A gene interaction network-based method to measure the common and heterogeneous mechanisms of gynecological cancer

MINGYUAN WANG, LIping LI, JINGLAN LIU and JINJIN WANG

The Affiliated Zhuzhou Hospital Xiangya Medical College CSU, Zhuzhou, Hunan 412000, P.R. China

Received November 26, 2017; Accepted March 1, 2018

DOI: 10.3892/mmr.2018.8961

Abstract. Gynecological malignancies are a leading cause of mortality in the female population. The present study intended to identify the association between three severe types of gynecological cancer, specifically ovarian cancer, cervical cancer and endometrial cancer, and to identify the connective driver genes, microRNAs (miRNAs) and biological processes associated with these types of gynecological cancer. In the present study, individual driver genes for each type of cancer were identified using integrated analysis of multiple microarray data. Gene Ontology (GO) has been used widely in functional annotation and enrichment analysis. In the present study, GO enrichment analysis revealed a number of common biological processes involved in gynecological cancer, including ‘cell cycle’ and ‘regulation of macromolecule metabolism’. Kyoto Encyclopedia of Genes and Genomes pathway analysis is a resource for understanding the high-level functions and utilities of a biological system from molecular-level information. In the present study, the most common pathway was ‘cell cycle’. A protein-protein interaction network was constructed to identify a hub of connected genes, including minichromosome maintenance complex component 2 (MCM2), matrix metalloproteinase 2 (MMP2), collagen type 1 α1 chain (COL1A1) and Jun proto-oncogene AP-1 transcription factor subunit (JUN). Survival analysis revealed that the expression of MCM2, MMP2, COL1A1 and JUN was associated with the prognosis of the aforementioned gynecological cancer types. By constructing an miRNA-driver gene network, let-7 targeted the majority of the driver genes. In conclusion, the present study demonstrated a connection model across three types of gynecological cancer, which was useful in identifying potential diagnostic markers and novel therapeutic targets, in addition to in aiding the prediction of the development of cancer as it progresses.

Introduction

Gynecological malignancies, particularly ovarian cancer, cervical cancer and endometrial cancer, are serious medical conditions in women and have been leading causes of cancer mortality in recent years. However, the use of cancer markers for early and progressive detection remain lacking (1). In addition, research has demonstrated that there are close associations across the three aforementioned types of cancer. It has been demonstrated that the progress and the development of the three aforementioned types of cancer are similar, which may be useful when diagnosing any one of these three cancer types. In the case of endometrial cancer, prior to the development of endometrial carcinoma, the endometrium undergoes progressive neoplastic alterations in a parallel fashion to the premalignant alterations observed in the cervix prior to the development of cervical carcinoma (2). The rationale of oophorectomy in surgical management is that endometrial cancer may metastasize to the ovary, in which women with endometrial cancer are at risk for synchronous and metachronous ovarian cancer, and the source of estrogen may be eliminated by oophorectomy (3,4). In cancer cells, oncogenic transformation is associated with major alterations in gene expression (5). With the advent of large-scale screening of cancer genomes, hundreds of genes with alterations in different types of tumors from patients with cancer have been identified (6-10), which revealed that cancer is a complex disease caused by genetic alterations in multiple genes (11,12). In order to elucidate the cancer marker genes and biological processes associated with each type of gynecological tumor, and the potential underlying mechanism of associations among gynecological tumors, the contribution of identified differentially expressed genes (DEGs) to the pathogenesis of gynecological tumors must be understood.

To analyze different DEGs, high-throughput experimental methods, including microarray analysis, have been widely used in a number of studies (13,14). A vast quantity of microarray data has been produced and deposited in publicly-available data repositories, including the Gene Expression Omnibus (GEO) (15). With the methods of integrated bioinformatics
analysis, researchers have been able to advance the identification of genetic signatures. This may provide insights into the underlying biological mechanisms of the development of gynecological tumors.

Chung et al (16) revealed that microRNA (miRNA)-200b/a is a direct transcriptional target of grainyhead like transcription factor 2, which is associated with development and overall survival in epithelial ovarian cancer. Halabi et al (17) demonstrated that 41 genes, including matrix metalloproteinase (MMP)7 and tumor protein 53, were involved in the potential underlying mechanisms of ovarian cancer. Espinosa et al (18) revealed that six genes encoding cyclin B2, cell-division cycle protein 20, protein regulator of cytokinesis 1, synaptonemal complex protein 2, nucleolar and spindle associated protein 1 and cyclin-dependent kinase inhibitor 2 belonging to the mitosis pathway, were potential markers for screening or therapeutic targets of cervical cancer. However, biomarkers which were identified in this way have had poor translation into actual clinical practices. Results have been non-concordant among studies due to small sample sizes. In addition, the studies into the associations of biomarker genes (driver genes) remain lacking among the different types of gynecological tumors.

A robust driver gene biomarker signature may be beneficial for the diagnosis and targeted treatment of gynecological tumors. In the present study, in order to identify a driver gene biomarker signature for the three types of gynecological tumors, data from the Metabolic Gene Rapid Visualizer database (MERAV, which is derived from GEO) was used (19). In MERAV, microarrays were normalized together to eliminate systematic errors caused by different batch experiments.

