

Cartilage oligomeric matrix protein (*COMP*)-mediated cell differentiation to proteolysis mechanism networks from human normal adjacent tissues to lung adenocarcinoma

Lin Wang^{a,b,*}, Juxiang Huang^{a,1}, Minghu Jiang^{c,1}, Haizhen Diao^a, Huilei Zhou^a, Xiaohe Li^a, Qingchun Chen^a, Zhenfu Jiang^d, Haitao Feng^e and Stefan Wolff^f

^aBioinformatics Center, School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing, China

^bState key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China

^cLab of Computational Linguistics, School of Humanities and Social Sciences, Tsinghua University, Beijing, China

^dSchool of Mechanical Electronic & Information Engineering, China University of Mining & Technology, Beijing, China

^eDean department, Heilongjiang University of Chinese Medicine, Harbin, China

^fInstitute of Pharmacy and Molecular Biotechnology, Heidelberg University, Heidelberg, Germany

Abstract.

BACKGROUND: To understand cartilage oligomeric matrix protein (*COMP*) mechanism network from human normal adjacent tissues to lung adenocarcinoma.

METHODS: *COMP* complete different activated (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) network were identified in higher lung adenocarcinoma compared with lower human normal adjacent tissues from the corresponding *COMP*-stimulated (≥ 0.25) or inhibited (Pearson $CC \leq -0.25$) overlapping molecules of Pearson correlation coefficient (CC) and GRNInfer, respectively. *COMP* complete different activated and inhibited (all no positive correlation, Pearson $CC < 0.25$) mechanisms networks of higher lung adenocarcinoma and lower human normal adjacent tissues were constructed by integration of Pearson CC, GRNInfer and GO. As visualized by integration of GO, KEGG, GenMAPP, BioCarta and Disease, we deduced *COMP* complete different activated and inhibited network in higher lung adenocarcinoma and lower human normal adjacent tissues.

RESULTS: As visualized by GO, KEGG, GenMAPP, BioCarta and disease database integration, we proposed mainly that the mechanism and function of *COMP* complete different activated network in higher lung adenocarcinoma was involved in *COMP* activation with matrix-localized insulin-like factor coupling carboxypeptidase to metallopeptidase-induced proteolysis,

*Corresponding author: Lin Wang (Prof. Dr.), Bioinformatics Center, School of Electronics Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China. Tel.: +86 13240981826; Fax: +86 10 62785736; E-mail: wanglin98@tsinghua.org.cn (Minghu Jiang c/o Lin Wang).

¹Equal contribution.

whereas the corresponding inhibited network in lower human normal adjacent tissues participated in *COMP* inhibition with nucleus-localized vasculogenesis, B and T cell differentiation and neural endocrine factors coupling pyrophosphatase-mediated proteolysis. However, *COMP* complete different inhibited network in higher lung adenocarcinoma included *COMP* inhibition with nucleus-localized chromatin maintenance, licensing and assembly factors coupling phosphatase-inhibitor to cytokinesis regulators-mediated cell differentiation, whereas the corresponding activated network in lower human normal adjacent tissues contained *COMP* activation with cytoplasm-localized translation elongation factor coupling fucosyltransferase to ubiquitin-protein ligase-induced cell differentiation.

CONCLUSION: *COMP* different networks were verified not only by complete and uncomplete *COMP* activated or inhibited networks within human normal adjacent tissues or lung adenocarcinoma, but also by *COMP* activated and inhibited network between human normal adjacent tissues and lung adenocarcinoma.

Keywords: Cartilage oligomeric matrix protein (*COMP*), cell differentiation to proteolysis mechanism networks, from human normal adjacent tissues to lung adenocarcinoma

1. Introduction

Cartilage oligomeric matrix protein (*COMP*) (our identified 500 significant molecules in 22,284 genes of 25 higher lung adenocarcinoma) was identified different from the corresponding 25 lower human normal adjacent tissues in GEO data set GSE7670. Now the function and mechanism studies with cartilage oligomeric matrix protein (*COMP*) positive relationships are presented in papers. Such as, *COMP* acts as a catalyst in collagen fibrillogenesis [1]. Ultrastructural immunolocalization of cartilage oligomeric matrix protein (*COMP*) in relation to collagen fibrils in the equine tendon [2]. Differential response of cartilage oligomeric matrix protein (*COMP*) to morphogens of bone morphogenetic protein/transforming growth factor-beta family in the surface, middle and deep zones of articular cartilage [3]. Cartilage oligomeric matrix protein (*COMP*) forms part of the connective tissue of normal human hair follicles [4]. Identification of cartilage oligomeric matrix protein (*COMP*) gene mutations in patients with pseudoachondroplasia and multiple epiphyseal dysplasia [5]. The functional studies with cartilage oligomeric matrix protein (*COMP*) negative relationships are presented in papers. Such as, the deficiency of cartilage oligomeric matrix protein (*COMP*) leads to vascular mineralization [6]. Microarray analysis identifies *COMP* as the most differentially regulated transcript throughout *in vitro* follicle growth [7]. The murine *COMP* (cartilage oligomeric matrix protein) promoter contains a potent transcriptional repressor region [8]. Transgenic mice expressing D469Delta mutated cartilage oligomeric matrix protein (*COMP*) show growth plate abnormalities and sternal malformations [9]. Chondrocyte cell death and intracellular distribution of *COMP* and type IX collagen in the pseudoachondroplasia growth plate

[10]. Delta 469 mutation in the type 3 repeat calcium binding domain of cartilage oligomeric matrix protein (*COMP*) disrupts calcium binding [11]. Retention of cartilage oligomeric matrix protein (*COMP*) and cell death in redifferentiated pseudoachondroplasia chondrocytes [12]. Yet novel *COMP*-mediated cell differentiation to proteolysis mechanism networks is not clear and remains to be elucidated from human normal adjacent tissues to lung adenocarcinoma.

