

Dynamic Network of Transcription and Pathway Crosstalk to Reveal Molecular Mechanism of MGd-Treated Human Lung Cancer Cells

Liyan Shao¹*, Lishan Wang¹*, Zhiyun Wei¹, Yuyu Xiong¹, Yang Wang¹, Kefu Tang¹, Yang Li¹, Guoyin Feng¹, Qinghe Xing^{3*}, Lin He^{1,2,3*}

1 Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai Jiao Tong University, Shanghai, China, **2** Institute for Nutritional Sciences, Shanghai Institutes of Biological Sciences, Chinese Academy of Sciences, Shanghai, China, **3** Institutes of Biomedical Sciences, Fudan University, Shanghai, China

Abstract

Recent research has revealed various molecular markers in lung cancer. However, the organizational principles underlying their genetic regulatory networks still await investigation. Here we performed Network Component Analysis (NCA) and Pathway Crosstalk Analysis (PCA) to construct a regulatory network in human lung cancer (A549) cells which were treated with 50 uM motexafin gadolinium (MGd), a metal cation-containing chemotherapeutic drug for 4, 12, and 24 hours. We identified a set of key TFs, known target genes for these TFs, and signaling pathways involved in regulatory networks. Our work showed that putative interactions between these TFs (such as ESR1/Sp1, E2F1/Sp1, c-MYC-ESR, Smad3/c-Myc, and NFkB1/RELA), between TFs and their target genes (such as BMP41/Est1, TSC2/Myc, APE1/Sp1/p53, RARA/HOXA1, and SP1/USF2), and between signaling pathways (such as PPAR signaling pathway and Adipocytokines signaling pathway). These results will provide insights into the regulatory mechanism of MGd-treated human lung cancer cells.

Citation: Shao L, Wang L, Wei Z, Xiong Y, Wang Y, et al. (2012) Dynamic Network of Transcription and Pathway Crosstalk to Reveal Molecular Mechanism of MGd-Treated Human Lung Cancer Cells. *PLoS ONE* 7(5): e31984. doi:10.1371/journal.pone.0031984

Editor: Yidong Bai, University of Texas Health Science Center at San Antonio, United States of America

Received: September 18, 2011; **Accepted:** January 16, 2012; **Published:** May 31, 2012

Copyright: © 2012 Shao et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the 973 Program (2010CB529600), the 863 Program (2009AA022701), the National Nature Science Foundation of China (81121001), the Shanghai Municipal Commission of Science and Technology Program (09DJ1400601), the National Key Technology Research & Development Program (2006BAI05A09), and the Shanghai Leading Academic Discipline Project (B205). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: xingqinghesjtu@gmail.com (QX); helin@bio-x.cn (LH)

† These authors contributed equally to this work.

Introduction

Lung cancer is a worldwide leading cause of cancer-related death with a 5-year survival rate of less than 15% [1]. Several molecular markers associated with lung cancer progression have been identified, including TGF, MET, TP53, HIF1A, APC, KRAS, and EGFR [2].

Transcription factors (TFs) and pathways play critical roles in etiologies of lung cancer. For example, the transcription factor E2F-1 is over-expressed in lung cancer cell, and the level is enhanced by deregulated pRb-p53-MDM2 circuitry [3]. Transcriptional regulation analysis has shown that the promoter activity and expression level of Sp1 are regulated by Ets-1 in A549 lung cancer cells. Functional analysis of two Ets-1-binding sites in Sp1 promoter suggests that only Ets-1-binding site -413 to -404 is involved in the activation of Sp1 promoter [4]. It has also been well-documented that the expression of cPLA2 is critical for the transformed growth of lung cancer and for phorbol 12-myristate 13-acetate (PMA)-activated signal transduction pathway which is involved in enzymatic activation of cPLA2. Studies reveal that c-Jun/nucleolin and c-Jun/Sp1 complexes play an important role in PMA-regulated cPLA2a gene expression [5]. In addition, several pathways involved in lung cancer progression have been

demonstrated, such as PI3K/AKT pathway, TGF-beta signaling pathway, Wnt pathway, JAK/STAT pathway, and MAPK/ERK pathway [6,7,8,9].

High-throughput techniques in biology, such as microarray, have generated a large amount of data that can potentially provide systems-level information regarding the underlying dynamics mechanisms [10]. To extract meaningful information (TFs and pathways information) from high-throughput expression data, we employed NCA and PCA to construct and analyze the dynamic regulatory network in MGd-treated human lung cancer cells.

NCA, developed by James Liao [10], is a network structure-driven framework for deducing regulatory signal dynamics. NCA models the expression of a gene as a linear combination of the activity of each transcription factor that controls the expression of the gene. NCA makes use of the connectivity structure from transcriptional regulatory networks and a set of gene expression data to infer dynamics of transcription factor activities. NCA has been successfully applied in inferring a transcriptional regulatory network of the cytokinesis-related genes [11] and phase-specific control elements of its cell cycle in yeast [12]. In this study, we built an integrated dynamic model of the human lung cancer in response to MGd, which consisted of the calculated transcription

factor activities, transcription factor regulatory influences on each gene.

Given the complex nature of biological systems, more than one pathway may be involved in any given complex disease. Two or several pathways may interact with each other to cause the disease. This is very likely because functional important proteins may be involved in multiple pathways [13]. Therefore, besides the identification of specific pathways, we also take a further step by exploring the interaction and crosstalk between pathways that related to MGD-treated lung cancer. In this study, we used a computational approach to detect crosstalk among pathways based on a protein-protein interaction (PPI) network, the co-expressed significance of each gene pair, and a scoring scheme which is used to define a function [14].

We defined the dynamic regulated network using NCA which requires two inputs: a set of gene expression profiles and a pre-defined matrix containing the influence of each transcription factor on its estimated or identified target genes. Two outputs of NCA (predicted factor activities and regulatory influences) have added additional insights to gene expression data where the underlying regulatory network structure is partially known. In order to interpret transcription factor activities and regulation strength (influences), the correlation between TF activities and expression, hierarchical clustering were calculated. Finally, the dynamic regulated networks were constructed. Beside, PCA was used to detect the relationship among pathways.

In brief, our study aims to reveal molecular mechanism of MGD-treated human lung cancer cells from a dynamic and systematic perspective by PCA and NCA. Our results should provide new avenues for more advanced investigation into the biological role of TFs and pathways in MGD-treated human lung cancer cells.

Methods

Human lung cancer (A549) cells [15] were treated with 50 uM metal cation-containing chemotherapeutic drug motexafin gadolinium (MGd) for 4, 12, or 24 hours. Their expression profiles were compared with those of the control cells treated by 5% mannitol with the same incubation time. The detail of the samples was shown in Table 1. The limma method [16] was used to identify differentially expressed genes (DEGs) in the expression profile (GSE2189). The DEGs with fold change >1.5 and p-value <0.05 were selected for further analysis. Each selected DEG must be differentially expressed in more than one stage. In addition, 6328 regulatory relationships between 250 TFs and 2255 target genes were collected from TRED [17] and TRANSFAC [18].

