

LURKING PATHWAY PREDICTION AND PATHWAY ODE MODEL  
DYNAMIC ANALYSIS

A Dissertation

by

RENGJING ZHANG

Submitted to the Office of Graduate and Professional Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of  
MASTER OF SCIENCE

Chair of Committee,	Zixiang Xiong
Committee Members,	Tie Liu
	Jim Ji
	Anxiao Jiang
Head of Department,	Chanan Singh

December 2013

Major Subject: Electrical and Computer Engineering

Copyright 2013 Rengjing Zhang

## ABSTRACT

Signaling pathway analysis is so important to study the causes of diseases and the treatment of drugs. Finding the lurking pathway from ligand to signature is a significant issue in studying the mechanism of how the cell response to the stimulation signal. However, recent literature based pathway analysis methods can only tell about highly differentially expressed pathways related to the experiment data, which may tell nothing about our interested specific ligand and signature.

In this paper, we designed an approach to successfully detect the most reliable pathways for specific ligand and signature by solving multi-objective optimization problem on the bridge connecting two signaling pathways where the ligand and signature locate. The pathway bridge consisted of enriched looping patterns refined the complicated entire protein interactions network with 39031 links, which made the approach time-saving. The approach was further applied to study the modulator mechanism of the signal molecule, receptor, intermediate transfer proteins, transcription factor, and signature.

With preliminary studied pathways, we then employed Ordinary Differential Equations(ODE) to modeling and dynamic analysis the signaling transduction. The biological reactions were represented in terms of differential equations, and the solutions to the group of equations were further be optimized to fit the RPPA experiment data. In order to find the potential signaling paths in specific disease and discovery the best therapy, coefficient variation analysis, system robustness study and system outcomes changes to perturbations were also utilized.

Our approach successfully predicted the lurking pathway for the signal molecule

*TGFβ1* and the nova protein *OCIAD2* in cancer microenvironment: *TGFβ1 – TGFβR1 – SMAD2/3 – SMAD4/AR – OCIAD2*, and this result was verified by literature. Better than recent pathway analysis tool, our predicted pathway also took care of significant but relatively less regulated proteins in the transduction process. And by modeling the CCL2 pathway in MTB infected cells, *JNK*, *cMYC* and *PLC* showed as the most significant modules. Hence, the drug treatments inhibiting *JNK*, *cMYC* and *PLC* would effectively obstruct the increasing of MMPs and further prevent the Mtb infections.

## DEDICATION

This Thesis is dedicated to my parents with love.

Without Dad and Mom's unreserved support, I would not have the chance to see the beautiful and different world. Thanks for always being there for me.

## ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Guangxu Jing for his early work on pathway modeling and the discussion with the members of Bioinformatics and Systems Biology, Dr. Chen Zhao for his experiment validation of the study on OCIAD2 and biological introduction/discussion supports, and Dr. Priyanka Kachroo for her RPPA experiments and background analysis on Mtb stimulated cells.

## NOMENCLATURE

pbMOO	Pathway Bridge Based Multi-objective Optimization
MC	Motifs Cluster
TF	Transcription Factor
RPPA	Reverse Phase Protein Array
HPRD	Human Protein Reference Database
TGFB1	Transforming Growth Factor Beta 1
OCIAD2	Ovarian Carcinoma Immunoreactive Antigen-Like Protein Domain Containing 2
Mtb	Tuberculosis
MMP	Matrix Metalloproteinase
MCP	Monocyte Chemotactic protein 1
CCL2	Chemokin Ligand 2

## TABLE OF CONTENTS

	Page
ABSTRACT . . . . .	ii
DEDICATION . . . . .	iv
ACKNOWLEDGEMENTS . . . . .	v
NOMENCLATURE . . . . .	vi
TABLE OF CONTENTS . . . . .	vii
LIST OF FIGURES . . . . .	ix
LIST OF TABLES . . . . .	x
1. INTRODUCTION . . . . .	1
2. METHODS . . . . .	4
2.1 Cell Lines and Drug Treatment . . . . .	6
2.2 RNA Extraction and Quantitative Real-time PCR . . . . .	6
2.3 Motif Detection of Protein-Protein Interaction Network . . . . .	6
2.4 Motifs Clustering and Enrichment in Cancer Related Signaling Pathways . . . . .	9
2.5 Ill-defined Protein Pathways Prediction . . . . .	11
2.6 Interacted Pairs Inference for Protein without PPI . . . . .	13
2.7 Multiple Micro-array Data Based Differential Expression Score . . . . .	14
3. RESULTS . . . . .	15
3.1 Dysregulation of OCIAD2 in Different Cancers and Its Induction by TGF $\beta$ in HCC and PC Cells . . . . .	15
3.2 Potential Protein Pathway Prediction . . . . .	17
3.2.1 Pathways in Prostate Cancer . . . . .	17
3.2.2 Pathways in Liver Cancer . . . . .	18
3.3 Modular Mechanism Exploration . . . . .	18
3.3.1 Speculation of Human TF Enrolled in TGF $\beta$ Signal . . . . .	21
3.3.2 Feasible TF of OCIAD2 in Cancer Cell Line . . . . .	23
3.3.3 Mechanism of TGF $\beta$ Induced OCIAD2's Expression . . . . .	24
4. MTB INDUCED CCL2 PATHWAY ODE MODEL DYNAMIC ANALYSIS . . . . .	27

4.1	Background Introduction . . . . .	27
4.1.1	Motivation and Experiments Design . . . . .	27
4.1.2	Experimental Result Observations . . . . .	28
4.2	ODE Model for CCL2 Pathway . . . . .	38
4.2.1	ODE According to the Law of Mass Action . . . . .	39
4.2.2	Parameters Initialization . . . . .	42
4.2.3	Solutions Optimization and Fitness . . . . .	42
4.3	Sensitivity Analysis of CCL2 Signaling Pathway . . . . .	45
4.4	Model Robustness Analysis . . . . .	47
4.5	System Perturbation and Drug Treatment Prediction . . . . .	49
5.	SUMMARY . . . . .	50
	REFERENCES . . . . .	54



## LIST OF FIGURES

FIGURE	Page
2.1 Pathway Prediction from Ligand to Signature by pbMOO Approach .	5
2.2 Four-vertex Motif Detections of FANMOD Outputs . . . . .	8
2.3 Looping Motifs Enriched in Cancer-related Signaling Pathways . . . .	9
2.4 Loop-shaped Motifs And Motifs Cluster Toy Model . . . . .	10
2.5 P-value Calculation Examples . . . . .	12
3.1 Expression of OCIAD2 and its induction by $TGF\beta$ . . . . .	16
3.2 One Predicted Mechanism for OCIAD2 Expression in miR205 Treatment	20
3.3 Transcription Factor Detection for OCIAD2 . . . . .	24
3.4 $TGF\beta$ Induced OCIAD2s Expression Mechanism . . . . .	26
4.1 Gene Expression in Different Groups of Cell Lines . . . . .	29
4.2 Proteins in RPPA List Analyzed by IPA . . . . .	31
4.4 Possible Signaling Molecules via PAR-1 influence MMP-1 by IPA . .	34
4.5 Stimulated Cells Responses to PAR-1 Inhibitor Addition by IPA . . .	36
4.6 CCL2 Pathway in Mtb Simulated Cell . . . . .	38
4.9 ODE Solutions vs. Experimental Data . . . . .	44
4.10 CCL2 Pathway ODE Model Parameters Fitness . . . . .	45
4.11 Output Changes When Parameter Increased 100 Times . . . . .	46
4.12 Coefficient of Variation Analysis of Parameters . . . . .	47
4.13 System Model Robustness Analysis . . . . .	48
4.14 System Outcome Changes by Perturbation . . . . .	49

## LIST OF TABLES

TABLE	Page
3.1 Possible Protein Paths from miR205 to OCIAD2 in Prostate Cancer .	19
3.2 OCIAD2's Transcription Factors Detection in TGF $\beta$ Signal . . . . .	22

## 1. INTRODUCTION

Tumor microenvironment has been largely studied as a dynamic system to define the behaviors of cancer. This system is orchestrated by cytokines, growth factors, inflammatory cells, cancer cells, stroma as well as the extracellular matrix [1]. Tumor-associated fibroblasts (TAFs) are major elements of tumor stroma and have been shown to play an important role in tumor growth and progression. Epithelial-to-mesenchymal transition (EMT) is a major source of TAFs. In tissue fibrosis it is well-established that epithelial cells contribute to the accumulation of fibroblasts by undergoing EMT in response to stimuli from the microenvironment [2].  $TGF\beta$  remains among the key factors responsible for the recruitment of Tumor Associated Fibroblasts (TAFs) and induction of EMT. TAFs, meanwhile, strongly contribute to the production and activation of  $TGF\beta$  in the activated stroma and thereby generate the autocrine feed-forward loop that is characteristic for persisting fibroblasts activities [3]. However, the exact regulation between  $TGF\beta$  signals and TAFs in tumor microenvironment is yet to be completely understood.

OCIAD2 was originally immunoscreened from ascites of a patient with ovarian cancer and found to be an immunoreactive antigen [4]. However, the function of OCIAD2 protein, involved pathways and molecular mechanisms has never been reported. Based on data mining and our preliminary data, we hypothesize that human OCIAD2 represent a potential tumor suppressor gene in some tumor types and its dysregulation involved in  $TGF\beta$  regulated signaling in tumor microenvironment with following reasons: 1) High-throughput profiling data and public database analyses showed that OCIAD2 is frequently methylated and/or downregulated in some kinds of cancers [5], [6], and [7], 2) GEO database revealed that the expressions

of OCIAD2 are induced by TGF $\beta$  signal in pancreatic (GSE23952), lung adenocarcinoma (GSE17708) and ovarian cancer cells (GSE6653), 3) Moreover, a computational analysis with TCGA database revealed that methylation site of OCIAD2 is top-ranked in ovarian Metastasis-Associated Fibroblasts (MAFs) signature [8]. These evidences indicated a potential biological milieu of OCIAD2. We hereby speculate that down-regulated OCIAD2 expression in tumor microenvironment facilitates deregulated TGF $\beta$  signaling. As a consequence of these changes, tumor cells escape immunesurveillance and exaggerate tumor progression and metastatic spread.

To predict molecular network of OCIAD2 in TGF $\beta$  regulated tumor microenvironment, a nova pathway analysis with bioinformatics approaches have been developed. Current signal analysis methods typically have three steps: build literature based preliminary signaling pathways model; generate gene expression experimental data; detect the shortest path as the specific signal and verify biological meaning. Pathways consist of highest differentially expressed genes and reported interactions would be shown as the results in this kind of pathway study. However, not all the targets or receptors of ligands are with top expression changes, i.e., TGF $\beta$  regulates numerous other growth factors positively and negatively, some of which are not the most obviously changed ones but still response to the stimulation of TGF $\beta$ . Moreover, new genes with seldom previous studies, such as OCIAD2, is hardly included in any pathways because the lack of known interactions with other proteins. To study the mechanism of OCIAD2 changes induced by TGF $\beta$  stimulation in cancer cell lines, a novel approach to inferring the signaling paths based on the pathway bridges between TGF $\beta$  and target gene OCIAD2 using the Multi-objective Optimization Approach, named pbMOO, was developed. Pathway Bridge was defined as a subset of protein interactions network that was consisted by clustering loop

motifs with extremely high frequency occurring in cancer related processes than by chance. All four-vertex motifs, among which the triangle and rectangle were shown with significantly higher occurring than randomized ones, were detected from a network generated from HPRD database with 12794 proteins and 39031 interactions. Rather than traversing the entire protein interaction network with enormous nodes and edges, all the loop motifs were clustered as a Pathway Bridge between TGF $\beta$  signaling pathway and cancer signaling pathway. Relatively, the time saving approach returned to highly reliable protein paths only by searching connecting nodes on the bridge. Moreover, motifs on the bridge were concentrated on cancer related processes, which guaranteed the nodes chosen for the path are specified for cancer microenvironment. Then, the cost of a protein path was defined by gathering up the cost of each edge, which is the p-value sum of two interacted protein nodes. According to the property of p-value, the path cost is the probability of obtaining a path whose cost is no more than the one that was actually observed so that the less cost, the more reliable the path is.

Applied pbMOO approach and treated the most tightly correlated genes with OCIAD2 as its potential interacts ones on Transcriptional Regulatory Element Database [9] with 177 transcription factors of homo sapiens, androgen receptor (AR) was discovered as the most credible transcription factor of OCIAD2. Applied the approach on GSE42357 and GDS3634 expression data from NCBI, the paths with the lowest cost were picked out as the responsible possible molecular mechanisms between TGF $\beta$  and OCIAD2 in hepatocellular carcinoma (HCC) samples, and prostate cancer cell lines. Verified the biological meaning of the low-cost paths and finally the signal TGF $\beta$ 1- TGF $\beta$ R1- SMAD2/3- SMAD4- AR- OCIAD2 was discovered and explained TGF $\beta$ s stimulation on OCIAD2 expression in cancer.

## 2. METHODS

In order to see how the signature gene be enrolled in the ligand stimulation signal, a pathway bridge based multi-objective optimization approach (pbMOO) was designed and summarized as Figure 2.1. Three kinds of data were chosen as the initial input data: protein interactions Figure 2.1A from HPRD [10] were applied for building the entire protein-protein interactions (PPI) network; signaling pathways Figure 2.1B from both KEGG [11] and IPA [12] were selected as the background pathway library based on which the pathway bridges were constructed; groups of microarray data Figure 2.1C were used for solving multi-objective optimization problem and presenting the correlation among genes. Calculated by FANMOD [13], loop motifs Figure 2.1A1, higher frequency occurring sub network of entire PPI network, were shown enriched in cancer and related pathways. Searching on the pathway bridge Figure 2.1B2, which was defined as a set of loop motif clusters Figure 2.1A2 connecting ligand and signature genes, a multi-objective optimization problem was solved by finding the pathways with the lowest path cost that was assigned by gene expression p-value. When multiple experimental gene expression data were used, the cost of each path was then defined as the summation of average p-value of connected genes in the optimization problem. Then the modular study Figure 2.1C1 was applied on the calculated results of the optimization problem and completed signals, which began with ligand and its receptor, passing through transduction proteins and targeting on transcription factor and finally the signature, were output as the most reliable predicted pathways Figure 2.1C2 explaining how the ligand changes affected the signature.