The present study devised a target network for ovarian cancer, cervical cancer and endometrial cancer using the selected driver genes, and further investigated the identified DEGs via functional enrichment analysis, pathway enrichment analysis and protein-protein interaction (PPI) networks. In addition, the present study extracted clinical information of ovarian cancer, cervical cancer and endometrial cancer from The Cancer Genome Atlas (TCGA) data portal. Subsequently, driver genes in each type of cancer were analyzed. It was important to investigate the underlying mechanism of each gynecological tumor and whether the identified driver genes contributed to these diseases. Subsequently, a network was generated between the miRNAs and the identified driver genes, using the method of mining the Mir2 disease and Tarbase databases which provide information on miRNAs, diseases and the interactions between miRNAs and genes. Finally, the present study determined hub-genes and hub-miRNAs across the gynecological tumors to study the potential underlying mechanisms of the developments of gynecological tumors, which may shed light on different strategies for the design of biological targets for cancer therapies.

Materials and methods

Identification of gene expression datasets. In the present study, DEGs were identified between normal tissues and tumors extracted from the MERAV database from the National Center for Biotechnology Information GEO database (MERAV, http://merav.wi.mit.edu). The experimental samples for the present study are presented in Tables I and II. The following information was extracted from each identified study: GEO accession number, sample type, number of cases and controls, and gene expression data. Studies in which the microarray data were uncertain were excluded. The experimental protocol for the present study is presented in Fig. 1.

Integrated analysis of DEGs identified in the extracted databases. Information was extracted from the microarray datasets in MERA V which are presented in Tables I and II, respectively. Following the intersection of the microarray datasets, the DEGs were established between the normal and cancer tissues. In the present study, the degree of differential gene expression was measured by fold-change based on the Student's t-test. A fold-change value >2 or <0.5 and t-test P<0.01 for a gene was considered to be significant. The differential expression analysis was conducted using the Linear Models for Microarray Data package in R (20).

Protein interaction network. The DEGs were subsequently applied to the Human Protein Reference Database (21) (HPRD, www.hprd.org), to identify the more complex functional interactive driver genes of separate cancer types. Genes with interactions with each other were extracted from the DEGs as mentioned above (presented in Tables III-X). The PPI network is a useful research tool for investigating the cellular networks of protein interactions, and was downloaded from the HPRD. Cancer-associated gene-gene interaction networks were constructed by mapping the DEGs into the HPRD PPI network for each cancer (cervix tumor, ovarian tumor and endometrium tumor). To make it easier to identify the driver genes, the present study calculated the lines attached to each node, which was defined as the degree of the node. The nodes that exhibited degrees ≥4 were defined as driver genes. The nodes whose degree was ≥4 were considered to serve more complex roles in the development of the diseases of interest. These nodes were then extracted for the PPI network (Fig. 2). The present study constructed a connected network which contained the driver genes across the three cancer types. Through this method, it was determined whether the driver genes of the separate cancer types had any interaction with each other. The networks were constructed using Cytoscape version 3.3.0 (www.cytoscape.org).

miRNAs regulating gene network construction. The present study analyzed the association between miRNAs and the identified driver genes (Fig. 3). This process was performed by extracting a list of miRNAs which were associated with the type of cancer (cervical tumor, ovarian tumor or endometrial tumor) from the Mir2 Disease database (www.mir2disease.org) (22). Following this step, a network was created regarding the regulatory associations between the miRNA and the specific driver gene of each type of cancer in order to identify the hub-miRNAs of the gynecological tumors. The associations of the regulation were extracted from Tarbase (diana.cslab.ece.ntua.gr/tarbase) (23).

Functional and pathway enrichment analysis. In order to assess the functional relevance of the aforementioned DEGs, a pathway analysis was created based on the Database for Annotation, Visualization and Integrated Discovery (DAVID) (24). DAVID provides a useful tool to analyze large gene lists, including gene
ontology (GO) and pathway analysis. DEGs in different diseases were applied to this database in order to detect potentially represented functions. GO-categories were organized based on the GO database (25) (www.geneontology.org). In addition, pathway analysis was based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (26) (genome.jp/kegg).

Significant categories were identified by expression analysis systematic explorer scores, a modified Fisher's exact P‑value. The threshold for significance for a category was considered to be P<0.01, with >4 genes for the corresponding term.

Survival analysis. The present study used TCGA database to extract clinical information and gene expression profile information. At the start of the analysis, the expression values of each driver gene were listed, which were identified via the PPI network. To find the median level of gene expression, the samples were divided into two groups by median of expression (high expression group and low expression group). Additionally, the corresponding clinical information of each sample was extracted. Survival data representing time between initial diagnosis and mortality were downloaded directly from TCGA data portal (tcga-data.nci.nih.gov/tcga/tcgaHome2.jsp) (27). With this information, the present study was able to estimate the association between the identified driver genes of the three types of cancer mentioned above and the survival rates of patients. All analyses were conducted using custom-written code in R (www.r-project.org).

