

文章编号 1006-8147(2022)01-0008-07

论 著

## 基于生物信息学分析筛选宫颈癌进展中的关键基因

刘玉林, 臧玉琴, 王颖梅, 薛凤霞

(天津医科大学总医院妇产科; 天津市女性生殖健康与优生重点实验室, 天津 300052)

**摘要** 目的: 通过生物信息学分析的方法, 筛选宫颈癌进展中的关键基因, 寻找潜在的分子标志物和治疗靶点。方法: 在 GEO (Gene Expression Omnibus) 数据库中检索宫颈癌进展相关的 mRNA 和 miRNA 基因芯片数据, 借助 GEO2R 分析差异表达基因 (DEG) 和差异表达 miRNA (DEM), 进行富集分析以及构建蛋白质-蛋白质相互作用 (PPI) 网络和 miRNA-靶基因调控网络。结果: 共筛选出 250 个 DEG 和 166 个 DEM, 并构建出由 123 个节点 (node) 和 283 项相互作用 (edge) 构成的 PPI 网络以及由 66 个节点和 137 个相互作用构成的 miRNA-靶基因调控网络。经分析, ATAD2、SMC4 和 POLQ 基因不仅是筛选出的表达上调的 DEG, 而且是 PPI 网络中枢纽蛋白的编码基因, 并在 miRNA-靶基因调控网络中同时受 DEM—miR-20A、miR-20B、miR-106B 和 miR-17-5P 的调控。结论: ATAD2、SMC4 和 POLQ 基因可能在宫颈癌进展过程中发挥着重要作用。

**关键词** 宫颈癌; 生物信息学分析; 差异表达基因; 差异表达 miRNA; 蛋白质-蛋白质相互作用网络; miRNA-靶基因调控网络  
中图分类号 R737.33 文献标志码 A

## Identification of key genes in cervical cancer progression by integrated bioinformatics analysis

LIU Yu-lin, ZANG Yu-qin, WANG Ying-mei, XUE Feng-xia

(Department of Gynecology and Obstetrics, The General Hospital, Tianjin Medical University; Tianjin Key Laboratory of Female Reproductive Health and Eugenics, Tianjin 300052, China)

**Abstract** **Objective:** To identify the key genes in cervical cancer progression by integrated bioinformatics analysis to search for potential biomarkers and therapeutic targets. **Methods:** The mRNA and miRNA microarray datasets related to the progression of cervical cancer were searched on the GEO (the Gene Expression Omnibus) database and differentially expressed gene (DEG) and differentially expressed-miRNA (DEM) were analyzed by GEO2R. Then, the functional enrichment analyses, protein-protein interaction (PPI) network and miRNA-target regulatory network construction were conducted. **Results:** Totally, 250 DEGs and 166 DEMs in the progression of cervical cancer were screened out, and a PPI network composed of 123 nodes and 283 edges and a miRNA-target regulatory network composed of 66 nodes and 137 edges were constructed. Based on a series of analyses, the genes ATAD2, SMC4 and POLQ were not only identified as the overexpressed DEGs, but also the coding genes of hub proteins in the PPI network and regulated by the DEMs, miR-20A, miR-20B, miR-106B and miR-17-5P in the miRNA-target gene regulatory network. **Conclusion:** ATAD2, SMC4 and POLQ genes may play important roles in the progression of cervical cancer.

**Key words** cervical cancer; bioinformatics analysis; differentially expressed gene; differentially expressed miRNA; protein-protein interaction network; miRNA-target regulatory network

宫颈癌作为常见的女性生殖系统恶性肿瘤, 严重威胁着女性的健康。据统计, 每年全球有超过 50 万名女性被诊断为宫颈癌、超过 30 万名女性死于宫颈癌。宫颈癌最常见的组织学亚型为宫颈鳞状细胞癌 (cervical squamous cell carcinoma, CSC) 和宫颈腺癌, 分别占有宫颈癌的 70% 和 25%<sup>[1]</sup>。绝大部分宫颈癌的发病是由高危型人乳头瘤病毒 (human papilloma virus, HPV) 感染引起的, 其发病过程多为正常宫颈感染高危型 HPV 后, 产生宫颈上皮内瘤变

