, H., Feng, P., Zhou, X., Wen, H., Xie, X., Shen, H., Zhu, X. (2010). Reduced expression of Toll-like receptor 4 inhibits human breast cancer cells proliferation and inflammatory cytokines secretion. Journal of Experimental and Clinical Cancer Research, 29, 92.

Yang, L., Wu, X., Wang, Y., Zhang, K., Wu, J., Yuan, Y.C., Deng, X., Chen, L., Kim, C.C., Lau, S., Somlo, G., Yen, Y. (2011) FZD7 has a critical role in cell proliferation in triple negative breast cancer. Oncogene, 30, 4437-4446.

Yao, J., Weremowicz, S., Feng, B., Gentleman, R.C., Marks, J.R., Gelman, R., Brennan, C., Polyak, K. (2006). Combined cDNA array comparative genomic hybridization and serial analysis of gene expression analysis of breast tumor progression. Cancer Research, 66(8), 4065-4078.

Yarden, Y., Sliwkowski, M.X. (2001) Untangling the ErbB signalling network. Nature Reviews Molecular Cell Biology, 2, 127-137.

Ye, B., Ji, C. Y., Zhao, Y., Li, W., Feng, J., Zhang, X. (20014) Single nucleotide polymorphism at alcohol dehydrogenase-1B is associated with risk of esophageal squamous cell carcinoma. Cancer Cell International, 14(1), 12.

Yi, M., Schnitzer, J.E. (2009). Impaired tumor growth, metastasis, angiogenesis and wound healing in annexin A1-null mice. Proceedings of the National Academy of Sciences U S A, 106(42), 17886-17891.

Yin, X., Dewille, J.W., Hai, T. (2008). A potential dichotomous role of ATF3, an adaptive-response gene, in cancer development. Oncogene, 27(15), 2118-2127.

Yin, X., Wolford, C.C., Chang, Y.S., McConoughey, S.J., Ramsey, S.A., Aderem, A., Hai, T. (2010). ATF3, an adaptive-response gene, enhances TGF{beta} signaling and cancer-initiating cell features in breast cancer cells. J Cell Sci, 123(Pt 20), 3558-3565.

Yokoyama, A., Omori, T. (2005) Genetic polymorphisms of alcohol and aldehyde dehydrogenases and risk for esophageal and head and neck cancers. Alcohol, 35(3), 175-185.

Yom, C.K., Han, W., Kim, S.W., Kim, H.S., Shin, H.C., Chang, J.N., Koo, M., Noh, D.Y., Moon, B.I. (2011). Clinical significance of annexin A1 expression in breast cancer. Journal of Breast Cancer, 14(4), 262-268.

Yu Jr., J., Finley, R.L. (2009) Combining multiple positive training sets to generate confidence scores for protein–protein interactions. Bioinformatics, 25, 105–111.

Yu, G., Li, F., Qin, Y., Bo, X., Wu, Y., Wang, S. (2010) GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. Bioinformatics, 26, 976–978.

Yu, H., Braun, P., Yıldırım, M.A., Lemmens, I., Venkatesan, K., Sahalie, J., Hirozane-Kishikawa, T., Gebreab, F., Li, N., Simonis, N., Hao, T., Rual, J.F., Dricot, A., Vazquez, A., Murray, R.R., Simon, C., Tardivo, L., Tam, S., Svrzikapa, N., Fan, C., de Smet, A.S., Motyl, A., Hudson, M.E., Park, J., Xin, X., Cusick, M.E., Moore, T., Boone, C., Snyder, M., Roth, F.P., Barabasi, A.L., Tavemier, J., Hill, D.E., Vidal, M. (2008) High quality binary protein interaction map of the yeast interactome network. Science, 322, 104–110.

Yu, H., Jove, R. (2004). The STATs of cancer--new molecular targets come of age. Nature Reviews Cancer, 4(2), 97-105.

Yu, H., Tu, K., Wang, Y.J., Mao, J.Z., Xie, L., Li, Y.Y., Li, X.Y. (2012) Combinatorial network of transcriptional regulation and microRNA regulation in human cancer. BMC Systems Biology, 6,61.