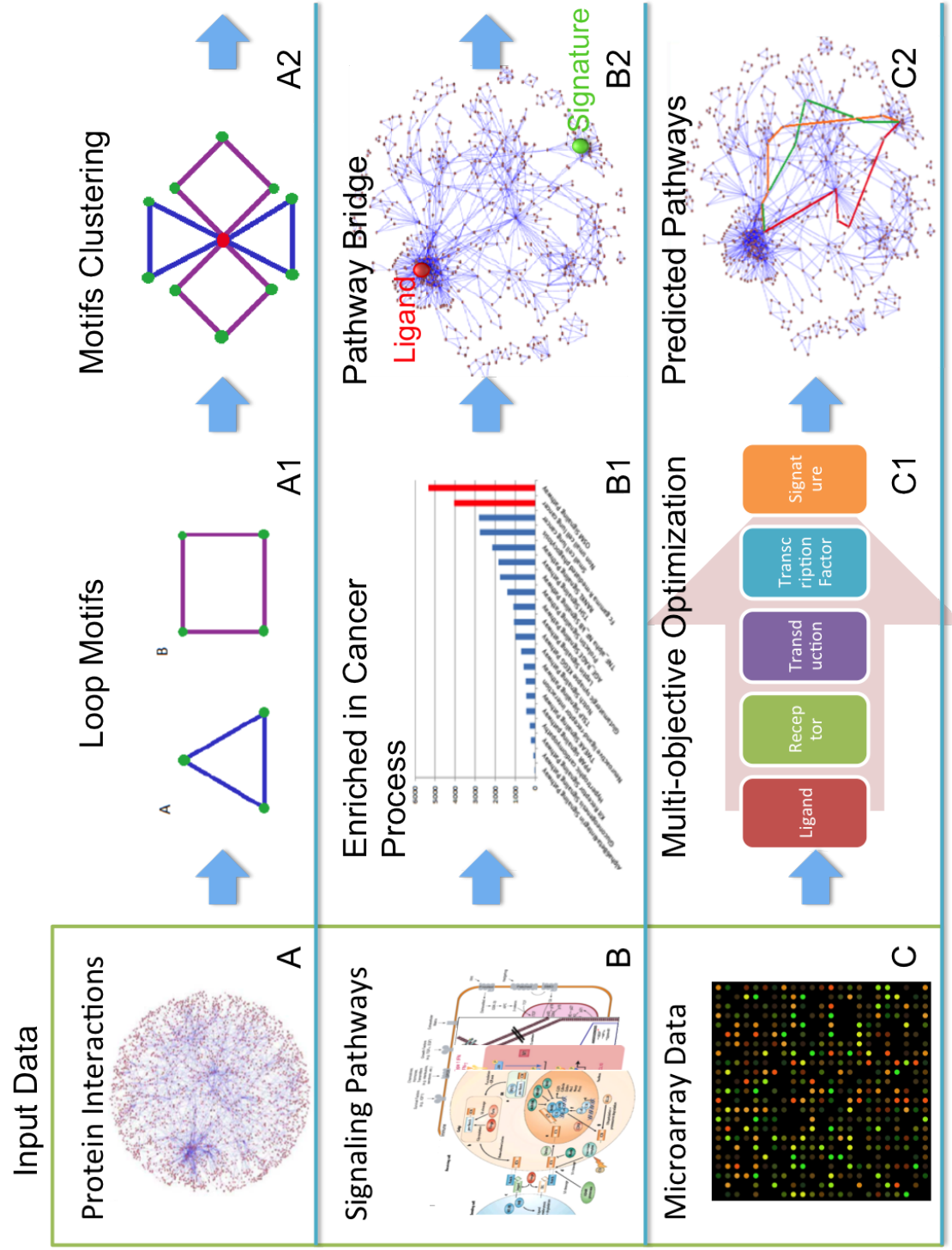


Figure 2.1: Pathway Prediction from Ligand to Signature by pbMOO Approach

## 2.1 Cell Lines and Drug Treatment

Hep-3B and Du-145 were obtained from American Type Culture Collection. All cell lines were cultured in DMEM with 10% fetal bovine serum (FBS) and antibiotics. TGF $\beta$ 1 (R&D Systems, Minneapolis, MN) were applied at concentrations of 5ng/ml. TGF $\beta$ R inhibitor LY2109761 were purchased from Selleck Chemicals LLC (Houston, TX), using at 2M. For the drug treatment, human liver and prostate cancer cell lines, Hep-3B and Du-145, were treated with 5ng/ml TGF $\beta$ 1, 2M LY2109761 and combination for 24 hours in serum free media, and OCIAD2 mRNA levels were determined by quantitative real-time RT-PCR analysis.

## 2.2 RNA Extraction and Quantitative Real-time PCR

Total RNAs were isolated from tumor cells using TRIZOL reagent (Life Technologies, USA) following the manufacturers recommendations. RNA concentration and purity were determined by measuring absorbance at 260 and 280 nm with a NanoDrop<sup>TM</sup> 1000 Spectrophotometer (Thermo Scientific, USA). cDNA synthesis was performed with Superscript III reverse transcriptase kit (Life Technologies, USA). Quantitative real-time reverse-transcription polymerase chain reaction (RT-PCR) was performed using an Applied Biosystems 7300 Sequence detection system (Applied Biosystems, Life Technologies, USA). The primer set of OCIAD2 are described below: 5'-TGCGAGAATGTCAGGAAGAA-3' and 5'-AAATCCCAAGAGACCAGCAA-3'.

## 2.3 Motif Detection of Protein-Protein Interaction Network

”Network Motifs” [14] are interconnected patterns (sub graphs) with significantly higher occurring in complicated networks than in randomized ones. Loop-structural motifs have been proofed to be enriched in a protein-protein interaction (PPI) net-



work generated from PPI database, i.e. HPRD. As a literature-collected public database, HPRD has 12794 proteins and 39031 pairs of interactions for 9605 of them. The sufficient data capacity helps a lot on unclear reciprocities prediction.

In this paper, outstanding tool Fast Network Motif Detection (FANMOD) [13] was applied to census four-vertex sub graphs in undirected PPI network by using the Randomized Enumeration (RAND-ESU) algorithm.

Motifs detections result from PPI network were shown as Figure 2.2. Motifs ID lied in column one and adjacency matrix presented in the second column; Frequency was the probability of each motif in original PPI network and Mean-Freq was of the motif occurred in random networks; the standard deviation from the mean frequency was listed in the fifth column; Z-score meant the value of the difference of two frequencies divided by the standard deviation; and p-value was the difference of motif number between random networks and original one then divided by the total number of random networks. For instant, ID 8598 in Figure 2.2 had relative higher occurring frequency in random networks than the original one, thus its Z-Score was negative and p-Value was large, indicated that the chain-looked structure was really normal among the entire PPI network; on the contrary, ID 31710 was more special in the original network structure than random showing that the combination of triangle and square sub network was enriched among PPI network. Similarly, motifs ID 4382, recurring much more often than random chosen sub network, are with negative Z-Score and large p-Value; ID 13278, ID 4958 and ID 27030 motif structures are also enriched from randomized network and obtain positive Z-Score and relative smaller p-Value. The outcome suggests loop-structural motifs, i.e. shapes like triangle, spoon and square, are special patterns with high occurrences in protein interactions network.

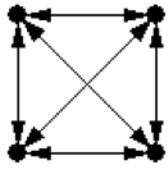
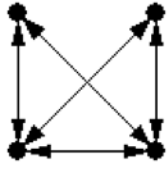
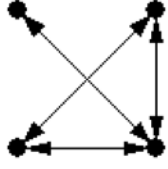
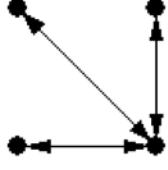
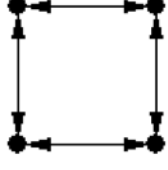
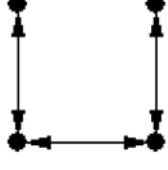
ID	Adj	Frequency [Original]	Mean-Freq [Random]	Standard-Dev [Random]	Z-Score	p-Value
31710		0.017858%	1.1634e-007%	1.6208e-008	11018	0
13278		0.27748%	6.9157e-005%	4.1539e-006	667.84	0
4958		4.7955%	0.0085102%	0.00042599	112.37	0
4382		50.482%	54.627%	0.0017771	-23.325	1
27030		0.31215%	0.24795%	7.7688e-005	8.264	0
8598		44.115%	45.117%	0.0016806	-5.9588	1

Figure 2.2: Four-vertex Motif Detections of FANMOD Outputs

20 signaling pathways derived from KEGG [11] were analyzed for motifs distribu-

tion. Results in Figure 2.3 induced that proteins on looping motifs are mainly from cancer and correlated signal pathways, in other word, motifs with loop structure are enriched in 14 type of carcinomatosis and related signaling pathways, such as cell cycle signaling pathway, immune system signaling pathways and etc.

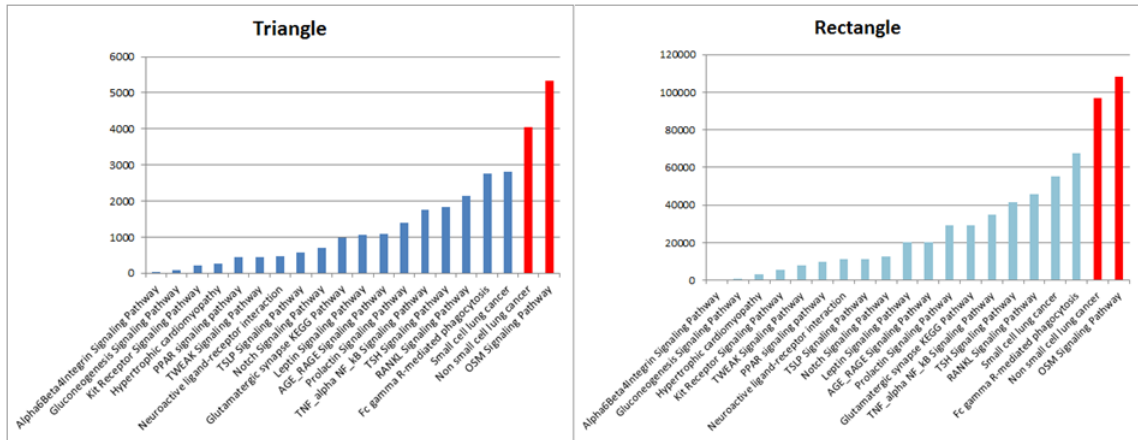


Figure 2.3: Looping Motifs Enriched in Cancer-related Signaling Pathways

#### 2.4 Motifs Clustering and Enrichment in Cancer Related Signaling Pathways

Loop-shaped motifs with no more than four vertex have the only two specific possibilities —triangle and square. Motifs Cluster (MC) is defined as converged cyclic motifs that sharing at least one protein. The common protein is called Center Point (CP), which is the identifier for distinguishing different motifs clusters. The toy model shows in Figure 2.4.

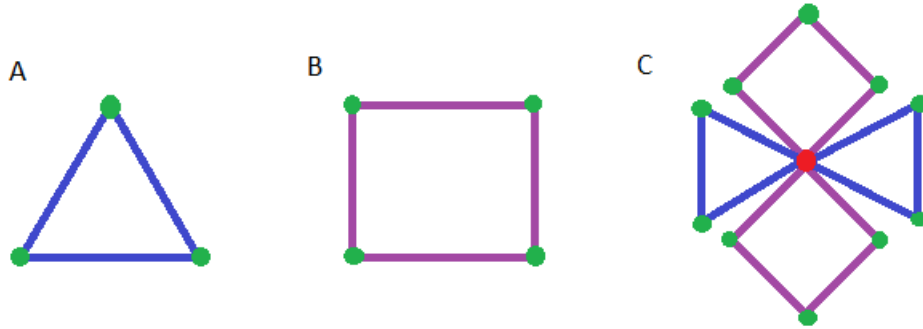


Figure 2.4: Loop-shaped Motifs And Motifs Cluster Toy Model

Since looping motifs were proved that occurring much more often in cancer and related signaling pathways, they can be treated as a bridge to link up cancer and its kinship pathways, which would provide a characteristic group of candidate protein interactions for future unclear links forecast.

Let  $P_1$  be a chosen cancer signaling pathway, and  $P_2$  be a cancer-related signaling pathway.  $\{MC_1^{P_1P_2}, MC_2^{P_1P_2}, \dots, MC_n^{P_1P_2}\}$  are the total  $n$  motif clusters between  $P_1$  and  $P_2$ , thus  $|MC^{P_1P_2}| = n$  by virtue of the number of identifiers  $|CP^{P_1P_2}| = n$ , where  $|\cdot|$  denotes the number of elements in a set. In order to evaluate the enrichment of MCs lying between  $P_1$  and  $P_2$ ,  $p$ -value was introduced as the probability of obtaining a larger number of MCs for a pair of randomly chosen protein sets, keeping the same sizes with  $P_1$  and  $P_2$  and the capacity of intersection, than for  $P_1$  and  $P_2$ .

$$p = \text{prob}\{n' > n \mid n' = |MC^{S_1S_2}|, n = |MC^{P_1P_2}|\} \quad (2.1)$$

where  $S_1$  and  $S_2$  are random proteins sets picked out from entire proteins of HPRD database with the same size as  $P_1$  and  $P_2$ , i.e  $|S_1| = |P_1|$  and  $|S_2| = |P_2|$ ,

and satisfy  $|S_1 \cap S_2| = |P_1 \cap P_2|$ . Repeating the sampling for 1000 times, a random distribution  $f$  for the 1000 numbers of MCs can be generated. The complementary set of cumulative probability density function  $F(|MC^{P_1 P_2}|)$  interprets the chances that a stochastic pair of protein sets has quantity of MCs being the bridges between them rather than the two chosen pathways, which is indeed motif clusters'  $p$ -value.

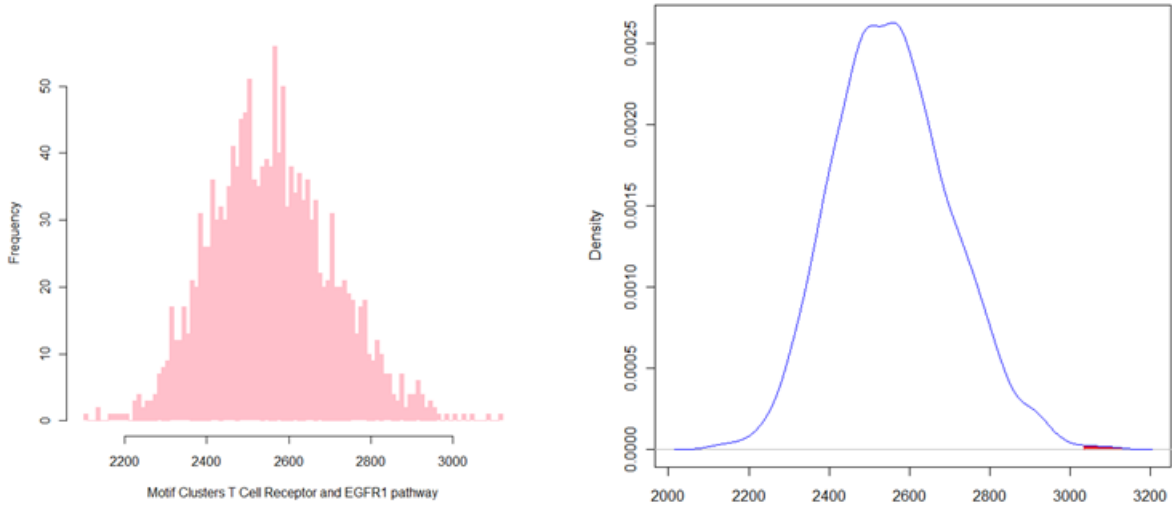
$$p = F'(|MC^{P_1 P_2}|) = 1 - F(|MC^{P_1 P_2}|) \quad (2.2)$$

MCs connecting  $P_1$  and  $P_2$  are enriched if the  $p$ -value is tiny, indicating that the bridging MCs linking up two protein sets are the main substructure of cancer signaling path  $P_1$  and cancer enrolled signaling path  $P_2$ . Comparing with searching the enormous and complex integrated PPI network, the enriched MCs bridge efficiently limits and specilized the traversing range for forecasting uncertain protein paths, which increases the calculation speed thoroughly.

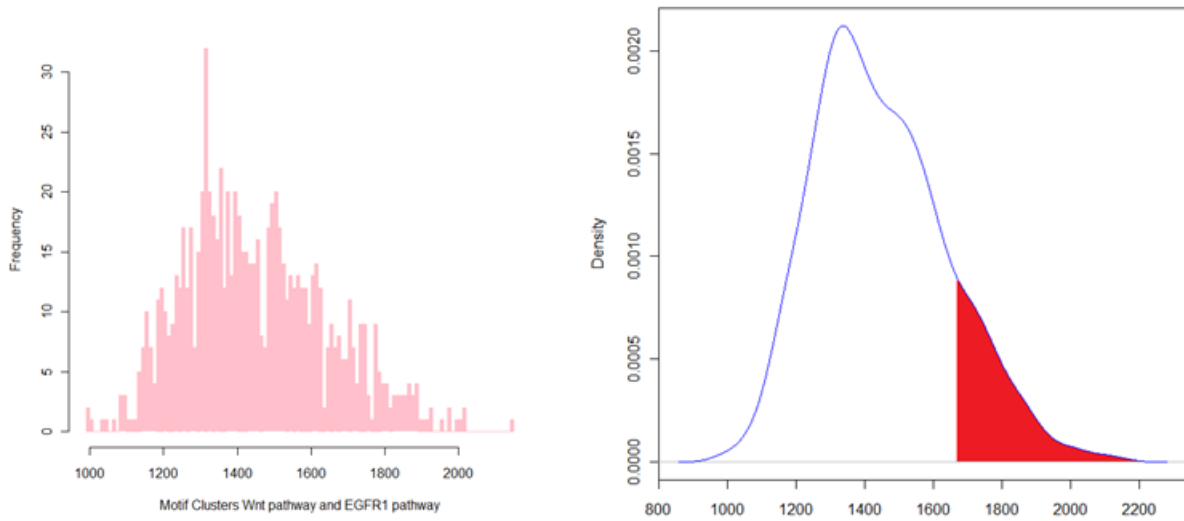
MCs'  $p$ -value for combinations of different types and subtypes of carcinomatosis and involved signaling pathways were calculated and exemplified in Figure 2.5. MCs bridges that have  $p$ -value less than 0.01 were chosen to be the candidate subnetwork, from which ill-defined protein pathways would be selected.

