RESEARCH ARTICLE

Shortest Path Analyses in the Protein-Protein Interaction Network of NGAL (Neutrophil Gelatinase-associated Lipocalin) Overexpression in Esophageal Squamous Cell Carcinoma

Ze-Peng Du[1&], Bing-Li Wu[2&*], Shao-Hong Wang[1], Jin-Hui Shen[1], Xuan-Hao Lin[1], Chun-Peng Zheng[3], Zhi-Yong Wu[3], Xiao-Yang Qiu[1], Xiao-Fen Zhan[1], Li-Yan Xu[4*], En-Min Li[2]

Abstract

   NGAL (neutrophil gelatinase-associated lipocalin) is a novel cancer-related protein involves multiple functions in many cancers and other diseases. We previously overexpressed NGAL to analyze its role in esophageal squamous cell carcinoma (ESCC). In this study, a protein-protein interaction (PPI) was constructed and the shortest paths from NGAL to transcription factors in the network were analyzed. We found 28 shortest paths from NGAL to RELA, most of them obeying the principle of extracellular to cytoplasm, then nucleus. These shortest paths were also prioritized according to their normalized intensity from the microarray by the order of interaction cascades. A systems approach was developed in this study by linking differentially expressed genes with publicly available PPI data, Gene Ontology and subcellular localizaton for the integrated analyses. These shortest paths from NGAL to DEG transcription factors or other transcription factors in the PPI network provide important clues for future experimental identification of new pathways.

Keywords: NGAL - protein-protein interaction network - shortest path - subcellular localization

Introduction

   NGAL (lipocalin2) or NGAL (neutrophil gelatinase-associated lipocalin), is a member of the lipocalin superfamily, which contains more than twenty members (Flower, 1996). Increased level of NGAL has been detected in the patients with various types of cancer (i.e., brain, breast, colon, ovarian, pancreatic and prostate) (Chakraborty et al., 2012). NGAL plays important role in cancer invasion and metastasis by different mechanisms. NGAL forms heterodimer with matrix metalloproteinase-9 (MMP-9) through a disulfide bond, thus stabilizes MMP-9 and increases its ability to degrade the extracellular matrix, thereby promoting metastasis (Yan et al., 2001). NGAL/MMP-9 complex is also associated with CD44, promoting the cleavage of E-Cadherin inducing epithelial-mesenchymal transition (EMT) (Abecassis et al., 2003). NGAL can tightly bind to bacterial siderophore, regulating innate immunity and inflammation by sequestering iron (Yang et al., 2002). Our previous studies showed NGAL was upregulated in esophageal squamous cell carcinoma (ESCC) cell lines and clinical samples, acting as an independent prognostic factor of (Zhang et al., 2007; Du

et al., 2011).

   However, the functional roles of NGAL in various cancer cells are controversial. NGAL facilitates gastrointestinal mucosal regeneration by promoting cell motility and invasion and decreasing E-cadherin mediated cell-cell adhesion in colon cancer (Hu et al., 2009). On the other hand, NGAL can act as a tumor suppressor gene. NGAL reduces adhesion/invasion partly by suppressing FAK activation and inhibits angiogenesis partly by blocking VEGF production in pancreatic cancer (Tong et al., 2008).

   To explore the biological role of NGAL in ESCC, we have previously overexpressed NGAL in EC109 ESCC cell line and applied Agilent whole genome oligo microarray to analyze its mRNA expression profile. We obtained hundreds of differentially expressed genes (DEGs) from NGAL overexpressed cell comparing with its control. It is interesting to understand why a gene overexpression or knockdown could cause a wide range alternation of expression profile. In this study, we applied shortest path algorithm to identify possible paths from NGAL to the transcription factor in the protein-protein interaction (PPI) network.

*[1]Department of Pathology, [3]Department of Oncology Surgery, Shantou Central Hospital, Affiliated Shantou Hospital of Sun Yat-sen University, [2]Department of Biochemistry and Molecular Biology, [4]Institute of Oncologic Pathology, Shantou University Medical College, Shantou, Guangdong Province, China &Equal contributors *For correspondence: blwu@stu.edu.cn; lyxu@stu.edu.cn*

Materials and Methods

Differentially expressed genes

We have submitted the microarray data of NGAL overpression in ESCC to GEO under accession number of GSE57630. Briefly, NGAL was overexpressed by pcDNA3.0 plasmid in EC109 ESCC cell line with an empty plasmid as a control. The mRNA expression profiles were analyzed by Agilent whole genome oligo microarray (Agilent Technologies, USA) according its manual. The raw data was treated by LOWESS (locally weighted scatterplot smoothing) normalization and log transformation. The differentially expressed genes (DEGs) were defined using a 2-fold threshold.