In order to add more regulation relationships between TF and target genes, a total of 250 TFs and 144 DEGs were selected to be

hierarchically clustered by hcluster of R language. For each pair of TF and its target gene, only the target gene in the sub-tree of the TF-node with a coefficient larger than 0.8 (threshold $|r| > 0.8$) was selected for NCA.

Finally, 627 TF-target genes regulation relationships (containing the TF-TF interactions) were identified based on 164 TFs and 83 DEGs.

Network Component Analysis

NCA uses the standard log-linear model to approximate the relationship between levels of TFs activity and that of the target-gene expression by assuming the Hill cooperation between TFs on the promoter region of target genes. Formally,

$$\frac{E_i(t)}{E_i(0)} = \prod_{j=1}^L \left(\frac{TFA_{ij}(t)}{TFA_{ij}(0)} \right)^{CS_{ij}} \quad (1)$$

where t represents the time stage, $E_i(t)$ is the gene expression level and $TFA_{ij}(t)$ is TF j 's activities and CS_{ij} reflects the control strength of TF j on gene i .

Then, the equation (1) is linearized as (in matrix form):

$$\log[Er] = [CS] \log[TFAr] \quad (2)$$

The matrix $[Er]$ consists of elements $[Er]_{ij} = E_{ij}(t)/E_{ij}(0)$, and similarly $[TFAr]_{ij} = TFA_{ij}(t)/TFA_{ij}(0)$ represents the relative gene expression levels and TFs' activities. The dimension of $[Er]$ is $N \times M$ (N genes and M samples or conditions) while that of $[TFAr]$ is $L \times M$ (L TFs). They respectively indicate the time courses of relative gene expression levels and TFs' activities. Finally, the size of $[CS]$ is $N \times L$, which is the control strength for L TFs on each of N genes. The equation (2) above can be further simplified as:

$$[E] = [S][A] \quad (3)$$

Here, we have the strength matrix, $[S]$, which corresponds to the term of $[CS]$ in equation (2) and the TFs' activity matrix $[A]$, which is the equivalent of $\log[TFAr]$ in the equation (2), and finally, the gene expression matrix of $[E]$ corresponding to the term of $\log[Er]$ in equation (2).

Based on the above preparation, the decomposition of $[E]$ into $[S]$ and $[A]$ can be achieved by minimizing the following object function:

$$\min \|[E] - [S][A]\| \quad (4)$$

Subject to. $S \in Z_0$

In NCA, the above target function is estimated by using the bootstrap algorithm and the value of $[S]$ and $[A]$ can be normalized through a nonsingular matrix of $[X]$ according to

$$[E] = [S][A] = [S][X][X^{-1}][A] \quad (5)$$

Specifically, to guarantee uniqueness of the solution for the matrix decomposition of Eq. 4, the network topology needs to satisfy some criteria [19]: (i) the connectivity matrix $[A]$ must have full-column rank. (ii) When a node in the regulatory layer is removed along with all of the output nodes connected to it, the

Table 1. The description of samples in GSE2189.

Samples	Sample replicates	Treated time (hr)	Treatment
4_Hr_+MGd	3	4	50 uM MGd
12_Hr_+MGd	3	12	50 uM MGd
24_Hr_+MGd	3	24	50 uM MGd
4_Hr_No_MGd	3	4	5% mannitol
12_Hr_No_MGd	3	12	5% mannitol
24_Hr_No_MGd	3	24	5% mannitol

doi:10.1371/journal.pone.0031984.t001

resulting network must be characterized by a connectivity matrix that still has full-column rank. (iii) [P] must have full row rank.

The NCA algorithm implemented in MATLAB by the authors is downloadable at <http://www.seas.ucla.edu/~liao/>. In this study, we followed the manual of this package.

Pathway Enrichment Analysis

Pathway information was collected from KEGG (Kyoto Encyclopedia of Genes and Genomes), a collection of online databases dealing with genomes, enzymatic pathways, and biological chemicals [20]. DAVID [21], a high-throughput and integrated data-mining environment, was utilized to analyze gene lists derived from high-throughput genomic experiments. Significant pathways that have at least two members and p-value <0.1 were thus selected.

Pathway Crosstalk Analysis

PPI data were collected from the HPRD [22] and BIOGRID [23]. A total of 326119 unique PPI pairs were collected to construct the PPI network. Limma eBayes method [16] was used to measure the differential expression status of genes. Pearson correlated coefficient test was employed to determine the co-expressed significance of each gene pair. The above two types of values were mapped to the nodes and edges in the PPI network. The following formula was used to define a function as the combination of statistical significance of an interaction by a scoring scheme [24]. The detail could be seen in Liu et al [14].

$$S(e) = f(\text{diff}(x), \text{corr}(x,y), \text{diff}(y)) \\ = -2 \sum_{i=1}^k \log e(pi)$$

The $\text{diff}(x)$ and $\text{diff}(y)$ are differential expression assessments of gene x and gene y , respectively. $\text{Corr}(x, y)$ represents the correlation between gene x and gene y . f is a general data integration method that can handle multiple data sources differing in statistical power. Where $k=3$, p_1 and p_2 are the p-values of differential expression of two nodes, and p_3 is the p-value of their co-expression.

To define the interaction significance between pathways, we summarized all the scores of edges $S(e)$ of all non-empty overlaps. Specifically, the interaction score between two pathways was estimated according to their overlapping status of weighted pathways in the following formula:

$$C(pi, pj) = \sum_{e \in O_{ij}} S(e)$$

where P_i and P_j are two pathways and O_{ij} is their overlapping. To estimate the significance of the overlapping between different pathways, we randomly sampled 1×10^6 times of the two same size pathways in the edges of pathway network and calculated their overlapping scores. The frequency larger than a C score denoted significant interaction between P_i and P_j .

Results and Discussion

Transcription Factor Activities

The schematic of our approach was shown in Figure 1. Based on NCA method, 16 TFs were screened to construct a dynamic regulatory network. Figure 2A and 2B showed the estimated activities of the 16 TFs. Transcription factor activities clearly

showed early-, mid-, and late-phase action in response to MGd. SP1, RARA, RELA, TP53, ETS1, and SMAD3 were activated within 4 hours after the MGd was injected. SP1 activation peaked at 4 hours and HIF1A, CREB1 and SPI1 were predicted to be somewhat deactivated over 12 hours (Figure 2A). Research found that Sp1 level accumulated strongly in early stage and then declined in late stage [25] and Aryl hydrocarbon receptor in association with RelA modulates IL-6 expression in non-smoking lung cancer [26]. These are evidence which could improve the reliable of research.

The calculated transcription factor activities were compared with the gene expression data for each transcription factor (Figure 2B). TP53, SMAD3, and HIF1A showed strong positive correlation between activities and expression (correlation coefficient $r > 0.8$) (Figure 2B). However, transcription factor activities were sometimes, but not always, correlated with the gene expression of the TFs.

We also compared the significant correlation between transcription factor activities with published protein-protein interactions catalogued in the HPRD [22] and BIOGRID [23]. Interestingly, TFs known to act together showed high correlation in their activity profiles (Figure 2C). For example, the highly correlated TFs SP1 and RELA regulated their target genes together [27].

