Degs in ovarian cancer with irf1 silencing

Bioinformatic analysis of the role of IRF-1 in ovarian cancer

Running title: DEGs in ovarian cancer with IRF1 silencing

Highlights

- We use microarray data to analyze DEGs of ovarian cancer with IRF1silencing.
- Total 242 up-regulated genes and 185 down-regulated genes were screened.
- Significant biology process and pathways involved in the progression are analyzed.
- Construct the PPI network and sub-network for hub gene screened
- TNF, CDH1, MMP2, DKK1 and GJA1 were identified as hub genes.

Abstracts

Background: ovarian cancer is a gynecological neoplastic disease with high death rate among women as its early detection is difficult due to lacking of specificity of clinical symptoms. Recently, researches on gene expression for searching efficient ways of ovarian cancer diagnose carried out and the transcription factor interferon regulatory factor 1 (IRF1), classified as a tumor suppressor, provides us a new insight on screening feature genes of ovarian cancer combined with gene expression method.

Methods: we extracted the microarray expression profile GSE38551 from gene expression omnibus (GEO) database and conducted gene expression analysis of samples with and without IRF1 silencing. After differentially expressed genes (DEGs) were identified, pathways analysis was performed on DEGs and Protein-protein interaction (PPI) network was constructed.

Afterwards, a sub-network was constructed to gain further information of genes scored in the PPI network.

Results: We totally identified 427 DEGs including 242 up-regulated genes and 185 down-regulated genes. Pathways analysis showed that metabolic pathway was the most significant pathway. Some nodes were obtained by PPI-network and sub-network, in which TNF and CDH1 were identified as hub genes.

Conclusion: We observed some remarkable modification on gene expression and some pathways associate with ovarian cancer, which may clarify the role of IRF1 in ovarian cancer.

Keywords: ovarian cancer, interferon regulatory factor 1, differentially expressed genes, pathways analysis, protein-protein interaction

Introduction

Ovarian cancer is known as the fifth most common cause of cancer deaths among women and has become the leading cause of death from gynecological neoplastic disease [1], especially the epithelial ovarian cancer, which is the most common in ovarian caner [2]. The efficient treatment is surgical and associated chemotherapy in early stage as survival is highly dependent on stage of disease: 5-year survival in patients with early-stage is 80-90% compared to 25% for patients with advanced-stage disease [4]. However, it is difficult to take treatment in time because of no obvious symptoms in early stage and lacking of effective early diagnose tool. During the past decades, many researches on ultrasound [5] and cancer antigen 125 (CA-125) [6] using as the primary test have been carried out and CA125

assay has been used as a first-line screening due to the relatively non-invasive nature of blood sampling [7]. However, CA125 is thought to be robust only in following the response or progression of the disease, rather than a diagnostic or prognostic marker [8]. Thus, it is necessary for the researching of new symptom and diagnosis tool in early stage.

General view is that the single-cell layer of surface epithelium that surrounds the ovary derives approximately 90% of ovarian cancers [3]. In the view of genetic, tumorigenesis is a multistep process involving mutations of dominantly acting proto-oncogenes and mutations and loss-of-function mutations of tumor suppressor genes. In case of ovarian cancer, both somatic mutations of oncogenes and somatic or germ-line inactivation of tumor suppressor genes are important genetic changes for etiology [9] and therefore, transcription factors is of vital importance in this process. Interferon regulatory factor 1 (IRF1) is a member of the interferon regulatory transcription factor (IRF) family and serves as an activator of interferons alpha and beta transcription, which is a tumor suppressor gene capable of preventing oncogene-mediated malignant transformation [10]. It has been proved to be required in mouse for a double-stranded RNA induction of these genes [11]. Because of the importance of IRF1 in ovarian cancer, it is likely that gene expression differs with IRF1 silencing, which may provide useful information for the researching of symptom in early stage for diagnosis.

Simona. P et al [10] have examined the role of IRF1 in the response of ovarian cancer cells to the front-line chemotherapeutic drug cisplatin, in which microarray GSE38551 and epithelial ovarian cancer cell SK-OV-3 were used. However, the mechanism of IRF1 in genesis in these ovarian cancer

cells has not been clarified. In this study, two samples, SK-OV-3 transfected with IRF-1 shRNA and SK-OV-3 transfected with scrambled shRNA, were used to analyze the role IRF1 plays in ovarian cancer and the mechanism in genesis. Significantly differentially expressed genes were identified and bioinformatics analysis was performed for further research.