(cervical intraepithelial neoplasia, CIN), 随后 CIN 逐渐由低级别病变进展为高级别病变, 最后进一步发展为侵袭性宫颈癌。HPV 是一种非包膜病毒, 具有双链环状 DNA 基因组, 可以整合到宿主基因组中, 随后导致 E2 蛋白失活, E6 和 E7 癌蛋白表达失调, 进而引起宫颈癌变。然而, 该过程的具体分子机制仍有待深入研究, 且晚期远处转移或复发患者的生存率仍较低, 新的诊断分子标志物及治疗靶点仍有待发掘。

生物信息学分析是研究人类疾病分子变化的常用方法, 可以方便地在高通量平台上比较不同样本之间基因表达的差异, 以获得疾病相关基因信息。基因芯片 (gene microarray), 又称基因微阵列, 是目前生物信息学中常用的工具, 通过将成千上万的核

基金项目 国家自然科学基金 (81972448); 天津市科委基金 (20JCZDJC00330)

作者简介 刘玉林 (1993-), 女, 硕士在读, 研究方向: 子宫内膜癌的基础与临床研究; 通信作者: 薛凤霞, E-mail: fengxiayue1962@tmu.edu.cn。

酸探针固定在该芯片上,可一次性检测大量基因的表达情况。关于宫颈癌与正常宫颈组织之间差异表达基因的生物信息学分析已有数项研究,但尚无对正常宫颈上皮-CIN-宫颈癌这一进展过程中的差异表达基因进行综合性生物信息学分析的研究。因此,笔者在本研究中分别比较了正常宫颈上皮与 CIN 之间以及 CIN 与宫颈癌之间差异表达的基因,综合二者结果,得到在宫颈癌进展过程中的差异表达基因,并对以上基因进行一系列的生物信息学分析,以筛选出最有可能在该过程中发挥重要作用的关键基因,为寻找宫颈癌分子标志物和治疗靶点提供新的思路。

## 1 材料与方法

**1.1 基因芯片数据的选择** 在 GEO (the Gene Expression Omnibus, <http://www.ncbi.nlm.nih.gov/geo/>) 数据库中检索宫颈癌进展相关的基因芯片数据,选取了数据集 GSE63514 和 GSE86100。GSE63514 是基于 Affymetrix Human Genome U133 Plus 2.0 Array 平台的 mRNA 表达数据集,包含 24 个正常宫颈上皮标本、76 个 CIN 标本(14 个 CIN1 标本、22 个 CIN2 标本、40 个 CIN3 标本)和 28 个 CSC 标本的 mRNA 表达信息。GSE86100 是基于 Agilent-046064 Unrestricted Human miRNA V19.0 Microarray 平台的 miRNA 表达数据集,包含 6 个非 HPV 感染的正常黏膜组织样本和 6 个高危型 HPV 感染的宫颈癌组织样本的 miRNA 表达信息。

**1.2 差异表达基因(differentially expressed gene,DEG)和差异表达 miRNA(differentially expressed miRNA,DEM)的筛选** 通过 GEO 数据库中的分析软件 GEO2R(<https://www.ncbi.nlm.nih.gov/geo/geo2r/>),对数据集 GSE63514 中的正常宫颈上皮标本与 CIN 标本以及 CIN 标本与 CSC 标本之间的基因表达差异进行比较,以  $|\log FC| \geq 1$  (即差异倍数大于等于 2 倍)为界, $P < 0.05$  视为差异有统计学意义,筛选出各组间具有表达差异的基因,并将在两组结果中均为表达上调或均为表达下调的基因,视为在宫颈上皮由正常组织至癌前病变再至癌这一过程中的 DEG;对数据集 GSE86100 中的正常组织标本与宫颈癌标本之间的 miRNA 表达差异进行比较,同样以  $|\log FC| \geq 1$  和  $P < 0.05$  为界,得到宫颈癌发生过程中的 DEM。随后,利用 Morpheus 网站(<https://software.broadinstitute.org/morpheus/>)做出基因差异表达的热图,以便更为直观地显示其在各标本中的表达情况。

**1.3 DEG 的功能及通路富集分析** 通过 DAVID 数据库(The Database for Annotation, Visualization and

Integrated Discovery,<http://david.abcc.ncifcrf.gov/>),对筛选得到的 DEG 进行了 GO(Gene Ontology,<http://geneontology.org/>)功能富集分析和 KEGG(Kyoto Encyclopedia of Genes and Genomes,<http://www.genome.jp/>)通路富集分析,从而获取以上 DEG 涉及的主要生物学过程(biological processes,BP)、分子功能(molecular functions,MF)和细胞定位(cellular components,CC)以及信号通路方面的信息。 $P < 0.05$  为差异有统计学意义。