Our results also revealed several interactions between TFs (Figure 2D). The transcription factor Sp1 regulates expression of numerous genes involved in various cellular processes, and dysregulation of Sp1 is observed in many cancers and diseases [28]. Involvement of ESR1 in lung cancer has also been observed [29]. Interaction of SP1 with ESR1 has been demonstrated in breast cancer cells [30]. In addition, E2F1 and SP1 participate in cell proliferation and viability via regulating phosphocholine cytidyltransferase alpha (CCT α) [31]. Thus the predicted ESR1/Sp1 and E2F1/Sp1 interactions may suggest their regulation role in the pathogenic process. Estrogen stimulation can enhance the c-MYC-ESR1 interaction and facilitate the association of ESR1, c-MYC, and the co-activator TRRAP with these estrogen-responsive promoters, resulting in chromatin remodeling and transcription increase in breast cancer. These suggest ESR1 and c-MYC may physically interact to stabilize the ESR1-coactivator complex, thereby permitting other signal transduction pathways to fine-tune estrogen-mediated signaling networks [32]. C-Myc, an oncogene, has also been demonstrated to interact specifically with Smad3, one of the signal transducers involved in TGF- β signaling which is involved in cancer development [33]. As for NFKB1/RELA, NFKB1 or NFKB2 could bind to RELA, RELB or REL to form the NFKB family of TFs. These heterodimers participate in controlling a wide variety of genes, and are important in embryonic development, apoptosis, immune, inflammatory and stress responses. The NFKB1/RELA complex is the most abundant form of NFKB. In HeLa cells, RELA phosphorylation could result in increased transcription of NFKB target genes and inhibiting apoptosis [34].

Significant Regulation Relationships between TFs and Target Genes

In Figure 3A, target genes were hierarchically clustered with the adjusted strengths of TFs and shown with gene expression. We identified several major clusters, which were correlated to the coordinated action of TFs to regulate gene expression. We found that MYBL2, DDX11, LAMP1, ETV4, and BMP4 were regulated by MYC and ETS1. The MYC independently regulated BAZ1B, ZFP36L2, DPM2, TSC2, ZNF274, and STAT6. MYOD1 and ACP5 were regulated by the NFKB1. APEX1

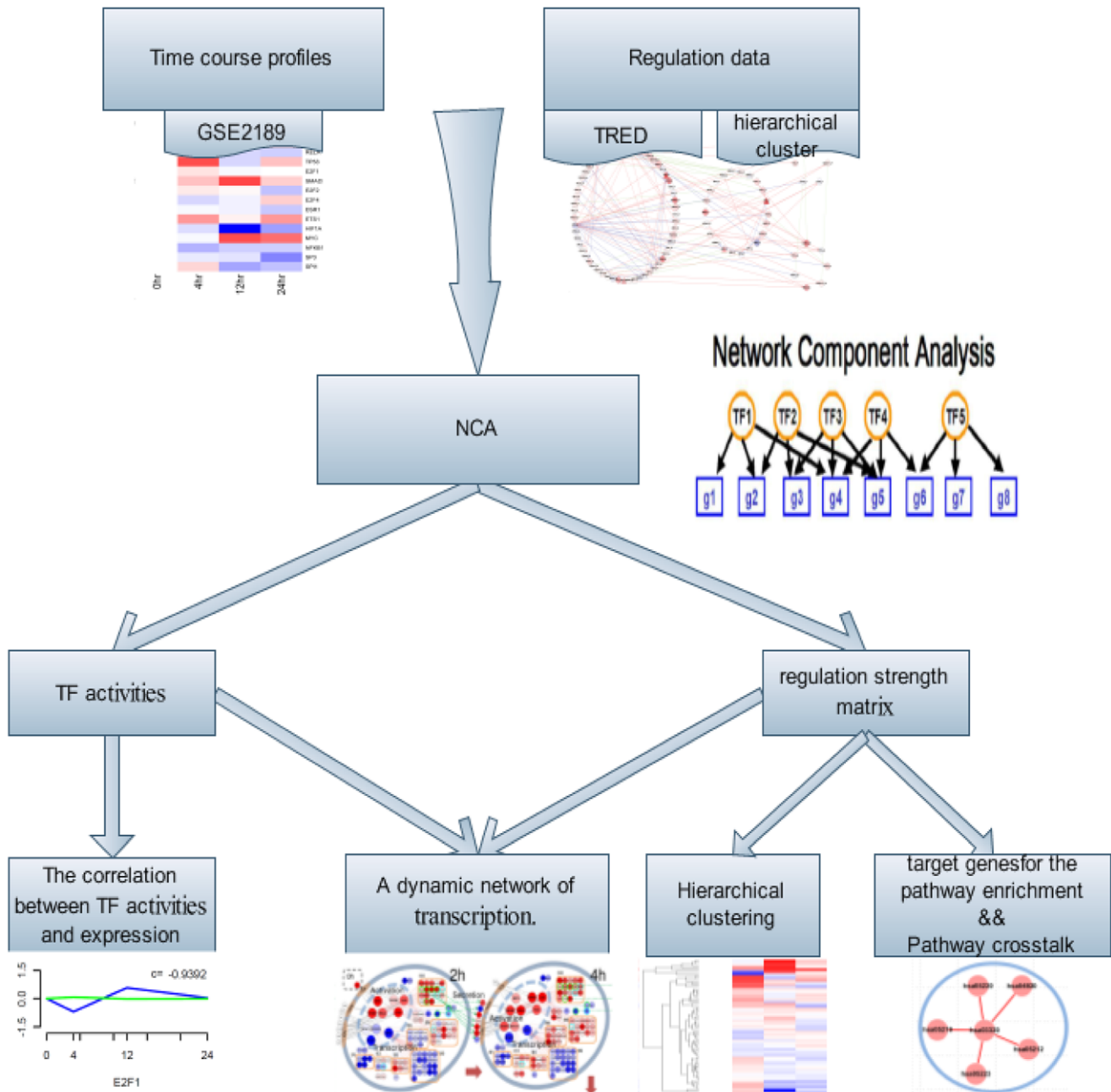


Figure 1. Flowchart of dynamic transcriptional regulatory network construction.
doi:10.1371/journal.pone.0031984.g001

and POLG2 were regulated by SP1 and CREB1. Target genes were hierarchically clustered with the expression of TFs Figure 3B. Some of them in Figure 3B haven't been emerge in Figure 3A. The clusters shown in Figure 3A suggested that we might be able to use our cluster information to discover new regulatory relationships.

We found five significant regulation relationships which were proved by the previous research: BMP4/Est1, TSC2/myc, APE1/Sp1/p53, RARA/HOXA1, and SP1/USF2. BMP4 signaling induces senescence and modulates the oncogenic phenotype of lung cancer cell [35]. BMP4 promoter has two Ets-1 binding sites, and Ets-1 activity is increased in hepatocellular carcinoma cells under hypoxic conditions. Thus over-expression of Ets-1 markedly

enhances BMP4 promoter activity [36]. In addition, BMP4 is associated with Smad and p38 MAPK pathway in lung cancer cell, which was also observed in our regulatory network [28].