Yu, K.D., Fan, L., Di, G.H., Yuan, W.T., Zheng, Y., Huang, W., Chen, A.X., Yang, C., Wu, J., Shen, Z.Z., Shao, Z.M. (2010) Genetic variants in GSTM3 gene within GSTM4-GSTM2-GSTM1-GSTM5-GSTM3 cluster influence breast cancer susceptibility depending on GSTM1. Breast Cancer Research and Treatment, 121, 485-496.

Yu, Y.L., Chou, R.H., Liang, J.H., Chang, W.J., Su, K.J., Tseng, Y.J., Huang, W.C., Wang, S.C., Hung, M.C. (2013) Targeting the EGFR/PCNA signaling suppresses tumor growth of triple-negative breast cancer cells with cell-penetrating PCNA peptides. PLoS One, 8, e61362.

Yuan, Y., Xue, L., Fan, H. (2014) Screening of differentially expressed genes related to esophageal squamous cell carcinoma and functional analysis with DNA microarrays. International Journal of Oncology, 44 (4), 1163-1170.

Zaika, A., Irwin, M., Sansome, C., Moll, U.M. (2001). Oncogenes induce and activate endogenous p73 protein. The Journal of Biological Chemistry, 276(14), 11310-11316.

Zeller, C., Hinzmann, B., Seitz, S., Prokoph, H., Burkhard-Goettges, E., Fischer, J., Jandrig, B., Schwarz, L.E., Rosenthal, A., Scherneck, S. (2003) SASH1: a candidate tumor suppressor gene on chromosome 6q24.3 is downregulated in breast cancer. Oncogene, 22(19), 2972-2983.

Zhang, A.H., Sun, H., Qiu, S., Wang, X.J. (2013). Metabolomics in noninvasive breast cancer. Clinica Chimica Acta, 424, 3-7.

Zhang, D.Y., Ye, F., Gao, L., Liu, X., Zhao, X., Che, Y., Wang H., Wang, L., Wu, J., Song, D., Liu, W., Xu, H., Jiang, B., Zhang, W., Wang, J., Lee, P. (2009) Proteomics, pathway array and signaling network-based medicine in cancer. Cell Division, 4, 20.

Zhang, F., Du, G. (2012). Dysregulated lipid metabolism in cancer. World Journal of Biological Chemistry, 3(8), 167-174.

Zhang, G.H., Mai, R.Q., Huang, B. (2010) Meta-analysis of ADH1B and ALDH2 polymorphisms and esophageal cancer risk in China. World Journal of Gastroenterology, 16(47), 6020-6025.

Zhang, W., Weissfeld, J.L., Romkes, M., Land, S.R., Grandis, J.R., Siegfried, J.M. (2007). Association of the EGFR intron 1 CA repeat length with lung cancer risk. Molecular Carcinogenesis, 46(5), 372-380.

Zhang, Y., Zhang, B. (2006) D4-GDI, a Rho GTPase regulator, promotes breast cancer cell invasiveness. Cancer Research, 66, 5592-5598.

Zhang, Z., Yamashita, H., Toyama, T., Sugiura, H., Ando, Y., Mita, K., Hamaguchi, M., Hara, Y., Kobayashi, S., Iwase, H. (2005). Quantitation of HDAC1 mRNA expression in invasive carcinoma of the breast*. Breast Cancer Research and Treatment, 94(1), 11-16.

Zhao, H., Yang, J., Fan, T., Li, S., Ren, X. (2012) RhoE functions as a tumor suppressor in esophageal squamous cell carcinoma and modulates the PTEN/PI3K/Akt signaling pathway. Tumour Biology, 33(5), 1363-1374.

Zhao, M., Sun, J., Zhao, Z. (2013) TSGene: a web resource for tumor suppressor genes. Nucleic Acids Research., (Database issue):D970-6.