## 2.5 Ill-defined Protein Pathways Prediction

If changing condition of a protein  $A$  results in the upregulated or downregulated protein  $B$ , while  $A$  and  $B$  have neither direct interaction with each other nor indirect upstream and downstream relationship on any authentic signaling pathway, the underlying protein pathways for them could be detected on those MCs enriched pathway bridges. An optimization model, which is described as following, was employed



A. Enriched Motif Clusters between EGFR1 and T Cell Receptor Pathways



B. Less Enriched Motif Clusters between EGFR1 and Wnt Pathways

Figure 2.5: P-value Calculation Examples

to acquire high-confidence potential protein pathways.

$$f(x) = \arg \min_{\vec{x} \in \{MC_1^{P_1 P_2}, MC_2^{P_1 P_2}, \dots, MC_n^{P_1 P_2}\}} \sum_{i=1}^N x_i \cdot DES_i + \lambda \sum_{i=1}^N x_i$$

$$\text{s.t.} \quad \begin{cases} 2 \leq \sum_{i=1}^N x_i \leq 7, \\ i = 1, 2, \dots, N, \\ \lambda \in |\mathfrak{R}|. \end{cases} \quad (2.3)$$

where  $N$  is the total number of proteins pertaining to MCs bridge for the selected pair  $P_1$  and  $P_2$ .  $\vec{x} = \{x_1, x_2, \dots, x_N\}$  is protein path vector implying which element was contributed to the path—if protein  $i$  was taken count into the lurking protein path, then  $x_i = 1$ , otherwise  $x_i = 0$ . Differential expression score (DES) was defined by each gene’s  $p$ -value from student  $t$ -test for gene expression experiment data in two conditions. The larger  $p$ -value, the larger DES, the less reliable the data. Thus, minimizing the first part of the object function  $\sum_{i=1}^N x_i \cdot DES_i$  could ensure the maximization of the reliability of the predicted protein paths. The length of protein pathway, i.e.  $\sum_{i=1}^N x_i$ , is an integer in the range of  $[2, 7]$ , which was decided by the fact that MCs were composed of looping structures up to 4 vertex. At this point, the latter part of the object function took the responsibility of controlling the length of analyzed protein pathway with the aid of distinct settings of nonnegative parameter  $\lambda$ . Large  $\lambda$  made for limited proteins and short connections, and optimization result was free to rope in proteins when  $\lambda = 0$ .

## 2.6 Interacted Pairs Inference for Protein without PPI

For those proteins have no canonical protein interaction supported, gene expression data conduced to providing indistinct mutual effects and pointing out candidate proteins with which the separated proteins were closely bound up by the correlation between pairs of genes.

## 2.7 Multiple Micro-array Data Based Differential Expression Score

As the matter of fact that the p-value of experimental gene expression data may vary a lot by different experiment designs and operators, a good inferred protein path is the one who gets rid of the destabilizing factors. Thus, multiple micro-array data sets were employed here for error deduction.



### 3. RESULTS

#### 3.1 Dysregulation of OCIAD2 in Different Cancers and Its Induction by TGF $\beta$ in HCC and PC Cells

To determine the OCIAD2 expression, different available microarray studies were analyzed by the Oncomine database and GEO gene microarray data analysis tools. A significant down-regulation of OCIAD2 mRNA expression was found in liver cancer and gastric stroma carcinoma tissues. ( $P < .001$  in both cases) (Fig 3.1; Left) based on Oncomine database analyses. With a GEO database, OCIAD2 expression in metastatic prostate tissues, but not in primary tumor tissues, is clearly lower than normal prostate gland. ( $P = 0.009$ ) (Fig 3.1; Right; GSE6919). Frequently down-regulated OCIAD2 expressions are also observed in CLL and malignant pleural mesothelioma [5] and [6]. In glioblastoma, OCIAD2 expression is being silenced via DNA methylation mechanism [7]. With the suggestion of the fact that OCIAD2 substantially unregulated in TGF $\beta$ 1 treated pancreatic (GDS4106), lung (GSE17708) and ovarian (GSE6653) cancer cells, we tested the possibility of OCIAD2 expression induced by TGF $\beta$  in HCC and PC cells. Human HCC and PC cell lines, Hep-3B and Du-145, were treated with 5ng/ml TGF $\beta$ 1, 2M LY2109761 and combination for 24 hr in serum free media, and OCIAD2 mRNA levels were determined by quantitative real-time RT-PCR analysis. OCIAD2 mRNA has increased 2.5 and 4.6 fold in Hep-3B and Du-145 cells by TGF $\beta$ 1 treatment respectively. This induction was totally suppressed by TGF $\beta$ 1 receptor inhibitor LY2109761 (Fig 3.1B).

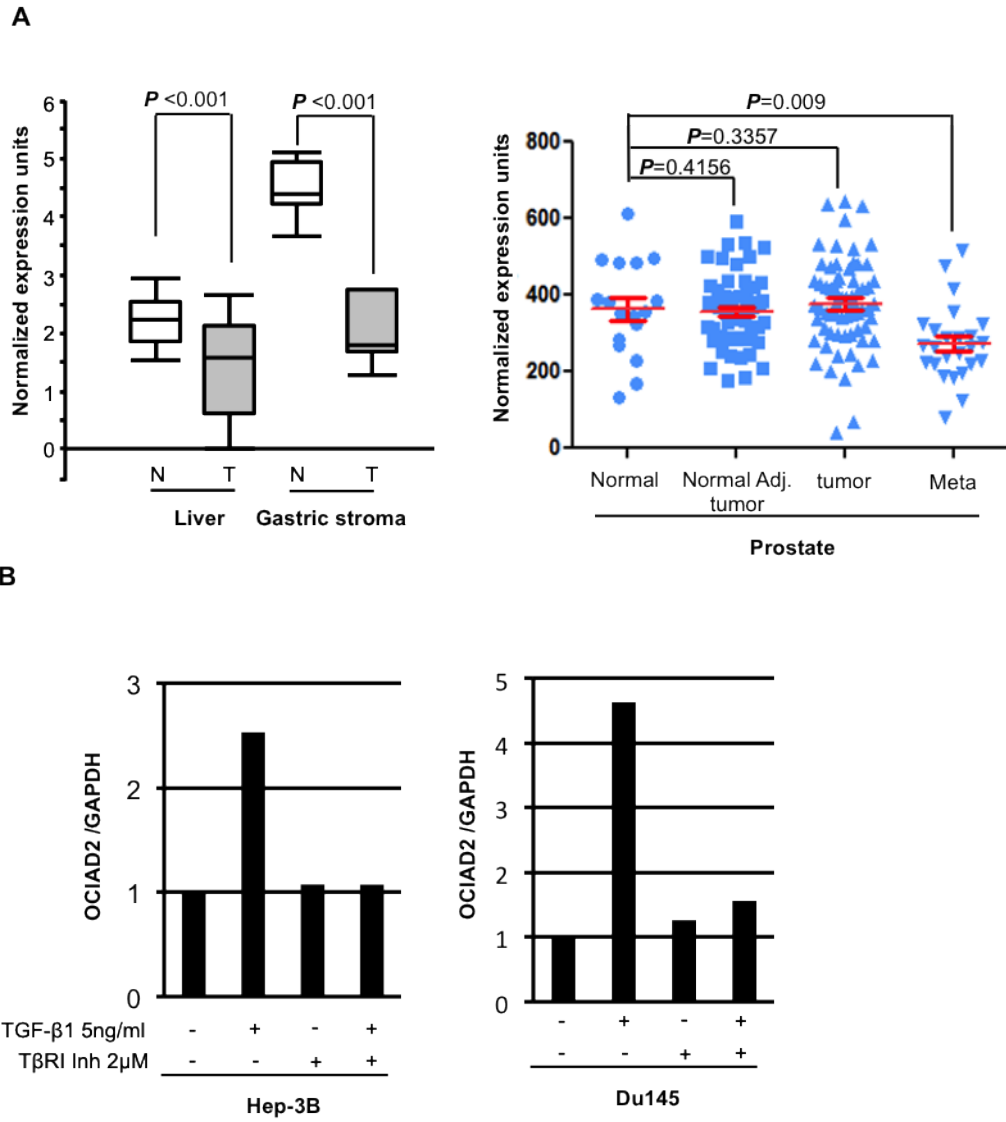


Figure 3.1: Expression of OCIAD2 and its induction by TGF $\beta$

## 3.2 Potential Protein Pathway Prediction

Remarkable gene array profiles from GEO database indicated the expression of OCIAD2 in several kinds of cancer, i.e., GDS3634 showed that OCIAD2 obviously unregulated in prostate cancer cell line transfected with 20nM miRNA Presursor Molecules miR-205. MiR-205 is selectively down-regulated in metastasic breast and prostate cancer and suppresses metastatic spread of a human breast cancer xenograft in nude mice. In addition to its function in the regulation of EMT, the loss of miR-205 in prostate cancer also reduced some tumor suppressor genes' expression.

Mesenchymal stem cells (MSC), like other bone marrow-resident cells, have the capacity to differentiate into fibroblasts-like cells that have been variably referred to as: myofibroblasts, tumor-associated fibroblasts (TAF), fibrocytes or pericytes within the tumor microenvironment [15]. Therefore, *pbMOO* approach was applied on both prostate cancer cell line GDS3634 and liver cancer associated mesenchymal stem cells GSE42357 gene expression data to study the possible molecular path involved in OCIAD2 by TGF $\beta$  stimulation.

### 3.2.1 Pathways in Prostate Cancer

Based on experimental data GDS3634, miR-205 expression effect on prostate cancer cell line, from NCBI public database browser, p-value of 8 samples Student T-test Micro-array data was applied as the link cost for predicted paths. Searched on the pathway bridge that has been enriched through prostate cancer pathways, the top 10 out of 88 forecasted paths were picked for further biological meaning verification (Table3.1). One reasonable forecast shown in Figure 3.2 was miR205- PRKCE- CNA13- CDH1- PTPN14- OCIAD2. High correlated genes with OCIAD2 were distinguished as dashed lines (red for positive correlation and green for negative

correlation). PRKCE, as one of the six target genes of miR205, was marked as the green box. Among the pathway bridge that consisted of protein nodes (dashed circles) and proteins interactions (green lines), a shortest protein path with least DES cost was emphasized by red lines.

Suggested by the significant association between TGF $\beta$ 1 and CDH1, *pbMOO* Approach was employed again aiming at finding out how TGF $\beta$  affects OCIAD2 across CDH1 in prostate cancer. Interestingly, "TGF $\beta$ 1 influenced CDH1 across SMADs" was observed after filtering the predicted protein paths and verified by [16].

### 3.2.2 Pathways in Liver Cancer

By the analysis of GSE42357 gene expression data, genes like C5, AG7, SDC2, FHL2, and etc., have been suggested to be the candidates of OCIAD2 by their tight correlations with it. Significantly, those candidate genes all play important roles in cancer related processes, for instant, C5 takes the responsibility in inflammatory and cell killing processes [17] and FHL2 acts as both tumor-promoter or tumor-suppressor depending on different types of cancer [18]. The calculated results of the approach were paths with credibility cost, i.e. TGF $\beta$ 1- TGF $\beta$ R1- CLU- C7- C5- OCIAD2 (cost 0.233181). This pathway was fully explained by the fact CLU is a novel modulator of TGF $\beta$ 1 signaling pathway by regulating Smad2/3 proteins [19] and the well-known protein interactions CLU- C7 and C7- C5.

## 3.3 Modular Mechanism Exploration

If a signaling transmission process, from extracellular through cytoplasm to nucleus, results in upregulation or downregulation of genes in the cell, then transcription

---

<sup>0</sup>The less Cost, the more credible the pathway is.

Table 3.1: Possible Protein Paths from miR205 to OCIAD2 in Prostate Cancer

No.								Cost <sup>1</sup>
1	PTPN14	CDH1	G-113	PRKCE				0.152141
2	PTPN14	TJP1	GRIN1	PRKCE				0.165309
3	PTPN14	YWHAG	TIAM1	PRKCE				0.182662
4	PTPN14	YWHAG	SRC	PRKCE				0.403788
5	PTPN14	JUP	SRC	PRKCE				0.505998
6	TAL2	MAPK3	PRKCE					0.513332
7	CAPN2	NFKBIA	CUL1	E2F1				0.513761
8	PKP3	DSC2	JUP	SRC	PRKCE			0.571772
9	CAPN2	NMT1	LYN	PLCG1	GRIN1	PRKCE		0.581042
10	CAPN2	NFKBIA	SRC	PRKCE				0.671004
11	PPP2R1B	RELA	COMMD1	CUL1	E2F1			0.680008
12	MYD88	IRAK2	HRAS	GRIN1	PRKCE			0.755849
13	PPP2R1B	RELA	HDAC3	CCND1	NPDC1	E2F1		0.761167
14	CAPN2	CDK5R1	CHN1					0.775821
15	PPP2R1B	SET	KLF5	PRKCD	YWHAG	MDM4	E2F1	0.786134
16	PPP2R1B	PPP2CA	CCNG1	CDK5	CABLES1	CC-11	E2F1	0.80783
17	PPP2R1B	RELA	PARP1	E2F1				0.817619
18	MYD88	IL1RAP	RAC1	CHN1				0.833815
19	MYD88	IL1RAP	RAC1	TIAM1	PRKCE			0.871396
20	PPP2R1B	SET	SGOL1	XRCC6	CC-11	E2F1		0.903496
21	PPP2R1B	RELA	HDAC3	NRIP1	VDR	CDK7	E2F1	0.92203
22	PTPN14	CDH1	IQGAP1	MAPK1	PRKCE			0.937152
23	MYD88	IRAK1	PRKCI	SRC	PRKCE			0.949189
24	PTPN14	CDH1	IRS1	MAPK1	PRKCE			0.980665
25	PPP2R1B	RELA	MYC	GSK3B	E2F1			0.997489
26	TAL2	MAPK3	MYC	RBL1	E2F5			1.011626
27	CAPN2	NMT1	LYN	MAPK3	PRKCE			1.026447
28	PPP2R1B	HDAC1	PHB	E2F1				1.030947
29	TAL2	TCF3	CREBBP	E2F1				1.047172
30	PPP2R1B	RELA	HDAC3	PPARG	NCOA6	E2F1		1.048389
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮



factor (TF) usually plays the downstream role in this signaling flow. Since OCIAD2 differentially expressed in prostate cancer cell line, liver cancer associated mesenchymal stem cells, and especially in TGF $\beta$  treated Panc-1 pancreatic adenocarcinoma cell line, the question how is OCIAD2 activated by TGF $\beta$  signaling was solved by studying the probable transcription factor of OCIAD2 acting the downstream of TGF $\beta$  signal.

Browsed Transcriptional Regulatory Element Database [9] with 30981 genes, 177 transcription factors of homo sapiens were picked out as background human transcription factors library. The algorithm to find the possible transcription factors of OCIAD2 in TGF $\beta$  treated signaling was divided into three main steps: first, *pbMOO* approach was employed to calculate the costs of all the shortest paths between TGF $\beta$  and human transcription factors; then the ones with the least costs and high correlations with OCIAD2 in gene expression data were filtered out and selected as candidate transcription factors for OCIAD2; finally, biological TGF $\beta$  induced OCIAD2s differential expression mechanism was concluded with literature verification.

### *3.3.1 Speculation of Human TF Enrolled in TGF $\beta$ Signal*

As a fresh gene with rare reported property, the discovery of transcription factor in TGF $\beta$ 1 signal is the main issue in OICAD2 study. Among those 177 human transcription factors, the ones with the least pathway cost, which was defined by the sum of gene expression experimental pvalues of proteins on the pathway, are the most credible TFs for OCIAD2. The start point of the pathway was chosen as TGF $\beta$ 1, and the end point was OCIAD2. All the pathway costs for those passing through TFs were calculated by applying *pbMOO* approach and top of them were listed in the following table3.2.

