Computational Modeling of Cancer-Associated Cellular Systems
by

Jian Li, M. Sc.

<Ph.D. Thesis>

Presented to the Faculty of
Biology, Chemistry and Pharmacy of
Free University of Berlin
in Partial Fulfillment of the Requirements
for the Degree of

Doctor Of Philosophy

The Free University of Berlin

<Summer 2013>

Computational Modeling of Cancer-Associated Cellular Systems

APPROVED BY

SUPERVISING COMMITTEE:

| Professor Burghardt Wittig |
| Professor Peter N. Robinson |

DISPUTATION on 30.10.2013

Dedication

I lovingly dedicate my dissertation to my wife, who supported me each step of the way and has always been taking care of our children so that I could concentrate on my work in daily life.

I also dedicate my dissertation to my many friends and church families who have supported me throughout my entire PhD study. I would never forget how they have supported me with their personal experience, spiritual encouragement, wise counsel and material support. Especially, Robert Weissman, Bjoern Gruenning for their excellent discussions related to Bioinformatical research field; Family Gentsch, Family Pimentel and Family Volker for their intensive spiritual teachings and encouragement; Birgit Adam and Dominik Jednoralski for their excellent professional work as PhD Net representatives.

I also want to dedicate this thesis to my parents in China, despite the fact that we have not spoken very often since I left for Germany 10 years ago. However, their selfless care and support are deeply embedded in the bottom of my heart.

Lastly, I dedicate this thesis to my wonderful daughters, Loi and Lorri, both of whom are very young. Most of the time, they demand my full love and attention. Their presence has always been my best true joy under any circumstance.

Acknowledgements

I wish that I could express my profound gratitude towards many colleagues and professors in person, who were more than generous with their expertise and precious time. A special thank to Dr. Andriani Daskalaki, the senior cancer researcher for her precious careful reading, professional feedback, wise counsel and faithful encouragement throughout my entire PhD process. A special thank to Dr. Hendrik Hache, the senior biophysical researcher for his outstanding technical discussion and support. A special thank for Prof. Burghardt Wittig and Prof. Ferdinand Hucho for their precious, occasional discussions, encouragement and counsel.

The work of this thesis is conducted in the Max-Planck-Institute for molecular genetics, in Berlin. I would like to thank Dr. Christoph Wierling, the leader of Systemsbiology group in this institute, for his discussion about the title of this thesis. The revised version of this thesis is based on the informative and constructive suggestions from Prof. Burghardt Wittig and Prof. Olaf Wolkenhauer.

I would like to thank my former colleagues in this institute, Vikash Pandey and Felix Heeger for their countless discussions and valuable collaboration. And I also would like to thank many other colleagues for their continuing support.

I would like to acknowledge and thank my faculty Biology, Chemistry and Pharmacy, at Free University of Berlin for allowing me to conduct my research and providing any assistance requested. Special thank goes to the Promotion-Office, Mrs. Kerstin Austen for her continuing advice and assistance.

Abstract

Computational Modeling of Cancer-Associated Cellular Systems

Jian Li, M. Sc.
The Free University of Berlin, 2013

Cancer is the consequence of disordered cellular systems. To date, more than hundred distinct types of cancer and subtypes of tumor have been identified in the human body. This complexity does not only provoke a number of questions such as the origin of cancer and its medical treatment, but also poses a high level of challenge to battle against them. The experience gained by diverse cancer-research studies suggests that all differences among traditional cancer and tumor types can be reduced at the molecular level. Thus, it is essential to understand the interplay and relation between oncogenes and tumor suppressors within the cancerous cellular system, which could help to identify the core function of molecular pathways for effective therapeutic intervention. This study tried to realize an important system biological concept that integrates modeling of different types of biological networks including signaling-, metabolic-, gene-regulatory-, and miRNA-regulation-network from the whole organism and generate responsive computational molecular models to describe the behavior of cells (or cancer cells). The models created during this thesis are demonstrated with application-examples to be able to incorporate the experimental and observational inquiry and receive the individual patient information to reflect/predict the dynamic behavior of the underlying cellular systems. This study also tried to elucidate that models of this kind can be utilized for early tumor detection, precise diagnose of the pathological state, the identification of anti-tumor/cancer drug and drug combination effect. Therefore, models of this kind might be able to form a "Virtual Patient" model to propose, verify and predict new therapeutic strategies towards personalized medicine. In addition, a systemsbiological software has been developed during this study. Its novel properties including its internal modeling database can be used to overcome some current limitations of systemsbiology research.

Computergestützte Modellierung von Krebs-Assoziierten Zellulären Systeme

Jian Li, M. Sc.
Freie Universität Berlin, 2013

Krebs ist eine Art von Krankheit, die durch funktionale Störungen innerhalb zellulärer Systeme verursacht wird. Bei Menschen und Tieren bezeichnet Krebs maligne Tumoren, die unkontrolliertes Wachstum zeigen, was auf den Verlust der Proliferationskontrolle zurückzuführen ist. Krebs kann aus allen sich teilenden Zelltypen entstehen, sodass bis heute mehr als 100 verschiedene Krebserkrankungen bekannt sind. Gegenstand der Krebsforschung heutzutage ist auf der einen Seite die Aufklärung der Krebs-Entstehung und auf der anderen Seite die Entwicklung von neuen bzw. verbesserten medizinischen Behandlungmöglichkeiten für Krebs. Die aus diversen Studien gewonnenen Erkenntnisse erwecken den Eindruck, dass alle Unterschiede zwischen traditionellen Krebs- und Tumortypen auf die molekulare Ebene zurückgeführt werden können. Daher ist es wichtig, das Zusammenspiel zwischen Onkogenen und Tumorsuppressorgenen in den molekularen Prozessen die zur einzelnen Krebszelle führen, noch genauer zu verstehen, um Krebs auf molekularer Ebene effektiver zu bekämpfen und schließlich zu besiegen.

Diese Arbeit versucht, ein wichtiges systembiologisches Konzept zu verwirklichen, indem ein integratives mathematisches Computer-Modell entwickelt wurde, das unterschiedliche biologische Netzwerke (Signal-, metabolische Netzwerke, genregulatorische Netzwerke und miRNA Regulationsnetzwerke) beinhaltet. Dieses integrative Modell besteht aus verschiedene Sub-modelle, die während der Arbeit aufgebaut und validiert worden sind. Die Arbeit hat auch aufgeführt wie man Modelle dieser Art mit Anwendung/Integration von der genetischen Information der individuellen

Zelllinien and experimentellen Daten verwenden kann, um das dynamische Verhalten von Krebszellen aufzuklären.

Diese Arbeit versucht auch, zu erklären, dass Modelle dieser Art auch für medizinische Anwendungen, z.B. die Früherkennung von Tumoren, die Identifizierung der Wirkstoffe und Diagnose der Wirkung von Anti-Krebs-Medikamenten oder Anti-Krebs-Arzneimittelkombinationen, einsetzbar sind. Daher könnten diese Modelle in der Lage sein, als ein "Virtueller Patient" zu funktionieren, und neue therapeutische Strategien zur personalisierten Medizin vorschlagen, überprüfen und vorhersagen. Zusätzlich wurde eine systemsbiologische Software während der Arbeit entwickelt, derer vorteilhafte Eigenschaften, inklusive deren internale Datenbank, die dazu gedient werden kann, Beiträge zur systemsbiologische Forschung zu leisten.

10

12

List of Tables

List of Figures

Abbreviations

ACVR   Activin A receptor

ADP   Adenosine Diphosphate

AMP   Adenosine Monophosphate

AMO   Anti-miRNA Oligonucleotide

AKT   Arotein Kinase B

AR   Androgen Receptor

ATF   Activating Transcription Factor

ATP   Adenosine Triphosphate

BAD   Bcl2-associated Agonist of Cell Death

BCL   B-cell CLL/Lymphoma

BCL2L1   BCL like 1

CAMK   Calcium/Calmodulin-dependent Protein Kinase

CASP9   Caspase 9, apoptosis-related Cysteine Peptidase

CCL   Chemokie (C-C motif) Ligand

CSF   Colony Simulating Factor

CDKN1A   Cyclin-dependent Kinase Inhibitor 1A (p21, Cip1)

CISH   Cytokine Inducible SH2-containing Protein

CNTF   Ciliary Neurotrophic Factor

COX   Cyclooxygenase

CREB   cAMP responsive Element Binding Protein

DGCR8   DiGeorge Syndrome Chromosome Region 8

DNA   Deoxyribonucleic Acid

EIF4EBP   Eukarytoic Transition Initialization Factor 4E Binding Protein

ELK1   Member of ETS Oncogene Family 1

EGF(R)   Epidermal Growth Factor (Receptor)

ERK	Mitogen-activated Protein Kinase
ESR	Estrogen Receptor
FAS	Fas (TNF receptor superfamily, member 6)
FBA	Flux Balance Analysis
FCGR1A	Fc Fragment of IgG, high affinity la, Receptor (CD64)
FGF(R)	Fibroblast Growth Factor (Receptor)
FOS	FBJ Murine Osteosarcoma Viral Oncogene Homology
FOXM	Forkhead Box M1
GBP1	Guanylate Binding Protein 1, Interferon-inducible
GDF	Growth Differentiation Factor
GLI	GLI Family Zinc Finger
GSTM	Cytosolic Glutathione Transferase
HGF(R)	Hepatocyte Growth Factor (Receptor)
HIF1	Hypoxia Inducible Factor 1
HMMA	Human Metabolic miRNA (model)
hTERT	human Telomerase Reverse Transcriptase
IEG	Induced Early Response Gene
IGF1R	Insulin-like Growth Factor 1 Receptor
IKBKB	Inhibitor of Kappa Light Polypeptide Gene Enhancer in B-cells, beta
IKBKE	Inhibitor of Kappa Light Polypeptide Gene Enhancer in B-cells, epsilon
IKBKG	Inhibitor of Kappa Light Polypeptide Gene Enhancer in B-cells, gamma
IL1	Interleukin1
IRF1	Interferon Regulatory Factor 1
ISG15	ISG15 Ubiquitin-like Modifier
JAK1	Janus Kinase 1
JUN	Jun Proto-oncogene
LEP(R)	Leptin (Receptor)

MacOS	Macintosh Operating System
MAPK	Mitogen-activated Protein Kinase
MAP2K	Mitogen-activated Protein Kinase Kinase 2
MCL	Myeloid Cell Leukemia Sequence 1 (BCL2-related)
MDM2	Mdm2, p53 E3 Ubiquitin Protein Ligase Homolog
MEK	Dual specificity Mitogen-activated Protein Kinase Kinase
MKI67	Antigen Identified by Monoclonal Antibody Ki-67
MTOR	Mechanistic Target of Rapamycin (serine/threonine kinase)
MVP	Major Vault Protein
MYB	v-Myb Myeloblastosis Viral Oncogene Homolog (avian)
MYC	v-Myc Myelocytomatosis Viral Oncogene Homolog (avian)
NFKB	Nuclear Factor of Kappa Light Polypeptide Gene Enhancer in B-cells
NRAS	Neuroblastoma RAS viral (v-ras) Oncogene Homolog
NSAID	Non-Steroidal Anti-Inflammatory Drug
OSM(R)	Oncostain M (Receptor)
PAK	p21 protein (Cdc42/Rac)-activated kinase
PDPK1	3-phosphoinositide dependent protein kinase-1
PI3K	Phosphatidylinosytol-3-kinase
PIK3CA	phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha
PIK3R1	phosphoinositide-3-kinase, regulatory subunit 1 (alpha)
PIM1	Pim-1 Oncogene
PG	Prostaglandin
PKC	Protein Kinase C
PLK1	Polo-like Kinase 1
PP	Protein Phosphatase
PTEN	Phosphatase and Tensin Homolog
pRb	Retinoblastom-Protein

p53	Tumor Protein 53
RNA	Ribonucleic Acid
Ras	Rat Sarcoma
Raf	Rapidly Accelerated Fibrosarcoma
RISC	RNA-induced silencing complex
RNASEN	RNase III DROSHA
RPS6KA	Ribosomal Protein S6 Kinase, 90kDa, Polypeptide
SBML	Systems Biology Markup Language
SBW	Systems Biology Workbench
SMAD	SMAD Family Member
STAT	Signal Transducer and Activator of Transcription
TARBP2	TAR (HIV-1) RNA binding protein 2
TIMP1	TIMP Metallopeptidase Inhibitor 1
TGFbeta	Transforming Growth Factor beta
TLR	Toll-like Receptor
TNF	Tumor Necrosis Factor
TOP2A	DNA topoisomerase 2-alpha
TP53	Tumor Protein p53
TRIM21	Tripartite Motif-containing Protein 21
TSC2	Tuberous Sclerosis 2
ODE	Ordinary Differentiation Equation
URGCP	Upregulation of Cell Proliferation
VEGF	Vascular Endothelial Growth Factor
ZOPE	Z Object Publishing Environment
WNT	Wingless-type MMTV integration site family

1. Introduction

1.1 TUMORIGENESIS

Different significant and fundamental discoveries in cell biology from the nineteenth century till now have led us to understand that all cells in the body of a multicellular organism including human, can be traced back to the fertilized egg. The fertilized egg develops later into a stem cell which is capable to produce all the cells in the body through repeated cycles of cell growth and division (Weinberg, 2007). These versatile, autonomic cellular processes span the time interval of multiple magnitude, which involves in cell growth, death, division, etc. Therefore, they pose different types of grave dangers, such as randomly occurring mutations and epigenetic alternations of DNA that can affect the functional ability of different cellular components to control proliferation, survival and other cellular processes. In addition, the failure of DNA-repair or DNA-damage sensing processes helps to accumulate relevant mutations (Boulton, 2006). All kinds of the cellular damaging processes that help or assistant normal cells to be transformed into cells with increasingly neoplastic phenotypes, are termed "tumor progression" or "tumorigenesis" in biology (Hanahan & Weinberg, 2000). Thanks to the rapid progression of cancer research in the past decades, we came to the knowledge that the "tumorigenesis" in human cells is a multiple-step-process that reflects the genetic alternations driving the progressive transformation of normal cells into highly malignant derivatives. The "tumorigenesis" process indicates that the body of a complex organism including human possess multiple lines of defense mechanisms developed and established during evolution. Each of these defense mechanisms is provided and maintained by the hard-wiring of a complex cellular regulatory circuit. The breaching of each defense mechanism symbolizes a certain cellular security breach and indicates an advance of the pathological state, until the final clinically detectable tumor or cancer occurs (Hanahan &

Weinberg, 2011). To the present day, more than several hundred types of tumors and cancers have been found within specific tissues and organs of individual patients. Nonetheless, one important question remains unknown: how many cellular regulatory pathways need to be altered or deregulated before a normal cell becomes cancerous?

1.1.1 Deregulation of Cellular Regulatory Circuits

Dr. Robert Weinberg states in his book "The Biology of Cancer" (Weinberg, 2007) that at least five distinct cellular regulatory circuits need to be altered or deregulated before a normal cell can be transformed into a cancer cell:

(a) The deregulation of the mitogenic signaling pathway,

(b) The deregulation of the cell cycle checkpoint,

(c) The deregulation of the alarm pathway,

(d) The deregulation of telomere maintenance,

(e) The deregulation of anti-growth signaling pathway,

(a) The deregulation of the mitogenic signaling pathway (e.g. gain of function of Ras):

Weber et al., (2000) investigated that signaling through the mitogenic Ras/Raf/MEK/ERK pathway induces cell growth and differentiation. Later, many studies provided the evidence that different types of oncogenic Ras can induce different types of cellular pathological states (Aoki et al., 2005; Niihori et al., 2006; Schubbert et al., 2006), which convincingly explained the pathological effect of deregulation of a putative mitogenic signaling pathway.

(b) The deregulation of the cell cycle checkpoint (e.g. loss of function of pRb):

The inappropriate completion of a cell phase can lead to catastrophic genetic damage. In order to minimize the potential of mistakes in cellcycle events, different control mechanisms implemented within four key-checkpoints ensure a smooth cell cycle

progression (Lodish et al., 2000 [chapter 13]). The results of diverse studies provide clear evidence that pRb with its regulation of the G1/S-phase transition (checkpoint), is frequently inactivated in human cancers (Manning et al., 2010; Santamaia & Pagano, 2007; Sinclair & Frost, 1999), which indicates the serious consequence of deregulation of a cell cycle checkpoint.

(c) The deregulation of the alarm pathway (e.g. loss of function of p53):

p53 is regarded as master guardian of cellular processes and its main activity is to respond to multiple signals of cellular alarm, including DNA damage (Kastan et al., 1991; Maltzman and Czyzyk, 1984; Lu and Lane, 1993), perturbations of cell cycle regulation (Diller et al., 1991; Kuerbitz et al., 1992; Martinez et al., 1991), programmed cell death (Clarke et al., 1993; Lowe et al., 1993; Yonish-Rouach et al., 1991), etc. Persons with the 'loss of function' mutations of p53 are predisposed to different types of cancer and usually develop several independent tumors in a variety of tissues in early adulthood (Hollstein et al., 1991).

(d) The deregulation of telomere maintenance pathway (e.g. gain of function of hTERT):

Telomeres are found at the ends of linear chromosomes and have important roles in eukaryotic cellular processes including chromatin organization and the control of cell proliferation (Price, 1999; Harley, 2002). The telomerase activity for telomere maintenance is absent from most human somatic cells beyond the early stages of fetal development (Wright et al., 1996; Prowse & Greider, 1995). The 'gain of function' of human TERT can maintain the balance between telomere loss and telomere addition, which leads to the sustained proliferative ability (Wong & Collins, 2003).

(e) The deregulation of anti-growth signaling pathway (e.g. loss of function of PP2A):

22

Anti-growth signaling pathways (protein phosphatases) function as the negative regulator of cellular processes to balance positive cellular regulators ability and therefore to ensure a kind of cellular system control. PP2A is an important negative regulator of the extracellular signal-regulated-kinase signaling pathway (Chen et al., 2001; Johnson & Lapadat, 2002; Ugi et al., 2002) and is involved in the control of many cellular processes including metabolism, transcription, translation, RNA splicing, DNA replication, cell cycle progression, transformation, and apoptosis (Heriche et al., 1997; Janssens et al., 2001; Silverstein et al., 2002; Sontag, 2001; Virshup, 2000). The 'loss of function' of PP2A or other protein phosphatases (such as PP1, PP2B, PP4, PP5, PP6 and PP7) leads to the aberrant protein phosphorylation, which has been linked to many human diseases (Cohen, 2002). As a consequence, protein phosphatases have widely gained recognition as future therapeutic targets (Cohen, 2006; Tonks, 2006; Zhang et al., 2006).

Furthermore, more than ten different cancer hallmarks are introduced to symbolize the different aspects of tumorigenesis and to provide a biological organization principle for cancer research (Hanahan & Weinberg, 2000; 2011). However, two important questions remain unknown: how exactly can these cancer hallmarks be applied in order to contribute to cancer research? Can these cancer hallmarks be used as potential indicators for personalized medicine?

1.1.2 Personalized Medicine

Personalized medicine is an advanced medical treatment approach by taking the specific genetic information (including gene-expression level, proteomic data, mutational status and others) of individuals into consideration, which can direct the course of treatment more in step with the way how the individual body functions. This way the medical care can be optimized based on the individualized assessment of pathological risk and, prescription of clinical treatments with a high likelihood of success (Chan & Ginsburg, 2011). In comparison to the traditional medicine, personalized medicine should be more

23

precise, accurate and effective by offering medical care and making clinical decisions on the individual level.

Although tailoring medical treatments down to the individual level can certainly benefit each of individuals, the ways for reaching this goal could be myriad and subtle. One way is to create classification of individuals to group them into different categories labeling with ages, weights, health conditions, susceptibility to a particular disease and responses to specific medical treatment. Despite the fact that the creation of those kinds of categories could raise plenty of ethical questions including human rights and, discrimination and misuse possibilities (McClellan et al., 2012), the huge advantage of personalized medical care should not be underestimated. To the present, different methods and approaches have been developed for this classification purpose. Most of them focus on the genetic map or the profile of individual genetic variation. For instance, Kroll (2008) describes a concept of DNA biomarker for preventive surgery of individuals. Manolio (2010) summarizes the relationship between genetic profile and personalized assessment of disease risk. Ball et al., (2012) create standard health genome sequence data and try to indicate the possibility of individual pathological risk, whose genome sequence data possesses a certain high level of variants compared to the standard ones.

However, personalized medicine is more than a genetic map or genetic variation. It includes individual preventive therapeutic intervention, individual identification of optimal drug combination, individual calculation of the optimal drug dosage, individual withholding treatment, individually targeted treatment and others. The individual genome is like an original blueprint and expresses and produces more than a million different proteins in order to construct a complex cellular system, the human body, which is further under the influence of environmental and genetic factors. Therefore, in order to perform a meaningful practice of personalized medicine, in this study I intend to introduce an advanced computing model combined with systemsbiological approaches and try to open

a more effective and sophisticated personalized medicine field for a revolution of medical care.

1.2 MICRORNA REGULATION REGARDING CANCER RESEARCH

MicroRNAs (miRNAs) are endogenous, non-protein-coding, 18-22 nucleotide RNAs that are evolutionary conserved with unique function to regulate gene expression at the post-transcriptional level in a sequence-specific manner (Cho, 2007; Esquela-Kerscher et al., 2006; Voorhoeve et al., 2007). To the present day, it is known that the human genome expresses more than 1,000 miRNAs, which is equivalent to about 3% of the total number of human genes (Bartel, 2004). Diverse studies provide compelling evidence that miRNAs play indispensable roles in nearly all cellular processes including developmental timing, proliferation, apoptosis, differentiation, stem cell maintenance, signaling pathways, and pathogenesis (Bueno et al., 2008; Hwang et al., 2006; Wang & Blelloch, 2009; Zhang et al., 2007a; Zhang et al., 2007b; Zhang et al., 2008). The expression of approximately 30-40% of human proteins appears to be regulated by miRNAs (Lewis et al., 2005). Via complementary base pairing between their 5' seed sequence and the target mRNA 3' untranslated region, miRNAs can control the expression level of target gene. The 5' "seed" region of the miRNA sequence (bases two to eight) is essential in mRNA target recognition (Lytle et al., 2007). With imperfect complementarity of both sequences, miRNA can bind to mRNA to repress its translation process, whereas miRNAs binding to the mRNA with perfect complementarity target it for destruction (Hatley et al., 2010). Due to the relatively few complementary base pairs, the target spectrum of miRNAs can be very promiscuous, which becomes a great obstacle to determine the precise number of targets of each miRNA (Li et al., 2012a). However, according to miRNA research results and target prediction softwares, it is reasonable to suggest that the number could not be less than hundreds, which implies that a single miRNA can target multiple components of

25

a single cellular pathway, or components of multiple pathways. As a matter of fact, this could confer an individual miRNA the unique function as a signaling amplifier, to convey signaling crosstalk between pathways or to confer signaling robustness of signaling pathways (Chiocca & Lawler, 2010).

1.2.1 MiRNA & Cancer Research

Cancer is the result of accumulation of relevant genetic mutations that multiplicatively generate a consequence of disordered cellular and genome functions. As a matter of fact, miRNAs play essential roles in diverse cellular processes such as proliferation, apoptosis, angiogenesis, and others (Li et al., 2012a). The results of diverse studies link miRNA to the aetiology, progression and prognosis of cancer and its deregulation is highly associated with important cellular activities that contribute to tumorigenesis (Brid et al., 2010; Cui et al., 2006). Moreover, several studies have attempted to utilize the miRNA expression profiling in order to identify and classify several types of cancers (Gaur et al., 2007; Tong & Nemunaiti, 2008). In addition, several recent studies demonstrated the different regulation mechanisms of miRNAs on signaling pathways (Huang et al 2010; Li et al., 2012a). However, a comprehensive model integrating a signaling network with a miRNA-regulation network (including miRNA biogenesis, miRNA-target regulation) has not been proposed yet and an interesting question would be whether the construction such a complex model incorporating diverse level of cellular regulation mechanisms could help us more to understand cellular pathological origin.

One key challenge in the field of cancer research is to understand how cancer cells can acquire abilities of uncontrolled cell growth, proliferation, aggressive invasion and damage of adjacent tissues, and simultaneously ignore and circumvent cellular security mechanisms such as apoptosis (Hanahan & Weinberg, 2000 and 2011). One approach in molecular cancer therapy is to specifically manipulate deregulated intracellular signaling pathways within the cancerous cellular system by balancing the aberrant cellular signal

produced by over-expressed oncogenes or by the expression level decrease of related tumor-suppressor genes. A newly proposed alternative approach might employ miRNA targeting drugs to exert a potential broad impact which a particular miRNA or group of miRNAs might possess on multiple components within the same deregulated signaling pathway to effectively balance cellular aberrant signals and activities (Li et al, 2012a).

1.2.2 **MiRNA & EGFR Signaling Pathway**

The EGFR signaling pathway belongs to the most important cellular signaling pathways and is able to regulate relevant cellular processes, including proliferation, differentiation, and development (Webster et al., 2008). Many studies show that in many cancer-related malignant cellular processes including uncontrolled cellular proliferation, evading apoptosis and autocrine signaling transduction, up-regulation and/or over-expression of the EGFR signaling have been often and clearly observed. And the over-expression of EGFR is frequently found in epithelial tumor entities such as gastric, colorectal, head-and-neck, breast, and lung cancers and is associated with advanced disease and poor clinical prognosis (Gross et al., 2004; Yano et al., 2003). Furthermore, the EGFR signaling pathway also appears to protect cancer cells from toxic actions generated by different clinical treatments including chemotherapy and radiotherapy, and to render these treatment modalities less effective and efficient (Blume-Jensen et al., 2001; Chen et al., 2000; Longley et al., 2006). Therefore, it would be very interesting to gather current available biological information related to this pathway for construction of a molecular model in order to elucidate those therapeutics-relevant cellular responses.

In the study of Li et al. (2012a), with colleagues, I applied systemsbiological approach to construct a molecular model of the EGFR signaling pathway and its related miRNA regulation network. All miRNA-target information defined in the model is manually collected from diverse published miRNA studies. Additionally, with colleagues, I performed an *in silico* analysis to investigate the impact of miRNAs and their specific

inhibitors on the EGFR signaling pathway, which successfully demonstrated the broad range of influences exerted through miRNA regulatory processes on the cellular behavior of this signaling pathway. Moreover, with colleagues, I quantitatively elucidated the newly introduced miRNA therapeutic concept "One hit -- multiple targets" from the study of Wurdinger & Costa (2007) with which we hope to open a new avenue for drug development in cancer research.

1.3 HUMAN SIGNALING MODEL

Living cells can sense and respond to internal and environmental stimuli and perturbation with the help of communication nets (such as signaling network). They coordinate diverse cellular responses according to the different types of stimuli and perturbation, and simultaneously monitor a wide range of parameters, including nutrient level, genetic mutation and deregulation of cellular regulatory circuits (Chen et al., 2008; Hyduke et al., 2009). Functioning as an interface between the environment, the genome and the metabolism, the signaling network tightly controls the decision-making processes of living cells and is responsible for the appropriate utilization and distribution of nutrient resources (Carter et al., 2007; Pawson et al., 1995). Different types of deregulation within the signaling network can help cancer cells to become extremely persistent and robust for growth and proliferation, to keep being driven by the engine of the cell cycle and to eliminate communication with their external environment and neighborhood. This in turn makes them insensitive against the anti-growth signal and easily evade the control of the immune system and apoptosis (Hanahan & Weinberg, 2000, 2011). Until now, many studies have tried to use different technologies to construct such a signaling network, however, all of them were mainly focused on the one particular signaling pathway such as Li et al., (2009), which constructed the Toll-like receptor signaling pathway presenting the largest signaling network in 2009; Oda et al., (2005), who constructed the EGFR signaling pathway that was considered the most comprehensive signaling network in

28

2005. However, a comprehensive model covering multiple signaling pathways has not been reported yet and another interesting question could be whether such a complex model of intracellular signaling network could help us to elucidate signaling transduction mechanisms that can contribute therapeutic intervention; whether such a complex signaling model could become a relevant step-stone for *in silico* therapeutic intervention testing approach.

One of the great challenges in medicine is to deliver therapies tailored to patients based on their molecular signatures (Auffray et al., 2011; Daskalaki A & Lazakidou A, 2012). In recent years, drugs specifically targeting deregulated signaling pathways are being increasingly used to treat cancer patients. Such mechanistic drugs in general have far fewer side effects than less specific chemotherapies, and often have very positive effects on the subset of patients responding to the therapy. With an average response rate of 25-35%, this is usually only a small fraction of the patients receiving the drug, with the majority only suffering side effects, but showing little or no benefit from the therapy (Gottesman, 2002; Tekir et al., 2007; Gan et al., 2010).

Systems biology concepts aim to integrate modeling of signaling pathways and regulatory networks at different biological levels in the whole organism. These concepts can provide key components in the field of personalized medicine (Auffray et al., 2009; Daskalaki A & Lazakidou A, 2012). The so-called "personalized medicine" would involve tools of patient response evaluation so as to select the effective drugs with optimal doses for the individual patient. The result of clinical outcome of a drug intervention can be calculated as the result of probabilistic interactions between drug-metabolizing enzyme genes included in regulatory networks and environmental factors (Kitano, 2007).

Hence, comprehending the dynamic behavior of the signaling network, one might be able to understand the causal connection between genotype and phenotype of cancer cells (Kitano, 2002). Therefore, I would like to investigate whether it is possible to create a

human signaling model, by means of which one could understand how the signaling network instantiated a specific pathological phenotype so that different hypotheses could be developed, verified or rejected and valuable predictions could be provided in order to develop new efficient therapeutic *in-silico* strategies and to reduce the medical-treatment-test cost (Kitano, 2002). Because of the major impact of miRNA regulation on diverse cellular processes, I will be curious about whether a global miRNA-regulation network integrated into a global signaling network could provide a more detailed and precise understanding of the cellular system of the individual cancer patient, so that personalized medicine fueled with systemsbiological approach could become more realistic and powerful.

Lastly, the identification of biomarkers associated with a particular disease such as a tumor, is in demand of enhancing early detection, diagnosis, and prognosis. However, many tumor biomarkers identified through the use of proteomic techniques in the past have failed to attain broad application for clinic use (Ludwig & Weinstein, 2005). And little is known about how the miRNA expression data can be used to identify the informative biomarker regarding an individual cancer cell line or even patient. Therefore, the next question, which is also one goal of personalized medicine, is how could one utilize miRNA expression data to identify miRNA biomarkers at the individual level.

1.4 HUMAN METABOLIC MODEL

Human metabolic pathways are usually in charge of degrading energy-rich materials from the environment, converting them into the cell's own characteristic metabolites, polymerizing monomeric precursors into macro-molecules such as nucleic acids and polysaccharides, and recycling or disposing intracellular metabolites to maintain the physiological health state of a body (Nelson & Cox, 2008). The deregulated metabolism is the precondition for a cell to become cancerous and many studies indicate that the disorder of metabolic pathways could help cancer cells dictate their own survival,

proliferation, and invasion (Smallbone et al., 2007; DeBerardinis et al., 2008). The study of Hanahan & Weinberg (2011) pointed out that major changes of metabolism (metabolic pathways) could support unstopped cell growth and proliferation, replace the metabolic program in normal tissues and fuel extreme physiological operations in a cancerous cell, which is represented by cancer hallmarks including deregulating cellular energetics and tumor-promoting inflammation. For instance, a recent study of Denkert et al. (2008) provided evidence that the deregulation of a metabolic pathway, TCA (Citrate Acid Cycle) cycle, could lead to the dysfunction of energy homeostasis, which helps the color cancer cells to grow, survive and invade.

Thanks to a variety of genome and molecular biological projects, a diverse and wide range of biological and clinical data has been generated, which includes transcriptomic, proteomic, metabolomic, and phenotypic data. In order to meet the demand for the dramatical increase of this molecular and clinic data, the large-scale metabolic models are needed to system-level understand those invaluable data in order to make full use of them. In fact, from the year 2000 on, computational modeling of cellular metabolism has been performed with different aspects and goals. Many studies have focused on the functional description of distinct metabolic pathways (Kanehisa & Goto, 2000; Romero et al., 2004) and constructing models of specific cell types or organelles (Wiback & Palsson, 2002; Chatziioannou et al., 2003; Vo et al., 2004). Nevertheless, the model development of genome-scale network has always been pursued as long-standing goal.

Recently, Duarte and his colleagues conducted a global reconstruction of the human metabolic network by applying the massive amount of genomic and bibliomic data (Duarte et al., 2007). They published their human metabolic model and described this reconstruction process with the emphasis of detailed and comprehensive evaluation of literature references. In this study, Duarte et al., have successfully demonstrated the genome-scale model's usages for the discovery of missing biological information, for the mathematical analysis of network structure related to implications of intracellular

31

compartmentalization, and for the potential use regarding the alternative drug target identification. Furthermore, based on the model, they introduced an integrated analysis of high-throughput data sets with the goal of a global assessment about the functional metabolic states. Following the study of Duarte et al., (2007), different studies have been conducted to extend the application of this high-quality genome-scale human metabolic model. For instance, Shlomi et al., (2009) utilized this human metabolic model and applied a flux variability analysis combined with the constraint-based approach to predict metabolic biomarkers of human inborn errors of metabolism. They quantitatively demonstrated the precision of their approach by utilizing the human metabolic model and reached a high level of prediction. Folger et al., (2011) applied the Model Building Algorithm (MBA) to generically generate a genome-scale model of cancer metabolism based on this human metabolic model for predicting common growth-supporting genes in different cancerous cellular environments, and this study also tried to propose and verify the anti-cancer drugs and targets.

Interestingly, in the same year as the study of of Duarte et al., another study of Ma et al., (2007) reported the reconstruction of the Edinburgh human metabolic network and tried to elucidate the necessity of applying different technologies including proteomics and metabolomics and collecting information of diverse databases for the purpose of the genome-scale network reconstruction. The aforementioned study also tried to discover generic mechanisms causing different abnormal metabolic states and to develop/identify drugs for targeting unknown/uncharacterized cellular components including genes and proteins for disease treatment.

However, to the present day, none of these published large-scale metabolic models contains the regulatory information related to miRNA-regulation. None of them computationally or in a similar way explains or even mentions the potential of miRNA modeling into the metabolic network. Therefore, the question is whether miRNA regulation mechanism could improve those already published large-scale human

metabolic network with regard to therapeutic intervention including drug discovery and drug response prediction.

Since many studies indicate the essential involvements of miRNAs in many metabolic processes such as lipid metabolism (Bernard et al., 2007; Lynn, 2009), glucose metabolism (Lynn, 2009), adipocyte differentiation (Kajimoto et al., 2006) and others, I am interested in investigating if the addition of miRNA-regulation network into the human metabolic network would help us to better understand how different types of cancer cells could adapt their metabolism to satisfy their own growth and proliferation demand; further to understand the complex implications in cancer cells, where drugs, such as anti-metabolites, that target nucleotide biosynthesis are used to counteract a variety of high cellular malignancies originated from disordered metabolism.

1.5 GENE-REGULATORY NETWORK

Gene-regulatory network includes cellular interactions between different groups of transcription factors and a collection of DNA segments of different genes. In general, most of the transcription processes are regulated directly or indirectly by corresponding transcription factors that bind to the promoter region of a DNA sequence to invoke or repress the transcription processes (MacNeil & Walhout, 2011). In single-celled organisms, gene-regulatory network can directly respond to the environmental stimuli and optimize the cellular activities by regulating transcriptional processes of diverse relevant genes for the survival in new environmental conditions (Blais & Dynlacht, 2005). However, in multicellular, highly developed organisms such as the human, the gene-regulatory network is coupled with signaling pathways and metabolic pathways. Its complexity has far exceeded its predecessor, for instance, many downstream components (such as ERK, JNKs) of signaling pathways can trigger the activation of gene-regulatory network by activating different transcription factors (such as AP-1, NfkB), which in turn transactivate many other genes to enhance the signal transduction intensity of different

33

signaling pathways. This way the positive feedback-loop can be established to enhance cellular activities. In a similar way, different negative feedback-loops can be formed to restrict certain cellular activities. In sum, the gene-regulatory networks play the key role in coupling with diverse cellular processes, function as carrier to translate message into the level of gene expression, and its deregulation leads to severe pathological states of living cells.

1.6 SYSTEMS BIOLOGY & SOFTWARE

From the days of Nobert Wiener onwards, the system-level understanding of biological systems became a long standing goal of biological science (Wiener, 1948). Unfortunately, at Wiener's time, only a phenomenological study was possible. In the past few decades, the study field of molecular biology with the rapid progress of genome sequence projects and a range of other molecular biology projects led to the accumulation of in-depth knowledge of molecular nature of biological systems. Therefore, now we are standing face to face with the golden opportunity to seek out the possibility of the system-level understanding solidly grounded on the molecular-level knowledge. However, the system-level understanding requires a shift in our notion from "what to look for" in biology to "the whole is better than the sum" in systems biology. It is widely known that a biological system is not just an assembly of genes, RNAs and proteins and that its properties including dynamic behavior cannot be fully understood merely by being clustered together (Kitano, 2002). Thus, it is desirable to get to know how those genes, RNAs and proteins are gathered together to form the structure of a cellular system. To say it in other words, insight into diverse cellular interactions of protein-protein, gene-protein and RNA-protein, which includes signaling network, metabolic network, gene-regulatory network and miRNA regulation network, will be enlightening and essential to understand the nature of a biological system.

34

Systems biology with the characteristics of synergistic integration of theory, computational modeling and experiments, has its primary goal in comprehensively gathering data from different levels of biological systems and subsequently to integrating this data so that diverse predictive computational models can be generated. Those types of models can serve as references for understanding the biological systems' behavior, guides for interpreting experimental results, and powerful ways to predict the dynamic behavior of cellular systems (Kitano, 2002). Therefore, with the system-level understanding of the dynamic behavior of cancerous cellular systems, one might be able to understand the causal connection between the cancerous genotype and the pathological phenotype and those systems-biological approaches are indeed helpful for cancer research.

To design and model cellular systems involving gene expression, gene regulation, signal transduction, vesicular transport and metabolism is not an easy task, which usually requires persistent effort and continuous attention. A common modeling approach for construction of cellular system consists of three steps: 1.) conducting an in-depth literature research including defining biochemical reactions and pathways from publicly available databases such as Reactom (Vastrik et al., 2007), Kegg (Kanehisa et al., 2000) and others; 2.) performing wet lab experiments to measure reaction kinetic parameters and their mechanisms; deriving appropriate kinetic equations; and estimating values of kinetic constants. Afterwards, when these data acquisition activities have been accomplished, 3.) the model construction including object and reaction definition (such as gene, RNA, protein, complex, compound, phosphorylation, degradation, complex formation, methylation), object compartmentalization (such as extra cellular region, cytoplasm, nucleoplasm) and pathway classification (such as EGFR signaling pathway, FGF signaling pathway, WNT signaling pathway). Currently, modeling of cellular system can be realized by applying softwares such as CellDesigner (Funahashi et al., 2003),

35

Copasi (Hopps et al., 2006), PyBioS2 (Li et al., 2012b), PySCeS (Olivier et al., 2004) and Virtual Cell (Slepchenko et al., 2003).

1.6.1 **CellDesigner**

The primary goal of the software CellDesigner is to provide users a process diagram editor with standardized annotations and technologies for the model construction, so that the information and knowledge presented by a model can be used repeatedly, improved and extended (Funahashi et al., 2003; Kitano et al., 2005). CellDesigner has a SBML compliant (http://sbml.org) for receiving/producing a simple, flexible text format in exchange for a wide variety of model data. At the same time, CellDesigner is a SBW-enabled application (http://sbw.kgi.edu), which enables all SBW-enabled modules to be integrated into this software. The simulation engine of CellDesigner is based on ODE. The software is implemented in java (http://celldesigner.org) and can be operated in different platforms such as Windows, Linux and MacOS X. This software has many advantages: one of them is the comprehensive graphical presentation, with which a clear and high-resolution network graph can be generated. The source-code of this software is written in java and the software is free of charge. The speed of ODE solver in cellDesigner is very satisfactory. The disadvantages of CellDesigner are that it does not provide different standard structural and dynamic analysis approaches (such as cluster analysis, in- and out-degree analysis and metabolic control analysis). Furthermore, an established model in cellDesigner cannot be reused, which means that an established model cannot be split into different sub-modules or integrated into anther model. This software has its limitation to deal with large-size networks and does not provide the user with the ability to organize his/her model archive.
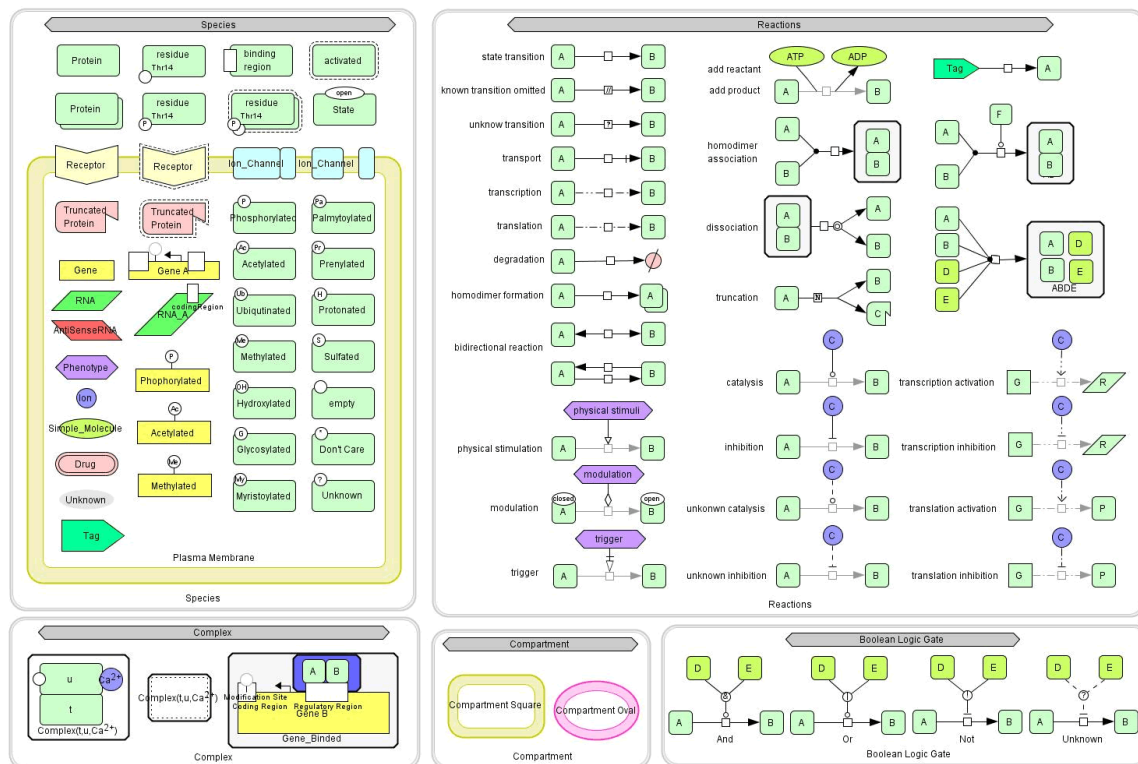
Figure 1: CellDesigner Annotation Interface.

1.6.2 **Copasi**

This software is a stand-alone program for modeling and simulation of complex biochemical processes. It possesses two types of executable version: a graphical user-friendly interface (CopasiUI) and a command line version (CopasiSE) containing only a calculation engine. Copasi is equipped with different simulation approaches (such as deterministic-, stochastic- and hybrid approach) and a number of diverse optimization algorithms for coping with the model analysis (Hopps et al., 2006). Furthermore, Copasi

provides many different analysis options including steady state calculation, stoichiometric network analysis, sensitivity analysis, parameter estimation and others. The complete software is available in binary code (http://www.copasi.org) for Windows, MacOS X, Linux and Sun Solaris. This software has many advantages, which includes handling complex tasks such as sensitivity analysis and parameter estimation, and a frequent update with strong background technical support. Furthermore, this software is facilitated with diverse biochemical and mathematical analysis approaches. The disadvantages are that Copasi is limited in the features, that Copasi itself provides. The user is not given the opportunity to develop his/her own feature within Copasi.

COPASI (4.0 Build 8)

#	Name	Equation	Kinetics	Flux
1	HXT	GLCo = GLCi	HXT kinetics	0
2	HK	GLCi + ATP = G6P + ADP	HK kinetics	0
3	PGI	G6P = F6P	PGI kinetics	0
4	PFK	F6P + ATP -> F16bP + ADP; AMP F26bP	PFK kinetics	0
5	ALD	F16bP = DHAP + GAP	ALD kinetics	0
6	TPI	DHAP = GAP	Mass action (reversible)	0
7	GAPDH	GAP + NAD = BPG + NADH	GAPDH kinetics	0
8	PGK	BPG + ADP = P3G + ATP	PGK kinetics	0
9	PGM	P3G = P2G	PGM kinetics	0
10	ENO	P2G = PEP	ENO kinetics	0
11	PYK	PEP + ADP = PYR + ATP	PYK kinetics	0
12	PDC	PYR -> AcAld + CO2	PDC kinetics	0
13	ADH	EtOH + NAD = AcAld + NADH	ADH kinetics	0
14	ATPase	ATP -> ADP	ATPase	0
15	AK	2 * ADP = ATP + AMP	Mass action (reversible)	0
16	G3PDH	DHAP + NADH -> Glycerol + NAD	G3PDH kinetics	0
17	Glycogen Branch	G6P + ATP -> ADP + Glycogen	Glycogen synthesis kinetics	0
18	Trehalose Branch	2 * G6P + ATP -> ADP + Trehalose	Trehalose synthesis kinetics	0
19	Succinate Branch	2 * AcAld + 3 * NAD -> Succinate + 3 * NADH	Succinate kinetics	0
20				

Copasi
▼Model
  ▼Biochemical
    ▶Compartments
    ▶Metabolites
      Moiety
    ▶Reactions
  ▼Mathematical
    Compartment Symbols
    Function Symbols
    Constant Symbols
    Fixed Metabolite Symbols
    Metabolite Symbols
    Differential Equations
▼Tasks
  ▶Steady-State
  ▼Stoichiometry
    Elementary Modes
    ▶Mass Conservation
  ▼Time Course
    Result
▼Multiple Tasks
  Optimization
  Scan
  ▶Fitting
▼Output
  Pathway
  Plots
  Reports
▶Functions
  Preferences

Commit   Revert   Delete/Undelete   New

Figure 2: Copasi User Interface.

The option of mathematical modeling within Copasi is not clear like a black box. In addition, modeling large networks by using Copasi is highly time-consuming and the same applies when an analysis-approach is carried out for a large size network. Copasi is not facilitated with an internal database, which makes it difficult for the user to organize his/her model archive.

1.6.3 **PySCeS**

PySCeS is developed to provide users a flexible and comfortable interface, which utilizes and extends the low-level capabilities derived from Python and SciPy. The main procedure of this software is to create/enter a biological model in terms of its stoichiometry and rate equation. Based on this information, the subsequent creation of differential equation will be processed automatically, so that different model-analysis approaches including steady-state calculation and bifurcation analysis can be performed according to the differential equation system derived from a model (Olivier et al., 2004; http://pysces.sourceforge.net). One of the advantages of this software is that its implementation language is python, a modern programming language. Python is famous for automatic memory management and flexible integration with other languages and possesses diverse libraries and packages. Therefore, PySCeS is well suited to accept many external applications that are developed with other programming languages. The disadvantages of PySCeS are that it is a console-based application and does not provide a graphic user interface, which is not user-friendly during the model construction. In addition, PySCeS is not facilitated with a SBML compliant and models created within PySCeS are not compatible to softwares with a SBML compliant. This software can handle models in small-size and medium size well, however, it has the limitation with regards to large-size models. PySCeS does not possess its own database, which makes it difficult for the user to organize his/her model archive.
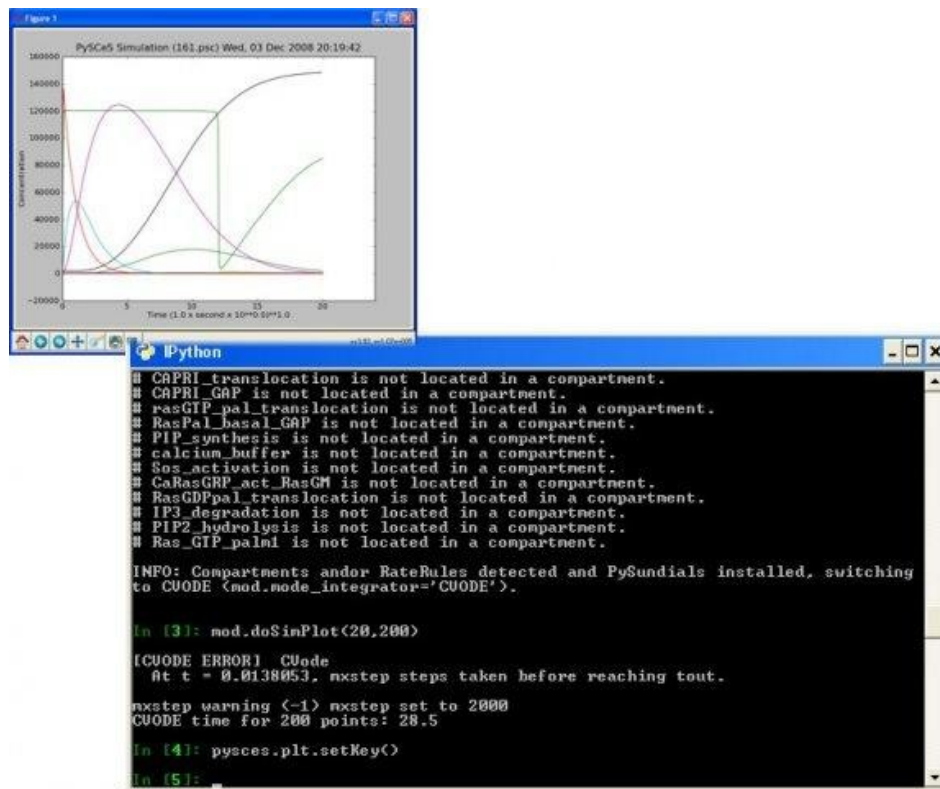
Figure 3: PySCeS Command Line and Output Graph.

1.6.4 **Virtual Cell**

Virtual Cell is a modular computational framework for design, construction and simulation of cellular processes. Its key features include the incorporation of realistic experimental geometries within a full 3D spatial model, application of diverse numerical solvers to perform simulation and different sophisticated analytical approaches on the simulation results. Virtual Cell is implemented in java and provides a user-friendly java interface, which can be accessed via a web browser and provides options including database access, geometry definition, specification of compartment topology, species definition and assignment, cellular reaction input and computational mesh. Its primary

user-group are cell biologists and mathematical biologists, who use it for interpretation and planning of biological experiments (Slepchenko et al., 2003; http://www.life.illinois.edu/plantbio/cell). The advantages of this software are that model annotation can strictly follow Minimal Information Requested In the Annotation of biochemical Models (MIRIAM) (Novère et al., 2005). It provides an easy option for the user to compare models defined within this software. Virtual Cell also provides multiple non-spatial applications and mapping structures. One important advantage of Virtual Cell is the access to many publicly available databases including (Reactome, BioGRID, HumanCyc, IntAct and MetaCyc) in order to import relevant biochemical information. The disadvantages of Virtual Cell are that it separates the definition of BioModels and MathModels which leads to the separation of mathematical modeling from the biological context and makes it more difficult for the user to perform model construction. The speed
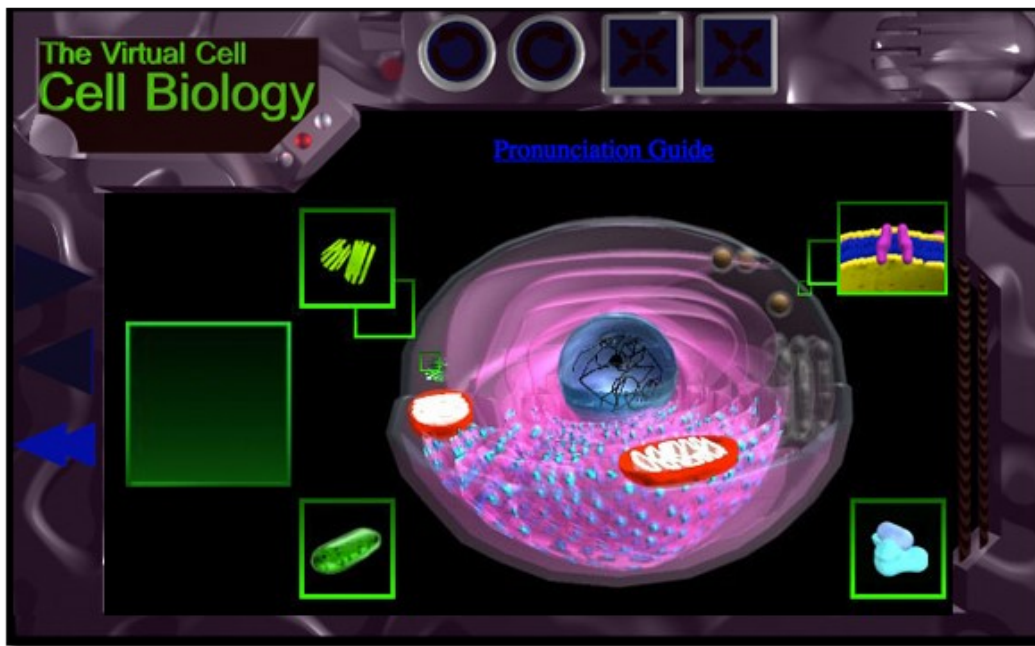
Figure 4: The Visualization of Virtual Cell Interface.