Results

Integrated analysis of multiple studies to establish the driver genes in cancer. There are multiple genes that contribute to
Table III. Driver genes identified by integrated analysis of the microarray datasets (cervical squamous cell carcinoma).

| Gene       | Gene  | Gene  | Gene       | Gene       | Gene       |
|------------|-------|-------|------------|------------|------------|
| RB1        | HTRA1 | MTOR  | CLDN5      | NARF       | PURA       |
| MCM7       | KPN2A | PLSCR4| CYBA       | NCAPD2     | RBM8A      |
| MCM2       | LMNB1 | PRKD1 | DCUN1D1    | NCF4       | RECK       |
| PLK1       | MEIS1 | PSMA5 | DDAD4      | NME4       | REV3L      |
| AR         | NCOA1 | PSMB10| DMPK       | NLOC4      | RFC3       |
| PPP1CA     | PBX1  | PSMB9 | EPS8       | N2F1       | RNF126     |
| ABL1       | PIAS3 | PSMD2 | EXOSC5     | N2F2       | RPA3       |
| LMNA       | POLA2 | RACGAP1| GABBR1     | NRAS       | RRM1       |
| PTN        | PPP1R14A | RTN3 | GAS6       | NTF3       | RRM2       |
| TRIP13     | AXL   | SNRPB | GCH1       | NTRK2      | SAT2       |
| CAV1       | BUB1B | TOR1AIP1| GCHFR      | NUB1       | SDC2       |
| CDC20      | CCL14 | TUBA4A| GLRX3      | NUP210     | SEC24A     |
| CDC6       | CCR5  | UBTF  | GMFB       | NUP50      | SELENBP1   |
| FLNA       | COL4A5| USP6NL| GOLGA2     | PAFAB1B3   | SERBP1     |
| FXR2       | CSNK1D| UTP3  | HOXD13     | PAK2       | SH3BP5     |
| ZHX1       | DBF4  | ACTN4 | ILK        | PAM        | SMC4       |
| CCNA2      | DVL3  | ADAM10| KANK1      | PCGF2      | SNRPD1     |
| DGKZ       | EFEMP2| ANTXR2| LAPT5M     | PHACTR4    | SNTB2      |
| MCM10      | EIF4EBP1| ARHGAP17| LDB2       | PLK2       | SNX27      |
| MCM6       | EZH2  | ASPM  | LDC01      | PNO1       | SPIN1      |
| PCNA       | FAM46A| BID   | LMO4       | PNP        | SSSCA1     |
| RBPMS      | HOXD10| BMP4  | LRP1       | PPIA       | STXB2P     |
| RPS6KA1    | HSPA4 | BNIP2 | LRP6       | PPH        | SUB1       |
| SAT1       | ITGB3BP| CIQA  | LRPC41     | PRPF18     | TALDO1     |
| BUB1       | KLF6  | CBX4  | LZZT2      | PSA6       | TGBR3      |
| CSNK1E     | MAD2L2| CCNE1 | MAGEH1     | PSMB7      | TNFRSF1A   |
| DCN        | MAP2K4| CCR1  | MELK       | PSMD4      | UFD1L      |
| FGFR1      | MAPK10| CDC42BPA| MPDZ       | PSM3E      | WSB2       |
| FXYD1      | MCM5  | CENPE | MTA1       | PSMF1      | XPNPEP1    |
| GMNN       | MITF  | CHFR  | MYCBP      | PSTTIP1    | YLPM1      |
| HOXA10     | MMP9  | CIB1  | MYL9       | PTTG1      | ZMIZ1      |

Figure 1. Experimental protocol of the present study. DEG, differentially express genes; GO, gene ontology; MERAV, Metabolic Gene Rapid Visualizer database; TGCA, The Cancer Genome Atlas.
the cause of the aforementioned cancer types and, therefore, no single gene is a determining factor in diagnosis. It was identified that each type of cancer was driven by different variations of genes that serve key roles during the development of pathology. However, no single gene may explain the heterogeneity of each type of cancer. In the case of cervical cancer, 186 genes in squamous cell carcinoma of the cervix (Table III), 107 genes in keratinized squamous cell carcinoma of the cervix (Table IV), 96 genes in cervical adenocarcinoma Grade 3 (Table V), 133 genes in non-keratinized squamous cell carcinoma of the cervix (Table VI) and 203 genes in cervical adenocarcinoma Grade 2 (Table VII) were identified to be important. In addition, 120 genes and 76 genes were established, respectively, in adenocarcinoma of the ovary Grade 2

Table IV. Driver genes identified by integrated analysis of the microarray datasets (cervical keratinized squamous cell carcinoma).

| Gene    | Gene    | Gene    | Gene    | Gene    | Gene    |
|---------|---------|---------|---------|---------|---------|
| FYN     | ADAM10  | ARHGAP17| HSPB2   | PHF1    | TMOD1   |
| ZHX1    | ADAM17  | ARMCX2  | ID4     | PIK3C2B | TMSB10  |
| ABL1    | ANXA6   | BIN1    | LDB2    | PIP5K1C | TPD52   |
| BCL2L1  | AXL     | CBX3    | LDOC1   | PNP     | UBTF    |
| FXR2    | BCL11A  | CLDN5   | LMO4    | PSME3   | ZHX2    |
| TBP     | CSNK1E  | CNN3    | LRP1    | PSMF1   | ZMIZ1   |
| AR      | DMPK    | CNNM3   | LRP6    | PTPN12  | ZNF76   |
| BARD1   | ITGB3BP | CNTNAP1 | LSM5    |         |         |
| BID     | KPNL6   | CRYAB   | MAGI2   | RAE1    |         |
| DDX24   | MAD2L2  | CSE1L   | MAPK10  | REV3L   |         |
| NCOA1   | MCL1    | CSTF1   | MIS12   | RUNX1T1 |         |
| PDGFRB  | NR2F1   | EFEMP2  | MPDZ    | SDC2    |         |
| PRKD1   | NTRK2   | EXOSC5  | MTA1    | SFRP1   |         |
| PSEN1   | PPP1R14A| FGFR1   | MYCBP   | SH3BP5  |         |
| RBPMS   | PTN     | FXYD1   | NPDC1   | TAF9    |         |
| SPTAN1  | RTN3    | FZD6    | NR2F1   | TCF7L2  |         |
| TCF4    | SYK     | GAS6    | NTF3    | TERF1   |         |
| TGFA    | VIM     | GD1     | NUDT21  | TFP1    |         |
| A2M     | ANTXR2  | GTF3C3  | PBX1    | TGFB3   |         |
| ACP1    | AQPI    | HOXA10  | PDGFD   | TLN2    |         |