**1.4 蛋白质-蛋白质相互作用(protein-protein interaction,PPI)网络构建和亚网络分析** 利用 STRING 10.0(Search Tool for the Retrieval of Interacting Genes,<http://string-db.org/>)获得 DEG 编码蛋白之间的相互作用关系,将总得分  $> 0.4$  设置为阈值,经由 Cytoscape 3.5.0(<http://www.cytoscape.org/>)构建出 PPI 网络,以可视化形式显示蛋白间的相互作用,并用 Network Analyzer 工具对 PPI 网络行进一步分析。随后,使用插件 MCODE(Molecular Complex Detection,<http://apps.cytoscape.org/apps/mcode>)进行亚网络分析,并对亚网络中涉及的基因行进一步的功能和通路富集分析。 $P < 0.05$  为差异有统计学意义。

**1.5 miRNA-靶基因调控网络的构建** 通过使用在线工具 WebGestalt(<http://www.webgestalt.org/>)对第一步筛选得到的 DEG 进行 miRNA-靶基因分析,并使用 Cytoscape 构建出 miRNA-靶基因调控网络。

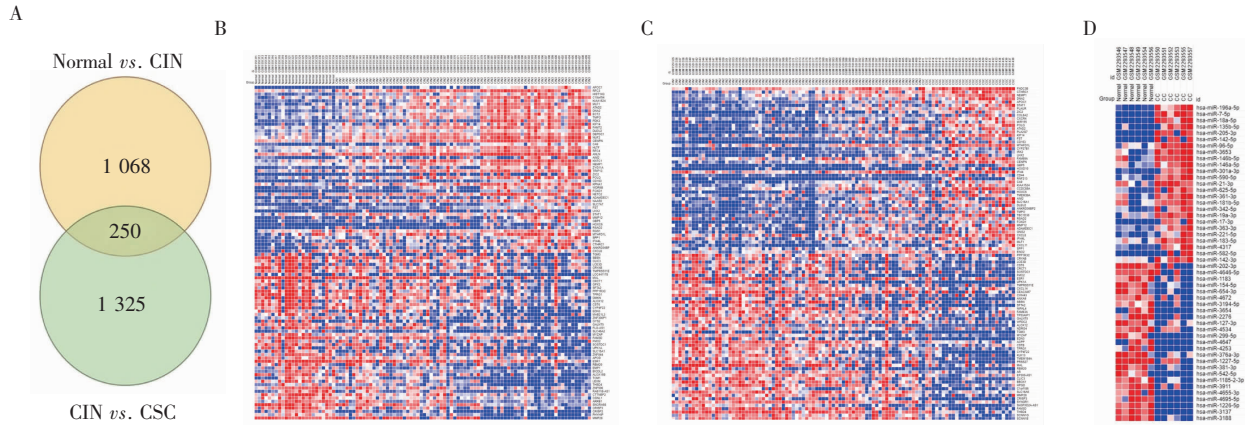
## 2 结果

**2.1 DEG 和 DEM 筛选结果** 基于数据集 GSE63514,在正常宫颈上皮样本与 CIN 样本间筛选出 1 318 个表达存在差异的基因,在 CIN 和 CSC 样本间筛选出 1 575 个表达存在差异的基因。其中 250 个基因(包括 118 个上调基因和 132 个下调基因)在两组结果中同为上调或下调(图 1A),表明该部分基因既参与了正常宫颈上皮向 CIN 转化的过程,又参与了 CIN 向宫颈癌转化的过程,故将其鉴定为本研究中宫颈癌进展过程中的 DEG,其前 100 个基因的表达热图图 1B、1C。基于数据集 GSE86100,共筛选出 166 个 DEM,包括 74 个上调的 DEM 和 92 个下调的 DEM,其前 50 个基因的表达热图见图 1D。

**2.2 GO 和 KEGG 功能和通路富集分析** 表 1 展示了针对 DEG 进行的 GO 和 KEGG 分析的主要结果。从表中可以看到,宫颈癌进展过程中 DEG 主要富集于“monocarboxylic acid transport(单羧酸运输)”、“lipoygenase pathway(脂氧合酶途径)”、“oocyte maturation(卵母细胞成熟)”、“oxidation-reduction process(氧化还原过程)”和“mitotic nuclear division