41

of run simulation and many other analysis-approaches including parameter estimation are not satisfactory. Construction of large-scale network using Virtual Cell is highly time-consuming.

1.6.5 **PyBioS2**

PyBioS2 is an advanced web-based software platform for the design, modeling and simulation of genome-scale cellular networks and has been developed during the course of writing this thesis (Li et al., 2012b; https://pybios.molgen.mpg.de). In the result part of this thesis, the detailed functionalities of the software will be explained. Its primary goal is to enable users to construct genome-scale networks with consistent annotations and within a short period of time. Therefore, PyBioS2 is connected with more than 30 different publicly available databases (Table 1). Moreover, PyBioS2 is the only software in the world, which possesses two types of simulation engines: ODE and Petri-net, to cope with diverse simulation conditions and different types of models. PyBioS2 also provides a unique modeling option: a semi-automatic reaction definition to reduce tedious common definition work, to improve construction efficiency and to enhance the consistency of model data. PyBioS2 is facilitated with a number of model analysis approaches including steady-state calculation, connectivity analysis, parameter scan and network-cluster analysis. The advantages of PyBioS2 are that it possesses access to multiple publicly available databases (listed in Table 1) to easily import relevant or necessary biological information. Models defined in PyBioS2 have the same unique and consistent structure pattern so that they can conveniently be re-used, e.g. merged and split. One important advantage of this software is that it uniquely provides the semi-automatic option for designing and constructing model, which improves efficiency during model construction and is especially suitable for constructing a large-scale network. PyBioS2 also provides a user-friendly graphic interface and has its specifically defined internal database to reduce redundant information via import process and enables the user

Database			
BioCarta*	BioGRID*	BioMart	BioModels
Bind*	ChEBI	Corum*	Ehmm*
DIP*	Ensembl	HumanCyc*	HPRD*
Innatedb*	Inoh*	Intact-ss*	Intact-Is*
IntAct*	KEGG*	Maxtrixdb*	miRBase
Mips-mppi*	MINT*	NetPath*	Pdzbase*
Pharmgkb*	PID*	PIG*	Signalink*
Smpdb*	SPIKE*	Reactome	UniProt

Table 1: PyBioS2 Database Connection. PyBioS2 is connected with different publicly available databases. ('*' indicates the connection via CPDB (Kamburov et al., 2011))
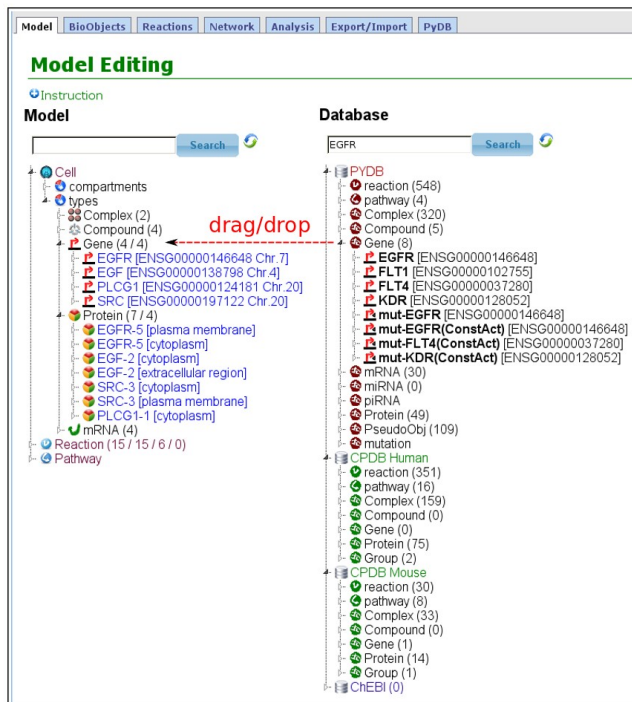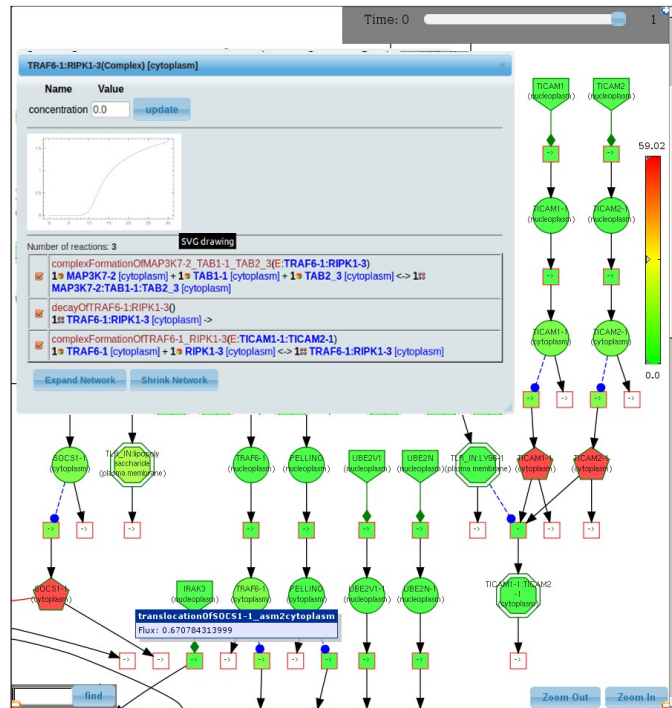
Figure 5 A: PyBioS2 User Interface. Hierarchical tree-like structure of the model

and the pathway databases. Models can be developed by drag-and-drop of components, reactions, or entire pathways from the databases into the model . **B**: The Network Visualization of a Model. Simulation results can be displayed directly into the model network graph, where a slider allows stepping through the time course data. (the figure and its legend is taken from Li et al., 2012b)

to comfortably organize their model archive. The disadvantage of PyBioS2 is that its source-code is not publicly available, therefore this software is limited by the features it provides. Furthermore, the current version of PyBioS can only accept the biological data from human and mice by import process.

1.7 OBJECTIVE AND FOCUS OF THIS THESIS

The primary goal of this thesis is to apply the latest knowledge from the field of cancer research to the design and construction of an effective computational framework. Within this framework, a 'virtual patient' model will be developed and validated to precisely understand the cancer development and therapy at the individual level. This thesis also aims at demonstrating that computational approaches can possess huge potential for the development of personalized medicine. With the application of this 'virtual patient' model combined with an effective *in silico* simulation strategy, personalized medicine including individual drug-targeted treatment, individual biomarker discovery and individual identification of optimal drug combination, should be made possible.

2. Results

2.1 ESTABLISHMENT OF AN INTEGRATED MIRNA-EGFR SIGNALING PATHWAY MODEL

During my research work (Li et al., 2012a), the miRNA-EGFR (ME) signaling model was constructed using the PyBioS2, a web-based modeling and simulation software (Li et

al., 2012b; http://pybios.molgen.mpg.de). Fig. 6A depicts the general overview of the EGFR signaling pathway of the model. It contains 1241 reactions and 901 entities and table 2 gives a statistical overview of the ME model components. The miRNA annotations with their corresponding target information are derived from miRBase (Griffiths-Jones et al., 2008) as well as extensive literature research. Besides common implemented component type such as gene, protein and RNA, there are also gene-set, protein-set and RNA-set for the simplification of modeling procedure. For instance, a gene-set entity "MEK" is defined as a set of genes including MAP2K1 and MAP2K2 that are functional closely related mitogen activated protein kinases. A miRNA-set entity "mir-TRDD" is defined in the model and presents 5 miRNAs (mir-631, mir-608, mir-604, mir-492, and mir-30a) that possess same targets TARBP2, RNASEN, DICER1 amd DGCR8 (Cummins et al., 2006; Clague et al., 2010; Ye et al., 2008). The ME model was constructed with a great emphasis on the basic biological relationship, meaning that each protein entity (or protein-set entity) is produced in the compartment cytoplasm by the translation reaction of a respective mRNA entity (or mRNA-set entity), which is generated in the compartment nucleus by the transcription reaction of a gene entity (or gene-set entity) (Fig. 6B).

Component	No.	Reaction	No.
gene	183	transcription	179
mRNA	78	translation	75
protein	133	decay	218
miRNA	241	complex-formation	115
compound	35	translocation	133
complex	131	phosphorylation	49
pseudo-object	100	dephosphorylation	40
		activation	10
		miRNA-binding	417
		other types	5
Sum:	901	Sum:	1241

Table 2: The summary of miRNA-EGFR model components / reactions. Compound is metabolite; complex includes protein-protein complex, protein-gene-complex, mRNA-miRNA complex; pseudo-object includes protein inhibitor and miRNA inhibitor. (this table and its legend is taken from Li et al., 2012a)
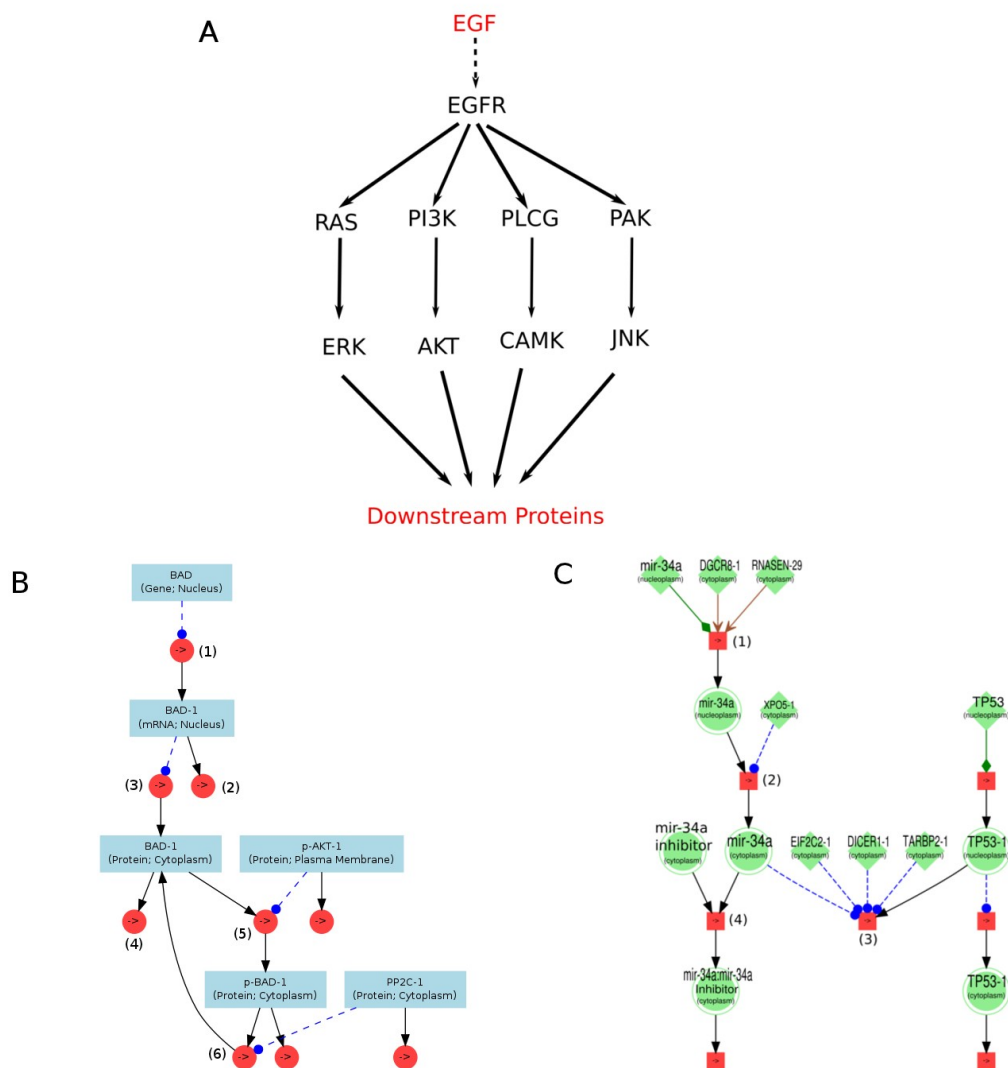
Figure 6 Overview of the miRNA-EGFR model network. A: The model of the EGFR signaling pathway contains 4 sub-pathways: MAPK pathway, PI3K-AKT pathway, Ca²⁺ signaling pathway, PAK signaling pathway. **B, C** are parts

of the model network visualized by PyBioS2. **B**: The mRNA BAD-1 is produced by the transcription reaction (1) of gene BAD and also takes part in decay reaction (2) and translation-reaction (3), which produces the protein BAD-1 (cytoplasm). The protein BAD-1 takes part in further three reactions that are the decay-reaction (4), phosphorylation reaction (5) and dephosphorylation reaction (6). The phosphorylated protein P-AKT (plasma membrane) catalyzes the phosphorylation reaction, in which protein BAD-1 is phosphorylated into P-BAD-1. Afterwards, the protein P-BAD-1 is then dephosphorylated; **C**: shows a simplified miRNA biogenesis, target recognition and competitive anti-miRNA effect. (1) miRNA-gene transcription; (2) miRNA translocation (from nucleus into cytoplasm); (3) miRNA-binding-target reaction; (4) miRNA binds to the miRNA inhibitor. (This figure and its legend is taken from the Li et al., 2012a)

It is noteworthy that during *in silico* simulation (explained in the Materials and Methods 3.1), the generated signal flow within the model is starting at the transcriptional level. In model, the gene transcription is defined as a one-step process assuming that basal transcription is generated by constitutive action of a single transcription factor, whereas the miRNA gene transcription simplifies two processes: (i) miRNA gene transcription catalyzed by DNA-polymerase II or DNA-polymerase III; (ii) cropping of the primary transcript (pri-miRNA) into a hairpin intermediate (pre-miRNA) by the nuclear 650 kDa microprocessor complex, comprising of the RNase III DROSHA (RNASEN) and the DiGeorge syndrome critical region gene 8 (DGCR8) in humans (Davis et al., 2008) (see Fig. 6C). The miRNA binding target reaction simplifies two processes: (i) mature miRNA forms the complex with DICER and TARB, then this complex associates with the Ago-complex and converts into the RNA-induced silencing complex (RISC); (ii) the complex RISC recognizes the target mRNA and binds to it (see Fig. 6C) (Li et al., 2012a). This way the modeling of miRNA biogenesis has been intensively addressed during the construction of the ME model.

In addition, distinct reactions defined in the ME model are involved in specific functions, such as signaling transduction of small GTPase signaling, MAPK cascade, phosphatidylinositol signaling, and Ca^{2+} signaling (Li et al., 2012a). The detailed model

47

information including literature reference is available in the appendix A.1 as a form of SBML (Hucka et al., 2003).

2.2 VALIDATION OF THE PREDICTIVE ABILITY OF THE MIRNA-EGFR MODEL

After the ME model has been accomplished, a validation approach was designed to test the predictive power of this model. The center ideal is to incorporate the genetic information (such as gene-expression data) into this model, then to perform *in silico* simulation (in this case, Petri-net simulation), afterwards the simulated result should be compared with experimental established result to investigate the predictive ability of this model. If it exist any inconsistency in this comparison, it indicates that the quality of model construction is not good enough. If simulated result act in line with experimental established result, then the model can be applied to perform prediction that could be tested by experiments, to explore interesting biological issues that are not directly amenable to experimental inquiry (Li et al., 2012a).

A recent study (Avraham et al., 2010) was conducted to investigate the miRNA's impact on the EGFR signaling pathway in MCF10A cells. The authors over-expressed each of 23 miRNAs (including mir-134, mir-212, mir-320 and others) and then measured the concentration of multiple immediated early genes (IEGs) to estimate the activity of the EGFR signaling pathway (Fig. 7A). In order to investigate whether the ME model could reflect the same or similar molecular activities as MCF10A cells did, the individual over-expression of the same 23 miRNA were set as initial conditions for the *in silico* simulation (Petri-net simulation, see Materials and Methods 3.1). It is noteworthy that the ME model construction was independent from the study of Avraham et al., (2010) and no data from this study was feed into the ME model. The simulation result is displayed in the Fig. 7B. The comparison of both results (Fig. 7A and 7B) indicates that the simulated result acts in agreement with the experimental established result, which implies that the literature information of EGFR signaling pathway and miRNA-target relationship used

48
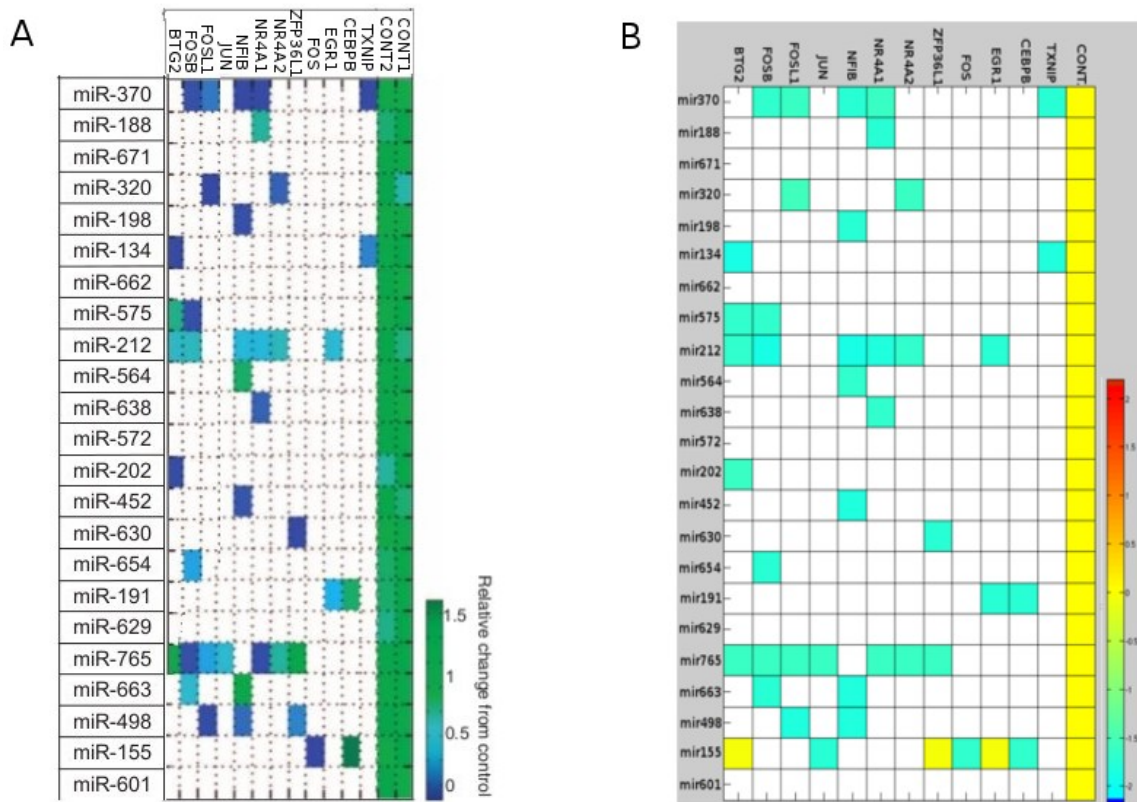
Figure 7. Comparison of miRNA-EGFR model predictions with experimental results. A: Experimental result of relative concentration changes of target mRNAs according to individual miRNA over-expression experiments from Avraham et al., (2010). (This figure was taken from Avraham et al., (2010) with the permission from the journal of Science Signaling. Its copy right remains in this journal); **B**: *in silico* prediction result of relative concentration changes (log2-ratio) of target-mRNAs according to each miRNA over-expression in the EGFR model. Both heatmaps show very similar qualitative results (protein down-regulation), the only discrepancies are for miR-155 and miR-498. Therefore, the comparison shows the high accuracy prediction of my model. mRNAs with low concentration changes (log2-ratio < 0.001) are ignored and shown in 'white'. (This figure and its legend is taken from the Li et al., 2012a)

for the ME model construction was proven to be sufficient to qualitatively recapitulate the IEG transcriptional response to a cell's miRNA expression status in the context of an

49

activated EGFR pathway (Li et al., 2012a). Moreover, this comparison also indicates that the ME model pass the initial validation and can be used for further validation experiments.

2.3 SIMULATION OF FUNCTIONAL ABILITY OF MIR-192 AND MIR-181

Different studies provide evidence that the expression levels of miRNA can be correlated with the developmental stages of various human cancers, which indicates the relevant functional roles of miRNA as tumor suppressors and oncogenes (Chen et al., 2005, Hwang et al., 2006). For instance, the down-regulation of miRNAs including mir-15a, mir-16 and let-7 are often observed in different types of cancer suggesting their tumor suppressor roles (Hwang et al., 2006). The dramatic up-regulation of mir-21 has been observed in multiple cancers indicating its oncogenic role (Liu et al., 2010; Zhang et al., 2012). Based on results of diverse studies, mir-192 tightly control the expression of following targets: MDM2 (Braun et al., 2008), EGFR (Yantiss et al., 2009), PIK3CA (Yantiss et al., 2009), TP53 (Braun et al., 2008; Georges et al., 2008;), PTEN (Kato et al., 2009), and CDKN1 (Braun et al., 2008). Therefore, an in silico experiment was designed during the study of Li et al. (2012a) to investigate the functional role of mir-192 and its impact on this signaling pathway. Considering the fact that expression level of an individual miRNA can vary widely in copy number per cell, with a few tissue-specific species up to more than 10,000 copies per cell (equivalently more than 20,000 nM) (Liang et al., 2007), the initial expression level of mir-192 varied between 0 and 10,000 nM during *in silico* simulation (Petri-net simulation, see Materials and Methods 3.1). According to the study of Subramanian et al. (2008), the number of copies of an individual mRNA present in a single cell is considered to vary over four orders of magnitude (1 to > 1,000 copies that is equivalent to 0.001 to 10 nM), with most mRNA species present in < 100 copies and a few exceeding 1,000 copies. Hence, in order to ensure a lasting and strong signal source for signal transduction of EGFR signaling

50
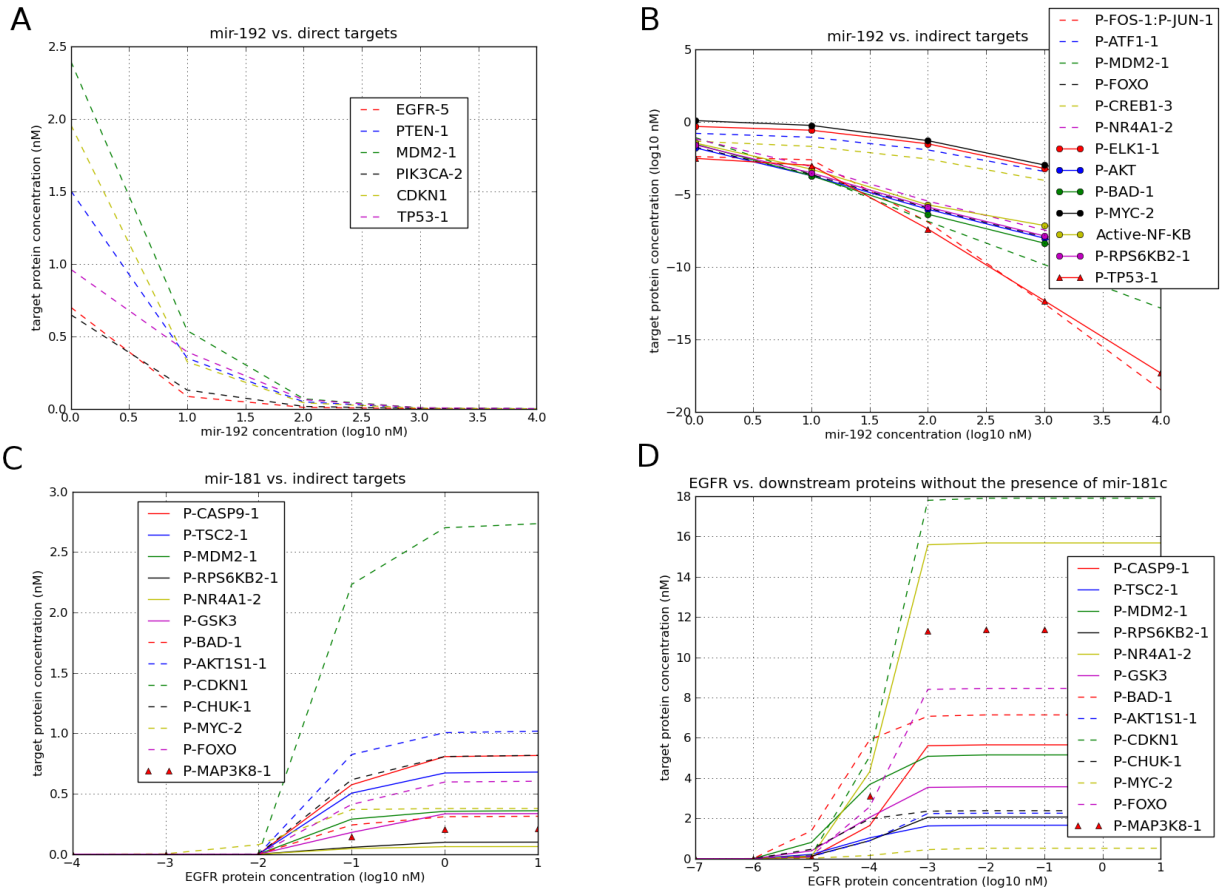
Figure 8. Modeling of mir-192 and mir-181c effects on the EGFR signaling pathway. A: Increased expression of mir-192 gene corresponds with a reduced level of targets' protein expression; **B**: Concentrations of EGFR downstream activated proteins are inversely correlated with the mir-192 gene expression level; **C**: simulation result with fixed mir-181c gene expression level (1 nM), whereas all other 240 miRNA genes are not expressed. All 13 AKT-dependent proteins can only be activated after the concentration of EGFR protein is getting higher than 10 pM. This activation threshold is due to the presence of mir-181c. **D**: represents another simulation result without the presence of mir-181c, additionally all other 240 miRNAs were not expressed, as well. At this condition, the activation threshold of these proteins is at 0.001 pM. By comparing these two results (**C**, **D**), one can understand that mir-181c raises the activation threshold of the EGFR signaling pathway significantly. (This figure and its legend is taken from the Li et al., 2012a)

pathway, the extracellular ligand protein EGF was fixed to 1 nM. The initial concentration of other entities including mRNAs, miRNAs, proteins, complexes, etc. were set to 0 nM. This way the signal effect caused by EGF ligand within this EGFR signaling pathway during the *in silico* simulation would be clear and contain no background noise information (Li et al., 2012a).

The effect by increasing the mir-192's expression level during *in silico* simulation shows the stepwise concentration decrease of its target proteins including TP53, PTEN, MDM2 and CDKN1 (Fig. 8A). This function of mir-192 by simultaneous suppressing those putative tumor suppressors strongly indicates its oncogenic role to provide a proliferative advantage for cancer cells, which is supported by the observation that upregulation of mir-192 often occurs in cancers (Georges et al., 2008). Hence, this result reveals the oncomir role of mir-192 and the first introduced oncomir definition is the study of Esquela-Kerscher et al. in 2006. The indirect effect of mir-192's over-expression can be elicited through the continuous decrease of multiple downstream protein components within the EGFR signaling pathway (Fig. 8B). This result indicates through regulation of the expression level of the EGFR receptor, mir-192 is able to control the signal intensity of this signaling pathway, which is reflected by signal reduction within the entire signaling pathway. This illustrates the impact of a given miRNA on a signaling pathway by directly targeting its receptor protein (Li et al., 2012a).

Considering both results (Fig. 8A and 8B) together, one might think that the oncogenic function of mir-192 is partially hampered by negatively regulating the EGFR protein level. However, constitutive activating mutations of different signaling component of the EGFR signaling pathway, such as NRAS(Q61L) or BRAF(V600E) occurs often in cancers (Forbes et al., 2008), which probably circumvent the negative regulation of mir-192 by constantly sustaining the EGFR signaling intensity. Therefore, the clear oncomir effect of mir-192 via targeting TP53, PTEN, MDM2 and CDKN1 remains clear and impaired. Hence, it would be interesting to analyze the expression level of miRNA in

cancers with respect to the mutational status in individual cancers. In addition, it is noteworthy that among 241 miRNAs defined in the ME model, 44 miRNAs target TP53, 57 miRNAs target PTEN, and 10 target MDM2 (see appendix A.2). The mir-192 is one of few miRNAs in the model that simultaneously regulate these 4 tumor suppressors (Li et al., 2012a).

Another important signaling branch of the EGFR signaling pathway is the activation of protein kinase PI3K, which in turn activates the powerful protein kinase AKT. This putative protein kinase can then activate its numerous downstream targets including famous oncogene MYC: EGFR+EGF → PI3K → PIP3 → PDPK1 → AKT → target protein (MYC). According to results of early studies, mir-181c direct regulate the expression level of AKT (Androulidaki et al., 2009) and MYC (Wong et al., 2008), which sparks the interest to investigate the regulation ability and impact of this miRNA on the EGFR signaling pathway. During the study of Li et al. (2012a), another *in silico* experiment was designed to achieve this goal. This experiment is divided into two parts. In the first part, the initial condition defined that the expression level of mir-181c was fixed to 1 nM and EGF ligand concentration was also fixed to 1 nM to provide a stable signal source. All other model components were initialized with 0 nM. During the *in silico* simulation, the expression level of EGFR varied from 0.1 pM to 10 nM, while the only difference between the second part in compare with the first part is that the mir-181c is knocked out and all other initial conditions remain the same. Its goal is to investigate whether certain miRNAs that target key components of signaling pathway can tightly control the threshold of signal transduction process. During the *in silico* simulation of the first part, the signal threshold was around 10 pM of EGFR, as shown in Fig. 8C, from this receptor concentration on, most downstream components began to become more active, while the increase of EGFR receptor concentration. However, during the *in silico* simulation of the second part, the signal threshold was lowered at 0.001 pM (Fig. 8D), which strongly indicates the important contribution of mir-181c for fine tuning of EGFR

signaling threshold in a range from 0.001 pM to 10 pM of EGFR protein. Both results (Fig. 8C and 8D) reveal a crucial functional role of this mir-181c by optimizing signal efficacy and limiting undesired signal fluctuation (Li et al., 2012a). This is functional ability is especially relevant for genes expressed at low levels, because stochastic changes of different factors, for instance, the change of activation level of corresponding promoter, which leads to affect their expression level, would be more genetically prominent (Raser and Shea, 2010).

2.4 EVALUATION OF THE "ONE HIT – MULTIPLE TARGETS" CONCEPT USING MIRNA MODELING

Diverse studies provide evidence that a normal miRNA expression is essential for proper development and differentiation in a tissue and cell-type specific manner (Wurdinger & Costa, 2007; Wu et al., 2009). Therefore, deregulation of miRNA expression is a common phenomenon occurring in many types of cancer (Shi et al., 2010). However, related to this issue, one important question raises: whether the deregulation of miRNAs lead to a certain pathologic cellular state, or whether the particular cancer or disease is a direct cause of the deregulated expression of miRNAs (Li et al., 2012a). No study until now could answer this question. Nevertheless, one recent study utilizes the unique property (possessing multiple targets) of miRNA and developed a therapeutic concept, named "One Hit – Multiple Targets" (Wurdinger & Costa, 2007).

The center idea present in this concept of Wurdinger & Costa (2007) is simple and straightforward: change the expression of deregulated miRNAs back to corresponding normal levels, which can directly or indirectly influence their protein-coding targets in a positive way or even restore normal expression levels of indirect target proteins. This way multiple key cellular components such as relevant oncogenes and tumor suppressors can simultaneously be restored with their normal expression level, which can have immense therapeutic benefit. In order to validate the therapeutic benefit of this concept,

54

an *in silico* experiment were designed during the study of Li et al. (2012a) to individually access potential effect of anti-miRNAs related to the EGFR signaling pathway. Generally speaking, an anti-miRNA could be designed in the form of antagomir, which is an anti-miRNA oligonucleotides (AMOs) conjugating with cholesterol (Krutzfeldt et al., 2005); alternatively, an anti-miRNA could be designed as locked-nucleic-acid anti-sense oligonucleotides (LNAs) (Orom et al., 2006; Vester et al., 2004). In both cases, specific miRNA inhibitory function can be hampered through association of the corresponding anti-miRNA according to sequence complementarity. The *in silico* experiment in the study of Li et al. (2012a) investigates the inhibitor effect of 100 specific anti-miRNAs individually and the readout components consist of 19 key downstream protein components of the EGFR signaling pathway (Fig. 9). A control state is created by the *in silico* simulation under condition of inactivation of all these anti-miRNAs. Each line of the heatmap displayed in Fig. 9 symbolizes a comparison between the activation of individual anti-miRNA and control state to show the impact of corresponding anti-miRNA activation. Clear differences of functional inhibitory effect can be observed in this heatmaps. For instance, in Fig. 9B, anti-mir-221 can clear reduce negative regulation of mir-221, which leads to increase of its direct and indirect target proteins' concentration. This positive effect is now displayed as the activation of multiple downstream protein

A
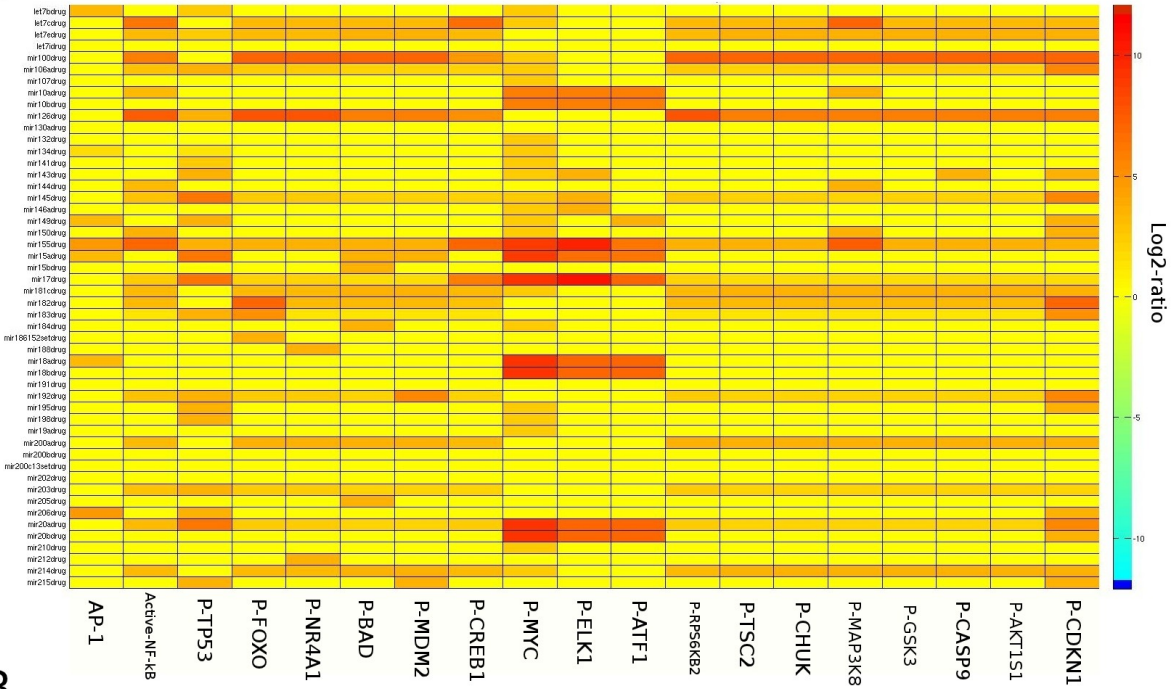
B
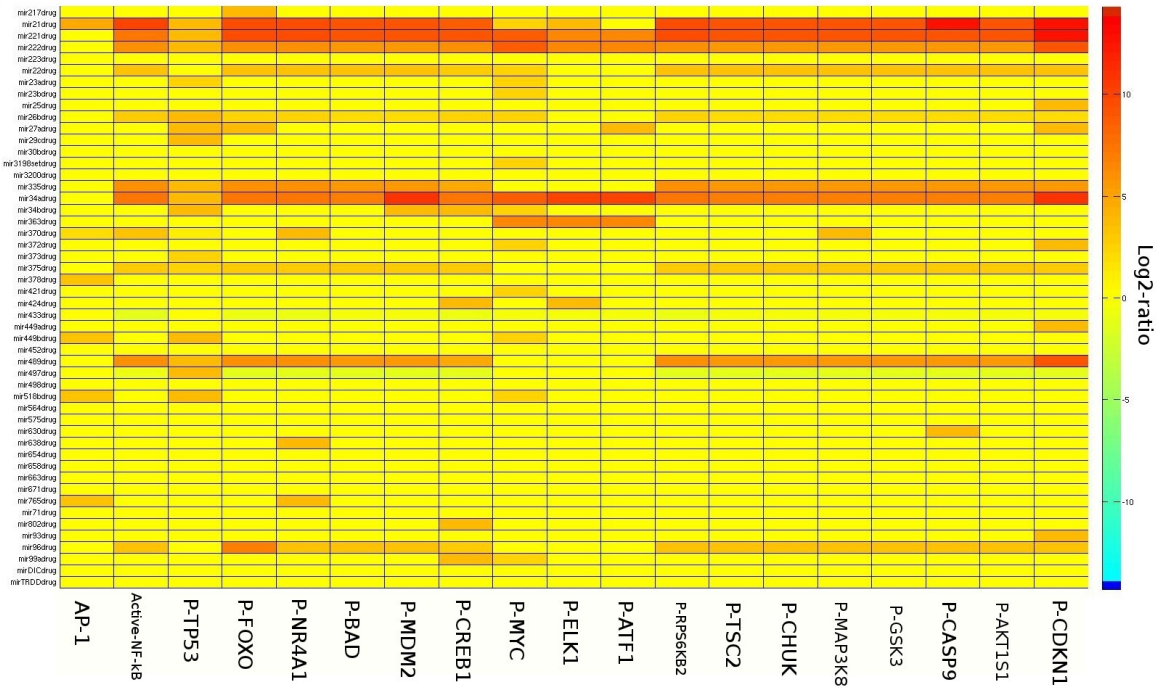
Figure 9. Modeling the individual effect of 100 anti-miRNAs. A and B: Each anti-miRNA is activated when the corresponding miRNA is overexpressed individually (a anti-mir-combination effect is omitted). Each row in this heatmap represents the predicted anti-miRNA effect and the more columns in the same row appears red or orange, the stronger is the predicted effect that this particular miRNA inhibitor can exert on the EGFR signaling pathway. In this manner, I can examine the impact of each anti-miRNA on this signaling pathway. (The simulation data is attached in the appendix C.4; the figure and its legend is taken from Li et al., 2012a)

components of the EGFR signaling pathway. In contrast, anti-mir-30b does not exert any impact on this pathway and all readout downstream components remain the same as in the control state (yellow, ratio = 0). Based on the simulation data from this *in silico* experiment, a statistical t-test was performed to analyze the global concentration change of model components due to each individual activation of 100 anti-miRNAs (see Materials and Methods 3.2). The aim of this statistical t-test is to find out anti-miRNAs with most significant impact on the EGFR signaling pathway. The top 15 anti-miRNAs with most significant p-value (<0.05) among 100 anti-miRNAs in the model were selected (Table 3). These 15 anti-miRNAs are considered to be more suitable as potential candidates for clinical cancer treatments, however, each of them needs first to be tested and verified in-vivo or in-vitro experiments.

Afterwards the miRNA-target relationship of the corresponding top 15 miRNAs and bottom 15 miRNAs whose anti-miRNAs have the least impact on the EGFR signaling pathway was analyzed. The target-number of those miRNA are countered individually, according to the information from appendix A.2 (Fig. 10). As expected, anti-miRNAs whose target miRNA possess more target proteins in the ME model, could exert greater impact on the signaling pathway. But anti-miRNAs whose target miRNA have less target protein, do not always have less impact. For instance, anti-mir-181c, anti-mir-214 and anti-mir-200a have have high influence on this signaling pathway, although the number of direct targets from these three miRNAs are low (Fig. 10A). The reasons for this

Figure 10. Histogram of miRNA/target relationship. A: The top ranked 15 miRNAs of Table 3 correlated with the amount of their corresponding targets in the EGFR signaling pathway (Table 2). B: The 15 miRNAs correlated with amount of their corresponding targets in EGFR signaling pathway. The anti-miRNAs of these miRNAs have least effect on this pathway (Fig. 9A and B). (This figure and its legend is taken from Li et al., 2012a)

phenomenon could be explained through the target relevance of miRNA. For instance, mir-181c negatively regulates the expression level of AKT (protein kinase B) and MYC (functional relevant transcriptional factor). Both play relevant roles in regulating cellular

Anti-microRNA	p-Value
Anti-mir-21	1.90e-06
Anti-mir-155	4.54e-06
Anti-mir-221	3.64e-04
Anti-mir-17	1.44e-03
Anti-mir-489	1.78e-03
Anti-mir-222	1.81e-03
Anti-mir-335	1.85e-03
Anti-mir-126	1.86e-03
Anti-mir-100	3.50e-03
Anti-mir-181c	1.17e-02
Anti-mir-214	1.18e-02
Anti-mir-200a	1.19e-02
Anti-mir-34a	1.26e-02
Anti-let-7c	3.22e-02
Anti-mir-182	4.22e-02

Table 3: Top 15 anti-miRNA Inhibitors. P-value gives the significant level of concentration-differences of all model components between the state

of the applied specific anti-miRNA inhibitor and control state (inactivated anti-miRNA inhibitor). (This table and its legend is taken from the study of Li et al., 2012a)

processes by controlling function activities of diverse cellular components (Fayard et al., 2005; Dang, 1999), which implies that mir-181c indirectly regulates a large number of cellular components. This phenomenon helps to understand the strength of the concept "One Hit – Multiple Target".

2.5 ESTABLISHMENT OF A NON-STEROIDAL ANTI-INFLAMMATORY DRUG MODEL

To date, it has been widely recognized that Non-Steroidal Anti-Inflammatory Drugs (NSAIDs) can exert considerable anti-neoplastic and anti-tumor effect with regard to many types of cancer such as colon (Sandler et al., 2003), lung (Randall et al., 2007), prostate (Wynne & Djakiew, 2010), head-and-neck (Jayaprakash et al., 2006) and stomach (Uefuji et al., 2000). It was estimated that the regular use of NSAIDs for a 10- or 15-year-period can reduce more than 40-50% of colon cancer occurrence (Smalley & DuBois, 1997). Moreover, in the United States alone, more than 20 billion aspirin (1st generation NSAID) tablets are purchased annually, and that more than 1% of the world population consumes at least one aspirin tablet daily (Lichtenberger LM., 2001). Unfortunately, the long-term use of NSAID can have a highly increase risk of side effect including abdominal pain, peptic ulcer and gastrointestinal bleeding, dyspepsia and others (Wolfe et al., 1999). Therefore, it sparks a lot interest and enthusiam to investigate the molecular mechanism, physiological and pathological functions of NSAID to reduce side effect and increase the drug efficiency. To the present day, a number of studies have been conducted to understand the NSAID-related mechanisms and functions. For instance, Dannenberg and Zakim (1999) conducted research to study the molecular function of first generation of NSAID, which inhibits the functions of cyclooxygenase (COX) 1 and 2. Both kinases are key enzymes playing key role of prostaglandin biosynthesis,

59

furthermore, they discovered relevant biological functions of prostaglandin and their derived products; Fosslien (2000) studied the functional activity of COX-2, which is only detectable in pathologic tissue and can be strongly induced by oncogenes, growth factors, cytokines and others. The function of COX-2 is highly connected with carcinogenesis. Subsequently, many studies discovered that COX-2 related signaling pathway has crosstalk and synergistic effect with other signaling pathways including epidermal growth factor receptor (EGFR)-signaling (Yoshimoto et al., 2002), nuclear receptor signaling (Wang et al., 2004), nuclear factor of kappa light polypeptide gene enhancer in B-cells (NfkB)-signaling (Poligone & Baldwin, 2001), rat sarcoma (Ras)-mitogen activated protein kinase (MAPK) signaling (Husain et al., 2001; Wong et al., 2004), vascular endothelial growth factor (VEGFR)-signaling (Iniguez et al., 2003), janus kinase/signal transducer and activator of transcription (JAK-STAT)-signaling (Shreedhar et al., 1998) and others.

Thanks to diverse studies, a large amount of molecular information related to NSAID has been produced. Hence, an integrative systematic approaches is needed, which is able to utilize these valuable information to analyze NSAID effect at individual level. In my recent study, the available molecular information was utilized to construct a molecule-based model and it is named the NSAID model. The aim of this study is to investigate whether the modeling of NSAID's molecular mechanisms can help to perform NSAID application at the individual level (Li & Mansmann, 2013). The model containing 3874 components and 6398 biochemical reactions. The definition of biochemical reactions to link different components including gene, RNA, protein within the NSAID model is similar to the ME model. Table 4 gives a statistical overview of the NSAID model components. The center part of this model is the COX-pathway (Fig. 11). This pathway starts with biosynthesis of COX genes. Their protein products can bind to available arachidonic acid in order to transform it into prostaglandin G_2 (PGG_2), which is further transformed into the unstable intermediate PGH_2 by the same enzyme. The unstable

60

Component	No.	Reaction	No.
gene	766	transcription	1145
mRNA	1530	translation	572
protein	1003	decay	1796
miRNA	18	complex-formation	363
compound	44	translocation	995
complex	603	phosphorylation	713
pseudo-object	21	dephosphorylation	212
siRNA	1	activation	240
		miRNA-binding	721
Sum:	3986	Sum:	6757

Table 4: The statistical summary of the NSAID model. Complex includes protein-protein complex, protein-metabolite complex, protein-mRNA and protein-gene complex. Compound is metabolite. PseudoObj includes protein-inhibitors (ABL inhibitor, AKT inhibitor, EGFR inhibitor, MEK inhibitor and other) and hallmarks including Evading Apoptosis, Angiogenesis, Proliferation and Tissue Invasion. (the table is taken from Li & Mansmann, 2013)
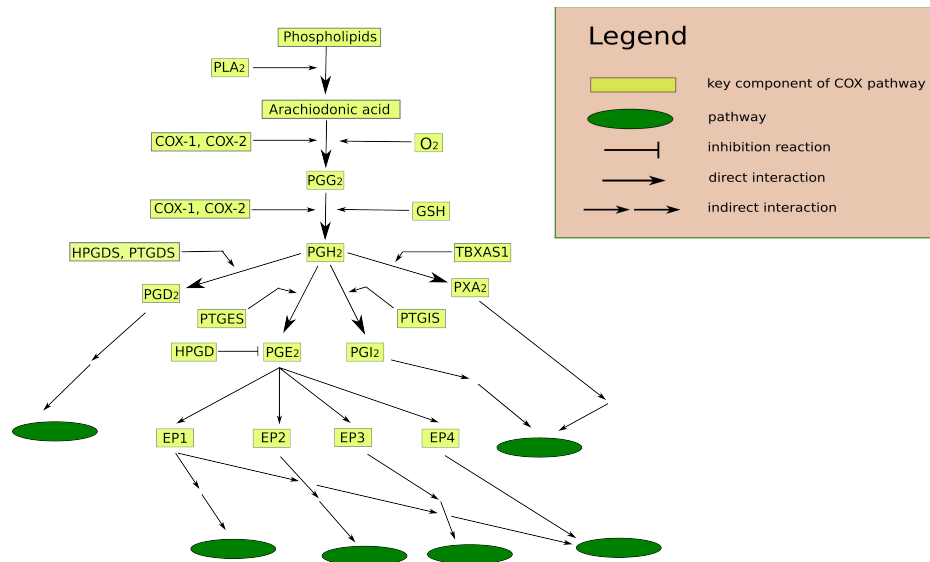
Figure 11. The network graph of COX pathway.

intermediate PGH_2 is converted into different prostaglandins such as PGE_2, PGI_2, PGD_2, PGF_2 and TXA_2 with catalyzation of corresponding prostaglandin synthases such as PTGES, PTGIS, PTGDS, PTGFS and TBXAS1 (Needleman et al., 1986). Those derived prostaglandins can associate with their receptor to invoke signaling cascades to interact with other signaling pathways and exert biological functions (Li & Mansmann, 2013). The NSAID model is exported as a XML file in appendix A.12.

2.6 INTEGRATION OF CANCER HALLMARKS INTO NSAID MODEL AND ITS VALIDATION

Different independent studies provide evidence that deregulation of COX-pathway can assist cellular malignant tranformation by at least following four mechanisms: (1) sustained angiogenesis (Tsujii et al., 1998); (2) tissue invasion and metastasis (Hiraga et al., 2006); (3) proliferation (Wang et al., 2008); (4) evading apoptosis (Jana, 2008). Moreover, Hanahan & Weinberg (2000, 2010) proposed the theory of cancer hallmarks in order to introduce a framework for a biological organization principle of tumorigenesis, which in detail explains that the human cancer development is a multiple-step biological malignant transformation. Based on those scientific finding and theory, these four cancer hallmarks (sustained angiogenesis, invasiveness and metastasis, proliferation and apoptosis) were created as pseudo-object and integrated into the NSAID model by applying the mass action law. The hallmark proliferation represents the proliferation strength of the model and is directly embedded within the MAPK-, wingless-type MMTV integration site family (WNT)-, mechanistic target of Rapamycin (MTOR)-, and hypoxia inducible factor 1 (HIF1)-pathway. This hallmarks is mathematically implemented as the sum of putative proliferative biomarkers including upregulation of cell proliferation (URGCP), antigen identified by monoclonal antibody Ki-67 (MKI67), tripartite motif-containing protein 21 (TRIM21), DNA topoisomerase 2-alpha (TOP2A), forkhead box M1 (FOXM1), polo-like kinase 1 (PLK1). The hallmark *tissue invasion* describes the

62

potential of invasion and metastasis process and is directly involved in MAPK-, transforming growth factor beta (TGFbeta)-, insulin-like growth factor 1 receptor (IGF1R)-, and NfkB-pathway. This hallmark is mathematically implemented as sum of all matrix metalloproteinases defined in the NSAID model. The hallmark *evading apoptosis* symbolizes the level of circumventing apoptosis and is directly involved within the Death-Receptor-, IGF1R-, Toll-like Receptor (TLR)- and COX-pathway. Its mathematical implementation is the sum of anti-apoptotic factors over pro-apoptotic factors. The hallmark sustained angiogenesis represents the strength of sustained angiogenesis in the model and is connected with the VEGF- and fibroblast growth factor (FGF)-pathway. Similar to the mathematical implementation of evading apoptosis, this hallmark is implemented the sum of pro-angiogenesis factors over anti-angiogenesis factors. The aim of this modeling is to underscore the essential contribution and influence on the tumorigenesis process in different aspects (Li & Mansmann, 2013). These four cancer hallmarks are the final readout components of following *in silico* simulation experiments.