Table V. Driver genes identified by integrated analysis of the microarray datasets (cervical adenocarcinoma G3).

| Gene    | Gene    | Gene    | Gene    | Gene    | Gene    |
|---------|---------|---------|---------|---------|---------|
| AR      | BAD     | PLD2    | CIB1    | MAPK10  | SERPIN1A|
| CAV1    | BAH1D1  | PPA1    | CLDN5   | MED14   | SF1     |
| FLNA    | C1QBP   | PRKD1   | CUL4B   | MPDZ    | SMO     |
| PPP1CA  | CPSF6   | SAT1    | DMPK    | MYL9    | SPINT2  |
| NCK2    | CSNK1D  | SMAD1   | EFN1B   | NR2F1   | SSBP3   |
| PLC5R1  | DOCK1   | SNAP23  | F3      | NTF3    | SRR1    |
| SUMO4   | DVL2    | TAF1D   | GDI1    | PGP2F   | STAM    |
| LMNA    | FXE2    | TAF9    | HOXA10  | PDPK1   | SYNE1   |
| LRP1    | FXYD1   | TCF4    | HOXDI0  | PHACTR4 | TCF7L2  |
| PSEN1   | ILK     | WIP1    | HOXDI3  | PHYHHP  | TGFBR3  |
| PTN     | LDB1    | ACVR2A  | HSPA1B  | PLC4R   | TMF1    |
| CSNK1E  | LMO4    | ANTXR2  | HSPBAP1 | PNPLA2  | UBF1    |
| DVL3    | MAP2K4  | ATG12   | KANK1   | PPP1R10 | VAMP8   |
| MMP14   | NCOA1   | CD82    | KPN1A   | PTCH2   | WASF1   |
| PPP1R14A| NTRK2   | CDC42BA | LDB2    | RNF138  | WASF2   |
| ALDOA   | PBX1    | CDC42EP1| LRP6    | RUNX1T1 | ZHX2    |
Table VI. Driver genes identified by integrated analysis of the microarray datasets (cervical non-keratinized squamous cell carcinoma).

| Gene   |
|--------|
| AR     |
| ABL1   |
| CAV1   |
| CHD3   |
| HIF1A  |
| PTPN6  |
| SAT1   |
| FLNA   |
| HOXA10 |
| PLSCR1 |
| RAF1   |
| DCN    |
| EZR    |
| MMP14  |
| PDGFRB |
| ABCA1  |
| C1QA   |
| CSNK1E |
| DMPK   |
| ELN    |
| SNX2   |
| TMEM8B |
| WASF3  |

and Grade 3 (Tables VIII and IX). A total of 168 genes were established in endometrial carcinoma (Table X).

**Integrated PPI (protein-protein interactions) network construction.** Based on the HPRD, the interaction network of the identified driver genes was constructed, which consisted of 101 nodes (genes that form associations) and 185 edges (biological association) (Fig. 2). Genes with a higher degree of association (degree ≥4) were observed to be larger in size, and included the genes CDK1, CAV1, ZBTB16, Jun proto-oncogene AP-1 transcription factor subunit (JUN), RAF1, RB1, mini-chromosome maintenance complex component 2 (MCM2), AR, ABL1, LMNA, FLNA, DCN, FYN, SMAD1, LRP1, PSEN1, EP300, CTNND1, collagen type I α1 chain (COL1A1) and FOS. Through this method, it was identified that driver genes in each gynecological cancer have contact interactions.

**Comprehensive analysis of miRNA regulation and the selected driver genes.** Fig. 3 illustrates that certain miRNAs serve important roles in regulating the driver genes. In the present study, it was demonstrated that a number of miRNAs regulate separate networks [for example the let7 family, miRNA (miR)-23b, miR-21, miR-214 and miR-218]. miRNAs that were confirmed to be significant in cervical cancer, including let7c and let7b, are also found to be associated with the other two cancers in this study. This information may be important in establishing the connections between the three gynecological cancer types, which may be used in the development of targets for further research and diagnosis.

**Functional and pathway enrichment analysis.** GO analysis revealed that the identified genes of cervical tumors, ovarian tumors and endometrial tumors were predominantly involved in the illustrated biological processes (Fig. 4). The top three significant biological processes of cervical cancer were 'mitotic cell cycle', 'cell cycle' and 'cell cycle process', while for ovarian cancer, the biological processes consisted of 'cell cycle process', 'cell cycle phase' and 'macromolecule metabolic process'. For the progression of endometrial cancer, the top three biological processes observed to be at fault for cancer progression were 'response to organic substance', 'regulation of cell proliferation' and 'skeletal system development'.