(有丝分裂核分裂)”等生物学过程;“single-stranded DNA-dependent ATPase activity (单链 DNA 依赖的 ATP 酶活性)”、“monocarboxylic acid transmembrane transporter activity (单羧酸跨膜转运体活性)”、“oxidoreductase activity(氧化还原酶活性)”、“protein homodimerization activity (蛋白同源二聚化活性)”和

“ATPase activity(ATP 酶活性)”等分子功能;以及“Arachidonic acid metabolism (花生四烯酸代谢)”、“Serotonergic synapse(血清素突触通路)”、“Chemokine signaling pathway(趋化因子信号通路)”、“Proteoglycans in cancer(肿瘤中的蛋白聚糖)”和“Mismatch repair(错配修复)”等通路。



注:A:基于芯片 GSE63514 筛选的 250 个 DEG;B:正常宫颈上皮与 CIN 标本差异表达前 100(50 个上调和 50 个下调)的 DEG 热点图,红色为上调,蓝色为下调;C:CIN 与 CSC 标本差异表达前 100(50 个上调和 50 个下调)的 DEG 热点图;D:基于芯片 GSE86100 筛选的正常宫颈上皮和宫颈鳞癌差异表达前 50(25 个上调和 25 个下调)的 DEM 热点图;DEG:差异表达基因;DEMs:差异表达 miRNA;CIN:宫颈上皮内瘤变;CSC:宫颈鳞状细胞癌

图 1 DEG 和 DEM

Fig 1 Identification of DEG and DEM

表 1 宫颈癌进展相关 DEG 的 GO 和 KEGG 富集分析结果

Tab 1 Results of GO and KEGG analyses of the DEG associated with the progression of cervical cancer