The study of Denkert et al. (2003) focused on inhibition effect of two different therapeutic interventions: COX-isoform-specific siRNA and COX-2 selective drug NS-398 in the human ovarian carcinoma cell line (OVCAR-3). For the first time, this study demonstrated that COX-isoform-specific siRNA does not have any proliferative inhibition effect on OVCAR-3, whereas the same cancer cell line had clear response for the drug NS-398. This is indeed an interesting phenomenon, however, the authors could not explain which molecular mechanism could be the major reason to cause this big difference of therapeutic responses. Therefore, it would be interesting to investigate and elucidate the underlying molecular mechanism by applying the NSAID model. An *in silico* experiment was designed during the study of Li & Mansmann (2013) to simulate the inhibition effect of both aforementioned therapeutic interventions with regard to the OVCAR-3 cell line. The first step of this experiment was to initialize the NSAID model

with the gene expression data of the OVCAR-3 (see Materials and Methods 3.5). Afterwards the NSAID model integrated the gene expression data was simulated until it reaches a steady state. The ultimate goal of this simulation procedure is to transfer the genetic information of the cancerous cellular system into the model system so that the model could reflect the dynamic behavior of corresponding cancer cell line by facing diverse therapeutic interventions. This type of simulation has been performed three times: 1.) tumor state (only with gene expression data of OVCAR-3 and without any therapeutic intervention); 2.) perturbation state (with COX-2-specific siRNA); 3.) perturbation state (with NS-398 drug). The inhibition effect of COX-2-specific siRNA can be analyzed through the comparison between the simulated data from 2.) state and the from 1.) state (abbreviated CT comparison) (Fig. 12A). Similarly, the drug effect of NS-398 can be analyzed through the comparison between the simulated data from 3.) state and the from 1.) state (abbreviated NT comparison) (Fig. 12B).
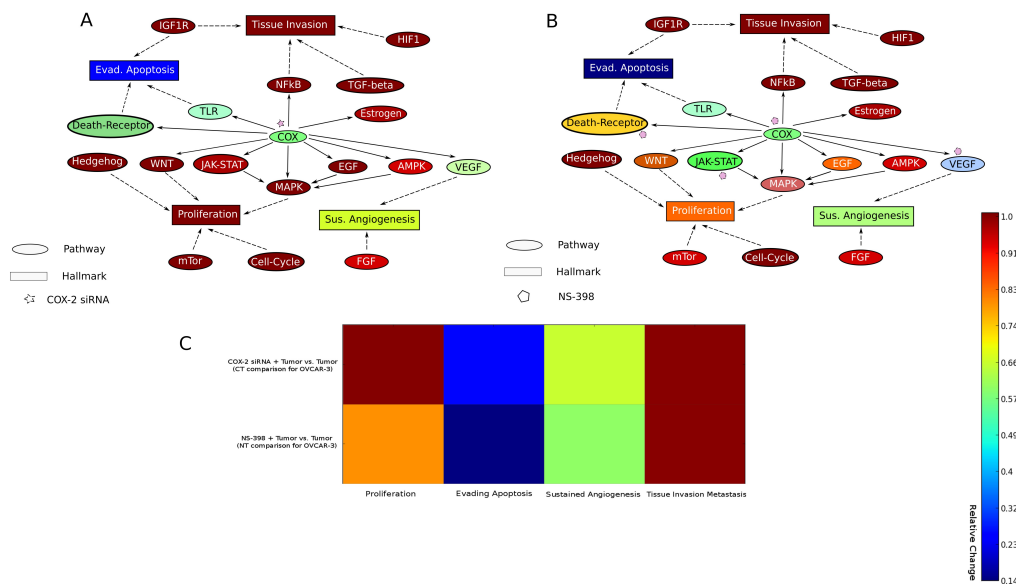
Figure 12. Simplified visualization of COX-2 inhibitions within the NSAID model network. **A**: Inhibition of COX-pathway via COX-2 specific siRNA interference. **B**:

Inhibition of COX-pathway via the drug NS-398 (COX-2 selective inhibitor). **C**: Heatmap result of state comparison includes CT comparison (first row), NT comparison (second row) and CT2 comparison (last row). The color bar is designed for **A**, **B** and **C.** (this figure and its legend is modified and taken from Li & Mansmann, 2013)

The comparison result shows that for the CT comparison, the readout component hallmark *proliferation* remains the same in both states, whereas the hallmark *proliferation* is reduced to 79.8% for the NT comparison (Fig. 12C). This result of hallmark *proliferation* acts in line with the result achieved by the study of Denkert et al. (2003) and shows that the model could differ the inhibition effects between COX-2-specific siRNA (no inhibition-effect) and drug NS-398 (with inhibition-effect) with regard to the OVCAR-3. The comparison results also reveal that both types of inhibitions could cause signal reduction within the COX-pathway, which in turn dramatically reduced the intensity of crosstalk between COX-pathway and other pathways. Unfortunately, the signal reduction of the COX-pathway did not reach a high level of influence on the entire model system, because the PGE_2, a key functional component, can bind to their cognate receptors EP1-4 due to their low expression in the OVCAR-3. Without this ligand-receptor association, PGE_2 can not fully forward signal. Therefore, the COX-pathway can only exert its biological function with less magnitude in OVCAR-2. This is also the reason that only targeting COX-2 in OVCAR-3 can not lead to reduction of high cellular activities of relevant signaling pathways including MAPK-, WNT- and NfkB-pathway (Fig. 12A), whose high activities enhance the expression level of transcription factors including v-myc myelocytomatosis viral oncogene homolog (MYC), jun proto-oncogene (JUN), fbj murine osteosarcoma viral oncogene homology (FOS), and specificity protein (SP)1,3. Those putative transcription factors have major involvement for activate signaling pathways including MTOR-, WNT-, and Hedgehog pathway to promote or sustain the high cellular proliferation. In contrast, the drug NS-398 possesses therapeutic effect to inhibit the functions of COX-2, VEGF, interleukin

65

(IL) 1 and tumor necrosis factor (TNF) (Ferrario et al., 2005), which leads to reduce the signal intensities within VEGF-, TLR- and JAK_STAT pathway; to prevent the kinase activities of phospholipase C gamma (PLCG) and PTK2 protein tyrosine kinase 2 (PTK2); to hamper cellular cycle activity. Therefore, the cellular proliferation was decreased during *in silico* simulation (Fig. 12B). With regard to hallmark evading apoptosis is reduced to 14.4% and 25.8% in the CT and NT comparison respectively. The result also shows that signal intensities of the TLR- and Death-Receptor-pathway is decreased (Fig. 12B). The hallmark sustained angiogenesis is reduced to 59.8% and 65.6% because of signal reduction within the VEGF- and FGF-pathway. These results suggest that both COX-2-specific siRNA and drug NS-398 can reduce cellular functions involved in the pathological processes of evading apoptosis and sustained angiogenesis. With regard to the OVCAR-3, drug-398 has a better inhibition effect to reduce both pathological processes than the COX-2-specific siRNA does. However, the comparison result shows that both therapeutic intervention can not influence the hallmark tissue invasion and metastasis.

2.7 Analysis of COX Based Synthetic Lethality for Breast, Colon and Lung Tumor

Synthetic lethality describes the gene-pair-relationship, when mutations occurring simultaneous in this gene pair could cause the cell death, however, mutations of each gene is still compatible to the normal cell viability. This concept has been intensively discussed in the study of Kaelin (2005), where results of many studies are introduced to demonstrate low risk and high efficiency of the combination therapy. The strength of this concept is to take advantage of essential molecular information for developing high quality of the combined anti-cancer drugs that can enhance the on-target efficiency and reduce the off-target effects to the minimum.

Many preclinical studies have investigated the combined effect for inhibiting COX-2 and a receptor tyrosine kinase such as EGFR, ERBB2. They reported that this type of combination inhibition is far more effective than each single agent alone (Mann et al., 2001; Tortora et al., 2003; Torrance et al., 2000). Therefore, another *in silico* experiment was designed during the study of Li & Mansmann (2013) to apply the NSAID model for detecting the combination inhibition. The gene expression data of 60 cancer cell lines from Cancer Genome Atlas (https://tcga-data.nci.nih.gov) was used as initial genetic input for the NSAID model. Those cancer cell lines include 20 breast cancer cell lines, 20 colon cancer cell lines and 20 lung cancer cell lines (appendix B.10). There are more than ten tyrosine kinases defined in the NSAID model including RET, EGFR, FGFR, PDGFRA, INSR, ERBB2-4 and others. All combination inhibition of COX-2 and receptor tyrosine kinases within the NSAID model were analyzed for each cancer cell
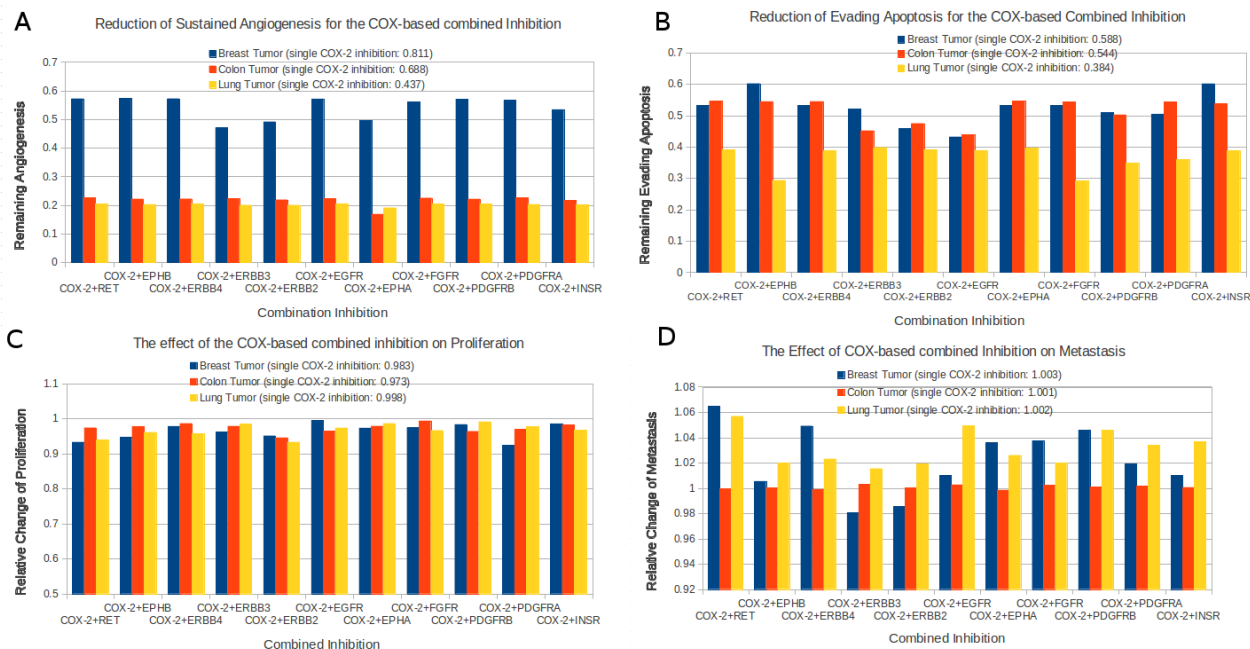
Figure 13. Effect of the COX-based combined inhibition on cancer hallmarks: Angiogenesis (**A**), Evading Apoptosis (**B**), Proliferation (**C**) and Metastasis (**D**). The

blue, red, and yellow columns in each histogram indicates averages of the relative changes of cancer hallmarks for 20 breast carcinoma samples, 20 colonrectal carcinoma samples, and 20 lung carcinoma samples respectively. The relative change of each cancer hallmark refers the comparison between therapeutic perturbation state (COX-based inhibition) and control state (tumor state). (this figure and its legend is taken from Li & Mansmann, 2013)

line. Each combination-inhibition presents a type of therapeutic intervention. The control state of each cancer cell line is the steady state reached by the NSAID model with initialization of gene expression from corresponding cancer cell line. The perturbation state of each cancer cell line is the steady state reached by the NSAID model with initialization of gene expression from corresponding cancer cell line and additional combined inhibition of COX-2 and a receptor tyrosine kinase (see Materials and Methods 3.12).

Angiogenesis is the cellular process responsible for the formation of new blood vessel. The maintenance and progression of tumor tissues including breast, colon and lung are highly relying on the angiogenesis during the course of tumor development (Hanahan & Folkman, 1996). The result shows that combination inhibitions of COX-2 and receptor tyrosine kinases could exert more significant reductions of angiogenesis than single COX-2 inhibitor did. This better effect is especially clear shown in colon and lung cancer cell lines, where combination inhibition reduced angiogenesis process to 68.84% and 43.72% of the from the control state in colon and lung cancer cell lines respectively (Fig. 13A). This result clearly demonstrates that the combination inhibition of COX-2 and a relevant receptor tyrosine kinase can suppress the highly promoted angiogenesis in cancerous cellular system to prevent the colon and lung cancer development (Li & Mansmann, 2013). Therefore, this result act in agreement with the results from different independent studies of Mann et al. (2001), Tortora et al. (2003) and Tuccillo et al., (2005). In those independent studies, it is demonstrated that combined anti-cancer drugs of ZD6474 (a EGFR and VEGFR inhibitor) and SC-326 (a selective COX-2 inhibitor)

have clearly additive effect on xenograft models of colon and lung cancer. The combined treatments of inhibiting COX-2 and ERBB2, and COX-2 and EGFR to achieve additive anti-angiogenic effect on colon cancer cell lines and xenograft models.

For breast cancer cell lines, COX-2 based combination inhibition could reduce angiogenesis process to 40-60% of the from the control state that are better than 81.13% achieved by application of the single COX-2 inhibition (Fig. 13A). This result suggests that combined therapeutic treatment related to COX-2 and receptor tyrosine kinase might present an effective way for reducing angiogenesis process in the breast cancer treatment. Interestingly, the breast cancer cell lines reflects stronger resistance against COX-2 based combination inhibition compared to colon and lung cancer cell lines (Fig. 13A). I conducted pathway analysis related to signal intensity within the NSAID model and found out that in average, under the same combination inhibitions compared to colon and lung cancer cells, breast cancer cells possess still relatively higher activities of Hedgehog-pathway, EGFR-pathway, JAK/STAT-pathway, WNT-pathway, cell-cycle-pathway, and ERBB-pathway. These remaining high activities of relevant cellular pathways could offer breast cancer cells to resist combination inhibition (Li & Mansmann, 2013).

Apoptosis is the cellular process of programmed cell death and provide a critical cellular mechanism for efficient removal of superfluous, damaged or infected cells. A number of pathological conditions including tumorigenesis have the precondition of deregulated apoptotic machinery, which results in excessive cell survival and accumulation of malignant mutations (Adams, 2003; Green, 1998). The result shows that all combination inhibition of COX-2 and receptor tyrosine kinase can reduce evading apoptosis below 61% of the from the control state. For the lung cancer cell lines, this reduction effect is even clearer, which is even below 40% (Fig. 13B). However, in comparison with the reduction effect on evading apoptosis by single COX-2 inhibition, the combination inhibition of COX-2+ERBB2, COX-2+ERBB3, COX-2+EGFR and COX-2+PDGFRB

69

(platelet-derived growth factor receptor beta), show clear additive effect for breast and colon cancer cells; for lung cancer cells, only the combination inhibition of COX-2+EPHB (EPH receptor) and COX-2+FGFR have a clear additive effect. The rest of combination inhibitions have the same or highly similar effect as the single COX-2 inhibitor does for these three cancer cells (Fig. 13B). Interestingly, the combination inhibition of COX-2+INSR even reduce the evading apoptosis process less than the single COX-2 inhibition. According to the component analysis within the NSAID model, the reason could be that inhibition of INSR dramatically prevents the insulin signaling, which results in decrease of phosphorylated protein kinase (PKC). The decrease of this putative protein kinase activity leads to the activity reduction of NfkB-pathway. Subsequently, the downstream target genes including many pro-apoptoic ligands of this pathway are down-regulated, hence, the evading apoptosis process is enhanced (Li & Mansmann, 2013).

Proliferation is the cellular process for cell growth and reproduction. Under suitable environmental conditions including sufficient nutrient, temperature, the rate of cellular proliferation becomes stable to enable the formation of cell population under the control of the normal cellular system (Berridge, 2012). The primary task of cancer development is to force the cellular system become highly proliferative. Unfortunately, with regard to proliferation the comparison of the control state and perturbation state shows that proliferation process in perturbation state remains more than 90% of the in control state for all those COX-2 based combination inhibitions within the three types of tumor. None of combination inhibitions does perform clear better (reduction 30%; $p < 0.05$) proliferative inhibition than the single COX-2 inhibition does (Fig. 13C). Only three combination inhibitions (COX-2+RET (ret proto-oncogene), COX-2+ERBB2 and COX-2+PDGFRA) show little improvement (reduction 4%; $p < 0.05$) in comparison with the single COX-2 inhibition (Li & Mansmann, 2013).

Tissue invasion is the primary course of cancer mortality and is also one of the most pertinent cancer hallmarks from a therapeutic perspective (Weinberg, 2008). The *in silico* result shows that except three combination inhibitions (COX-2+ERBB2 for breast cancer cells, COX-2+ERBB3 for breast cancer cells and COX-2+EPHA for colon cancer cells), all combination inhibitions do not have any effect on the metastasis of these three types of cancer cells (Fig. 13D). The single COX-2 inhibition does not have any effect on the metastasis either (Li & Mansmann, 2013).

2.8 THE CHALLENGES OF CONSTRUCTION OF A COMPREHENSIVE SIGNALING MODEL IN HUMAN

Although many early studies have been conducted to standardize the formalisms for creating signaling networks (Gianchandani et al., 2006; Klamt et al., 2006; Papin et al., 2005), the global reconstruction of human signaling network covering multiple cancer-relevant signaling pathways has not been reported yet, hence the challenges to achieve this goal should not be underestimated and can be divided into three parts: (i) identifying the model components, (ii) reassembling model components focused on the system-biological properties and (iii) revealing the dynamic behavior of the underlying biological system (Bhalla et al., 1999). During the course of writing this thesis, I constructed a human signaling model and demonstrated how to apply this model for the purpose of personalized medicine.

2.9 COMPONENTS IDENTIFICATION OF THE HUMAN SIGNALING (HS) MODEL

At initial stage of this thesis, I conducted an in-depth literature research and defined the corresponding model components and reactions according to diverse published data (appendix A.3). Meanwhile, I also utilized the information provided by different publicly accessible databases such as Reactome (Vastrik et al., 2007) and Kegg (Kanehisa et al., 2000), as well as literature and defined more than 50 cancer relevant signaling pathways

71

(appendix C.1, C.2 and C.3); furthermore, I applied the knowledge from the Transfac database (Wingender et al., 1995) to define the transcription factor related network in the model; lastly, I incorporated literature-based miRNA-regulation information into the model extending my previous work (Li et al., 2012a; appendix A.4). Table 5 summarizes the statistical overview of this model. Figure 14 shows the simplified graphic model overview.
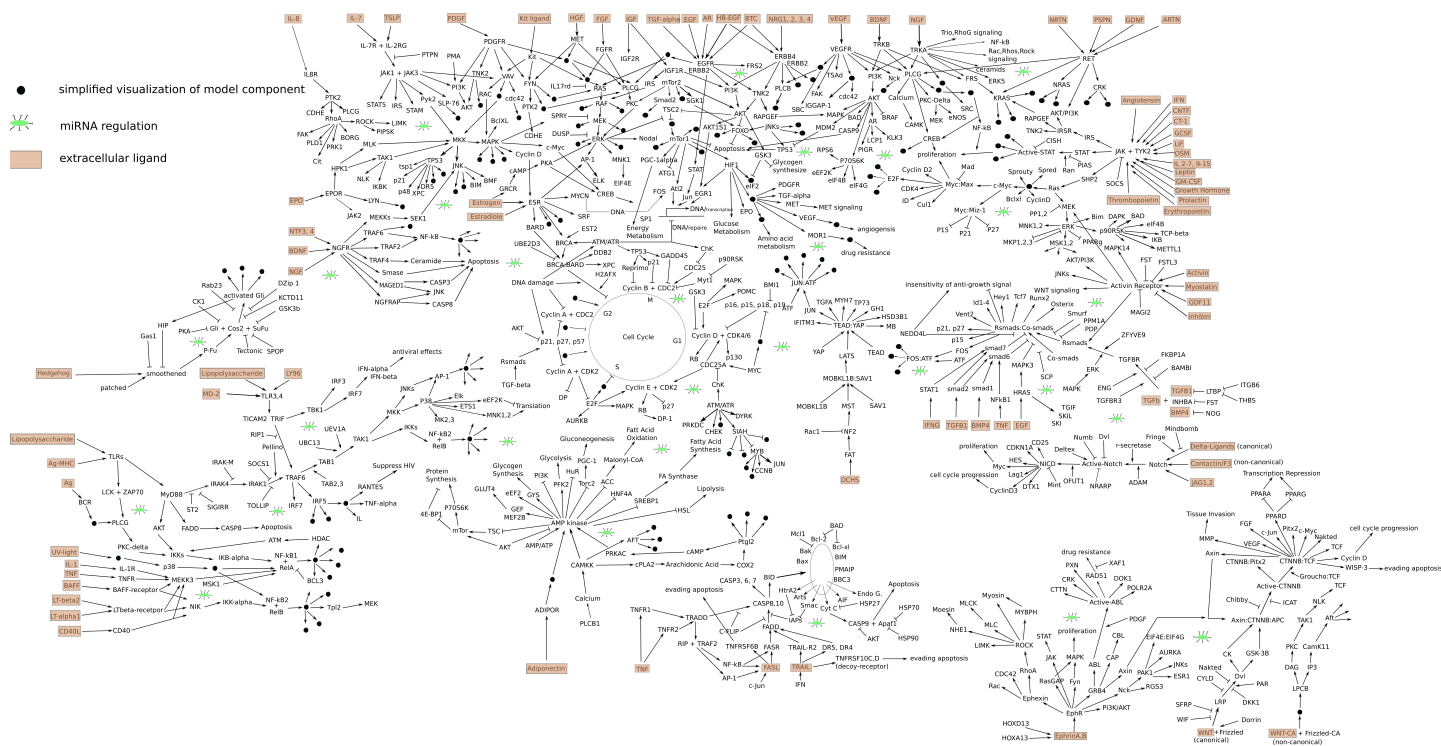
Figure 14. A simplified network-overview of the implemented human signaling model.

During model construction, substantial effort has been made to underscore the biological relation between gene, mRNA and protein through the transcription and translation reaction defined in the model, e.g. the concentration of a protein entity in the model depends only on the corresponding mRNA entity and reactions, in which this protein

entity takes part. The model possesses a consistent data structure (see Fig. 26). For instance, a gene entity and an mRNA entity in the model are associated with ensembl-Id, an miRNA entity is associated with miRNA-accession, a protein entity is associated with UniProt and a metabolite entity is associated with CHEBI-Id. These associations of the model component with unambiguous Ids enable the model to be initialized with experimental data such as gene-expression data in a standardized manner (see Materials and Methods 3.5).

Component	No.	Reaction	No.
gene	2718	transcription	2700
mRNA	2223	translation	863
protein	2040	decay	3828
miRNA	1048	complex-formation	657
compound	974	translocation	2616
complex	45	phosphorylation	2537
pseudo-object	63	ubiquitination	37
		acetylation	12
		activation	173
		dimerization	47
		GTP exchange	269
		methylation	11
		miRNA-binding	4347
		other types	26
Sum:	9111	Sum:	18123

Table 5: The statistics of the human signaling model. Complex includes protein-protein complex, protein-metabolite complex, protein-mRNA and protein-gene complex. Compound is metabolite. PseudoObj includes protein-inhibitors (ABL inhibitor, AKT inhibitor, EGFR inhibitor, MEK inhibitor and other) and hallmarks (Apoptosis, Angiogenesis, Proliferation and other)

2.10 REASSEMBLING MODEL COMPONENTS FOCUSED ON THE SYSTEMS-BIOLOGICAL PROPERTIES

It is a non-trivial matter to reassemble model components in order to reflect the systems-biological properties, whereby I focused on different properties of the biological system

which includes: (1) feedback control, (2) functional redundancy, (3) structural stability and (4) modularity (Kitano, 2002).

System control ensures that system-level behavior of model can not be dramatically changed due to slight perturbation (such as slight change of temperature, pressure). This type of biological property can be realized by different types of feedback loops within the HS model. I defined the feedback-loop within a signaling pathway as direct feedback (positive, negative) and the feedback-loop within the crosstalk of different signaling pathways as indirect feedback (positive, negative). Because identification of intracellular feedback loops is a NP-hard problem and no effective algorithm has been reported yet, I manually counted the number of 4 types of feedback loops (direct and indirect positive-feedbacks, direct and indirect negative-feedbacks). Table 6 lists a statistical overview of these feedback-loops (appendix A.5).

Feedback Loop	No.
Direct positive-feedback	35
Direct negative-feedback	20
Indirect positive-feedback	92
Indirect negative-feedback	49

Table 6: The statistics of feedback loop within the human signaling network.

The biological property of functional redundancy in the model reflects a kind of fail-safe mechanism in cell biology. In case one cellular component fails to exert its function, there is another component with similar function, which can replace this one. Thus, cellular processes related to this component will not be inactive after their failure. Different studies indicated that the combination of MEK inhibitor and AKT inhibitor can achieve much better therapeutic effects than a "stand-alone" MEK or AKT inhibitor application (Williams et al., 2012; Meng et al., 2010). To study this concept of combined MEK/AKT inhibition an *in-silico* experiment was set up with four experimental conditions aiming to investigate the inhibitor effect regarding the functional redundancy of MEK and AKT in

the model (Table 7): <1> Deactivation of both inhibitors (control state: CS) <2> Activation of MEK inhibitor (active state1: AS1); <3> Activation of AKT inhibitor (active state2: AS2); <4> Activation of both inhibitors (active state3: AS3).

Nr.	Simulation Condition
<1>	Deactivation of MEK and AKT inhibitors (CS)
<2>	Activation of MEK inhibitor (AS1)
<3>	Activation of AKT inhibitior (AS2)
<4>	Activation of MEK and AKT inhibitors (AS3)

Table 7: Four simulation conditions for analysis of the functional redundancy within the human signaling model. The deactivation is realized by fixing the initial concentration of the corresponding inhibitor to 0. The activation is realized by setting the initial concentration of the corresponding inhibitor by the same value as the average gene expression from the corresponding cell line.

The expression of genes in the model was initialized based on the values derived from the gene expression data of the human cell line H23. this cell line was chosen because both inhibitors have an average effect on this cell line according to the Meng et al., (2010). The initial concentrations of other model components including mRNAs, miRNAs, proteins, complexes, etc. were set to 0 nM, so that the signal generated during signal transduction process is created in the transcriptional level at first place. Figure 15 shows the network graph related with the proteins MEK and AKT and their downstream components in the model. The functional redundancy of MEK and AKT is seen by their cooperative effect on the downstream phosphorylated targets TSC and RPS6KB1. The heatmap (Fig. 16) displays the concentration ratios of downstream to AKT, MEK or combined model components between different simulation conditions (AS1 vs. CS; AS2 vs. CS; AS3 vs. CS). The concentration ratios of direct to AKT, MEK downstream phosphorylated proteins (TSC, RPS6KB1) are much lower by the activation of both inhibitors than those of each single activation, which reveals the functional redundancy towards the combination of drug effects. The specific downstream targets such as ESR

75
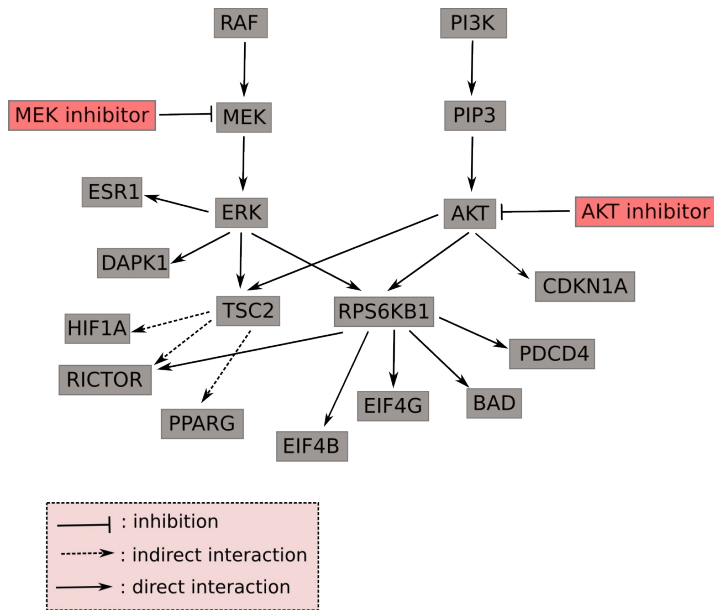
Figure 15: Schematic representation of the concept for functional redundancy in the human signaling model. The interaction network with downstream components derived from MEK and AKT sub-pathways.

and CDKN1A, are only affected by the activation of the corresponding inhibitor. However, it shows that the indirect downstream proteins (HIF1A, PPARG) can only be affected by MEK inhibitor, but not AKT inhibitor. This could be a reason for intensive crosstalk between different signaling pathways such as MAPK, mTor and AMPK.

The structural stability of the signaling network is associated with the intrinsic structure integrated in this model. The entire network structure represents the bow-tie structure (including Fan-in, Core, Fan-out and GSC; Fig. 17A), which is an important biological organization principle for complex systems to be robust and flexible (Csete and Doyle, 2004; Kitano, 2007). Based on the classification of the model components, the extracellular ligand proteins and their related reactions are classified into the Fan-in group indicating the input part of the model system. The downstream transcription factors
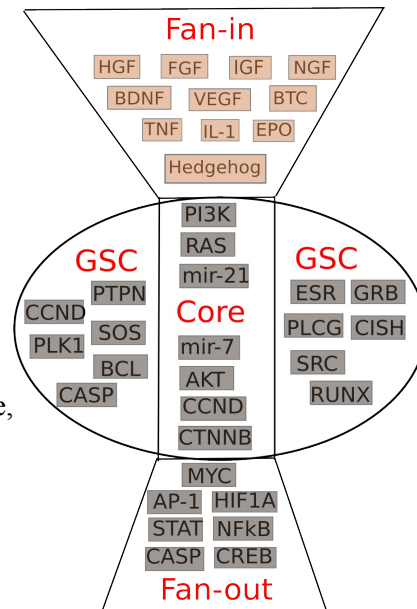
Figure 17A: The 'bow-tile' structure (Fan-in, Core, Fan-out and GSC) of the human signaling model.

77

Group	No. of Components	No. of related Reactions
Fan-in	390	1021
Core	156	2531
Fan-out	301	911
GSC	8256	23466

Table 8: The statistics of model componets in the 'bow-tile' structure.

of signaling pathways and their related reactions are classified into the Fan-out group denoting the output part of the model system. The Core group includes the model components with key-functions in related signaling pathways and their corresponding reactions. The remaining model components with their reactions are classified into the GSC (giant strong component) group. The number of the model components and reactions in these four groups are summarized in table 8.

With respect to modularity of the biological system integrated in the model, I incorporated seven cancer hallmarks (evading-apoptosis, sustained-angiogenesis, insensitivity-to-anti-growth-signals, limitless-replicative-potential, tissue-invasion, self-sufficiency-in-growth-signals, apoptosis) described in Hanahan & Weinberg (2011) and extended them further to eleven hallmarks by creating four cellular-process hallmarks (proliferation, drug-resistance, cell-cycle-progression, tumorigenesis). These eleven hallmarks in the HS model symbolize the distinct functional modularity with respect to the tumorigenesis (Fig. 17B). The hallmarks "Proliferation", "Cell-cycle-progression", "Apoptosis" and "Tissue-invasion" indicate the strength of the proliferative potential, the cell cycle promotion, the ability of invoking apoptosis and tissue invasion, respectively. Table 9 lists these functional hallmarks and related pathways. For instance, the pathway "MAPK_signaling" can contribute directly to the "Proliferation" process. The pathway "Death Receptor" can invoke the "Apoptosis" process and suppress the ability of "Evading-apoptosis" in a cell. The pathway "TGF_beta" can prevent a cell from becoming insensitive to anti-growth signals. The pathways "VEGF_signaling" and

"FGF_signaling" help a cell to sustain the angiogenesis process. The pathway "Cell_Cycle" delegates the cell cycle progression and at the same time provides a cell with the possibility of limitless replicative potential.
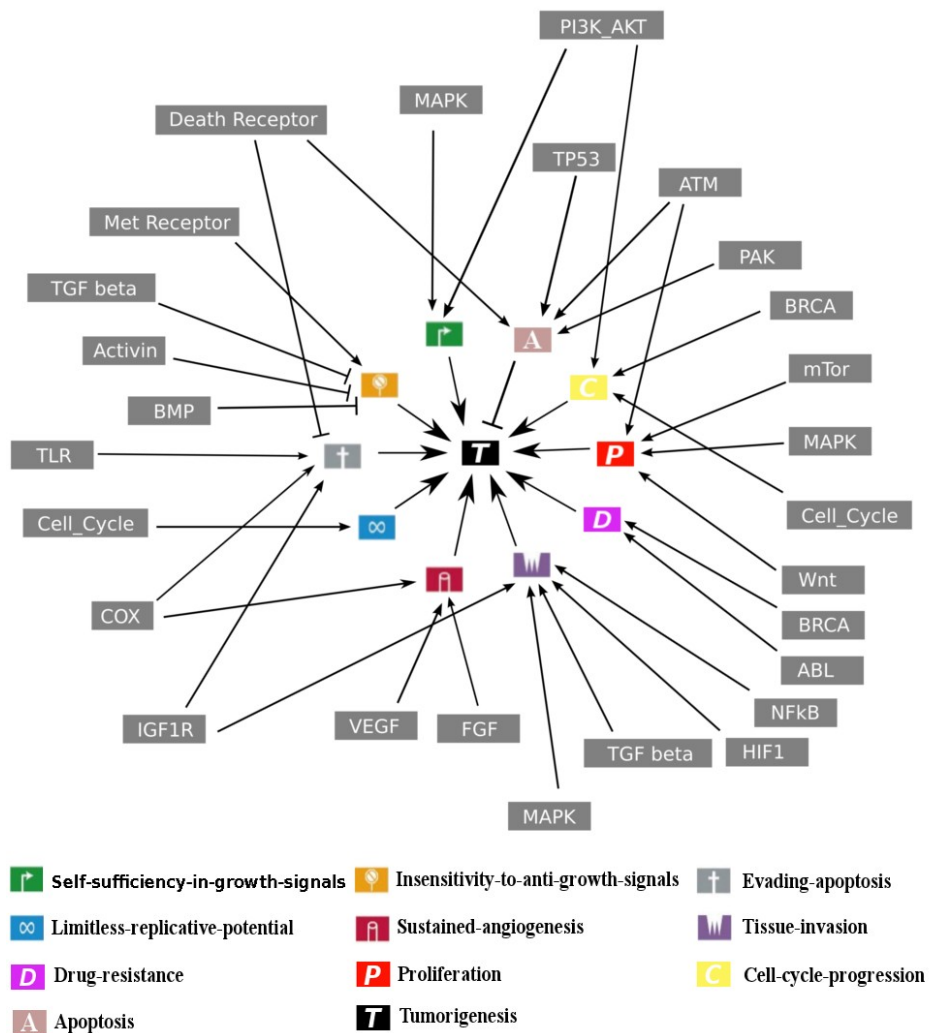
Figure 17B: The functional modularity regarding the relationship between pathways and hallmarks. The detailed implementation of these hallmarks in the human signaling model is explained in Materials and Methods.

After the construction of this integrated signaling model, Vikash Pandey and I analyzed the model with respect to signaling connectivity and isolated islands by conducting a connectivity-analysis by applying iGraph (Csardi and Nepusz, 2006). The connectivity-degree distribution of the model components after filling gaps shows a linear regression with R^2 value of 0.664 (Log-log scale; appendix B.1) and follows a power-law distribution, which is esteemed as the standard of a biological network with scale-free topology (Choura & Rebai, 2010). As discussed in previous studies, a scale-free biological network is robust to random errors from a structural point of view (Jeong et al., 2000).

2.11 ANALYSIS OF DYNAMIC BEHAVIOR OF THE BIOLOGICAL SYSTEM UNDERLYING THE HUMAN SIGNALING MODEL

At the advanced stage of this thesis, the dynamic behavior of the biological system according to an individual cell line (or patient) was analyzed. At first step, the basic property of signal transduction in this model should be validated, which is the creation of cause-and-effect relationship of signaling inputs and signaling outputs (Papin et al.,, 2005). The model contains more than 90 extracellular ligands that bind to the corresponding receptors to activate signaling pathways individually. A set of *in-silico* experiments were performed to verify signaling transduction processes within different signaling pathways to generate a model-wide signal dependency matrix (Fig. 18A). The results of this set of *in silico* experiments are recorded in the appendix B.8. In each of these experiments, extracellular ligands were activated individually as signaling inputs in the model and key components of corresponding signaling pathways were considered readout components (signaling outputs). For instance, the activation of extracellular ligand EGF can lead to the activation of MAPK signaling, which includes the activation of MEK and ERK. The activated ERK can in turn activate many target transcription

factor proteins such as MYC and JUN. In this case, the phosphorylated MYC is considered a signaling output.

The difference of calculated readout component concentration with the activation of corresponding ligands versus control state (no ligand activation) is visualized as a heatmap in Fig. 18A. Each row of the heatmap presents an *in silico* experiment with the condition of one ligand activation. Each column is the readout component under corresponding ligand activation condition. The both lists are arranged in the same order. Therefore, the diagonal line in this heatmap reveals that each ligand in the model can activate the corresponding signaling pathway, which verifies the basic property of the signal transduction in this model. Additionally, the result shows that several ligands activate or repress readout components of other designated pathways, allowing the readers to assess the amount of crosstalk implemented in the HS model.

In the second set of the simulation experiments, seven out of 93 ligands with the combination of miRNAs were activated. The selected seven ligands (BDNF, EGF, FASLG, GH1, IGF1, TGFB1 and VEGFA) can invoke signal transduction processes related to the relevant cellular processes such as proliferation, apoptosis, angiogenesis, etc. and therefore are deregulated in different tumor cell lines. The selected seven miRNAs (including mir-8, let-7a, mir-143, mir-15a, mir-7, mir-155 and mir-21) can specifically regulate the corresponding post-transcriptional processes of ligand genes in this model, respectively (appendix A.4). The heatmap (Fig. 18B) shows that the signal intensities were reduced during simulation when the ligand proteins were down-regulated by the corresponding miRNAs.

Table 9: The defintion of cancer- and process-hallmarks and corresponding relationship with pathways in the human signaling model. The detailed information of pathway information is in the appendix C.1, C.2 and C.3.

Hallmark	Mechanism	Related Pathways
Apoptosis	• The pro-apoptotic ligand (FASLG, TNF, TNFSF10) binds to the corresponding pro-apoptotic receptor (FAS, TNFRSF1A, DR4/5) to trigger the activation of caspase proteins and invoke apoptosis (extrinsic apoptosis). • In response to the signals resulting from DNA damage, loss of cell-survival factors or others, pro-apoptotic proteins including BCL2, BCL2L11, PMAIP, BBC3, BAX, BAK can promote the mitochondrian release of other pro-apopotoic factors such as CYCS, AIFM, EndoG, SMAC to activate caspase proteins and trigger apoptosis (intrinsic apoptosis).	Death_Receptor, TP53_signaling, ATM_signaling, PAK_signaling
Cell-cycle-progression	• **G0**: The first resting phase for cells who just have completed a cycle and accomplished dividing process. • **G1**: The begin phase with G1 checkpoint mechanism (CCND1/2/3 + CDK4/6) prepare the cellular readiness for the process of DNA synthesis. • **S**: The DNA replicative phase. The complex CCNE1/2 + CDK2 and CCNA1/2 + CDK2 organize and inspect the DNA replication process. • **G2**: The second resting phase with G2 checkpoint mechanism (CCNA1/2 + CDC2) prepare the cellular readiness for entering mitosis phase and dividing process. • **M**: The mitosis phase where the cell growth stops and the metaphase checkpoint (CCNB1 + CDC2) prepares the cellular readiness for cell division completion.	Cell_Cycle, BRCA_signaling, PI3K_AKT
Drug-resistance	• The protein products of genes (XAF1, RAD51) can regulate the strength of cellular drug resistance.	ABL_signaling, BRCA_signaling, JAK_STAT
Evading-apoptosis	• The ligand-receptor binding (TNF:TNFRSF1B, IGF1R:IGF1, IGF1R:IGF2) activates pathways to provide the ability of evading apoptosis. • The pro-apoptotic ligand receptor binding (such as TNF:FASLG, TNF:TNFRSF1A) triggers apoptosis and	IGF1R_signaling, TLR_signaling, Death_Receptor

	suppresses this evading ability.	
Insensitivity-to-anti-growth-signals	• The degradation of SMADs catalyzed by proteins (such as SMURF, NEDD4L) leads to the insensitivity of anti-growth signals.	TGF_beta, Activin_signaling, BMP_signaling, MET_Receptor
Limitless-replicative-potential	• The enzyme TERT can extend the end of chromsomes by adding repeating sequence to prolong the replicative potential of a cell. The constitutively active activity of this enzyme enables the replicative potential of a cell in a limitless manner. The transcriptional activity of the corresponding gene is regulated by diverse cellular components (MYC:MAX, TP53, RUNX2, E2F1, SMAD2/3:SMAD4).	Cell_Cycle
Proliferation	• The proliferation potential is regulated by diverse components (MYC:MAX, URGCP, MKI67, TRIM21, EIF4 initial factors).	MAPK_signaling, mTor_signaling, WNT_signaling, ATM_signaling
Self-sufficiency-in-growth-signals	• The constitutively active state of proteins (BRAF, ARAF, KRAS, HRAS) enables the ability of self-sufficiency in growth-signal of a cell.	MAPK_signaling, PI3K_AKT
Sustained-angiogenesis	• The ligand-receptor-binding (FGF:FGFR, VEGFC:KDR, VEGFA:KDR, FIGF:KDR) sustains the angiogenesis process of a cell. • The ligand-receptor-binding (CD36:THBS1) antagonizes the angiogenesis process.	VEGF_signaling, FGF_signaling
Tissue-invasion	• Proteins of the matrix metalloproteinase (MMP) familiy play important roles in the physiological processes related to extracellular matrix. The deregulation of these proteins conveys the ability of tissue invasion. The transcriptional activity of the related genes are highly regulated by diverse transcriptional factors (such as JUN, CREB, STAT, HIF1A, TP53, ETS2, NfkB, CTNNB).	MAPK_signaling, TGF_beta, HIF1_signaling, NFkB_signaling
Tumorigenesis	• The qualitative summary of all other hallmarks.	
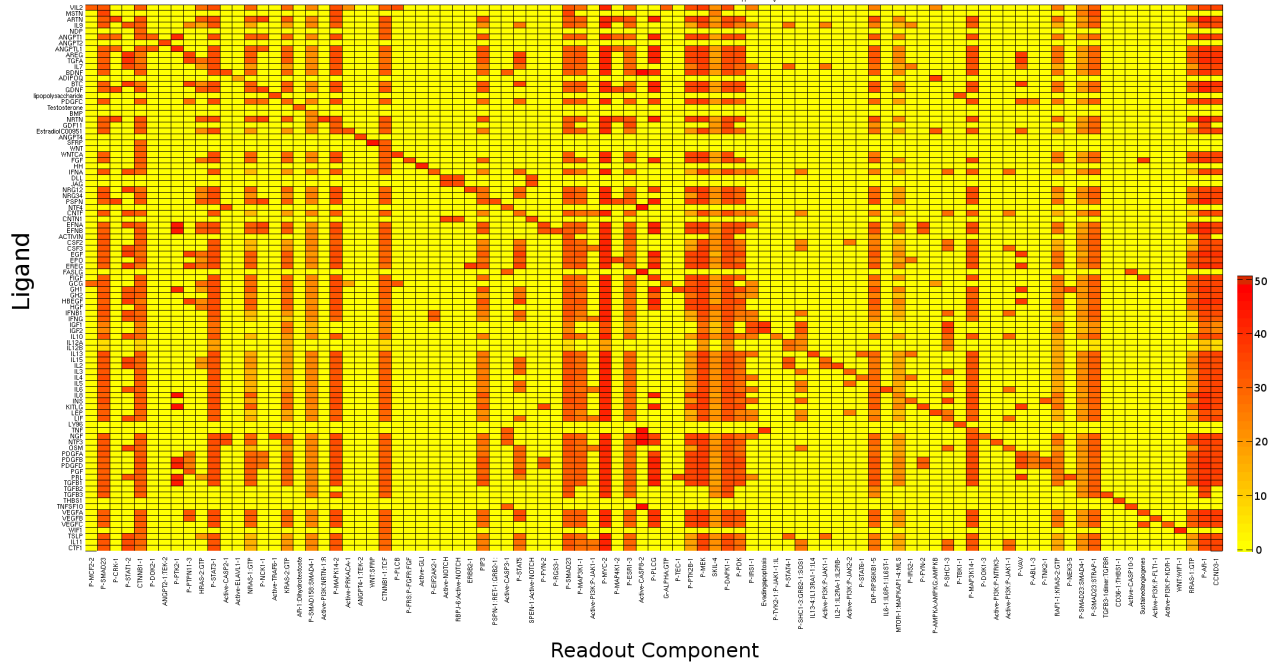
Figure 18A: Ligand-Activation-Test. Each row of the heatmap presents a condition of the activation related to a specific ligand. Each column includes the readout components activation based on the specific ligand activation. Rows and columns in the data matrix represented in the heatmap are arranged in the same order. Therefore, the diagonal line in this heatmap reveals that each ligand activation leads to the corresponding activation of signaling pathways. The color scale symbolizes the value of log2-ratio. Red: up-regulation state, Yellow: normal state. (The simulation data is attached in appendix C.6).

Ligand Activation and miRNA regulation
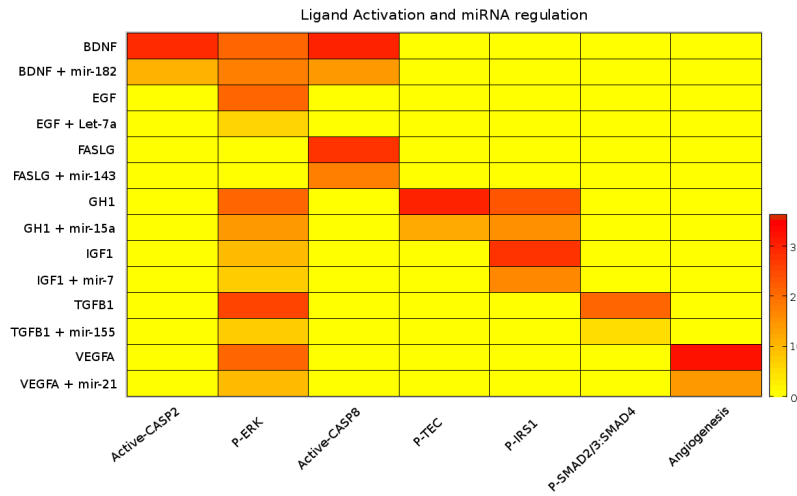
Figure 18B: miRNA regulation regarding the effect of ligand activations.
The down-regulation of ligand concentrations through the corresponding
miRNA regulation reduces the signal intensity of the corresponding signaling
pathway as visualized in this figure. The color scale symbolizes the value of
log2-ratio. Red: up-regulation state, Yellow: normal state. (The simulation data
is attached in appendix C.7).

The dynamic behavior of a biological system refers to, for instance, how a biological
system responses when facing different types of the perturbations including genetic
perturbations (e.g. prooncogene overexpression, tumor suppressor downregulation,
mutation) and environmental perturbations (e.g. drug treatment, temperature change). In
order to analyze how well the HS model can predict dynamic behavior of corresponding
biological system underlying individual cell line (or patient), an *in silico* simulation
pipeline was developed to enable the initialization of the HS model with genetic
information including gene-expression data and miRNA-expression data (Fig. 19). The
gene-expression data and miRNA-expression data of an individual cell line (or patient)
contain such rich information that it cannot only be used for the pathological state
classification of an individual cell line (Golub et al., 1999; Tong and Nemunaiti, 2008),
but it can also provide information related to the dynamic behavior of the underlying
cellular system to perform therapeutic prediction (Chang et al., 2003). At the stage of the
model initialization (Fig. 20A), the existing genetic information from the biological
system of an individual cell line (or patient) was copied into the model-system.
Afterwards, the simulation process was performed with the application of the model. The
goal of the simulation strategy was to propagate the signal from genetic data to the entire
model system so that in the steady state the model system can behave dynamically like
the underlying biological system of the corresponding individual as similarly as possible
(Fig. 20). During this signal propagation process, different types of perturbations can be
inserted into the model and the corresponding dynamic responses of the model system
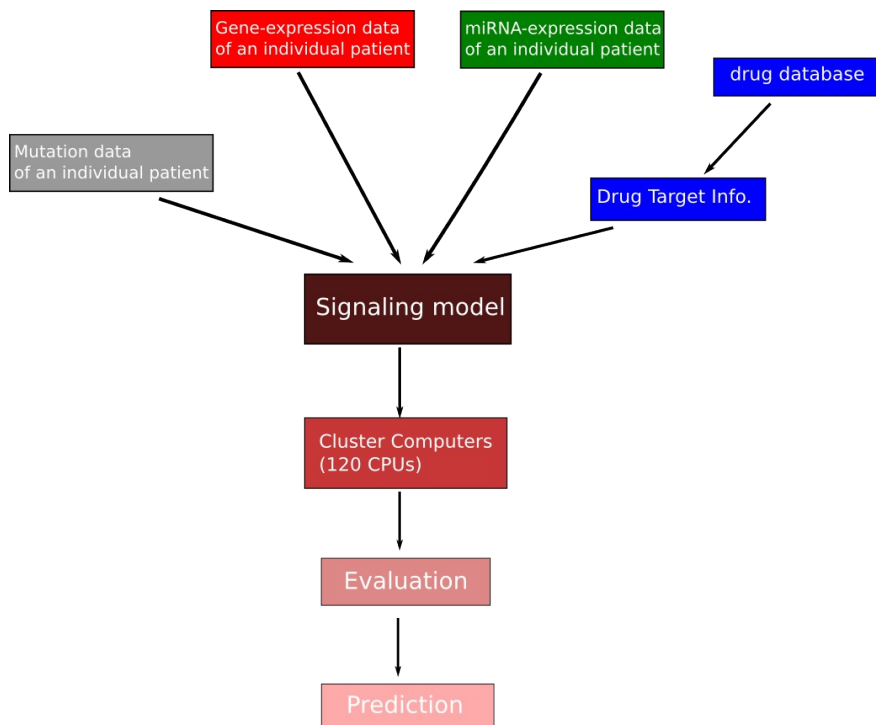
Figure 19: The workflow of *in-silico* simulation pipeline. The input information includes gene experssion data, miRNA-expresssion data, mutational data and drug target dissociation value. The simulation process is based on Petri net, which is explained in the Materials and Methods part.

should be able to reflect the responses of the underlying biological system under the same perturbation condition. For instance, a type of drug perturbation is inserted into the model to prevent the signal transduction from component A to component B (Fig. 20C). The following components C, D and E are all affected by this perturbation reflected by a different concentration reduction (Fig. 20D). However, D and E are only partially affected since they both receive a signal from another part of the network. Fig. 20D quantitatively also shows the difference in component concentrations between these two states: control state (Fig. 20B) and perturbation state (Fig. 20C). Thereby, the different

types of predictions can be made according to the genetic information of an individual cell line (or patient) and different types of perturbations.