PPI sub-network construction

The newest versions of human protein-protein interaction dataset were available from HPRD, BioGRID and IntAct, respectively (Goel et al., 2011; Kerrien et al., 2012; Chatr-Aryamontri et al., 2013). These interactions are derived from literatures of experimentally validation and reliable for widely application in disease researches combined with human PPI network. The integration data of three dataset were used as the parental PPI network in this study. The open software Cytoscape was applied for visualization and analysis of PPI network, which provides various plugins for different analyses (Smoot et al., 2011). In Cytoscape, PPI network is illustrated as a graph with the nodes represent the proteins and the edges represent their interactions. The total DEGs of NGAL overexpression were mapped to the parental PPI network and generated DEG PPI sub-networks. To confine the interactions only close to the DEGs and reduce the unrelated nodes, only the directly interacting proteins of DEGs and their all edges were extracted and showed by the "Spring Embedded" layout in Cytoscape.

Identification transcription factor in PPI network

DAVID Bioinformatics was applied to annotate and find the DEGs with transcription ability. The known transcription factors were collected from TRANSFAC database and TRED database (Matys et al., 2003; Jiang et al., 2007). The curated transcription factor list was compared with protein members of DEG PPI sub-network to identify the transcription factors presented in the sub-network, in turn whose transcription targets were retrieved from TRED.

Shortest path analyses

In a network, shortest path measures the least nodes need to pass through from one node to another. The Dijkstra's algorithm in the igraph R program was applied to find the shortest path between NGAL protein and interested transcription factors in the DEG PPI sub-network (Dijkstra 1959; Csardi et al., 2006). These shortest paths were prioritized listed according to the normalized intensity of genes in NGAL overexpression microarray by their order in the signaling cascade. We assumed the normalized intensity of the probes in the microarray presented their mRNA levels, and consequently protein levels to a certain extent. These shortest paths were prioritized listed according to the normalized intensity of genes in NGAL overexpression microarray by their order in the signaling cascade. For example, the four NGAL-interacting proteins were ranged as MMP9, MMP2, HGF and LRP2 according their normalized intensity. The MMP9-interacting proteins were also arranged according their normalized intensity.

Subcellular localization of proteins in PPI network

The subcellular localization for the proteins in the DEG PPI sub-network were extracted by a custom R program from the latest Gene Ontology (GO) annotation of Homo sapiens which was released on 4/15/2014 at http://www. geneontology.org/GO.downloads.annotations.shtml. For the protein is annotated with multiple localizations, especially the proteins localizing in the nucleus (e.g. cytoplasm and nucleus), its localizations were integrated nto cytoplasm/nucleus in this study. Cerebral plugin was applied to re-distribute the proteins in Cytoscape according to their subcellular localizations without changing their interactions, which provides a pathway-like diagram (Barsky et al., 2007).

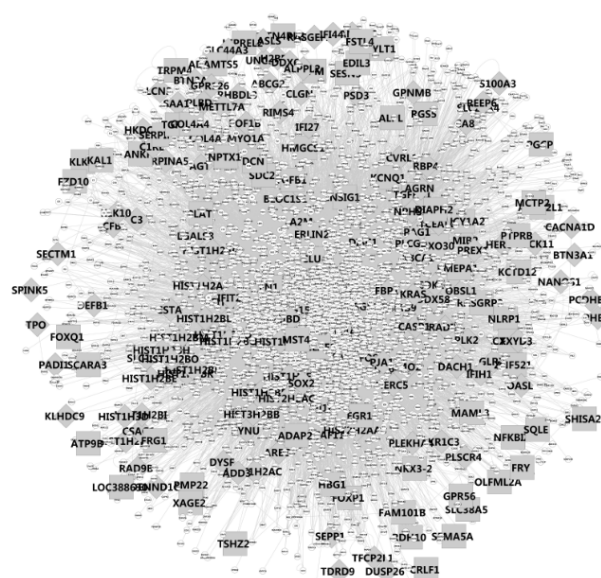**Figure 1. The Heat Map of mRNA Expression Profile**

Figure 2. PPI Sub-Networks were Generated by Mapping DEGs to HPRD, BioGRID and InAct Parent PPI Network. Nodes were labeled by different types to indicate the expression trend of proteins. The square gray nodes represent proteins encoded downregulated genes and the diamond gray nodes represent the upregulated genes. The other interacting proteins without significantly differentially expression were represented as white nodes.