In the lower expression of human normal adjacent tissues, we detected 366 and 361 as *COMP*-activated (Pearson $CC \geq 0.25$) and -inhibited (Pearson $CC \leq -0.25$) targets for *COMP*-mediated cell differentiation to proteolysis mechanism networks from human normal adjacent tissues to lung adenocarcinoma by GRNInfer (Table S1A), respectively. Validated 53 and 41 as *COMP*-activated and -inhibited targets were also quantified for *COMP*-mediated cell differentiation to proteolysis mechanism networks by correlation coefficient analysis, respectively (Table S1B). 41 and 31 as *COMP*-activated and -inhibited targets were identified for *COMP*-mediated cell differentiation to proteolysis mechanism networks by overlapping molecules of Pearson and GRNInfer, respectively (Table S1C). Taken together, the studies suggest that *COMP* includes maybe directly or indirectly at least potential 378 and 371 as *COMP*-activated and -inhibited distinct proteins for *COMP*-mediated cell differentiation to proteolysis mechanism networks (Table S1D). Mutual 9 and 7 as *COMP*-activated and -inhibited targets were identified from overlapped analysis for *COMP*-mediated cell differentiation to proteolysis mechanism networks by correlation coefficient, respectively (Table S1E). Positive mutual complete (all no positive correlation, Pearson $CC < 0.25$) 4 and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different 3 of *COMP*-activated targets, and positive

mutual complete (all no positive correlation, Pearson $CC < 0.25$) 5 and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different 2 of *COMP*-inhibited targets were further obtained in lower human normal adjacent tissues compared with higher lung adenocarcinoma by correlation coefficient analysis with fold changes seen (Table S1F and G and H).

In the higher expression of lung adenocarcinoma, we detected 313 and 402 as *COMP*-activated (Pearson $CC \geq 0.25$) and -inhibited (≤ -0.25) targets for *COMP*-mediated cell differentiation to proteolysis mechanism networks from human normal adjacent tissues to lung adenocarcinoma by GRNInfer (Table S1A), respectively. Validated 93 and 124 as *COMP*-activated and -inhibited targets were also quantified for *COMP*-mediated cell differentiation to proteolysis mechanism networks by correlation coefficient analysis, respectively (Table S1B). 47 and 84 as *COMP*-activated and -inhibited targets were identified for *COMP*-mediated cell differentiation to proteolysis mechanism networks by overlapping molecules of Pearson and GRNInfer, respectively (Table S1C). Taken together, the studies suggest that *COMP* includes maybe directly or indirectly at least potential 359 and 442 as *COMP*-activated and -inhibited distinct proteins for *COMP*-mediated cell differentiation to proteolysis mechanism networks (Table S1D). Mutual 12 and 40 as *COMP*-activated and -inhibited targets were identified from overlapped analysis for different *COMP* networks for *COMP*-mediated cell differentiation to proteolysis mechanism networks by correlation coefficient, respectively (Table S1E). Positive mutual complete (all no positive correlation, Pearson $CC < 0.25$) 5 and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different 6 of *COMP*-activated targets, and positive mutual complete (all no positive correlation, Pearson $CC < 0.25$) 10 and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different 23 of *COMP*-inhibited targets were further obtained in higher lung adenocarcinoma compared with lower human normal adjacent tissues by correlation coefficient analysis with fold changes seen (Table S1F and G and H).

2. Result

Mutual relationships of novel complete (all no positive correlation, Pearson $CC < 0.25$) and uncom-

plete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-activated (Pearson $CC \geq 0.25$) and -inhibited (Pearson $CC \leq -0.25$) networks grouped vertical bars quantification chart between lower human normal adjacent tissues and higher lung adenocarcinoma by Pearson correlation coefficient and our programming, respectively, as shown in Fig. 1 and Table S2.

Mutual relationships of novel complete (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-activated (Pearson $CC \geq 0.25$) and -inhibited (Pearson $CC \leq -0.25$) networks were computed between lower human normal adjacent tissues and higher lung adenocarcinoma by scatterplot (matrix) of linear regression and our programming, respectively, as shown in Fig. 2 and Table S3.

Novel complete (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-activated (Pearson $CC \geq 0.25$) and -inhibited (Pearson $CC \leq -0.25$) networks including autocorrelations were constructed between lower human normal adjacent tissues and higher lung adenocarcinoma by GRNInfer and our programming, respectively, as shown in Fig. 3 and Table S4.

Terms and occurrence numbers of GO (Cellular Component, Molecular Function, Biological Process), KEGG, GenMAPP, BioCarta and Disease in novel complete (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-activated (Pearson $CC \geq 0.25$) and -inhibited (Pearson $CC \leq -0.25$) networks were identified between lower human normal adjacent tissues and higher lung adenocarcinoma by MAS 3.0 and our programming, respectively, as shown in Fig. 4 and Table S5.

3. Discussion

Gene Ontology (GO) analysis of *COMP*-regulated proteins revealed extracellular region, extracellular matrix (sensu Metazoa), extracellular matrix structural constituent, calcium ion binding, protein binding, skeletal development, cell adhesion, organ morphogenesis, TGF-beta signaling pathway, Focal adhesion, ECM-receptor interaction,