In order to investigate the structure model and predictive ability of the HS model, the following *in silico* simulation experiments were performed by applying the aforementioned simulation pipeline (Fig. 19). A recent study (Nagaraj et al., 2011) achieved deep coverage (> 90%) of the functional transcriptome and the proteome of a cancer cell line (HeLa) by the application of an advanced mass spectrometry technology.
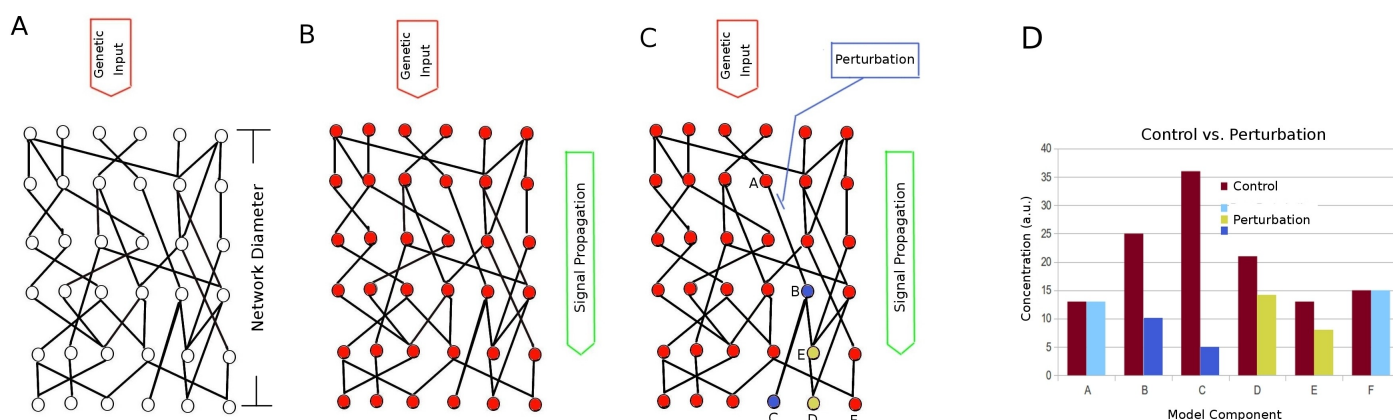
Figure 20: The signal propagation strategy for inheriting dynamic properties of the underlying biological system. A: before the HS model is initialized with the cell line data (gene-expression data), the initial values of all model-components are assigned with zero. **B**: during the signal propagation process, the genetic information is 'copied' into the model system, which is reflected as signal intensities, from the underlying cellular system of a corresponding cell line (or patient). **C:** meanwhile, the perturbation related to the specific model components is inserted during the signal propagation process, the model-system should be able to response to this perturbation in a way which should reflect the dynamic behavior of the underlying cellular system of the cell line (or patient) under the same perturbation condition. **D:** the comparison of concentration-change of model-component between perturbation-state and control-state. The 'brown' bar symbolizes the concentration from the control state. The 'light blue' bar stands for the concentration of model components, which are not affected by the drug treatment. The 'yellow' and 'dark blue' bars stand for concentrations of model

87

components which are moderately and highly affected by the drug treatment, respectively. (this figure is taken from Li & Mansmann, 2013)
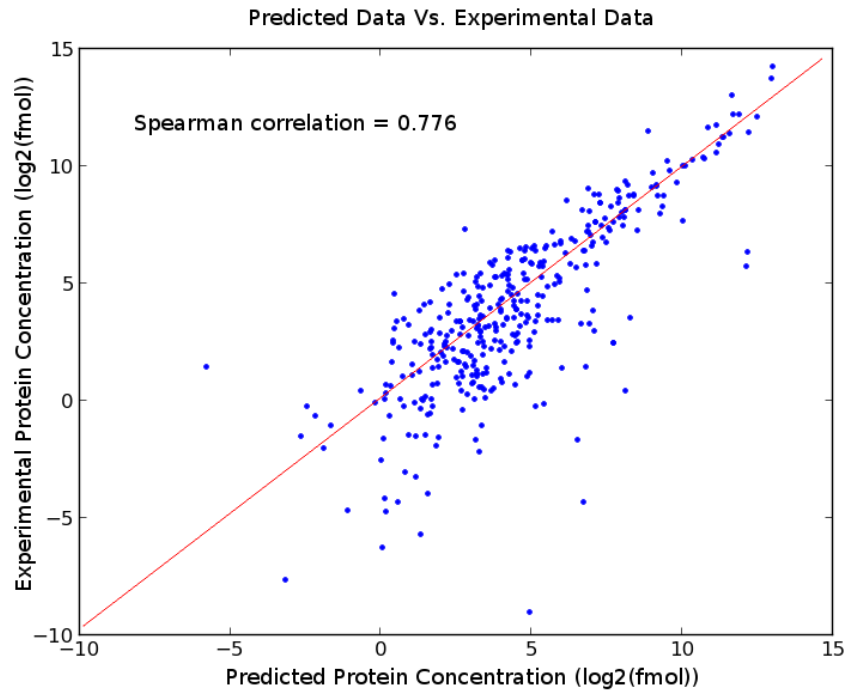
A. (top) B. (bottom)
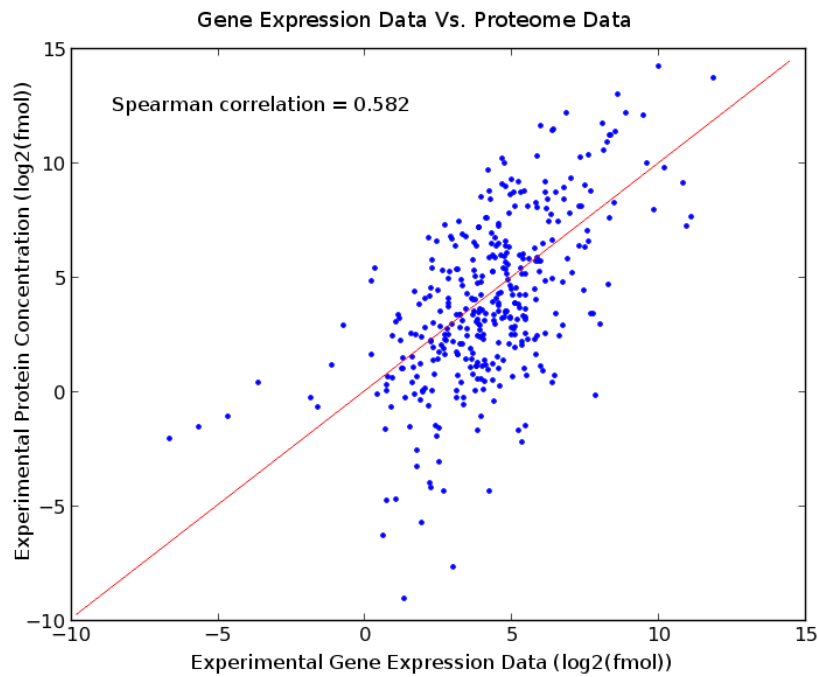
Figure 21. Visualization of correlation measurements. **A**: the correlation analysis resulted to a value of 0.776 for the spearman correlation between the predicted proteome data and the experimental proteome data from the HeLa cell line published by the Nagaraj et al., (2011). **B:** the analysis from the comparison of transcriptome data and proteome data derived from the same study, resulted in a spearman correlation of value of 0.582. For this analysis, I took into consideration genes and proteins that are defined in the HS model. (The data is attached in appendix C.8).

The gene-expression data from Nagaraj's study was utilized to initialize this model (see Material and Methods 3.5), thereby 'copied' the genetic information from this cellular system into the system of the HS model and performed *in silico* simulation. Afterwards, the simulated protein data output of the model system was compared with the proteome data of Nagaraj's study. The comparison between both data shows a spearman correlation of 0.776, whereas the spearman correlation of gene expression data and proteome data from the same study is below 0.6 (Fig. 21A and B).

In addition, another independent study (Lundberg et al., 2010) performed a global analysis of both mRNA and protein data in three functionally different human cell lines (U-2 OS, A-431 and U-251 MG). The same approach was applied for each of these cell lines to transfer genetic information into model and perform *in silico* simulation. Afterwards, the correlations between predicted proteome data and experimental proteome data are far more than corresponding correlations between transcriptome data and proteome data from the same study (appendix B.2-B.7). All these comparison results indicate that through this simulation strategy, the HS model can inherit the genetic information from a biological system and thereby can reflect the dynamic behavior of this biological system underlying a cell line by translating the gene-expression data input into protein concentration with high accuracy. Therefore, the HS model might possess a meaningful structure, provide predictive power and prove useful for biomarker prediction in cancer cell lines.

2.12 ESTABLISHMENT OF A TUMORIGENESIS CALCULATION FORMULA

The remarkable progresses in cancer research in the past two decades indicate that the tumorigenesis in the human is a multiple-step process and that these steps reflect genetic alternations that drive the progressive transformation of normal human cells into highly malignant derivatives (Hanahan & Weinberg, 2000; Hanahan & Weinberg, 2011). To date, more than a hundred distinct types of cancer and subtypes of tumor have been identified in the human body. The experience gained by diverse studies strongly suggests that all differences among tumor and cancer types can be reduced at molecular level, namely molecules within the tumor tissues and their interactions (Su et al., 2001; Golub et al., 1999; Chang et al., 2003). Therefore, it might be possible to explain the dynamic behavior of cancerous cells with a comprehensive molecular model, in this case, the human signaling model. Generally speaking, common choices for cellular objective functions in systems biological models include biomass production (Edwards and Palsson, 2000), energy (Ramakrishna et al., 2001), and byproduct production (Varma et al., 1993). Moreover, until now it has been difficult or impossible to define a putative objective-function for a canonical signaling model. In order to overcome this limitation, I incorporated seven cancer hallmarks and four process hallmarks into the HS model (explained in the section 2.10) and developed a tumorigenesis calculation formula to make the hallmark 'tumorigenesis' as the putative readout component of the HS model to predict therapeutic responses. If the component 'tumorigenesis' is not down-regulated under a certain therapeutic inhibition effect during simulation, then the corresponding cell line would be evaluated non-response for the corresponding type of therapeutic intervention. This tumorigenesis calculation formula is displayed as following:

$$T = (k_1*X_1+k_2*X_2+k_3*X_3+k_4*X_4+k_5*X_5+k_6*X_6+k_7*X_7+k_8*X_8+k_9*X_9)/(1+k_{10}*X_{10})$$

where: X_1 = evading-apoptosis (EA); X_2 = sustained-angiogenesis (SA);

   X_3 = proliferation (P); X_4 = drug-resistance (DR);

   X_5 = Insensitivity-to-anti-growth-signals (IS);

90

X_6 = limitless-replicative-potential (LP);

X_7 = tissue-invasion (TI); X_8 = self-sufficiency-in-growth-signals (SS);

X_9 = cell-cycle-progression (CC); X_{10} = apoptosis (A);

T = Tumorigenesis (T);

k_1... k_{10} are the weight-factors of the corresponding hallmarks;

The ultimate goal of this calculation-formula in the model is to quantitatively summarize the signal intensities from other hallmarks into the final readout component 'tumorigenesis'. The weight-factors k_i symbolize the different influence-degrees that the corresponding hallmarks can exert on the readout component "tumorigenesis". In order to estimate those weight-factors k, I applied the parameter optimization algorithm, simulated annealing (Kirkpatrick et al., 1983), by applying the published data of NCI-60 cancer cell lines regarding the cell line drug-responsiveness to Gefitinib (Holbeck et al., 2010). The embedded Gefitinib effect into the HS model is explained in appendix A.8. I chose the drug Gefitinib for the weight-factor optimization procedure because it has in average the moderate drug effect on the NCI-60 cancer cell lines so that the possibility of parameter-overfitting can be reduced. The detailed parameter estimation procedure is explained in the Materials and Methods 3.8. The estimated range of weight-factors are displayed in the following:

k_1, k_2, k_3, k_5, k_{10} = (0, 30);

k_4 = (300, 600);

k_6, k_8 = (800, 1300);

k_7 = (200, 500);

k_9 = (20, 70);

In order to test how well this HS model with cancer– and process-hallmarkers can predict the dynamic behavior of cancerous cellular systems, I applied the data of the NCI-60 cell lines to qualitatively predict the cell line drug response for Imatinib and Temsirolimus.

The embedded drug effect of Imatinib and Temsirolimus into the HS model is shown in appendix A.8. The simulation procedure is visualized in Fig. 19 and explained in Materials and Methods 3.5. The results of both drug response predictions are visualized in the Fig. 22.

2.13 UTILIZATION OF MIRNA-EXPRESSION DATA TO DISCOVER AND DETECT THE BIOMARKERS OF INDIVIDUAL CANCER CELL LINES

Biomarkers are defined as molecular, cellular or functional measurable parameters that can indicate a particular genetic, physiological or functional status of a cellular system (Ludwig & Weinstein, 2005). The successful identification of biomarkers associated with an individual tumor or cancer can help to early detect the tumor, to precisely diagnose the pathological state as well as the progression of the tumor and to clearly identify the anti-tumor drug effect related to an individual cell line (or patient). Recent studies indicate that the miRNA biomarker has been shown to be more specific, sensitive and robust (Mo et al., 2012).

As mentioned above, the HS model integrates a global miRNA regulation network containing 1048 miRNA with their corresponding miRNA target information (appendix A.4). Therefore, I intended to investigate whether protein biomarkers of an individual cancer cell line can be identified by the utilization of the individual miRNA-expression data, respectively. The study of Gaur et al., (2007) examined the characterization of miRNA expression level in different human cancer cell lines and identified the clear difference of miRNA expression profile between the normal and cancer cell lines. From Gaur's study, I applied their published miRNA expression data of three cancer types: Colon, Hematologic, and CNS. By applying this miRNA data, I initialized the HS model and simulated individually. The miRNA initialization process was to assign the miRNA-gene entities in the model with the input miRNA-expression data (Materials and Methods

92

3.5). All gene objects are switched on and expressed at the same level and other objects including RNA, protein, complex and others were set to zero. Its goal was to activate the miRNA-regulation network within the model under this 'real life' condition to detect those proteins that are most up- or down-regulated due to the miRNA regulation effect.

A.
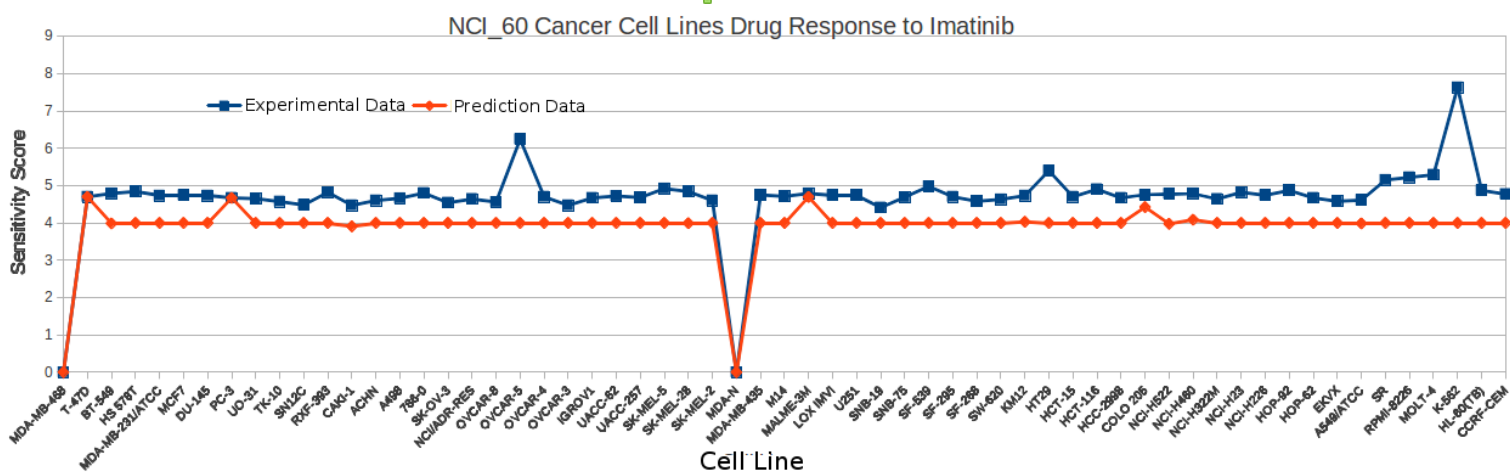
B.
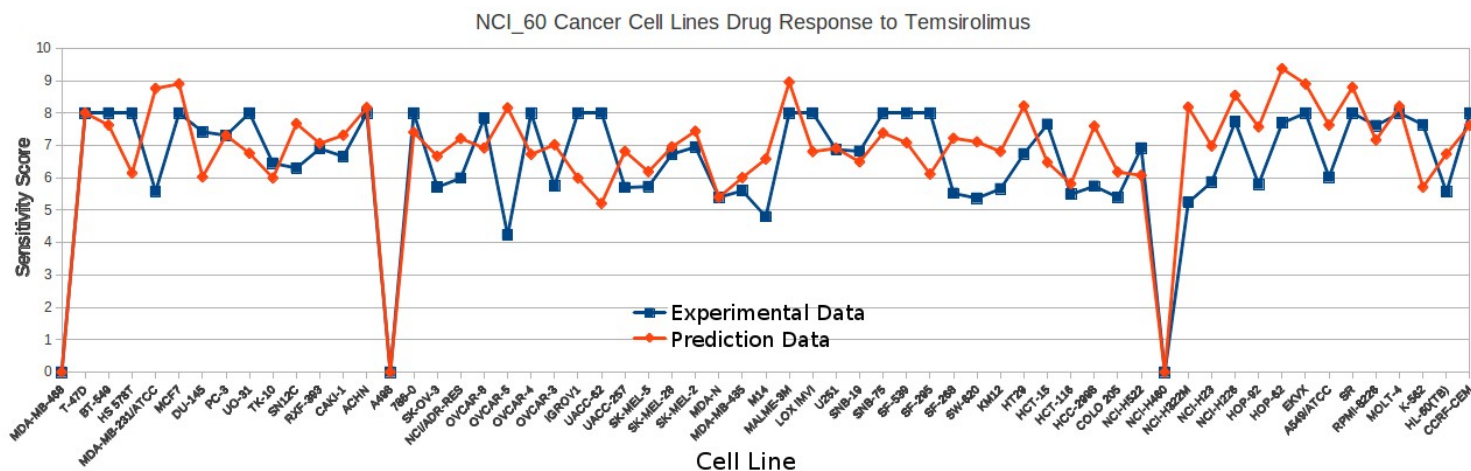
Figure 22. Drug response prediction of NCI-60 cancer cell line. A: visualizes the predicted sensitivity score and experimental established sensitivity score of NCI-60 cancer cell lines regarding the Imatinib drug perturbation. The correlation between predicted sensitivity score and experimentally established sensitivity score is 86.18% **B:** visualizes the predicted sensitivity score and experimental established sensitivity score of NCI-60 cancer cell lines regarding to the Temsiromulis drug perturbation. The correlation between predicted sensitivity score and experimentally established sensitivity score is 74.41%. (The data is attached in the appendix C.9)

By doing so, I enter the individual miRNA expression information to the model and translated this information into proteome data based on the miRNA regulation during simulation. After reaching the steady state, the proteins with most differential (up- and down-regulated) value might be considered as potential biomarker candidates, because they represent the unique information originated from the miRNA expression data. Based on the simulation data, I listed the top 10 up- and down-regulated proteins of each of these three cancer cell lines and proposed them as biomarkers of the corresponding cancer cell line (Table 10).

Colon		Hematologic		CNS	
up(fold)	**down(fold)**	**up(fold)**	**down(fold)**	**up(fold)**	**down(fold)**
EREG(98.87)	MAPK11(0.06)	RUNX2(250.2)	CCNE(0.02)	VEGFA(187.8)	E2F4/5(0.04)
LEPR(83.85)	CCNL2(0.05)	TCF(211.2)	PTPN11(0.03)	IRS2(145.7)	CSNK1D(0.06)
NR4A2(66.33)	SFRP(0.22)	TSC1(175.34)	CK1(0.04)	MKI67(125.8)	CK1(0.06)
DYRK1A(59.8)	CASP9(0.28)	MAD(153.3)	CSNK1D(0.05)	ROCK(92.84)	GSK3B(0.08)
Gli(55.12)	EPOR(0.35)	mTor(119.6)	EPOR(0.08)	APAF(77.41)	RHEB(0.11)
TERT(49.56)	KLF2(0.41)	BCL2(99.76)	GSK3B(0.09)	DR6(58.19)	RAPGEF(0.11)
IL4R(45.89)	PPARG(0.42)	PKC(88.67)	MAX(0.11)	PTPN12(60.62)	MAX(0.2)
WNT(42.67)	TP53(0.43)	CSK(58.56)	HSP90(0.19)	RB1(60.18)	RRAS(0.25)
BMP(39.45)	FOXO(0.44)	TEAD(58.34)	NFKB(0.21)	PTPLAD(45.3)	ERBB2(0.36)
CCND(20.68)	CEBPB(0.45)	APAF(48.34)	MYC(0.25)	CAMK(44.05)	CTNNB1(0.38)

Table 10: Identification of the biomarkers of individual cancer cell lines with the help of miRNA-expression data.

2.14 RECONSTRUCTION OF THE HUMAN METABOLIC MODEL AND EXTENSION WITH THE MIRNA-REGULATION NETWORK.

Here, I reconstructed the published human metabolic model (Duarte et al., 2007) by adding the transcription- and translation-reaction for each gene in this model. The advantage of this model modification is that different transcriptional, post-transcriptional and translational regulations can be added into it. I integrated the global miRNA-regulation network into this model. All the miRNA targets were validated and are listed in appendix A.9. The miRNA biogenesis includes miRNA gene transcription- and translocation-reaction in the model. I named the new version "Human Metabolic miRNA" (HMMA). Model HMMA includes 6765 components and 12586 reactions. It contains 5 types of object entities: Gene, mRNA, Protein, Compound and miRNA, which are labeled Ensembl-ID, Ensembl-ID, Uniprot-ID, ChEBI-ID and miRNA accession, respectively. Table 11 lists the statistics of model components and reactions in HMMA.

Component	No.	Reaction	No.
gene	2383	transcription	2389
mRNA	2670	translation	1335
protein	1380	decay	3724
miRNA	1048	miRNA-binding	1123
compound	1377	translocation	2389
complex	3	complex-formation	3
		metabolic reaction	1623
Sum:	8861	Sum:	12586

Table 11: The statistics of HMMA model components / reactions. The gene includes the normal gene and the miRNA gene; the compound is metabolite. The transcription reaction includes the normal gene transcription and the miRNA gene transcription. The miRNA gene transcription simplifies two processes: (i) miRNA gene transcription catalyzed by DNA-polymerase II or DNA-polymerase III; (ii) cropping of the primary transcript (pri-miRNA) into a hairpin intermediate (pre-miRNA) by the nuclear 650 kDa microprocessor complex, comprising in humans of the RNase III DROSHA (RNASEN) and the DiGeorge syndrome critical region gene 8 (DGCR8) (see Fig. 6C). The miRNA target-binding reaction simplifies two processes: (i) mature miRNA in complex with DICER and TARB binds to the Ago-complex and turns it into the RNA-induced silencing complex (RISC); (ii) RISC recognizes the target mRNA and binds to it (see Fig. 6C).

The observation of diverse studies reveals that cancer cells usually adapt their metabolism to meet the high cellular demand for proliferation and other cellular activities (Hanahan & Weinberg, 2011). Therefore, the creation and integration of biomass into a molecular metabolic model could present an effective way to predict the characteristic alternations in cancer metabolism. Recently, such a genome-scale computational modeling approaches has been successfully applied to predict metabolic states of fast growing microorganisms (Price et al., 2004). The flux balance analysis (FBA) is a constraint-based modeling (CBM) approach that is particularly suitable for analyzing the cancer metabolism because cancer cells need to dramatically increase their growth rate and constantly looking for energy input, mathematically equivalent to metabolic flux redistribution that can produce a high level of biomass (Price et al., 2003). Unfortunately, the FBA analysis is not applicable for analyzing model HMMA, because the model contains detail information of miRNA regulations, which can not be considered during FBA analysis. In contrast, the simulation approach elucidated in Fig. 20 in section 2.11 might be very suitable to analyze the different metabolic states (normal state vs. tumor state) within the model. The HMMA model is available in appendix A.10.

2.15 FEATURES AND FUNCTIONALITIES OF PYBIOS2

PyBioS2 is developed in an object-oriented manner for easier and more conformable design, modeling, and simulation of cellular systems. This software is connected with many public available databases such as BioMart, BioModels, ChEBI, ENSEMBL, KEGG, Reactome, UniProt. Therefore, large and complex models can easily be generated based on these diverse data sources and can comfortably be reused due to the simple model data storage hierarchy. The dynamic behavior of the underlying biological system can be analyzed and predicted with two different simulation strategies implemented in PyBioS2: (i) the deterministic integration of ordinary differential equations (ODEs); (ii) timed and stochastic simulation of Petri net. The simulation process can be reviewed

stepwisely, which enables the user to observe the dynamic changes in the model components during the simulation. This function allows the user to investigate the dynamic behavior of each model component at any time point of relevance. In appendix A.6, there is a step-by-step tutorial of PyBioS2. Table 12 summarizes the comparisons between different systemsbiological softwares. In comparison with other four relevant softwares/tools, PyBioS2 possesses many advantages that are explained in the following.

Systems Biological Software Comparison						
Software	Operating System	ODE Engine	Petri net Engine	Automatic-definition option	Model re-usefulness	Database Connection
CellDesigner	Windows, Linux and MacOS X	Yes	No	No	Middle	Middle
Copasi	Windows, MacOS X, Linux and Sun Solaris	Yes	No	No	Middle	Low
PySCeS	Windows, Linux and MacOS X	Yes	No	No	Low	Low
Virtual Cell	Windows, Linux and MacOS X	Yes	No	No	Low	Middle
PyBioS2	Operating system independent	Yes	Yes	Yes	High	High

Table 12: The functionalities comparison of different systems biological softwares.

2.15.1 Multiple Database-Connection

PyBioS2 possesses three different internal databases: ZOPE database, PyBioS2 database and CPDB database (Fig. 23). ZOPE is the Z Object Publishing Environment, the function environment of PyBioS2. Each model instance in PyBioS2 is created in ZOPE as well as the user account and permission. PyBioS2 database contains specific information about each model-component. Each component has a type including gene,

RNA, protein, compound, reaction, pathway and others. For instance, the chromosome location of genes and miRNA genes are stored in the PyBioS2 database. Furthermore, PyBioS2 utilizes PyCogent (Knight et al., 2007) to connect PyBioS2 database with the Ensembl Database to obtain biological information (such as chromosome location and sequence) about gene, mRNA and protein. Via the Soap protocol, PyBioS2 is connected to BioModels and ChEBI. Using HTTP-request, PyBioS2 can access information from UniProt. The connection with BioMart is implemented in PyBioS2 via xml-querys. The CPDB database collects data from many publicly available databases listed in Table 1 and updates every three months. PyBioS2 provides the user the option of multiple searches in different databases at the same time, which dramatically reduces the time consumption of data acquisition.
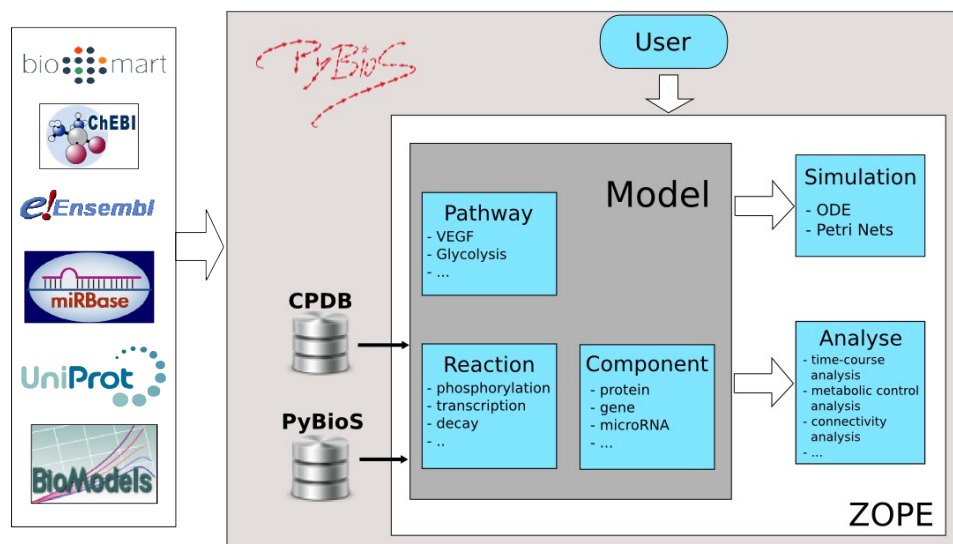
Figure 23. Schematic overview of PyBioS2 functions. PyBioS2 connects with many publicly available databases such as BioModels, BioMart, via different services provided by those databases like Soap-protocol, HTTP-request, xml-query and others (the figure is taken from Li et al., 2012b).

2.15.2 Data Integration

PyBioS2 functions as model repository receiving data from different data sources and deals with data integration problems. Therefore, it is vital to keep each object-entity unique and avoid importing redundant information. In PyBioS2, for instance, a gene entity is associated with Ensembl-ID and is linked to the Ensembl Genome Browser, UCSC Genome Browser and GeneCards. An mRNA entity, a miRNA entity, a protein entity and a compound entity are associated with Ensembl-ID, miRNA accession, UniProt and ChEBI-ID, respectively. There are other entities such as complex and pseudo object, that are assigned with the corresponding PyBioS2 internal Ids. All of the different entities are saved in the PyBioS2 interval database, which is responsible for data integration of different data sources (Fig. 24). According to these different types of putative IDs, PyBioS2 avoids importing redundant information and enhances the consistency of the model-data.

Figure 24. The diagram of PyBioS2 internal database recording relevant biological entities including gene, RNA, protein, complex, compound, reaction, pathway and others. One tab option within the PyBioS2 user interface is the database administration interface, where the PyBioS2 administrator can conduct data control, user-access permission and other routine regulation works. Among those data tables, there are three major tables: 'Bio_Obj', 'Reaction' and 'Pathway'. They are linked with each other through the 'id2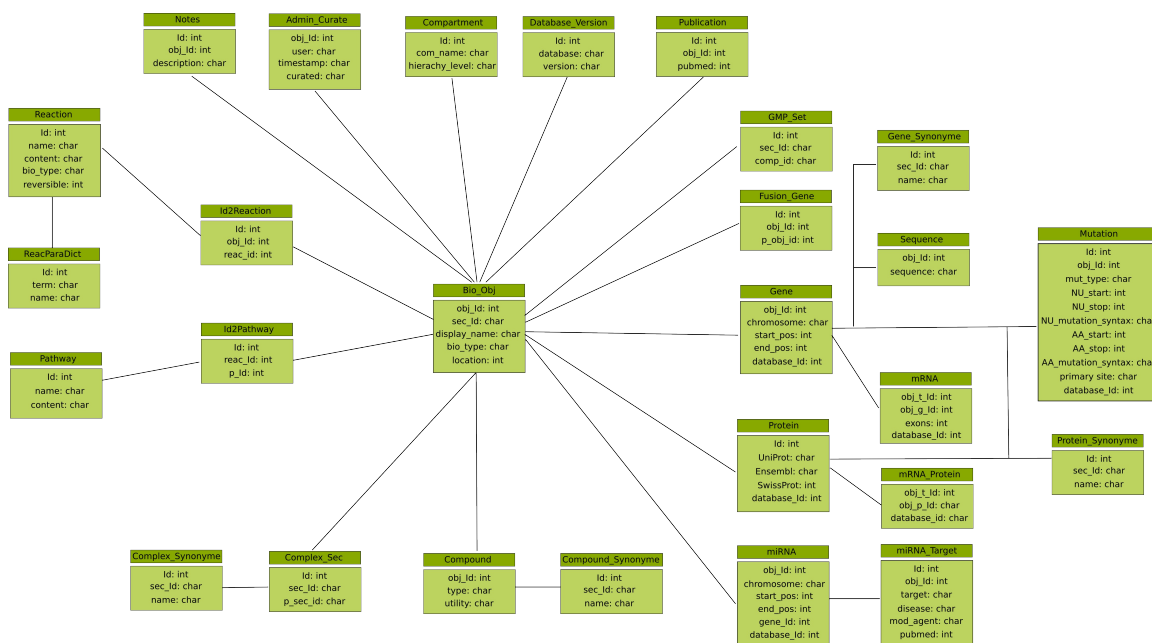Reaction' and 'id2Pathway' data table. The data table 'Bio_Obj' records all bio-object entities including gene, RNA, protein, complex, compound and others. Most of these object entities have a corresponding synonym data table for recording comprehensive synonym information. The data tables 'Reaction' and 'Pathway' save the reaction and pathway information, respectively. The data table 'Admin_Curate' records the user information regarding data import and creation. The data table 'Notes' and 'Publications' save the literature reference, publication related to model components. The data table 'GMP_Set' saves the group object information, for instance, MEK1 and MEK2 possess very similar biological function. In many cases of modeling, it is not required to make a detailed functional difference of both. Therefore, the user can use a group object MEK to symbolize both in order to reduce the complexity of modeling. The data table 'ReacParaDict' records the kinetic parameter distribution information according to each biochemical reaction. Because of current limitations of biological technology, most kinetic parameters of biochemical reactions are unknown. What is mostly known, is the magnitude of speed of biochemical reactions. In order to apply models in PyBioS2 with realistic simulation conditions, the information saved in the data table 'ReacParaDist' is essential.

2.15.3 Graphical Interface

In order to achieve the optimal visualization of the organization in model-components and database-connections, PyBioS2 defines the graphical tree structure. As depicted in Fig. 25, the 'Model' tree contains three root-nodes: 'Cell', 'Reaction' and 'Pathway', in which different types of model-components are organized. The root-node 'Cell' contains two child-notes: 'compartments' and 'types', in which all model-components including genes, proteins and others are sorted according to the compartment-information and the type-information. This enables users a comfortable browser and search for components either by type or by location. The 'Database' tree also contains three root-nodes for the PyBioS2 database, CPDB database and ChEBI database. The detailed database

100

information of search-results is displayed as the leaf nodes that are organized in different sub-categories protein, gene, compound, reaction and others.
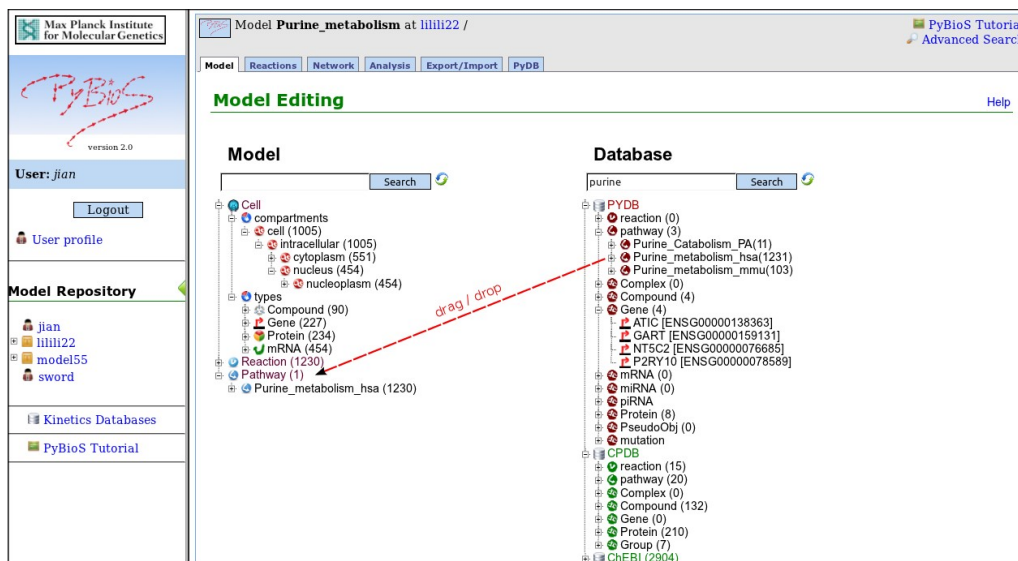
Figure 25. The navigation-frame (left) and content-frame (right) of PyBioS2.
The modeling procedure: search in the 'Database' tree; find the object-entity of interest; drag the node of this object-entity and drop it into the 'Model' tree.

2.15.4 Model Architecture

Considering the storage, processing, and utilization of the massive amount of biological knowledge/information, I designed a simple but efficient '3-level' hierarchy as model basic structure for PyBioS2 (Fig. 26). Each model is registered as an instance in ZOPE and contains two children instances: reaction and component. Reaction-entities and component-entities are organized in two different arrays. Therefore, there is no order restriction. Both arrays can be enlarged or reduced at any time, which makes it possible to merge or split any model. This structure avoids the disadvantages of conventional approaches to build a biological model by organizing objects according to their compartments, which complicates the model structure for merging and splitting

processes. As a matter of fact, many individual researchers may possess a great deal of knowledge about genes, molecular interactions, and other biological information underlying one particular pathway, however, no one can be familiar with the extremely large and complex number of interactions in an entire cell. Therefore, for constructing genome-scale models, it is considerably necessary to merge different models based on the excellent knowledge of many experts together into a bigger and more complex model. With the help of the model structure realized in PyBioS2, a model can be easily built by merging sub-models or it can simply be split into different sub-models (Fig. 26). Each of these sub-models can be utilized as a stand-alone model or incorporated into another model. This maximizes the reuseability of each defined object-entity, dramatically increases the efficiency of model construction and avoids redundant object definition. The model structure further ensures that during simulation the failure of a sub-model will not jeopardize the entire model simulation. Thus, it enhances the robustness of a model.
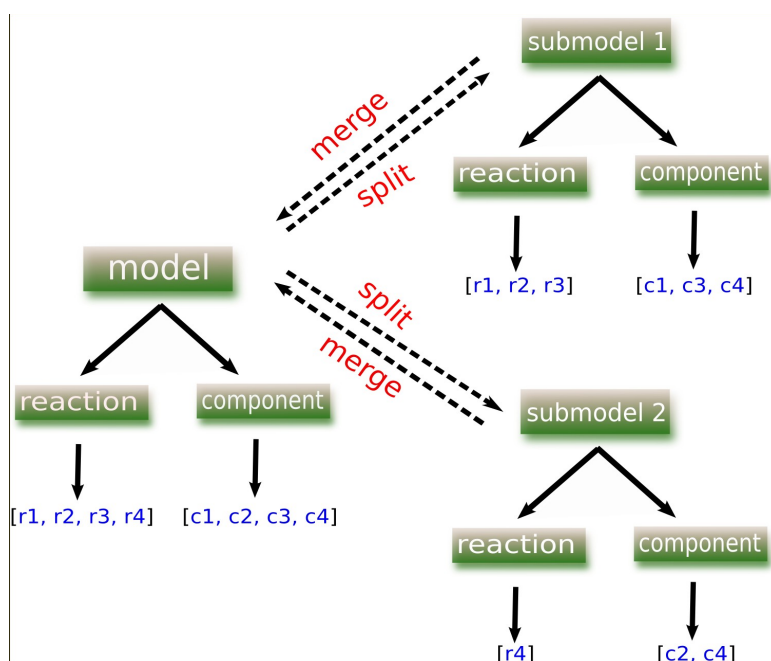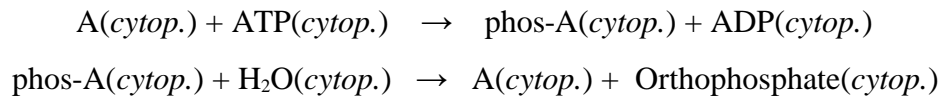
Figure 26. The internal structure of a model in PyBioS2. Based on this tree structure, a model in PyBioS2 can be easily split into different sub-models and vice versa.

102

2.15.5 Semi-Automatic Definition of Reaction

One of the major goals for PyBioS2 is to construct genome-scale networks and to analyze the dynamic behavior of the underlying biological system. Connections with multiple databases can certainly help to achieve this goal. Furthermore, PyBioS2 enables the user to define biochemical reactions with semi-automatic options, which reduces tedious definition work, improves construction efficiency and enhances the consistency of model data. This is an unique option of PyBioS2 for constructing large-scale network of cellular systems.

For instance, if the user selects protein A (*cytoplasm*) for defining a phosphorylation reaction, the phosphorylation and the corresponding dephosphorylation reaction will be automatically defined in the model:

$$A(cytop.) + ATP(cytop.) \rightarrow phos\text{-}A(cytop.) + ADP(cytop.)$$

$$phos\text{-}A(cytop.) + H_2O(cytop.) \rightarrow A(cytop.) + Orthophosphate(cytop.)$$

If the user selects protein B (*cytoplasm*) for defining a translocation reaction with destination *nucleoplasm,* the translocation reaction will be automatically defined in the model:

$$B(cytoplasm) \rightarrow B(nucleoplasm)$$
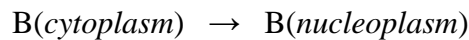
Table 13 lists all types of reactions with a semi-automatic option. (The corresponding definition patterns are explained in appendix A.7). More importantly, this option can be extended to all kinds of reactions.

Reaction Type for Semi-Automatic Definition			
Activation	ActivationGTP	Complex Formation	Self Formation
DeActivationGTP	Decay	Dephosphorylation	miRNA_binding
Methylation	Oxidization	Auto.Phosphorylation	Phosphorylation
Transcription	Translocation	Ubiquitination	

Table 13: PyBioS2 enables the user to define certain types of reactions semi-automatically. The list contains all types of reactions that are allowed to be created semi-automatically in PyBioS2. With this option, it is comfortable for the users to manually and systematically construct a large size model within a short time period. (the table and its legend is taken from Li et al., 2012b)

2.15.6 ODE System

Ordinary Differential Equations (ODEs) are the most common approach for modeling systems of biochemical reactions (Sachs et al., 2001; Kestler et al., 2008). By selecting ODE for simulation, PyBioS2 can automatically translate the reaction-entities into ODE mathematical formulas, based on pre- or user-defined kinetics from global or user-specific kinetics database respectively.
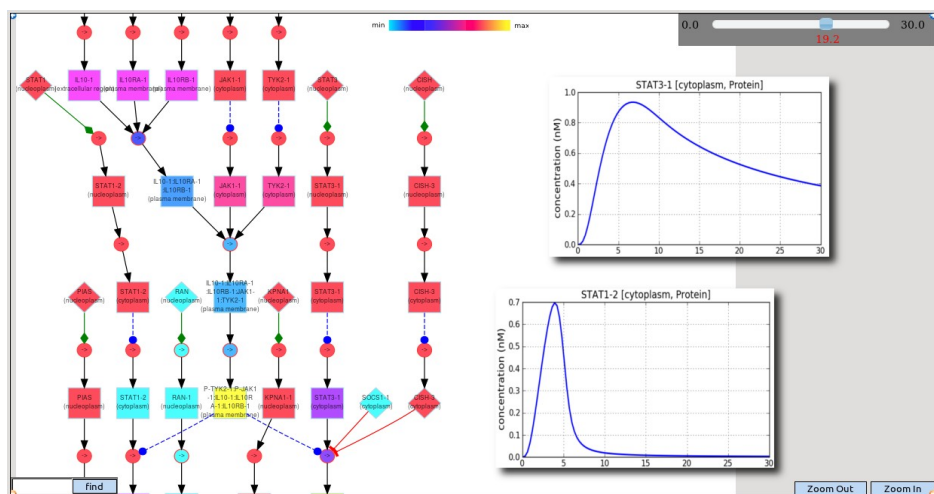
Figure 27. The visualization of model network under ODE system module. The slider on the top right is used to review any point in time of the simulation process and the concentrations of model components will be changed according to the selected point in time.

Each equation in ODE describes the rate of change in the concentration of a particular enzyme, substrate, or signaling molecule as a function of its production, degradation, and reaction with other network components. The created ODE system can be used for the subsequent simulation of time course series and for further analysis of the dynamic behavior of the underlying system. PyBioS2 deploys the LSODE solver to run the numerical integration process of ODE system (Hindmarsh, 1983). Users can review the entire simulation process step by step and observe the dynamic behavior of a model (Fig. 27). Furthermore, with the help of the ODE system, PyBioS2 is able to perform different

mathematical analyses including metabolic control analysis, stead-state analysis, parameter scanning, and others (details in appendix A.6).

2.15.7 Petri net

Petri net is a graphical and mathematical modeling language developed in the early 1960s by Carl Adam Petri (Petri, 1962). The Petri net approach has been subsequently adapted and extended into many fields such as systems biology (Hardy & Robillard, 2007). Many extensions to Petri net have been developed for various modeling and simulation purposes (Bernardinello et al., 1992; Hardy & Robillard, 2004; Li et al., 2006). With Vikash Pandy together, I developed a Petri net extension for PyBioS2, which includes the characteristics of hybrid, timed and stochastic Petri net. Therefore, it is particularly suitable for the simulation of large scale networks. In PyBioS2, if one selects Petri net for simulation, each component-entity (gene, RNA, protein, compound, complex and others) and reaction-entity will be automatically defined as 'place' and 'transition', respectively. The formal definition of Petri net (Murata T, 1988) is explained in the section Materials and Methods.

Each place is marked with a value representing the concentration of component-entity in a model and each place has its decay function with mass action law as its default kinetics. Moreover, pre- or user-defined kinetics are associated with corresponding transitions. During the simulation, when a transition is evaluated as 'active', the flux is then calculated according to the kinetics of the transition, the weight function of all participants and the marking function of all substrates. For instance, if a mass action law with kinetic parameter k is assigned to a phosphorylation reaction catalyzed by an enzyme:

$$\text{BAD(cytop.)} + \text{ATP(cytop.)} \quad \xrightarrow{\text{E}} \quad \text{phosp-BAD(cytop.)} + \text{ADP(cytop.)}$$
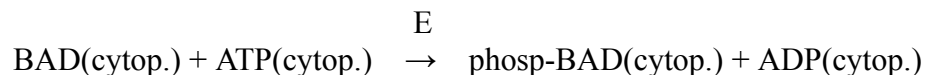
the corresponding flux is F = k * m(BAD, t)/W(BAD) * m(ATP, t)/W(ATP) * m(E, t). This flux will thus change according to the current concentrations of participant-entities at the specific point in time. We also considered the reversibility of a reaction-entity and if a flux becomes negative, the backward of a reaction occurs. Thereby, Petri net is endowed with the hybrid property.

Many cellular processes occurring in reality are inherently stochastic, therefore stochastic methods are more suitable and accurate for system-wise analysis of a biological system because they consider the intrinsic stochasticity of a biological system and do not presume continuous concentrations of biological components (Hoops et al., 2006). Therefore, I added the stochastic property to Petri net for PyBioS2. If one assigns the stochastic effect to a reaction-entity in a model, then the corresponding transition gets a threshold "g" and will only fire duration simulation when a randomly sampled value exceeds g (the sample-distribution is currently a normal-distribution). For default setting, only transcription- and translation-reaction in a model have the stochastic effect with a threshold of 0.5.

It is noteworthy that time-scales of biochemical interaction within intracellular networks generally span multiple orders of magnitude. Typically, signaling and metabolic reactions occur fast (less than a second), for example, kinase/phosphatase reactions, and most metabolic reactions occur on the order of fractions of a second to seconds (Papin et al., 2005). In contrast, receptor internalization (Lauffenburger et al., 1993) and regulatory events (Kaznessis et al., 2006) usually require several minutes to several hours for completion. In order to reflect the biology of these multiple time scale, we endowed Petri net with the delay property. In general, each transition has a delay attribute with a default value of 1. If delay-attribute of a transition is set to 20, this transition will fire every twenty time points during the simulation. In this way, the effect of multiple time scales can be created during the simulation process (Fig. 28).
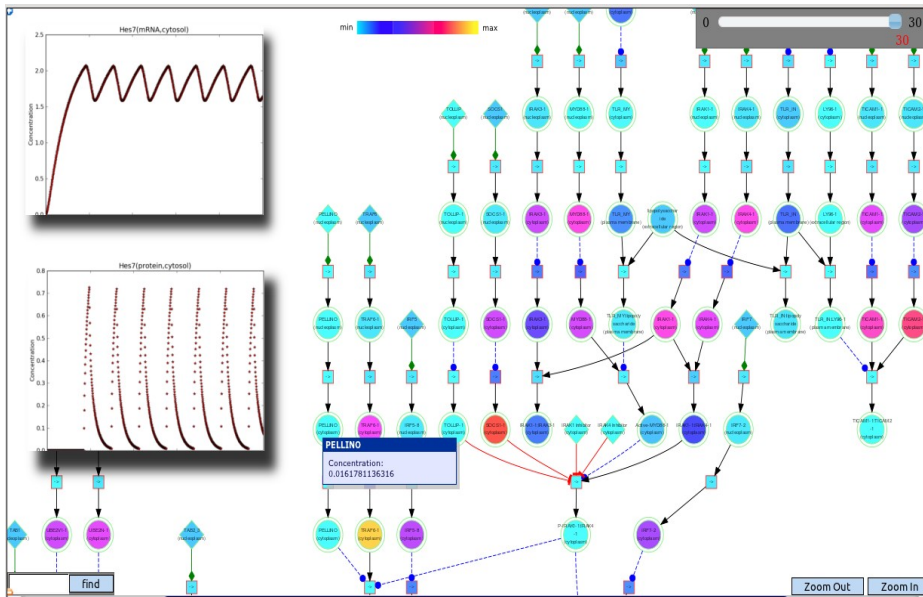
Figure 28. The visualization of model network under Petri nets module. Similar to the ODE system, the slide on the top right is used to review the simulation process at any time point. (The unit of concentration in the time-course graphs is nM)

2.15.8 Utilization of Gene-Expression Data

Thanks to the recent advances in high throughput technologies, an abundance of biological information has been generated such as gene expression-, metabolic-, proteomic-data. This leads to an immediate task: to develop powerful, effective analysis and simulation strategies for integration and analysis of these biological data to make discoveries and to predict behaviors of biological systems under different perturbation conditions (Smith et al., 2011). Initializing a model in PyBioS2 with different expression data is similar as giving the underlying biological system different specific perturbations, which may be genetic (e.g., gene over-expressions, gene mutation, or loss of gene function) or environmental (e.g., change environment condition, adding hormones or drugs). The corresponding response to each perturbation can be closely examined or

107

revealed throughout well-designed simulation processes. All model components can be monitored due to the advanced simulation facility of PyBioS2. The comparison between simulation output and experimental data can evaluate the predictive ability of a model which can subsequently be used as a predictive guidance to discover new properties of the biological system. Furthermore, PyBioS2 provides the user with the opportunity to visualize the expression data in the model network. By means of this mapping, the user will be able to explore specific useful information and interesting questions from the network point of view.

2.15.9 Model Analysis Approaches

PyBioS provides diverse approaches to enable users to conduct detailed analyses of the model system such as a time-course analysis, metabolic control analysis, parameter-scan, connectivity analysis, network-cluster analysis, closeness- and betweenness-analysis. The IGRAPH (http://igraph.sourceforge.net/) graph library is integrated in PyBioS2 and applied to analyze graph-theoretic properties like degree, closeness- and betweenness-analysis of a network. The network format of a model is changed into the IGRAPH format based on list of directed edges of a network. Table 14 summarizes all relevant functions in PyBioS2 for analyzing the model network structure.

Analysis	Functions
Connectivity Analysis	Analyzing the gaps of a model and how well different model components connect with each other (if the result shows only one island of an entire model, it means that there is no gap in the model).
Summary of Network Information	Summarizing the basic network statistic of the current model including the amount of nodes, edges; the diameter of network; the density of network; the reciprocity of the network; the average length of network

	paths; the clustering coefficient of network.
Indegree- and Outdegree Analysis	Analyzing the number of edges to a vertex, since the interaction network is directed, therefore, there are indegree and outdegree values of each vertex.
Closeness- and Betweeness Analysis	Analyzing how easily other vertexes can be reached from a vertex (closeness) and a measure of a vertex's centrality in a network equal to the number of shortest paths from all vertexes to all others that pass through this vertex (betweenness).
Clusters Analysis	Calculating the clusters of the current model network.
Burt Constraint Scores	Calculating the Burt's constraint scores for given vertexes, also known as structural holes.
Eigenvector Centrality	Measuring the influence of a node in the model network.
Farthest Vertex of Network	Finding farthest source-node and target-node of the longest path of the current network.
Shortest Circle of Network	Finding the shortest circle of the current network.
Self-loops and Multiple-edges	Detecting the self-loops and multiple edges of the current network.
Math. Info	Obtaining different types of mathematical information about a model system such as Stoichiometric Matrix of model reactions and Conservation Relations of a model system.
Identification of External Components	For a model system (especially a large-scale model), it is of advantage to identify the model components that have only out-going edges or only in-coming edges indicating

	that those components are only producing or being consumed.

Table 14. PyBioS2 integrates different in IGRAPH defined functions to provide users with diverse network analysis approaches.