Results

PPI sub-network of DEGs derived from NGAL overexpression

The cluster of microarray was shown in Figure 1. We obtained 167 upregulated genes and 96 downregulated genes from the mRNA expression profile following NGAL overexpression using a 2-fold threshold (Supplementary File 1). We assumed that a full screen of their interactions with other proteins would provide important clues for the functions of these DEGs. The DEGs were mapped to the parental PPI network from well-known HPRD,

BioGRID and InAct database to generate the specific DEG sub-network. The DEG PPI sub-network was composed of 3423 nodes and 54036 edges, including 200 DEGs (Figure 2). This DEG sub-network suggested that the overexpression of NGAL has greatly disturbed the PPI network in ESCC as hundreds of DEGs interacts with other thousands of proteins to enlarge the biological consequences of its overexpression.

Transcription factors in the DEGs and PPI sub-network

We assumed the transcription factors play critical roles in the alternation of mRNA expression profile

Table 1. The Differentially Expressed Genes Annotated with Transcription Regulation Ability

Gene	GO Biological process annotation
Upregulated DEG	
AGRN	GO:0006355~regulation of transcription, DNA-dependent,
	GO:0006357~regulation of transcription from RNA polymerase II promoter,
	GO:0045449~regulation of transcription
DACH1	GO:0006350~transcription, GO:0045449~regulation of transcription
FOXN4	GO:0006350~transcription, GO:0051252~regulation of RNA metabolic process
SMAD9	GO:0006350~transcription, GO:0045941~positive regulation of transcription
SOX2	GO:0006325~chromatin organization, GO:0006350~transcription,
	GO:0006357~regulation of transcription from RNA polymerase II promoter
TAF11	GO:0006350~transcription, GO:0006352~transcription initiation,
	GO:0006354~RNA elongation, GO:0032774~RNA biosynthetic process
TCEAL2	GO:0006350~transcription, GO:0045449~regulation of transcription
TFCP2L1	GO:0000122~negative regulation of transcription from RNA polymerase II promoter,
	GO:0006355~regulation of transcription, DNA-dependent,
	GO:0006357~regulation of transcription from RNA polymerase II promoter,
	GO:0045449~regulation of transcription
TP63	GO:0006350~transcription, GO:0006355~regulation of transcription, DNA-dependent,
	GO:0006357~regulation of transcription from RNA polymerase II promoter,
	GO:0045941~positive regulation of transcription,
	GO:0045944~positive regulation of transcription from RNA polymerase II promoter
Downregulated DEG	
BCLAF1	GO:0006350~transcription
EGR1	GO:0006355~regulation of transcription, DNA-dependent,
	GO:0006357~regulation of transcription from RNA polymerase II promoter
FOS	GO:0006355~regulation of transcription, DNA-dependent,
	GO:0006357~regulation of transcription from RNA polymerase II promoter,
	GO:0045449~regulation of transcription,
	GO:0045944~positive regulation of transcription from RNA polymerase II promoter
FOXP1	GO:0000122~negative regulation of transcription from RNA polymerase II promoter,
	GO:0045449~regulation of transcription, GO:0045892~negative regulation of transcription, DNA-dependent
FOXQ1	GO:0006350~transcription, GO:0006355~regulation of transcription, DNA-dependent,
	GO:0051252~regulation of RNA metabolic process
FRY	GO:0006350~transcription, GO:0045449~regulation of transcription
IFI16	GO:0006350~transcription, GO:0006355~regulation of transcription, DNA-dependent,
	GO:0051252~regulation of RNA metabolic process
MAML3	GO:0006355~regulation of transcription, DNA-dependent,
	GO:0006357~regulation of transcription from RNA polymerase II promoter,
	GO:0045449~regulation of transcription,
	GO:0045944~positive regulation of transcription from RNA polymerase II promoter
NFKBIZ	GO:0006350~transcription, GO:0045449~regulation of transcription
NKX3-2	GO:0006351~transcription, DNA-dependent, GO:0006355~regulation of transcription, DNA-dependent
TGFB1	GO:0051090~regulation of transcription factor activity,
	GO:0051091~positive regulation of transcription factor activity,
	GO:0051092~positive regulation of NF-kappaB transcription factor activity
TSHZ2	GO:0006350~transcription, GO:0006355~regulation of transcription, DNA-dependent,
	GO:0045449~regulation of transcription, GO:0051252~regulation of RNA metabolic process
ZNF521	GO:0006350~transcription, GO:0045449~regulation of transcription
ZNF783	GO:0006350~transcription, GO:0006355~regulation of transcription, DNA-dependent,
	GO:0045449~regulation of transcription, GO:0051252~regulation of RNA metabolic process