MYC could directly affect transcription of tuberous sclerosis 2 (TSC2), as shown by quantitative mRNA analyses and by Myc binding to its promoter in chromatin immunoprecipitation assays. Importantly, myc-null experiments have shown that Myc acts as a strong and direct repressor for TSC2 expression because its loss results in increased TSC2 mRNA. This finding shows that regulation of TSC2 may contribute to the effects of MYC on cell proliferation and neoplastic growth [37,38].

The putative promoter region of the Apex1 gene contains CCAAT boxes and a CpG island possessing putative binding sites

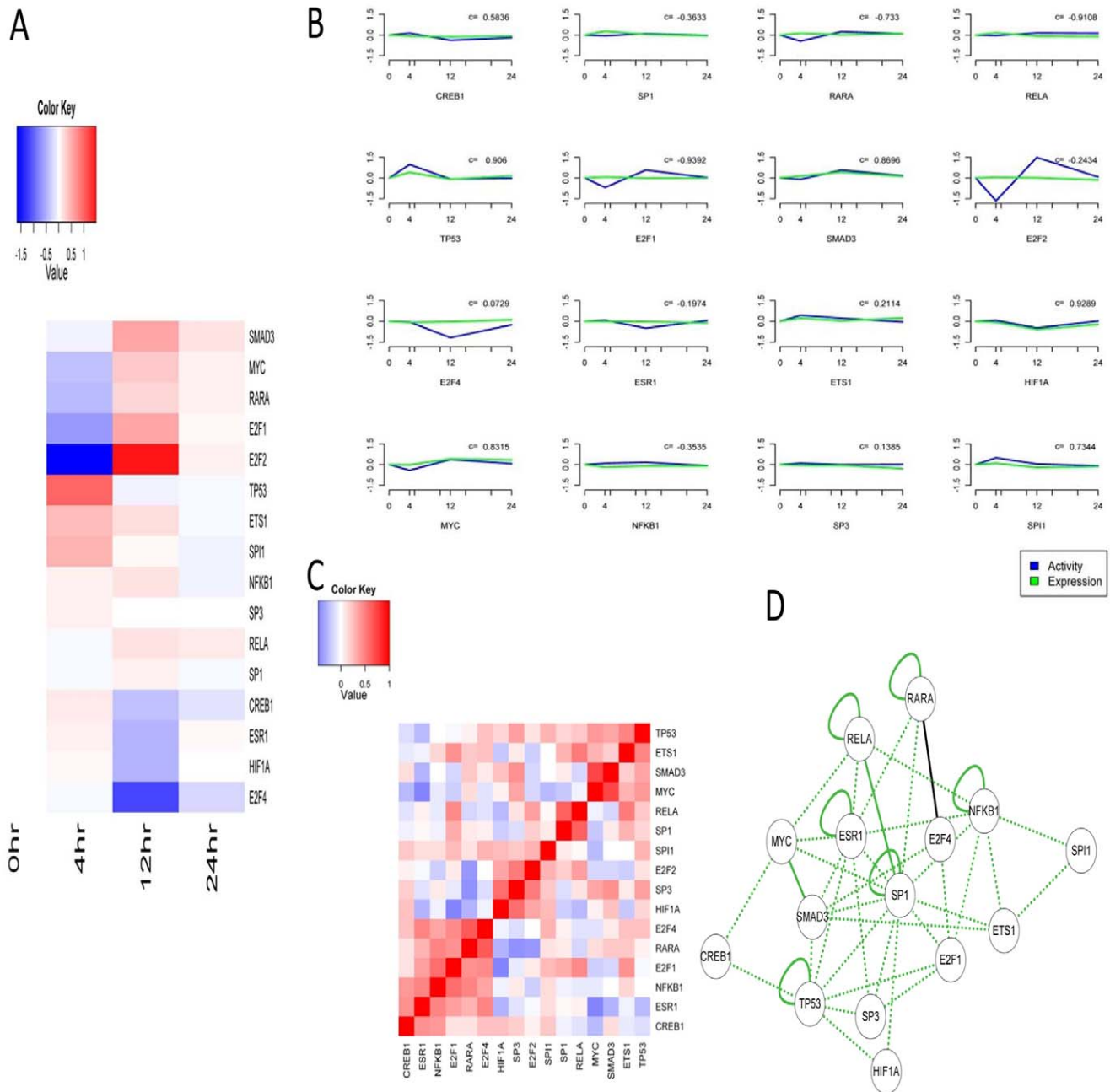


Figure 2. Transcription factor activities calculated using NCA. (A) Predicted activities of the 16 TFs used in this study. Rows represented progression in time and columns corresponded to the activities. Activities of each column were normalized to the zero time point. (B) Transcription factor activities (blue) compared with gene expression (green), with Pearson correlation coefficients noted. Both activity and expression at each time point were averages normalized to the time zero values. (C) Correlation matrix between transcription factor activities. Red represented positive correlation and blue represented negative correlation. (D) Inferred combinatorial regulation pairs of TFs. Green solid line indicated that the pair was supported by protein-protein interaction information from HPRD, BIOGRID and high correlation of their activities (>0.6). Black solid line indicated that the pair was only supported by high correlation, and a green dotted line indicated that the pair was only supported by the interaction database. doi:10.1371/journal.pone.0031984.g002

for several TFs, such as Sp1 [39]. The Sp1 site upstream of the transcription start, together with an adjacent CCAAT element, establishes a protein-DNA complex required for basal transcription of APEX1 [40]. Further study indicates that p53 provides a mechanism for the down-regulation of APE1 by interfering with Sp1 binding to the APEX1 promoter. These findings demonstrates that p53 is a negative regulator of APE1 expression in response to DNA damage [41].

RARA is one ligand dependent inducible transcription factor. The RARs family can activate gene expression directly through RA responsive elements (RAREs) localized in their target genes. Functional RAREs are currently known for only a few HOX genes, including HOXA1, HOXB and HOXC [42].

As a member of the bHLH family, USF-2 has been demonstrated to specifically bind with E-box motif A, located between -147 and -142 in the human [Arginine]vasopressin