Zheng, R., Blobel, G.A. (2010) GATA Transcription Factors and Cancer. Genes Cancer, 1, 1178-1188.

Zhi, H., Zhang, J., Hu, G., Lu, J., Wang, X., Zhou, C., Ding, F., Zhang, W., Sun, B., Liu, Z. (2003) The deregulation of arachidonic acid metabolism-related genes in human esophageal squamous cell carcinoma. International Journal of Cancer, 106(3), 327-333.

Zhou, L., Li, K., Luo, Y., Tian, L., Wang, M., Li, C., Huang, Q. (2013) Novel prognostic markers for patients with triple-negative breast cancer. Human Pathology, 44, 2180-2187.

Zhu, C., Qi, X., Chen, Y., Sun, B., Dai, Y., Gu, Y. (2011). PI3K/Akt and MAPK/ERK1/2 signaling pathways are involved in IGF-1-induced VEGF-C upregulation in breast cancer. Journal of Cancer Research and Clinical Oncology, 137(11), 1587-1594.

Zhu, H., Bhaijee, F., Ishaq, N., Pepper, D.J., Backus, K., Brown, A.S., Zhou, X., Miele, L. (2013) Correlation of Notch1, pAKT and nuclear NF-kappaB expression in triple negative breast cancer. American Journal of Cancer Research, 3, 230-239.

Zhu, H., Snyder, M. (2003) Protein chip technology. Current Opinion in Chemical Biology, 7, 55–63.

Kubra Karagoz

E-mail: kubra.karagoz@marun.edu.tr

Date of Birth: Feb/26/1985

Language: Turkish (Native), English (Fluent)

Education:

2010- 2015 **PhD in Bioengineering, Marmara University**
Istanbul, Turkey

- G.P.A: 3.71 (of 4)

Thesis title:

"Systematic and integrative analysis of breast cancer and other associated diseases using transcriptome data and interactome networks"

Thesis Advisor:

Kazim Yalcin Arga Assoc Prof., BioEngineering, Marmara University

2008- 2010 **MSc in Chemical Engineering, Boğaziçi University**
Istanbul, Turkey

- G.P.A: 3.25 (of 4)

Thesis title:

"Investigation of Relationship Between Sphingolipid and Insulin Signaling Pathways"

Thesis Advisor:

Kutlu Ulgen Prof., Chemical Engineering, Boğaziçi University

2003- 2007 **BSc in Chemical Engineering, Marmara University**
Istanbul, Turkey

- G.P.A: 3.28 (of 4)

Final Project:

"Ethanol Production from Agricultural Waste and Design The Process"

Project Advisor:

Dilek Kazan, Prof., Chemical Engineering, Marmara University

Research Experience:

2013-2014 *Visiting Scholar,* Department of Biochemistry and Molecular Biology, Penn State Hershey Cancer Institute, Chemical Carcinogenesis and Chemoprevention Laboratory, Penn State University, Hershey, PA, USA

2009-2010 *Research Assistant of Scientific and Technological Research council of Turkey (TUBITAK) Research Project*
Research Assistant to Prof Kutlu Ulgen, at Bogazici University
'Investigation of the Relationship between Sphingolipid and Insulin Signaling Pathways'

July-August 2008 *Visiting Researcher,* Medical University of South Carolina Charleston, Biochemistry and Molecular Biology Laboratory, Charleston, USA

Publications

- Karagoz, Kubra; Sinha, Raghu and Arga, Kazim Yalcin, Proteomic and Metabolic Signatures of Esophageal Squamous Cell Carcinoma, Current Cancer Drug Targets, submitted.

- Sinha, Indu; Karagoz, Kubra; Fogle, Rachel L.; Hollenbeak, Christopher S.;Zea, Arnold H.; Arga, Kazim Y.; Stanley, Anne; Hawkes, Wayne C.; Sihna, Raghu, Changes in Plasma Proteome Network following selenized-yeast supplementation in healthy men, PLOS ONE, submitted.