2.15.10 A Case Study: Comparison of Petri-net simulation strategy and ODE simulation strategy.

To the present day, PyBioS2 is the only software, which provides users two simulation approaches: ODE- and Petri net-based simulation. Therefore, it is essential to demonstrate a basic comparison between these two simulation approaches and investigate if simulation results could be different due to the principle-difference of these two simulation approaches. Vikash Pandey and I constructed a simple biological network consisting of transcription, translation, complex formation and decay reactions (appendix B.9A). The applied kinetic law was mass action law. We simulated the model with these two simulation approaches for a period of 500 time steps. The simulation results (appendix B.9B and appendix B.9C) show that all components in the model behave in a very similar manner during the simulation. Furthermore, we also performed this type of comparison for several other biological networks (data not shown). All comparison results indicate that both simulation approaches provide qualitatively identical results.

3. Materials and Methods

3.1 PETRI NET EXTENSION

The simulation processes applied during this thesis are based on the Petri net, which is a graphical and mathematical modeling language developed in the early 1960s by Carl Adam Petri (Petri C.A., 1962). Subsequently, the Petri net approach has been world-

widely extended into diverse study fields (Heiner et al., 2008) To the present day, Petri net-based approaches are adapted and extended in the field of systems biology to order to perform mathematical modeling for better understanding of the nature of biology (Bernardinello et al., 1992; Hardy & Robillard, 2004; Li et al., 2006). During the study of Li et al. (2012a), a Petri net extension was implemented, which was incorporated into the PyBioS2 software (Li et al., 2012b). The properties of this petri net extension in PyBioS2 includes the characteristics of hierarchical-, hybrid- and timed-Petri net. Its goal is to enable users for analyzing large-scale networks. In the following, the detailed simulation process is explained. The Petri net definitions from the study of Murata T. (1988) are applied:

A Petri net is a 6-tuple, $PN = (P, T, F, W, m, D)$

where:

$P = \{ p_1, p_2, \ldots, p_x \}$ is a finite set of places (model contains x bio-objects),

$T = \{ t_1, t_2, \ldots, t_y \}$ is a finite set of transitions (model contains y reactions),

F is subset of $[(P \times T) \cup (T \times P)$ is a set of arcs]

$W: F \rightarrow \{0, 1, 2, \ldots x \}$ is a weight function,

$m: P \rightarrow \{0, 1, 2, \ldots x \}$ is the initial marking,

$P \cap T = \varnothing$ (meaning that the sets P and T are disjointed),

$D: m \rightarrow P \rightarrow \{p_1, p_2, \ldots, p_x\}$ is the decay rate function of each component.

A biological reaction system (model system) can be translated into a Petri net-based model system, where each reaction in the biological model is replaced with the transition the model system. Similarly, each object (gene, protein, complex, etc) from the biological system is symbolized with the place of the model system (Fig. 29) (Li et al., 2012a). The following is a simple translation example:

A biochemical reaction system   Petri net

  $A + B \dashrightarrow C$ (v1)   <====>   Places: [A, B, C]
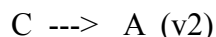
C ---> A (v2)   Transitions: [v1, v2]

Each place is signed with a value presenting the concentration of this corresponding model component, and each transition has a value presenting the flux of this corresponding reaction. Suppose that mass action law kinetics is applied to the reaction v1 and v2, the flux intensity calculated by the time point t for the reaction v1: flux $F1(t) = k1 * (m(A, t) / W(A)) * (m(B, t) / W(B))$, where k1: the kinetic parameter of this reaction; m(A,t): the concentration of model component A at point in time t; W(A): the weight function of model component A.
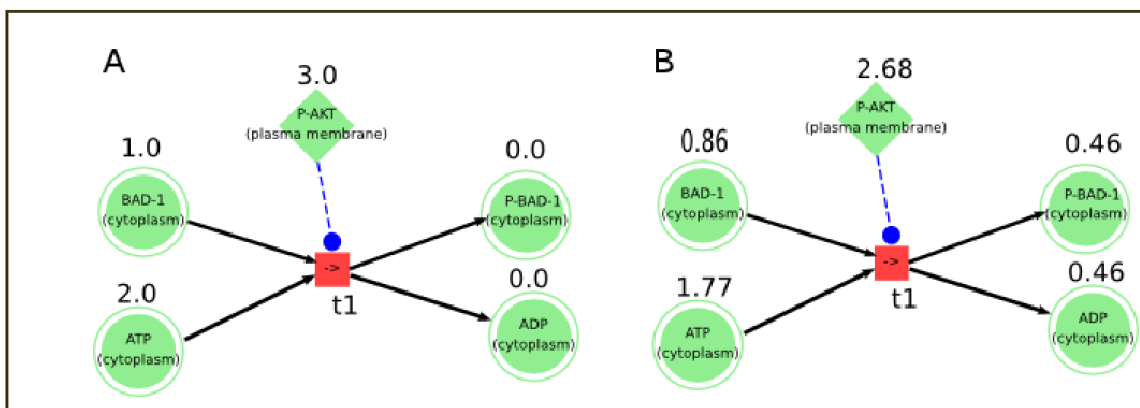
Figure 29. Transition fire of a Petri net. An exemplary phosphorylation reaction of the Petri nets. BAD-1 (p1), ATP (p2), P-BAD-1 (p3), ADP (p4) and P-AKT (p5) are places and phosphorylation reaction (t1) is transition. **A:** the concentrations of model components at point int time 0 (before transition fire); **B:** the concentrations of components in model at point in time 1 (after transition fire) (This figure and its legend is taken from the study of Li et al., 2012a).

During *in silico* simulation process, each transition is evaluated with a certain criterion defined within the Petri net-based model. When a transition is evaluated as 'activated', then it will fire during a time step within the simulation, referring to the biological sense, the corresponding reaction takes place, which is calculated by individual functions F_y(S,

E, I) (S: substrate, E: enzyme and I: inhibitor) (Li et al., 2012a). For instance, the speed of transition t1:

F(m(BAD-1), m(ATP) and m(P-AKT)) =

(m(BAD-1)/W(BAD-1)) * (m(ATP/W(ATP)) * (m(P-AKT)/W(P-AKT))

where m(BAD-1), m(ATP) and m(P-AKT) are concentrations of places BAD-1, ATP and P-AKT; W(BAD-1), W(ATP) and W(P-AKT) are the weights of the places. If the following conditions hold true:

speed = F(m(BAD-1), m(ATP) and m(P-AKT)) > 0;

m(BAD-1))>=speed; m(ATP)>=speed

the transition t1 fires in forward direction, meaning that this phosphorylation reaction takes place. If the following conditions hold true:

speed = F(m(BAD-1), m(ATP) and m(P-AKT)) < 0;

m(BAD-1))<=speed; m(ATP)<=speed

the transition t1 fires in backward direction, meaning that the dephosphorylation reaction takes place. If one of the two conditions above is fulfilled, but the enzyme P-AKT is present at a low concentration, the reaction can take place. However, its reaction flux will be low, too.

For example, at time point 0, initial concentrations of BAD-1, ATP and P-AKT are 1.0, 2.0 and 3.0, and I assume that their weight function values are

W(BAD-1) = 5,

W(ATP) = 5,

W(P-AKT) = 5,

then the decay function values will be

D(BAD-1) = m(BAD-1) * 0.09 = 0.09,

D(ATP) = m(ATP) * 0.09 = 0.18,

D(P-AKT) = m(P-AKT) * 0.09 = 0.27,

and the speed is calculated by f1:

113

speed = F(m(BAD-1), m(ATP) and m(P-AKT)) = 0.05

Tables 15, 16 show the concentration-changes of components of reaction t1 from the point in time 0 to the point in time 1. The unit of all concentrations used in this example is arbitrary unit.

Component	Concentration (a.u.)
BAD-1	1.0
ATP	2.0
P-AKT	3.0
P-BAD-1	0.0
ADP	0

Table 15: Transition t1 at the point in time 0 (before firing). (this table and its legend is taken from Li et al., 2012a)

Component	Concentration (a.u.)
BAD-1	0.95 – 0.09
ATP	1.95 – 0.18
P-AKT	2.95 – 0.27
P-BAD-1	0.05 – 0.05*0.09
ADP	0.05 – 0.05*0.09

Table 16: Transition t1 at the point in time 1 (after firing). (this table and its legend is taken from Li et al., 2012a)

It is noteworthy that multiple time scales of a biochemical reaction system have been taken into consideration. Therefore, the degradation processes of different model components are differed. For example, protein components have degradation function D(m(proteins)*0.09) with 0.09 as degradation kinetic parameter. This degradation function D is associated with its corresponding degradation reaction which is taking place each time step; while complex components have degradation function D(m(ligand-receptor complex)*0.2) with 0.2 as degradation kinetic parameter. This degradation function is associated with its corresponding degradation reaction, which is taking place

every 10 time steps. Currently, these different kinetic parameters are based on the empirical experience and only consider the speed magnitude of biochemical reactions listed in the study of Papin et al., (2005). All types of kinetic parameters applied in Petri net simulation are summarized in following Table 17 (Li et al., 2012a).

Reaction Type	Kinetic Parameter	Reaction Type	Kinetic Parameter
Complex Formation	0.55	Translocation	0.8
Phosphorylation	0.35	Translocation	0.5
Dephosphorylation	0.01	Dephosphorylation (enzym)	0.15
Transcritpion	0.5	Inhibitory	1.5
Decay	0.02	other parameter	0.35

Table 17: The summary of kinetic parameters applied in the Petri nets simulation. (this table and its legend is taken from Li et al., 2012a)

3.2 ANALYSIS OF SIGNIFICANT CHANGES OF MODEL COMPONENTS DUE TO THE EFFECT OF ANTI-MIRNAS

During the study of Li et al. (2012a), significance analysis of concentration changes with regard to all model components were performed using t-test (Markowski and Markowski, 1990). Assuming that the model contains n components and C_X symbolizes the simulated concentration of one model component and for the state of the application of anti-miRNA (e.g. anti-mir-21)

$$V_{anti-mir} = [\ C1, C_2, C_3, \ldots C_n\ ],$$

where C_x is the concentration of one model component in the current state, and for the corresponding control state (e.g. mir-21)

$$V_{control} = [C'_1, C'_2, C'_3, \ldots C'_n\ ]$$

where C'_x is the concentration of one model component in the control state the p-value is copmuted by the t-test:

$$P\text{-value} = t.test\ (V_{anti-mir}, V_{control})$$

3.3 MODEL CONSTRUCTION

In this thesis, I implemented a global human signaling network consisting of more than 50 different cancer-related signaling pathways with more than 18,000 biochemical reactions and 1,084 human miRNAs with their validated target information (appendix A.4). The modeling of the biogenesis and regulation mechanism of miRNA was explained in the previous study (Li et al., 2012a). The model integrated three types of networks: signaling, gene-regulator and miRNA-regulation network. Different object types including gene, RNA (mRNA, miRNA), protein, protein complex and compound are defined in the model. I translated the model network into a mathematical model by applying a Petri net application with mass action kinetics. The model is available in appendix A.11.

3.4 MODEL PARAMETERIZATION

The in-silico simulation approach for revealing or predicting the dynamics of a biological model system are largely based on the specific information such as kinetic parameters and environmental conditions. However, the kinetic parameter of a certain biochemical reaction is tightly dependent on the related environmental or experimental condition and can change time by time (Krebs et al., 2007). Due to this limitation of many unknown kinetic parameters of biochemical reactions, I applied the Monte Carlo-based strategy explained in the study of Wierling et al., 2012 and extended it into Petri net simulation process for the parametrization of the model.

3.5 MODEL INITIALIZATION WITH GENE-EXPRESSION OR MIRNA-EXPRESSION DATA AND MODEL SIMULATION PROCEDURE

The model contains different object types including gene, RNA (mRNA, miRNA), protein, complex compound and pseudo-object. Each of these types is associated with a corresponding Id. For instance, gene- and mRNA-object are Ensembl-Id; protein- and compound-object are associated with uniProt-Id and ChEBI-Id, respectively; miRNA-

and miRNA gene-object are associated with the miRNA accession; other objects are associated with internal model Id. The basic biological relation is ensured by defining that each gene object participates in a corresponding transcription reaction, which generates an mRNA object. An mRNA object takes part in a translation reaction, which creates a corresponding protein object. All objects except the gene object are set to 0. During the simulation process, the model signal is only generated from the transcriptional level and forwarded to the translational level. Afterwards, the signal can be propagated to the rest of the model.

The initialization procedure is preformed according to different types of Ids. If the gene-expression data is available, gene objects in the model can be iteratively assigned with a corresponding gene expression value according to Ensembl-Ids. Similarly, for the miRNA expression data, the miRNA gene objects in the model can be iteratively assigned with the data value according to miRNA accession. I initialized the expression data with the gene-objects, not with RNA-objects, because the transcriptional and post-transcriptional regulation needed to be considered in simulation experiments.

Pseudo-source Code:

```
1.      m ← signaling model
2.      Initialize(m){
3.          if gene expression data is available{
4.              gene-entities ← m.getAllGene_Entities()
5.              for gene-entity in gene-entities{
6.                  set gene-entity with data according to the Ensembl id
7.              }
8.          }
9.        if miRNA expression data is available{
10.             miRNA-entities ← m.getAllmiRNA_Entities()
11.             for miRNA-entity in miRNA-entities{
12.                 set miRNA-entities with data according to the miRNA id
13.             }
14.        }
15.      set all other entities in the model to zero.
16.    }
17.      simulate(m){     # signal propagation process
18.      perform_petri_net_simulation()  }
```

3.6 COMPARISON OF **P**REDICTED **P**ROTEOME AND **E**XPERIMENTAL **P**ROTEOME **D**ATA

For the comparison of protein concentrations from the simulated steady-state with the protein concentrations measured in experiments, I considered all modified states of corresponding proteins including methylation, phosphorylation and complex formation. The sum of steady-state concentrations from all states of a corresponding protein was compared to the concentration derived from experimental measurement (Nagaraj et al., 2011 and Lundberg et al., 2010).

3.7 ESTABLISHMENT OF **A** **D**RUG **D**ATABASE

In the course of this thesis, I designed and constructed a drug database, which contains the dissociation constant value of the interaction between drugs and their corresponding molecular targets. Most of this drug inhibition information is based on the quantitative analysis result of a number of kinase inhibitor selectivities from the study done by Karaman et al., (2008). Fig. 30 shows the diagram of this drug database. The dissociation constant value of binding affinity between inhibitors and their corresponding protein target (Kd_Value) of this database is used as a part of the input for the simulation pipeline (Fig. 19).
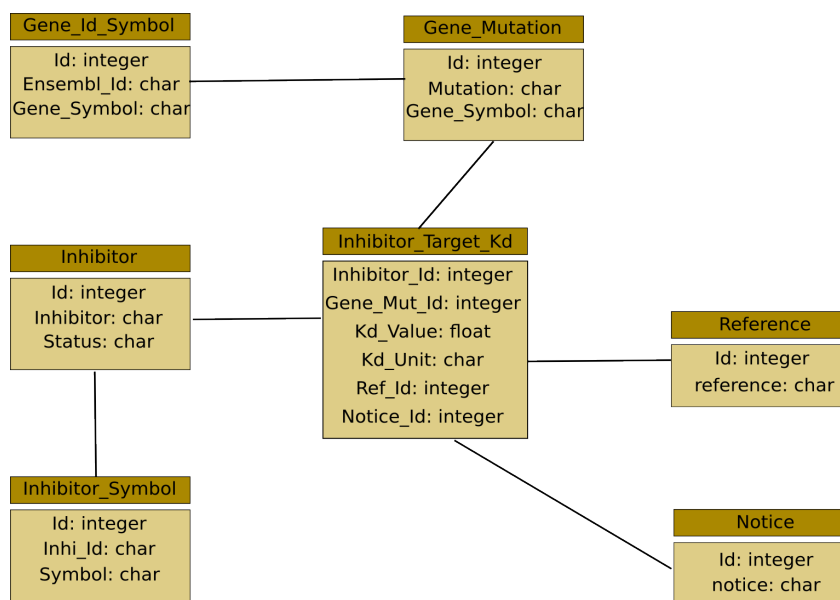
Figure 30. The drug database diagram. The major data table of this database is 'Inhibitor_Target_Kd', which records the dissociation constant value of binding affinity between inhibitors and their corresponding protein target. Other data tables such as Inhibitor, Gene_Mutation, Gene_Id_Symbol record corresponding information.

3.8 WEIGHT-FACTOR ESTIMATION FOR THE TUMORIGENESIS CALCULATION FORMULA

<1> m = signaling model
$$T = (k_1*X_1+k_2*X_2+k_3*X_3+k_4*X_4+k_5*X_5+k_6*X_6+k_7*X_7+k_8*X_8+k_9*X_9)/(1+k_{10}*X_{10})$$

$k_1,k_2...k_{10} = 1$
Initialize(m) # set gene-entities with gene-expression data through the Ensembl ID
　　　　　　　# set all other entities in the model to zero.

<2> Define the best solution: $F = f_1 \wedge f_2 \wedge ... \wedge f_{59} \wedge f_{60}$
f_1 = minimize(T)　　　　# for the cell line EKVX　　　　(responder)
f_2 = maximize(T)　　　　# for the cell line NCI-H23　　　(non-responder)
...
f_{59} = minimize(T)　　　　# for the cell line IGROV1　　　(responder)
f_{60} = maximize(T)　　　　# for the cell line OVCAR-5　　　(non-responder)

<3> While not meet the conditions of F:
k_1 = random_generate_number()
k_2 = random_generate_number()
...
k_{10} = random_generate_number()
continue to test the randomly generated weight-factors

<4> Return the optimzied weight factors $k_1,k_2...k_{10}$

3.9 CANCER CELL LINES FROM NCI-60 STDUY AND FROM CANCER GENOME ATLAS

The gene-expression data of NCI-60 cancer cell lines can be downloaded under

http://discover.nci.nih.gov/datasets.jsp

The gene-expression data of breast, colon and lung cancer cell lines from Cancer Genome Atlas can be downloaded under http://tcga-data.nci.nih.gov/tcga/tcgaHome2.jsp

3.10 SENSITIVITY SCORE DEFINITION

1. Sensitivity score (experimental data) = MGI_{50} (MGI_{50}: mean value of GI_{50} (the log mol/L drug concentration yielding a growth percent of 50, i.e. 50% growth inhibiton)).

2. Sensitivity score (simulation data) = K * P (K: multiple factor, currently estimated value is 4; P: relative change of readout component 'proliferation' to its value of the no-drug (control) state). In this case, I use the cancer hallmark 'proliferation' as readout component instead of the putative readout component 'tumorigenesis' because the drug response prediction here is referring the growth potential of cancer cell lines.

3.11 THE IMPLEMENTATION OF CANCER- AND PROCESS-HALLMARKS IN THE HUMAN SIGNALING MODEL.

The cancer- and process-hallmarks are defined as pseudo components in the human signaling model. Each of them should be able to summarize the signal from signaling pathways and presents a developmental aspect of the cancer cell physiology. The modeling implementation of these hallmarks is listed below with the literature references. The mass action law is applied for the integration implementation. Currently, the influence degrees of components on cancer hallmarks have not been taken into consideration.

Hallmark	Implementation	Reference (PubMed)
Apoptosis (A)	A= Active-CASP3 * (AIFM + ENDOG + TP53 + TP63 + TP73 + P-PAK2_P34 + HIPK2)	10200555,16239930,17626635,21248071,18767140,20193641
Cell-cycle-progression (CC)	CC = (P-CDKN1B + P-CDKN1A + CCND1:CDK4_6 + CCDN2:CDK4_6 + CCDN3:CDK4_6 + CCNE:CDK2_3 + Active-CCNB:CDK1) / (1 + INK_ACD + CDKN2B + BRCA2:RAD51 + Active-CHEK1 + BRCA1:BARD1)	21045237,19106607,18406353,16522651,8840967,20202217,15665856,19238148
Drug-resistance (D)	D = P-RAD51_T315 / (1 + XAF1)	11087668,2085543,21807066,17329253
Evading-apoptosis (EA)	EA = (TNF1:TNFRSF1B + P-IGF1R:IGF1 + P-IGF1R:IGF2) / (1+ FAS:FASLG + TNFSF10:DR4_5 + TNF1:TNFRSF1A)	14634624,17846171,21608150,20182539,8524870
Insensitivity-to-anti-growth-signals (IS)	IS = (SMAD2_3:UBB + SMAD1_5_8:UBB + I_SMAD:SMAD4) / (1 + P-SMAD2_3:SMAD4 + P-SMAD1_5_8:SMAD4)	19114990,14534577,22710166
Limitless-replicative-potential (LP)	LP = TERT / (1 + SETD2)	16869755,21792193,22396899
Proliferation (P)	P = P-EIF4EBP1 + P-MYC:MAX + URGCP + MKI67 + TRIM21 + MYBL2 + TOP2A + STK6 + PLK1	15184677,10430922,12454650,12620412,11018017,17217616
Self-sufficiency-in-growth-signals (SS)	SS = mut-BRAF + BRAF:mut-KRAS + BRAF:mut-HRAS + BRAF:mut-NRAS + mut-EGFR + mut-ERBB2	19370421,22613949,19855393,21779505,16397024
Sustained-angiogenesis (SA)	SA = (P-FGFR:FGF + P-KDR_dimer:VEGFC + P-KDR_dimer:FIGF + P-KDR_dimer:VEGFA) / (1 + CD36:THBS1)	18560389,17933680,20010945,21248359,21742222

| Tissue-invasion (TI) | TI = MMP7 + MMP9 + MMP2 + MMP1 + MMP13 + MMP10 | 14967450,16680569,1134 9215,11344033 |

Table 18: Cancer hallmark implementation. P-PAK2_P34 symbolizes that the PAK2 protein is phosphorylated at 34th site. CDK4_6 indicates an object-entity containing CDK4 and CDK6. mut-BRAF signifies the mutant version of BRAF (in this case, the gain-of-function). The reference ID is PubMed ID.

3.12 MATHEMATICAL IMPLEMENTATION OF COX-ISOFORM SPECIFIC SIRNA AND NS-398 DRUG EFFECT

CT Comparison	Tumor + COX-2 siRNA State	Tumor State
Kinetic Rate Law	$v = [COX\text{-}2(mRNA)] * K_{translation} / [siRNA]$	$v = [COX\text{-}2(mRNA)] * K_{translation}$

NT Comparison	Tumor + NS-398 State	Tumor State
Kinetic Rate Law	$v = [COX\text{-}2(Protein)] * K_{decay} * [Inhibit1]$ $v = [VEGF(Protein)] * K_{decay} * [Inhibit2]$ $v = [IL1(Protein)] * K_{decay} * [Inhibit3]$ $v = [TNF(Protein)] * K_{decay} * [Inhibit4]$	$v = [COX\text{-}2(Protein)] * K_{decay}$ $v = [VEGF(Protein)] * K_{decay}$ $v = [IL1(Protein)] * K_{decay}$ $v = [TNF(Protein)] * K_{decay}$

Table 19: Mathematical implementation of therapeutic intervention. The difference of mathematical implementation of comparative analysis of different states (perturbation state vs. control state). The effect of drug NS-398 is now composed of the effect of inhibit1-4. (this table is taken from Li & Mansmann, 2013)

4. Discussion

4.1 THE MIRNA-EGFR MODEL

In 1993, a new class of non-coding RNAs with 18~22 nucleotide long has emerged, named microRNA (miRNA), was discovered by the study of Lee et al. (1993). Since then, the biological function and role of these tiny molecules in tumorigenesis and cancer development has been a subject of intense investigation (Ryan et al., 2010). The study of Hornstein & Shomron (2006) proposes an interesting concept for investigating biological roles of miRNAs is to detect miRNA impact in the context of an integrated network from the summation of the interactions of miRNAs and their targets. Therefore, during the course of this thesis, a molecule-based miRNA-EGFR (ME) model was created by applying the system-biological software PyBioS2 (Li et al., 2012b). The validated miRNA-target information was manually collected from results of diverse independent studies (appendix A.2) and successfully integrated into the ME model (Li et al., 2012a). For a cellular system, a signal induction through the receptor-ligand association can induce specific cellular instructions, however, the specificity is entirely dependent on the histories and actual environments that the cell has. The aim of the study of Li et al. (2012a) is to reveal the potential that a molecule-based model of signaling pathways might be able to help us to understand the cellular instruction making process. At first step, the experimental conditions related to miRNA's function from the *in vitro* study of Avraham et al. (2010) were applied for an *in silico* simulation process based on the ME model. The simulated data was then compared with the measured experimental data with high accuracy, which validated the predictive ability and structure quality of this model (Li et al., 2012a). Furthermore, during the study of Li et al. (2012a), different *in silico* experiments were designed for demonstrating that miRNAs targeting specific pathway components, play essential roles in the signaling cascade triggered by EGF, for instance, by increasing the expression level of mir-192 within the model, it is shown that not only

123

the expression of its direct targets were decreased, but also the activities of downstream pathway components were dramatically reduced, which indicates the impact of this miRNA on the entire signaling pathway. Moreover, it is also shown that mir-181c was able to tightly control the threshold of the signaling cascade of EGFR pathway, which implies the critical role of this miRNA for optimization of cascade signal efficacy and restriction of undesired signaling fluctuation during the signaling transduction process. These results and phenomenon demonstrate the effect of this concept "One Hit – Multiple Targets" (Li et al., 2012a).

In order to further reveal the potential of the therapeutic concept "One Hit – Multiple Targets", global effects of 100 individual anti-miRNAs on the EGFR signaling pathway were analyzed through an *in silico* experiment. As partially expected, the result of this *in silico* experiment shows that more targets a miRNA possess within the EGFR signaling pathway, more likely this miRNA can exert considerable influence on this pathway. However, it is also shown that miRNA targeting specific key pathway components can also have significant effect on the pathway, even its number of targets is less. Therefore, the conclusion is the impact of miRNA is not entirely dependent on the number of its targets. In addition, according to results of the in silico experiment for analysis of anti-miRNA, a top-15 list of anti-miRNAs with significant impact on the EGFR signaling pathway was proposed during the study of Li et al. (2012a), which also prepared the excellent miRNA candidates for the follow-up studies to access the distinct impact of miRNA on the EGFR pathway with regard to cancer research. One of those particular interesting candidates is mir-21. Diverse studies provide evidence that mir-21 acts a critical oncogenic role in many types of cancer including breast-, colon- and cervical-cancer. Its high expression level is often correlated with poorer survival of patients (Chan et al., 2005; Tong et al., 2008; Iguchi et al., 2010). The result of this *in silico* experiment indicates that inhibition of this mir-21 could result in a significant up-regulation of pro-apoptotic proteins such as CASP9, BAD, TP53, which is in line with the result achieved

by the *in vitro* study of Chan et al. (2005) for demonstrating that mir-21 is indeed an anti-apoptotic factor. In addition, it is also shown the up-regulation of a putative tumor suppressor, CDKN1, which regulates the cell cycle progres (Iravani et al., 1997; Hemmati et al., 2010), can be caused by the inhibition of mir-21 to support the oncogenic activities of this miRNA. The study of Mei et al. (2010) provide strong evidence that inhibition of mir-21 can enhance the chemotherapeutic effect in breast carcinoma cells, which again reveals the predictive ability of the ME model (Li et al., 2012a).

Another of the top-15 anti-miRNAs resulted from this *in silico* experiment is anti-mir-335, whose activation within the ME model could strongly the upregulation of many essential tumor suppressors including CDKN1, TP53, CASP9 and TSC2. This result indicates the relevant oncogenic role of mir-335, which act in agreement with the result achieved by the independent study of Shi et al. (2011). In the study of Shi et al. (2011), authors validated that inhibition of growth and invasion of malignant astrocytoma cells can be achieved by repressing the expression level of mir-335 and supported the potential oncogenic role of this miRNA. Another two of top-15 anti-miRNAs are anti-mir-221 and anti-mir-222 and both were predicted to be significant influential in the EGFR signaling pathway. Interestingly, the study of Stinson et al. (2011) substantiated the potential role of both miRNAs for promoting malignant transformation into more aggressive cancer phenotypes with poorer diagnosis. All those results from different independent studies strongly suggest that the predicted results produced by ME model are highly relevant in cancer treatment, which indicates high predictive ability of this model (Li et al., 2012a).

It is noteworthy, according to statistical analysis, the ME model contains 130 miRNA targeting DICER, 45 miRNAs targeting RNASEN (appendix A.2). Both target proteins play key roles within the miRNA biogenesis. The misregulation of both proteins can induce a globally impaired expression of mature miRNAs, which can result in different pathological states in the human body (Lu et al., 2005; Kumar et al., 2007). Therefore, it

is reasonable to consider the potential of these 175 miRNAs for regulating the mature miRNA level in the entire cellular system. Interestingly, diverse studies provide evidence that deregulation of DICER and RNASEN is associated with different types of cancer (Grelier et al., 2009; Merritt et al., 2008; Kumar et al., 2009).

Additionally, the study of Li et al. (2012a) reports that the current ME model has several limitations. One of them is that three other members of ErbB family have not been considered while construction of this model. Therefore, the *in silico* analysis of signal cascade was performed under the condition of a single and isolated cell surface receptor. The study address the necessity that future work should focus on the effect delivered by combination of homo- and hereto-dimerization of these four ErbB family members. Furthermore, the study of Li et al. (2012a) also emphasized that all *in silico* experiments in this study were under the assumption that each gene defined in the model possesses a basal expression pattern. Therefore, the effect or influence of specific transcription factors have not been taken into consideration. Furthermore, the differences of specific phosphorylation sites within a certain protein kinase defined in the mode have not been taken into consideration, either. However, despite of these mentioned limitations of the ME model, it demonstrates that the potential of molecular modeling could become a powerful and significant tool for cancer research and it can help to step further towards the personalized medicine.

Recently, RNAi based methodologies have been widely used to silence a single target gene in order to investigate specific inhibition effect on pathological tissue (Moffat et al., 2006; Behlke, 2006). However, various studies applied microarray based expression screen technologies and reported that the single gene inhibition in pathological tissues could cause an imbalanced gene expression pattern involving many genes (Chung et al., 2002; Hoheisel, 2006) and therefore reduce the major-effect created by RNAi technologies (Li et al., 2012a). In contrast, the ability of miRNA being able to target multiple key components of signaling pathways, avoid to create imbalanced effect within

126

the entire cellular system. The study of Iguchi et al. (2010) explained the less possibility that miRNA-based therapeutic interventions can have severe adverse effect. However, some studies also indicate the passenger strand effect of miRNA-based therapeutic intervention to explain the resistance to miRNA regulation. Therefore, the conclusion drawn from the study of Li et al. (2012a) is that combined therapeutic strategies might be needed to succeed in the clinical application of anti-miRNA drug

4.2 THE NON-STEROIDAL ANTI-INFLAMMATORY DRUG (NSAID) MODEL

The NSAID model is the first molecule-based model, whose construction is based on literature references with regard to COX-pathway and its related pathways. The model integrates four cancer hallmarks to realize a biological organization principle for tumorigenesis. The integrated cancer hallmarks (sustained angiogenesis, tissue invasion, proliferation and evading apoptosis) symbolize the potential and ability of corresponding cancer developmental aspects. The data from different studies has been applied to validate functional indications of these cancer hallmarks within the model. The result of this validation experiment strongly indicates that the NSAID model can act with highly similar responses as individual cellular systems do when facing therapeutic interventions (siRNA interference and NS-398 drug). Therefore, the NSAID model possess potential to serve as "Virtual Patient" for prediction of clinical treatments (Li & Mansmann, 2013).

In addition, the novel therapeutic concept of synthetic lethality based on COX-pathway was analyzed by applying the NSAID model. The gene expression data of breast, colon and lung cancer cell lines were imported from the Cancer Genome Atlas to investigate the COX-2 based combination of anti-cancer drugs. The results from *in silico* experiments show that the combination inhibition (COX-2 and a receptor tyrosine kinase) could reach far better angiogenesis reduction than a single COX-2 inhibition can for colon- and lung-cancer cell lines, which acts in agreement with many independent *in vitro* studies (Mann et al., 2001; Tortora et al., 2003; Torrance et al., 2000; Tuccillo et al., 2005). Furthermore,

the study of Li & Mansmann (2013) suggests that combined inhibition therapy related to COX-pathway might present the unique potential for breast cancer treatment. However, this suggestion needs to be validated in *in vivo* or *in vitro* studies.

Unfortunately, different types of combination inhibition based on COX-pathway have not shown clear and strong effect with regard to other three cancer hallmarks. However, the possible molecular reasons could be found through the pathway analysis within the NSAID model, which presents another interesting application of this model for study the underlying molecular mechanism of cellular systems in a pathological state. In sum, the study for construction of the NSAID model successfully demonstrates that a molecule-based model involving gene expression, gene regulation, signal transduction, protein interaction and other cellular processes, is able to predict the individual cellular responses to different therapeutic interventions (in this case NS-398 and COX-2 specific siRNA inhibition) and this type of model can be applied for the purpose of therapeutic development including new potential drug target identification. Although the construction of the NSAID model presents the current biological knowledge and information related to COX-pathway, there are still some limitations of this model. For instance, concentrations of some metabolites including ADP, H_2O, Orthophosphate, O_2 and others, are fixed to avoid metabolic shortage during signal transduction process, which can cause the signal drop-down or signal termination (Li & Mansmann, 2013).

To the present day, the tumor-xenograft model represents the current standard procedure for preclinical testing of anti-cancer agents. Unfortunately, this type of model has too many limitations to remain an acceptable technology for clinical trials (Troiani et al., 2008). This study presents a systems-biological approach for molecular modeling and indicates that a 'predictive' molecule-based model could form the consolidated basis of 'virtual patient', being able to predict both effects and side effects of specific therapeutic intervention on individual patients.

4.3 THE HUMAN SIGNALING MODEL

Different neoplastic genome-types with various characteristics are expected to possess different traits and phenotypes. However, a recently suggested approach that identifies molecular commonalities based on perturbations in the pathway level could help to understand the common mechanisms that are responsible for transforming normal physiological cells to malignant ones (Lambrou et al., 2012). Furthermore, due to the low response rate of many drugs in cancer, modeling the biological effects of drug treatment on the tumor/cancer, will become an essential step towards personalized cancer treatment including a better control of effects produced by therapeutic interventions. This idea sparks increasing attentions from the research field of drug development and could lead to develop new, more global, therapeutic approaches (Hait, 2010; Kitano, 2007).

Although many studies have already been conducted to standardize the formalisms for creating signaling networks, a global human signaling network has not been reported yet (Gianchandani et al., 2006; Klamt et al., 2006; Papin et al., 2005). Many attempts have been made to construct such a signaling-network, however, all of them were mainly focused on a particular signaling pathway such as Li et al., (2009); Oda et al., (2005). However, an integrated signaling network covering multiple cancer relevant pathways with miRNA regulation has not been reported yet. One of main objectives of this thesis is to present the construction of a comprehensive human signaling model including signaling network, gene-regulatory network, and miRNA-regulation network (consisting of more than 50 signaling pathways and 1048 miRNAs in total), which plays a strategic role in the cellular functions and coordinates the phenotypic output by integrating various signals into an appropriate decision and then affecting the relevant cellular processes (Bouwmeester et al., 2004). The construction of a miRNA-regulation network was extended from the previous study (Li et al., 2012a). In addition to the model construction, I demonstrated four important system-biological modeling issues: (1) feedback control; (2) functional redundancy; (3) structural stability; (4) modularity.

129

I conducted a network analysis of the four different types of feedback loops within the implemented model to ensure control of the model system. For the functional redundancy, I demonstrated that the combination of drug effects on common downstream targets is much stronger that the single drug effect during the simulation, which has been explained and validated by different independent studies (Winter et al., 2010; Lee et al., 2010). This indicates not only the importance of functional redundancy within the implemented model, but also reveals that *in silico* experiment for the effect of therapeutic intervention towards personalized medicine could become one potential of this model application.

The 'bow-tie' structure of a biological network can provide robust yet flexible responses to various stimuli and inhibition of molecules involved owing to the redundancy of pathways in input and output modules (Kitano, 2007). I elucidated the bow-tie structure of the implemented model and classified model components into four groups: Fan-in, Core, Fan-out and GSC (Fig. 17A; Table 8). The functional modularity of the eleven hallmarks in the implemented model is intended to realize and improve the organizing principles of a logical framework constituted by Hanahan & Weinberg (2011), which can help us understand the high diversity of the tumorigenesis process (Fig. 17B). Nevertheless, there are "modular crosstalks" between the eleven hallmarks that have to be taken into consideration such as the pathway "Death Receptor" can invoke the "Apoptosis" process and suppress the ability of "Evading-apoptosis" in a cell. The pathway "Cell Cycle" delegates the cell cycle progression and at the same time provides a cell the possibility of limitless replicative potential.

In order to validate the basic principle of a signal transduction process in the implemented model, an *in silico* simulation experiment was performed to activate more than 90 extracellular ligands individually (Fig. 18A) and checked the states of the corresponding signaling pathways. As shown in Fig. 18A, each ligand can successfully activate its corresponding signaling pathways during the simulation, which validates the

basic function of signal transduction in the implemented model. Furthermore, among these ligands, seven ligands were selected to demonstrate that the miRNA regulation can have influence on the signal transduction process. The selected 7 ligands have major impact on the different types of cellular processes such as proliferation, apoptosis and are deregulated in many different cancer types. Therefore, the result displayed in Fig. 18B computationally reveals the possibility of miRNA regulation on different cellular processes.

In order to validate the predictive ability of this implemented human signaling model, the gene-expression data derived from the cell line published in the independent study of Nagaraj et al., (2011) was utilized to initialize the implemented model for in-silico simulation. The comparison between predicted proteome and experimental measured proteome reached a high correlation (77.6%), which indicates the high predictive ability of the implemented model (Fig. 21). Nevertheless, a correlation of less than 60% between the published gene-expression data and proteome data derived from the same study can be explained through the complexity of transcriptional and post-transcriptional regulation in the human cell. In addition, the same approach has been used to perform *in silico* simulation based on the mRNA and protein data in three functionally different human cell lines (U-2 OS, A-431 and U-251 MG; Lundberg et al., 2010) and reached high correlation between simulated proteome data and experimental proteome data (appendix B.2-B.7), which again reveals that the implemented human signaling model can predict the dynamic behavior of underlying cellular systems with high accuracy. These satisfactory results strongly indicate that the implemented model possesses a high quality of network structure including crosstalks among signaling pathways, feedback control loops and different levels of regulation (transcriptional and translational). Furthermore, the applied Monte Carlo-based approach for sampling kinetic parameters by considering the speed magnitude of biochemical reactions listed in the study of Papin et al., (2005),

helps the model system to reach the biological relevant steady state during *in-silico* simulation.

An enrichment-analysis was conducted to investigate the list of proteins of high correlation by applying the CPDB (Kamburov et al., 2011). Among these proteins, 17.4% (57 out of 328) belong to cancer-related pathways based on KEGG database, for example, extracellular ligand proteins such as TGFB1, EGF and transcription activators such as MYC. In order to demonstrate the therapeutic utilization of this human signaling model, I applied the gene-expression data from NCI-60 cancer cell lines and performed the drug response simulation to predict the drug response of these cancer cell lines regarding the drugs Imatinib and Temsirolimus. The Peasron's correlation between predicted sensitivity scores and experimental sensitivity scores for drug Imatinib is 86.18% and for drug Temsirolimus is 74.14%, which implies a potential application of this model in the field of anti-cancer drug development by predicting the dynamic behavior of cancerous cellular systems when facing drug treatments with high accuracy.

However, the fact that different correlations could not reach the ideal 100% can have many reasons. For instance, the functional difference was not considered between the different phosphorylated sites of kinases in phosphorylation reactions in the model. Nor was the functional differences considered between paralogs in the model such as MEK1 and 2. The study of Nagaraj et al., (2011) did not provide the genome mutation information. Moreover, enzyme-mediated interaction between metabolic pathways and signaling pathways has not been considered yet. Thus, many intensive crosstalks between different signaling- and metabolic pathways are absent during the simulation, for instance, regulation of signaling pathway AMPK on lipid metabolism, Fatty acid metabolism, Gluconeogenesis, GLUT4 biogenesis and others (Gaidhu et al., 2010; Hardie, 2004); the intensive crosstalk between activin signaling pathway and metabolic processes (Tsuchida et al., 2009).

Still, the modeling approach present in this thesis might be able to open up a path to a better understanding of the tumor dynamics, since by discovering common mechanisms of action, a deeper analysis of the cancer disease would be possible. One future goal of extend the current modeling work is to merge the HS model with the human metabolic model published by Duarte et al., (2007), in order to create a whole-cell model consisting of signaling network, metabolic network, gene-regulatory network and miRNA-regulation network.

The experience gained by diverse studies (Su et al., 2001; Golub et al., 1999; Chang et al., 2003) suggests that the major differences among tumor or cancer types could be reduced to cellular differences concerning molecules within these tumors or cancers and concerning the interactions among the molecules. In this thesis, I proposed a tumorigenesis calculation-formula with the incentive to better understand the causal connection between signaling pathways and tumorigenesis. Due to the fact that not all oncogenes and tumor-suppressor-genes contribute equally to cancer development (Izquierdo et al., 2005), the integrated cancer and process hallmarks in the model should be able to reflect the deregulation-degree of molecular interactions according to the given genetic information of an individual cell line (or patient). The tumorigenesis calculation-formula can help to summarize the signal from the cancer hallmarks into the final readout component "tumorigenesis", which might be redeemed as an informative readout component of this model system. I utilized this implemented human signaling model with integrated hallmarks to predict the drug response of NCI-60 cancer cell lines and reached a considerably high correlation rate, which implies the high quality of this model. For these drug response predictions, I only used process hallmark 'proliferation' as a readout component due to the reason that the experimentally established sensitivity score was based only on the growth potential of cancer cell lines (Holbeck et al., 2010). A future study should put emphasis on how to apply and improve this tumorigenesis calculation formula within this model for personalized medicine.

To date, different studies have indicated that the aetiology, progression and prognosis of tumors can be linked to the miRNA regulation (Cui et al., 2006; Brid et al., 2010). Moreover, several studies have tried to utilize the miRNA expression profiling to identify and classify several types of cancers (Gaur et al., 2007; Tong and Nemunaiti, 2008). However, little is known about how the miRNA expression data can be used to identify the informative biomarker regarding an individual cancer cell line (or patient). Many tumor biomarkers identified through the use of proteomic techniques in the past years, have failed to attain broad application in clinic. Therefore, I applied miRNA-expression data from three different cancer cell lines reported in the Gaur's study, took the advantage of miRNA-regulation within the signaling pathways and demonstrated how to identify and discover the biomarkers of an individual cancer cell lines by simulating this implemented model with miRNA-expression data as input. Biomarkers that are identified in this thesis, need to be tested in follow-up verifications and clinical validations. I listed only the top 10 up- and down-regulated proteins after the simulation and this top-range might not be ideal. However, a limitation of my current approach could be that the applied miRNA-expression data (from Gaur's study) contains only information about around 120 miRNAs. Taking into consideration that the miRNA-regulation network in the implemented model possesses 1048 miRNA, only a partial miRNA-regulation network has been activated during the simulation. Thus, I assume that this limitation could lead to false positively identified biomarkers. Furthermore, I think that this approach could even work better with the combination of gene-expression data. In the past years, many cancer biomarkers identified through the use of proteomic techniques have failed to attain broad application in clinic. Therefore, the current approach might shed light on the development of an alternative strategy for overcoming the limitation of proteomic techniques.

Modeling approaches are integrated in the concept of personalized medicine as they are able to generate detailed predictions for the treatment of complex diseases like cancer and

could be integrated in routine diagnostics in oncology (Daskalaki et al., 2009). This *in silico* analysis study applied a systems biology approach, integrated experimental data into a mathematical model, and intended to demonstrate a relevant way to establish an *in silico* biomarker-identification protocol related to a patient based on his or her molecular signature. One focus in my future research is to extend the approach to drug discovery and the optimization of medical treatment regimes for individual patients.

4.4 THE HUMAN METABOLIC MODEL (HMMA)

The model HMMA is the first large-scale human metabolic network with miRNA-regulation and the result of adding global miRNA regulation network into the published human metabolic model (Duarte et al., 2007). The HMMA could be applied to computational approaches in order to understand the functionality of miRNAs in the different metabolic processes better. The HMMA is created in a way that all components are labeled with public nomenclature (e.g. a gene object entity is labeled with Ensembl-ID, a protein object entity with UniProt-ID, a metabolic object entity is labeled with ChEBI-ID and others). Thus, the approach for inheriting genetic information of an individual cell line/xenograft (described in Fig. 20 in section 2.11), can be easily applied to the HMMA. A future study should be firstly focused on how to design and perform the validation experiment to check the predictive ability of the HMMA. For this purpose, one can utilize the published results of diverse studies, which investigated the miRNAs' role in metabolic processes including lipid metabolism (Fernández-Hernando et al., 2011), glucose metabolism (Bravo-Egana et al., 2008; Correa-Medina et al., 2009), energy metabolism (Esau et al., 2004) and compare these data results with HMMA simulation data to validate its predictive ability.

Before the concept of synthetic lethality and synthetic sickness was introduced in the study of Kaelin, (2005), a lot of effort had already been made to invent or identify an anti-cancer drug, which could kill rapidly dividing cancer cells and at the same time

tolerate normally dividing host cells. Unfortunately, most (>90%) anticancer drugs such as doxorubicin, bleomycin and cytarabine can successfully inhibit or kill cancer cells, however, at the same time, they are toxic to heart, lung and cerebellum, respectively (Kaelin, 2005). The key idea of this concept of synthetic lethality and synthetic sickness is to identify or detect gene-gene interactions, which can serve as effective drug targets to minimize the 'off-target' effect and maximize the 'on-target' effect. The special FBA approach applied in the study of Folger et al., (2011), could achieve a certain degree of the synthetic-lethality identification process. However, its result could be further improved by integrating the post-transcriptional regulaton, in this case, global miRNA regulation network. Therefore, future studies should put the emphasis on extending the limitation of the current FBA approach to incorporate the miRNA regulation information, so that the synthetic lethality identification process would be able to achieve a better result.

Because the HMMA includes the detailed model information about metabolic processes and miRNA-regulation, I surmised this model might possess the huge potential for personalized medicine related to many metabolic syndromes and metabolic disorders including obesity and diabetes. But it lacks information about gene-regulatory mechanism and the crosstalk effect between putative signaling pathways and metabolic pathways (mentioned in section 4.3). It also lacks the translational regulation information including methylation, oxidization, ubiquitination, etc. Therefore, future works could be focused on the merging process of the human signaling model and the human metabolic model (HMMA) from this study. Both models are manually created based on literature references, possess the same consistent model annotations (explained in sections 2.9 and 2.14) and saved in the PyBioS databases (Fig. 24), which records the biological data with the emphasis on the uniqueness and avoids recording any redundant information. Those reasons could certainly contribute to this merging process. The merged model will serve as a human generic model, which can be used to generate a tissue-specific model based

on diverse biological data including metabolomics, transcriptomics, proteomics, phenotypical data and others. The study conducted by Jerby et al., (2010) introduces a model-building algorithm (MBA) to cope with such kind of tissue-specific model generation and the same or similar approach might also be applicable for the human generic model.

4.5 PYBIOS 2

During this thesis study, I have developed PyBioS2, a web-based software platform for the design, modeling, and simulation of cellular systems. One of PyBioS2's novel features is that this software is connected to many public databases and possesses several data resources, which provides a consolidate precondition for constructing large-scale networks. The design of the internal PyBioS2 database ensures the biological data integration and avoids the redundant biological information, which contributes to conquering one important challenge of systems biology, namely the utilization of biological data from diverse levels, and enables users to apply the PyBioS2 for the model construction in a parallel way. One flaw about PyBioS2 is that this software is currently only dealing with human and mouse data. In the future, it should be extended to other species such as zebrafish, chicken, chimpanzee.

The PyBioS2 functions as a model repository and records not only information about the model components/reactions, but also diversely applied biochemical reaction kinetics. It is important and crucial for users to maximize the reutilization of their models. The special 3-level model architecture allows users to reuse their models more easily or create large scale networks, for instance, by merging different sub-models. The global construction of the genome-scale human signaling model involving forming and merging multiple pathways (sub-models) that connect signaling inputs to signaling outputs, was a good application of PyBioS2. Indeed, all models presented in this thesis are constructed

within PyBioS2 and saved in the PyBioS2 internal database. All of them can be re-used, merged and divided for any other research purposes.

Both ODE and Petri net simulation strategies implemented in PyBioS2 enable users to cope with specific simulation conditions and specific reaction systems, which empowers the predictive ability of PyBioS2 models. To the present day, PyBioS2 is the only software that can provide both of the simulation strategies at the same time. However, both simulation engines can be improved further. For instance, the ODE simulation engine can be facilitated with a hybrid algorithm to partition a model reaction system into two parts: a stochastic one and a deterministic one. The integrated hybrid algorithm can help to accelerate the entire simulation process, because only the stochastic part will be applied to the time-consuming stochastic methods (Hoops et al., 2006). The Petri net simulation engine can be facilitated with colored Petri net (CPN). The CPN provides a well-defined semantics, a hierarchical description, the synchronization with the description of data manipulation and the stability towards minor changes of a model system. For instance, it is possible in CPN to add a set of new biological processes without changing the current model structure of the existing biological system (Jensen et al., 2007).

The unique option of the semi-automatic reaction definition of PyBioS2 provides users with a comfortable way of designing/constructing networks, especially large-scale ones. In the near future, this semi-automatic option should be improved into a fully-automatic one, so that a user can easily obtain a tissue-specific model by uploading the corresponding tissue-specific data. PyBioS2 should be able to construct a comprehensive generic model based on models saved in the repository. The biological data of different levels (including gene-expression data, miRNA-expression data, and proteomic data) can be used in PyBioS2 to predict the dynamic behavior of the biological system and to demonstrate the therapeutic benefit of systems biology approaches. Because this data contains such rich information, it can be utilized for many different purposes such as

138

pathological state classification and dynamic analysis of underlying cellular systems (Golub et al., 1999; Chang et al., 2003). In the future, PyBioS2 should be improved to be able to absorb the phenotypical data and environmental factors for analyzing the dynamic behavior of cellular systems.

In conclusion, PyBioS2 is powerful yet user-friendly software for constructing large-scale models serving as references for interpreting experimental results and meaningful guidance for understanding and predicting cellular systems.

4.6 CONCLUSIONS AND FUTURE OUTLOOK

In this thesis, different computational molecular models have been introduced and their applications have been demonstrated with regard to molecular targeted therapeutic interventions. The ME model is the first molecular model including detailed information about miRNA regulation. Its *in-silico* application has successfully demonstrated an important therapeutic miRNA concept, "One Hit Multiple Targets" (Wurdinger & Costa, 2007). The *in-silico* simulation of the inhibition effect of 100 distinct miRNA inhibitors has revealed the predictive ability of this ME model (Li et al.,, 2012a).

The NSAID model is first molecular model explaining the biological function of COX-pathway, which is the key pathway related to the NSAID. I applied different gene-expression data from the Cancer Genome Atlas to analyze their corresponding drug responses. Through well designed *in silico* experiments, I demonstrated that the NSAID model could reflect and predict the dynamic behavior of different cancer cell lines when facing same therapeutic interventions. Those results were validated by results of different independent studies. Furthermore, the NSAID model has also been applied to explain the therapeutic benefit of the concept "Synthetic Lethality", where different types of combination inhibitions have been investigated. The results are also in agreement with results of different studies for conducting research on effect of combined anti-cancer

treatments. The center idea of those applied in silico experiments in my thesis is to transfer the dynamic properties hidden inside of individual genetic information (including gene-expression and miRNA expression) into the model system so that a personalized preclinical services including identification of optimal drug combination and optimal calculation of drug dose can be flawlessly performed. The conclusion of the study of Li & Mansmann (2013) explains that a molecule-base model involving significant information related to biological system is able to make a great contribution for personalized medicine.

In the course of researching for this thesis, the new version of PyBioS has been successfully developed, which is facilitated with an internal modeling database (Li et al.,, 2012b). With the help of the special designed model structure in PyBioS2 and the internal modeling database, the utilization and re-usefulness of each model defined in PyBioS2 has reached the maximum, which also overcomes one current limitation (model re-usefulness regarding merge and split) in the field of systems biology research. In this way, I merged the ME model and the NSAID model into a genome-scale signaling model. This genome-scale signaling model has been validated in different ways to investigate its quality with satisfactory result. It is the first time that a signaling model has been created in such a big scale and its applications for personalized medicine has been demonstrated with the prediction of different drug responses based on individual cancer cell lines whose cancer types include breast, lung, colon, CNS, leukemia, melanoma, renal and prostate (Holbeck et al.,, 2010). Although the prediction result could be regarded as satisfactory, there are many improvements that should be carried out on this human signaling model, as stated in the discussion part 4.3 of this thesis. The future goal is to improve this model with a prediction rate higher than 80%.