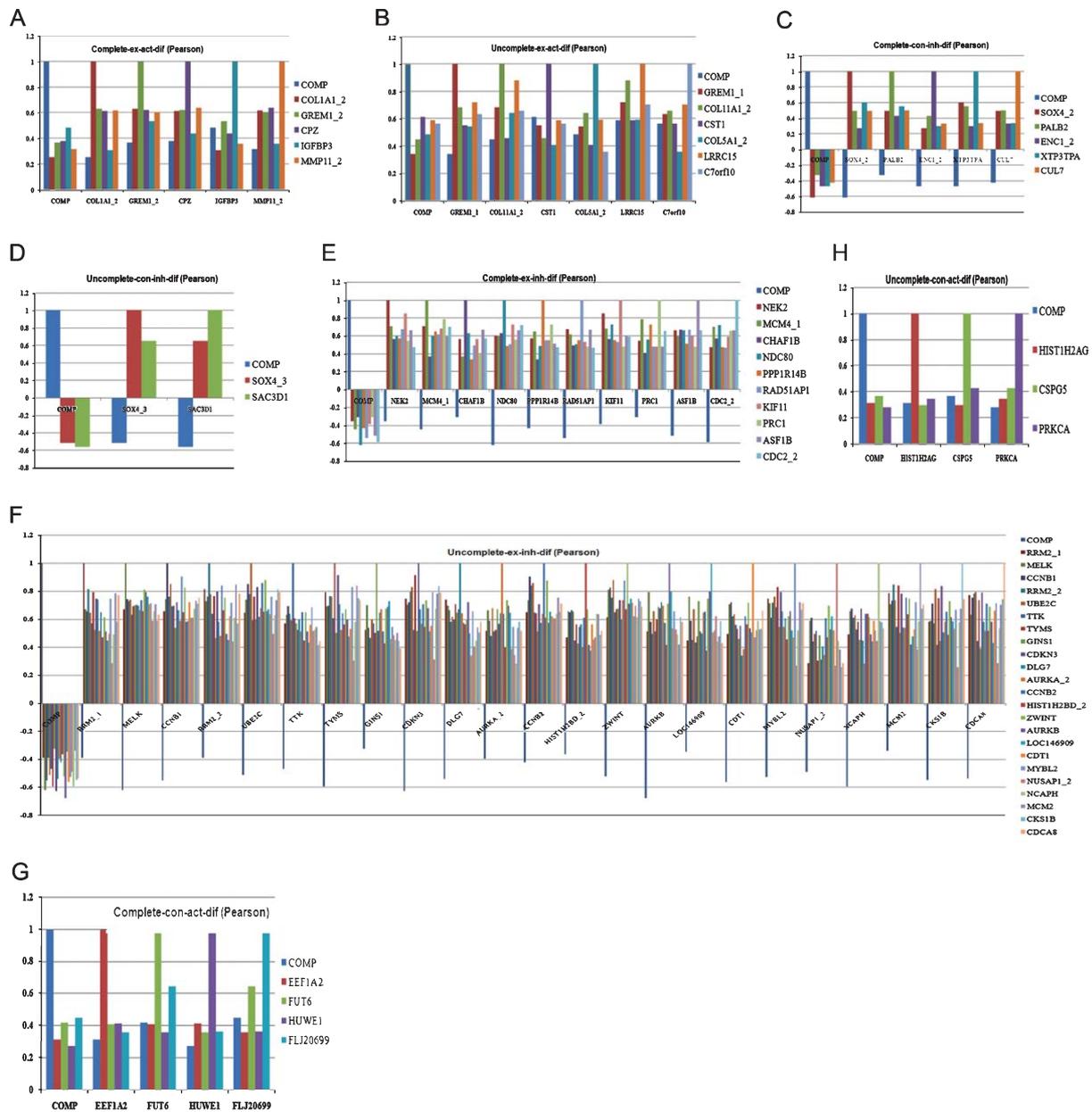


Fig. 1. Novel different *COMP*-activated and -inhibited networks by correlation coefficient (CC) and our programming. (A) Complete (all no positive correlation, Pearson $CC < 0.25$) and (B) uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-activated networks grouped vertical bars quantification chart in higher lung adenocarcinoma compared with lower human normal adjacent tissues by Pearson correlation coefficient (CC) and our programming, respectively. Vertical scale bars indicate CC value, $n = 25$. (C) Complete and (D) uncomplete different *COMP*-inhibited network grouped vertical bars quantification chart in lower human normal adjacent tissues compared with higher lung adenocarcinoma by Pearson CC and our programming, respectively. Vertical scale bars indicate CC value, $n = 25$. (E) Complete and (F) uncomplete different *COMP*-inhibited network grouped vertical bars quantification chart in higher lung adenocarcinoma compared with lower human normal adjacent tissues by Pearson CC and our programming, respectively. Vertical scale bars indicate CC value, $n = 25$. (G) Complete and (H) uncomplete different *COMP*-activated network grouped vertical bars quantification chart in lower human normal adjacent tissues compared with higher lung adenocarcinoma by Pearson CC and our programming, respectively. Vertical scale bars indicate CC value, $n = 25$. con, human normal adjacent tissues; ex, lung adenocarcinoma; act, activation; inh, inhibition; dif, different.