- Calimlioglu, Beste; Karagoz, Kubra; Sevimoglu, Tuba; Gov, Esra and Arga, Kazim Yalcin, Tissue-Specific Molecular Biomarker Signatures of Type 2 Diabetes: An Integrative Analysis of Transcriptomics and Protein-Protein Interaction Data, 2015, Omics: A Journal of Integrative Biology, 19(9): 563-573.

- Karagoz, Kubra; Sinha, Raghu and Arga, Kazim Yalcin, Triple Negative Breast Cancer: A Multi-Omics Network Discovery Strategy for Candidate Targets and Driving Pathways, 2015, Omics: A Journal of Integrative Biology, 19(2): 115-130.

- Karagoz, Kubra and Arga, Kazim Yalcin, Assessment of high-confidence protein-protein interactome in yeast, 2013, Computational Biology and Chemistry, 45,1-8.

Conference Presentation(s) ***Academic***

Poster

October, 2014 Kubra Karagoz, Raghu Sinha, Kazim Yalcin Arga
"OncoOmics Approach to Reveal Promising Candidate Biomarkers of Esophageal Squamous Cell Carcinoma", Beyond the Genome: Cancer Genomics Conference, October 8-10, 2014, Boston, MA, USA

August, 2014 Kubra Karagoz, Raghu Sinha, Kazim Yalcin Arga
"Deciphering Triple Negative Breast Cancer Driven Pathways via Multi-Omic Networks", FEBS-EMBO Conference, August 30-September 4, 2014, Paris ,France.

April, 2014 Kubra Karagoz, Tuba Sevimoglu, Kazim Yalcin Arga
"Complementary Analysis of Transcriptome Data Enlightens an Unexplored Link between Psoriasis and Prostate Cancer", Bio-IT World Conference, April 29- May1, 2014, Boston, USA.

June, 2013 Kubra Karagoz, Kazim Yalcin Arga, Gary Stoner, Raghu Sinha
"Hypotheses Generation for Genetic Mechanism of Esophageal Cancer Development Via Integrated Transcriptome and Interactome Analysis", International Conference on Genomics in Europe-ICG 2013, June 25-28, 2013, Ghent, Belgium.

December, 2012 Kubra Karagoz and Kazim Yalcin Arga
"Investigation of Colon Cancer Associated Genes and Pathways With Systems Bioinformatics Approaches", IDIBELL Cancer Conference on Personalized Medicine-ICC 2012, December 3-4, 2012, Barcelona, Spain.

March, 2012 Kubra Karagoz and Kazim Yalcin Arga
"Towards an Almost Complete, Highly Reliable Protein-Protein Interaction Network of Yeast", Visualizing Biological Data-VizBi 2012, March 6-8, 2012, Heidelberg, Germany.

Oral

September, 2012 Kubra Karagoz and Kazim Yalcin Arga
"Combining Various Heterogeneous Biological Features to Obtain a Highly Reliable Almost Complete Protein Interaction Network of Yeast", 15th European Congress on Biotechnology, September 23-26, 2012, Istanbul,Turkey

April, 2010 Kubra Karagoz and Kutlu Ulgen
"Investigation of the Relationship Between Sphingolipid and Insulin Signaling Pathways", 15th National Biomedical Engineering Meeting, April 21-24, 2012, Antalya, Turkey.

Honors& Awards

- Selected as a young scientist for the 14th Federation of European Biochemical Societies (FEBS) Young Scientists' Forum (YSF)
- Full scholarship awarded by TUBITAK (The Scientific and Technical Research Council of Turkey) 108M649 Research Project
- Graduated from the Department of Chemical Engineering in Marmara University as the second highest GPA

Computer Skills

- Microsoft Word, Excel, PowerPoint
- MATLAB, R Program

References

Assoc Prof Kazim Yalcin Arga

Bioengineering Department, Marmara University, Istanbul/TURKEY

kazim.arga@marmara.edu.tr

Assoc Prof Raghu Sinha

Department of Biochemistry and Molecular Biology, Penn State College of Medicine, Hershey, PA/USA

rus15@psu.edu