With the successful development of the systemsbiological software PyBioS2, an *in silico* simulation pipeline and the human signaling model (considered as a 'virtual patient' model), the primary objective of this thesis has been fulfilled.

140

Biomarker is defined as molecular, cellular or functional measurable parameters indicating a particular genetic, epigenetic or functional status of a biological system (Ludwig & Weinstein, 2005). In the case of a pathological state of a biological system, such as cancer, biomarkers can be used for diagnosis, staging, prognosis and treatment selection. Unfortunately, most biomarkers discovered to the present fail to attain a broad application for clinical use and all of them are discovered based solely on statistical approaches. I will name those biomarkers discovered by statistical approaches "first generation biomarker". The biggest disadvantage of the first generation biomarkers is that they are not capable of precluding inclusion of associated statistical bystanders that will have no predictive power in another patient groups. This means that first generation biomarkers have the clear limitation and might even lose their predictive and prognostic function when they are applied for a new patient group.

During research of this thesis, I developed an *in silico* method for biomarker discovery at cancer specific level by applying the miRNA-expression data derived from individual tumor patients. This method can be extended further to reach the individual level, and the type of biomarker created for individual patients (I will name those "second generation biomarker") should possess high indicative and predictive value regarding its corresponding individual tumor/cancer patient. In follow-up studies, *in-vivo* and *in-vitro* validation processes should be designed and performed to confirm the predictive precision of this *in silico* method.

In the past few years, two genome-scale human metabolic models have been published and their applications on the drug development have achieved outstanding successes (Duarte et al.,, 2007; Ma et al.,, 2007). After carefully studying these two models, I came to realize the common weakness of them, which is the lack of transcriptional and translation regulation within both of model networks. Therefore, substantial effort has been applied to merging these two genome-scale metabolic models in PyBioS2 and extending the merge model (HMMA) with one type of transcriptional regulation,

"miRNA regulation". In the discussion part 4.4, different validation and application concepts have been made for this HMMA model. In follow-up studies, these concepts will be further improved and applied for this model.

Personalized medicine is currently still in its infancy. In the near future a lot of more effort and work need to be undertaken in order to improve the public health issue in that direction. In this thesis, different models and *in silico* approaches regarding drug development and therapeutic intervention at individual level have been constructed and developed, which might shed some light on the potential usefulness of systems biology and computational biology towards personalized medicine.

Reference

Adams,J.M. (2003) Ways of dying: multiple pathways to apoptosis. Genes Dev. 17, 2481-2495.

Aoki Y., Niihori T., Kawame H., Kurosawa K., Ohashi H., Tanaka Y., Filocamo M., Kato K., Suzuki Y., Kure S. & Matsubara Y. (2005) Germline mutations in HRAS proto-oncogene cause Costello syndrome. Nature Genet. 37, 1038-1040.

Androulidaki A., Iliopoulos D., Arranz A., Doxaki C., Schworer S., Zacharioudaki V., Margioris A.N., Tsichlis PN. & Tsatsanis C. (2009) The kinase Akt1 controls macrophage response to lipopolysaccharide by regulating microRNAs. Immunity. 31(2):220-31.

Asslaber D., Piňón JD., Seyfried I., Desch P., Stőcher M., Tinhofer I., Egle A., Merkel O. & Greil R. (2010) microRNA-34a expression correlates with MDM2 SNP309 polymorphism and treatment-free survival in chronic lymphocytic leukemia. Blood. 115(21):4191-7.

Avraham R., Sas-Chen A., Manor O., Steinfeld I., Shalgi R., Tarcic G., Bossel N., Zeisel A., Amit I., Zwang Y., Enerly E., Russnes H., Biagioni F., Mottolese M., Strano S., Blandino G., Borresen-Dale A., Pilpel Y., Yakhini Z., Segal E. & Yarden Y. (2010) EGF decreases the abundance of microRNAs that restrain oncogenic transcription factors. Science Signaling. 3(124).

Ball MP, Thakuria JV, Zaranek AW, Clegg T, Rosenbaum AM, Wu X, Angrist M, Bhak J, Bobe J, Callow MJ, Cano C, Chou MF, Chung WK, Douglas SM, Estep PW, Gore A, Hulick P, Labarga A, Lee JH, Lunshof JE, Kim BC, Kim JI, Li Z, Murray MF, Nilsen GB, Peters BA, Raman AM, Rienhoff HY, Robasky K, Wheeler MT, Vandewege W, Vorhaus DB, Yang JL, Yang L, Aach J, Ashley EA, Drmanac R, Kim SJ, Li JB, Peshkin L, Seidman CE, Seo JS, Zhang K, Rehm H.L. & Church G.M. (2012) A public resources facilitating clinical use of genomes. Proc Natl Acad Sci U S A. doi/10.1073/pnas.1201904109.

Bartel D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116:281-297.

Bartel D.P. (2009) MicroRNAs: Target recognition and regulatory functions. Cell 136(2): 215-233.

Barh D., Bhat D. & Viero C. (2010) miReg: a resource of microRNA regulation. J Integr Bioinform. 7(1): 144.

Behlke M.A. (2006) Progress towards in vivo use of siRNAs. Mol Ther 13: 644-670.

Bernard R.W., Wang W.X. & Nelson P. (2007) Energizing miRNA research: a review of the role of miRNAs in lipid metabolism, with a prediction that miR-103/107 regulates human metabolic pathways. Mol Genet Metab, 91(3), 209-217.

Bernardinello L. & de Cindio F. (1992) A Survey of Basic Net Models and Modular Net Classes. Springer Verlag, Berlin, Germany.

Berridge M.J. (2012) Cell Cycle and Proliferation. Cell Signaling Biology. doi:10.1042/csb0001009.

Bhalla S. & Iyengar R. (1999) Emergent properties of networks of biological signaling pathways. *Science* **283**, 381–387.

Blais A. & Dynlacht BD. (2005) Constructing transcriptional regulatory networks. Genes & Dev., 19:1499-1511.

Boulton S.J. (2006) Cellular functions of the BRCA tumour-suppressor proteins. Biochem Soc Trans. 34, 633-45.

Bouwmeester T., Bauch A., Ruffner H., Angrand P.O., Bergamini G., Croughton K., Cruciat C., Eberhard D., Gagneur J., Ghidelli S., Hopf C., Huhse B., Mangano R., Michon A.M., Schirle M., Schlegl J., Schwab M., Stein M.A., Bauer A., Casari G., Drewes G., Gavin A.C., Jackson D.B., Joberty G., Neubauer G., Rick J., Kuster B. & Superti-Furga G. (2004) A physical and functional map of the human TNF-α/NF-κB signal transduction pathway. *Nature Cell Biol.* **6**, 97–105.

Braun CJ., Zhang X., Savelyeva I., Wolff S., Moll UM., Schepeler T., Ōrntoft TF., Andersen CL. & Dobbelstein M. (2008) p53-Responsive micrornas 192 and 215 are capable of inducing cell cycle arrest. Cancer Res. 68(24):10094-104.

Bravo-Egana V., Rosero S., Molano R.D., Pileggi A., Ricordi C., Domínguez-Bendala J. & Pastori R.L. (2008) Quantitative differential expression analysis reveals miR-7 as major islet microRNA. Biochem Biophys Res Commun. 22, 366(4), 922-6.

Brid R., Robies A. & Harris C. (2010) Genetic variation in microRNA networks: the implications for cancer research. Nature Review Cancer, 10, 389-402.

Bueno M.J., de Castro I.P. & Malumbres M. (2008) Control of cell proliferation pathways by microRNAs. Cell Cycle 7:3143-3148.

Blume-Jensen, P. & Hunter, T. (2001) Oncogenic kinase signalling. Nature 411:355–365.

Calin GA. & Croce CM. (2006) MicroRNA signatures in human cancers. Nat Rev Cancer, 6, 857-66.

Carter C., Prinz S., Neou C., Shelby P., Marzolf B., Thorsson V. & Galitski T. (2007) Prediction of phenotype and gene expression for combinations of mutations. Mol. Syst. Biol. 3, 96.

Chan I.S. & Ginsburg G.S. (2011) Personalized medicine: progress and promise. Annu Rev Genomics Hum Genet, 12, 217-44.

Chan J.A., Krichevsky A.M. & Kosik K.S. (2005) MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. Cancer Res, 65, 6029–6033.

Chang J.C., Wooten E.C., Tsimelzon A., Hilsenbeck S.G., Gutierrez M.C., Elledge R., Mohsin S., Osborne C.K., Chamness G.C., Allred D.C. & O'Connell P. (2003) Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet.* **362**(9381), 362-9.

Chatziioannou A., Palaiologos G. & Kolisis F.N. (2003) Metabolic flux analysis as a tool for the elucidation of the metabolism of neurotransmitter glutamate. Metab Eng 5: 201-210.

Chen C.Z. (2005) MicroRNAs as oncogenes and tumor suppressors. N. Engl. J. Med. 353:1768-1771.

Chen Z., Gibson T. B., Robinson F., Silvestro L., Pearson G., Xu, B., Wright A., Vanderbilt C. & Cobb M. H. (2001) MAP kinases. Chem. Rev. 101, 2449 –2476.

Chen X., Yeung T.K. & Wang Z. (2000) Enhanced drug resistance in cells coexpressing ErbB2 with EGF receptor or ErbB3. Biochem Biophys Res 277:757–763.

Chen X., Xu H., Yuan P., Fang F., Huss M., Vega V.B., Wong E., Orlov Y.L., Zhang W., Jiang J., Loh Y.H., Yeo H.C., Yeo Z.X., Narang V., Govindarajan K.R., Leong B., Shahab A., Ruan Y., Bourque G., Sung W.K., Clarke N.D., Wei C.L. & Ng H.H. (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 133, 1106–1117.

Chiocca E.A. & Lawler S. E. (2010): The Many Functions of MicroRNAs in Glioblastoma. World Neurosurgery. 73(6): 598-60.

Cho, W.C (2007) OncomiRs: The discovery and progress of microRNAs in cancers. Mol. Cancer 6:60.

Choura M. & Rebai A. (2010) Application of computational approaches to study signaling networks of nuclear and Tyrosine kinase receptors. Biology Direct, 5, 58.

Christoffersen NR., Shalgi R., Frankel LB., Leucci E., Lees M., Klausen M., Pilpel Y., Nielsen FC., Oren M., & Lund AH. (2010) p53-independent upregulation of miR-34a during oncogene-induced senescence represses MYC. Cell Death Differ. 17(2):236-45.

Chung C.H., Bernard P.S. & Perou CM. (2002) Molecular portraits and the family tree of cancer. Nat Genet. 32: 16-26.

Clague J., Lippman S.M., Yang H., Hildebrandt M.A., Ye Y., Lee JJ. & Wu X. (2010) Genetic variation in MicroRNA genes and risk of oral premalignant lesions. Mol Carcinog. 49(2):183-9.

Clarke A., Purdie C., Harrison D., Morris R., Bird C., Hooper M. & Wyllie A. (1993) Thymocyte apoptosis induced by p53-dependent and independent pathways. Nature, 362, 849-52.

Cohen P. (2002) Protein kinases — the major drug targets of the twenty-first century? Nature Rev. Drug Discov. 1, 309–315.

Cohen P. (2006) The twentieth century struggle to decipher insulin signalling. Nature Rev. Mol. Cell Biol. 7, 867–873.

Correa-Medina M., Bravo-Egana V., Rosero S., Ricordi C., Edlund H., Diez J. & Pastori R.L. (2009) MicroRNA miR-7 is preferentially expressed in endocrine cells of the developing and adult human pancreas. Gene Expr Patterns. 9(4), 193-9.

Csardi G. & Nepusz T. (2006) The igraph software package for complex network research, InterJournal, Complex Systems, P p. 1695.

Csete M. & Doyle J. (2004) Bow ties, metabolism and disease. *Trends Biotechnol.* **22**, 446-450.

Cui Q.H., Yu Z.B., Purisima E. & Wang E. (2006) Principles of microRNA regulation of a human cellular signaling network. Mol Syst Biol. 2:(46).

Cummins JM., He Y., Leary RJ., Pagliarini R., Diaz LA. Jr., Sjoblom T., Barad O., Bentwich Z., Szafranska AE., Labourier E., Raymond CK., Roberts BS., Juhl H., Kinzler KW., Vogelstein B. & Velculescu VE. (2006) The colorectal microRNAome. Proc Natl Acad Sci U S A. 103(10):3687-92.

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

Dannenberg AJ. & Zakim D. (1999) Chemoprevention of colorectal cancer through inhibition of cyclooxygenase-2. Sem Oncol. 26, 499-504.

Daskalaki A. & Lazakidou A. (2012) Preface. Quality Assurance in Healthcare Service Delivery, Nursing and Personalized Medicine: Technologies and Processes. IGI Global.

Davis BN., Hilyard AC., Lagna G. & Hata A. (2008). SMAD proteins control DROSHA-mediated microRNA maturation. Nature, 454(7200): 56-61.

DeBerardinis, R.J., Sayed, N., Ditsworth, D. & Thompson, C.B. (2008). Brick by brick: Metabolism and tumor cell growth. Curr. Opin. Genet. Dev. 2008, 18, 54-61.

Denkert, C., Budczies, J., Weichert, W., Wohlgemuth, G., Scholz, M., Kind, T., Niesporek, S., Noske, A., Buckendahl, A., Dietel M. & Fiehn O. (2008) Metabolite profiling of human colon carcinoma – deregulation of TCA cycle and amino acid turnover. Molecular Cancer, 7:72.

Duarte N., Becker S., Jamshidi N., Thiele I., Mo M., Vo T., Srivas R. & Palsson B. (2007) Global reconstruction of the human metabolic network based on genomic and bibliomic data. PNAS, 104, 1777-1782.

Diller L., Kassel J., Nelson C.E., Gryka M.A., Litwak G., Gebhardt M., Bressac B., Ozturk M., Baker S.J. & Vogelstein B. (1991) p53 functions as a cell cycle control protein in osteosarcomas. Mol Cell Biol. 10, 5772-81.

Edwards J.S. & Palsson B.O. (2000) The Escherichia coli MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. Proc Natl Acad Sci, 97, 5528-5533.

Esau C., Kang X., Peralta E., Hanson E., Marcusson E.G., Ravichandran L.V., Sun Y., Koo S., Perera R.J., Jain R., Dean N.M., Freier S.M., Bennett C.F., Lollo B. & Griffey R. (2004) MicroRNA-143 regulates adipocyte differentiation. J Biol Chem, 279(50):52361-5.

Esquela-Kerscher A. & Slack F.J.(2006) Oncomirs-microRNAs with a role in cancer. Nat. Rev. Cancer 6: 259-269.

Fayard E., Tintignac L.A., Baudry A. & Hemmings B.A. (2005) Protein kinase B/AKT at a glance. J Cell Sci., 118:5675-5678.

Ferrario A., Fisher AM., Rucker N. & Gomer CJ. (2005) Celecoxib and NS-398 enhance photodynamic therapy by increasing in vitro apoptosis and decreasing in vivo inflammatory and angiogenic factors. Cancer Res, 65(20), 9473-8.

Fernández-Hernando C., Suárez Y., Rayner K.J. & Moore K.J. (2011) MicroRNAs in lipid metabolism. Curr Opin Lipidol. 22(2), 86-92.

Filipowicz W., Bhattacharyya S.N. & Sonenberg N. (2008) Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight? Nature Reviews. Genetics 9(2):102-114.

Folger O., Jerby L., Frezza C., Gottlieb E., Ruppin E. & Shlomi T. (2011) Predicting selective drug targets in cancer through metabolic networks. Molecular Systems Biology, 7:501.

Fosslien E. (2000) Molecular pathology of cyclooxygenase-2 in neoplasia. Ann Clin Lab Sci. 30, 3-21.

Forbes S.A., Bhamra G., Bamford S., Dawson E., Kok C., Clements J., Menzies A., Teague J.W., Futreal P.A. & Stratton M.R. (2008) The Catalogue of Somatic Mutations in Cancer (COSMIC). Curr Protoc Hum Genet., 10, 11.

Funahashi, A., Morohashi, M., Kitano, H. & Tanimura, N. (2003). CellDesigner: a process diagram editor for gene-regulatory and biochemcial networks. *BIOSILICO,* I, 5.

Gaidhu M., Anthony N., Patel P., Hawke T., & Ceddia R. (2010) Dysregulation of lipolysis and lipid metabolism in visceral and subcutaneous adipocytes by high-fat diet: role of ATGL, HSL, and AMPK. Am J Physiol Cell Physiol, 298, C961–C971.

Gan B., Lim C., Chu G., Hua S., Ding Z., Collins M., Hu J., Jiang S., Fletcher-Sananikone E., Zhuang L., Chang M., Zheng H., Wang Y.A., Kwiatkowski D.J., Kaelin W.G., Signoretti S. & DePinho R.A. (2010) FoxOs enforce a progression checkpoint to constrain mTORC1-activated renal tumorigenesis. Cancer Cell, 18(5) 472:484.

Gaur A., Jewell D., Liang Y., Ridzon D., Moore J., Chen C., Ambros V. & Israel M. (2007) Characterization of MicroRNA Expression Levels and Their Biological Correlates in Human Cancer Cell Lines. Cancer Research, 67, 2456.

Georges SA., Biery MC., Kim SY., Schelter JM., Guo J., Chang AN., Jackson AL., Carleton MO., Linsley PS., Cleary MA. & Chau BN. (2008) Coordinated regulation of cell cycle transcripts by p53-Inducible microRNAs, miR-192 and miR-215. Cancer Res. 68(24):10105-12.

Gianchandani E. P., Papin J.A., Price N.D., Joyce A.R. & Palsson B.Ø. (2006) Matrix formalism to describe functional states of transcriptional regulatory systems. PLoS Comput. Biol. 2, e101.

Golub T.R., Slonim D.K., Tamayo P., Huard C., Gaasenbeek M., Mesirov J.P., Coller H., Loh M.L., Downing J.R., Caligiuri M.A., Bloomfield C.D. & Lander E.S. (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, **286**, 531-537.

Gottesman MM. (2002) Mechanisms of Cancer Drug Resistance. *Annual Review of Medicine*, **53**, 615-627.

Green D.R. (1998) Apoptotic pathways: the roads to ruin. Cell, 94, 695-698.

Grelier G., Voirin N., Ay A.S., Cox D.G., Chabaud S. & Treilleux I. (2009) Prognostic value of Dicer expression in human breast cancers and association with the mesenchymal phenotype. British Journal of Cancer 101(4):673-683.

Griffiths-Jones S., Saini H.K., van Dongen S. & Enright A.J. (2008) miRBase: tools for microRNA genomics. NAR 34(Database Issue):D140-D144.

Gross, M.E., Shazer, R.L. & Agus, D.B. (2004) Targeting the HER-kinase axis in cancer. Semin Oncol 31:9–20.

Gurdon J.B. (2006) From nuclear transfer to nuclear reprogramming: the reversal of cell differentiation. Annu. Rev. Cell Dev. Biol. 22:1-22.

Hait W. (2010) Anticancer drug development: the grand challenges. Nature Reviews Drug Discovery 9, 253-254.

Hanahan D. & Folkman J. (1996) Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. Cell, 86, 353-364.

Hanahan, D. & Weinberg, R.A. (2000) The hallmarks of cancer. Cell, 100, 57-70.

Hanahan D., & Weinberg R.A. (2011) Hallmarks of Cancer: The Next Generation. Cell 144(5): 646-74.

Hardie G. (2004) The AMP-activated protein kinase pathway – new players upstream and downstream. Journal of Cell Science, 117, 5479-5487.

Hardy S. & Robillard P. (2004) Modeling and Simulation of Molecular Biology Systems using Petri Nets: Modeling Goals of Various Approaches. J Bioinform Comput Biol. 2004, 2-4.

Hardy S. & Robillard P. (2007) Petri net-based method for the analysis of the dynamics of signal propagation in signaling pathways. Bioinformatics, 24(2): 209-217.

Harley C. (2002) Telomerase is not an oncogene. Oncogene, 21, 494–502.

Hatley M.E., Patrick D.M., Garcia M.R., Richardson J.A., Bassel-Duby R., Rooij E.V. & Olson E.N. (2010) Modulation of K-Ras-Dependent Lung Tumorigenesis by MicroRNA-21. Cancer Cell 18:282-293.

Heiner M., Gilbert D. & Donaldson R. (2008) Petri Nets for Systems and Synthetic Biology in Proceedings of the Formal Methods for the Design of Computer, Communication, and Software Systems 8th International Conference on Formal Methods for computational Systems Biology. pp. 215-264, Berlin, Heidelberg: Springer Verlag.

Hemmati P.G., Müer A., Gillissen B., Overkamp T., Milojkovic A., Wendt J., Dörken B. & Daniel P.T. (2010) Systematic genetic dissection of p14ARF-mediated mitochondrial cell death signaling reveals a key role for p21CDKN1 and the BH3-only protein Puma/bbc3. J Mol Med (Berl), 88(6), 609-22.

Heriche J.K. Lebrin F., Rabilloud T., Leroy D., Chambaz E.M. & Goldberg, Y. (1997) Regulation of protein phosphatase 2A by direct interaction with casein kinase 2alpha. Science, 276, 952–955.

Hindmarsh A.C. (1983) ODEPACK, A systematized collection of ODE solvers, in Scientific Computing, Stepleman R.S. Et al. (eds), North-Holland, Amsterdam.

Hiraga T., Myoui A., Choi ME., Yoshikawa H. & Yoneda T. (2006) Stimulation of cyclooxygenase-2 expression by bone-derived transforming growth factor-beta enhances bone metastases in breast cancer. Cancer Res. 66, 2067-2073.

147

Hoheisel J.D. (2006) Microarray technology: beyond transcript profiling and genotype analysis. Nat. Rev. Genet 7:200-210.

Holbeck S., Collins J. & Doroshow J. (2010) Analysis of Food and Drug Administration-Approved Anticancer Agents in the NCI60 Panel of Human Tumor Cell Lines. Mol Cancer Ther. 9, 1451.

Hoops S., Sahle S., Gauges R., Lee C., Pahle J., Simus N., Singhal M., Xu L., Mendes P. & Kummer U. (2006). COPASI—a COmplex PAthway SImulator. *Bioinformatics*. 22(24), 3067-74.

Hornstein E. & Shomron N. (2006) Canalization of development by microRNAs. Nature Genetics, 38, S20-S24.

Huang K., Zhang J.X., Han L., You Y.P., Jiang T., Pu P.Y. & Kang C.S. (2010) MicroRNA roles in beta-catenin pathway. Molecular Cancer, 9, 252.

Hucka et al.,, (2003) The systems biology makeup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics, 19(4), 524-531.

Husain SS. Szabo IL. Pai R. Soreghan B. Jones MK. & Tarnawski AS. (2001) MAPK (ERK2) kinase--a key target for NSAIDs-induced inhibition of gastric cancer cell proliferation and growth. *Life Sci.* 69, 3045-54.

Hyduke R., Amundson A. & Fornace J.R. (2009) Handbook of Cell Signaling. (Eds bradshow, R. A. & Dennis, E. A.), 3, 2107-2125.

Hwang H.W. & Mendell J,T. (2006) MicroRNAs in cell proliferation, cell death, and tumorigenesis. Br J Cancer 94:776-780.

Iguchi H., Kosaka N. & Ochiya T. (2010) Versatile Applications of microRNA in Anti-Cancer Drug Discovery: From Therapeutics to Biomarkers. Current Drug Discovery Technologies, 7, 95-105.

Iniguez MA., Rodríguez A., Volpert OV., Fresno M. & Redondo JM. (2003) Cyclooxygenase-2: a therapeutic target in angiogenesis. Trends Mol Med. 9,73-8.

Inui M., Martello G. & Piccolo S. (2010) MicroRNA control of signal transduction. Nature Reviews. Molecular Cell Biology 11(4): 252-263.

Iravani M., Dhat R. & Price C. (1997) Methylation of the multi tumor suppressor gene-2 (MTS2, CDKN1, p15INK4B) in childhood acute lymphoblastic leukemia. Oncogene, 15, 2609-2614.

Israel A., Sharan R., Rupping E. & Galun E. (2009) Increased microRNA activity in human cancer. PLoS ONE 4(6): e6045.

Izquierdo M. (2005) Short interfering RNAs as a tool for cancer gene therapy. Cancer Gene Therapy 12, 217-227.

Jana NR. (2008) NSAIDs and apoptosis. Cellular and Molecular Life Sciences. 65(9), 1295-1301.

Jayaprakash V., Rigual R., Moysich K., Loree T., Nasca M., Menezes R. & Reid M. (2006) Chemoprevention of head and neck cancer with aspirin: A case-control study. Arch Otolaryngol Head Neck Surg. 132(11), 1231-1236.

Jeong H., Tombor B., Albert R., Oltvai Z.N. & Barabasi A.L. (2000) The large-scale organization of metabolic networks. Nature 407, 651-654.

Jensen K., Kristensen L. & Wells L. (2007) Coloured Petri Nets and CPN Tools for modelling and validation of concurrent systems. International Journal on Software Tools for Technology Transfer, 9, 3-4.

Jerby L., Shlomi T. & Ruppin E. (2010) Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism. *Mol Sys Bio.* 6, 401.

Kaelin Jr WG. (2005) The concept of synthetic lethality in the context of anticancer therapy. *Nat Rev Cancer.* **5**, 689-698.

Kajimoto K., Naraba H. & Iwai N. MicroRNA and 3T3-L1 pre-adipocyte differentiation. Rna 2006;12 (9):1626-32.

Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H. & Herwig R. (2011) ConsensusPathDB: toward a more complete picture of cell biology. *Nucleic Acids Res.* 39,*D712-D717.*

Kanehisa M. & Goto S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.*, 28, 27–30.

Kato M., Putta S., Wang M., Yuan H., Lanting L., Nair I., Gunn A., Nakagawa Y., Shimano H., Todorov I., Rossi JJ. & Natarajan R. (2009) TGF-beta activates Akt kinase through a microRNA-dependent amplifying circuit targeting PTEN. Nat Cell Biol. 11(7):881-9.

Karaman M.W., Herrgard S., Treiber D.K., Gallant P., Atteridge C.E., Campbell B.T., Chan K.W., Ciceri P., Davis M.I., Edeen P.T., Faraoni R., Floyd M., Hunt J.P., Lockhart D.J., Milanov Z.V., Morrison M.J., Pallares G., Patel H.K., Pritchard S., Wodicka L.M. & Zarrinkar P.P. (2008) A quantitative analysis of kinase inhibitor selectivity Nat. Biotechnol. 26(1), 127-32.

Kaznessis,Y. (2006) Multi-scale models for gene network engineering. *Chem Eng Sci,* 61, 940–953.

Keese M., Offterdinger M., Tischer C., Girod A., Lommerse P., Yagublu V., Magdeburg R. & Bastiaens P. (2007) Quantitative imaging of apoptosis commitment in colorectal tumor cells. Book Differentiation 75(9):809-818.

Kestler H.A., Wawra C., Kracher B. & Kuehl M. (2008) Network modeling of signal transduction establishing the global view. BioEssays, 30(11-12):1110-1125.

Kirkpatrick S., Gelatt C.D. & Vecchi J.r. (1983) Optimization by simulated annealing. Science 220, 4598.

Kitano H. (2002) Systems Biology: A Brief Overview. Science, 295.

Kitano H., Funahashi A., Matsuoka Y. & Oda K. (2005) The process diagram for graphical representation of biological networks. Nature Biotechnology, 23(8): 961–966.

Kitano H. (2007) A robustness-based approach to systems-oriented drug design. Nature Reviews Drug Discovery 6, 202-210.

Klamt S., Saez-Rodriguez J., Lindquist A., Simeoni L. & Gilles D. (2006) A methodology for the structural and functional analysis of signaling and regulatory networks. *BMC Bioinformatics* **7**, 56.

Knight R., Maxwell P., Birmingham A., Carnes J., Caporaso G., Easton B., Eaton M., Hamady M., Lindsay H., Liu Z., Lozupone C., McDonald D., Robeson M., Sammut R., Smit S., Wakefield M., Widmann J., Wikman S., Wilson S., Ying H. & Huttley G. (2007) PyCogent: a toolkit for making sense from sequence. *Genome Biol.* **8**(8), R171.

Krebs O., Golebiewski M., Kania R., Mir S., Saric J., Weidemann A., Wittig U. & Rojas I. (2007) SABIO-RK: A data warehouse for biochemical reactions and their kinetics. Journal of Integrative Bioinformatics.

Kroll W. (2008) Biomarkers—predictors, surrogate parameters--a concept definition. In: Schmitz G., Endres S. & Gotte D. (eds) Biomarker. Schattauer, Stuttgart, pp 1–14

Kumar M.S., Lu J., Mercer K.L., Golub T.R. & Jacks T. (2007) Impaired microRNA processing enhances cellular transformation and tumorigenesis. Nature Genetics 39(5): 673-677.

Kumar M.S., Pester R.E., Chen C.Y., Lane K., Chin C. & Lu J. (2009) Dicer1 functions as a haploinsufficient tumor suppressor. Gene & Development 23(23): 2700-2704.

Krutzfeldt J, Rajewsky N, Braich R, Rajeev KG, Tuschl T, Manoharan M. & Stoffel M.(2005) Silencing of microRNAs in vivo with 'antagomirs'. Nature 438: 685–689.

Kumamoto K., Spillare EA., Fujita K., Horikawa I., Yamashita T., Appella E., Nagashima M., Takenoshita S., Yokota J. & Harris CC. (2008) Nutlin-3a activates p53 to both down-regulate inhibitor of growth 2 and up-regulate mir-34a, mir-34b, and mir-34c expression, and induce senescence. Cancer Res. 68(9): 3193-203.

Lambrou I., Adamaki M., Koultouki E. & Moschovi M. (2012). Systems Biology Methodologies for the Understanding of Common Oncogenetic Mechanisms in Childhood Leukemic and Rhabdomyosarcoma Cells. Quality Assurance in Healthcare Service Delivery, Nursing and Personalized Medicine: Technologies and Processes. IGI Global. Lazakidou A. & Daskalaki A. (ed).

Lauffenburger D.A. & Linderman J.J. (1993) Receptors: models for binding, trafficking, and signaling pp x, 365. New York: Oxford University Press.

Lee R.C., Feinbaum R.L. & Ambros V. (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell, 75, 843-54.

Lee Y., Ahn C., Han J., Choi H., Kim J., Yim J., Lee J., Provost P., Rådmark O., Kim S. & Kim V.N. (2003) The nuclear RNase III Drosha initiales microRNA processing. Nature 425:415-9.

Lee P.S., Wilhelmson A.S.K., Hubner A.P., Reynolds S.B., Gallacchi D.A., Chiou TT. & Kwiatkowski D.J. (2010) mTORC1-S6K Activation by Endotoxin Contributes to Cytokine Up-Regulation and Early Lethality in Animals. PLoS ONE 5(12): e14399. doi:10.1371/journal.pone.0014399.

Lewis B.P., Burge C.B. & Bartel D.P. (2005) Conserved seed pairing, often flanked by adenosine, indicates that thousands of human genes are microRNA targets. Cell 120, 15-20.

Li N., Fu H., Tie Y., Hu Z., Kong W., Wu Y. & Zheng X. (2009) miR-34a inhibits migration and invasion by down-regulation of c-Met expression in human hepatocellular carcinoma cells. Cancer Lett. 275(1): 44-53.

Li C., Ge Q-W., Nakata M., Matsuno H. & Miyano S. (2006) Modelling and simulation of signal transductions in an apoptosis pathway by using timed Petri nets. J. Biosci. 32: 113-127.

Li J. & Mansmann U. (2013) Modeling of non-steroidal anti-inflammatory drug effect within signaling pathways and miRNA-regulation pathways. PloS ONE (in press).

Li J., Maschke-Dutz E., Heeger F., Hache H., Pandey V., Kamburov A., Kuehn A., Herwig R., Lehrach H. & Wierling C. (2012b) PyBioS2.4: An Advanced Tool for the Design, Modeling and Simulation of Genome-Scale Network. Genome Biology, (in submission).

Li F., Thiele I., Jamshidi N. & Pallson B. (2009) Identification of potential pathway mediation targets in Toll-like receptor signaling. PLoS Comput. Biol. 5(2), e1000292.

Li J, Pandey V, Kessler T, Lehrach H. & Wierling C (2012a) Modeling of miRNA and Drug Action in the EGFR-Signaling Pathway. PloS ONE. 7(1), e30140.

Liang Y., Ridzon D., Wong L. & Chen C. (2007) Characterization of microRNA expression profiles in normal human tissues. BMC Genomics 8: 166.

Lichtenberger LM. (2001) Where is the evidence that cyclooxygenase inhibition is the primary cause of nonsteroidal anti-inflammatory drug (NSAID)-induced gastrointestinal injury? Topical injury revisited.Biochem Pharmacol. 61, 631-637.

Liu CZ., Yu J., Yu SG., Lavker R., Cai L., Liu W., Yang KG., He X. & Chen S. (2007) MicroRNA-21 acts as an oncomir through multiple targets in human hepatocellular carcinoma. Journal of Hepatology, 53:98-107.

Longley, D.B. & Johnston P.G. (2005) Molecular mechanisms of drug resistance. J Pathol 205: 275-292.

Lu J., Getz G., Miska E.A., Alvarez-Saavedra E., Lamb J. & Peck D. (2005) MicroRNA expression profiles classify human cancers. Nature 435(7043):834-838.

Ludwig J.A. & Weinstein J.N. (2005) Biomarkers in cancer staging, prognosis and treatment selection. *Nat Rev Cancer*, 5:845-856.

Lundberg E., Fagerberg L., Klevebring D., Matic I., Geiger T., Cox J., Algeng C., Lundberg J., Mann M. & Uhlen M. (2010) Defining the transcriptome and proteome in three functionally different human cell lines. *Mol Syst Biol*, **6**, 450.

Lynn F.C. (2009) Meta-regulation: microRNA regulation of glucose and lipid metabolism. Trends in Endocrinoloy and Metabolism. 20:9.

Lytle R.J., Yario T.A. & Steitz J.A. (2007) Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR. PNAS. 104: 9667-9672.

MacNeil LT. & Walhout A. (2011) Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression. Genome Res. 21(5): 9667-9672.

Ma H.W., Sorokin A., Mazein A., Selkov A., Selkov E., Demin O. & Goryanin I. (2007) The Edinburgh human metabolic network reconstruction and its functional analysis. Molecular Systems Biology, 3:135.

Markoswski A. & Markoswski P. (1990) Conditions for the effectiveness of a preliminary test of variance. The American Statistician. 44(4), 322–326.

Mann M., Sheng H., Shao J., Williams CS., Pisacane PL., Sliwkowski MX. & DuBois RN. (2001) Targeting cyclooxygenase 2 and her-2/neu pathways inhibits colorectal carcinoma growth. Gastroenterology. 120, 1713-19.

Manolio T.A. (2010) Genomewide association studies and assessment of the risk of disease. N Engl J Med 363, 166-176. doi: 10.1056/NEJMra0905980.

Markey M. & Berberich SJ. (2008) Full-length hdmX transcripts decrease following genotoxic stress. Oncogene. 27(52):6657-66.

Martello G., Ferrari F., Manfrin A., Cordenonsi M. & Dupont S. (2010) A MicroRNA targeting dicer for metastasis control. Cell 141(7): 1195-1207.

McClellan K., Avard D., Simard J. & Knoppers B. (2012) Personalized medicine and access to health care: potential for inequitable access? European Journal of Human Genetics, 1-5.

Mei M., Ren Y., Zhou X., Yuan XB., Han L., Wang G.X., Jia Z., Pu P.Y., Kang C.S. & Yao Z. (2010) Downregulation of miR-21 enhances chemotherapeutic effect of taxol in breast carcinoma cells. Technol Cancer Res Treat. 9(1):77-86.

Meng J., Dai B., Fang B., Bekele B.N., Bornmann W.G., Sun D., Peng Z., Herbst R.S., Papadimitrakopoulou V., Minna J.D., Peyton M. & Roth J.A. (2010) Combination Treatment with MEK and AKT Inhibitors Is More Effective than Each Drug Alone in Human Non-Small Cell Lung Cancer In Vitro and In Vivo. PLoS ONE 5(11): e14124. doi:10.1371/journal.pone.0014124.

Merritt W.M., Lin Y.G., Han L.Y., Kamat A.A., Spannuth W.A., Schmandt R., Urbauer D., Pennacchio L.A., Cheng J.F., Nick A.M., Deavers M.T., Mourad-Zeidan A., Wang H., Mueller P., Lenburg M.E., Gray J.W., Mok S., Birrer M.J., Lopez-Berestein G., Coleman R.L., Bar-Eli M. & Sood AK. (2008) Dicer, Drosha, and Outcomes in Patients with Ovarian Cancer. N Engl J Med. 359(25):2641-50.

Miller TE., Ghoshal K., Ramaswamy B., Roy S., Datta J., Shapiro CL., Jacob S. & Majumder S. (2008) MicroRNA-221/222 confers tamoxifen resistance in breast cancer by targeting p27Kip1. J Biol Chem. 283(44): 29897-903.

Mo MH., Chen L., Fu Y., Wang W. & FU SW. (2012) Cell-free circulating miRNA biomarkers in cancer. J. Cancer, 3, 432-48.

Moffat J. & Sabatini DM. (2006) Building mammalian signaling pathways with RNAi screens. Nat. Rev. Mol Cell Biol 7:177-187.

Murata T. (1988) Petri Nets: Properties, Analysis and Applications. Proceedings of ThE IEEE 77(4):541-580.

Nagaraj N., Wisniewski J., Geiger T., Cox J., Kircher M., Kelso J., Paabo S. & Mann M. (2011) Deep proteome and transcriptome mapping of a human cancer cell line. Molecular System Biology, 7, 548.

Needleman P., Truk J., Jakschik BA., Morrison AR. & Lefkowith JB. (1986) Arachidonic acid metabolism. Annu Rev Biochem. 55, 69-102.

Nelson D. & Cox M. (2008) Lehninger Principles of Biochemistry. W.H. Freeman & Company.

Novère NL., Finney A., Hucka M., Bhalla US., Campagne F., Collado-Vides J., Crampin EJ., Halstead M., Klipp E., Mendes P., Nielsen P., Sauro H., Shapiro B., Snoep JL., Spence HD. & Wanner BL. (2005) Minimum information requested in the annotation of biochemical models (MIRIAM). Nature Biotechnology, 23, 1509-1515.

Oda K., Matsuoka Y., Funahashi A. & Kitano H. (2005) A comprehensive pathway map of epidermal growth factor receptor signaling. Molecular System Biology, 1, 2005.0010.

Olivier B.G., Rohwer J. & Hofmeyr J. (2004) Modelling cellular systems with PySCeS. Bioinformatics. 21(4), 560-561.

Orom U.A., Kauppinen S. & Lund A.H. (2006) LNA-modified oligonucleotides mediate specific inhibition of microRNA function. Gene 372:137–141.

Parra HS., Cavina R., Latteri F., Zucali PA., Campagnoli E., Mornghi E., Grimaldi GC., Roncalli M. & Santoro A. (2004) Analysis of epidermal growth factor receptor expression as a predictvie factor for response to gefitinib in non-small-cell lung cancer. British Journal of Cancer 91:208-212.

Papin J.A., Hunter T., Palsson B.O. & Subramaniam S. (2005) Reconstruction of cellular signalling networks and analysis of their properties. Nat Rev Mol Cell Biol 6: 99–111.

Pawson T. (1995) Protein modules and signalling networks. Nature 373, 573–580.

Petri, C.A. (1962) Kommunikation mit Automaten (PhD Thesis), Institut für Instrumentelle Mathematik, Bonn, Germany.

Pineau P., Volinia S., McJunkin K., Marchio A., Battiston C., Terris B., Mazzaferro V., Lowe S.W., Croce C.M. & Dejean A. (2010) miR-221 overexpression contributes to liver tumorigenesis. Proc Natl Acad Sci U S A. 107(1):264-9.

Poligone B. & Baldwin AS. (2001) Positive and negative regulation of NF-kappaB by COX-2: roles of different prostaglandins. J Biol Chem. 276, 38658-64.

Price N.D., Papin J.A., Schilling C.H. & Palsson B.O. (2003) Genome-scale microbial in silico models: the constraints-based approach. Trends Biotechnol 21, 162-169.

Price N.D., Reed J.L. & Palsson B.O. (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. Nat Rev Microbiol 2, 886-897.

Raser J. & Shea E. (2010) Noise in Gene Expression: Origins, Consequences, and Control. Science, 309, 2010-2013.

Ramakrishna R., Edwards J.S., McCulloch A. & Palsson BO (2001) Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints. Am J Physiol Regul Integr Comp Physiol, 280, R695-704.

Randall H., Beebe-Donk J. & Alshafie G. (2007) Reduced risk of human lung cancer by selective cyclooxygenase 2 (Cox-2) Blockade: results of a case control study. Int J Biol Sci. 3(5), 328-334.

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

Romero P., Wagg J., Green M., Kaiser D., Krummenacker M., & Karp P. (2004) Computational prediction of human metabolic pathways from the complete human genome. Genome Biol 6: R2.

Ryan B.M., Robles A.I. & Harris C.C (2010) Genetic variation in microRNA networks: the implication of cancer research. Nat Rev Cancer, 10(6):389-402.

Sachs R.K., Hlatky L.R. & Hahnfeldt P. (2001) Simple ODE models of tumor growth and anti-angiogenic or radiation treatment. Mathematical and Computer Modeling, 33: 1297-1305.

Sandler R., Halabi S., Baron J., Budinger S., Paskett E., Keresztes R., Petrelli N., Pipas M., Karp D., Loprinzi C., Steinbach G. & Schilsky R. (2003) A randomized Trial of Aspirin to prevent colorectal adenomas in patients with previous colorectal cancer. N Engl J Med. 348 (10), 883-90.

Shi M., Liu D., Duan HJ., Shen BF. & Guo N. (2010) Metastasis-related miRNAs, active players in breast cancer invasion, and metastasis. Cancer Metastasis Rev. 29(4): 785-799.

Shlomi T., Cabili M. & Ruppin E. (2009) Predicting metabolic biomarkers of human inborn errors of metabolism. Molecular Systems Biology, 5:263.

Shreedhar V., Giese T., Sung VW. & Ullrich SE. (1998) A cytokine cascade including prostaglandin E2, IL-4, and IL-10 is responsible for UV-induced systemic immune suppression. J Immunol. 160, 3783-9.

Slepchenko B.M., Schaff J.C., Macara I. & Loew L.M. (2003) Quantitative cell biology with the Virtual Cell. Trends Cell Biol. 13(11), 570-6.

Smallbone, K., Gatenby, R.A., Gillies, R.J., Maini, P.K. & Gavaghan, D.J. (2007) Metabolic changes during carcinogenesis: Potential impact on invasiveness. J. Theor. Biol. 244, 703-713.

Smalley W. & DuBois RN. (1997) Colorectal cancer and non steroidal anti-inflammatory drugs. Adv Pharmacol. 39, 1-20.

Smith A. Balazinska M. Baru C., Gomelsky M., McLennan M., Rose L., Smith B., Stewart E., Kolker E. (2011) Biology and data-intensive scientific discovery in the begninning of the 21st century. OMICS. 15(4), 209-12.

Stinson S., Lackner M.R., Adai A.T., Yu N., Kim H.J., O'Brien C., Spoerke J., Jhunjhunwala S., Boyd Z., Januario T., Newman R.J., Yue P., Bourgon R., Modrusan Z., Stern H.M., Warming S., de Sauvage F.J., Amler L., Yeh R.F. & Dornan D. (2011) TRPS1 targeting by miR-221/222 promotes the epithelial-to-mesenchymal transition in breast cancer. Sci Signaling, 4, ra41.

Su A.I., Welsh J.B., Sapinoso L.M., Kern S.G., Dimitrov P., Lapp H., Schultz P.G., Powell S.M., Moskaluk C.A., Frierson H.F. & Hampton G.M. (2001) Molecular classification of human carcinomas by use of gene expression signatures. Cancer Res. 61(20), 7388-93.

Subramanian S., Thayanithy V., West RB., Lee CH., Beck AH., Zhu S., Downs-Kelly E., Montgomery K., Goldblum JR., Hogendoorn PC., Corless CL., Oliveira AM., Dry SM., Nielsen TO., Rubin BP., Fletcher JA., Fletcher C.D. & van de Rijn M. (2010) Genome-wide transcriptome analyses reveal p53 inactivation mediated loss of miR-34a expression in malignant peripheral nerve sheath tumours. J Pathol. 220(1):58-70.

Subkhankulova T, Gilchrist M.J. & Livesey F.J. (2008) Modelling and measuring single cell RNA expression levels find considerable transcriptional differences among phenotypically identical cells. BMC Genomics 9: 268.

Tekir D., Arga Y. & Uelgen K. (2007) Drug targets for tumorigenesis: Insights from structural analysis of EGFR signaling network. Journal of Biomedical Informatics, 42(2), 228-236.

Tong A.W. & Nemunaities J (2008) Modulation of miRNA activity in human cancer: a new paradigm for cancer gene therapy? Cancer Gene Thre, 15(6), 341-55.

Torrance CJ., Jackson PE., Montgomery E., Kinzler KW., Vogelstein B., Wissner A., Nunes M., Frost P. & Discafani CM. (2000) Combinatorial chemoprevention of intestinal neoplasia. Nat Med. 6, 1024-8.

Tortora G, Caputo R, Damiano V, Melisi D., Bianco R., Fontanini G., Veneziani BM., De Placido S., Bianco AR. & Ciardiello F. (2003) Combination of a selective cyclooxygenase-2 inhibitor with epidermal growth factor receptor tyrosine kinase inhibitor ZD1839 and protein kinase A antisense causes cooperative antitumor and antiangiogenic effect. Clin Cancer Res. 9, 1566-72.

Tsuchida K., Nakatani M., Hitachi K., Uezumi A., Sunada Y., Ageta H. & Inokuchi K. (2009) Activin signaling as an emerging target for therapeutic interventions. Cell Communication and Signaling, 7, 15.

Tsujii M., Kawano S., Tsuji S., Sawaoka H., Hori M. & DuBioS RN. (1998) Cyclooxygenase regulates angiogenesis induced by colon cancer cells. Cell. 93, 705-716.

Uefuji K., Ichikura T. & Mochizuki H. (2000) Cyclooxygenase-2 expression is related to prostaglandin biosynthesis and angiogenesis in human gastric cancer. Clin Cancer Res. 6, 135-8.

Varma A., Boesch B.W. & Palsson B.O. (1993) Stoichiometric interpretation of Escherichia coli glucose catabolism under various oxygenation rates. Appl Environ Microbiol 59, 2465-2473.

Vastrik I. D'Eustachio P., Schmidt E., Joshi-Tope G., Gopinath G., Croft D., Bono B., Gillespie M., Jassal B., Lewis S., Matthews L., Wu G., Birney E. & Stein L. (2007) Reactome: a knowledge base of biologic pathways and processes. Genome Biol., 8, R39.

Vester B. & Wengel J. (2004) LNA (locked nucleic acid): high-affinity targeting of complementary RNA and DNA. Biochemistry 43: 13233–13241.

Vo T.D., Greenberg H.J. & Palsson B.O. (2004) Reconstruction and functional characterization of the human mitochondrial metabolic network based on proteomic and biochemical data. J Biol Chem 279: 39532–39540.

Voorhoeve, P.M. & Agami, R.(2007) Classifying microRNAs in cancer: The good, the bad and the ugly. Biochim. Biophys Acta 1775: 274-282.

Wang Y. & Blelloch R. (2009) Cell cycle regulation by microRNAs in embryonic stem cells. Cancer Res 69:4093-4096.

Wang R., Wang X., Lin F., Gao P., Dong K. & Zhang HZ. (2008) shRNA-targeted cyclooxygenase (COX)-2 inhibits proliferation, reduces invasion and enhances chemosensitivity in laryngeal carcinoma cells. Molecular and Cellular Biochemistry. 317(1-2), 179-188.

Wang D., Wang H., Shi Q., Katkuri S., Walhi W., Desvergne B., Das S.K., Dey S.K. & DuBois R. (2004) Prostaglandin E(2) promotes colorectal adenoma growth via transactivation of the nuclear peroxisome proliferator-activated receptor delta. Cancer Cell. 6, 285-95.

Webster RJ., Giles KM., Price KJ., Zhang PM., Mattick JS. & Leedman PJ. (2008) Regulation of Epidermal Growth Factor Receptor Signaling in Human Cancer Cells by MicroRNA-7. J Biol Chem. 284(9):5731-41.

Weinberg R.A. (2007) The Biology of Cancer. Garland Science, Taylor & Francis Group, LLC.

Weinberg R.A. (2008) Mechanisms of malignant progression. Carcinogenesis. 29, 1092-1095.

Wiener N. (1948) Cybernetics or Control and Communication in the Animal and the Machine. MIT Press, Cambridge, (MA).

Wiback S.J. & Palsson B.O. (2002) Extreme pathway analysis of human red blood cell metabolism. Biophys J 83: 808–818.

Wierling C., Kuehn A., Hache H., Daskalaki A., Maschke-Dutz E., Peycheva S., Li J., Herwig R. & Lehrach H. (2012) Prediction in the face of uncertainty: a Monte Carlo-based approach for systems biology of cancer treatment. Mutation Research - Genetic Toxicology and Environmental Mutagenesis; Special Issue – Systems Biology. doi:10.1016/j.mrgentox.2012.01.005

Williams T., Flecha A., Keller P., Ram A., Karnak D., Galbán S., Galbán C., Ross B., Lawrence T., Rehemtulla A. & Sebolt-Leopold J. (2012) Cotargeting MAPK and PI3K Signaling with Concurrent Radiotherapy as a Strategy for the Treatment of Pancreatic Cancer. *Mol Cancer Ther,* **11**, 1193.

Wingender E., Dietze P., Karas H. & Knpel R. (1995) TRANSFAC: A Database on Transcription Factors and Their DNA Binding Sites. Nucleic Acids Res., 24, 238–241.

Winter J., Jefferson L. & Kimball S (2010) ERK and Akt signaling pathways function through parallel mechanisms to promote mTORC1 signaling. Am J Physiol Cell Physiol, 300(5), C1172-C1180.

Wolfe MM., Lichtenstein DR. & Singh G. (1999) Gastrointestinal toxicity of nonsteroidal antiinflammtory drugs. N Engl J Med. 340, 1888-99.

Wong BC, Jiang XH, Lin MC, Tu SP, Cui JT, Jiang SH, Wong WM, Yuen MF, Lam SK, Kung HF. (2004) Cyclooxygenase-2 inhibitor (SC-236) suppresses activator protein-1 through c-Jun NH2-terminal kinase. Gastroenterology. 126, 136-47.

Wong TS., Liu XB., Wong BY., Ng RW., Yuen AP. & Wei WI. (2008) Mature miR-184 as Potential Oncogenic microRNA of Squamous Cell Carcinoma of Tongue. Clin Cancer Res. 14(9):2588-92.

Wu CX., Lin JK., Hong M., Choudhury Y., Balani P., Leung D., Dang LH., Zhao Y., Zeng J. & Wang S. (2008) Combinatorial control of suicide gene expression by tissue-specific promoter and microRNA regulation of cancer therapy. Molecular Therapy, 17(12):2058-2066.

Wurdinger T. & Costa F., (2007) Molecular therapy in the microRNA era. The Pharmacogenomics Journal 7, 297-304.

Wynne S. & Djakiew D. (2010) NSAID inhibition of prostate cancer cell migration is mediated by Nag-1 induction via the p38 MAPK-p75NTR pathway. Mol Cancer Res. 8, 1656.

Yano, S., Kondo, K., Yamaguchi, M., Richmond, G., Hutchison, M., Wakeling, A., Averbuch, S. & Wadsworth, P. (2003) Distribution and function of EGFR in human tissue and the effect of EGFR tyrosine kinase inhibition. Anticancer Res 23:3639–3650.

Yantiss RK., Goodarzi M., Zhou XK., Rennert H., Pirog EC., Banner BF., and Chen YT. (2009) Clinical, pathologic, and molecular features of early-onset colorectal carcinoma. Am J Surg Pathol. 33(4): 572-82.