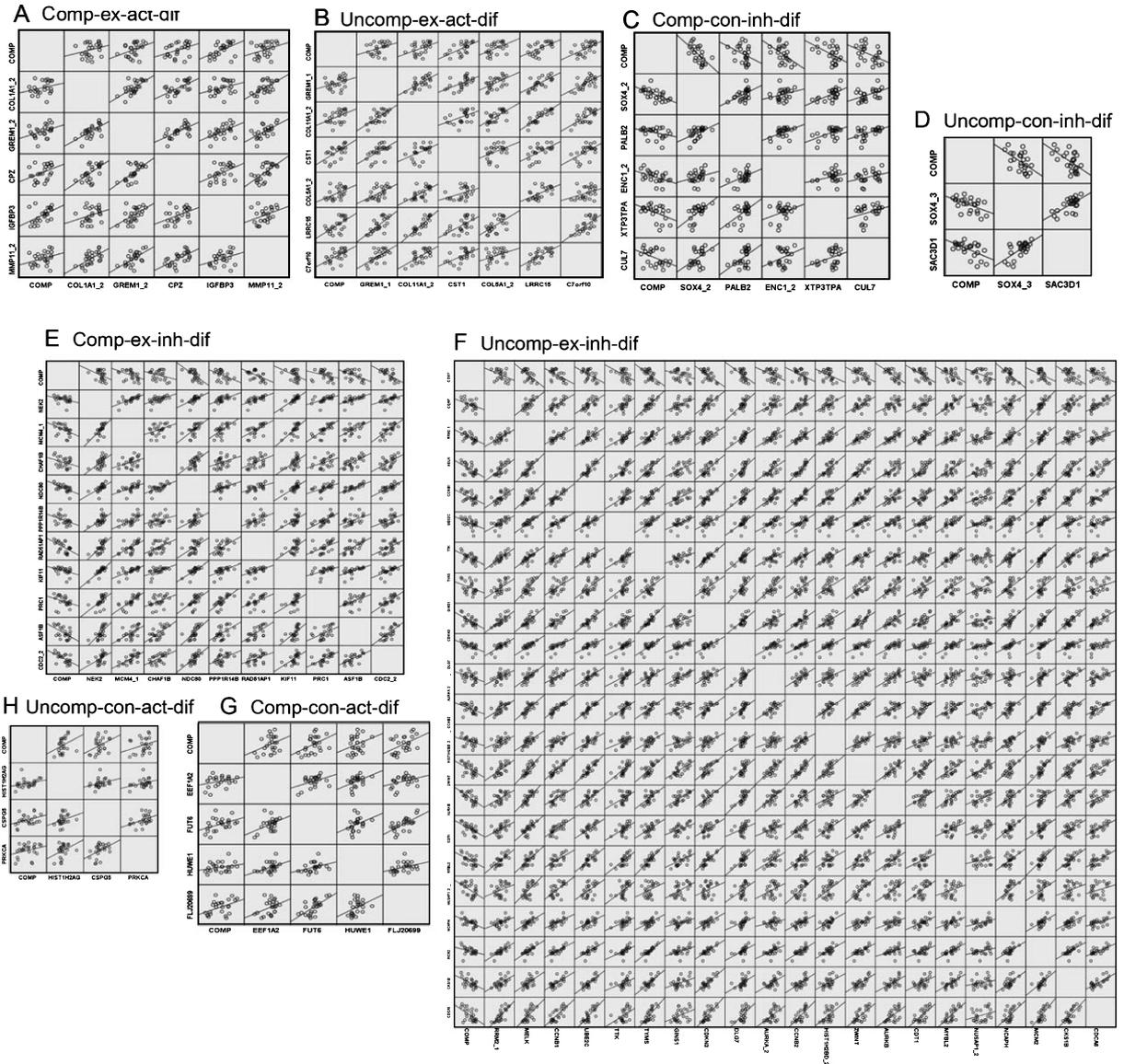


Fig. 2. The mutual relationships of novel different *COMP*-activated and -inhibited networks by scatterplot (matrix) and our programming. (A) Complete (all no positive correlation, Pearson $CC < 0.25$) and (B) uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different mutual relationships of *COMP*-activated network in higher lung adenocarcinoma compared with lower expression of human normal adjacent tissues by scatterplot (matrix) and our programming, respectively. $n = 25$. (C) Complete and (D) uncomplete different mutual relationships of *COMP*-inhibited network in lower human normal adjacent tissues compared with higher lung adenocarcinoma by scatterplot (matrix) and our programming, respectively. $n = 25$. (E) Complete and (F) uncomplete different mutual relationships of *COMP*-inhibited network in higher lung adenocarcinoma compared with lower expression of human normal adjacent tissues by scatterplot (matrix) and our programming, respectively. $n = 25$. (G) Complete and (H) uncomplete different mutual relationships of *COMP*-activated network in lower expression of human normal adjacent tissues compared with higher lung adenocarcinoma by scatterplot (matrix) and our programming, respectively. $n = 25$. Comp, complete; Uncomp, uncomplete; con, human normal adjacent tissues; ex, lung adenocarcinoma; act, activation; inh, inhibition; dif, different.

skeletal development, Tissue-specific-Hs_1-Tissue-Muscle_fat_and_connective, extracellular matrix (sensu Metazoa), Contributed-cellular_process-Hs_

Focal_adhesion_KEGG, osteoarthritis, thanatophoric dysplasia, dwarfism, ptosis, rheumatoid arthritis, osteoarthropathy, achondroplasia, dysplasia,

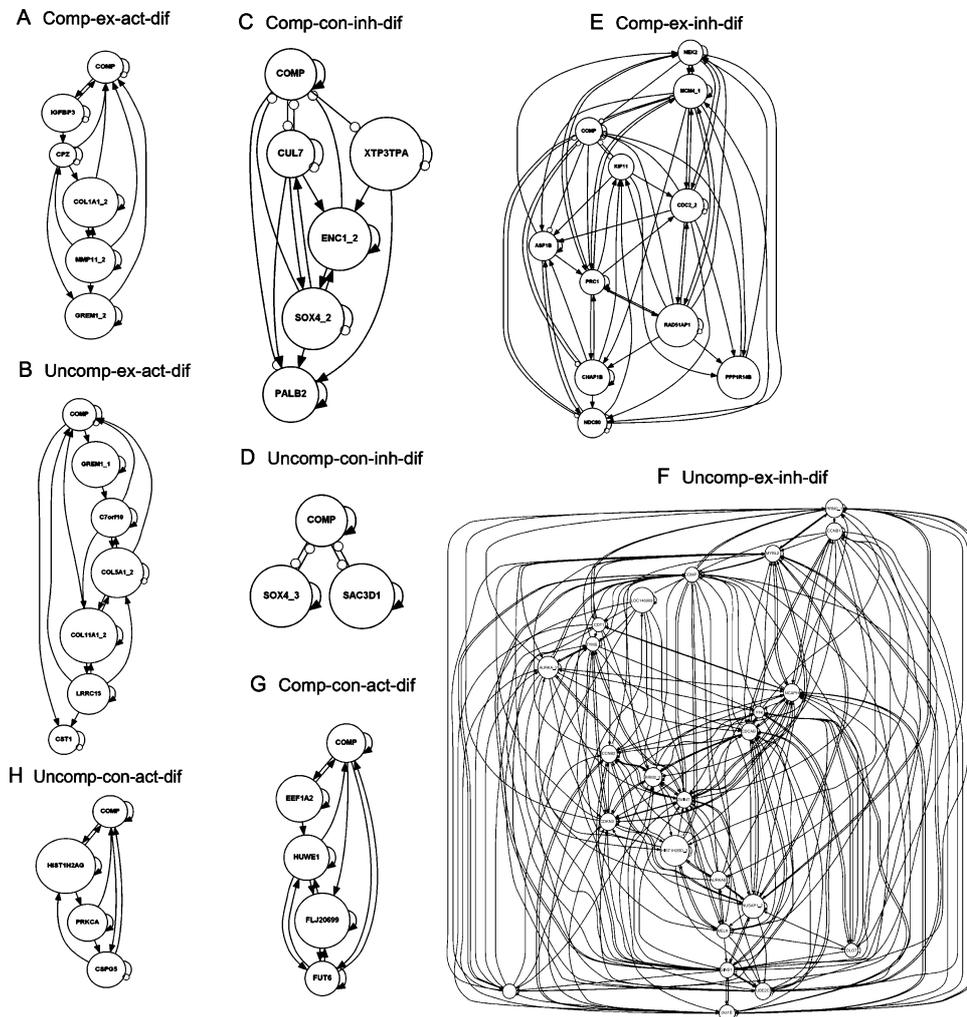


Fig. 3. The novel different *COMP*-activated and -inhibited directionality networks by GRNInfer and our programming. (A) Complete (all no positive correlation, Pearson $CC < 0.25$) and (B) uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-activated network in higher lung adenocarcinoma compared with lower expression of human normal adjacent tissues by GRNInfer and our programming, respectively. $n = 25$. (C) Complete and (D) uncomplete different *COMP*-inhibited network in lower human normal adjacent tissues compared with higher lung adenocarcinoma by GRNInfer and our programming, respectively. $n = 25$. (E) Complete and (F) uncomplete different *COMP*-inhibited network in higher lung adenocarcinoma compared with lower expression of human normal adjacent tissues by GRNInfer and our programming, respectively. $n = 25$. (G) Complete and (H) uncomplete different *COMP*-activated network in lower expression of human normal adjacent tissues compared with higher lung adenocarcinoma by GRNInfer and our programming, respectively. $n = 25$. Black arrow represents as activated relationship and empty circle as inhibited relationship. Comp, complete; Uncomp, uncomplete; con, human normal adjacent tissues; ex, lung adenocarcinoma; act, activation; inh, inhibition; dif, different.

tetanus, pseudoachondroplastic dysplasia, polychondritis relapsing, osteoarthritis hip, disability, osteochondrodysplasias, myopathy, inflammation, joint diseases, malignant neoplasm of pancreas, bone mass, atherosclerosis, osteoarthritis knee, sclerosis, skeletal dysplasia, short stature, brachydactyly, necrosis, effusion, autoimmune diseases, hereditary

diseases, chondrosarcoma, chronic pancreatitis, collagen fibril, arthritis reactive, acquired clubfoot, bone diseases developmental, spondyloepiphyseal dysplasia, diastrophic dysplasia, arthritis, bone destruction, hypertrophy, osteoporosis, vascular diseases, hydrarthrosis, synovitis, multiple epiphyseal dysplasia, swollen joint, connective tissue

diseases, Multiple Epiphyseal Dysplasia, Dominant, Pseudoachondroplasia, multiple epiphyseal dysplasia|pseudoachondroplasia, osteoarthritis, pseudoachondroplasia and multiple epiphyseal dysplasia, pseudoachondroplasia, multiple epiphyseal dysplasia (Fig. 4 and Table S5), yet few novel *COMP*-mediated cell differentiation to proteolysis mechanism networks are known from human normal adjacent tissues to lung adenocarcinoma.