Table 2. Shortest Paths from NGAL to RELA

| No.	The protein members of the path
1	NGAL → MMP9 → THBS1 → COL2A1 → RELA
2	NGAL → MMP9 → AREG → UBC → RELA
3	NGAL → MMP9 → TGFB1 → UBC → RELA
4	NGAL → MMP9 → TGFB1 → SMAD3 → RELA
5	NGAL → MMP9 → TGFB1 → SNIP1 → RELA
6	NGAL → MMP9 → EPHB2 → FOS → RELA
7	NGAL → MMP9 → COL4A3 → UBC → RELA
8	NGAL → MMP9 → COL4A4 → UBC → RELA
9	NGAL → HGF → SDC2 → UBC → RELA
10	NGAL → HGF → SDC2 → CSNK2A2 → RELA
11	NGAL → MMP2 → A2M → TP63 → RELA
12	NGAL → MMP2 → A2M → TGM2 → RELA
13	NGAL → MMP2 → THBS1 → UBC → RELA
14	NGAL → MMP2 → THBS1 → COL2A1 → RELA
15	NGAL → MMP2 → DCN → BRCA1 → RELA
16	NGAL → MMP2 → TGFB1 → UBC → RELA
17	NGAL → MMP2 → TGFB1 → SNIP1 → RELA
18	NGAL → MMP2 → EPHB2 → FOS → RELA
19	NGAL → LRP2 → CLU → UBC → RELA
20	NGAL → LRP2 → CLU → NFKB1 → RELA
21	NGAL → LRP2 → CLU → COMMD1 → RELA
22	NGAL → LRP2 → CLU → FOS → RELA
23	NGAL → LRP2 → CLU → MDM2 → RELA
24	NGAL → LRP2 → CLU → CHUK → RELA
25	NGAL → LRP2 → THBS1 → UBC → RELA
26	NGAL → LRP2 → THBS1 → COL2A1 → RELA
27	NGAL → LRP2 → APP → TAF11 → RELA
28	NGAL → LRP2 → APP → FOS → RELA
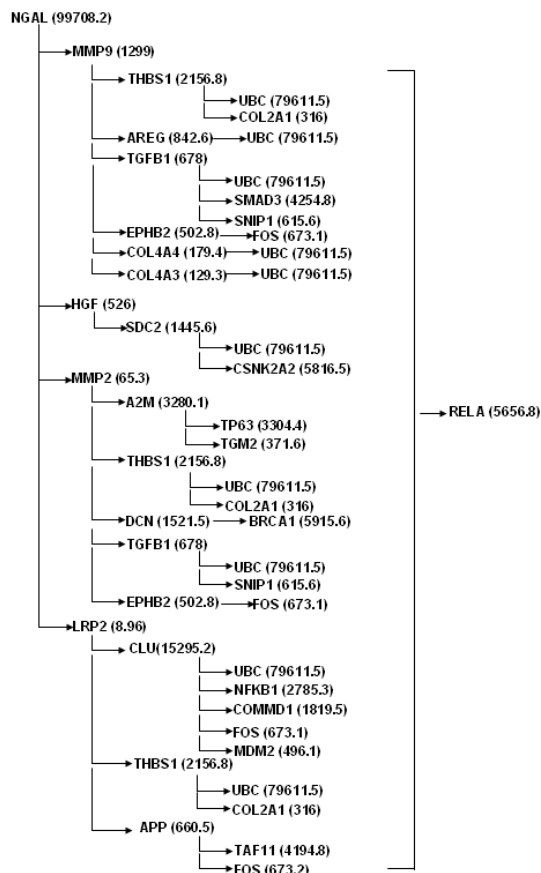
Figure 3. The Schematic Diagram of Prioritization of Shortest Paths from NGAL to RELA According to Their Normalized Intensity from The Microarray by The Order of Interaction Cascades