Materials and Methods

Microarray data

The microarray expression was obtained under accession number GSE38551 based on the platform data of GPL10558 (IlluminaHumanHT-12 V4. 0 expression beadchip) from National Center of Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database (http://www. ncbi. nlm. nih. gov/geo/), which was deposited by Pavan S et al [10]. This microarray was obtained based on the analysis of a total of 12 samples, including 3 SK-OV-3 transfected with scrambled shRNA samples, 3 SK-OV-3 transfected with scrambled shRNA with CDDP samples, 3 SK-OV-3 transfected with IRF-1 shRNA samples, and 3 SK-OV-3 transfected with IRF-1 shRNA with CDDP samples. Thereinto, SK-OV-3 transfected with IRF-1 shRNA and SK-OV-3 transfected with scrambled shRNA were chosen in this study.

Data preprocessing and DEGs analysis

The original microarray data were converted into expression measures and performed background correction and quartile data normalization by the robust multiarray average (RMA) algorithm [12] in Limma package. T-test [13] was used to identify the significantly differentially expressed genes between samples of SK-OV-3 transfected with IRF-1 shRNA and SK-OV-3

transfected with scrambled shRNA. The false discovery rate (FDR) < 0. 05 and | log2FC (fold change)| > 1 was used as the threshold.

Gene ontology (GO) and pathway enrichment analysis

GO analysis has become a widely used approach for functional studies of large-scale genomic or transcriptomic data [14]. Kyoto encyclopedia of genes and genomes (KEGG) is a collection of online database dealing with genomes, enzymatic pathways, and biological chemicals, which become widely used [15]. In this study, we identified over-represented GO categories in biological process of the selected DEGs with the p-value < 0. 05, and also identify significant pathways of the selected DEGs with the p-value < 0. 05.

Functional annotation of DEGs

Functional annotation of DEGswas performed for the detection of transcription factors and tumor associated genes. Two databases, tumor suppressor genes (TSGs) database [16] and tumor associated genes (TAG) database [17] were used for screening tumor suppressor genes and oncogenes.

Protein-protein interaction (PPI) network construction

The Search Tool for the Retrieval of Interacting Genes (STRING) database [18] is a database of known and predicted protein interactions, including both experimental and predicted interaction information. Cytoscape [19] is an open source bioinformatics software platform for visualizing molecular interaction networks and integrating with gene expression profiles and other state data. In this study, STRING was used to determine the interaction of a pair of proteins [18] with Required Confidence (combined score) > 0. 4.

Cytoscape was used to visualize molecular interaction networks and identify highly connected nodes (Hubs) [20].

Construction and enrichment analysis of sub-network

To gain further information of genes scored in the PPI network, sub-network was constructed by using BioNet tool [21] in R language with FDR = 0. 0001. GO and KEGG enrichment were performed for the identification of significant pathways related to the DEGs in the sub-network.

Results

DEGs selection

To identify the genes specifically expressed with IRF-1 silencing, we obtained publicly available microarray dataset GSE38551 from GEO database and carried out the Limma package for probe-data preprocessing. The results showed that totally 442 transcript counts were observed, including 250 up-regulation and 192 down-regulation, in which 467 DEGs were obtained, including 242 up-regulated genes and 185 down-regulated genes.

GO and KEGG enrichment analysis for DEGs

We performed GO and KEGG enrichment with p-value < 0. 05 for the functional and pathway analysis of DEGs, respectively. The top 10 GO categories of DEGs were listed in table 1. The up-regulated genes (table 1A) were involved in single-organism, single-organism cellular process, and response to stimulus, while the down-regulated genes (table 1B) participated in the single-organism process and single-organism cellular process.

In addition, the enriched pathways of the DEGs were showed in table 1. We can see from the results that the up-regulated genes (table 1C) were mainly enriched in the metabolic pathways, systemic lupus erythematosus and tight junction, while the down-regulate genes (table 1D) were involved in axon guidance and protein digestion and absorption.