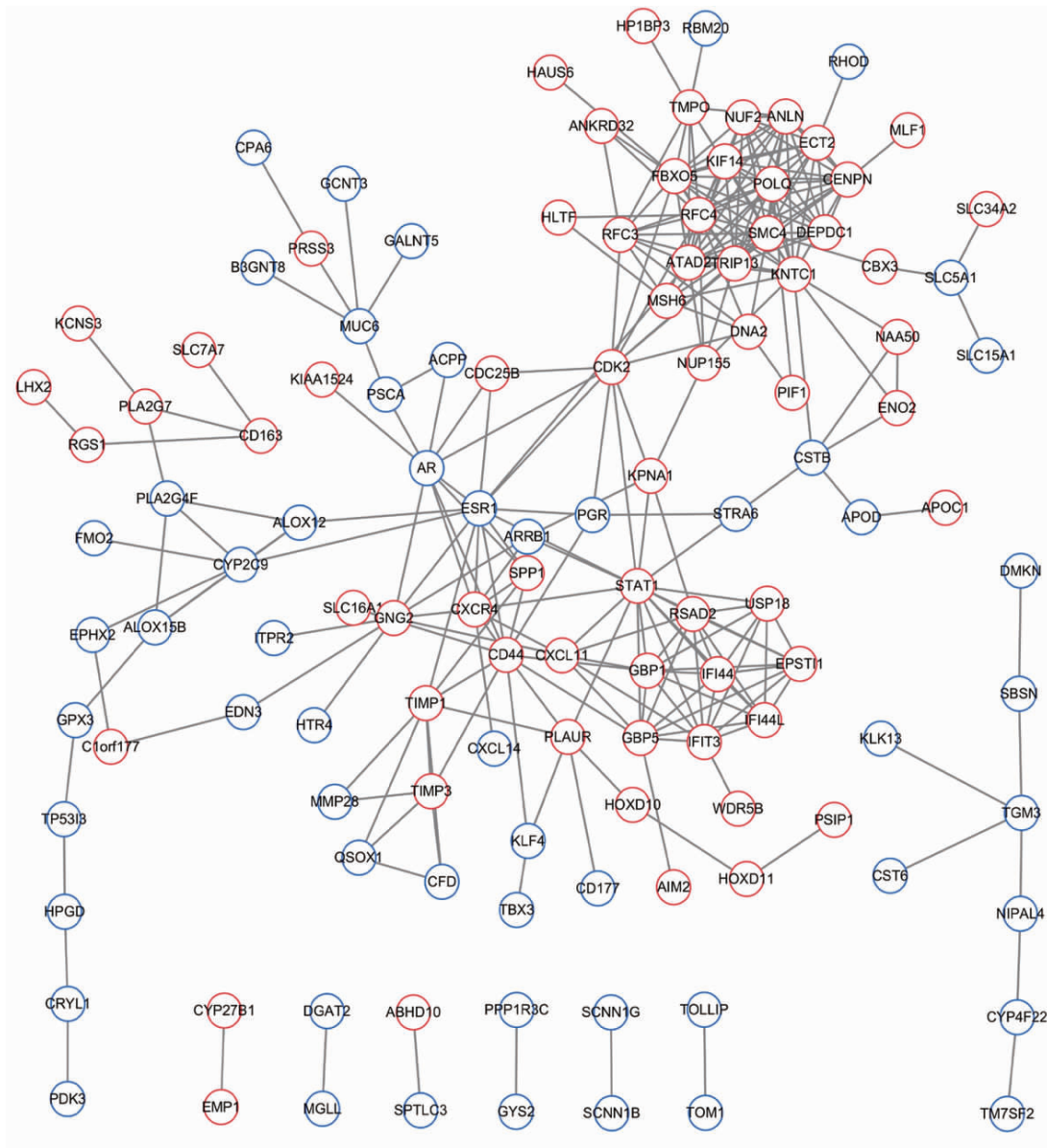
| 分类            | 分组   | 计数 | P        |
|---------------|--|----|----------|
| GOTERM_BP_FAT | GO:0015718~monocarboxylic acid transport   | 3  | 6.32E-03 |
| GOTERM_BP_FAT | GO:0019372~lipoxygenase pathway  | 3  | 1.07E-02 |
| GOTERM_BP_FAT | GO:0001556~oocyte maturation   | 3  | 2.01E-02 |
| GOTERM_BP_FAT | GO:0055114~oxidation-reduction process   | 14 | 3.11E-02 |
| GOTERM_BP_FAT | GO:0007067~mitotic nuclear division  | 8  | 3.39E-02 |
| GOTERM_MF_FAT | GO:0043142~single-stranded DNA-dependent ATPase activity   | 4  | 1.85E-04 |
| GOTERM_MF_FAT | GO:0008028~monocarboxylic acid transmembrane transporter activity  | 3  | 1.00E-02 |
| GOTERM_MF_FAT | GO:0016702~oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen | 3  | 1.16E-02 |
| GOTERM_MF_FAT | GO:0042803~protein homodimerization activity   | 17 | 1.34E-02 |
| GOTERM_MF_FAT | GO:0016887~ATPase activity   | 7  | 2.20E-02 |
| GOTERM_CC_FAT | GO:0070062~extracellular exosome   | 57 | 5.60E-05 |
| GOTERM_CC_FAT | GO:0005637~nuclear inner membrane  | 5  | 1.71E-03 |
| GOTERM_CC_FAT | GO:0005576~extracellular region  | 31 | 9.95E-03 |
| GOTERM_CC_FAT | GO:0005615~extracellular space   | 25 | 3.25E-02 |
| GOTERM_CC_FAT | GO:0005578~proteinaceous extracellular matrix  | 8  | 4.26E-02 |
| KEGG_PATHWAY  | hsa00590: Arachidonic acid metabolism  | 6  | 5.29E-04 |
| KEGG_PATHWAY  | hsa04726: Serotonergic synapse   | 7  | 1.35E-03 |
| KEGG_PATHWAY  | hsa04062: Chemokine signaling pathway  | 7  | 1.65E-02 |
| KEGG_PATHWAY  | hsa05205: Proteoglycans in cancer  | 7  | 2.27E-02 |
| KEGG_PATHWAY  | hsa03430: Mismatch repair  | 3  | 2.62E-02 |

注:GO:基因本体;KEGG:京都基因和基因组百科全书;DEG:差异表达基因



2.3 PPI 网络和亚网络构建和分析 如图 2 所示, PPI 网络由 123 个节点(node)和 283 项互相作用(edge)构成,其中 123 个节点包括 67 个上调的 DEG 和 56 个下调的 DEG 编码的蛋白;网络中与其他蛋白作用最多的前 15 个枢纽蛋白的编码基因分别为:RFC4、SMC4、STAT1、KNTC1、ATAD2、FBXO5、TRIP13、ESR1、CD44、CENPN、ANLN、CDK2、ECT2、KIF14 和 POLQ(表 2)。结合 DEG 的 GO 和 KEGG 富集分析,可以得到以上枢纽基因参与的富集功能与通路的情况:富集于“有丝分裂核分裂”的有 FBXO5、KNTC1、CDK2、ANLN 和 CENPN,富集于“单链 DNA 依赖