Ye Y., Wang K.K., Gu J., Yang H., Lin J., Ajani J.A. & Wu X. (2008) Genetic variations in microRNA-related genes are novel susceptibility loci for esophageal cancer risk. Cancer Prev Res(Phila). 1(6): 460-469.

Yoshimoto T., Takahashi Y., Kinoshita T., Sakashita T., Inoue H. & Tanabe T. (2002) Growth stimulation and epidermal growth factor receptor induction in cyclooxygenase-overexpressing human colon carcinoma cells. Adv Exp Med Biol. 507, 403-7.

Zenz T., Häbe S., Denzel T., Mohr J., Winkler D., Bühler A., Sarno A., Groner S., Mertens D., Busch R., Hallek M., Döhner H. & Stilgenbauer S. (2009) Detailed analysis of p53 pathway defects in fludarabine-refractory chronic lymphocytic leukemia (CLL): dissecting the contribution of 17p deletion, TP53 mutation, p53-p21 dysfunction, and miR34a in a prospective clinical trial. Blood. 114(13):2589-97.

Zhang B.H. & Farwell M.A. (2008) A new emerging class of player for disease diagnostics and gene therapy. J Cell Mol Med 12:3-21.

Zhang B.G., Li J.F., Yu B.Q., Zhu Z.G., Liu B.Y. & Yan M. (2012) MicroRNA-21 promotes tumor proliferation and invasion in gastric cancer by targeting PTEN. Oncology Reports 27:1019-1026.

Zhang B., Pan X., Cobb GP. & Anderson T.A. (2007a) MicroRNAs as oncogenes and tumor suppressors. Dev Biol 302:1-12.

Zhang B.H., Wang QL. & Pan X.P. (2007b) MicroRNAs and their regulatory roles in animals and plants. J Cell Physiol 210:279-289.

Zhou R., Yuan P., Wang Y., Hunsberger JG., Elkahloun A., Wei Y., Damschroder-Williams P., Du J., Chen G. & Manji HK. (2009) Evidence for selective microRNAs and their effectors as common long-term targets for the actions of mood stabilizers. Neuropsychopharmacology. 34(6):1395-405.

Publication

- Jörg Schaber, Max Flöttmann, **Jian Li**, Cari-Fredrik Tiger, Stefan Hohmann, Edda Klipp (2011) Automated Ensemble Modeling with *modelMaGe*: Analyzing Feedback Mechanisms in the Sho1 Branch of the HOG Pathway. *PloS ONE*, **6**(3): e14791. doi:10.1371/journal.pone.0014791

  (Jian Li extended the software modelMaGe so that different results stated in this publication can be validated and reproduced)

- **Jian Li**, Vikash Pandey, Thomas Kessler, Hans Lehrach, and Christoph Wierling (2012) Modeling of miRNA and Drug Action in the EGFR Signaling Pathway. *PloS ONE,* **7**(1): e30140. doi:10.1371/journal.pone.0030140

  (Jian Li conceived the idea of the miRNA modeling and supervised the entire project. He also designed, performed and analyzed all experiments stated in this publication and composed the manuscript)

- Wierling C, Kühn A, Hache H, Daskalaki A, Maschke-Dutz E, Peycheva S, **Li J**, Herwig R, and Lehrach H (2012) Prediction in the face of uncertainty: a Monte Carlo-based approach for systems biology of cancer treatment. *Mutation Research - Genetic Toxicology and Environmental Mutagenesis; Special Issue - Systems Biology*

  (Jian Li helped to develop the simulation pipeline of Monte Carlo-based approach)

- **Jian Li**, Elisabeth Maschke-Dutz, Felix Heeger, Hendrik Hache, Vikash Pandey, Atanas Kamburov, Alexander Kühn, Ralf Herwig, Hans Lehrach, and Christoph Wierling (2012) *PyBioS 2.4: A Web-based Tool for Design, Modeling, Simulation, and Analysis of Cellular Systems. Genome Biology* (in Submission).

  (Jian Li conceived the concept for developing new version of the software PyBioS and performed major contribution to the software development. Jian Li also wrote the main paper, performed most experiments and analyzed the results)

- *Jian Li*, Andriani Daskalaki, Vikash Pandey, Hans Lehrach and Christoph Wierling (2012) Global Reconstruction of Human Signaling Network. *Mol Sys Bio* (in Preparation).
  (Jian Li conceived the concept for developing genome-scale signaling network and performed stand-alone contribution to the model construction. He composed the main paper and conducted almost all experiments and analyzed different results)

- **Jian Li** and Ulrich Mansmann (2013) Modeling of Non-Steroidal Anti-Inflammatory Drug Effect within Signaling Pathways and miRNA-regulation Pathways. *PloS ONE* (in press)

  (Jian Li and Ulrich Mansmann conceived the idea for this publication. Jian Li performed contribution to the model construction. He composed the main paper and conducted almost all experiments and analyzed different results)

Major parts of my thesis are representations of the contents of the following publications:

- **Jian Li**, Vikash Pandey, Thomas Kessler, Hans Lehrach, and Christoph Wierling (2012) Modeling of miRNA and Drug Action in the EGFR Signaling Pathway. *PloS ONE,* **7**(1): e30140. doi:10.1371/journal.pone.0030140

 Jian Li and Ulrich Mansmann (2013) Modeling of Non-Steroidal Anti-Inflammatory Drug Effect within Signaling Pathways and miRNA-regulation Pathways. *PloS ONE* (in press)

 Jian Li, Elisabeth Maschke-Dutz, Felix Heeger, Hendrik Hache, Vikash Pandey, Atanas Kamburov, Alexander Kühn, Ralf Herwig, Hans Lehrach, and Christoph Wierling (2012) *PyBioS 2.4: A Web-based Tool for Design, Modeling, Simulation, and Analysis of Cellular Systems. Genome Biology* (in Submission).

A. Supplemental Information

A.1 MIRNA-EGFR MODEL

Attached in CD

A.2 MICRORNA TARGET VALIDATION REFERENCES FOR MIRNA-EGFR MODEL

Attached in CD

A.3 PATHWAY REFERENCES FOR HUMAN SIGNALING MODEL

Attached in CD

A.4 MICRORNA TARGET VALIDATION REFERENCES FOR HUMAN SIGNALING MODEL

Attached in CD

A.5 FEEDBACK LOOPS IN THE HUMAN SIGNALING MODEL

Attached in CD

A.6 PYBIOS2 TUTORIAL (STEP BY STEP)

PyBioS2

PyBioS2 is a Web-based platform for design, simulation and analysis of models for biological systems on cellular level. PyBioS2 is free of charge for academic use and it can be used with or without registration. While unregistered users can only access to the basic features of PyBioS2, registered users can have full-access to PyBioS2's functions and they have a user-specific model directory. The PyBioS2 main page consists of two frames: a navigation frame (left frame, Fig. 1) and a content

Figure 1: Navigation frame and content frame.

frame (right frame, Fig. 1). The navigation menu gives access to an example models folder, the user models folder (leading to a models folder where the user can create new models or model-folders for model organization) and the PyBioS2 tutorial. Clicking on a model in the content frame, it will be directed to the model view site giving access to model specific functions arranged under five different tabs (Model, BioObjects, Reactions, Network, Analysis, Import/Export). In the following I outline how to create a new model with PyBioS2, how to simulate and analyse the model, and finally I briefly describe some specific use cases.

How to construct a model in PyBioS2?

In the following, a brief introduction on the model construction using PyBioS2 is given. Suppose I are going to construct a model of EGF signal transduction (Wang et al., 2006):

EGF ligand binds to the EGF receptor to form the EGFR:EGF complex, which conducts the autophosphorylation process and become active. The activated P-EGFR:EGF can phosphorylate proteins such as PLCG, and it can also interact with many other proteins such as SRC to form complex to phosphorylate the downstream proteins of EGF signaling pathway. Each protein is generated by a corresponding translation reaction and each mRNA (not shown here) is generated by a corresponding transcription reaction.
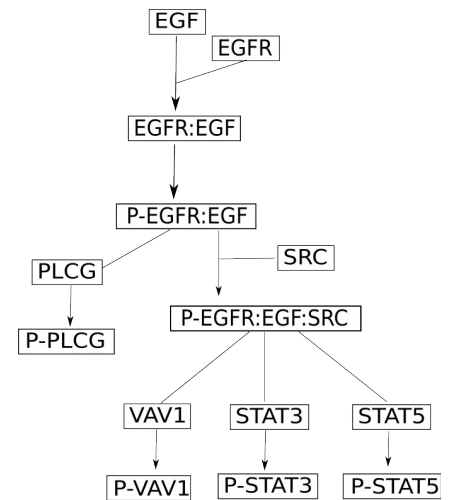
Figure 2: EGF signaling diagram

(1.) Login into your account. If you do not have an account yet, please register!	
	Figure 3: Login site

161

(2.) Click on 'User: <YOUR NAME>' in the navigation frame and then click the 'new model' option in the content frame to create a new model. Enter an identifier and optionally a titel for your new model.
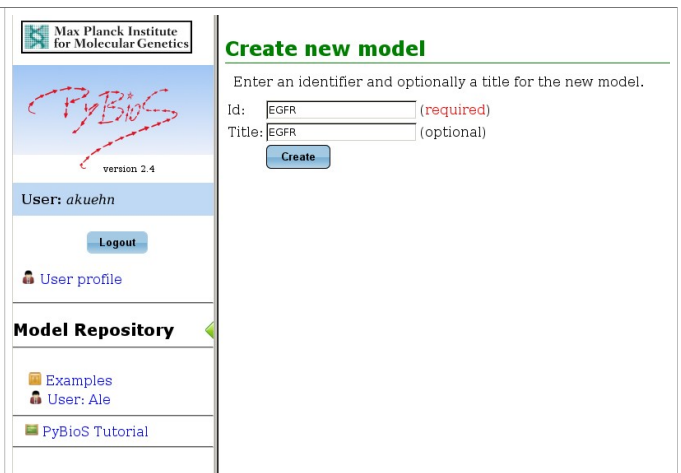
Figure 4: Create new model

(3.) Search for an EGF gene in the Database tree (the tree at the right hand side) by entering the search keyword 'EGF' in the text field on the top of this tree and click the 'search' button. The number appearing in each category node (Gene, Protein, Complex, Reaction, Pathway and others) indicates the number of matching hits in the corresponding category according to the keyword. Searching for objects, reactions and pathways is done for the internal PyBioS2 database (PyDB), the ConsensusPathDB database (CPDB, for mouse and human), and the ChEBI database (reference resource for small molecules and metabolite compounds).

(4.) Drag the EGF gene (ENSG00000138798) object from the 'Gene' category node of the Databese tree and drop it into the 'Cell' within the Model tree. Alternatively, you can click with the right mouse button on the EGF gene object in the
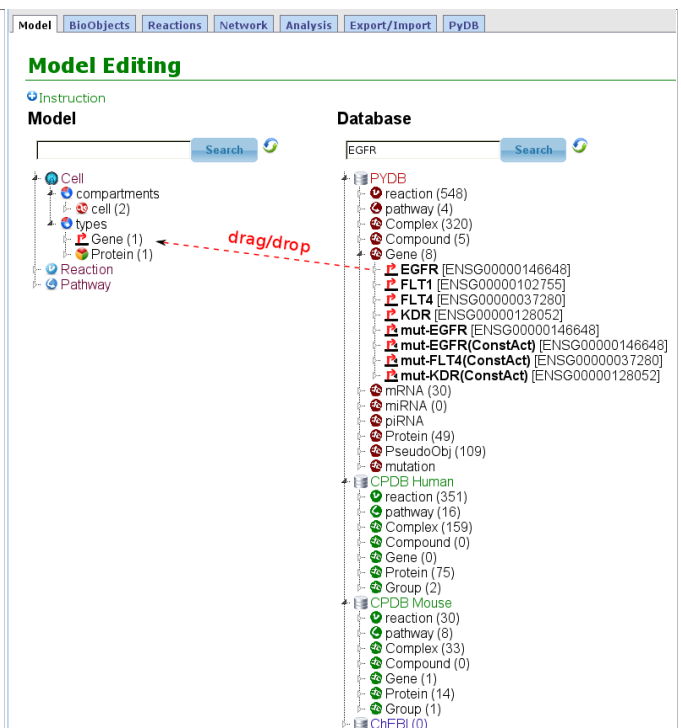
Figure 5: Search, drag and drop

Database tree to get the context-specific menu and select the 'Add to model' option to add this gene object into the currently model.	

(5.) After adding the EGF gene object to the model, this gene object appears twice within the Model tree: under the 'compartments' subtree and the 'types' subtree. This is because objects in a model are sorted twice: based on object-compartment and object-type. Clicking on a gene object opens a dialog window with three sub-tabs: View, Action and Property.
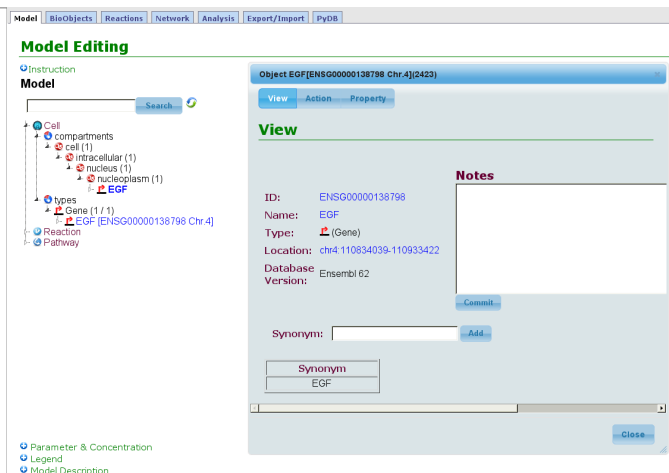
Figure 6: Object dialog window |

(6.) Clicking on the 'Action' tab within the dialog window to display all reactions the current object takes part in. Furthermore, there are two ways to create a new reaction in the model: (i) Manual Action Creation (MAC) and (ii) Automatic Action Generation (AAG). The difference between both options is that for MAC you first need to define all components and then put components into different roles such as substrate (S), product (P), inhibitor (I), enzyme (E) and other, whereas for AAG these two steps are done automatically to speed up the model design/construction process and making sure a consistent component structure of the model.

Figure 7: Manual Action Creation (MAC) and Automatic Action Generation (AAG) |

(7.) In the following the automated method AAG is used to construct this model. Therefore, the option 'transcription' is selected from the drop-down menu to generate the respective mRNA object from the EGF gene object. Multiple different mRNAs will be displayed accounting for alternative splicing effects. Normally, the mRNA with the longest sequence, indicated by '**', should be selected (in this case: 'EGF-2' (ENST00000265171)). Furthermore, a table is shown to list all model objects for being selected as transcriptional activator and repressor, in that case when a certain transcription reaction can only take place in the presence of a certain transcription factors. Clicking on the 'generate' button invokes the automatic generation process. Afterwards, three reactions (transcription, translocation, and decay) and two objects (EGF mRNA in nucleoplasm and EGF mRNA in cytoplasm) are automatically created. Here, the mRNA translocation reaction (from nucleoplasm to cytoplasm) is simplified.

Object EGF[ENSG00000138798 Chr.4](2423)

View | **Action** | Property

Action

There are currently no actions being participated by this object.

EGF ⤴

1. Manual Action Creation: [_____] [Create]

2. Automatic Action Generation: [Transcription ▼] [Generate]

mRNA selection:
- ○ EGF-3 (ENST00000503392)
- ○ EGF-1 (ENST00000509793)
- ○ EGF-2 (ENST00000265171)**
- ○ EGF-4 (ENST00000502723)
- ○ EGF-5 (ENST00000503305)
- ○ EGF-6 (ENST00000504633)
- ○ EGF-7 (ENST00000502579)
- ○ EGF-8 (ENST00000511228)
- ○ EGF-9 (ENST00000509996)

** indicates the longest RNA product

Transcription Activators and Repressors selection (default:None):

Show [10 ▼] entries   Search: [_____]

Transcription-Activator(s)	Transcription-Repressor(s)	(Symbol)Name ▲	Location
No data available in table			
Transcription-Activator(s)	Transcription-Repressor(s)	(Symbol)Name	Location

Showing 0 to 0 of 0 entries

[First] [Previous] [Next] [Last]

Figure 8: Select a mRNA for transcription reaction

164

(8.) To define the EGF ligand protein in the model, please click on the EGF mRNA with the location cytoplasm located in the mRNA category within the Model tree. Then go to 'Action' tab and select the 'translation' in the drop-down menu of AAG. Afterwards, it is required to choose one of three protein types: cytoplasmic protein, secreted protein and plasma membrane protein, to determine the protein location. Since the EGF ligand should be defined within the extracellular space, you need to choose secreted protein type and then click the 'generate' button to invoke the generation process. In this way, the 'translation' reaction as well as the corresponding simplified protein translocation and decay-reactions are defined automatically.
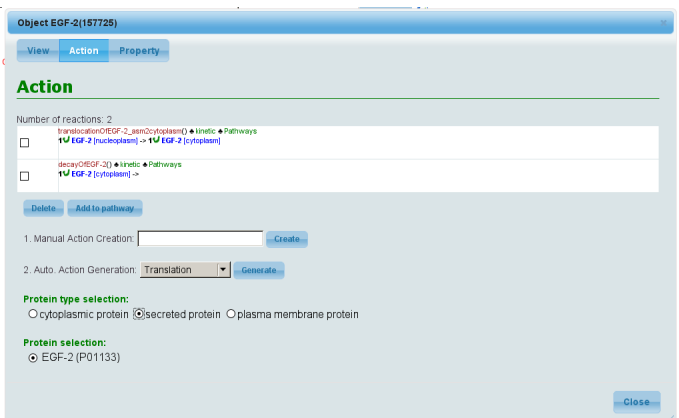
Figure 9: Protein type selection for translation reaction

(9.) Repeat the steps (3) to (8) to define EGFR receptor proteins in the model. Here, it is noteworthy to mention that for the translation reaction of the EGFR protein, you should choose the plasma membrane protein as it represents a receptor protein located in the plasma membrane.

(10.) To define the EGF ligand and receptor binding reaction, click on one of the two proteins (e.g. EGF in extracellular region) located in the protein category within the Model tree and go to the 'Action' tab, choose 'Complex Formation' from the drop-down menu for AAG and select as interacting partner, the other of these two proteins (e.g. EGFR in plasma menbrane), and finally press the 'generate' button to invoke the process. Then the complex 'EGF:EGFR' is created.
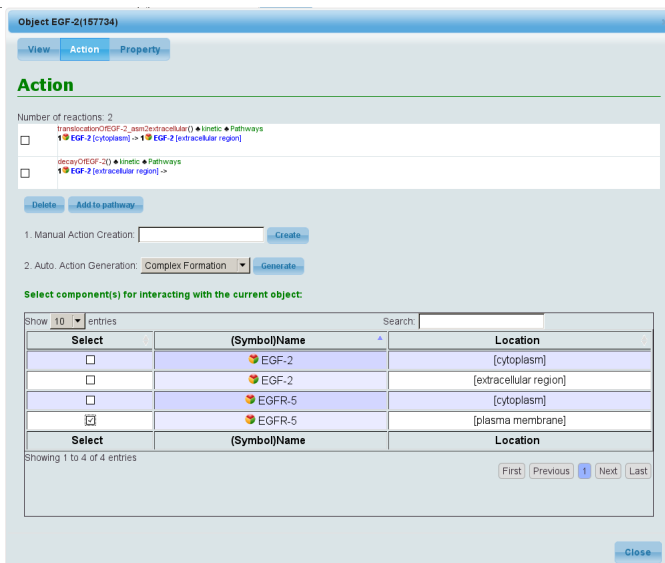

Figure 10: Interaction partner selection for complex formation reaction

(11.) After the complex EGF:EGFR has been created, the autophosphorylation process of this complex can take place. Please select this complex EGF:EGFR and go to the 'Action' tab. Select 'Autophosphorylation' in the drop-down menu of AAG. Here, you have to select which part of the complex is going to be phosphorylated. If the detail information about phosphorylation sites and amount is available, you can also choose the options 'phosphorylation-site amount' and 'Site'. Finally press the 'generate' button to invoke the process. Meanwhile, not only this autophosphorylation reaction but also the corresponding dephosphorylation reaction is defined in the model automatically.

Figure 11: Autophosphorylation reaction

(12.) Following the autophosphorylation of the EGF:EGFR complex, the activated complex P-EGFR:EGF ('P-' symbolizes the phosphorylation site) is going to interact or activate different proteins such as PLCG, SRC and others, to transduce the EGF signal further. Repeat the steps (3) to (8) to define the cytoplasmic protein PLCG and plasma membrane protein SRC.

Figure 12: complex formation reaction

(13.) In order to create an interaction between the complex P-EGFR:EGF and SRC, please go back to the complex P-EGFR:EGF by clicking this complex object located in the complex category within the Model tree and go to the 'Action' tab. Here, you select again the 'Complex Formation' in the drop-down menu of AAG, choose the interacting partner plasma membrane protein SRC for defining the new complex P-EGFR:EGF:SRC and press the button to invoke the process.
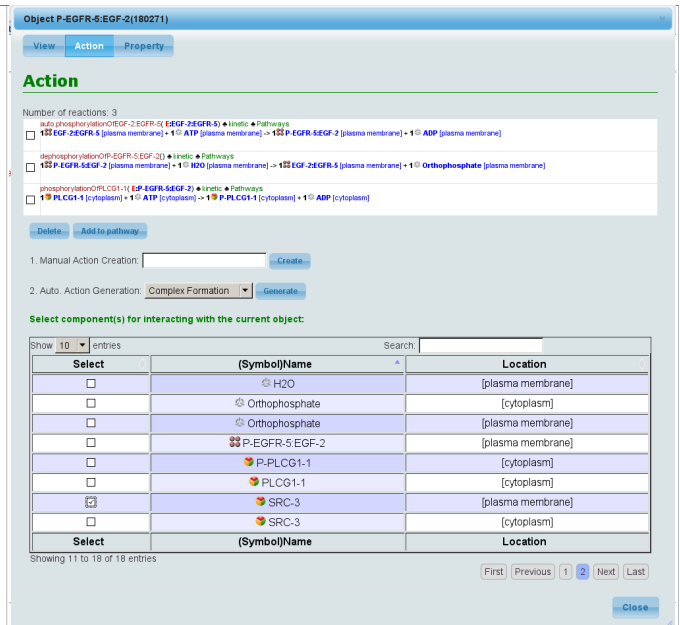

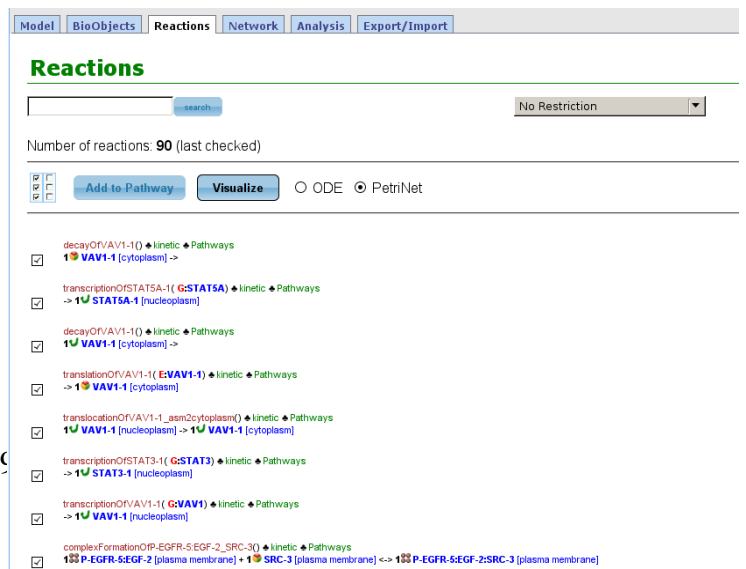*Figure 13: complex formation reaction*

(14.) Repeat the steps (3) to (8) to define the cytoplasmic proteins such as VAV1, STAT3, STAT5A. Apply the complex P-EGFR:EGF:SRC as enzyme to phosphorylate those protein using the step (13).

The final "EGF signaling model" can also be found in the Examples folder.

How to visualize a model in PyBioS2?

A graphical representation of the interaction network of a model can be created in the Network-tab. The network graph symbolizes mass- and information-flow defining either classical chemical reactions with substrates, modifiers and products or physical processes like the flux of a substance across a membrane.

Different types of nodes presenting different types of objects such as gene, protein etc., are connected by directed edges (reaction-

objects). Different types of network components should be selected from the BioObjects tab or Reactions tab. Two types of networks for ODE- (ordinary differential equation system) and Petri net-based models are provided by PyBioS2. The number of object and reaction nodes of the network graph can be selectively extended or shrank by clicking on any network component and selecting 'Expand Network' or 'Shrink Network'.

In each corner of the network-graph additional menus can be found:

(1.) In the left-upper corner, the menu provides options to download the network graph as a SVG file, to upload experimental data to visualize it within the interaction network, and to handle simulation data.

(2.) The menu in the right-upper corner is specifying the simulation time interval. Simulations are performed according to the type of the mathematical model (ODE or Petri net) that was selected beforehand for network visualization.

(3.) In the left-bottom corner, there is a 'Find' option for searching components according to the name within the network graph.

(4.) In the right-bottom corner, there are two options to zoom in and zoom out of the network graph.
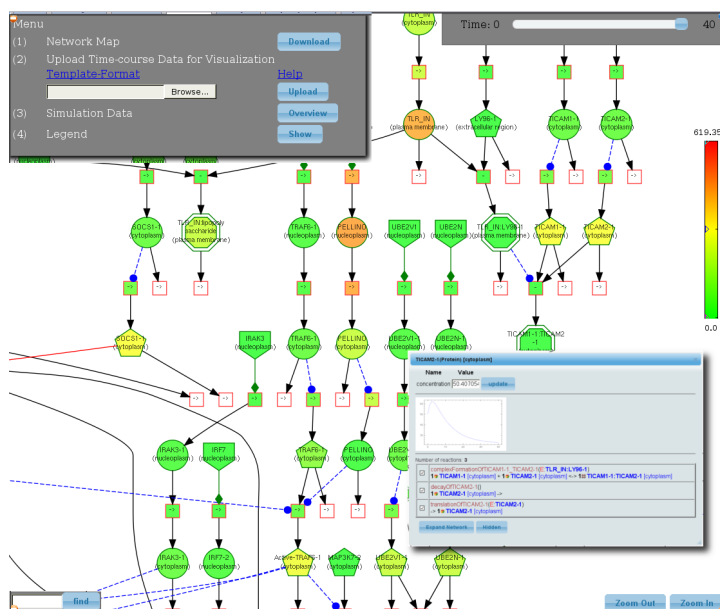
How to simulate a model in PyBioS2?

After the network graph of a model is visualized in the Network tab, you can enter the simulation duration in the option at the right-upper corner and press the 'simulation' button to perform a model simulation process. When a simulation process was successfully accomplished, the network components will be filled with a certain color spanning from green to red for symbolizing the component concentration intensity. By clicking a network component, the simulation time course will be visualized and in the same dialog window users can also expand and shrink the network graph.

Furthermore, a time-slide appears in the right-upper corner, which enables users to go step by step through the entire simulation time course results to follow the dynamic behavior of all model components in the context of the visualized interaction network.

The simulation data including component concentration and flux intensity can be downloaded, recorded and removed from the option menu in the left upper corner. Clicking on the 'overview' button opens a dialog window of handling simulation data. Here, PyBioS2 provides different options for dealing with simulation datasets , comparing them with each other based on concentration change of model components.

Furthermore, this option menu allows it also to initialize the model components with a given simulation data set by clicking on the 'initial' button. By doing so, an extended network simulation and dynamic analysis becomes possible.
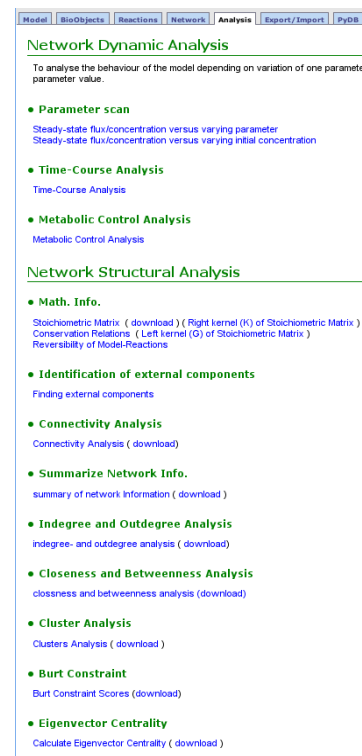
PyBioS2 also provides an option for data visualization within the network graph. Download the template file in menu of the left-upper corner, fill the content with your own data and upload it into PyBioS2 within this menu option. The data will be then visualized within the network graph.

How to analyze a model in PyBioS2?

PyBioS2 is equipped with different mathematical analysis methods that are generally divided into two categories: network dynamic analysis and network structural analysis:

(1) Network dynamic analysis

- Parameter Scan: to analyze the dynamic behavior of a model system, one can perform steady-state analysis verses kinetic parameters.

- Time-Course Analysis: to perform the time course simulation and steady-state analysis. Steady-state search can be performed by two different root-finding methods or via a direct search methods, that uses a large presimulation interval and repeated control calculations to find a stationary state of the system. Further, an analysis for determining the stability of the steady state is performed. The results of these methods are provided as a graphic in textform, as well as for downloading.

- Metabolic Control Analysis: to analyze the sensitivity of network dependent steady-state properties to perturbations of the systems parameters. This method analyzes the relation between individual reaction rates and variations in concentrations and fluxes. The

171

elasticity matrix, response- and control coefficient matrix of the according network graph can be calculated and visualized by this method.

(2) Network structural analysis

Math. Info: to obtain different types of mathematical information about a model system such as Stoichiometric Matrix of model reactions and Conservation Relations of a model system.

Identification of External Components: for a model system (especially a large-scale model), it is of advantage to identify the model components that have only out-going edges or only in-coming edges indicating that those components are only producing or being consumed.

Connectivity Analysis: to analyze the gaps of a model and how well different model components connect with each other (if the result shows only one island of an entire model, it means that there is no gap in the model).

Summary of Network Information: to summarize the basic network statistic of the current model including the amount of nodes, edges; the diameter of network; the density of network; the reciprocity of the network; the average length of network paths; the clustering coefficient of network.

Indegree- and Outdegree-Analysis: to analyze the number of edges to a vertex, since the interaction network is directed, therefore, there are indegree and outdegree value of each vertex.

Closeness- and Betweeness Analysis: to analyze how easily other vertexes can be reached from a vertex (closeness) and a measure of a vertex's centrality in a network equal to the number of shortest paths from all vertexes to all others that pass through this vertex (betweenness).

Clusters Analysis: to calculate the clusters of the current network.

Burt Constraint Scores: to calculate the Burt's constraint scores for given vertexes, also known as structural holes.

Eigenvector Centrality: to measure the influence of a node in the model network.

Farthest Vertex of Network: to find farthest source-node and target-node of the longest path of the current network.

Shortest Circle of Network: to find the shortest circle of the current network.

Self-loops and Multiple-edges: to detect the self-loops and multiple edges of the current network.

Please go to the Analysis-tab to invoke and perform those analysis.

A.7 SEMI-AUTOMATIC REACTION DEFINITION

The following formulas are applied when corresponding semi-automatic options are selected in the PyBioS2 (A, B and C stand for model components including gene, RNA, protein, complex, compound, pseudo object and others).

(1) Activation:

$$A \rightarrow active\text{-}A$$
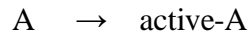
(2) ActivationGTP:

$$A{:}GDP + GTP \rightarrow A{:}GTP + GDP$$

(3) Complex Formation:

$$A + B + C \rightarrow A{:}B{:}C$$
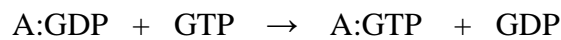
PyBioS2 considers all possible permutation of complex parts:

$$B + A + C \rightarrow B{:}A{:}C, \quad C + B + A \rightarrow C{:}B{:}A, \quad C + A + B \rightarrow C{:}A{:}B$$
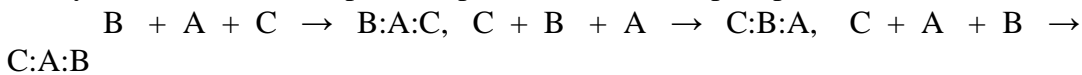
the complex object: A:B:C, B:A:C, C:B:A and C:A:B are considered as the same.

(4) Self Formation (including dimerization, trimerization and tetramerization):
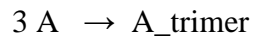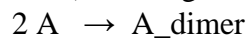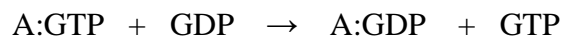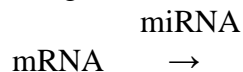
$$2\,A \rightarrow A\_dimer$$
$$3\,A \rightarrow A\_trimer$$
$$4\,A \rightarrow A\_tetramer$$

(5) DeActivationGTP:

$$A{:}GTP + GDP \rightarrow A{:}GDP + GTP$$

(6) Decay:

$$A \rightarrow$$

(7) Dephosphorylation:

$$phos\text{-}A + H_2 0 \rightarrow A + Orthophosphate$$

(8) miRNA_binding:

$$mRNA \xrightarrow{miRNA}$$

(9) Methylation:

$$A \quad \rightarrow \quad \text{methy-A}$$

(10) Oxidization:

$$A \quad + \quad O_2 \quad \rightarrow \quad \text{hydroxy-A}$$

(11) Auto.Phosphorylation:

$$\overset{A}{A \quad + \quad ATP \quad \rightarrow \quad \text{phos-A} \quad + \quad ADP}$$

(12) Phosphorylation:

$$A \quad + \quad ATP \quad \rightarrow \quad \text{phos-A} \quad + \quad ADP$$

(13) Transcription:

$$\text{Gene} \quad \rightarrow \quad \text{mRNA-1}$$
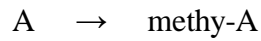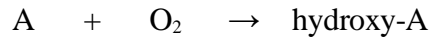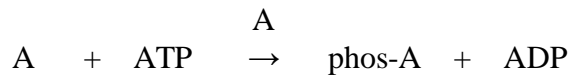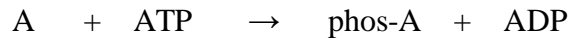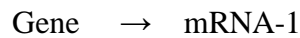
PyBioS2 considers different mRNAs due to alternative splicing. The example above indicates the
first mRNA variant being modulated.

(14) Translocation:

$$B\,(\textit{compartment 1}) \quad \rightarrow \quad B\,(\textit{compartment 2})$$
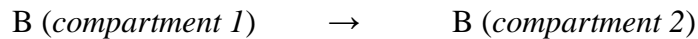
(15) Ubiquitination:

$$A \quad + \quad \text{Ubiquitin} \quad \rightarrow \quad \text{A:Ubiquitin}$$

A.8 Gefitinib, Imatinib and Temsirolimus Drug Association Constant

ID	Display Name	Gefitinib	Imatinib	Temsirolimus
153812	ABL Inhibitor	0	1	0
	ABL Kd	0	8.333	0
154121	CRAF Inhibitor	0	1	0
	CRAF Kd	0	0.058	0
154125	EGFR Inhibitor	1	0	0
	EGFR Kd	100	0	0
211122	ERBB2 Inhibitor	1	0	0

ID				
	ERBB2 Kd	0.028	0	0
157407	ERBB4 Inhibitor	1	0	0
	ERBB4 Kd	0.243	0	0
154129	FYN Inhibitor	0	1	0
	FYN Kd	0	0.032	0
154140	KIT Inhibitor	0	1	0
	KIT Kd	0	7.142	0
154147	mTor Inhibitor	0	0	1
	mTor Kd	0	0	45.454
154149	PDGFRA Inhibitor	0	1	0
	PDGFRA Kd	0	3.225	0
154150	PDGFRB Inhibitor	0	1	0
	PDGFRB Kd	0	7.142	0
154161	SRC Inhibitor	1	0	0
	SRC Kd	0.026	0	0
154749	VEGFR1 Inhibitor	0	0	1
	VEGFR1 Kd	0	0	1.449
154750	VEGFR2 Inhibitor	0	0	1
	VEGFR2 Kd	0	0	1.477
154751	VEGFR3 Inhibitor	0	0	1
	VEGFR3 Kd	0	0	0.172

Table A.8: The summary information of drugs Gefitinib, Imatinib, Temsirolimus. The column 'ID' indicates the model component Id of the inhibitor object in the human signaling model. '1' indicates the corresponding inhibitor object active and '0' indicates inactive. Each Inhibitor has the specific association value (Kd) for the target protein. All these information is derived from the drug database (Materials and Methods 3.7)

A.9 MICRORNA TARGET VALIDATION REFERENCES FOR HUMAN METABOLIC MODEL (HMMA)

Attached in CD

A.10 THE HUMAN METABOLIC MIRNA MODEL (HMMA)

Attached in CD

A.11 The Human Signaling Model

Attached in CD

A.12 The Non-steroidal Anti-Inflammatory Drug (NSAID) Model

Attached in CD

B. Supplemental Figure

B.1 The connectivity-degree distribution of model components with a linear regression line.

Connectivity-degree distribution of model components follows the power-law distribution and the linear regression R2 value is 0.664. The red line presents the perfect case with the regression R2 value of 1.0.

Degree distribution
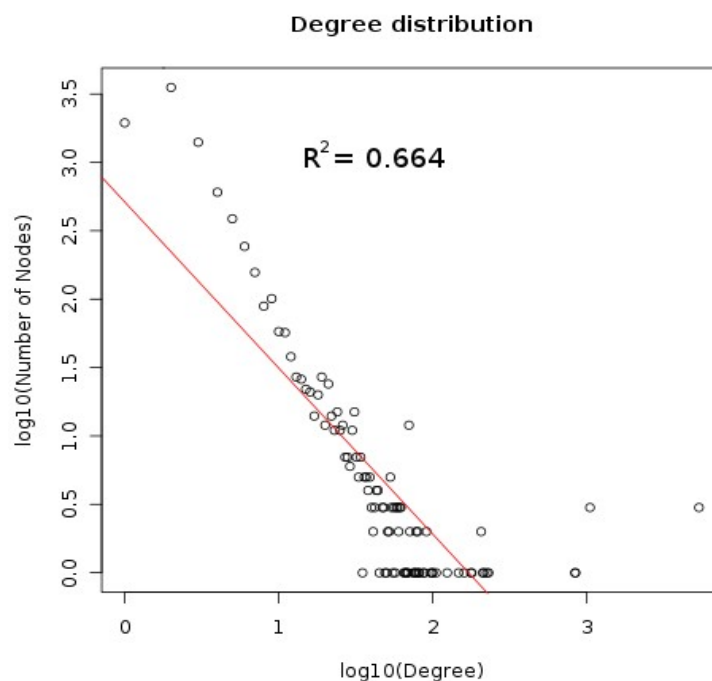
B.2 THE CORRELATION PLOT BETWEEN EXPERIMENTAL PROTEOM DATA AND PREDICTED PROTEOME DATA FROM THE U-2 OS CELL LINE (ATTACHED IN CD).
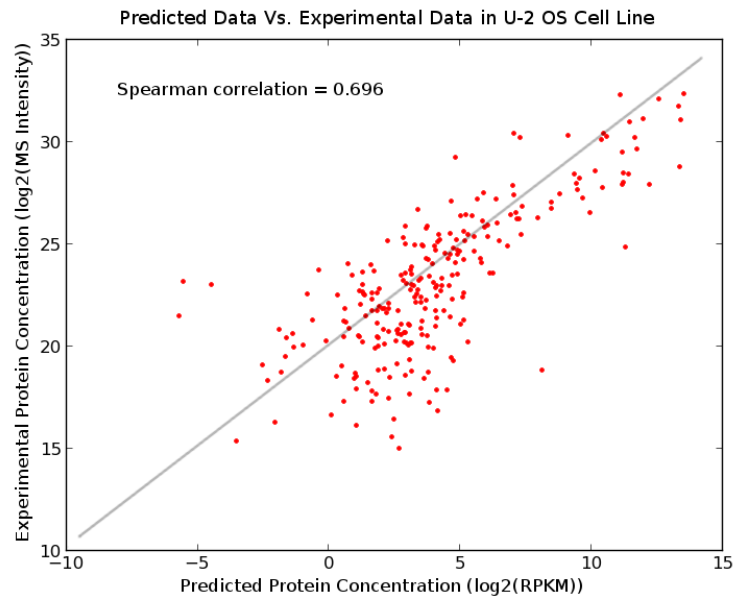
B.3 THE CORRELATION PLOT BETWEEN EXPERIMENTAL PROTEOM DATA AND TRANSCRIPTOME DATA FROM THE U-2 OS CELL LINE (ATTACHED IN CD).
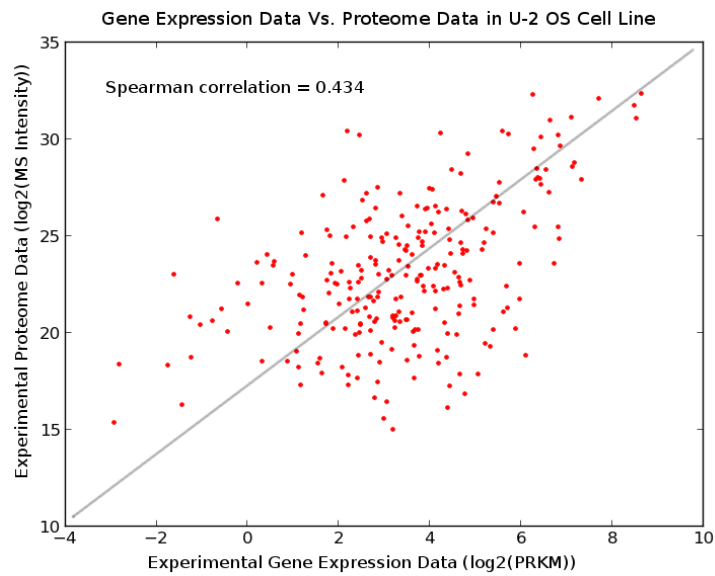
B.4 THE CORRELATION PLOT BETWEEN EXPERIMENTAL PROTEOM DATA AND PREDICTED PROTEOME DATA FROM THE A-431 CELL LINE (ATTACHED IN CD).
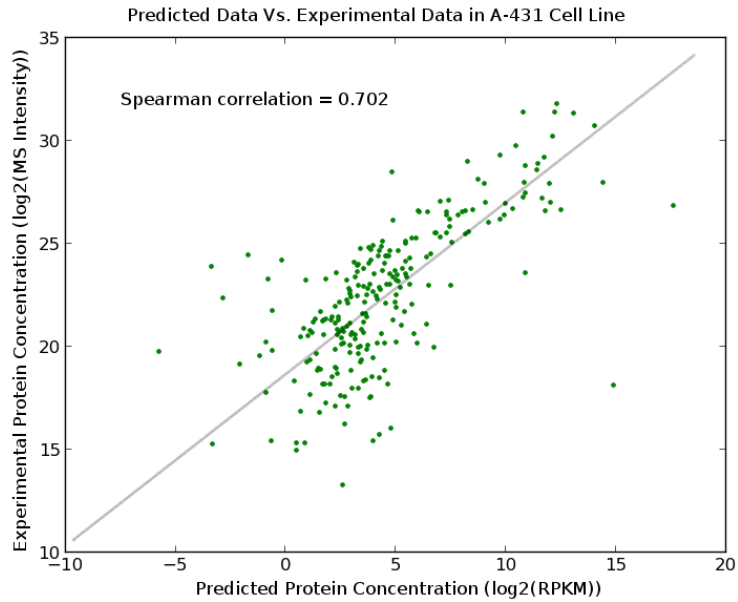
B.5 THE CORRELATION PLOT BETWEEN EXPERIMENTAL PROTEOM DATA AND TRANSCRIPTOME DATA FROM THE A-431 CELL LINE (ATTACHED IN CD).
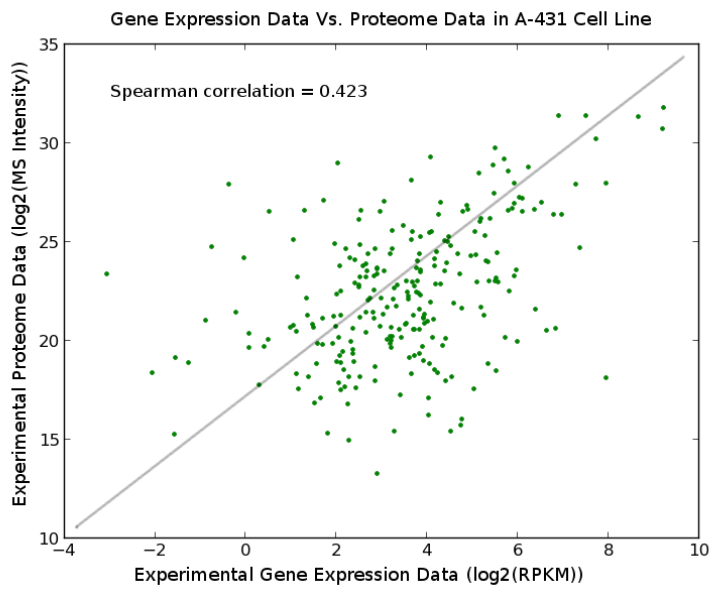
B.6 The correlation plot between experimental proteom data and predicted proteome data from the U-251 MG cell line (attached in CD).

Predicted Data Vs. Experimental Data in U-251 MG Cell Line

Spearman correlation = 0.649

B.7 THE CORRELATION PLOT BETWEEN EXPERIMENTAL PROTEOM DATA AND
TRANSCRIPTOME DATA FROM THE U-251 MG CELL LINE (ATTACHED IN CD).

Gene Expression Data Vs. Proteome Data in U-251 MG Cell Line

Marker Ratios (Log2 Of Geo Mean) bysummary$_r$eadout$_F$GF.tsv
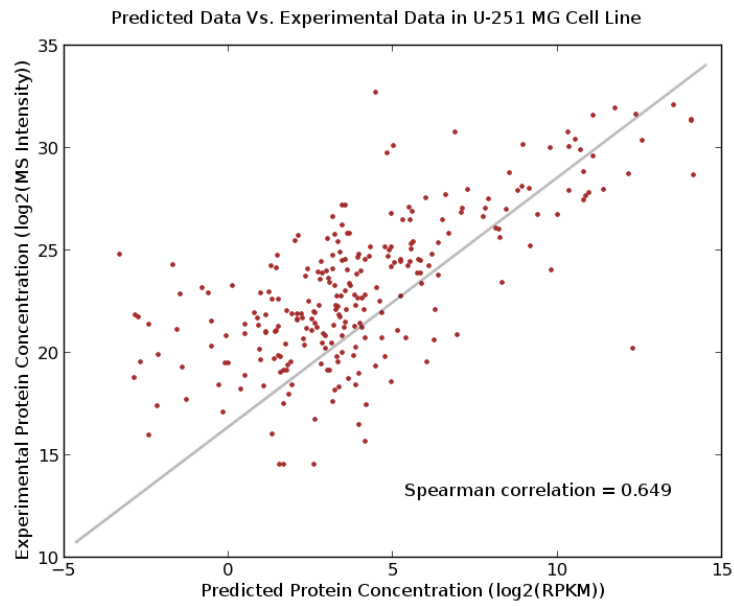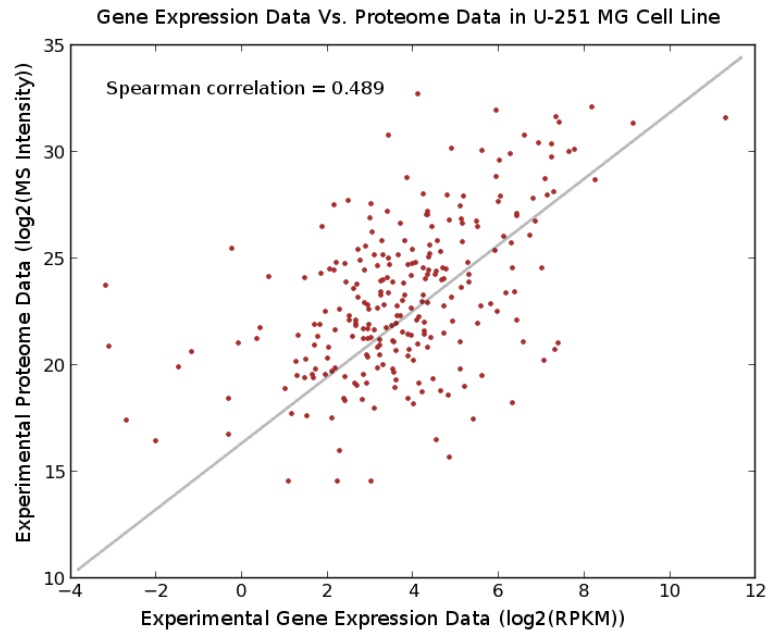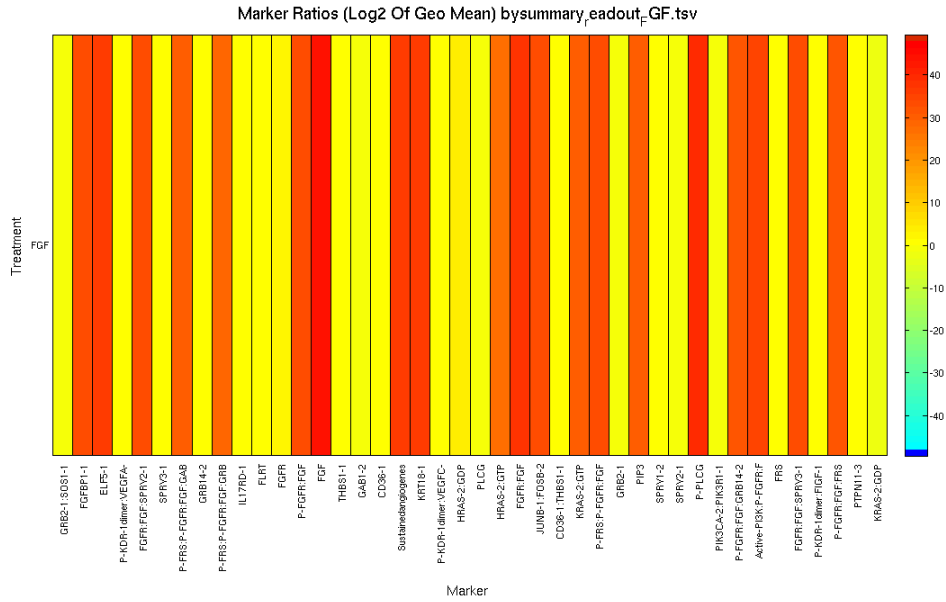
Figure B.8

The heatmap visualizes the comparison result between the FGF-active state and the control state (no activation-signal) in the FGF signaling pathway. The scale value (>0; from yellow to red) indicates that upregulation-degree of a corresponding model component by comparing its concentration in the FGF-active state with it in the control state. This comparison result vigorously indicates that the FGF ligand activation can invoke the activation of many key components in this signaling pathway. The heatmaps of the other signaling pathways are attached in CD.