3.1. *COMP*-stimulated proteolysis network in lung adenocarcinoma

COMP complete different activated (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) network were identified in higher lung adenocarcinoma compared with lower human normal adjacent tissues from the corresponding *COMP*-stimulated (≥ 0.25) overlapping molecules of Pearson and GRNInfer (Fig. 1A, B and Table S2). This result was verified by the corresponding scatter matrix (Fig. 2A, B and Table S3). *COMP* complete different activated network was constructed with *COL1A1_2*, *GREM1_2*, *CPZ*, *IGFBP3*, *MMP11_2* (Fig. 3A and Table S4) and the corresponding uncomplete network with *GREM1_1*, *COL11A1_2*, *CST1*, *COL5A1_2*, *LRRC15*, *C7orf10* (Fig. 3B and Table S4). As visualized by GO, KEGG, GenMAPP, BioCarta and disease database integration (Fig. 4A, B and Table S5), we proposed mainly that the mechanism and function of the corresponding *COMP* network was involved in *COMP* activation with matrix-localized insulin-like factor coupling carboxypeptidase to metallopeptidase-induced proteolysis in higher lung adenocarcinoma. This hypothesis was verified by the corresponding uncomplete network in lung adenocarcinoma and *COMP*-inhibited proteolysis network in human normal adjacent tissues.

3.2. *COMP*-inhibited proteolysis network in human normal adjacent tissues

COMP complete different inhibited (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) network were identified in lower human normal adjacent tissues compared with higher lung adenocarcinoma from the corresponding *COMP*-

inhibited (≤ -0.25) overlapping molecules of Pearson and GRNInfer (Fig. 1C, D and Table S2). This result was verified by the corresponding scatter matrix (Fig. 2C, D and Table S3). *COMP* complete different inhibited network was constructed with *SOX4_2*, *PALB2*, *ENC1_2*, *XTP3TPA*, *CUL7* (Fig. 3C and Table S4) and the corresponding uncomplete network with *SOX4_3*, *SAC3DI* (Fig. 3D and Table S4). As visualized by GO, KEGG, GenMAPP, BioCarta and disease database integration (Fig. 4C, D and Table S5), we proposed mainly that the mechanism and function of the corresponding *COMP* network participated in *COMP* inhibition with nucleus-localized vasculogenesis, B and T cell differentiation and neural endocrine factors coupling pyrophosphatase-mediated proteolysis in lower human normal adjacent tissues. This hypothesis was verified by the corresponding uncomplete network in human normal adjacent tissues and *COMP*-activated proteolysis network in lung adenocarcinoma.

3.3. *COMP*-repressed cell differentiation network in lung adenocarcinoma

COMP complete different inhibited (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) network were identified in higher lung adenocarcinoma compared with lower human normal adjacent tissues from the corresponding *COMP*-repressed (≤ -0.25) overlapping molecules of Pearson and GRNInfer (Fig. 1E, F and Table S2). This result was verified by the corresponding scatter matrix (Fig. 2E, F and Table S3). *COMP* complete different inhibited network was constructed with *NEK2*, *MCM4_1*, *CHAF1B*, *NDC80*, *PPP1R14B*, *RAD51API*, *KIF11*, *PRCI*, *ASF1B*, *CDC2_2* (Fig. 3E and Table S4) and the corresponding uncomplete network with *RRM2_1*, *MELK*, *CCNB1*, *RRM2_2*, *UBE2C*, *TTK*, *TYMS*, *GINS1*, *CDKN3*, *DLG7*, *AURKA_2*, *CCNB2*, *HIST1H2BD_2*, *ZWINT*, *AURKB*, *LOC146909*, *CDT1*, *MYBL2*, *NUSAP1_2*, *NCAPH*, *MCM2*, *CKS1B*, *CDCA8* (Fig. 3F and Table S4). As visualized by GO, KEGG, GenMAPP, BioCarta and disease database integration (Fig. 4E, F and Table S5), we proposed mainly that the mechanism and function of the corresponding *COMP* network contained *COMP* inhibition with nucleus-localized chromatin maintenance, licensing and assembly factors coupling phosphatase-inhibitor to cytokinesis

regulators-mediated cell differentiation in higher lung adenocarcinoma. This hypothesis was verified by the corresponding uncomplete network in lung adenocarcinoma and *COMP*-activated cell differentiation network in human normal adjacent tissues.

3.4. *COMP*-activated cell differentiation network in human normal adjacent tissues

COMP complete different activated (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) network were identified in lower human normal adjacent tissues compared with higher lung adenocarcinoma from the corresponding *COMP*-activated (≥ 0.25) overlapping molecules of Pearson and GRNInfer (Fig. 1G, H and Table S2). This result was verified by the corresponding scatter matrix (Fig. 2G, H and Table S3). *COMP* complete different activated network was constructed with *EEF1A2*, *FUT6*, *HUWE1*, *FLJ20699* (Fig. 3G and Table S4) and the corresponding uncomplete network with *HIST1H2AG*, *CSPG5*, *PRKCA* (Fig. 3H and Table S4). As visualized by GO, KEGG, GenMAPP, BioCarta and disease database integration (Fig. 4G, H and Table S5), we proposed mainly that the mechanism and function of the corresponding *COMP* network included *COMP* activation with cytoplasm-localized translation elongation factor coupling fucosyltransferase to ubiquitin-protein ligase-induced cell differentiation in lower human normal adjacent tissues. This hypothesis was verified by the corresponding uncomplete network in human normal adjacent tissues and *COMP*-inhibited cell differentiation network in lung adenocarcinoma.