following NGAL overexpression. Then we first identified the transcription factors in the DEGs by DAVID Gene Ontology annotation and 24 DEGs are annotated with the ability of transcription (Table 1). Second, we compared the protein members in the DEG PPI sub-network to the list from TRANSFAC and TRED and found other transcription factors that did not change their expression level after NGAL was overexpression. By the method, we found other 52 transcription factors presented in the DEG sub-network (Supplementary File 2).

Shortest Path analyses result

We found total 245 shortest paths of from NGAL to 24 DEG transcription factors and 2499 shortest paths of from NGAL to other 52 transcription factors in the DEG PPI sub-network. To illustrate the power of this kind analysis, we took RELA and CEBPB for example to analyze the shortest path from NGAL to RELA and NGAL to CEBPB, respectively. To reduce the noise, we remained the shortest paths containing at least one DEG. There were 28 shortest paths from NGAL to RELA (Table 2), while 42 shortest paths from NGAL to CEBPB (Supplementary File 3). In Table 2, the shortest paths were prioritized according
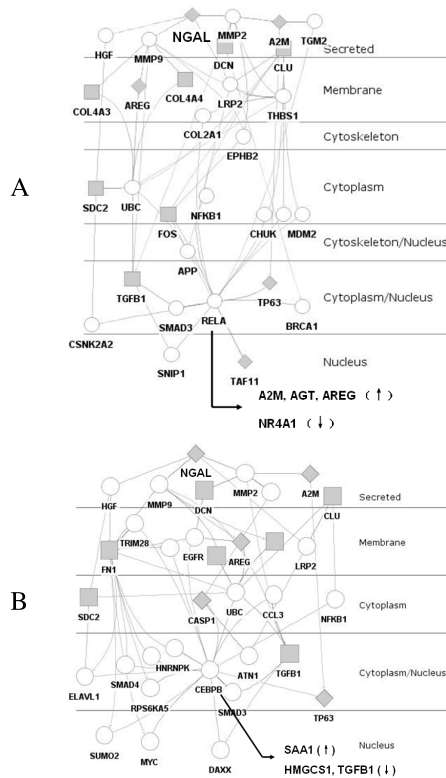

Figure 4. (A) Twenty-eight shortest paths from NGAL to RELA with the protein members are distributed according to their subcellular localizations. The square gray nodes represent proteins encoded downregulated genes, the diamond gray nodes represent the upregulated genes. The target genes of RELA (upregulated A2M, AGT, AREG and the downregulated NR4A1) were also shown. (B) Forty-two shortest paths from NGAL to CEBPB with the protein members are distributed according to their subcellular localizations. The CEBPB targets including the upregulated SAA1, the downregulated HMGCS1 and TGFB1 were shown.

to their normalized intensity from the microarray by the order of interaction cascades (Figure 3). After searching the TRED and TRANSFAC database, we found the four confirmed RELA targets in our DEG list, including upregulated A2M, AGT, AREG and the downregulated NR4A1. And the CEBPB targets include the upregulated SAA1 and the downregulated HMGCS1 and TGFB1. To illustrate the information flow from cell surface till into nucleus, we integrated the subcellular localization of the proteins in the paths to re-arrange their distribution (Figure 4). According to their subcellular localizations, most of these shortest paths obey the principle of from extracellular to cytoplasm, till nucleus.

Discussion

It is usually caused a wide range alternation of mRNA expression profile when one gene is overexpressed or knockdown analyzed by microarray or other high through-put methods. It is interesting to understand how these are happened. Sometimes these could be explained by the traditional pathway enrichment analyses applying KEGG database or other pathways database (Huang et al., 2009). However, many pathway databases are only collecting canonical pathways, those non-canonical pathways have been addressed so far.