Using the method of pathway analysis, it was revealed that genes in cervical cancer were significantly enriched in 'cell cycle', 'pathways in cancer' and 'DNA replication'. Ovarian cancer was observed to be significantly enriched in 'MAPK signaling pathway', 'cell cycle' and 'DNA replication'. Ovarian cancer was observed to be significantly enriched in 'MAPK signaling pathway', 'cell cycle' and 'DNA replication'. Endometrial cancer was observed to be significantly enriched in 'pathways in cancer', 'focal adhesion' and 'complement and coagulation cascades' (Fig. 5).

**Survival analysis of patients with gynecological tumor.** Fig. 6 illustrates the association between survival time and survival rate in the high and low expression groups. The genes MCM2,
MMP2, COL1A1 and JUN are presented in the figure, and it was observed that the driver genes of the expression groups were able to divide each of the target cancer types into two groups, one of which contained the high expression group with the other containing the low expression group. Therefore, in order to determine whether the driver genes had a key role in the development of gynecological tumors and the connective function of separate cancer types, the present study aimed to identify the association between the target cancer driver genes and other types of gynecological cancer.

Discussion

The principal challenge of high-throughput cancer genomics is to identify specific driver genes and the underlying mechanisms of carcinogenesis, apart from the vast quantity of heterogeneous genomic alteration data. Numerous studies have focused on identifying individual functional modules or pathways involved in cancer (28-30). Based on this methodology, the analysis of the present study focused specifically on DEGs in order to reveal the transcriptional responses of gynecological tumors. The results of this analysis suggested that the common biological processes of cancer of the cervix, ovary and endometrium were those involved in the cell cycle and the regulation of macromolecule metabolism.

The cell cycle is the progression of biochemical and morphological phases and events that occur in a cell during successive cell replication or nuclear replication. Research has shown that interference with cell cycle components may...
lead to tumor formation (31). Certain cell cycle inhibitors, including retinoblastoma protein and tumor protein 53 may mutate during replication, causing the cell to proliferate uncontrollably, ultimately resulting in a tumor. Furthermore, the proportion of active cell division in tumors is much higher compared with the rate in normal tissue.

To clarify the hub genes in ovarian cancer, cervical cancer and endometrial cancer, DEGs were predicted to be biomarkers for each cancer using PPI networks. It is considered that hub nodes are genes that are highly connected with other genes and have been predicted to serve key roles in numerous networks. In addition, highly connected hub genes were proposed to have a considerable role in biological development. Hub nodes have more complex interactions compared with those of other nodes, which indicates that they have pivotal roles in the underlying mechanisms of disease. In addition, certain identified biomarkers of each type of cancer were extracted from each network and these driver genes were placed into one PPI network.

Table VIII. Driver genes identified by integrated analysis of the microarray datasets (adenocarcinoma of the ovary Grade 2).

| Gene     | CDK1       | HLA-DRA | MEFC2 | HSPA1A | JUN       | MEF2C |
|----------|------------|---------|-------|--------|-----------|-------|
|          | FOS        | GNE     | COX5A | GNE    | CNNM3     | PHF1  |
|          | GCA        | SOX9    |        |        |           |       |
|          | RAF1       | FCGR2B  |        |        | PDGFD     |       |
|          | RAF1       |         |        |        | NR2F2     |       |

Table IX. Driver genes identified by integrated analysis of the microarray datasets (adenocarcinoma of the ovary Grade 3).

| Gene     | CDK1       | HLA-DRA | MEFC2 | HSPA1A | JUN       | MEF2C |
|----------|------------|---------|-------|--------|-----------|-------|
|          | FOS        | GNE     | COX5A | GNE    | CNNM3     | PHF1  |
|          | GCA        | SOX9    |        |        |           |       |
|          | RA1        | FCGR2B  |        |        | PDGFD     |       |
|          | RA1        |         |        |        | NR2F2     |       |
Table X. Driver genes identified by the integrated analysis of the microarray datasets (endometrial carcinoma).

| Gene  | Gene  | Gene  | Gene  | Gene  |
|-------|-------|-------|-------|-------|
| EP300 | CDKN2A| F2R   | AMFR  | EPN3  |
| JUN   | COL3A1| FZD5  | AXL   | EPR1  |
| CAV1  | EGR1  | HLA-DMB| BCL1A | FOSB  |
| CTNNB1| ERBB4 | HOXA10| BCL2A1| GALNT10|
| ABL1  | FBLN1 | ID1   | BIK   | GAS6  |
| AR    | FBN1  | ID4   | BLNK  | GATA2 |
| TCF4  | FLNA  | IDE   | C1R   | GCH1  |
| THBS1 | FOXO1 | INADL | C1S   | GCHFR |
| TUBA4A| HLA-DRA| JUND | C3AR1 | GPI   |
| ATXN1 | ID3   | LMO4  | CCND2 | GPRASP1|
| COL1A1| IGFBP5| LNX1  | CDH11 | HLA-DQB1|
| DCN   | LAMB3 | NCALD | CDKN1A| HLA-DRB1|
| LRPI  | MIF   | NCF2  | CDKN2C| HLF   |
| C3    | MYC   | NR2F2 | CFB   | HOXA9 |
| COL7A1| PLAT  | PDGFRA| CGN   | ID2   |
| FBLN2 | RUNX1T1| PLEKHF2| CLEC3B| IGFBP4 |
| FOS   | S100A8| PTPN13| CLK1  | IGFBP6 |
| GNAI2 | SERPINA1| RAB8B| CXADR | IL33   |
| IGF1  | SYK   | RABAC1| CXCL10| IRS1  |
| LAMC2 | TGFBI1| ROR2  | DNMT1 | KLF5   |
| MUC1  | CD14  | SFN   | DPYSL2| LAPT5M|
| NID1  | COL5A1| SFRP1 | ECM1  | LDB2  |
| PRKD1 | CRMP1 | TFAP2A| EDNRA | LUC7L3 |
| PTPN12| DBP   | TJP2  | EFGMP2| MAFB  |
| VCAN  | DDR2  | TRPC1 | EFS   | MAL2  |
| CD74  | F10   | WNT5A | ENO2  | MAPK10|
| SYTL1 | TJP3  | TLR3  | TRO   | WASP2 |
| TBL1X | TLR2  | TPD52 | USP54 | WNT2  |