3.5. *COMP* networks autocorrelations

In complete (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive

correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-activated network of higher lung adenocarcinoma compared with lower human normal adjacent tissues, there is *COL1A1_2* auto-activation including *upstream CPZ*, *MMP11_2*, *downstream COMP*, *MMP11_2*; There is *GREM1_2* auto-activation including *upstream CPZ*, *MMP11_2*, *downstream COMP*; There is *MMP11_2* auto-activation including *upstream COL1A1_2*, *downstream COL1A1_2*, *GREM1_2*, *COMP*, *CPZ* by GRNInfer (Fig. 3A and Table S4). Thus, our results reflected the stimulation of *GREM1_2* upstream; *MMP11_2* downstream; *COL1A1_2* the same molecular numbers in lung adenocarcinoma. There is *COMP* auto-inhibition containing *upstream COL1A1_2*, *GREM1_2*, *CPZ*, *IGFBP3*, *MMP11_2*, *downstream IGFBP3*; There is *CPZ* auto-inhibition containing *upstream IGFBP3*, *MMP11_2*, *downstream COL1A1_2*, *GREM1_2*, *COMP*; There is *IGFBP3* auto-inhibition containing *upstream COMP*, *downstream COMP*, *CPZ* by GRNInfer (Fig. 3A and Table S4). Thus, our results reflected the repression of *CPZ*, *IGFBP3* upstream; *COMP* downstream molecular numbers in lung adenocarcinoma.

In complete (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-inhibited network of lower human normal adjacent tissues compared with higher lung adenocarcinoma, there is *COMP* auto-activation including *upstream SOX4_2*, *ENC1_2*, *CUL7*, *downstream PALB2*, *XTP3TPA*, *CUL7*; There is *PALB2* auto-activation including *upstream XTP3TPA*, *CUL7*, *COMP*, *SOX4_2*; There is *ENC1_2* auto-activation including *upstream SOX4_2*, *XTP3TPA*, *CUL7*, *downstream SOX4_2*, *COMP* by GRNInfer (Fig. 3C and Table S4). Thus, our results reflected the stimulation of *PALB2* *ENC1_2*, upstream; *COMP* the same molecular numbers in human normal adjacent tissues. There is *SOX4_2* auto-inhibition containing *upstream ENC1_2*, *CUL7*, *downstream ENC1_2*, *CUL7*, *COMP*,

Fig. 4. Terms and occurrence numbers of GO (Cellular Component, Molecular Function, Biological Process), KEGG, GenMAPP, BioCarta and Disease by MAS 3.0 and our programming, respectively. (A) Complete (all no positive correlation, Pearson $CC < 0.25$) and (B) uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-activated network in higher lung adenocarcinoma compared with lower expression of human normal adjacent tissues by MAS 3.0 and our programming, respectively. $n = 25$. (C) Complete and (D) uncomplete different *COMP*-inhibited network in lower human normal adjacent tissues compared with higher lung adenocarcinoma by MAS 3.0 and our programming, respectively. $n = 25$. (E) Complete and (F) uncomplete different *COMP*-inhibited network in higher lung adenocarcinoma compared with lower expression of human normal adjacent tissues by MAS 3.0 and our programming, respectively. $n = 25$. (G) Complete and (H) uncomplete different *COMP*-activated network in lower expression of human normal adjacent tissues compared with higher lung adenocarcinoma by MAS 3.0 and our programming, respectively. $n = 25$. Comp, complete; Uncomp, uncomplete; con, human normal adjacent tissues; ex, lung adenocarcinoma; act, activation; inh, inhibition; dif, different.

PALB2; There is *XTP3TPA* auto-inhibition containing *upstream COMP*, *downstream PALB2*, *ENC1_2*; There is *CUL7* auto-inhibition containing *upstream SOX4_2*, *COMP*, *downstream SOX4_2*, *PALB2*, *ENC1_2*, *COMP* by GRNInfer (Fig. 3C and Table S4). Thus, our results reflected the repression of *SOX4_2*, *XTP3TPA*, *CUL7* upstream molecular numbers in human normal adjacent tissues.

In complete (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-inhibited network of higher lung adenocarcinoma compared with lower human normal adjacent tissues, there is *NEK2* auto-activation including *upstream MCM4_1*, *PPP1R14B*, *RAD51API*, *PRC1*, *ASF1B*, *CDC2_2*, *downstream MCM4_1*, *NDC80*, *PRC1*, *CDC2_2*, *COMP*; There is *MCM4_1* auto-activation including *upstream NEK2*, *NDC80*, *RAD51API*, *PRC1*, *CDC2_2*, *COMP*, *downstream NEK2*, *CHAF1B*, *PPP1R14B*, *RAD51API*, *PRC1*, *ASF1B*, *CDC2_2*, *COMP*; There is *CHAF1B* auto-activation including *upstream MCM4_1*, *RAD51API*, *PRC1*, *COMP*, *downstream NDC80*, *KIF11*, *PRC1*, *ASF1B*, *COMP*; There is *PRC1* auto-activation including *upstream NEK2*, *MCM4_1*, *CHAF1B*, *RAD51API*, *ASF1B*, *COMP*, *downstream NEK2*, *MCM4_1*, *CHAF1B*, *RAD51API*, *KIF11*, *CDC2_2*; There is *ASF1B* auto-activation including *upstream MCM4_1*, *CHAF1B*, *NDC80*, *KIF11*, *CDC2_2*, *COMP*, *downstream NEK2*, *NDC80*, *PRC1* by GRNInfer (Fig. 3E and Table S4). Thus, our results reflected the stimulation of *NEK2*, *ASF1B* upstream; *MCM4_1*, *CHAF1B* downstream; *PRC1* the same molecular numbers in lung adenocarcinoma. There is *COMP* auto-inhibition containing *upstream NEK2*, *MCM4_1*, *CHAF1B*, *NDC80*, *PPP1R14B*, *RAD51API*, *KIF11*, *CDC2_2*, *downstream MCM4_1*, *CHAF1B*, *NDC80*, *KIF11*, *PRC1*, *ASF1B*; There is *NDC80* auto-inhibition containing *upstream NEK2*, *CHAF1B*, *ASF1B*, *CDC2_2*, *COMP*, *downstream MCM4_1*, *KIF11*, *ASF1B*, *COMP*; There is *RAD51API* auto-inhibition containing *upstream MCM4_1*, *PRC1*, *CDC2_2*, *downstream NEK2*, *MCM4_1*, *CHAF1B*, *PPP1R14B*, *KIF11*, *PRC1*, *CDC2_2*, *COMP*; There is *CDC2_2* auto-inhibition containing *upstream NEK2*, *MCM4_1*, *RAD51API*, *KIF11*, *PRC1*, *downstream NEK2*, *MCM4_1*, *NDC80*, *PPP1R14B*, *RAD51API*, *ASF1B*, *COMP* by GRNInfer (Fig. 3E and Table S4). Thus, our results reflected the repression of *RAD51API*, *CDC2_2* upstream; *COMP*, *NDC80*

downstream molecular numbers in lung adenocarcinoma.