B.9 COMPARISON OF ODE SIMULATION AND PETRI NET SIMULATION BASED ON A SIMPLE NETWORK.
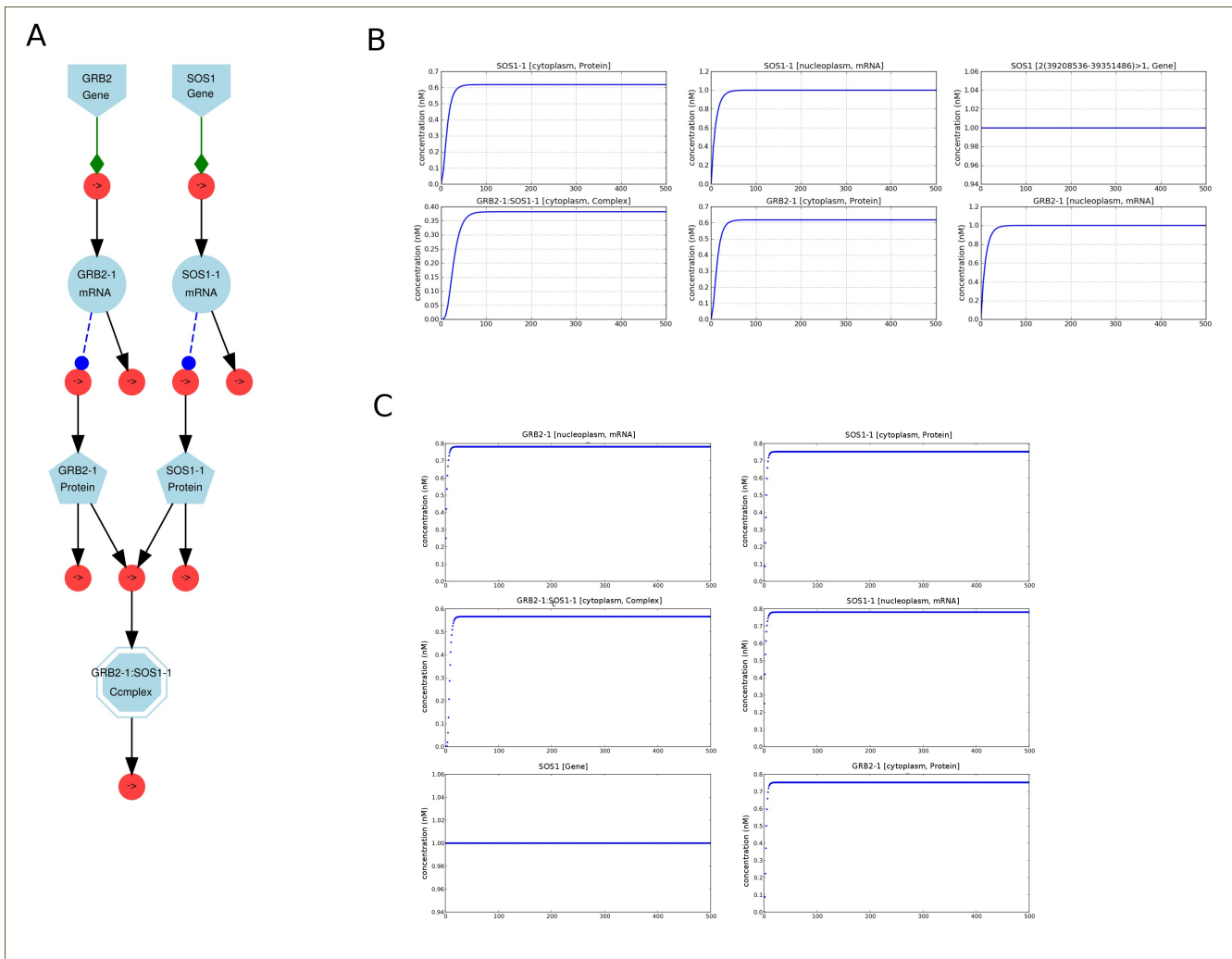
Figure B.9. **A:** a simply biological network consisting of transcription, translation, complex formation and decay reactions. **B:** The plot of time course of model components during the ODE simulation. **C:** The plot of time course of model components during the Petri net simulation.

B.10 THE CANCER CELL LINES FROM CANCER GENOME ATLAS.

Attached in CD

C. Supplemental Table

Pathway	Description	Mechanism
MAPK_signaling	A mitogen-activated protein kinase (MAPK) cascade, which composes of diverse protein phosphorylation cascades that function to control different cellular processes related to proliferation, angiogenesis, apoptosis and other.	The activation of this pathway triggers the phosphorylation of protein kinase RAS and its cousin that in turn phosphorylate MEK1/2. Afterwards, ERK1/2 is phosphorylated by MEK1/2 and activates diverse down-stream targets (such as RPS6KA2, TSC2, AR, ESR1, SMAD2/3, EIF4EBP1)
p38MAPK	Functions in a similar way as MAPK_signaling, possesses protein phosphorylation cascades based on protein kinase MAPK14 (p38) and control different cellular processes. This pathway has intensive crosstalk with signaling pathways such as Activin_signaling, TGF_beta.	The phosphorylation of the protein kinase MAPK14 can be triggered by the different activated cellular components (such as MKK3/6, MKK4/7, ACVR:MSTN, ACVR:ACTIVIN, ACVR:GDF11). This activated protein kinase can in turn phosphorylate many target proteins (such as ELK1, ATF2, TP53).
NFkB_signaling	Includes relevant transcriptional regulations of NF-kappa-Bs and has a multitude of functions (initiating inflammatory response and invading pathogens as a part of an innate immune response).	The activation of diverse signaling pathways leads to phosphorylation of I-kappa-B kinases (IKBKA, IKBKB, IKBKE, IKBKG) that phosphorylate the inhibitors of NF-kappa-Bs to free and activate them. The activated NF-kappa-Bs transcriptional regulate many genes (such as IL2, IL2RA, IL13RA2, IL9, IL13, IL15, CSF2, CSF3, VEGFC)

JAK_STAT	The Janus kinases (JAKs) phosphorylate the signal transducers and activators of transcription (STATs) that convey the messeage from cell-surface into the nucleus.	Diverse ligands (such as IL2, IL4, IL5, IL9, CSF, CSF3, CTF1, IL12A, IL12B, IL6, LIF, OSM and other) bind to their corresponding receptors and phosphorylate different JAKs or combination of them. After the activation of STATs, the activated STATs can be translocated into nucleus to carry the signal information into nucleus by transactivating many target genes (such as SOCS3, CISH, MCL1, PIM1,MYC,TIMP1,IRF1,CD40,FOS, FCGR1A,GBP1,IL21,ISG15,CCL19,IRF7,PRF1,IFNG,IL2RA,IRF4,S100P, CDKN1A,BCL2L1,BCL2,MVP and others).
GP130	Includes a number of interaction related to functional receptor complexes of pathway JAK_STAT that share a common signal transducing receptor component (IL6ST) belonging to the GP130 family.	The ligand-receptor binding (CNTF:CNTFR:IL6ST, LIF:LIFR:IL6ST, IL11:IL11RA:IL6ST, IL6:IL6R:IL6ST, CTF1:IL6ST, OSM:OSMR:IL6ST, LEPR:LEP:IL6ST, CSF3:CSF3R:IL6ST, IL12A:IL12B:IL12RB1:IL12RB2:IL6ST) enables the activation of different JAKs or combination of them (JAK1:JAK2, JAK1:TYK2, JAK2:TYK2, JAK2) that in turn activates STATs (STAT1/3/4/5). The activated STATs translocate into nucleus and activate the corresponding target genes (including IRF1, IFNG, MYC, IL21, PIM1, IRF4, VIP, CDKN1A, FOS, CD40).

BETA_chain	Includes diverse reactions related to functional receptor complexes of pathway JAK_STAT that share a common signal transducing receptor component (CSF2RB) belonging to the Beta chain family.	The ligand-receptor binding (CSF2:CSF2RA:CSF2RB, IL5:IL5RA:CSF2RB, IL3:IL3RA:CSF2RB) enables the activation of JAK2 that in turn activates STAT5. The activated STAT5 translocate into nucleus and activate the corrresponding target genes (including CCND1, MYC, BCL2L1, CISH).
GAMMA_chain	Includes diverse reactions related to functional receptor complexes of pathway JAK_STAT that share a common signal transducing receptor component (IL2RG) belonging to the Gamma chain family.	The ligand-receptor binding (IL2:IL2RA:IL2RB:IL2RG, IL7:IL7R:IL2RG, IL4:IL4R:IL2RG, IL9:IL9R:IL2RG, IL15:IL15RA:IL2RG) leads to the different combination of JAKs (JAK1:JAK3, JAK1:JAK2, JAK1:TYK2), which in turn activate different STATs (STAT1/3/4/5/6). The activated STATs migrate into nucleus and regulate their target genes (including IL21, MMP2, CCND1, IL6ST, HMOX1, FAAH1, REG1A, SOCS1, VEGFA).
TGF_beta	Includes the actions of the transforming growth factor-beta family, which controls the activity of Smad transcription factors.	The ligand-receptor binding (TGFB:TGFBR) activates the SMAD2/3 and other protein kinase (such as MAP3K7, MAPK14, ERK). The activated SMAD2/3 can then translocate into nucleus with the help of SMAD4 to conduct the transcription regulation.
BMP_signaling	This pathway is involved in broad spectra of biological activities in different tissues such as bone, blood vessels, heart, kidney and others. BMPs are members of the transforming growth factor-beta (TGF-beta)	The ligand-receptor binding (BMP:BMPR1:BMPR2) enables the activation of SMAD1/5/8 and subsequent translocation into nucleus with the help of SMAD4. In nucleus,

185

	family, therefore this pathway has a similar signal transduction structure as TGF_beta.	the SMAD1/5/8 transactivate diverse target genes (such as MSX1/2, IGHA, RUNX2, TCF7, BAMBI, CREB3L2).
Activin_signaling	This pathway has a similar function to the previous described pathway TGF_beta.	The ligand-receptor binding (ACTIVIN:ACVR, GDF11:ACVR, MSTN:ACVR) activates the SMAD1/2/3/5/8 and other protein kinases (such as ERK, MAPK14, JNK).
WNT_signaling	Involves a large number of molecular reactions regulating the production of WNT signaling molecules, interactions with receptors and activation of the downstream molecules. It includes the canonical Wnt pathway and non-canonical Wnt pathway.	**Canonical WNT pathway**: the ligand-receptor binding (WNT:FZD:LRP5/6) leads to the activation of DVL, which in turn inhibits the CTNNB degradation complex. This inhibition of degradation enables the CTNNB freely translocate into the nucleus where this transcription activator regulate many genes transcription (including MYC, CCND1, WISP, JUN, PITX2). **Non-Canonical WNT pathway**: the ligand-receptor binding (WNT_CA:FZD_CA:LRP5/6) activates the LPCB, which leads to production of IP3 and DAG. Both key compounds can then activate the corresponding downstream cellular components (including CAMK11, ATF2, MAP3K7, NLK)
Hedgehog_signaling	Functions in a very similar way as the WNT_signaling pathway. It controls the cell proliferation and development begins with the interaction between ligand Hedgehog and transcription factor GLI.	The ligand-receptor binding (PTCH:HH) enables the SMO, SUFU become active, which leads to the activation of GLI. The active GLI migrates into the nucleus and activates a number of target genes

		(including WNT, CCND, PTCH, FOXs)
SWH_signaling	Possesses a core protein kinase cascade in that a series of serine/threonine protein kinases function to regulate diverse cellular processes such as proliferation, apoptosis and other.	The complex formation of FAT:DCHS leads to the subsequent activation of NF2 and MST, which then enables the complex formation of MOBKL1B:SAV1. This complex activates the LATS, which helps YAP to enter into nucleus to activate many target genes (including JUN, TP73, TGFA, IFITM3, GH1).
Notch_signaling	This is a highly conserved signaling pathway that is involved in the proper development of organs (such as brain, immune system, pancreas, heart, blood vessels). This pathway is also a communication channel in which cells decide to overtake different fate of differentiation.	Via the ligand binding (DLL:NOTCH, JAG:NOTCH, CNTN1:NOTCH), the receptor Notch is cleaved away by the complex (PSEN:APH1A:NCSTN:PSENEN) from the cell surface. The internal portion of the Notch receptor is set free and migrates into the nucleus to transcriptional regulate many target genes (including IL12RA, CDKN1A, PTCRA, DTX1, HEYL).
ABL_signaling	The Abelson non-receptor tyrosine kinase (Abl) signaling pathway is involved in a broad range of biological processes such as oncogenesis, cell growth, adhesion, migration, neurite extension. The pathway PDGF_signaling has considerable influence on this pathway.	The key protein kinase ABL1/2 can be phosphorylated by cellular components (such as PDGFA:PDGFB:SRC, PDGFB:PDGFB:SRC). This activated kinase can phosphorylate in turn its downstream target proteins (such as MDM2, DOK1, CTTN, RAC1, CRK, PXN, POLR2A).
AMPK_signaling	Functions as a sensor of cellular energy status such as regulates food intake and energy expenditure, mediate the effects of hormones and cytokines such as leptin, adiponectin. Therefore, this pathway	Upon the ligand-receptor binding (ADIPOR:ADIPOQ, LEPR:LEP ??) or the cellular ATP:ADP ratio falls, the kinase AMP-activated kinase protein (AMPK) is activated, which

	controls the different cellular processes such as proliferation.	in turn activates diverse downstream cellular components (including TP53, PFKFB2, SREBF1, ELAVL1, EEF2K, MEF2, CRTC2) and negatively regulate the activity of mTor signaling pathway.
ATF_signaling	Including a number of transcriptional regulation based on the functions of activating transcription factor ATF1/2/3/4.	Different members of this transcription factor family exert their corresponding functions to regulate target genes (including JUN, CCL5, TNF, FOS1, RB1, VIP, TP53).
ATM_signaling	Including a number of reactions related to the key regulator ATM (Ataxia Telangiectasia Mutated Protein), which respond to DNA damage processes.	ATM functions as a 'caretaker' and suppresses tumorigenesis in different cellular processes. After activation by DNA damage or auto-phosphorylation, this tumor suppressor regulates many targets including CHK1/2, TP53, RAD51, BRCA, NF-kB.
BDNF_signaling	BDNF and its receptors are broadly expressed in mammalian brains. And the corresponding simulated intracellular signaling is critical for neuronal survival, morphogenesis, and plasticity.	The ligand-receptor binding (BDNF:NTRK2) invokes various intracellular signaling pathways including MAPK_signaling, PI3K_AKT. The ligand-receptor binding (BDNF:NGFR) triggers the activation of CASP proteins.
BRCA_signaling	The genes BRCA1/2 hold the major responsibility for developing breast or ovarian cancer, which are the leading cause of cancer-related deaths allover the world. The pathway including many relevant molecular interactions related to the BRCA1/2.	The BRCA1 forms a complex with BARD1 with help of the activated H2AFX. Within this complex, BARD1 helps BRCA1 to exert its biological function, e.g., regulating target genes (including CDKN1B, DDB2, XPC) and response to the DNA damage signal to halt the cell cycle. The BRCA2 forms a complex with

		RAD51 helps to halt the cell cycle by receiving DNA damage signal.
Cell_Cycle	A series of cellular events (G0, G1, S, G2 and M) that leads to cell devision and duplication for proliferation.	**G0**: the resting phase where the cell has accomplished a cycle and stopped dividing. **G1**: the begin phase where G1 checkpoint mechanism (CCND1/2/3 + CDK4/6) ensures the cellular readiness for DNA synthesis. **S**: the DNA replicative phase. The complex CCNE1/2 + CDK2 and CCNA1/2 + CDK2 organize and inspect the DNA replication process. **G2**: the second gap phase where G2 checkpoint mechanism (CCNA1/2 + CDC2) ensures the cellular readiness for enter mitosis phase and divide. **M**: the mitosis phase where cell growth stops and the metaphase checkpoint (CCNB1 + CDC2) ensures the cellular readiness for cell division completion.
DYRK_signaling	Dual specificity tyrosine-regulated kinase (DYRKs) play key roles on the cellular processes such as proliferation and apoptosis and is one of major sensors to a variety of cellular stress conditions including DNA damage, cell cycle arrest. This pathway including many relevant molecular interactions based on functions of DYRKs.	DRYK1A phosphorylates its target protein such as EIF2B5, CCNL2, JNK, CREB, CASP9, FOXO1 and other. DYRK2 can be activated by the phosphorylated ATM and then phosphorylate TP53. DRYK4 can activate protein including MAPT, STAT3 and EIF2B5.
Death_Receptor	Includes the extrinsic and intrinsic apoptosis pathways that are defined as a common property of multicellular organisms.	Upon the ligand-receptor binding (FAS:FASLG, TNFSF10:DR4/5, TNF:TNFRSF1A), different capase proteins (CASP8/10) can be activated to trigger the apoptosis.

		In response to signals resulting from DNA damage, loss of cell-survive factors or other, pro-apoptotic proteins (including BCL2, BCL2L11, PMAIP, BBC3, BAX, BAK) can promote the mitochondrian release of other pro-apopotoic factors including CYCS, AIFM, EndoG, SMAC.
EGFR_signaling	Epidermal Growth Factor (EGF) signaling pathway is one of the most well studied Receptor Tyrosine Kinases (RTKs) pathway. This pathway regulates activities of a wide range downstream targets and plays key roles in diverse cellular processes including cell growth, proliferation. The mutations occurring in different pathway-components such as EGFR, SRC are highly cancer-related.	Binding of ligands (AREG, BTC, EGF, EREG, TGFA) to the dimer receptor (EGFR:EGFR, EGFR:ERBB2) enables the auto phosphorylation of the ligand-receptor complex, which in turn invokes signaling cascades, in which diverse cellular proteins (including SRC, GRB2, SOS1, PLCG, CBL, JAK1/2) take part in.
EPO_signaling	Including relevant reactions related to erythropoiesis in which a pluripotent hematopoietic stem cell can give rise to mature and stage cells. Erythropoietin (Epo) is therefore usually required for survival, proliferation and differentiation of erythoid progenitor cells.	The ligand-receptor binding (EPO:EPOR) enables the activation of diverse cellular components (including LYN, SYK, JAK2, MAP4K1), which in turn activates their corresponding downstream targets (such as MKK3/6, MAPK14, MYC, CCND, PRKCA, CAMK).
ERK5_signaling	A pathway member of mitogen-activated protein kinase (MAPK) cascades family, which is highly conserved module that is involved in various cellular processes, including cell grow, proliferation, differentiation and migration.	The key component ERK5 can be activated by diverse upstream components such as RAS, MEKK2/3, MEK5, SRC and MAP3K8. The activated ERK5 can then activate the downstream components including CREB, MYC, FOS, SAP1 and SGK.
Ephrin_Eph_signaling	Eph receptors and ephrins form the largest receptor-ligand complexes in the receptor tyrosine kinase family. The signaling	Upon the ligand-receptor binding (including EFNA:EPHA, EFNB:EPHB) a bi-directional

	pathway influence the cell-cell interaction and migration. The deregulation of this pathway is highly related to the tumor growth, survival, angiogenesis and metastasis.	signaling is taking place to transform the signal into downstream targets including ERK, RASA, ACP1, NCK, CTNND, FYN, PTK2.
ERBB_signaling	Including reactions based on receptors (ErbB2/3/4) that regulate signals of cell differentiation, proliferation, migration and survival. This pathway has a very similar structure as the EGFR_signaling.	The ligand-receptor binding (including ERBB4:BTC:ERBB2, ERBB4:NRG3/4:ERBB2, ERBB3:NRG1/2:ERBB2, ERBB4:EREG:ERBB2, ERBB4:HBEGF:ERBB2) enable the activation of diverse cellular components such as KRAS, HRAS, RAF, STAT5, PI3K, PKC, S6K, PICTOR.
Estrogen_signaling	This pathway maintains functions of a wide range of mammalian tissues and plays important roles in cellular processes related to growth, development and reproduction. This pathway including signaling transduction process starting from ligand binding of receptor (including ESR1/2, GPER) to the activation of downstream targets.	The ligand-receptor binding (including Estradiol:ESR1, Estradiol:GPER) enables complex formation with SRC, GRB2:SOS1 or the production of cAMP, which in turn leads to activation of RAC1:GTP, CDC42:GTP, KRAS:GTP, HRAS:GTP, PRKACA, RAP1:GTP, RAPGEF:cAMP and other.
FGF_signaling	Fibroblast Growth Factor (FGF) is one of the well studied Receptor Tyrosine Kinases (RTKs) pathway. The FGF binding can lead to activation of a plethora of signaling pathways involved with cell growth, proliferation and contribute to important functions for normal development, tissue maintenance.	The ligand-receptor binding (FGFR:FGF) enables auto.phosphorylation and the subsequent complex formation with FRS, GRB2, GRB14, GAB1, PI3K, which in turn activates PRKCG, PRKCB, KRAS, HRAS, AKT, ERK, MAP3K7, RASGRP1/2/3/4, CAMK, CREB.
FOXO_signaling	FOXOs belong to a sub family of transcription factor FOXs family which plays important roles in regulating	The transcription factors FOXOs regulate the expression of their target genes (including CCND, PDK4,

	expression of genes that are involved in different cellular processes such as cell growth, proliferation, differentiation and longevity. This pathway including relevant transcription regulations based on the FOXOs function.	CEBPB, IGFBP, PPARG, TGFB2, G6PC, IRS2, ESR1, NOS2, AR). However, their functional activities can be negatively regulated by different protein kinases (such as AKT, DYRK, SGK, JNK).
Growth_Hormone	The effects of growth hormone (GH) can directly influence the cell growth, proliferation and metabolism. This pathway includes major signal transduction processes starting by the GH receptor and mediated by multiple signaling pathways.	The ligand-receptor binding (GHR:GH1, GHR:GH2) enables the complex formation with JAK2, which leads to auto.phosphorylation of JAK2 within the complex. The activated complex GHR:GH1/2:JAK2 can activate diverse downstream components including ERK, AKT, MAP3K7, RASGRP1/2/3/4, CAMK, MAP2K4, MAP2K6, STAT1/3/5.
HIF1_signaling	Directly influencd by the mTor_signaling and MAPK_signaling. Includes a large number of transcriptional regulations based on the function of HIF1.	The activation of HIF1A by ERK or MTOR:MLST8:RPTOR enables the complex formation of ARNT:HIF1A and ARNT2:HIF1A, which then transcriptional regulate many target genes (including ABCB1, IGF2, END1, ADM, VEGFA, HK1, HK2, TGFA, MET).
IGF1R_signaling	This signaling pathway influences the cellular processes including proliferation, cell motility, adhesion and contributes to the apoptosis protection and hypoxia signaling. This pathway is often deregulated in different types of cancer and therefore the receptor IGF1R becomes a putative therapeutic target.	The ligand-receptor binding (IGF1R:IGF2, IGF1R:IGF1) leads to the recruitment of IRS and SHC and subsequent activation of corresponding complexes that in turn regulate the downstream components such as RAF, ERK, PKC, ATF2, MAP3K7, PRKCB, PRKCA, CAMK, SMAD4.
InsulinR_signaling	Insulin is the major hormone controlling energy metabolisms and plays important roles by triggering different signaling	Insulin (INS) binds to its receptor (INSR), which leads to the tyrosine phosphorylation of IRS. This

	pathways.	activation of IRS enables the activation of PI3K and subsequent the PDK, which activates AKT. The activated AKT can regulate a number of its downstream components (including GSK3B, MDM2, PAK1/2, AR, MAP3K8, CASP9). On the other hand, GRB2 and SOS1 are recruited to the activated IRS to trigger the signal cascade MAPK_signaling from RAS, RAF to ERK.
Integrin_signaling	Plays vital roles in cellular processes including cell survival, growth, differentiation, migration, inflammatory responses. This pathway includes relevant interactions related to different integrins, collagens and laminins.	The interactions between integrins and collagen (such as ITGA1:ITGB1:COLLAGEN, ITGA2:ITGB1:COLLAGEN) leads to the phosphorylation of SYK, ILK, MAP3K1 and the subsequent phosphorylation of downstream components (such as GSK3B, MDM2, AKT1S1, PAK). The interaction between integrins and laminin (such as ITGA6:ITGB4:LAMA:LAMB:LAMC, ITGA6:ITGB1:LAMA:LAMB:LAMC) leads to the activation of RAF and RHOA:GTP, which in turn activates the downstream components including ERK, ROCK, MYLK, ETV5, LIML.
KIT_signaling	This pathway helps progenitor cells differentiate into blood and/or vascular endothelial cells and plays important roles in amplification/mobilization processes.	The ligand-receptor binding (KITLG:KIT) enables the recruitment of SRC, GRB10, FYN and subsequent the activation of corresponding complexes that in turn activate MAP3K7, RASGRP1/2/3/4,

193

		CAMK, MAP2K4, MAP2K6, ATF2, ERK.
MET_Receptor	Induce mitogenic, motogenic, and morphorgenic cellular responses and the activation of this pathway is required for normal cellular development.	The ligand-receptor binding (HGF:MET) recruits GRB2, SOS1, SHC1, GAB1, which in turns activates the downstream components such as CRKL, STAT, RAS, ERK. On the other hand, the ligand-receptor complex can phosphorylate PKC, PRKCA, PRKCB, CAMK, MAP3K7 via the activation of PLCG.
NGF_signaling	This pathway contributes the cellular maintenance between survival and death and has a similar structure as BDNF_signaling.	Followed by the ligand-receptor binding (NGF:NTRK1), the auto.phosphoyrlation of this complex takes place, which leads to the activation of downstream components (including IRS, RAS, ERK, PKC, PRKCA, PRKCB, CAMK, MAP3K7). On the other hand, the ligand-receptor binding (NGF:NGFR) leads to the recruitment of NGFRAP1 and MAGED1, which leads to activation of CASP2/3 to trigger the apoptosis.
Neurotrophin_signaling	Including relevant interactions related to the two members (NTF3, NTF4) of neurotrophin family. This pathway has a similar structure as BDNF_signaling and NGF_signaling.	The ligand-receptor bindings (NTF3:NGFR, NTF3:NTRK3, NTF4:NGFR) enable the activation of downstream components (including IRS, RAS, ERK, PKC, PRKCA, CASP2/3).
TP53_signaling	Receive signals from diverse pathways (such as AMPK_signaling, Death_Receptor, p38MAPK and PI3K_AKT) and transform these information into the transcriptional regulation level. Includes diverse	The activation of TP53 by different cellular components such as MAPK14, DAPK, AMPK, JNK, HIPK2, CHEK, disables the functional restriction of MDM2 and enables the migration of TP53 into

194

	transcriptional regulations based on function of TP53, TP63 and TP73.	the nucleus to transactivate target genes (including CDKN1A, DUSP1, TGFA, EPHA2, TNFRSF10B, APC).
PAK_signaling	The p21-activated kinases (PAKs) are evolutionarily conserved and widely expressed in diverse tissues (including heart, kidney, lung and other). This pathway includes relevant reactions related to the functions of PAK family members and plays important roles in cellular processes such as mitosis, transcription, translation, cell survival.	PAKs (including PAK1-7) can be phosphorylated by different protein kinases such as AKT, CDC42:GTP, AR). The activated PAKs can phosphorylate downstream targets such as MYLK, STMN, LIMK, AURKA, ESR, JNKs, RHOA, RAF, to exert their biological functions.
PDGF_signaling	This pathway plays key roles in different cellular processes such as angiogenesis, proliferation, migration and is highly implicated in tissue remodelling, patterning and morphogenesis.	The different combination bindings between ligands (such as PDGFA, PDGFB, PDGFC, and PDGFD) and receptors (such as PDGFRA and PDGFRB) transduce the signal and convey it to the downstream components. For instance, the activated receptors interact with SRC, PI3K, SHP2, SHC, and PLCG, to transfer the signal further via MAPK14, JNKs, ERK.
PI3K_AKT	Includes relevant reactions starting from PI3K, transfers signal via the activation of AKT into the downstream target components. This pathway has key roles in regulation of cellular processes such as cell growth, proliferation, cell cycle and metabolism.	The PI3Kcomplex (PIK3CA:PIK3R) can be activated by diverse components including different receptor tyrosine kinases (such as EGFR:EGF, KDR:VEGFA, FLT1:PGF), JAKs, EPHA:EFNA and other. The activated PI3KCA can catalyze the production of PIP3 to activate PDK, which in turn activates the AKT. The powerful protein kinase can regulate many downstream proteins including GSK3beta, RHEB, TP53, JNK, FOXO, mTor, CDKN1A, RPS6KB1 and other.

Prolactin_Receptor	Contributes to the mammary development processes by regulating many important genes expression level.	The ligand-receptor binding (PRLR:PRL, PRLR:GH1) lead to the activation of protein kinases such as JAK2, NEK3, TEC, FYN and recruit SRC, GRB2, SOS1, which in turn activate downstream components such as RAC1:GTP, CDC42:GTP, RAS, ERK, STAT1/3/5, IRF, CDH, PRKCA, PRKCB, PKC.
RAR_signaling	Contributes the regulation of different cellular processes and developments via binding of several heterodimeric nuclear receptors. This pathway including relevant reactions related to transcriptional regulation of those nuclear receptor complexes.	The receptor complex formation (RARA:RXR, RARB:RXR, RARG:RXR) regulate the expression level of many target genes including THBD, MAOB, GRNH2, RARG, RARB, PLCG, SFTPB, APOA1.
RET_signaling	RET proto-oncogene encodes a receptor tyrosine kinase, which can invoke a signaling cascade by binding to corresponding ligands. The signaling cascade plays essential roles in diverse cellular processes. The loss/gain of function of RET gene are associated with various types of human cancer.	The ligand-receptor bindings (GDNF:RET, NRTN:RET, PSPN:RET, ARTN:RET) invoke the activation of NCK, CRK, PLCG. These activated components transfer signal further by activating corresponding downstream targets including PRKCA, PRKCG, CAMK, MAP3K7, RASGRP1/2/3/4, MAP4K5. On the other hand, the ligand-receptor complex can recruit GRB2:SOS1 to activate the downstream target such as RAS, ERK.
TEK_signaling	TEK is non-receptor tyrosine kinase, which is expressed in diverse types of cells including hematopoietic cell, T-cell, liver cell and other. The signaling pathway invoked by this kinase exert considerable effect on the regulation of distinct cellular processes such as angiogenesis,	The ligand-receptor bindings (ANGPT1:TEK, ANGPTL1:TEK) leads to the auto.phosphorylation of TEK kinase, which in turn recruits and activates various components such as GRB2/7/14, DOK2, PI3K and other. The signal is transferred further

		via those activations resulted in phosphorylation of MAP4K5, MAP4K1, AKT, PXN and CDH.
TLR_signaling	This pathway plays a crucial role in inflammatory responses and is highly involved also in many other cellular processes such as apoptosis.	After ligand binding, TLRs undergo the conformational changes required for recruitment of downstream signaling components including MyD88, IRAKs, TAK1, TAB2 and other, which in turn activates the downstream components including IRFs, IFN-beta, AP-1, NF-kB, JNKs and other.
VEGF_signaling	Plays key roles in cellular processes related to blood vessels such as angiogenesis, vasculogenesis.	The ligand-receptor bindings (VEGFA:KDR, VEGFA:FLT1, VEGFB:FLT1, VEGFC:KDR, VEGFC:FLT4) enables the auto.phosphorylation of corresponding ligand-receptor complex, which recruits the components such as GRB2, SOS1, SHC, PLCG. These recruitments lead to the subsequent activation of diverse downstream targets including CALM, PRKCA, ATF2, PTK2B, RASGRP1, MAP3K7, MAP3K.
mTor_signaling	Function as an important sensor to coordinate different cellular processes with the nutrient level and cellular growth factors. This pathway including the signal transduction processes started from mTor complex 1 (mTORC1) and mTor complex 2 (mTORC2).	After activation by RHEB:GTP, the mTORC1 (MTOR:MLST8:RPTOR) is created, which in turn activates diverse downstream components including PPARGC1, HIF1A, EIF4BP1, EIF4G1, RPS6, PPARG, RPS6KB1, PDCD4. The mTORC2 (MTOR:MAPKAP1:MLST8:PRR5:PICTOR) enables the activation of components including SGK1, PRKCG, PRKCA, NEDD4L, IKBK,

		RAF.

C.2 PATHWAY, KEY COMPONENT, TRANSCRIPTIONAL TARGET AND MAJOR INVOLVED MICRORNA

Attached in CD

C.3 PATHWAY, LIGAND, RECEPTOR, DOWNSTREAM TARGETS AND CROSSTALK PATHWAYS

Pathway	Ligand	Receptor	Downstream Targets	Crosstalk Pathways
ACTIVIN_signaling	ACTIVIN, GDF11, MSTN	ACVR, FST, FSTL	ERK, JNK, SMAD1, 2, 3, 5, 8, MAPK14, IGHA1	MAPK_signaling, TGF_beta
AMPK_signaling	ADIPOQ, GCG, LEP	ADIPOR, GCGR, LEPR	TP53, ELAVL1, RPTOR, EEF2K, HNF4A, CRTC2, GYS, MEF2, SLC2A4, CCNG2, SLC2A4RG, PFKFB2, EEF2, HMGCR, AKT, ACACB, SREBF1, CREB, ACACA, TSC	Estrogen_signaling, Cell_Cycle, P53_signaling, p38MAPK, NfkB_signaling, mTor_signaling, Death_Receptor
TEK_signaling	ANGPT1, 2, 4, ANGPTL1	TEK	KRAS, NRAS, HRAS, MAP4K5, MAP4K1, MAP3K17, MAP3K7, AKT, PLCG, PKC	EPO_signaling, Integrin_signaling, MAPK_signaling, VEGF_signaling
PAK_signaling			MYLK, LIMK, EIF4E, AURKA, EIF4G1, STMN1, PCBP, AR, DNAL1,	Ephrin_Eph_signaling, PI3K_AKT_signaling, p38MAPK

			BCL2L11, CDC42, RAF1, ESR1	
DYRK_signaling			MAPT, STAT3, EIF2B5, CASP9, FOXO1, CCNL2, JNK, CREB1, TP53, NOS3, VIP, TRH, IL21, MMP1, CCND3, BCL2L, REG1A, PRF1, TWIST1	WNT_signaling, ATM_signaling, p38MAPK, JAK_STAT, FOXO
EGFR_signaling	AREG, BTC, EGF, EREG, HBEGF, TGFA	EGFR, ERBB2	ERK, AKT, DAG, IP3, PKC, ATF2, MAP3K7, IKBK, KRAS, HRAS, NRAS, STAT1, 3, 5, ERBB2, JUN, RASGRP2,3,4, VAV, PTPN11, CISH, MMP7, GRB2, ERRFI1, RAF1, SRC, IKBK, ARAF, JAK1,2, ARNT, PTK2B, PRKCG, PTPN11, PRKCA, PRKCB, MMP13, RAF, CAV1,	JAK_STAT, InsulinR_signaling, PI3K_AKT_signaling, HIF1_signaling, VEGF_signaling, WNT_signaling, NfkB_signaling
RET_signaling	ARTN, GDNF, NRTN, PSPN	RET	RPS6KB1, GSK3B, AR, MDM2, BRAF, CDKN1A, CDKN1B, MAP3K8, BAD, KRAS, HRAS, RRAS, NCK, CRK, SP1, JNK	Integrin_signaling, p38MAPK, PI3K_AKT_signaling

BDNF_signaling	BDNF	NTRK2, NGFR	KRAS, PKC, BAD, GSK3B, AKT1S1, MAP3K8, TSC, BRAF, CASP, JUN, FOS	MAPK_signaling, p38MAPK
BMP_signaling	BMP	BMPR1, BMPR2	SMAD1, 5, 8, TCF7, SP7, MSX, IBSP, HEY1, HES1	Hedgehog_signaling, MAPK_signaling
ERBB_signaling	BTC, EREG, HBEGF, NRG1,2,3,4	ERBB4, ERBB2, ERBB3	AKT, ERK, PKC, PRKCA, PRKCG, KRAS, BRAF, CAMK, MEF2, CREB, PTGS2, RRAS, STAT1, 3, 5, ERRFI1, HRAS, PLCG, CREB, ETS1, MYC, MAX	EGFR_signaling, MAPK_signaling, WNT_signaling, Cell_Cycle, NfkB_signaling
NOTCH_signaling	CNTN1, DLL, JAG	NOTCH1-4	NRARP, HES5, DTX1, LASS1, HEY1, 2, HEYL, PTCRA, HES1, RBPJ	Cell_Cycle, MAPK_signaling
PI3K_AKT_signaling			CCND1,2, EIF4B, PKC, CDKN1A, RPS6KB1, AR, PDCD4, MDM2, KLK3, EEF2K, INK, PIGR, RPS6, NKX3, EIF4G1, TSC2, CCNE, mTor, CASP9, PIP3, LCP1, FOXO, PTPN11, CCNB, BRAF, CHEK1,	MAPK_signaling, EGFR_signaling, mTor_signaling, FOXO_signaling, Cell_Cycle, WNT_signaling, InsulinR_signaling, VEGF_signaling, NfkB_signaling, PDGF_signaling, Death_Receptor

			CDKN1B, MAP3K8, AKT1S1, GSK3B, PDK, TP53	
p38MAPK			TP53, CDC42BPA, FOS, BBC3, TCF, MAP3K7, DUSP4, PERP, RAC1, JUN, FOS, MAPK14, EI24, ACTA2, TNFRSF10B, PMAIP1, VAV, MYC, ERK, MKK3,6	MAPK_signaling, EPO_signaling, AMPK_signaling, NfkB_signaling, TP53_signaling, WNT_signaling, TLR_signaling, Cell_Cycle
MAPK_signaling			MITF, ELK4, CFTR, RAF, NRAS, TRH, DNAJB4, CCND1, CREB1, PKC, RPL10, CAMK4, ACHE, PTH, RPS6KB1, PSEN1, PYGO2, EGR2, SULT1A3, STAT3,5, RPS6KA2, SLC25A3, KRT18, ERBB2, AR, NOS3, DBH, CRH, ERK, ARAF, TIMP1, BRAF, ELK1, RIN1, EIF4G1, CISH, TSC2, NR1H2, HIF1A, SULT1A1, HRAS, TYR, NRAS, JUNB, ADRB, FOSB,	ERBB_signaling, Cell_Cycle, BDNF_signaling, Estrogen_signaling, PI3K_AKT_signaling, FOXO_signaling, ATF_signaling, mTor_signaling, InsulinR_signaling, p38MAPK, WNT_signaling, PDGF_signaling, JAK_STAT

			KRAS, ACTB, TYRP, DMD, SP1, APOE, MYC, RPS6KA5, KCNK5, PRKCA, MMP13, TLR9, MYL3, MEK, MAPK11, HLA-DRB1, ELK1,4, EIF4EBP1, ETS2, RAF, MKK3,6, ATF1, SRC,	
GP130	CNTF, CSF3, CTF1, IL11, IL12A, IL12B, IL6, LIF, OSM	CNTFR, CSF3R, IL6ST, IL11RA, IL12RB1, IL12RB2, IL6R, LIFR, OSMR	STAT1, 2, 3, 4, 5, IRS1, PKC, PIP3, MAP3K1, ETS2, SHC, JAK1,2, CISH, TYK2, SOCS1, NfkB	TLR_signaling, NfkB_signaling
BETA_chain	CSF2, IL3, IL5	CSF2RA, CSF2RB, IL3RA, IL5RA	STAT5, CISH, SHC, CSF2, MAP3K1, NfkB, LYN, JAK2	
GAMMA_chain	IL13, IL15, IL2, IL4, IL7, IL9, TNF, TSLP	IL13RA1, IL13RA2, TNF, IL4R, IL15RA, IL2RG, IL2RA, IL2RB, IL7R, IL9R, CRLF2, IL7R	ALOX15, JUN, ICAM1, OPRM1, IGHE, STAT1, 3, 4, 6, TIMP1, CISH, KRAS, NRAS, PTK2B, IRS1, 2, SOCS1, FCGR1A, NFkB2, PIP3, JAK1, 3, IVL, PTPN11	NfkB_signaling, InsulinR_signaling, TGF_beta
JAK_STAT	IFNA, IFNB, IFNG, IL10	IFNAR1, IFNAR2, IFNGR1, IFNGR2, IL10RA, IL10RB	CISH, TAP1, ISG15, LY6E, IL1R1, FCER2, BACE1, CD274, TWIST1, MVP, REG1A, S100P,	EGFR_signaling, MAPK_signaling, Death_Receptor, Cell_Cycle

			SOAT1, CRP, CYBB, STAT1,2,3,5, AGT, IDO1, IVL, MUC1, AP-1, PIM1, MMP2, CCR5, TIMP1, CD40, FCGRT, SHC, GBP1, FCGR1A, IRF1,7, SOCS1, NOX1, IFNA, OPRM1, IL21, FGG, SP1, FAAH, ICAM1, CCL19, PRF1, NOS3	
ATM_signaling			SIM2, HIPK2, JUN, SND1, CHI3L1, CTBP1, SLC34A2, MDM2, MAT2A, BCL2, KLF4, TRHR, SKIL, PML, ETS1, CYLD, SP1, MYC, RUNX	ATF_signaling, TP53_signaling, BRCA_signaling, Ephrin_Eph_signaling, ABL_signaling, Death_Receptor
Ephrin_Eph_signaling	EFNA, EFNB	EPHA2, 3, 4, 5, 7, EPHB	RRAS, MYLK, ACP1, ETV5, RAC1, CTNND1, NCK1, LIMK, ROCK, STAT3, FYN, PTK2, RHOA, CDC42, PAK1, TP53	ABL_signaling, ATM_signaling
EPO_signaling	EPO	EPOR	MAP3K7, MAP4K1, MAP3K13, STAT5, KRAS, HRAS, NRAS,	HIF1_signaling, KIT_signaling, p38MAPK

			HIF1A, IRS2, MKK4, 7	
Estrogen_signaling	Estradiol	ESR1, GPER	ERK, cAMP, KRAS, HRAS, RAP1, PIP3, VAV, BRAF, ESR, CREB	AMPK_signaling, MAPK_signaling, FOXO_signaling, Cell_Cycle
Death_Receptor	FASLG, TNF, TNFSF10	FAS, TNFRSF6B, TNFRSF1A, TNFRSF1B, TNFRSF21, TNFRSF10A, TNFRSF10B, TNFRSF10C, TNFRSF10D	CASP3, 6, 7, 8, 10, PAK2, HTRA2, DIABLO, ENDOG, BID, CYCS, MAP3K14, HSPB	AMPK_signaling, TP53_signaling, PI3K_AKT_signaling, JAK_STAT, NGF_signaling, FOXO_signaling, ATM_signaling, Cell_Cycle
FOXO_signaling			PDK4, CEBPB, RBL2, KLF2, TSC22D3, NR2C2, ABCB, PRDX, SOD2, BCL2L11, FOXO3, HIF1A, PPARGC, CDKN1B, NOS2, PPARG, IRS2	Death_Receptor, IGF1R_signaling, PI3K_AKT_signaling, Cell_Cycle, MAPK_signaling, TGF_beta, Estrogen_signaling, mTor_signaling
FGF_signaling	FGF1-10	FGFR1-4	ELF5, FGFBP, KRT18, KRAS, HRAS, JUNB, FOSB	VEGF_signaling
VEGF_signaling	FIGF, PGF, VEGFA, VEGFB, VEGFC	KDR, FLT1, 4	IP3, FIGF, HIF1A, KRAS, SP1, PLCG, NfkB2, TP53, ARNT2, PIP3, PTPN11, DAG, CSK, ARNT1,2, RELB	HIF1_signaling, PI3K_AKT_signaling, EGFR_signaling, FGF_signaling, P53_signaling
Prolactin_Receptor	GH1, GH2, PRL	GHR, PRLR	RRAS, IRF1,	

			HRAS, KRAS, NEK3, STAT1, 3, 5, CISH, TEC, VAV, PTK2, SOCS1, 3, ETS2, IRS1, JAK2, FYN, SHC, SOS1	
ATF_signaling			CCND, VIP, CIITA, FOS, ETS2	MAPK_signaling, ATM_signaling, HIF1_signaling, SWH_signaling
MET_Receptor	HGF	MET	RAC1, RRAS, RRAS2, KRAS, STAT3, DOCK1, 8, HIF1A, RAPGEF, CRKL, SKIL, PLCG	HIF1_signaling
Hedgehog_signaling	Hedgehog	HIP1, PTCH, GAS1	Gli, FOXC2, FOXF1, SELP, PTCH, FOXE, WNT2B, SUFU, SMO, FOXA, SOX9, MMP3, JUP, NHLH, HIP1, STK36, FOXL1, IGFBP6, CCND2, SFTPB, KIF7	WNT_signaling, BMP_signaling, Cell_Cycle
IGF1R_signaling	IGF1,2	IGF1R	IGFBP3, IGF1, SHC, IRS1, TP53	HIF1_signaling, FOXO_signaling
ABL_signaling			PTPRO, CRK, XAF1, PXN, RAC1, CRK, NCK1, POLR2A, MDM2, YWHAB, PAK2, PLCG, PTPN11, PIK3AP1, JNK, DOK1, YWHAZ	Ephrin_Eph_signaling, ATM_signaling

Growth_Hormone	VIL2, GCG	GCGR	MITF, ELK4, SRF, CFTR, RAF, MMP2, TRH, DNAJB4, CREB1, PKC, RPL10, CAMK, ACHE, PTH, RPS6KB1, MIR132, PSEN, PYGO2, EGR2, HRAS, SULT1A3, MITF	MAPK_signaling, Ephrin_Eph_signaling
InsulinR_signaling	INS	INSR	WWOX, RRAS, NRAS, SHC, AR, AKT, IRS2, KRAS, PIP3, RAPGEF, CREB, TNK2	EGFR_signaling, PI3K_AKT_signaling, MAPK_signaling, GAMMA_chain, Growth_Hormon
KIT_signaling	KITLG	KIT	FYN, AKT, KRAS, PIP3, VAV, PLCG, EST2, SHC, GRB10	EPO_signaling
TLR_signaling	LY96, lipopolysaccharide	TLR1-10	ISG15, IL28AB9, IRF3, 5, 7, TBK1, MAP3K7, TAB1,2, TICAM1,2, CD40, CXCL, TRAF6, ABCC2, RIPK1, CCL19, IRAK1, 4	NfkB_signaling, p38MAPK
WNT_signaling	NDP, WNT1-11, WIF1, SFRP	FZD, LRP5, 6	CTNNB1, CK1, MMP7, NKD, PITX2, PPARD, SIPA1L1, Active-DVL, CSNK1D, RUVBL, CBY1, GSK3B, TCF	EGFR_signaling, PI3K_AKT_signaling, MAPK_signaling, p38MAPK, Cell_Cycle, Hedgehog_signaling
NGF_signaling	NGF	NTRK1, NGFR	TRAF6, CASP2, 3,	Death_Receptor

			8, FRS, DOK1, KRAS, PIP3, PLCG, RAPGEF, IRS1	
Neurotrophin_signaling	NTF3,4	NTRK3, NGFR	SHC, KRAS, PIP3, PLCG, CASP2,3	
mTor_signaling	IL8, THBS	CXCR1, 2, RRAGA, RRAGC, RRAGD	RHEB, HIF1A, RPS6KB1, EIF4E, EIF4EBP1, PDCD4, FOXO3, ERK, REDD1,2, TSC2, ARNT, PRKCG, RICTOR, NEDD4L, PRKCA, PPARG, EIF4EBP1, PPP1CC, SGK1, MYC, PPARGC1, AKT1S1, GSK3B, TP53	PI3K_AKT_signaling, AMPK_signaling, MAPK_signaling, HIF1_signaling, FOXO_signaling, TGF_beta, Cell_Cycle
PDGF_signaling	PDGFA, PDGFB, PDGFC, PDGFD	PDGFRA, PDGFRB	SRC, CBL, KRAS, PIP3, SP1, VAV, PLCG, RELB, SOS1, SHC, CDH1, FYN, NRAS, PTK2, TNK2	PI3K_AKT_signaling, MAPK_signaling
TGF_beta	TGFB1, 2, 3	TGFBR1-3	CAMK, TNC, IGHA1, SERPINE1, PKC, ITGA5, MAP3K7, CYR61, TGFB1, MAPK14, ERK, FBLN5, SMAD1, 2, 3, 5, PTK2, MMP13, PTHLH, ACTA2, TERT,	mTor_signaling, HIF1_signaling, FOXO_signaling, Cell_Cycle, GAMMA_chain

207

			ELN, SKIL, MYC	
TP53_signaling			HDAC5, S100A2, SCN3B, RCHY1, CX3CL1, PERP, PCNA, PLAGL1, PTTG1, SERPINB5, NDRG1, PRODH, SESN1, TP53AIP1, CHEK1, PMAIP1, TP53INP1, CHEK, PMAIP, FANCC, CD82, ACTA2, TERT, EI24, CASP1, 3, 6, SUB1, POLD1, GADD45A, DDB2, TRIM22,	PI3K_AKT_signaling, AMPK_signaling, p38MAPK, NfkB_signaling, VEGF_signaling, ATM_signaling, Cell_Cycle, Death_Receptor
HIF1_signaling			CD99, NT5E, EDN1, PKM2, GPI, HK2, ENO1, PGK1, BNIP3, LRP1, VEGFA, LDHA, BNIP3L, IGFBP3, CP, TFR2, SLC2A1, TF, CCNG2, ETS1, PROK1, REDD1_2, VIM, HK1, ADM, DDIT4, TNP1, ADRA1B, KRT14, GAPDH, IGFBP2, ENG, NR4A1, CITED2, ERK, MYC, PFKL,	IGF1R_signaling, Cell_Cycle, TGF_beta, EGFR, EPO, MET_Receptor, mTor_signaling, VEGF_signaling, ATF_signaling
Integrin_signaling		ITGA1,2,4,5,6,11,	TNC, MAP4K1,	Ephrin_Eph_signali

208

		D,V, ITGB1,2,3,4,6,7	RAC1, ERK, SYK, CRK, HRAS, AKT, KRAS, ITGA5, PLCG, MAP3K1, CDC42, RAF1, GSK3B	ng, MAPK_signaling, PI3K_AKT_signaling, PAK_signaling
NfkB_signaling			MAP3K14, PKC, TNFRSF10B, MAP3K7, IKBK, CYLD, MMP9, MAP3K8, PIGR, TAP1,CCND1	PI3K_AKT_signaling, ERBB_signaling, Cell_Cycle, AMPK_signaling, p38MAPK, TP53_signaling, EGFR_signaling, TLR_signaling, GAMMA_chain
Cell_Cycle			CDKN1C, E2F, MCM7, MAP3K5, TNFSF11, CCND1,2,3, DIRAS3, CAMK4, HIST1H2AD, CAD, POMC, RB, RBL1, CDKN1A, XAF1, HSP90, GADD45B, BIK, BRD7, H19, PCNA, NOP16, CDC25C, ERBB2, EIF4E, CDC25A, CUL1, BMI1, CDC25C, CCNE, CCND3, CCNG, ODC1, DP1,2, STMN, SMAD2,3, TP53, CHEK1, HIF1A, CDKN1B, MYC, PMAIP1, TCF, TERT, SP1,	PI3K_AKT_signaling, AMPK_signaling, MAPK_signaling, HIF1_signaling, FOXO_signaling, TGF_beta, NOTCH_signaling, mTor_signaling, WNT_signaling, JAK_STAT, NfkB_signaling, SWH_signaling, p38MAPK, P53_signaling

209

Pathway	Ligand	Receptor	Downstream Targets	Crosstalk Pathways
			CYC1, MYC, CCNB, CREB, YBX1, CDKN1B, EST2, DUSP4, STAT5, AURKB, STAT3, PIGR,	
BRCA_signaling			H2AFX, FANCD2, XPC, ESR1, CCND1, BARD1, CDKN1B, ETS2, DDB2, SFN	ATM_signaling
RAR_signaling		RARA, RARB, RARG, RXRA-C	MYH6, THPO, MGP, BGLAP, MAOB, GNRH2, SP1, ADH1B	
SWH_signaling		FAT1-4	TGFA, TP73, JUN, MB, NF2, MST, LATS, MYH7, RAC1, GH1, HSD3B1, HRAS, SP1, MOBKL1B, IFITM3	ATF_signaling, Cell_Cycle, MAPK_signaling, TGF_beta
ERK5_signaling			FOSL1, CREB, MEF2B, FOS, PTPRH, MYC, SGK1	MAPK_signaling, NGF_signaling
Pathway	**Ligand**	**Receptor**	**Downstream Targets**	**Crosstalk Pathways**

C.4 THE SIMULATION DATA OF 100 SPECIFIC MIRNA INHIBITORS

Attached in CD

C.5 THE SIMULATION DATA OF AKT AND MEK INHIBITOR

Attached in CD

C.6 THE SIMULATION DATA OF SINGLE LIGAND-ACTIVATIONS IN THE HUMAN SIGNALING MODEL.

Attached in CD

C.7 THE SIMULATION DATA OF 100 SPECIFIC MIRNA INHIBITORS

Attached in CD

C.8 THE COMPARISON RESULT OF THE SIMULATION DATA WITH THE DATA FROM NAGARAJ'S STUDY

C.8.A The predicted proteome data and the experimental proteome data from the HeLa cell (attached in CD).

C.8.B The transcriptome data and proteome data derived from this study (attached in CD).

(only take into consideration genes and proteins that are defined in the human signaling model)

C.9 THE DRUG RESPONSE PREDICTION DATA

Attached in CD