Table 3.2: OCIAD2's Transcription Factors Detection in TGF $\beta$  Signal

No.									Cost
1	TGFB1	ITGAV	ITGB1	ITGA3	FHL2	AR	OCIAD2		0.022027
2	TGFB1	ITGAV	ITGB1	FHL2	AR	OCIAD2			0.024967
3	TGFB1	THBS1	PDGFB	PDGFRB	STAT4A	OCIAD2			0.029893
4	TGFB1	THBS1	SPARC	PLAT	MAPK3	CAV1	AR	OCIAD2	0.033335
5	TGFB1	VTN	FGF2	CASP1	AR	OCIAD2			0.035584
6	TGFB1	VTN	SERPINE1	KRT18	RAF1	AR	OCIAD2		0.036216
7	TGFB1	ACVRL1	XIAP	TRAF6	CAV1	AR	OCIAD2		0.047305
8	TGFB1	THBS1	LRP5	SMAD9	OCIAD2				0.062151
9	TGFB1	DCN	EGFR	AR	OCIAD2				0.063297
10	TGFB1	TGFBR1	CAV1	AR	TRIM24	RXRA	PPARG	OCIAD2	0.078664
11	TGFB1	TGFBR1	CAV1	AR	PNRC1	RARG	OCIAD2		0.081610
12	TGFB1	TGFBR1	CAV1	AR	TRIM24	RXRA	PPARG	OCIAD2	0.081804
13	TGFB1	TGFBR1	CAV1	AR	FHL2	WT1	OCIAD2		0.082158
14	TGFB1	TGFBR1	SMAD7	PIAS1	AR	OCIAD2			0.083669
15	TGFB1	TGFBR1	SMAD7	PIAS1	MSX1	OCIAD2			0.08367
16	TGFB1	TGFBR1	SMAD7	HEYL	SMAD9	OCIAD2			0.083735
17	TGFB1	TGFBR1	PIK3R1	PDGFRB	STAT5A	OCIAD2			0.083851
18	TGFB1	TGFBR1	PIK3R1	IL2RB	STAT5A	OCIAD2			0.083851
19	TGFB1	TGFBR1	CAV1	AR	OCIAD2				0.086596
20	TGFB1	TGFBR1	CAV1	MAPK3	STAT5A	OCIAD2			0.087053
21	TGFB1	TGFBR1	CAV1	AR	HIF1A	OCIAD2			0.08864
22	TGFB1	TGFBR1	SMAD7	OCIAD2					0.089536
23	TGFB1	TGFBR1	SMAD6	OCIAD2					0.089536
24	TGFB1	TGFBR1	CTNNB1	AR	OCIAD2				0.09008
25	TGFB1	TGFBR1	CAV1	AR	NCOA1	FOS	USF2	OCIAD2	0.090869
26	TGFB1	TGFBR1	CAV1	AR	NCOA1	FOS	OCIAD2		0.093771
27	TGFB1	TGFBR1	CTNNB1	HIF1A	OCIAD2				0.095064
28	TGFB1	TGFBR1	CTNNB1	MITF	FOS	OCIAD2			0.099917
29	TGFB1	TGFBR1	PAK1	CASP1	AR	OCIAD2			0.10004
30	TGFB1	TGFBR1	CAV1	AR	FOXO1	PPARG	OCIAD2		0.106862
:	:	:	:	:	:	:	:	:	:



### 3.3.2 Feasible TF of OCIAD2 in Cancer Cell Line

In this part, verification of the observation that AR might be the transcription factor of OCIAD2 in TGF $\beta$ 1 signal, and SMAD group might enrolled this process, was made in the light of gene expression data from both HCC and Prostate Cancer cell lines.

Due to the fact that GSE42357 was gene expression comparison data between liver cancer associated mesenchymal stem cells (LC8-MS-C) and normal ones (LN8-MS-C) from the same patient, genes like OCIAD2 had only two experimental values one for condition LC8-MS-C and one for control LN8-MS-C. The sample space was too tiny for Pearson Correlation calculation. For better result, the distribution of the fold change of each gene was plotted as the following Figure 3.3, and evidently, AR, which had ten pairs of experiment data in the range [1.09315, 1.74845], highly differentially expressed in HCC microenvironment. The fold change value of OCIAD2 in the same array data is -0.377255, which implied that AR must have negative effect on OCIAD2, in the other word, it should be the inhibitor of OCIAD2. Not like the top obviously expressed genes with fold changes close to 3.0, the SMAD group showed the relative lower differentially expression- most of them had slight positive changes less than 0.1. However, SMAD4 with fold change 0.06034, SMAD2 0.4908 and SMAD3 0.02102 still survived as pathway proteins in *pbMOO* predictions, which were ignored by other pathway analysis methods.

Analyzing the results (top 30 were detailed in Supplementary Table 1.3), two interesting facts were observed: AR was shown with the highest frequency as the transcription factor of OCIAD2 in 17 pathways out of the top 50, while STAT5A was the second recurrent one that was the transcription factor of 10 pathways; AR appeared 34 times, and SMAD group proteins appeared 16 times in the top 50

pathways with lowest cost, in which other transcription factors had less occurrences. The observation insinuated AR was the most reliable transcription factor of  $TGF\beta 1$  signal induced OCIAD2, and SMAD group proteins had the closest relationships with this signaling process.

In DU145 prostate cancer cell line with restored miR-205 expression, unfortunately no data mapping with AR was found. However, this 8 samples experimental data was still powerful to analysis how SMADs enrolled in OCIAD2 expression. As the figure showed, SMAD4 had the largest expression value as well as the highest negative correlation -0.696476 with differential expressed OCIAD2 among SMADs, followed by SMAD2 with correlation value -0.595238 and SMAD6 -0.571429.

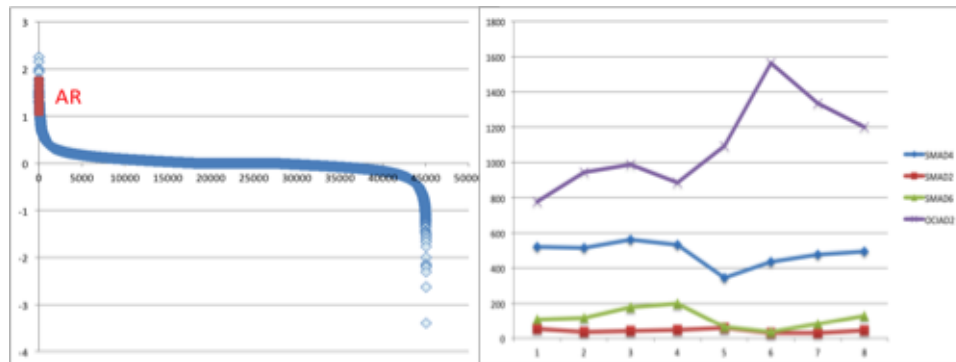


Figure 3.3: Transcription Factor Detection for OCIAD2

### 3.3.3 Mechanism of $TGF\beta$ Induced OCIAD2's Expression

Analyzed the observations comprehensively on modular study and referred to related literature, the signaling pathway from  $TGF\beta 1$  targeting to OCIAD2 was concluded as the Figure 3.4: the signaling transmits from  $TGF\beta 1$ -  $TGF\beta R1$ - AR-

OCIAD2 in liver cancer mesenchymal stem cell, then differentiates into Tumor-Associated-Fibroblasts (TAFs) in tumor stroma. As the only known mammalian coSMAD, SMAD4 transferred signaling from cytoplasm to in TGF $\beta$  signal. AR, the symbol of androgen receptor, mainly functioned as a DNA-binding transcription factor that regulates target gene expression from cytoplasm into nucleus. Specifically, the path TGF $\beta$ 1- TGF $\beta$ R1- SMAD2/3- SMAD4 was reported in [20] and [21], and TGF $\beta$ 1s influence on OCIAD2 through AR was also proved in [22].

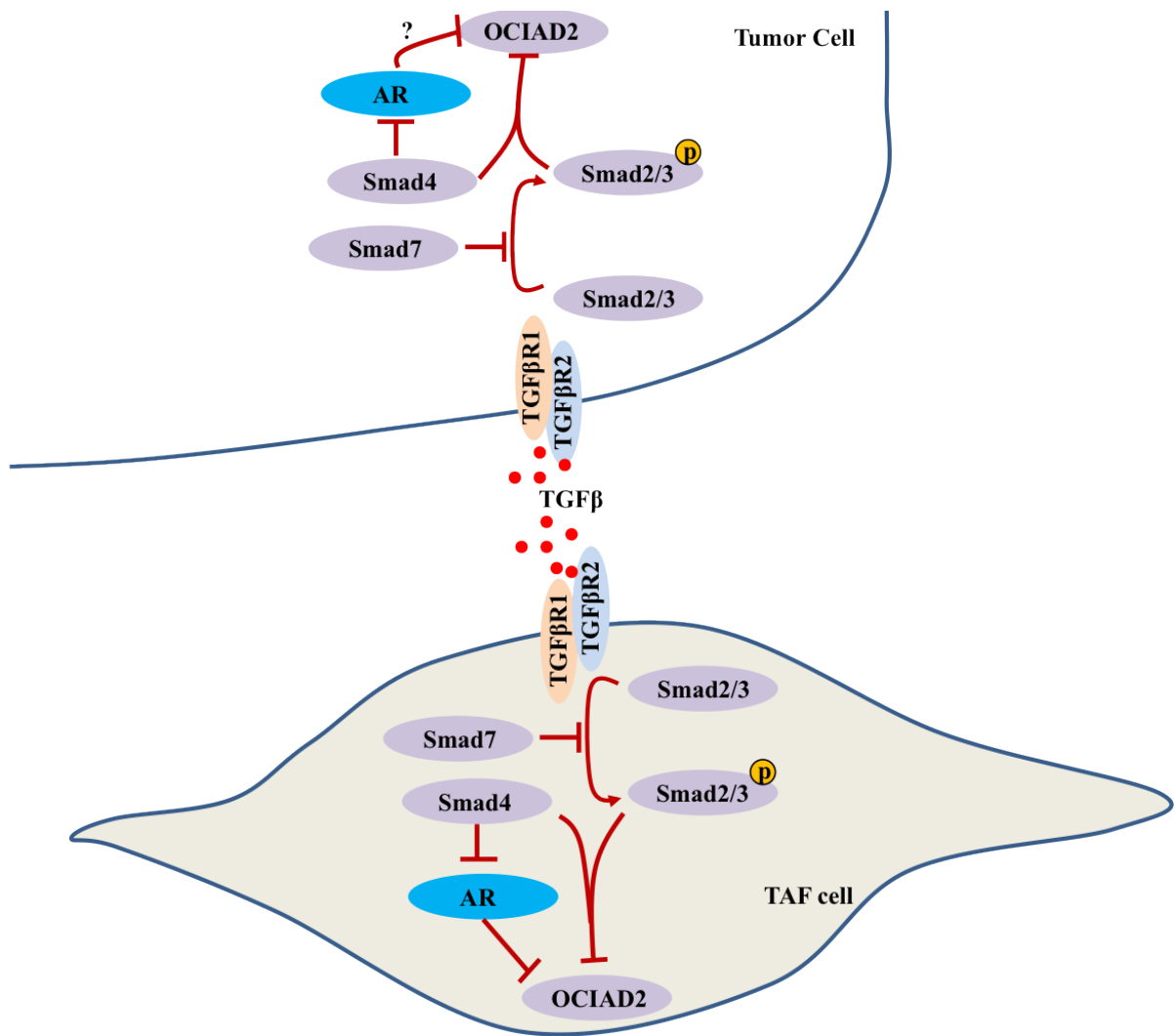


Figure 3.4: TGFβ Induced OCIAD2s Expression Mechanism

## 4. MTB INDUCED CCL2 PATHWAY ODE MODEL DYNAMIC ANALYSIS

### 4.1 Background Introduction

#### *4.1.1 Motivation and Experiments Design*

Infection with *Mycobacterium tuberculosis* leads to recognition of the bacteria/bacterial components by the macrophage surface receptors causing activation of signaling cascades leading to pro-inflammatory response elicited by innate and adaptive immunity. Pathogen recognition leads to activation of MAP kinases like ERK1/2, JNK and p38 (Howard Yim 2011 PNAS) further terminating in activation of transcription factors for e.g. NFkB and AP-1 that regulate expression of critical cytokines like TNF-a, IL-6, IL-1b etc.

In order to understand the effect of mycobacterial tuberculosis induced signaling in monocytes, we challenged THP-1 cells with sonicated Mtb antigen as follows:

a) THP-1 monocyte cells challenged with Mtb for 0, 5, 10, 15, 30, 60 and 90 minutes;

b) THP-1 cells+ Mtb incubated with PAR-1 inhibitor (SCH79797) for 0, 5, 10, 15, 30, 60 and 90 minutes;

c) THP-1 cells+Mtb incubated with exogenous MCP-1 for 0, 5, 10, 15, 30, 60 and 90 minutes.

In this thesis, we mainly focused on Mtb stimulated cells responses to exogenous MCP-1 treatment comparing with Mtb alone treated cells.

## *4.1.2 Experimental Result Observations*

### *4.1.2.1 Cells Challenged with Mycobacterial Antigen Alone*

Macrophages when exposed to Mtb activates various surface receptors to induce expression of pro-inflammatory responses e.g. production of cytokines. Mtb stimulation of the THP-1 cells lead to increased protein expression (40% more than control) of EF4E, MEK1, NFKB, P38, P70SK, PDCD4, S6235 and S6240 as indicated in Table 1. It was interesting to note that the increased expression of these proteins was sustained with additional treatments like MCP-1 and PAR-1 addition (highlighted in yellow in Figure 4.1). We also observed increased expression of cMYC and Caveolin protein in cells treated with Mtb alone. cMYC has been reported to regulate anti-mycobacterial responses via induction of TNF-a and IL-6. Mycobacteria induce Myc expression via activation of ERK1/2 and JNK1/2 and myc also increases NF-kb pathway [23].

a) Mtb stimulation		b) Mtb+PAR-1 inhibitor		c) Mtb+exogenous MCP-1	
Gene symbol	Protein	Gene symbol	Protein	Gene symbol	Protein
MYC	cMYC	CASP7	Casp7	AKT	Akt R.V
CAV1	Caveolin	PECAM1	CD31	BCL2L1	Bcl XL
EIF4E	EIF4E	EIF4E	EIF4E	CTNNB1	Beta-Catenin
MAP2K1	MEK1	EIF4G	Eif4G	MYC	CMyc
NFKB1	NFKB	FANCD2	FANCD2	PECAM1	CD31
MAPK14	P38	GSK3	GSK3	CHEK2	Chk2
RPS6KB1	P70SK		GSK3	EEF2	EEF2
PDCD4	PDCD4	MAP2K1	MEK1	EGFR	EGFR
RPS6	S6235	NFKB1	NFKB1	EIF4E	EIF4E
	S6240	MAPK14	P38	G6PD	G6PD
		RPS6KB1	P70	ERBB2	HER2
		PDCD4	PDCD4	MAPK9	JNK2
		RPS6	S6	NF2	NF2
			S6	MAPK14	P38
		YBX1	YB1	RPS6KB1	P70
				PCNA	PCNA
				PDCD4	PDCD4
				PEA15	PEA15
				PRKCA	Pkca
				PRKCD	Pkcd
				PKC	PKC pan
				RB1	RB
				RBM15	RBM
				RPS6	S6
					S6
				SYK	Syk
				TRFC	TRFC
				TTF1	TTF1

Figure 4.1: Gene Expression in Different Groups of Cell Lines

Gene expression analysis shows that Mtb challenge led to increase in immune cell

trafficking and inflammatory response. The genes were involved in pathways central to tuberculosis pathogenesis and host immune response including IL-6, IL-10 pathway, TNF receptor signaling pathway, pathways involved in pattern receptor recognition of bacteria and viruses and pathways involved in communication between innate and adaptive immune cells and acute phase response signaling.