In complete (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-activated network of lower human normal adjacent tissues compared with higher lung adenocarcinoma, there is *COMP* auto-activation including *upstream EEF1A2*, *FUT6*, *HUWE1*, *downstream EEF1A2*, *FUT6*, *FLJ20699*; There is *EEF1A2* auto-activation including *upstream COMP*, *downstream COMP*, *HUWE1*; There is *FUT6* auto-activation including *upstream COMP*, *HUWE1*, *FLJ20699*, *downstream COMP*, *HUWE1*, *FLJ20699*; There is *HUWE1* auto-activation including *upstream EEF1A2*, *FUT6*, *FLJ20699*, *downstream COMP*, *FUT6*, *FLJ20699*; There is *FLJ20699* auto-activation including *upstream COMP*, *FUT6*, *HUWE1*, *downstream FUT6*, *HUWE1* by GRNInfer (Fig. 3G and Table S4). Thus, our results reflected the stimulation of *FLJ20699* upstream; *EEF1A2* downstream; *COMP*, *FUT6*, *HUWE1* the same molecular numbers in human normal adjacent tissues.

3.6. Summary

As visualized by GO, KEGG, GenMAPP, BioCarta and disease database integration, we proposed mainly that the mechanism and function of *COMP* complete different activated network in higher lung adenocarcinoma was involved in *COMP* activation with matrix-localized insulin-like factor coupling carboxypeptidase to metallopeptidase-induced proteolysis, whereas the corresponding inhibited network in lower human normal adjacent tissues participated in *COMP* inhibition with nucleus-localized vasculogenesis, B and T cell differentiation and neural endocrine factors coupling pyrophosphatase-mediated proteolysis. However, *COMP* complete different inhibited network in higher lung adenocarcinoma included *COMP* inhibition with nucleus-localized chromatin maintenance, licensing and assembly factors coupling phosphatase-inhibitor to cytokinesis regulators-mediated cell differentiation, whereas the corresponding activated network in lower human normal adjacent tissues contained *COMP* activation with cytoplasm-localized translation elongation factor coupling fucosyltransferase to ubiquitin-protein ligase-induced cell differentiation. *COMP* different

networks were verified not only by complete and uncomplete *COMP* activated or inhibited networks within human normal adjacent tissues or lung adenocarcinoma, but also by *COMP* activated and inhibited network between human normal adjacent tissues and lung adenocarcinoma.

4. Materials and methods

Cartilage oligomeric matrix protein (*COMP*) (our identified 500 significant molecules in 22,284 genes of 25 higher lung adenocarcinoma) was identified different from the corresponding 25 lower human normal adjacent tissues in GEO data set GSE7670 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE7670>) [13]. The raw microarray data was processed by log base 2.

Novel different *COMP*-activated and -inhibited networks were detected and quantified from 500 significant molecules between lower human normal adjacent tissues and higher lung adenocarcinoma using significant analysis of microarrays (SAM) (<http://www-stat.stanford.edu/tibs/SAM/>) [14] and our programming. Two classes paired and minimum fold change ≥ 2 were selected (the false-discovery rate 0%).

Mutual relationships of novel complete (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-activated (Pearson $CC \geq 0.25$) and inhibited (≤ -0.25) network were computed between lower human normal adjacent tissues and higher lung adenocarcinoma by SPSS of Pearson correlation coefficient quantification about measures of the correlation (linear dependence) between two variables X and Y, giving a value between +1 and -1 inclusive and our programming, respectively.

Mutual relationships of novel complete (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-activated and -inhibited networks were investigated between lower human normal adjacent tissues and higher lung adenocarcinoma by SPSS of scatterplot (matrix) based on linear regression and our programming, respectively.

Novel complete (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-activated and -inhibited networks were constructed between lower human normal adjacent tissues

and higher lung adenocarcinoma by GRNInfer [15], our articles [16–33] and GVedit tool and our programming, respectively.

Terms and occurrence numbers of GO (Cellular Component, Molecular Function, Biological Process), KEGG, GenMAPP, BioCarta and Disease in novel complete (all no positive correlation, Pearson $CC < 0.25$) and uncomplete (partly no positive correlation except *COMP*, Pearson $CC < 0.25$) different *COMP*-activated and -inhibited networks were identified between lower human normal adjacent tissues and higher lung adenocarcinoma, respectively, by Molecule Annotation System, MAS (CapitalBio Corporation, Beijing, China; <http://bioinfo.capitalbio.com/mas3/>) and our programming. The primary databases of MAS integrated various well-known biological resources, such as Gene Ontology (<http://www.geneontology.org>), KEGG (<http://www.genome.jp/kegg/>), BioCarta (<http://www.biocarta.com/>), GenMapp (<http://www.genmapp.org/>), HPRD (<http://www.hprd.org/>), etc.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 61171114), the State Key Laboratory of Drug Research (SIMM1302KF), Automatical Scientific Planning of Tsinghua University (20111081023 and 20111081010), State Key Lab of Pattern Recognition Open Foundation.

Conflict of interests

We have no conflict of interests.

References

- [1] K. Halasz, et al., *COMP* acts as a catalyst in collagen fibrillogenesis, *J Biol Chem* **282**(43) (2007), 31166–31173.
- [2] F. Sodersten, et al., Ultrastructural immunolocalization of cartilage oligomeric matrix protein (*COMP*) in relation to collagen fibrils in the equine tendon, *Matrix Biol* **24**(5) (2005), 376–385.
- [3] S.C. Motaung, P.E. Di Cesare and A.H. Reddi, Differential response of cartilage oligomeric matrix protein (*COMP*) to morphogens of bone morphogenetic protein/transforming growth factor-beta family in the surface: middle and deep zones of articular cartilage, *J Tissue Eng Regen Med* **5**(6) (2011), e87–e96.