Figure 2. Protein–protein interaction networks of the DEGs identified by integrated analysis of the microarray databases throughout cancer of the cervix, ovary or endometrium. Each cancer holds a number of DEGs. Driver genes were extracted from the DEGs, whose degree (the number of lines attached to each node) was ≥4. The orange dots represent cervical carcinoma, green dots represent ovarian carcinoma and blue dots represent endometrial carcinoma. Genes with a higher degree of association exhibit a larger node size. Each biological association (an edge) between two genes (nodes) was supported by at least one reference from the literature or information stored in the Human Protein Reference Database. DEGs, differentially expressed genes.
network with the duplication hub genes eliminated. Therefore, the particular hub genes of each gynecological cancer and the connection nodes across the three types of cancers may be identified. Accordingly, the identification of hub genes and hub connected genes involved in each gynecological cancer may lead to the discovery of the association across ovarian cancer, cervical cancer and endometrial cancer, and may lead to the development of effective diagnostic and therapeutic approaches.

In order to ascertain a causal association across the three types of gynecological cancer, the present study extracted clinical information and gene expression profile information from TCGA database, and used the hub connected genes identified in the PPI network to perform survival analysis. In the present study, four noteworthy genes were identified, including MCM2, MMP2, COL1A1 and JUN.

The present study demonstrated that MCM2 may serve a key role in cervical cancer. A poor prognosis was associated with lower expression. Furthermore, MCM2 was highly connected with ovarian cancer and endometrial cancer. The results suggested that MCM2 is a component of the DNA replication licensing complex, with a rich binding surface that directs multiple regulatory interactions of cancer significance, marking DNA replication origins during the G1 phase of the cell cycle for use in the subsequent S-phase. A deficiency of MCM2 results in death or morbidity in the absence of an overt tumor (32). These processes of DNA replication have been studied and used as therapeutic targets. Simon and Schwacha (33) suggested that MCM2 was a promising target for blocking the proliferation of cancerous and precancerous cells.

In the present study, MMP2 was identified to be essential in causing cervical cancer. MMPs are zinc-containing endopeptidases with an extensive range of substrate specificities. These enzymes are able to degrade various components of extracellular matrix (ECM) proteins. In photocarcinogenesis,
degradation of the ECM is the initial step towards tumor cell invasion, to intrude in the basement membrane and the surrounding stroma that primarily comprises fibrillary collagens. Additionally, MMP2 is involved in angiogenesis, which promotes cancer cell growth and migration (34).

COL1A1 and COL1A2 encode the $\alpha_1$ and $\alpha_2$ chains of type I collagen, respectively (35). The primary constituents of the ECM are collagens, adhesive glycoproteins and proteoglycans (36). Specific interactions between cells and ECM-mediated cell-surface-associated components and transmembrane molecules result in the control of cellular activities, including adhesion and migration (37). Collagen is the primary component of the ECM, which serves pivotal roles in maintaining skin and vessel elasticity, and increasing cartilage lubricity (38). Upregulation of type II collagen expression may contribute to ovarian cancer metastasis and biological processes, including cell proliferation, invasion and migration (39). The oncogene JUN is the putative transforming gene of avian sarcoma virus 17, which is the most extensively studied protein of the activator protein-1 complex and is involved in numerous cell activities, including proliferation, apoptosis, survival, tumorigenesis and tissue morphogenesis. The present study identified that COL1A1 was important in ovarian cancer, which was highly connected with cervical and endometrial cancer. Therefore, COL1A1 and JUN may be potentially important associated genes of the three types of gynecological malignancies.

miRNAs are small noncoding regulatory RNAs that downregulate transcription by targeting specific mRNAs. Furthermore, the present study identified that certain miRNAs were highly associated with hub connected genes, including let7, which is one of the founding members of the miRNA family. This miRNA was first identified in Caenorhabditis elegans. Lee and Dutta (40) identified six functional let7 target sites in the 3'-untranslated region of high mobility group AT-hook 2 (HMGA2), which reduced HMGA2 expression and cell proliferation in a lung cancer cell line. Using genome-wide mRNA expression analysis, Mi et al (41) identified that miRNA let7B was downregulated in acute lymphoblastic leukemia (ALL) compared with acute myeloid leukemia (AML). Quantitative polymerase chain reaction analysis confirmed the downregulation of let7B in ALL samples compared with AML samples and normal controls.