Using the Ingenuity pathway analysis (IPA) of the gene array we observed that the Mtb challenge modulated various downstream genes that were involved in cell migration and chemotaxis of immune cells. Gene profile also indicated decrease in cell death and increase in cell survival. In addition to the downstream functional analysis, IPA based on array expression data predicted activation of various upstream regulators like ERK1/2, IRF3, JUN, JNK, MAP2K 1/2, NFkB, p38 MAPK and various Toll like receptors (TLR 2, 3,4, 9,8,7). Many of the predicted upstream regulators of the differentially expressed genes were also present on the protein array. Out of the predicted upstream regulators, RPPA data confirmed Mtb induced enhanced activity of MAP2K 1 (MEK1), Nfkb and p38 MAPK (labeled as blue arrows in Figure 4.2). Previous literature showed that CXCL10, MCP-1, IL-15 can differentiate active TB patients for infected but not sick individuals. CXCL10 and MCP-1 upregulated in our experiment.



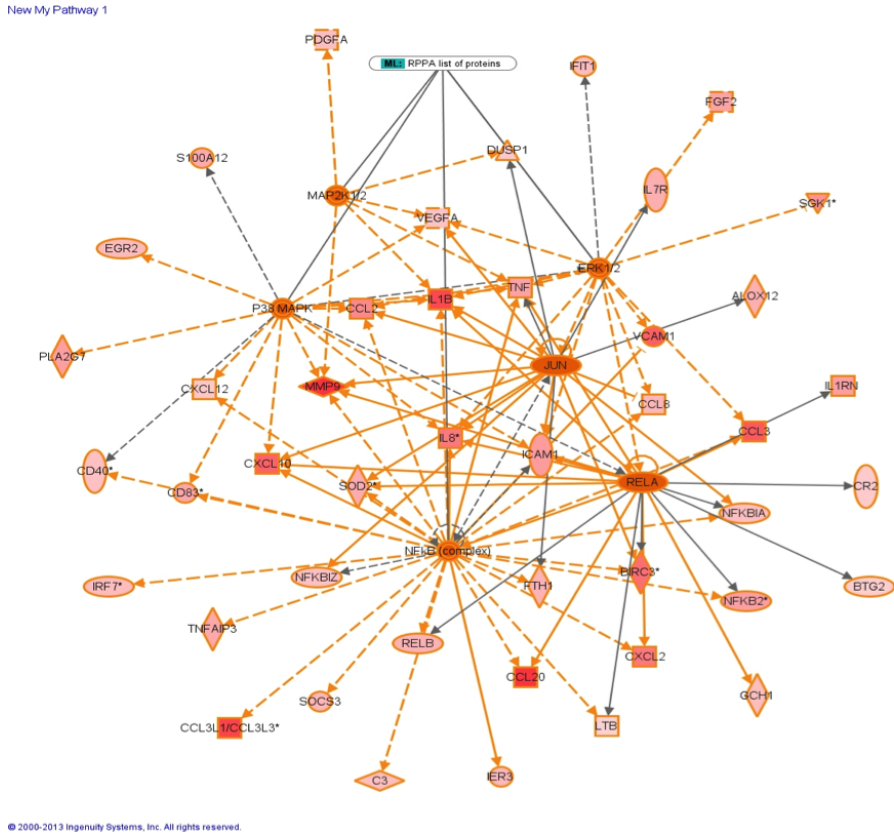


Figure 4.2: Proteins in RPPA List Analyzed by IPA

We observed that proteins activated by the Mtb antigen as obtained by RPPA analysis: MAP2K 1 (MEK1), Nfkb and p38 MAPK, leads to the regulation downstream genes that were a part of top 5 networks from array IPA analysis: 1) cellular movement and system development, 2) hematological development and inflammation 3) antimicrobial response. These proteins regulate signaling cascades like: TREM1 signaling, IL-6 signaling, Acute phase response signaling, IL-8, IL-17 and apoptosis signaling.

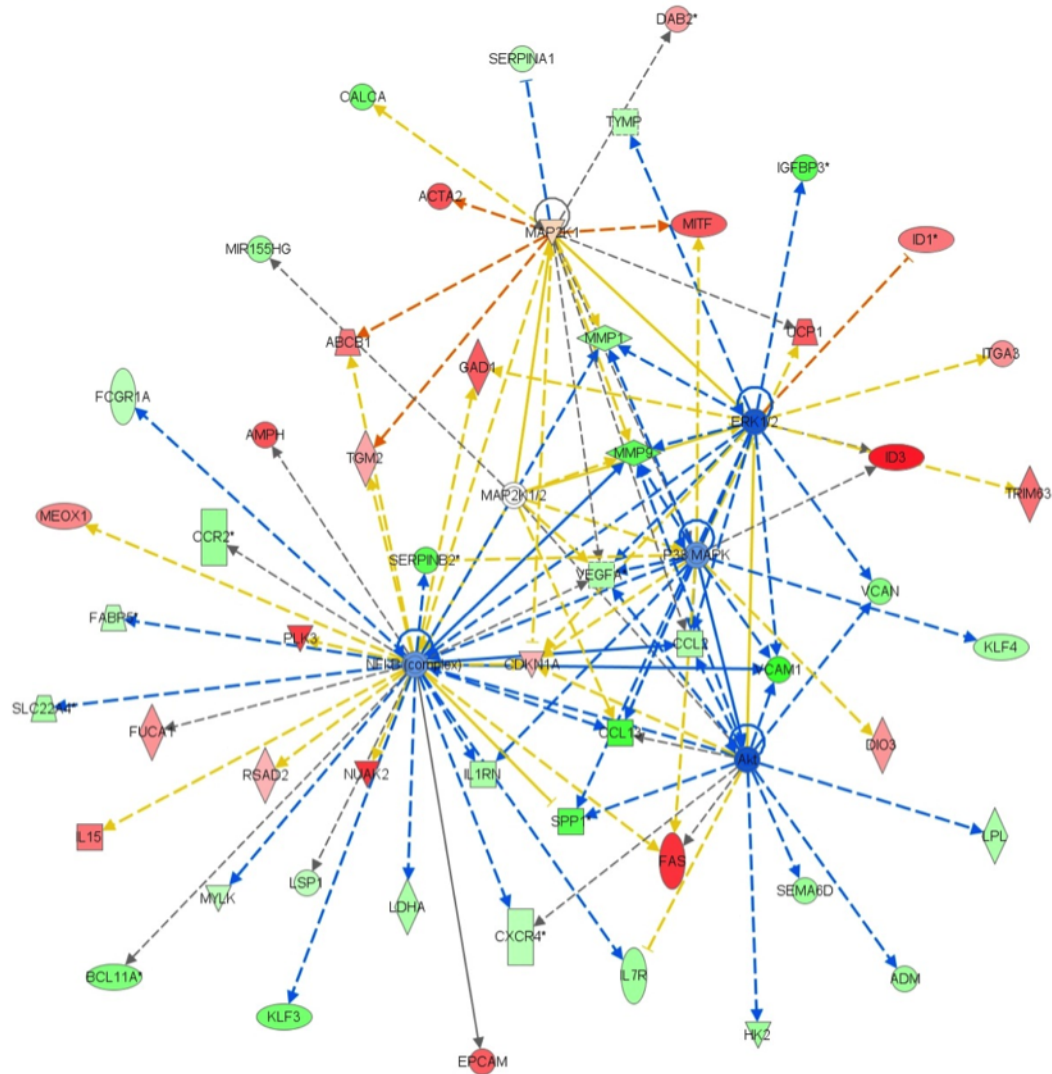
Hence our RPPA data depict that early MTB stimulation of monocytes/macrophages leads to p38 MAPK, NFkB and MAP2K activation. Microarray data supports this

observation and shows that activation of these upstream elements likely via pattern recognition receptors (like TLRs) will lead to induction of various acute phase response molecules for e.g. IL1b, TNF, IL6, MCP-1/CCL-2 etc. It has been reported that pro-inflammatory signals and other stimuli can lead to expression of matrix metalloproteinases from macrophages in tuberculosis. TNF $\alpha$  has been shown to influence elevated levels of MMP-1 [24]. TLR8 was found to be up-regulated after Mtb stimulation in the array results. TLR8 polymorphism has been associated with TB and this gene has been reported to be up-regulated in the acute phase of the disease [25].

#### *4.1.2.2 Influence of PAR-1 Inhibitor on Mtb Challenged Monocytes*

PAR-1 is a receptor for thrombin, plasmin, MMP-1 etc. Many studies in cancer biology have shown that MMP-1 can cleave PAR-1, activating it, leading to downstream signaling influencing MMP-1 levels. Study by [26] showed that MMP-1 via PAR-1 influences levels of MMP-1, MCP-1 and MMP-9. Studies so far have reported that MMP-1 acts via PAR-1 by activating signaling pathways like Rho-GTPase, MAPK signaling [27] with molecules like p38, MEK1/2 and Akt [28] in various cell types leading to diverse functions including cellular proliferation [29] and invasion. While PAR-1 blockage by its inhibitor is expected to decrease the levels of MMP-1, MMP-9 and CCL2 [30], the signaling pathways implicated in this process are still not known in macrophages stimulated by tuberculosis. Upon receptor activation by MMP-1, PAR-1 inhibitor has been shown to reduce the activity of p38, MEK1/2 and Akt. In our RPPA data where the cells were simulated by PAR-1 inhibitor, we observed activation of MAP2K1 (MEK1), NF $\kappa$ B, p38 MAPK (Figure 4.1). We still confirmed the effect of PAR-1 inhibitor at the gene expression level by microarray analysis and found downregulation of MMP-1, CCL2, CCR2 (receptor for

CCL2), MMP-9 and TIMP3 (Figure 4.4). It is possible that Mtb stimulation leads to activation of the above mentioned signaling molecules e.g. due to TNF $\alpha$ , IL-1 $\beta$  etc. Based on the predicted upstream regulators by Ingenuity analysis of the array data, it is possible that signaling molecules other than p38, MEK1/2 or Nf $\kappa$ B may influence MMP-1 levels via PAR-1 activation like Akt, JNK or ERK, as reported in other PAR-1 studies (Figure 4.4). From our RPPA data, we gathered that phosphorylated AKT, JNK or ERK protein levels did not change across the time point by PAR-1 treatment, which could potentially be due to PAR-1 inhibition leading to reduced MMP-1 and MMP-9 gene expression. In order to determine the effect of MMP-1 induced activation of PAR-1 signaling, we would need to perform further experiments (experiments underway). Based on gene expression profile of genes Ingenuity analysis also predicted inhibition of Akt, Ca $^{+}$ , JUN, ERK 1/2 and JNK. The next step will be to explore these molecules for PAR-1 inhibition.

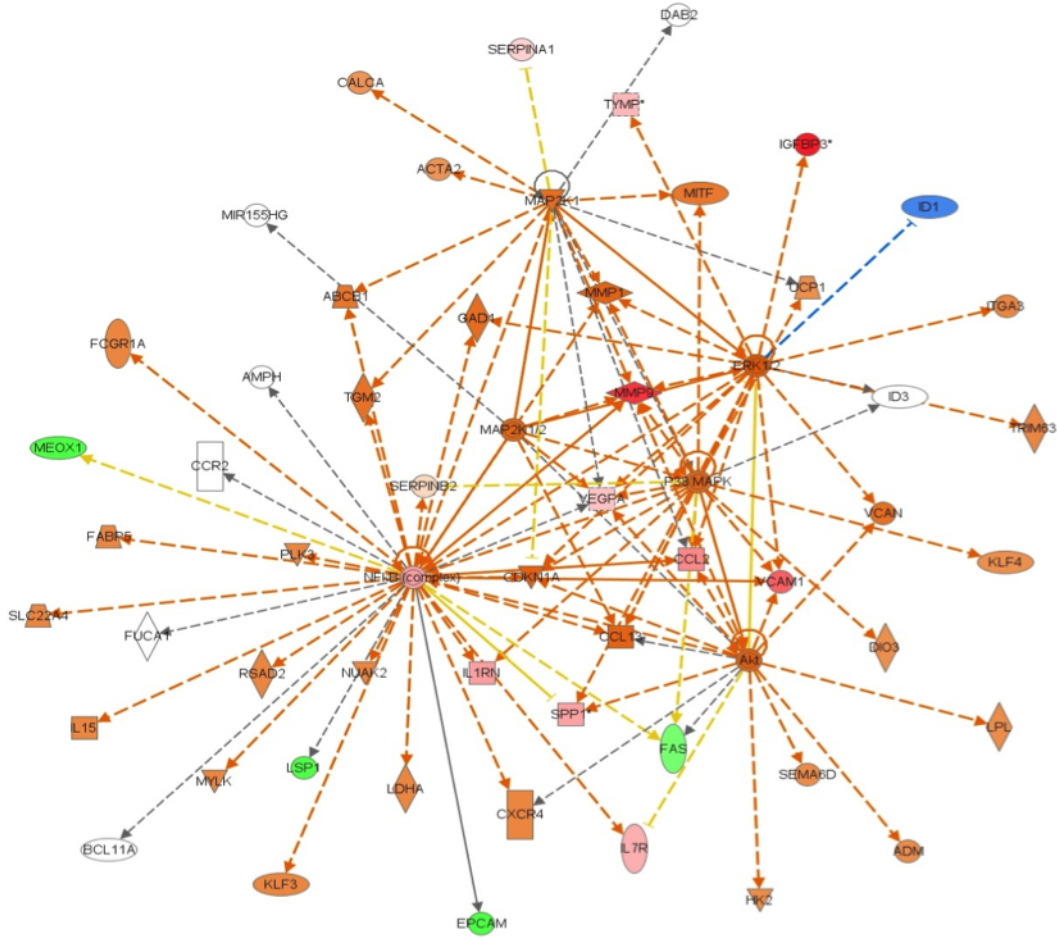


© 2000-2013 Ingenuity Systems, Inc. All rights reserved.

Figure 4.4: Possible Signaling Molecules via PAR-1 influence MMP-1 by IPA

We found GSK3 protein to be activated in the cells with PAR-1 inhibitor. The activity of GSK3 gets inhibited by phosphorylation at the N-terminal serine residues [31]. It has been reported that mycobacterium inhibits GSK3 activity leading to

IL-10 production. IL-10 is anti-inflammatory cytokine and suppresses IFN- $\gamma$  and HLA-DR expression. Mycobacterium uses IL-10 as an evasion mechanism from host immunity [32]. This study reported that live BCG was able to inhibit GSK3 leading to IL-10 upregulation. We observed that GSK3 phosphorylation, and in-turn inhibition was high in cells stimulated by Mtb and treated with PAR-1 inhibitor. GSK3 phosphorylation was observed in cells challenged with Mtb alone as well, however, to a lesser extent. It will be interesting to further explore the role of MMP-1, PAR-1 signaling of IL-10. GSK3 inhibition could also just a non-specific side effect of SCH79797. As mentioned in the previous section, Mtb led to up-regulation of TLR8 gene. Treatment with PAR-1 inhibitor led to its downregulation. TLR8 is located in the endosomal/lysosomal compartment of cells and lead to signaling of pro-inflammatory cytokines and IFN. Since the up-regulation of TLR8 has been associated with TB active disease, its reduced expression by PAR-1 inhibition could have important therapeutic implications. Also, IL-15 was found to be up-regulated in cells treated with PAR-1. IL-15 has been found to be elevated in healthy control individuals compared to TB patients [33]. Functional analysis of downstream genes in microarray with PAR-1 treated cells showed that mainly down-regulation of various biological functions has occurred, for e.g. Recruitment of monocytes (down-regulation of CCL13, CCL2, CCR2), reduced accumulation of macrophages (down-regulation of CBS, CCL13, S100A9, CCL2, CCR2, IL1RN, VEGFA, SPP1 etc). IPA analysis showed that PAR-1 inhibitor addition to stimulated cells leads to increased biological functions like increased quantity of interferon, increased cell death and atrophy.



© 2000-2013 Ingenuity Systems, Inc. All rights reserved.