- [4] A. Ariza de Schellenberger, et al., Cartilage oligomeric matrix protein (COMP) forms part of the connective tissue of normal human hair follicles, *Exp Dermatol* **20**(4) (2011), 361–366.
- [5] H.R. Song, et al., Identification of cartilage oligomeric matrix protein (COMP) gene mutations in patients with pseudoachondroplasia and multiple epiphyseal dysplasia, *J Hum Genet* **48**(5) (2003), 222–225.
- [6] Y. Du, et al., Upregulation of a disintegrin and metalloproteinase with thrombospondin motifs-7 by miR-29 repression mediates vascular smooth muscle calcification, *Arterioscler Thromb Vasc Biol* **32**(11) (2012), 2580–2588.
- [7] R.M. Skory, et al., Microarray analysis identifies COMP as the most differentially regulated transcript throughout *in vitro* follicle growth, *Mol Reprod Dev* **80**(2) (2013), 132–144.
- [8] F. Han, et al., The murine COMP (cartilage oligomeric matrix protein) promoter contains a potent transcriptional repressor region, *Osteoarthritis Cartilage* **10**(8) (2002), 638–645.
- [9] M. Schmitz, et al., Transgenic mice expressing D469Delta mutated cartilage oligomeric matrix protein (COMP) show growth plate abnormalities and sternal malformations, *Matrix Biol* **27**(2) (2008), 67–85.
- [10] J.T. Hecht, et al., Chondrocyte cell death and intracellular distribution of COMP and type IX collagen in the pseudoachondroplasia growth plate, *J Orthop Res* **22**(4) (2004), 759–767.
- [11] J. Hou, J.A. Putkey and J.T. Hecht, Delta 469 mutation in the type 3 repeat calcium binding domain of cartilage oligomeric matrix protein (COMP) disrupts calcium binding, *Cell Calcium* **27**(6) (2000), 309–314.
- [12] J.T. Hecht, et al., Retention of cartilage oligomeric matrix protein (COMP) and cell death in redifferentiated pseudoachondroplasia chondrocytes, *Matrix Biol* **17**(8-9) (1998), 625–633.
- [13] L.J. Su, et al., Selection of DDX5 as a novel internal control for Q-RT-PCR from microarray data using a block bootstrap re-sampling scheme, *BMC Genomics* **8** (2007), 140.
- [14] J.D. Storey, A direct approach to false discovery rates, *J Roy Stat Soc Ser B* **64** (2002), 479–498.
- [15] Y. Wang, et al., Inferring gene regulatory networks from multiple microarray datasets, *Bioinformatics* **22**(19) (2006), 2413–2420.
- [16] L. Wang, et al., Activated PTHLH Coupling Feedback Phosphoinositide to G-Protein Receptor Signal-Induced Cell Adhesion Network in Human Hepatocellular Carcinoma by Systems-Theoretic Analysis, *Scientific World Journal* (2012), 428979.
- [17] L. Wang, et al., Inhibited PTHLH downstream leukocyte adhesion-mediated protein amino acid N-linked glycosylation coupling Notch and JAK-STAT cascade to iron-sulfur cluster assembly-induced aging network in no-tumor hepatitis/cirrhotic tissues (HBV or HCV infection) by systems-theoretical analysis, *Integr Biol (Camb)* **4**(10) (2012), 1256–1262.
- [18] L. Wang, et al., Tissue-specific transplantation antigen P35B (TSTA3) immune response-mediated metabolism coupling cell cycle to postreplication repair network in no-tumor hepatitis/cirrhotic tissues (HBV or HCV infection) by biocomputation, *Immunol Res* **52**(3) (2012), 258–268.
- [19] L. Wang, et al., Signal transducer and activator of transcription 2 (STAT2) metabolism coupling postmitotic outgrowth to visual and sound perception network in human left cerebrum by biocomputation, *J Mol Neurosci* **47**(3) (2012), 649–658.
- [20] H. Lin, et al., P-glycoprotein (ABCB1) inhibited network of mitochondrion transport along microtubule and BMP signal-induced cell shape in chimpanzee left cerebrum by systems-theoretical analysis, *Cell Biochem Funct* **30**(7) (2012), 582–587.
- [21] J. Huang, et al., PTHLH coupling upstream negative regulation of fatty acid biosynthesis and Wnt receptor signal to downstream peptidase activity-induced apoptosis network in human hepatocellular carcinoma by systems-theoretical analysis, *J Recept Signal Transduct Res* **32**(5) (2012), 250–256.
- [22] L. Wang, et al., Cyclin-dependent kinase inhibitor 3 (CDKN3) novel cell cycle computational network between human non-malignancy associated hepatitis/cirrhosis and hepatocellular carcinoma (HCC) transformation, *Cell Prolif* **44**(3) (2011), 291–299.
- [23] L. Wang, et al., AFP computational secreted network construction and analysis between human hepatocellular carcinoma (HCC) and no-tumor hepatitis/cirrhotic liver tissues, *Tumour Biol* **31**(5) (2011), 417–425.
- [24] L. Wang, et al., Survivin (BIRC5) cell cycle computational network in human no-tumor hepatitis/cirrhosis and hepatocellular carcinoma transformation, *J Cell Biochem* **112**(5) (2011), 1286–1294.
- [25] L. Wang, et al., MYBPC1 computational phosphoprotein network construction and analysis between frontal cortex of HIV encephalitis (HIVE) and HIVE-control patients, *Cell Mol Neurobiol* **31**(2) (2011), 233–241.
- [26] L. Wang, J. Huang and M. Jiang, CREB5 computational regulation network construction and analysis between frontal cortex of HIV encephalitis (HIVE) and HIVE-control patients, *Cell Biochem Biophys* **60**(3) (2011), 199–207.
- [27] L. Wang, J. Huang and M. Jiang, RRM2 computational phosphoprotein network construction and analysis between no-tumor hepatitis/cirrhotic liver tissues and human hepatocellular carcinoma (HCC), *Cell Physiol Biochem* **26**(3) (2011), 303–310.
- [28] L. Sun, et al., Glycogen debranching enzyme 6 (AGL): enolase 1 (ENOSF1), ectonucleotide pyrophosphatase 2 (ENPP2_1), glutathione S-transferase 3 (GSTM3_3) and mannosidase (MAN2B2) metabolism computational network analysis between chimpanzee and human left cerebrum, *Cell Biochem Biophys* **61**(3) (2011), 493–505.
- [29] J.X. Huang, L. Wang and M.H. Jiang, TNFRSF11B computational development network construction and analysis between frontal cortex of HIV encephalitis (HIVE) and HIVE-control patients, *J Inflamm (Lond)* **7** (2011), 50.
- [30] Y. Sun, et al., Secreted Phosphoprotein 1 Upstream Invasive Network Construction and Analysis of Lung Adenocarcinoma Compared with Human Normal Adjacent Tissues by Integrative Biocomputation, *Cell Biochem Biophys* **56**(2-3) (2010), 59–71.
- [31] J. Huang, et al., Interferon α -Inducible Protein 27 Computational Network Construction and Comparison between the

- Frontal Cortex of HIV Encephalitis (HIVE) and HIVE-Control Patients, *The Open Genomics Journal* **3**(1875-693X) (2010), 1–8.
- [32] L. Wang, et al., Integrative decomposition procedure and Kappa statistics for the distinguished single molecular network construction and analysis, *J Biomed Biotechnol* (2009), 726–728.
- [33] L. Wang, et al., FOS proliferating network construction in early colorectal cancer (CRC) based on integrative significant function cluster and inferring analysis, *Cancer Invest* **27**(8) (2009), 816–824.