Since NGAL can distribute both extracellularly and intracellularly, it is interesting to explore how NGAL signal is transduced from the cell exterior or within the cytoplasm to the nucleus to cause large scale alternation in gene expression profile. There three points inspire us to carry out the analyses in this paper. First, it is well known that all proteins in living body perform their specific functions through the interactions with other proteins in specific biological contexts (Wu et al., 2009). Second, the idea of PPI network has been widely applied in the analyses of microarray data or large scale cancer-related genes analyses (Zhu et al., 2012; Pan 2012). We assumed the transcription factors or transcriptional regulators in the PPI network play critical roles in this expression alternation. Third, we also presumed the information transmitted from one protein to another in the network would also adopt the most economic ways, through the cascade of PPI activities.

So we paid attention to the transcription factors or transcriptional regulators in our PPI sub-network and applied the shortest path algorithm to illustrate how NGAL reaches a specific transcription factor. We took transcription factors RELA and CEBPB for example to illustrate the strength of shortest path analyses. RELA (v-rel avian reticuloendotheliosis viral oncogene homolog A), also named p65 and NFKB3, plays a significant role in multiple tumor progression and metastasis (Hayden, et al., 2008). RELA promotes the invasion and metastasis of ESCC through attenuating the expression of MMP-9 and epithelial-to-mesenchymal transition (EMT) (Wang et al., 2013). CEBPB (CCAAT/enhancer binding protein (C/EBP), beta), is a pluripotent transcription factor that controls inflammation, proliferation, and differentiation. For example, CEBPB is a critical intermediate for IL-1B-dependent MMP gene activation in SW1353

chondrosarcoma cells (Petrella et al., 2011).

According to TRED and TRANSFAC, these two cancer-related transcription factors regulate the transcription of the DEGs following NGAL overexpression, such as A2M, NR4A1 and TGFB1. Many target genes of these transcription factors play crucial role in ESCC. For example, TGFB1 is a key gene that mediates epithelial to mesenchymal transition (EMT) and cell invasion in ESCC (Rees et al., 2006). These indicated that the mRNA expression profile alternation following NGAL overexpression were through some critical transcription factors. Except the transcription factors in the DEG PPI network with their expression did not change, we also paid attention to the DEGs with ability of transcription. On the other hand, we found a downregulated DEG FOXP1 (forkhead box P1), which could potential regulate the transcription of six downregulated DEGs (COL4A4, EGR1, FOS, PGCP, PMP22 and TGFBI). These indicated that when the expression level of a transcription factor is changed, it might also change the expression its target genes, thus a transcription regulational cascade signals were formed and genome-wide expression was changed.

The composition and biological role of proteins might vary according their subcellular localization, suggesting the signaling is transduced by sequential PPIs (Friedman et al., 2007; Kumar et al., 2010). In this study, we also integrated subcellular localization information into the shortest path analyses, which offers important clues for proteins. We also found many of these shortest paths from NGAL to RELA obey the principle of from extracellular to cytoplasm, till nucleus. These results suggested that overexpression of NGAL might influence the PPI network directly or indirectly, affecting the signaling of extracellular-membrane-cytoskeleton/cytoplasm-nucleus cascades to cause the altered expressions of DEGs and consequent alterations in cell proliferation, cell morphology, invasion and metastasis.

In summary, a systems approach was developed in this study by linking differentially expressed genes following NGAL overexpression with publicly available PPI data to build a sub-network that delineate the mechanism of large scale alternation of mRNA expression profile. These shortest paths from NGAL to DEG transcription factors or other transcription factors in the PPI network provide important clues for future experimental identification of new paths.

References

Abecassis I, Olofsson B, Schmid M, et al (2003). RhoA induces MMP-9 expression at CD44 lamellipodial focal complexes and promotes HMEC-1 cell invasion. *Exp Cell Res*, **291**, 363-76.

Barsky A, Gardy JL, Hancock RE, et al (2007). Cerebral: a Cytoscape plugin for layout of and interaction with biological networks using subcellular localization annotation. *Bioinformatics*, **23**, 1040-2.

Chakraborty S, Kaur S, Guha S, et al (2012). The multifaceted roles of neutrophil gelatinase associated lipocalin (NGAL) in inflammation and cancer. *Biochim Biophys Acta*, **1826**, 129-69.

Chatr-Aryamontri A, Breitkreutz BJ, Heinicke S, et al (2013).

The BioGRID interaction database: 2013 update. *Nucleic Acids Res*, **41**, 816-23.