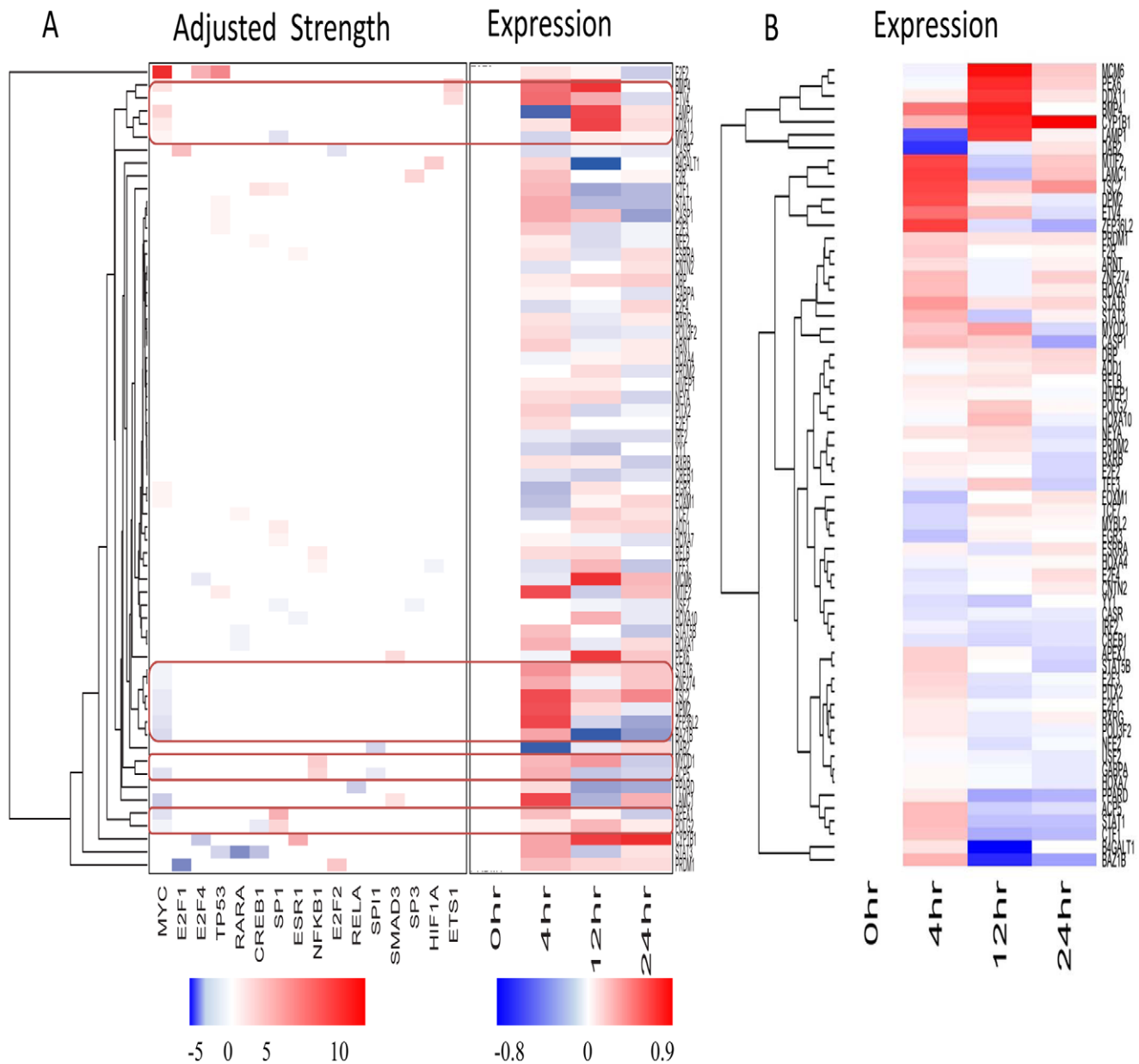


Figure 3. Hierarchical clustering in the context of a defined regulatory network. (A) The adjusted strength matrix was used for clustering, with the gene expression matrix appended. Four major clusters, which have more than three associated genes, were highlighted. In the adjusted strength matrix heatmap, white indicated a weak regulatory influence. (B) Clustering with gene expression.
doi:10.1371/journal.pone.0031984.g003

promoter to be involved in small cell lung cancer [43,44]. However, there was evidence that a physical interaction between USF2 and Sp1/Sp3 [45], suggesting USF-2 may exert important roles in lung cancer through interaction with E-Box or GC box.

Overall Regulatory Dynamics in Response to MGD

We built an integrated dynamic model of the human lung cancer in response to MGD (Figure 4), which consisted of the calculated transcription factor activities, transcription factor regulatory influences on each gene, subcellular location, and the gene expression data. During the first 4-hour period, TP53, SP1, E2F1, ETS1, SMAD3 and RELA were activated and interacted to regulate gene expression. These TFs had already affected gene expression including the genes in the Nucleus and Cytoplasm after

4 hours. SMAD3, also expressed in the Nucleus and Cytoplasm, showed peak activity at 12 hours and then at 24 hours returned to the previous 4 hours level. By contrast, E2F1 activation rapidly returned to the base level of activity.

Pathway Crosstalk Analysis

Most of the significant pathways (p -value < 0.1 using the hypergeometric test) were cancer related signaling pathways (Table 2), including Pathways in cancer, Small cell lung cancer, Non-small cell lung cancer, Pancreatic cancer, Jak-STAT signaling pathway, PPAR signaling pathway and so on. These pathways have been demonstrated involved in lung cancer in previous works. For example, BMP4 treatment (enriched in the Pathways in cancer) has been suggested to induce a senescent morphology in