The present study identified that let7a, let7b and let7c had strong connections with the hub genes and that these miRNAs may serve an important part of the potential mechanism, which may explain the connections across the hub genes.

Overall, the present study identified a number of DEGs associated with gynecological cancer, in addition to the functions and signaling pathways in which these genes were involved. Comprehensive network analyses of the dysregulated gene expression in gynecological cancers identified a series of hub genes and the connection genes across ovarian cancer, cervical cancer and endometrial cancer in a PPI network. Subsequently, this study confirmed the driver genes by survival analysis using the TCGA database. Comprehensive network analyses
of miRNAs and connection driver genes identified certain miRNAs which may be potential therapeutic and prevention targets of gynecological cancer. In addition, the present study demonstrated the associations across the different gynecological cancers, which may be useful for identifying potential useful diagnostic markers and novel therapeutic targets. The results of this study may provide an insight into the underlying mechanism of the aforementioned gynecological cancers and may lead to further improvement in diagnosis and treatment of them.

Acknowledgements

The authors would like to thank Professor Yunyan Gu (College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China.) for technical support and critically reviewing the manuscript.

Funding

The present study was supported by grant no. RC2013QN004112 from Harbin Science and Technology Innovation Talents, China.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors’ contributions

MY and JW conceived and designed the study; MY, LL and JL performed the experiments and analyzed the data. MY wrote the paper, and JW revised the manuscript and gave final approval of the version to be published.

Ethics approval and consent to participate

The present study was approved by the Clinical Research Ethics Committee of the Affiliated Zhuzhou Hospital Xiangya Medical College CSU (Zhuzhou, China), and written informed consent was obtained from all participants.

Consent for publication

Written informed consent was obtained from all volunteers for the publication of any associated data.
Competing interests
The authors declare that they have no competing interests.