Functional annotation of DEGs

According to the results (table 2), 8 up-regulated transcriptional factors, such as aryl hydrocarbon receptor nuclear translocator (ARNT), CCAAT/enhancer binding protein, delta (CEBPD) and ets homologous factor (EHF), and 11 down-regulated transcriptional factors, such as E2F transcription factor 7 (E2F7), forkhead box F2 (FOXF2) and LIM homeobox 1 (LHX1), were screened.

In addition, In the up-regulated genes, 26 genes were detected, including 2 oncogenes, 20 tumor suppressor genes and 4 genes whose functions were unknown were screened in the iup-regulated genes, while , 15 genes were detected, including 3 oncogenes, 15 tumor genes and 1 gene whose function was unknown were screened in the down-regulated genes.

PPI network construction

The network constructed with all DEGs was shown in figure 1, in which 173 interactions were obtained. In this network, tumor necrosis factor (TNF), cadherin 1, type 1, E-cadherin (CDH1), matrix metallopeptidase 2 (MMP2), collagen type I alpha 1 (COL1A1) and serpin peptidase inhibitor, clade E member 1(SERPINE1) were identified as hub genes as their degrees were relatively larger than other nodes (table 3).

Analysis of sub-network

According to the results, 48 nodes and 68 interactions were obtained (Figure 2). Among them, TNF, CDH1, MMP2, gap junction protein, alpha 1 (GJA1) and dickkopf WNT signaling pathway inhibitor 1 (DKK1), were nodes with higher degrees, which is consistent with that in PPI network.

GO enrichment analysis of the DEGs in the sub-network were listed in table 4, in which morphogenesis of an epothelium, cell junction assembly and heart development were main categories nodes were enriched in.. Additionally, KEGG enrichment was also performed, which was listed in table 4. According to the KEGG analysis, genes were mainly enriched in the pathways of Pathways in cancer, Tight junction, Wnt signaling pathway, Bladder cancer, p53 signaling pathway, and so on.

Discussion

Ovarian cancer is a serious disease with high death rate as lacking of efficient diagnosis in early stage. IRF1 is a member of the interferon regulatory transcription factor family, which is capable of preventing oncogene-mediated malignant transformation and deletion or point mutations of human IRF genes have been reported in some human cancer samples [22]. In the present study, we have identified DEGs and their functions categories that are altered in SK-OV-3 samples with IRF1 silencing for researching the mechanism of IRF1 in ovarian cancer. With IRF1 knocking out, 427 DEGs were observed, including 242 up-regulated genes and 185 down-regulated genes, in which TNF, CDH1, MMP2 and DKK1 were identified as hub genes and play major roles in pathways of morphogenesis of an

epithelium, response to drug, cell-cell adhesion, Wnt signaling pathway and cell junction assembly.

TNF encodes a multifunctional proinflammatory cytokine that belong to the tumor necrosis factor superfamily from the cadherin superfamily, which is highly expressed in ovarian cancer cells, indicating the importance of TNF as a regulator of the proinflammatory tumor microenvironment in this malignancy [23]. For some members of TNF, such as TNF-α, TNF-α mRNA stability is displayed to be increased in malignant ovarian surface epithelial cells [25], and include a central role in a cancer-cell autonomous tumor-promoting network of other cytokinase and chemokines, stimulation of epithelial to mesenchymal translation in malignant cells [26], or further DNA damage to malignant cells [27]. TNF-α binds to 2 receptors to induce a signaling cascade that induces transcriptional regulation of mediators that are key to cell survival, angiogenesis, invasion and impairment of immune surveillance in tumor biology, which promote the occurrence of tumor [28]. With IRF1 silencing, the expression of TNF has been observed up-regulated, which may increase the risk of tumorgenesis.

CDH1, known as a tumor suppressor gene, is a classical cadherin from the cadherin superfamily and is expressed predominantly on the surface of epithelial cells, where it works in the establishment and maintenance of normal tissue architecture [29]. CDH1 encodes a calcium dependent cell-cell adhesion glycoprotein, which consists of an extracellular domain and an intracellular domain [30]. The extracellular domain plays a major role in adhesion specificity and the intracellular cytoplasmic domain interacts with the actin cytoskeleton through a group of membrane associated proteins

[31]. Loss or reduced expression of this gene is associated with poor historical dedifferentiation and increased local invasion and metastases frequency in some human cancer [32] [33] which make it a marker of metaplastic changes in ovarian epithelium. This theory may be able to explain why expression of CDH1 in epithelial ovarian cancer cells went up-regulated with IRF1 silencing. In this study, CDH1 was enriched in the process of cell junction assembly, a cellular process that E-cadherin play a major role in and results in the aggregation, arrangement and bonding together of a set of components to form a cell junction, which is thought to contribute to epithelial tumorigenesis [35].