Csardi G, Nepusz T (2006). The igraph software package for complex network research. *InterJournal Complex Systems*, **5**, 1695.

Dijkstra E (1959). A note on two problems in connection with graphs. *Numerische Mathematik*, **1**, 269-71.

Du ZP, Lv Z, Wu BL, et al (2011.) Neutrophil gelatinase-associated lipocalin and its receptor: independent prognostic factors of oesophageal squamous cell carcinoma. *J Clin Pathol*, **64**, 69-74.

Flower DR (1996). The lipocalin protein family: structure and function. *Biochem J*, **318**, 1-14.

Friedman A, Perrimon N (2007). Genetic screening for signal transduction in the era of network biology. *Cell*, **128**, 225-31.

Goel R, Muthusamy B, Pandey A, et al (2011). Human protein reference database and human proteinpedia as discovery resources for molecular biotechnology. *Mol Biotechnol*, **48**, 87-95.

Hayden MS, Ghosh S (2008). Shared principles in NF-kappaB signaling. *Cell*, **132**, 344-62.

Hu L, Hittelman W, Lu T, et al (2009). NGAL decreases E-cadherin mediated cell-cell adhesion and increases cell motility and invasion through Rac1 in colon carcinoma cells. *Lab I nvest*, **89**, 531-48.

Huang da W, Sherman BT, Lempicki RA (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*, **37**, 1-13.

Jiang C, Xuan Z, Zhao F, et al (2007). TRED: a transcriptional regulatory element database, new entries and other development. *Nucleic Acids Res*, **35**, 137-40.

Kerrien S, Aranda B, Breuza L, et al (2012). The IntAct molecular interaction database in 2012. *Nucleic Acids Res*, **40**, 841-6.

Kumar G, Ranganathan S (2010). Network analysis of human protein location. *BMC Bioinformatics*, **11 Suppl 7**, S9.

Matys V, Fricke E, Geffers R, et al (2003). TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res*, **31**, 374-8.

Pan XH (2012). Pathway crosstalk analysis based on protein-protein network analysis in ovarian cancer. *Asian Pac J Cancer Prev*, **13**, 3905-9.

Petrella BL, Armstrong DA, Vincenti MP (2011). CCAAT-enhancer-binding protein beta activation of MMP-1 gene expression in SW1353 cells: independent roles of extracellular signal-regulated and p90/ribosomal S6 kinases. *J Cell Physiol*, **226**, 3349-54.

Rees JR, Onwuegbusi BA, Save VE, et al (2006). In vivo and in vitro evidence for transforming growth factor-beta1-mediated epithelial to mesenchymal transition in esophageal adenocarcinoma. *Cancer Res*, **66**, 9583-90.

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

Tong Z, Kunnumakkara AB, Wang H, et al (2008). Neutrophil gelatinase-associated lipocalin: a novel suppress or of invasion and angiogenesis in pancreatic cancer. *Cancer Res*, **68**, 6100-8.

Wang F, He W, Fanghui P, et al (2013). NF-κBP65 promotes invasion and metastasis of oesophageal squamous cell cancer by regulating matrix metalloproteinase-9 and epithelial-to-mesenchymal transition. *Cell Biol Int*, **37**, 780-8.

Wu J, Vallenius T, Ovaska K, et al (2009). Integrated network analysis platform for protein-protein interactions. *Nat Methods*, **6**, 75-7.

Yan L, Borregaard N, Kjeldsen L, et al (2001). The high molecular weight urinary matrix metalloproteinase (MMP) activity is a complex of gelatinase B/MMP-9 and neutrophil gelatinase-associated lipocalin (NGAL). Modulation of MMP-9 activity by NGAL. *J Biol Chem*, **276**, 37258-65.

Yang J, Goetz D, Li JY, et al (2002). An iron delivery pathway mediated by a lipocalin. *Mol Cell*, **10**, 1045-56.

Zhang H, Xu L, Xiao D, et al (2007) Up regulation of neutrophil gelatinase associated lipocalin in oesophageal squamous cell carcinoma: significant correlation with cell differentiation and tumour invasion. *J Clin Pathol*, **60**, 555-61.

Zhu XL, Ai ZH, Wang J, et al (2012). Weighted gene co-expression network analysis in identification of endometrial cancer prognosis markers. *Asian Pac J Cancer Prev*, **13**, 4607-11.