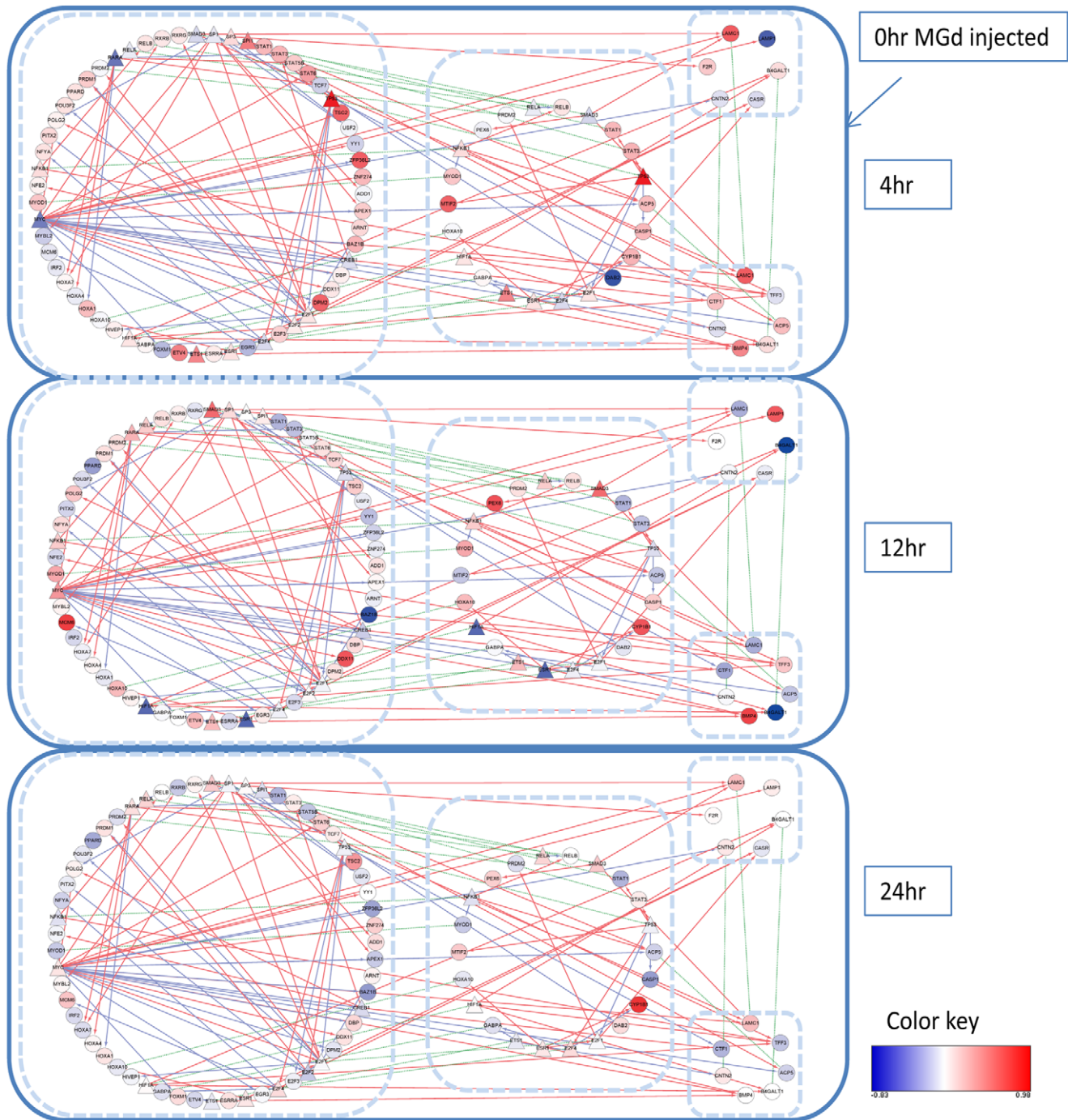


Figure 4. A dynamic network of transcription in response to MGd. Target genes were noted with circles, and TFs with triangle. The 4 subcellular locations (Nucleus; Cytoplasm; Plasma membrane; Extracellular) were grouped in 4 cycles. Green dotted lines denoted of a target gene which may transfer between two subcellular locations. Red and blue lines showed the influence of a transcription factor on a target gene. doi:10.1371/journal.pone.0031984.g004

A549 lung cancer cells [35]. Unphosphorylated signal transducers and activators of STAT6 (enriched in the Jak-STAT signaling pathway) may transcriptionally up-regulate cyclooxygenase-2 expression and protect against apoptosis in NSCLC cells [46]. PPAR γ has been suggested to modulate the proliferation and apoptosis of lung cancer cell through interaction with its ligand. PPAR γ expression is found higher in lung cancer cell patients when compared with normal surrounding tissue. The treatment of

lung adenocarcinoma cells (A549) with troglitazone (a PPAR γ ligand) can enhance PPAR γ transcriptional activity and induce a dose-dependent inhibition of A549 cell growth [47,48,49]. In brief, activation of PPAR γ impedes lung tumor progression and PPAR γ ligands may serve as potential therapeutic agents for lung cancer. In this study, we found PPAR signaling pathway was an important pathway in response to MGd-treatment, suggesting MGd may be one potential PPAR γ ligand as troglitazone.

Table 2. Significant pathways.

Pathway ID	Description	P-value
hsa05200	Pathways in cancer	2.97E-07
hsa05222	Small cell lung cancer	1.55E-04
hsa05223	Non-small cell lung cancer	3.28E-04
hsa05212	Pancreatic cancer	9.85E-04
hsa05215	Prostate cancer	0.002168
hsa05221	Acute myeloid leukemia	0.005524
hsa04110	Cell cycle	0.007346
hsa05220	Chronic myeloid leukemia	0.011247
hsa05216	Thyroid cancer	0.014013
hsa04630	Jak-STAT signaling pathway	0.015367
hsa05219	Bladder cancer	0.028249
hsa05214	Glioma	0.059075
hsa04920	Adipocytokine signaling pathway	0.065857
hsa03320	PPAR signaling pathway	0.069342
hsa05218	Melanoma	0.072888

doi:10.1371/journal.pone.0031984.t002

We further analyzed the pathway's interactions and calculated a C score of each pair of pathways. PPAR signaling pathway (hsa03320) was found cross-talking with the pathway of Non-small cell lung cancer (hsa05223; p-value = 0.056135), Pancreatic cancer (hsa05212; p-value = 0.056145), Bladder cancer (hsa05219; p-value = 0.056165), Adipocytokine signaling pathway (hsa04920; p-value = 0.056214), and Chronic myeloid leukemia (hsa05220; p-value = 0.056214) after 4-hour MGD treatment (Figure 5). But this crosstalking was not significant at the 12-hour and 24-hour stages with the p-value >0.1.

Overall, we suggest PPAR signaling pathway (hsa03320) plays an important role in the pathways crosstalk.

There was evident of interaction relationship between PPAR signaling pathway and Adipocytokines signaling pathway in previous study [50,51]. Among them, adipocytokines, secreted from adipocytes, such as tumor necrosis factor-alpha (TNF- α), plasminogen activator inhibitor type 1 (PAI-1), interleukin 6 (IL-6), leptin, resistin, and adiponectin, play a significant role in normal metabolic homeostasis and in the development of several diseases. Leptin could be decreased regulation by PPAR- γ agonists. PPAR- γ and liver receptor homolog-1 (LRH-1) play significant roles in adiponectin transcriptional activation by means of PPRE and LRH-RE in its promoter [50,51].

However, there are still some limitations in our research. Our study is based on an assumption that mRNA expression levels are controlled entirely by transcriptional regulation. However, mRNA stability is also a less informative factor to the mRNA expression levels [52,53]. If an mRNA whose expression level is determined by mRNA stability can also express in the control cells, we can exclude the influence of mRNA stability by comparison when their degradation rates are similar. If such an mRNA cannot express in the control cells, we ignore the mRNA stability. This may bring systematic errors to this study. In addition, the quality and quantity of Protein-protein interaction (PPI) data is one of the problems for the PCA. PCA was based on a PPI interaction data [24]. Protein-protein interactions provide valuable information about how genes carry out their biological functions. It is expected that protein-protein interaction data information will be widely accessible in the near future by using various experiment methods.

Conclusion

We managed to interpret the molecular mechanism of lung cancer from a systematic and dynamic perspective by NCA. We took the control strength (only as positive or negative) as the regulatory relationships between TFs and their target genes (including TFs), and the TFs activities was substituted for their

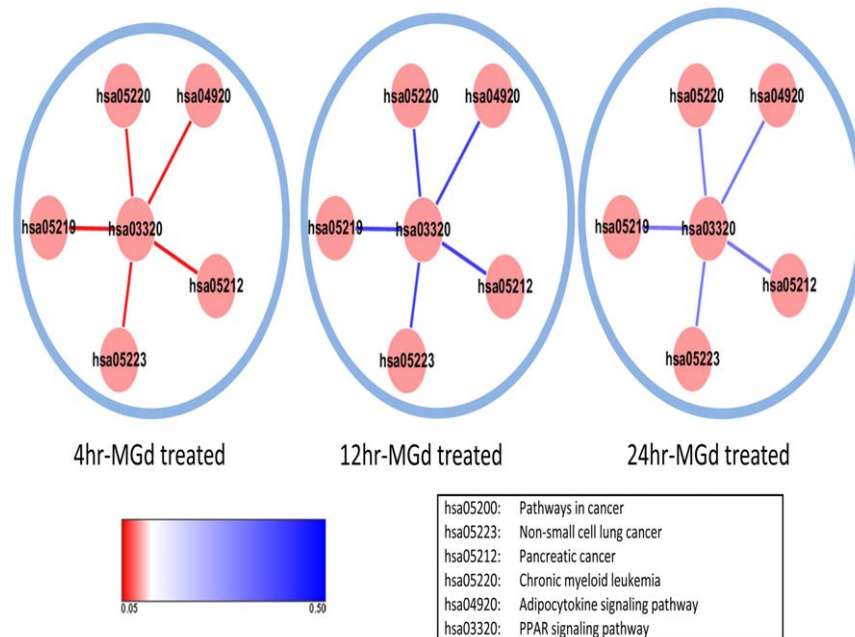


Figure 5. Dynamic of pathway-crosstalk. The red line indicated the p-value of cross-talking between two pathways less than 0.1. The blue line indicated the cross-talking was not significant with the p-value large than 0.3.

doi:10.1371/journal.pone.0031984.g005

gene expression to construct the dynamic network. Using NCA, the significant TFs and their target genes were detected, the control strength of TFs to their target genes was recalculated, and the activities of the TFs were estimated.