Both TNF and CDH1 are involved in biological processes of morphogenesis of an epithelium, response to drug and cell-cell adhesion, which are all related closely with cancer cells survive and grow. Morphogenesis of an epithelium is the process in which the anatomical structures of epithelia are generated and organized. One of hypotheses of ovarian cancer is that incessant ovulation causes repeat minor trauma to the surface of the ovary, leading to proliferation of ovarian epithelium [36], during which the process morphogenesis of an epithelium plays a major role. Cell-cell adhesion is disturbed by genetic and epigenetic changes, resulting in changes in signaling, loss of contact inhibition and altered cell migrarion and stromal interactions [37]. Therefore, TNF and CDH1 are the most important genes in the sub-network with high degree and main target genes of IRF1 in ovarian cancer. These results suggest that IRF1 may be able to target TNF and CDH1 and IRF1 inhibit tumoigenesis and ovarian cancer growth by regulating the expression of these two genes.

MMP2 is a member of the matrix metalloproteinase family, which is involved in the breakdown of extracellular matrix in normal physiological process and the function in ovarian cancer has been postulated based on the observation that some members of this family are up-regulated in during neoplastic progression [40]. MMP2 has been associated with the malignant phenotype of tumor cells because of the unique ability to degrade type IV collagen, which is a major component of the basement membrane [41] and the invasiveness of ovarian cancer cell lines will be enhanced with expression of MMP2 [42]. According to the results of pathways analysis on sub-network, MMP2 is enriched in pathways of pathways in cancer and bladder cancer, and biological process of extracellular matrix organization. It has been reported that reorganization of the extracellular matrix has been implicated in ovarian cancer [43], and this may be an approach for IRF1 functions in ovarian cancer.

DKK1 encodes a secreted protein belong to the dickkopf family and is involved in embryonic development through its inhibition of the Wnt signaling pathway, which plays an important role in normal follicular development and ovarian functions [44]. Wnt signaling is a major factor in oncogenesis in human liver, cervix, ovary, colon and other tissues [45] and involved in the epithelial-mesenchymal transition process [46], which is now implicated in the progression of cancer [47]. During the progression of cancer, Wnt-1 protein specifically binds to the frizzled receptor (Fz) and the low-density lipoprotein receptor-related protein-5/6 (LRP5/6), triggering signals important for proliferation via β-catenin [48]. DKK1 binds to LRP5/6 and blocks interaction with Wnt-1, which results in β-catenin degradation and

effects on proliferation [49]. As expected in this study, DKK1 is enriched in the Wnt signaling pathway, whose P value is relatively lower, indicating that DKK1 is a target gene of IRF1, and IRF1 functions in ovarian cancer by impact the expression of DKK1 in the Wnt signaling pathway.

GJA1 is a member of the connexin gene family, encoding a protein as a component of gap junctions, which are composed of intercellular channels arrays that provide a route for the diffusion of low molecular weight materials from cell to cell [50]. Connexin family lines the intercellular channels or gap junctions that connect adjacent cells and exchange of ions, small molecules and secondary messengers [51] and disruption of these cell connections is a cause of tumor [52]. The main pathway GJA1 participates in is cell junction assembly, which is in accordance with the function of connexin family. With IRF1 silencing, the expression of GJA1 goes down-regulated, which may increase the possibility of tumorigenesis.

Conclusion

In conclusion, our study proposed the use of bioinformatics methods to analyzing the role of IRF1 and screening feature genes in ovarian cancer with IRF1 silencing. We observed some significant DEGs, including TNF, CDH1, MMP2, DKK1 and GJA1, which play major roles in some important pathways, such as morphogenesis of an epithelium, cell-cell adhesion and Wnt signaling pathway, suggesting the mechanism of IRF1 in ovarian cancer. However, sample size is relatively less in this study and there is no experiments performed to confirm the conclusion of our study. Further analysis on the microarray data should be performed for the research of ovarian cancer diagnosis in early stage.