Figure 4.5: Stimulated Cells Responses to PAR-1 Inhibitor Addition by IPA

4.1.2.3 Influence of Exogenous MCP-1 on Mtb Challenged Monocytes

We found that addition of exogenous MCP-1 led to change in levels of around 25 proteins (40% more than expression than control in Mtb challenged monocytes. As mentioned earlier, TLRs recognize bacterial ligands and induce immune response by host cells. TLRs like TLR-1, 2, 3, 4 and 9 in macrophages induce pro-inflammatory mediators like MCP-1. Induction of MCP-1 in human monocytes by M. tuberculosis

requires NF-kB, ERK, and p38 MAPK signaling. Inflammatory mediators like TNF-1, IL-1B can also lead to induction of mcp-1 [34].

MCP-1 is a chemokine involved in trafficking of leukocytes to the site of inflammation and also induces cytokine expression from monocytes (Martina Werle, 2002). MCP-1 also induces expression of matrix metalloproteinases like MMP-1, MMP-9 [30] and MMP-2 [35] and [36] in various cells. Matrix metalloproteinases (MMPs) degrade extracellular matrix and have been implicated in TB chronic pathology. Tissue destruction associated with chronic TB may lead to spread of the mycobacterium and disease progression. Polymorphisms in MCP-1 genes causes increased expression of MCP-1 and associated with increase in MMP-1 levels, these factors may increase susceptibility to develop TB [37]. MCP-1 increases MMP-1 levels in Mtb stimulated cells of monocyte lineage [37] and [30]. In patients with active TB, a polymorphism in the promoter of MCP-1 leads to elevated expression levels of this protein [38].

From the RPPA data we observed that addition of MCP-1 to the Mtb stimulated cells led to activation of additional molecules like total AKT protein, JNK2, ERBB2, PRKCA/PRKCD and other translation machinery related protein (EEF2, EIF4E, RPS6KB1) etc. Activation of JNK2, AKT, P38, MEK1 by MCP-1 may lead to increased expression of MMP-1 and MMP-9 as depicted in Figure 4.1. MCP-1/CCL-2 has been reported to play a role in autophagy. MCP-1 via PI3K/AKT dependent signaling up-regulate surviving and leads to inhibition of autophagy in cancer cells [39]. Autophagy is also critical in TB since it has been reported to be involved in killing intracellular mycobacterium and regulating other pro-inflammatory cytokines [40]. As depicted in Figure 4.6, exogenous MCP-1 added to stimulated cells may further enhance mycobacterium survival by inhibition of autophagy. However, no experimental evidence supports that so far. From our RPPA data we observe that

exogenous MCP-1 addition leads to increased levels of AKT, P70Sk (RPS6KB1).

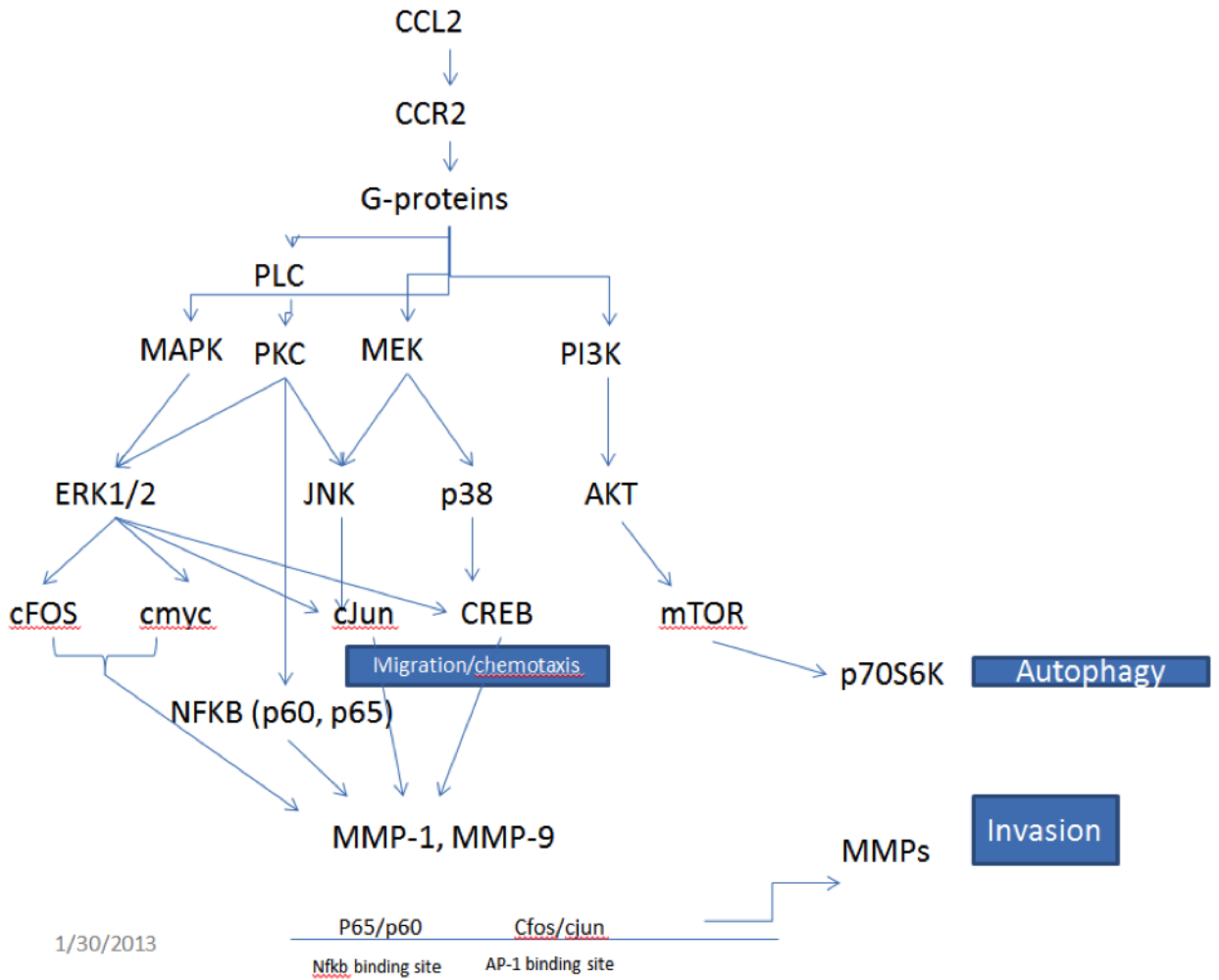


Figure 4.6: CCL2 Pathway in Mtb Simulated Cell

#### 4.2 ODE Model for CCL2 Pathway

In order to dynamic study how MCP1/CCL2 signaling works in Mtb stimulated cells, we built the Ordinary Differential Equations (ODE) model for CCL2 pathway.



With the 10 proteins that available in RPPA data list, PLC, MEK, PI3K, ERK, JNK, P38, AKT, mTOR, cMYC and P7OS6K were chosen as the components of the CCL2 pathway ODE model.

#### 4.2.1 ODE According to the Law of Mass Action

Assuming all the components are uniformly distributed and can access to each other with equal probability, applying the Law of Mass Action, Equation 1 describes the changes on the concentration of complex  $A$  due to binding (with rate  $a_0$ ) to the complex  $pB$  and dissociation (with rate  $d_0$ ) from complex  $pA$ , i.e.  $A + pB \xrightleftharpoons[d_0]{a_0} pA + B$ .

$$\frac{dA}{dt} = -a_0 * [A][pB] + d_0 * [pA] \quad (4.1)$$

Similarly, the whole pathway model for CCL2 pathway model for Mtb stimulated cells were presented by the following ordinary differential equations.

$$\frac{dCCL2_{Input}}{dt} = d_1 * [pCCL2_{Input}] \quad (4.2)$$

In this model, reactions "CCL2-CCR2-G Proteins" were treated as a whole and named  $CCL2_{Input}$ , which had no up stream protein. Therefore, the only change of the quantity of  $CCL2_{Input}$  was the decomposition with rate  $d_1$ .

$$\frac{dPLC}{dt} = -k_1 * [pCCL2_{Input}][PLC] + d_2 * [pPLC] \quad (4.3)$$

The quantity change of  $PLC$  came from the association of  $pCCL2_{Input}$  and  $PLC$  with rate  $k_1$ , and the dephosphorylation of  $pPLC$  with rate  $d_2$ .

$$\frac{dMEK}{dt} = -k_2 * [pCCL2_{Input}][MEK] + d_3 * [pMEK] \quad (4.4)$$

The quantity change of  $MEK$  came from the association of  $pCCL2_{Input}$  and  $MEK$  with rate  $k_2$ , and the dephosphorylation of  $pMEK$  with rate  $d_3$ .

$$\frac{dPI3K}{dt} = -k_3 * [pCCL2_{Input}][PI3K] + d_4 * [pPI3K] \quad (4.5)$$

The quantity change of  $PI3K$  came from the association of  $pCCL2_{Input}$  and  $PI3K$  with rate  $k_3$ , and the dephosphorylation of  $pPI3K$  with rate  $d_4$ .

$$\frac{dERK}{dt} = -k_4 * [pPLC][ERK] - k_5 * [pMEK][ERK] + d_5 * [pERK] \quad (4.6)$$

The quantity change of  $PI3K$  came from the association of  $pPLC$  and  $ERK$  with rate  $k_4$ , the association of  $pMEK$  and  $ERK$  with rate  $k_5$ , and the dephosphorylation of  $pERK$  with rate  $d_5$ .

$$\frac{dJNK}{dt} = -k_6 * [pPLC][JNK] - k_7 * [pMEK][JNK] + d_6 * [pJNK] \quad (4.7)$$

The quantity change of  $JNK$  came from the association of  $pPLC$  and  $JNK$  with rate  $k_6$ , the association of  $pMEK$  and  $JNK$  with rate  $k_7$ , and the dephosphorylation of  $pJNK$  with rate  $d_6$ .

$$\frac{dP38}{dt} = -k_8 * [pMEK][P38] + d_7 * [pP38] \quad (4.8)$$

The quantity change of  $P38K$  came from the association of  $pMEK$  and  $P38$  with rate  $k_8$ , and the dephosphorylation of  $pP38$  with rate  $d_7$ .

$$\frac{dcMYC}{dt} = -k_9 * [pERK][cMYC] + d_8 * [pcMYC] \quad (4.9)$$

The quantity change of  $cMYC$  came from the association of  $pERK$  and  $cMYC$  with rate  $k_9$ , and the dephosphorylation of  $pcMYC$  with rate  $d_8$ .

$$\frac{dAKT}{dt} = -k_{10} * [pPI3K][AKT] + d_{10} * [pAKT] \quad (4.10)$$

The quantity change of  $AKT$  came from the association of  $pPI3K$  and  $AKT$  with rate  $k_{10}$ , and the dephosphorylation of  $AKT$  with rate  $d_9$ .

$$\frac{dmTOR}{dt} = -k_{11} * [pPI3K][mTOR] + d_{10} * [pmTOR] \quad (4.11)$$

The quantity change of  $mTOR$  came from the association of  $pPI3K$  and  $mTOR$  with rate  $k_{11}$ , and the dephosphorylation of  $pmTOR$  with rate  $d_{10}$ .

$$\frac{dP70S6K}{dt} = -k_{12} * [pmTOR][P70S6K] + d_{11} * [pP70S6K] \quad (4.12)$$

The quantity change of  $AKT$  came from the association of  $pPI3K$  and  $AKT$  with rate  $k_{12}$ , and the dephosphorylation of  $AKT$  with rate  $d_{11}$ .

$$\frac{dMMPs}{dt} = -k_{13} * [pcMYC][MMPs] - k_{14} * [pJNK][MMPs] - k_{15} * [pAKT][MMPs] \quad (4.13)$$

The system output was defined as the quantity change of  $MMPs$  (the  $MMP$  family including  $MMP - 1$  and  $MMP - 9$ ), which came from the associations of  $pcMYC$  and  $MMPs$  with rate  $k_{13}$ ,  $pJNK$  and  $MMPs$  with rate  $k_{14}$ , and  $pAKT$  and  $MMPs$  with rate  $k_{15}$ .

### 4.2.2 Parameters Initialization

There were 10 RPPA available proteins in the system model, 11 dephosphorylation parameters and 15 association parameters in total. RPPA experimental data were normalized by dividing the original values at time zero. At time zero, all the proteins value were equal to  $RPPA_{N_0} = \frac{RPPA_0}{RPPA_0} = 1$ . Hence, the initial values of all the proteins in the system model were set to be 1. The system input,  $CCL2_{Input}$ , was also set as 1, in order to find the group of parameters when all the proteins initialized on the same level.

$$RPPA_{N_i} = \frac{RPPA_i}{RPPA_0}, i = \{0, 3, 10, 15, 30, 60, 90\}min \quad (4.14)$$

### 4.2.3 Solutions Optimization and Fitness

The system mean square error(MSE) was defined as the second moment of the differences between the estimated values and experimental data.

$$MSE = \frac{1}{nt} \sum_{i=1}^n \sum_{j=1}^t (\hat{R}_{ij} - R_{ij})^2 \quad (4.15)$$

Where  $n$  was the total number of proteins enrolled in the system pathway model, and  $t$  was the number of RPPA experiment time points, in our case  $n = 10$  and  $t = 7$ . Then, an optimization process was employed to finding the group of parameters best fitting the experiment data, in the other word, the system parameters with the minimum MSE (MMSE) value, and the objective function was defined as:

$$MMSE = arg \min MSE = arg \min \frac{1}{nt} \sum_{i=1}^n \sum_{j=1}^t (\hat{R}_{ij} - R_{ij})^2 \quad (4.16)$$

Solving the group of differential equations by genetic algorithm (GA) with 200

generations on Matlab, the optimized estimation values for the proteins in the model comparing with the normalized experimental RPPA data were shown as the following Figure 4.9. The fitness of the optimization process with 200 generations for CCL2 pathway model was presented as Figure 4.10. And the best estimated system had the minimum MSE  $MMSE = 0.008425$ .

Figure 4.9 showed that the estimations of PLC, PI3K, P38, MTOR, MEK, JNK and AKT were fitting the experimental data well. The other 3 out of 10 total proteins, i.e, cMYC, ERK and P70S6K did not match the experiment values exactly. The reason was those three proteins all had sharp changes in the first few minutes that they increased/decreased rapidly in the first 5 minutes and suddenly reversed their direction in the next few minutes. This made the total system MSE focused optimization hard to track their rapid changes. The genetic algorithm performed better in tracking relative long term tendency and global optimization for groups of parameters than those specific 5-min reversing points. Eventually, 7 out of 10 estimations fitting the experimental data guaranteed the reliability of the ODE pathway model.