NCA and PCA methods were applied to explore the transcription response mechanism in MGD-treated human lung cancer cells based on the assumption that lung cancer is a contextual attribute of distinct patterns of interactions between multiple elements. The results identified a set of key TFs, target genes for these TFs and signaling pathways involved in regulatory networks. Through the activity of TFs, we found that transcription factor activities clearly showed early-, mid-, and late-phase action in response to MGD. We also identified several major clusters, which were correlated to the coordinated action of TFs to regulate gene expression. Besides, pathway-crosstalk analysis indicated there was an interaction relationship between PPAR signaling pathway and Adipocytokines signaling pathway in our study. Finally, an integrated dynamic model of the human lung cancer was built in response to MGD (Figure 4), which consisted of the calculated transcription factor activities, transcription factor

regulatory influences on each gene, subcellular location, and the gene expression data.

The development of new high-throughput technologies greatly produces great amounts of biology data. Then how to mine meaning information from the data become necessary. Our studies revealed that NCA and PCA could be successfully applied for inferring the transcriptional regulatory network of MGD-treated human lung cancer.

Acknowledgments

We wish to express our warm and sincere thanks to Fenghe Information and Technology Inc. Their ideas and help gave a valuable added dimension to our research.

Author Contributions

Conceived and designed the experiments: LYS LH QHX. Performed the experiments: LYS LSW ZYW. Analyzed the data: YYX YW KFT. Contributed reagents/materials/analysis tools: YL GYF. Wrote the paper: LYS LSW.

References

- Boyle P, Chapman C, Holdenrieder S, Murray A, Robertson C, et al. (2011) Clinical validation of an autoantibody test for lung cancer. *Annals of Oncology* 22: 383.
- Wang L, Xiong Y, Sun Y, Fang Z, Li L, et al. (2010) HLungDB: an integrated database of human lung cancer research. *Nucleic Acids Res* 38: D665–669.
- Gorgoulis VG, Zacharatos P, Mariatos G, Kotsinas A, Bouda M, et al. (2002) Transcription factor E2F-1 acts as a growth-promoting factor and is associated with adverse prognosis in non-small cell lung carcinomas. *The Journal of Pathology* 198: 142–156.
- Sato T, Furukawa K (2007) Sequential action of Ets-1 and Sp1 in the activation of the human beta-1,4-galactosyltransferase V gene involved in abnormal glycosylation characteristic of cancer cells. *J Biol Chem* 282: 27702–27712.
- Tsou JH, Chang KY, Wang WC, Tseng JT, Su WC, et al. (2008) Nucleolin regulates c-Jun/Sp1-dependent transcriptional activation of cPLA2alpha in phorbol ester-treated non-small cell lung cancer A549 cells. *Nucleic Acids Res* 36: 217–227.
- Nguyen DX, Chiang AC, Zhang XH, Kim JY, Kris MG, et al. (2009) WNT/TCF signaling through LEF1 and HOXB9 mediates lung adenocarcinoma metastasis. *Cell* 138: 51–62.
- Kim TK, Mo EK, Yoo CG, Lee CT, Han SK, et al. (2001) Alteration of cell growth and morphology by overexpression of transforming growth factor beta type II receptor in human lung adenocarcinoma cells. *Lung Cancer* 31: 181–191.
- Borcuzak AC, Kim HK, Yegen HA, Friedman RA, Powell CA (2005) Lung adenocarcinoma global profiling identifies type II transforming growth factor-beta receptor as a repressor of invasiveness. *Am J Respir Crit Care Med* 172: 729–737.
- Ji H, Wang Z, Perera SA, Li D, Liang MC, et al. (2007) Mutations in BRAF and KRAS converge on activation of the mitogen-activated protein kinase pathway in lung cancer mouse models. *Cancer Res* 67: 4933–4939.
- Liao JC, Boscolo R, Yang YL, Tran LM, Sabatti C, et al. (2003) Network component analysis: reconstruction of regulatory signals in biological systems. *Proc Natl Acad Sci U S A* 100: 15522–15527.
- Chen SF, Juang YL, Chou WK, Lai JM, Huang CYF, et al. (2009) Inferring a transcriptional regulatory network of the cytokinesis-related genes by network component analysis. *BMC Systems Biology* 3: 110.
- Bushel PR, Heard NA, Gutman R, Liu L, Peddada SD, et al. (2009) Dissecting the fission yeast regulatory network reveals phase-specific control elements of its cell cycle. *BMC Systems Biology* 3: 93.
- Li Y, Agarwal P, Rajagopalan D (2008) A global pathway crosstalk network. *Bioinformatics* 24: 1442–1447.
- Liu ZP, Wang Y, Zhang XS, Chen L (2010) Identifying dysfunctional crosstalk of pathways in various regions of Alzheimer's disease brains. *BMC systems biology* 4: S11.
- Magda D, Lecane P, Miller RA, Lepp C, Miles D, et al. (2005) Motexafin gadolinium disrupts zinc metabolism in human cancer cell lines. *Cancer Res* 65: 3837–3845.
- Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3: Article3.
- Jiang C, Xuan Z, Zhao F, Zhang MQ (2007) TRED: a transcriptional regulatory element database, new entries and other development. *Nucleic acids research* 35: D137–140.
- Matys V, Fricke E, Geffers R, Goessling E, Haubrock M, et al. (2003) TRANSFAC(R): transcriptional regulation, from patterns to profiles. *Nucleic acids research* 31: 374.
- Liao JC, Boscolo R, Yang YL, Tran LM, Sabatti C, et al. (2003) Network component analysis: Reconstruction of regulatory signals in biological systems. *Proceedings of the National Academy of Sciences of the United States of America* 100: 15522–15527.
- Kanehisa M (2002) The KEGG database. *Novartis Found Symp* 247: 91–101; discussion 101–103, 119–128, 244–152.
- Da Wei Huang BTS, Lempicki RA (2008) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* 4: 44–57.
- Keshava Prasad T, Goel R, Kandasamy K, Keerthikumar S, Kumar S, et al. (2009) Human protein reference database; 2009 update. *Nucleic acids research* 37: D767.
- Stark C, Breitkreutz BJ, Chatr-aryamontri A, Boucher L, Oughtred R, et al. (2011) The BioGRID Interaction Database: 2011 update. *Nucleic acids research* 39: D698.
- Liu ZP, Wang Y, Zhang XS, Chen L (2010) Identifying dysfunctional crosstalk of pathways in various regions of Alzheimer's disease brains. *BMC Syst Biol* 4 Suppl 2: S11.
- Hsu TI, Wang MC, Chen SY, Yeh YM, Su WC, et al. (2011) Sp1 expression regulates lung tumor progression. *Oncogene*.
- Chen PH, Chang H, Chang JT, Lin P (2011) Aryl hydrocarbon receptor in association with RelA modulates IL-6 expression in non-smoking lung cancer. *Oncogene*.
- Perkins ND, Agranoff AB, Pascal E, Nabel GJ (1994) An interaction between the DNA-binding domains of RelA(p65) and Sp1 mediates human immunodeficiency virus gene activation. *Mol Cell Biol* 14: 6570–6583.
- Su D, Peng X, Zhu S, Huang Y, Dong Z, et al. (2011) Role of p38 MAPK pathway in BMP4-mediated Smad-dependent premature senescence in lung cancer cells. *Biochem J* 433: 333–343.
- Lin Q, Geng J, Ma K, Yu J, Sun J, et al. (2009) RASSF1A, APC, ESR1, ABCB1 and HOXC9, but not p16INK4A, DAPK1, PTEN and MT1G genes were frequently methylated in the stage I non-small cell lung cancer in China. *J Cancer Res Clin Oncol* 135: 1675–1684.
- Kim K, Barhoumi R, Burghardt R, Safe S (2005) Analysis of estrogen receptor alpha-Sp1 interactions in breast cancer cells by fluorescence resonance energy transfer. *Mol Endocrinol* 19: 843–854.
- Elena C, Banchio C (2010) Specific interaction between E2F1 and Sp1 regulates the expression of murine CTP: phosphocholine cytidyltransferase alpha during the S phase. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* 1801: 537–546.
- Cheng AS, Jin VX, Fan M, Smith LT, Liyanarachchi S, et al. (2006) Combinatorial analysis of transcription factor partners reveals recruitment of c-MYC to estrogen receptor-alpha responsive promoters. *Mol Cell* 21: 393–404.
- Feng XH, Liang YY, Liang M, Zhai W, Lin X (2002) Direct interaction of c-Myc with Smad2 and Smad3 to inhibit TGF-beta-mediated induction of the CDK inhibitor p15(Ink4B). *Mol Cell* 9: 133–143.
- Hansson S, Chen Y, Brodzski J, Chen M, Hernandez-Andrade E, et al. (2006) Gene expression profiling of human placentas from preeclamptic and normotensive pregnancies. *Molecular human reproduction* 12: 169.