References

1. Li XY and Wang X: The role of human cervical cancer oncogene in cancer progression. Int J Clin Exp Med 8: 8363-8368, 2015.
2. Niskar JA: Screening for endometrial cancer. Can Fam Physician 29: 961-965, 1983.
3. Setiawan VW, Yang HP, Pike MC, McCann SE, Yu H, Xiang YB, Wolk A, Wentzensen N, Weiss NS, Webb PM, et al: Type I and II endometrial cancers: Have they different risk factors? J Clin Oncol 31: 2607-2618, 2013.
4. Wright JD: Take ‘em or leave ‘em: Management of the ovaries in young women with endometrial cancer. Gynecol Oncol 131: 287-288, 2013.
5. Lukk M, Kapushesky M, Nikkilä J, Parkerin H, Goncalves A, Hube W, Ukkonen E and Brazma A: A global map of human gene expression. Nat Biotechnol 28: 322-324, 2010.
6. Cancer Genome Atlas Research Network; McM lendon R, Friedman A, Bignier J, Van Meir EG, Brat DJ, Mastrogiannakis GM, Olson JJ, Mikkelson T, Lehman N, et al: Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 455: 1061-1068, 2008.
7. Cancer Genome Atlas Research Network: Integrated genomic analyses of ovarian carcinoma. Nature 474: 609-615, 2011.
8. Guichard C, Amaddeo G, Imbeaud S, Ladeiro Y, Pelletier L, Maad IB, Calderaro J, Bioulac-Sage P, Letexier M, Degos F, et al: Integrated analysis of somatic mutations and clonal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nat Genet 44: 604-609, 2012.
9. Liu P, Morrison C, Wang L, Xiong D, Vedell P, Cui P, Hua X, Ding F, Lu Y, James M, et al: Identification of somatic mutations in non-small cell lung carcinomas using whole-exome sequencing. Carcinogenesis 33: 1270-1276, 2012.
10. Nie-Cristofaro S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stephens LA, et al: Mutational processes molding the genomes of 21 breast cancers. Cell 149: 979-993, 2012.
11. Baudoit A, Real FX, Izarzagaza JM and Valencia A: From cancer genomes to cancer models: Bridging the gaps. EMBO Rep 10: 259-366, 2009.
12. Kreeger PK and Lauffenburger DA: Cancer systems biology: A network modeling perspective. Carcinogenesis 31: 2-8, 2010.
13. Lee S, Stewart S, Nagtegaal I, Luo J, Wu Y, Colditz G, Medina D and Allred DC: Differentially expressed genes regulating the progression of ductal carcinoma in situ to invasive breast cancer. Cancer Res 65: 694-698, 2005.
14. Wray CJ, Ko TC and Tan FK: Secondary use of existing public microarray data to predict outcome for hepatocellular carcinoma. J Surg Res 188: 137-142, 2014.
15. Barrett T, Troup DB, Wilhite SE, L Edenbux P, Evangelista C, Wray CJ, Ko TC and Tan FK: Secondary use of existing public microarray data to predict outcome for hepatocellular carcinoma. J Surg Res 188: 137-142, 2014.
16. Chung VY, Tan TZ, Tan M, Wong MK, Kuay KT, Yang Z, Ye J, Muller J, Koh CM, Guccione E, et al: GRHL2-mir-200-ZEB1 maintains the epithelial status of ovarian cancer through transcriptional regulation and histone modification. Sci Rep 6: 19943, 2016.
17. Halabi NM, Martinez A, Al-Farsi H, Mery E, Puyden L, Pujol P, Khalak HG, McLurcan C, Ferron G, Querleu D, et al: Preferential allele expression analysis identifies shared germline and somatic driver genes in advanced ovarian cancer. PLoS Genet 12: e1005892, 2016.
18. Espinosa AM, Alfaro A, Roman-Basaure E, Guardado-Estrada M, Palma I, Serralde C, Medina J, Juñer E, Bermúdez M, Márquez E, et al: Mitosis is a source of potential markers for screening and survival and therapeutic targets in cervical cancer. PLoS One 8: e55975, 2013.
19. Shaul YD, Yuan B, Thiru P, Nutter-Upham A, McCallum S, Lanzkron C, Bell GW and Sabatini DM: MERAV: A tool for comparing gene expression across human tissues and cell types. Nucleic Acids Res 44: D560-D566, 2016.
20. R Development Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2017.
21. Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S, Telikicherla D, Raju R, Shafreen B, Venugopal A, et al: Human protein reference database-2009 update. Nucleic Acids Res 37 (Database Issue): D767-D772, 2009.
22. Jiang Q, Wang Y, Hao Y, Juan L, Teng M, Zhang X, Li M, Wang G and Liu Y: miR2Disease: A manually curated database for microRNA deregulation in human disease. Nucleic Acids Res 37 (Database Issue): D98-D104, 2009.
23. Vlachos IS, Paraskevopoulos M, Karagouni D, Georgakilas G, Vergoulis T, Kanellos I, Anastasopoulos IL, Maniou S, Karathanou K, Kalfakakou D, et al: DIANA-TarBase v7.0: Indexing more than half a million experimentally supported miRNA:miRNA interactions. Nucleic Acids Res 43 (Database Issue): D153-D159, 2015.
24. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC and Lempicki RA: DAVID: Database for annotation, visualization, and integrated discovery. Genome Biol 4: P3, 2003.
25. Kaimal V, Bardes EB, Tabar SC, Jegga AG and Aronov BJ: ToppCluster: A multiple gene list feature analyzer for comparative enrichment clustering and network-based dissection of biological systems. Nucleic Acids Res 38: W96-W102, 2010.
26. Schummer M and Patilharm-intE: PathwayVoyager: Pathway mapping using the Kyoto encyclopedia of genes and genomes (KEGG) database. BMC Genomics 6: 60, 2005.
27. Stamoulis C and Betensky RA: A novel signal processing approach for the detection of copy number variations in the human genome. Bioinformatics 27: 3238-3245, 2011.
28. Miller CA, Settle SH, Sulpin EM, Alpade KD and Miloslavjvic A: Discovering functional modules by identifying recurrent and mutually exclusive mutational patterns in tumors. BMC Med Genomics 4: 34, 2011.
29. Cirriello G, Cerami E, Sander C and Schultz N: Mutual exclusivity analysis identifies oncogetic network modules. Genome Res 22: 398-406, 2012.
30. Vandin F, Upfal E and Raphael BJ: De novo discovery of mutated driver pathways in cancer. Genome Res 22: 375-385, 2012.
31. Champeris T, Sinaras S, Kanelakis N, Symeonidou IE, Nikolopoulou P, Lygerou Z and Taraviras S: Licensing of DNA replication, cancer, pluripotency and differentiation: An interlinked world? Semin Cell Dev Biol 30: 174-180, 2014.
32. Pruitt SC, Bailey KJ and Freeland A: Reduced Mcm2 expression results in severe stem/progenitor cell deficiency and cancer. Stem Cells 25: 3121-3132, 2007.
33. Simon NE and Schwacha A: The Mcm2-7 replicative helicase: A promising chemotherapeutic target. Biomed Res Int 2014: 549719, 2014.
34. Peterspruc P, Meephanan P, Prapanon O, Komine M and Ohtsuki M: Role of matrix metalloproteinases in photoaging and photocarcinogenesis. Int J Mol Sci 17: pii:E868, 2016.
35. Chan TF, Poon A, Basu A, Addleman NR, Chen J, Phong A, Byers PH, Klein TE and Kwok PF: Natural variation in four human collagen genes across an ethnically diverse population. Genomics 91: 307-314, 2008.
36. Bosman FT and Stamenkovic I: Functional structure and composition of the extracellular matrix. J Pathol 200: 423-428, 2003.
37. Uitto VJ and Larjava H: Extracellular matrix molecules and their receptors: An overview with special emphasis on periodontal tissues, Crit Rev Oral Biol Med 2: 323-354, 1991.
38. Deyl Z, Miskik I and Eckhardt A: Preparative procedures and purity assessment of collagen proteins. J Chromatogr A 914: 275-278, 2001.
39. Day J, Wang T, Wang W, Zhang S, Liao Y and Chen J: Role of MAPK7 in cell proliferation and metastasis in ovarian cancer. Int J Clin Exp Pathol 8: 1044-1045, 2015.
40. Lee YS and Dutta A: The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. Genes Dev 21: 1025-1030, 2007.
41. Mi S, Lu J, Sun M, Li Z, Zhang H, Neilly MB, Wang Y, Qian Z, Jin J, Zhang Y, et al: MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia. Proc Natl Acad Sci USA 104: 19971-19976, 2007.