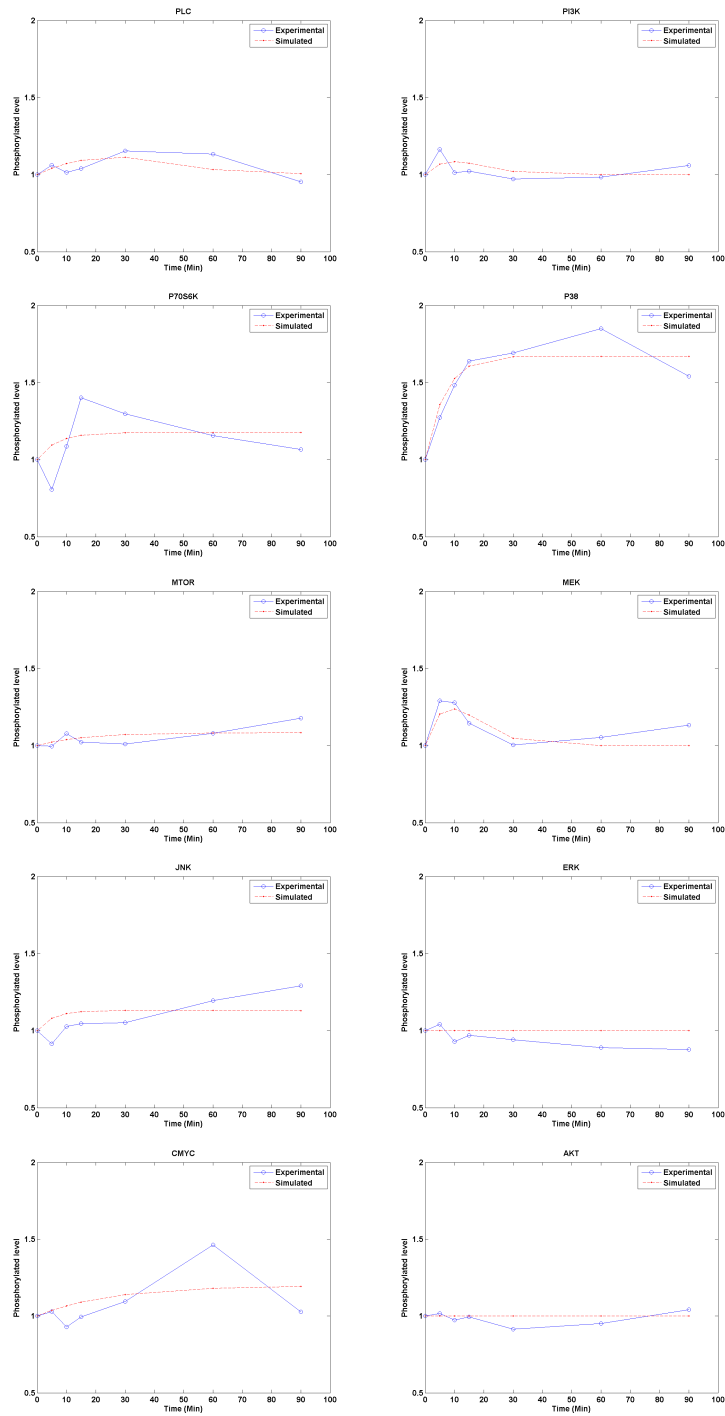


Figure 4.9: ODE Solutions vs. Experimental Data

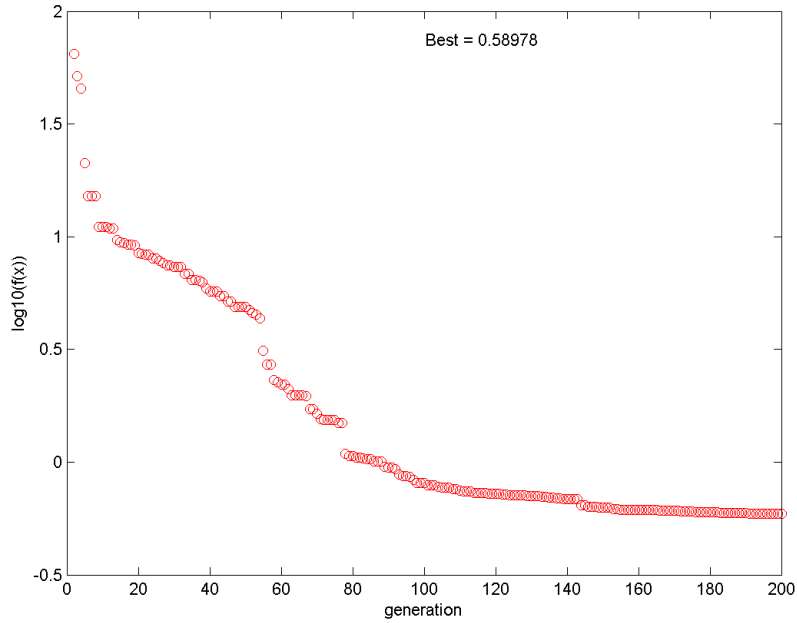


Figure 4.10: CCL2 Pathway ODE Model Parameters Fitness

### 4.3 Sensitivity Analysis of CCL2 Signaling Pathway

Sensitivity analysis can tell how the model output changing with the input or structures variation. Hence, it can further identify the significant and redundant parts of the model structure.

The sensitivity was defined as the ration of system output change rate and parameter change rate.

$$S = \frac{\Delta Output}{Output} / \frac{\Delta Parameter}{Parameter} \quad (4.17)$$

The parameter sensitivity quantitatively told the significance of the related signal transmission. A sensitive parameter represented a main signaling path that affecting

the system output much.

In CCL2 pathway model, a 100 parameter change rate was employed in testing the parameter sensitivity and the results were listed as Figure 4.11. Increased each parameter in turn by 100 times, the changes of the system output were plotted as the blue columns. Comparing to the other parameters changing the system output with nearly 0% and regardless of parameters  $(k_{13}, k_{14}, k_{15})$  without RPPA data and directly assigned to system output,  $(k_6, k_9, k_{10})$  and  $(d_6, d_8, d_9)$  were proved to be the most sensitive parameters for the CCL2 pathway, and obviously represented the top important links and nODE in the pathway ODE model. That meant the associations of  $pPLC - JNK$ ,  $pERK - cMYC$ , and  $pPI3k - AKT$ , the dephosphorylation of  $pJNK$ ,  $pcMYC$  and  $pAKT$  were the most important processes of the CCL2 signaling in Mtb stimulated cell.

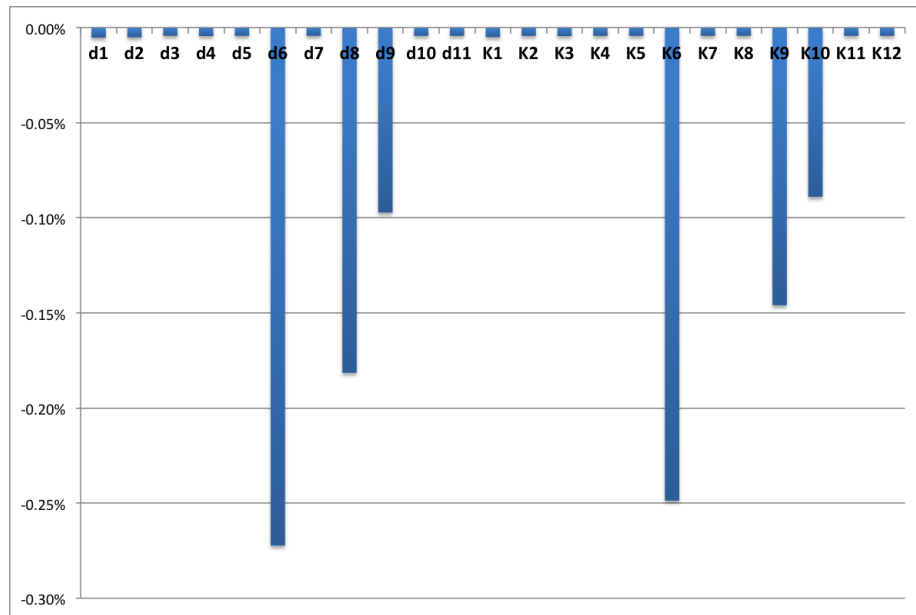


Figure 4.11: Output Changes When Parameter Increased 100 Times



To further determine the identification of the parameters, the coefficient of variation (CV), defined as the ratio of standard deviation to the mean, i.e.,  $Cv = \frac{\sigma}{\mu}$ , was employed the 26 system parameters. When a given parameter with coefficient of variation greater than 1, the parameter was unidentifiable, and vice versa. From Figure 4.12, the coefficients of variation were all less than 1. Hence, the parameters of ODE model for CCL2 pathway were all identifiable.

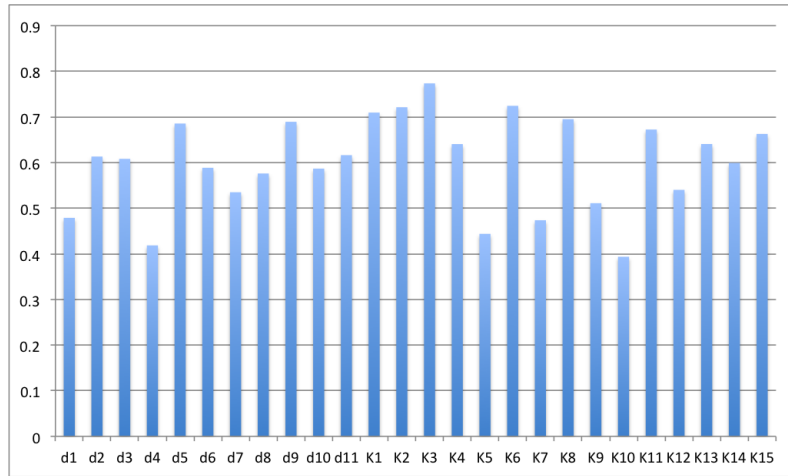


Figure 4.12: Coefficient of Variation Analysis of Parameters

#### 4.4 Model Robustness Analysis

To investigate the robustness of the model, the most sensitive parameters ( $k_6, k_9, k_{10}$ ) and ( $d_6, d_8, d_9$ ) were selected to test the system responses to parameter variations, while other parameters were fixed.

Figure 4.13 described the changes of system output related to  $10^{0.1}, 10^2, 10^5, 10^{10}, 10^{100}$  fold variations of the key parameters. The results showed that the system slightly adjusted its outcomes as the responses to dephosphorylation parameters change in

the range  $[10^{0.1}, 10^2]$  and association parameters change in the range  $[10^{0.1}, 10^5]$ , and then kept stable as the parameters continuously increasing. Thus, the CCL2 pathway ODE model was comparably robust with relatively large ranges of parameter variations.

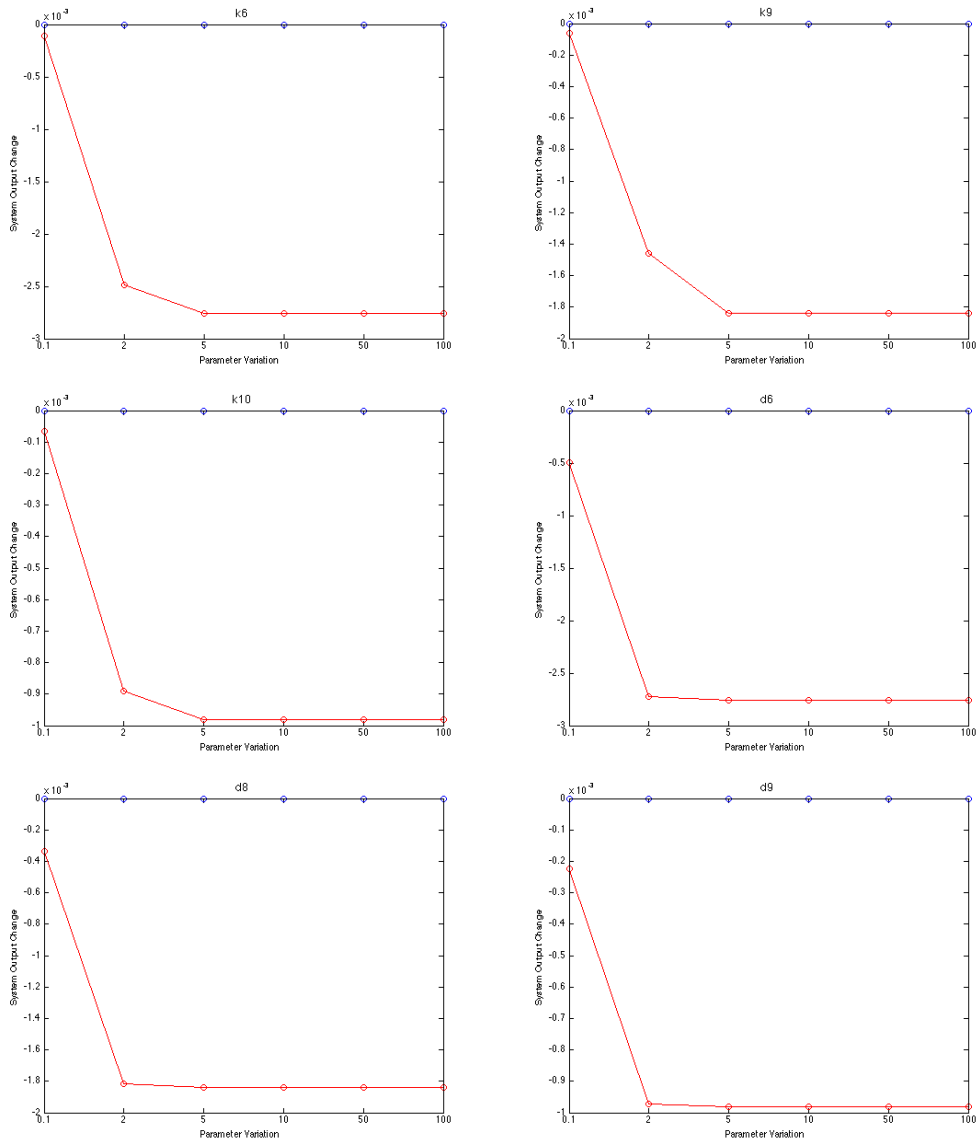


Figure 4.13: System Model Robustness Analysis

#### 4.5 System Perturbation and Drug Treatment Prediction

To study the system responses to over expression or inhibition of the proteins, the perturbation process by increasing/decreasing initial value of each protein slightly was employed, and the system outcome changes were represented as Figure 4.14. When increased JNK by 1%, the system gave the strongest response than to other proteins and relatively increased output by 0.002%, which was followed by cMYC with 0.00014% and PLC with 0.00013% system output changes. Hence, JNK, cMYC and PLC were the most significant modules in the CCL2 pathway ODE model that affecting the system outcome much. The discoveries also suggested that the drug treatment targeting on JNK, cMYC and PLC could be the most effective therapy for preventing Mtb infections.

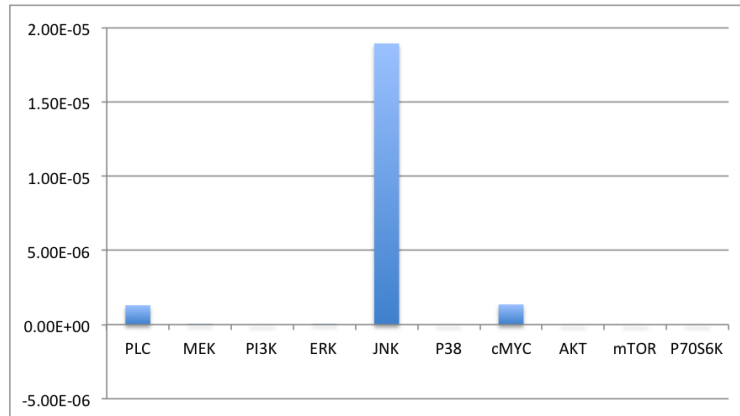


Figure 4.14: System Outcome Changes by Perturbation

## 5. SUMMARY

In this study, a bioinformatics approach was developed to demonstrate that a novel and function unknown protein, Ovarian Carcinoma Immunoreactive Antigen-like protein 2 (OCIAD2), is probably regulated by TGF $\beta$  and AR signals in the tumor EMT process. OCIAD2 is an immunoreactive antigen and its functions and involved pathways, molecular mechanisms have never been reported. Current popular signaling analysis tools like IPA [12] are focusing on the highest differentially expressed genes with sufficient literature supported, but have nothing to do with novel signatures such as OCIAD2. Moreover, as the comprehensive analysis of wide-field database, the output pathways of those tools can hardly be specific for an interesting stimulation or disease. By predicting the modular signaling mechanism from the given ligand to the pointed signature, the pbMOO approach successfully answered the question how the observed signature gene OCIAD2 got involved in ligand TGF $\beta$  stimulation signal, overcoming the insufficiency of previous reports of new gene OCIAD2 and detailing the predicted pathways into the tumor microenvironment.

Loss of cell adhesions or polarity is widely associated with CDH1 (E-cadherin). This process, referred to as EMT, enhances motility and invasiveness of many cell types and is often considered as a prerequisite for tumor infiltration and migration. TGF $\beta$  mediated induction of EMT processes is associated with specific stages of morphogenesis and during tumorigenesis by activating downstream signaling pathways in both Smad-dependent and -independent mechanisms. The up-regulation of OCIAD2 expression by TGF $\beta$  stimulation, and downregulated OCIAD2 expression in metastatic prostate tissues, made us to postulate that OCIAD2 played roles in

TGF $\beta$  promoted tumor cell migration, invasion and mobility. Indeed, OCIAD2 has been enrolled in TGF $\beta$  signal across CDH1 by pbMOO Approach in prostate cancer cells which potently testified our postulation that OCIAD2 could act as a downstream effector of TGF $\beta$  signals. In our predicted path, SMAD4 had the largest expression value as well as the highest negative correlation -0.696476 with differential expressed OCIAD2 among SMADs families. Previous study reported that Smad3/4 cooperated with Snail1 acted as co-repressors of CDH1 in the EMT process [41]. Since there lack protein interaction information related with OCIAD2, future biological assay need to investigate potential relationships between CDH1 and OCIAD2 in tumor EMT. In addition, smad signaling is required to maintain epigenetic silencing of some key EMT related proteins in breast cancer progression [42]. Because OCIAD2 frequently methylated in some kinds of cancers [5], [6] and [7], we speculate that activated TGF $\beta$ -Smad signaling provides an epigenetic memory to maintain silencing of OCIAD2 in EMT as well. Thus, disruption of TGF $\beta$ -Smad4-OCIAD2 signaling may be a useful therapeutic strategy to target tumor progression.