35. Buckley S, Shi W, Driscoll B, Ferrario A, Anderson K, et al. (2004) BMP4 signaling induces senescence and modulates the oncogenic phenotype of A549 lung adenocarcinoma cells. *Am J Physiol Lung Cell Mol Physiol* 286: L81–86.
36. Maegdefrau U, Amann T, Winklmeier A, Braig S, Schubert T, et al. (2009) Bone morphogenetic protein 4 is induced in hepatocellular carcinoma by hypoxia and promotes tumour progression. *J Pathol* 218: 520–529.
37. Ravitz MJ, Chen L, Lynch M, Schmidt EV (2007) c-myc Repression of TSC2 contributes to control of translation initiation and Myc-induced transformation. *Cancer Res* 67: 11209–11217.
38. Schmidt EV, Ravitz MJ, Chen L, Lynch M (2009) Growth controls connect: interactions between c-myc and the tuberous sclerosis complex-mTOR pathway. *Cell Cycle* 8: 1344–1351.
39. Yao M, Akiyama K, Tan Y, Sarker AH, Ikeda S, et al. (1999) Genomic structure of the rat major AP endonuclease gene (Apex) with an adjacent putative O-sialoglycoprotease gene (Prsmg1/Gcp11) and a processed Apex pseudogene (Apexp1). *Acta Med Okayama* 53: 245–252.
40. Fung H, Bennett RA, Demple B (2001) Key role of a downstream specificity protein 1 site in cell cycle-regulated transcription of the AP endonuclease gene APE1/APEX in NIH3T3 cells. *J Biol Chem* 276: 42011–42017.
41. Zaky A, Busso C, Izumi T, Chattopadhyay R, Bassiouny A, et al. (2008) Regulation of the human AP-endonuclease (APE1/Ref-1) expression by the tumor suppressor p53 in response to DNA damage. *Nucleic Acids Res* 36: 1555–1566.
42. Flagiello D, Gibaud A, Dutrillaux B, Poupon MF, Malfoy B (1997) Distinct patterns of all-trans retinoic acid dependent expression of HOXB and HOXC homeogenes in human embryonal and small-cell lung carcinoma cell lines. *FEBS letters* 415: 263–267.
43. Coulson JM, Fiskerstrand CE, Woll PJ, Quinn JP (1999) E-box motifs within the human vasopressin gene promoter contribute to a major enhancer in small-cell lung cancer. *Biochemical Journal* 344: 961.
44. Ocejó-García M, Baokbah TA, Ashurst HL, Cowlshaw D, Soomro I, et al. (2005) Roles for USF-2 in lung cancer proliferation and bronchial carcinogenesis. *J Pathol* 206: 151–159.
45. Liu HC, Shen JT, Augustin LB, Ko JL, Loh HH (1999) Transcriptional regulation of mouse δ -opioid receptor gene. *Journal of Biological Chemistry* 274: 23617.
46. Cui X, Zhang L, Luo J, Rajasekaran A, Hazra S, et al. (2007) Unphosphorylated STAT6 contributes to constitutive cyclooxygenase-2 expression in human non-small cell lung cancer. *Oncogene* 26: 4253–4260.
47. Keshamouni VG, Reddy RC, Arenberg DA, Joel B, Thannickal VJ, et al. (2004) Peroxisome proliferator-activated receptor- γ activation inhibits tumor progression in non-small-cell lung cancer. *Oncogene* 23: 100–108.
48. Chang TH, Szabo E (2000) Induction of differentiation and apoptosis by ligands of peroxisome proliferator-activated receptor γ in non-small cell lung cancer. *Cancer Research* 60: 1129.
49. Han SW, Roman J (2006) Rosiglitazone suppresses human lung carcinoma cell growth through PPAR γ -dependent and PPAR γ -independent signal pathways. *Molecular Cancer Therapeutics* 5: 430.
50. Vettor R, Milan G, Rossato M, Federspil G (2005) Review article: adipocytokines and insulin resistance. *Alimentary pharmacology & therapeutics* 22: 3–10.
51. Choi KC, Ryu OH, Lee KW, Kim HY, Seo JA, et al. (2005) Effect of PPAR-alpha and -gamma agonist on the expression of visfatin, adiponectin, and TNF-alpha in visceral fat of OLETF rats. *Biochem Biophys Res Commun* 336: 747–753.
52. Garcia-Martinez J, Aranda A, Perez-Ortin JE (2004) Genomic run-on evaluates transcription rates for all yeast genes and identifies gene regulatory mechanisms. *Mol Cell* 15: 303–313.
53. Hayles B, Yellaboina S, Wang D (2010) Comparing transcription rate and mRNA abundance as parameters for biochemical pathway and network analysis. *PLoS One* 5: e9908.