With *pbMOO* approach, a new pathway TGF $\beta$ 1- TGF $\beta$ R1- AR- OCIAD2 in liver cancer mesenchymal stem cell was predicted, which will differentiate into Tumor-Associated-Fibroblasts (TAFs), one of the major components of tumor stroma. Stromal-epithelial crosstalk regulates all phases of cancer metastasis. In prostate cancer, Androgen signaling is central to stromal-epithelial cross-talk in tumor progression. Tissue-based studies of human prostate cancer have shown that stromal AR expression and transcriptional activity downstream of the AR are lower in stromal cells derived from carcinomas. Androgen Receptor (AR) may promote hepato-carcinogenesis or suppress HCC metastasis. These opposite roles of AR also occur in prostate cancer [43]. The potential mechanisms for the AR dual roles are possibly caused by

the differential AR signals in different cellular types: having an oncogenic role in stroma and epithelial cells, but a suppressive role in basal intermediate epithelial cells. As DU-145 is an AR-independent cell lacking of AR protein expression, the predicted path from AR to OCIAD2 in prostate cancer can't be guaranteed. Further biological experiments are still needed to explore the existence of pathway TGF $\beta$ -Smad4-OCIAD2 signaling in AR-dependent cell models as well.

The signal from TGF $\beta$ , via the AR, play a critical role in the deregulation of TGF $\beta$  signaling in prostate and/or liver tumorigenesis and that TGF $\beta$  effectors (Smads 3 and 4) serve as negative regulators of AR-mediated transcription in cancer cells has been established by several investigations [44]. With *pbMOO* approach, poMOO, a functional unknown protein, OCIAD2, was also enrolled into a signal pathway TGF $\beta$ 1- TGF $\beta$ R1- SMAD2/3- SMAD4- AR- OCIAD2 in tumor and adjacent microenvironment. Currently, clinical studies using anti-androgens had disappointing results, with few beneficial effects on patients, or even worse survival. Understanding the molecular mechanisms of AR in tumor microenvironment will undoubtedly further improve the results obtained with antitumor therapeutic strategies.

By dynamic analyzing the CCL2 pathways in MTB infected cell lines within ODE model, model parameter sensitivity analysis revealed that the associations of *pPKC – JNK*, *pERK – cMYC*, and *pPI3k – AKT*, and the dephosphorylation of *pJNK*, *pcMYC* and *pAKT* played key roles in the CCL2 signaling of Mtb stimulated cell, which strongly supported our suggested pathways in Figure 4.6 that MCP1/CCL2 enhanced MMPs via *ERK*, *JNK*, and *AKT* signaling pathways. Furthermore, the system dynamic responses to the perturbation of each protein showed that JNK, cMYC and PLC were the most significant modules that dominated the

concentration of MMPs as the system outcome much more than others. Hence, drug therapies targeting on the inhibitions of JNK, cMYC and PLC could be the most effective treatment for frustrating the MMPs increasing in MCP-1/CCL2 induced Mtb cells and further preventing Mtb infections.

## REFERENCES

- [1] I. P. Witz and O. Levy-Nissenbaum, “The tumor microenvironment in the post-paget era,” *Cancer Lett*, vol. 242, no. 1, pp. 1–10, 2006.
- [2] M. Zeisberg, J. Hanai, H. Sugimoto, T. Mammoto, D. Charytan, F. Strutz, and R. Kalluri, “Bmp-7 counteracts tgf-beta1-induced epithelial-to-mesenchymal transition and reverses chronic renal injury,” *Nat Med*, vol. 9, no. 7, pp. 964–8, 2003.
- [3] Y. Kojima, A. Acar, E. N. Eaton, K. T. Mellody, and et, “Autocrine tgf-beta and stromal cell-derived factor-1 (sdf-1) signaling drives the evolution of tumor-promoting mammary stromal myofibroblasts,” *Proc Natl Acad Sci U S A*, vol. 107, no. 46, pp. 20 009–14, 2010.
- [4] L. Y. Luo, A. Soosaipillai, and E. P. Diamandis, “Molecular cloning of a novel human gene on chromosome 4p11 by immunoscreening of an ovarian carcinoma cdna library,” *Biochem Biophys Res Commun*, vol. 280, no. 1, pp. 401–6, 2001.
- [5] F. Gueugnon, S. Leclercq, C. Blanquart, C. Sagan, L. Cellerin, M. Padiou, C. Perigaud, A. Scherpereel, and M. Gregoire, “Identification of novel markers for the diagnosis of malignant pleural mesothelioma,” *Am J Pathol*, vol. 178, no. 3, pp. 1033–42, 2011.
- [6] M. Kulis, S. Heath, M. Bibikova, A. C. Queiros, and et, “Epigenomic analysis detects widespread gene-body dna hypomethylation in chronic lymphocytic leukemia,” *Nat Genet*, vol. 44, no. 11, pp. 1236–42, 2012.



- [7] H. Noushmehr, D. J. Weisenberger, K. Diefes, H. S. Phillips, and et, “Identification of a cpg island methylator phenotype that defines a distinct subgroup of glioma,” *Cancer Cell*, vol. 17, no. 5, pp. 510–22, 2010.
- [8] H. Kim, J. Watkinson, V. Varadan, and D. Anastassiou, “Multi-cancer computational analysis reveals invasion-associated variant of desmoplastic reaction involving inhba, thbs2 and col11a1,” *BMC Med Genomics*, vol. 3, p. 51, 2010.
- [9] C. Jiang, Z. Xuan, F. Zhao, and M. Q. Zhang, “Tred: a transcriptional regulatory element database, new entries and other development,” *Nucleic Acids Res*, vol. 35, no. Database issue, pp. D137–40, 2007.
- [10] S. Peri, J. D. Navarro, T. Z. Kristiansen, and R. Amanchy, “Human protein reference database as a discovery resource for proteomics,” *Nucleic Acids Res*, vol. 32, no. Database issue, pp. D497–501, 2004.
- [11] M. Kanehisa and S. Goto, “Kegg: kyoto encyclopedia of genes and genomes,” *Nucleic Acids Res*, vol. 28, no. 1, pp. 27–30, 2000.
- [12] IngenuitySystems, 2013.
- [13] S. Wernicke and F. Rasche, “Fanmod: a tool for fast network motif detection,” *Bioinformatics*, vol. 22, no. 9, pp. 1152–3, 2006.
- [14] G. Jin, K. Cui, X. Zhou, and T. Wong, “Unraveling the signal-transduction networks in cancer metastasis,” *IEEE*, vol. 26, no. 5, pp. 129–32, 2009.
- [15] M. Ogawa, C. LaRue, and J. Drake, “Hematopoietic origin of fibroblasts /myofibroblasts: Its pathophysiologic implications,” *Blood*, vol. 108, no. 9, pp. 2893–6, 2006.

- [16] J. Stegmuller, M. A. Huynh, Z. Yuan, Y. Konishi, and A. Bonni, "Tgfbeta-smad2 signaling regulates the cdh1-apc/snnon pathway of axonal morphogenesis," *J Neurosci*, vol. 28, no. 8, pp. 1961–9, 2008.
- [17] J. Varani, M. J. Bendelow, D. E. Sealey, S. L. Kunkel, D. E. Gannon, U. S. Ryan, and P. A. Ward, "Tumor necrosis factor enhances susceptibility of vascular endothelial cells to neutrophil-mediated killing," *Lab Invest*, vol. 59, no. 2, pp. 292–5, 1988.
- [18] Y. Wei, C. A. Renard, C. Labalette, Y. Wu, L. Levy, C. Neuveut, X. Prieur, M. Flajolet, S. Prigent, and M. A. Buendia, "Identification of the lim protein fhl2 as a coactivator of beta-catenin," *J Biol Chem*, vol. 278, no. 7, pp. 5188–94, 2003.
- [19] K. B. Lee, J. H. Jeon, I. Choi, O. Y. Kwon, K. Yu, and K. H. You, "Clusterin, a novel modulator of tgf-beta signaling, is involved in smad2/3 stability," *Biochem Biophys Res Commun*, vol. 366, no. 4, pp. 905–9, 2008.
- [20] R. Vogelmann, M. D. Nguyen-Tat, K. Giehl, G. Adler, D. Wedlich, and A. Menke, "Tgfbeta-induced downregulation of e-cadherin-based cell-cell adhesion depends on pi3-kinase and pten," *J Cell Sci*, vol. 118, no. Pt 20, pp. 4901–12, 2005.
- [21] T. H. Lin, S. O. Lee, Y. Niu, D. Xu, L. Liang, L. Li, S. D. Yeh, N. Fujimoto, S. Yeh, and C. Chang, "Differential androgen deprivation therapies with anti-androgens of casodex or mdv3100 vs anti-androgen receptor of asc-j9 lead to promote vs suppress prostate cancer metastasis," *J Biol Chem*, 2013.
- [22] M. R. Cardillo, E. Petrangeli, L. Salvatori, L. Ravenna, and F. Di Silverio,

- “Transforming growth factor beta 1 and androgen receptors in prostate neoplasia,” *Anal Quant Cytol Histol*, vol. 22, no. 5, pp. 403–10, 2000.
- [23] C. H. Y. Howard, C. B. L. James, and S. Y. L. Allan, “A role for c-myc in regulating anti-mycobacterial responses,” *PNAS*, vol. 108, no. 43, pp. 11 749–54, 2011.
- [24] C. M. OKane, P. Elkington, and F. JS, “Monocyte-dependent oncostatin m and tnf- synergize to stimulate unopposed matrix metalloproteinase-1/3 secretion from human lung fibroblasts in tuberculosis,” *Eur J Immunol*, vol. 38, no. 5, pp. 1321–30, 2008.
- [25] S. Davila, M. Hibberd, D. R. Hari, and H. Wong, “Genetic association and expression studies indicate a role of toll-like receptor 8 in pulmonary tuberculosis,” *PLoS Genet*, vol. 4, no. 10, 2008.
- [26] M. Ganachari, H. Guio, N. Zhao, and P. Flores-Villanueva, “Host gene-encoded severe lung tb: from genes to the potential pathways,” *Genes Immun*, vol. 13, no. 8, pp. 605–20, 2012.
- [27] M. A. Karyn, C. Lidija, and K. Athan, “Matrix metalloproteases and par1 activation,” *Blood*, 2012.
- [28] E. Yang, A. Boire, A. Agarwal, N. Nguyen, K. O’Callaghan, P. Tu, A. Kuliopulos, and L. Covic, “Blockade of par1 signaling with cell-penetrating pepducins inhibits akt survival pathways in breast cancer cells and suppresses tumor survival and metastasis,” *Cancer Research*, vol. 69, no. 15, pp. 6223–6231, 2009.
- [29] W. Luo, Y. Wang, and G. Reiser, “Protease-activated receptors in the brain:

- receptor expression, activation, and functions in neurodegeneration and neuroprotection,” *Brain Research Reviews*, vol. 56, no. 2, pp. 331–345, 2007.
- [30] M. Ganachari, H. Guio, N. Zhao, and P. O. Flores-Villanueva, “Host gene-encoded severe lung tb: from genes to the potential pathways,” *Genes and Immunity*, vol. 13, no. 8, pp. 605–620, 2012.
- [31] S. Frame and P. COHEN, “Gsk3 takes centre stage more than 20 years after its discovery,” *Biochem. J*, vol. 359, pp. 1–16, 2001.
- [32] M. M. Chan, B. K. Cheung, J. C. Li, L. L. Chan, and A. S. Lau, “A role for glycogen synthase kinase-3 in antagonizing mycobacterial immune evasion by negatively regulating il-10 induction,” *Journal of Leukocyte Biology*, vol. 86, no. 2, pp. 283–291, 2009.
- [33] S. S. Pydi, A. R. Bandaru, S. Venkatasubramanian, S. Jonnalagada, and V. L. Valluri, “Vaccine for tuberculosis: Up-regulation of il-15 by ag85a and not by esat-6,” *Tuberculosis*, vol. 91, no. 2, pp. 136–139, 2011.
- [34] T. Jia, N. V. Serbina, K. Brandl, M. X. Zhong, I. M. Leiner, I. F. Charo, and E. G. Pamer, “Additive roles for mcp-1 and mcp-3 in ccr2-mediated recruitment of inflammatory monocytes during listeria monocytogenes infection,” *The Journal of Immunology*, vol. 180, no. 10, pp. 6846–6853, 2008.
- [35] M. Werle, U. Schmal, K. Hanna, and J. Kreuzer, “Mcp-1 induces activation of map-kinases erk, jnk and p38 mapk in human endothelial cells,” *Cardiovascular Research*, vol. 56, no. 2, pp. 284–292, 2002.
- [36] P. Sheen, C. M. OKane, K. Chaudhary, M. Tovar, C. Santillan, J. Sosa, L. Caviedes, R. H. Gilman, G. Stamp, and J. S. Friedland, “High mmp-9 ac-

- tivity characterises pleural tuberculosis correlating with granuloma formation,” *European Respiratory Journal*, vol. 33, no. 1, pp. 134–141, 2009.
- [37] M. Ganachari, J. A. Ruiz-Morales, J. C. G. de la Torre, J. Dinh, J. Granados, P. O. Flores-Villanueva *et al.*, “Joint effect of mcp-1 genotype gg and mmp-1 genotype 2g/2g increases the likelihood of developing pulmonary tuberculosis in bcg-vaccinated individuals,” *PloS one*, vol. 5, no. 1, p. e8881, 2010.
- [38] P. O. Flores-Villanueva, J. A. Ruiz-Morales, C.-H. Song, L. M. Flores, E.-K. Jo, M. Montaña, P. F. Barnes, M. Selman, and J. Granados, “A functional promoter polymorphism in monocyte chemoattractant protein–1 is associated with increased susceptibility to pulmonary tuberculosis,” *The Journal of Experimental Medicine*, vol. 202, no. 12, pp. 1649–1658, 2005.
- [39] H. Rikiishi, “Novel insights into the interplay between apoptosis and autophagy,” *International Journal of Cell Biology*, vol. 2012, 2012.
- [40] C. Ni Cheallaigh, J. Keane, E. Lavelle, J. Hope, and J. Harris, “Autophagy in the immune response to tuberculosis: clinical perspectives,” *Clinical & Experimental Immunology*, vol. 164, no. 3, pp. 291–300, 2011.
- [41] T. Vincent, E. P. Neve, J. R. Johnson, A. Kukalev, and et, “A snail1-smad3/4 transcriptional repressor complex promotes tgf-beta mediated epithelial-mesenchymal transition,” *Nat Cell Biol*, vol. 11, no. 8, pp. 943–50, 2009.
- [42] P. Papageorgis, A. W. Lambert, and F. e. Ozturk, S.and Gao, “Smad signaling is required to maintain epigenetic silencing during breast cancer progression,” *Cancer Res*, vol. 70, no. 3, pp. 968–78, 2010.

- [43] Y. Niu, S. Altuwaijri, K. P. Lai, C. T. Wu, and et, “Androgen receptor is a tumor suppressor and proliferator in prostate cancer,” *Proc Natl Acad Sci U S A*, vol. 105, no. 34, pp. 12 182–7, 2008.
- [44] H. Y. Kang, K. E. Huang, S. Y. Chang, W. L. Ma, W. J. Lin, and C. Chang, “Differential modulation of androgen receptor-mediated transactivation by smad3 and tumor suppressor smad4,” *J Biol Chem*, vol. 277, no. 46, pp. 43 749–56, 2002.