的 ATP 酶活性”的有 RFC4 和 POLQ,富集于“ATP 酶活性”的有 KIF14 和 ATAD2,富集于“肿瘤中的蛋白聚糖”的有 CD44 和 ESR1,富集于“蛋白同源二聚化活性”的有 ECT2 和 STAT1,富集于“卵母细胞成熟”的有 TRIP13 和 FBXO5,富集于“错配修复”的有 RFC4,富集于“趋化因子信号通路”的有 STAT1。图 3A、3C、3E 展示的是从 PPI 网络中提取的得分大于 4 的 3 个亚网络。可用看到,除 ESR1 和 AR 以外,3 个亚网络中其他所有的 DEG 均为表达上调者。图 3B、3D、3F 展示的分别是亚网络中涉及 DEG 的 GO 和 KEGG 分析结果(仅显示前 3 位)。



注:红色圆圈代表上调基因,蓝色圆圈代表下调基因,圆圈间线段代表蛋白间存在相互作用

图 2 PPI 网络

Fig 2 PPI network

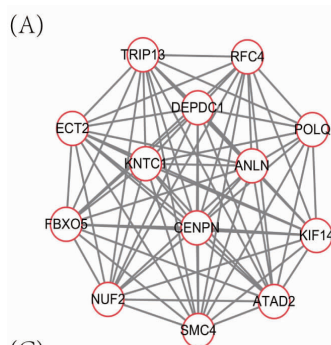
表 2 PPI 网络中的枢纽基因

Tab 2 The hub genes in the PPI network

| 枢纽基因   | 基因名  | 节点数 |
|--------|--|-----|
| RFC4   | replication factor C subunit 4                     | 19  |
| SMC4   | structural maintenance of chromosomes 4            | 18  |
| STAT1  | signal transducer and activator of transcription 1 | 17  |
| KNTC1  | kinetochore associated 1                           | 17  |
| ATAD2  | ATPase family, AAA domain containing 2             | 16  |
| FBXO5  | F-box protein 5                                    | 16  |
| TRIP13 | thyroid hormone receptor interactor 13             | 15  |
| ESR1   | estrogen receptor 1                                | 14  |
| CD44   | CD44 molecule (Indian blood group)                 | 13  |
| CENPN  | centromere protein N                               | 13  |
| ANLN   | anillin actin binding protein                      | 13  |
| CDK2   | cyclin dependent kinase 2                          | 12  |
| ECT2   | epithelial cell transforming 2                     | 12  |
| KIF14  | kinesin family member 14                           | 12  |
| POLQ   | DNA polymerase theta                               | 12  |

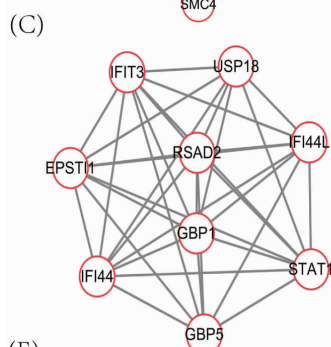
2.4 miRNA-靶基因调控网络 miRNA-靶基因调控网络如图 4 所示,由 66 个节点组成,包括 30 个 DEG 和 36 个 miRNA,节点间共有 137 个相互作用。

由图可见,该网络中的大多数的 DEG 为表达上调者,且同时为 PPI 网络中的枢纽基因;网络中约三分之一的 miRNA 为第 1 部分结果中筛选出的 DEM,且与枢纽基因存在密切联系。例如,miR-20A、miR-20B、miR-106C 和 miR-17-5P 均为 DEM,而且能靶向调控 PPI 网络中的枢纽基因 ATAD2(AT-Pase family,AAA domain containing 2,ATP 酶家族蛋白 2)、SMC4 (structural maintenance of chromosomes 4, 染色质结构维持蛋白 4) 和 POLQ (DNA polymerase theta,DNA 聚合酶  $\theta$ )。



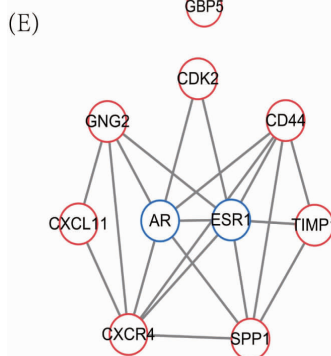
(B)

| Category      | Term                          | P-Value  | Genes                            |
|---------------|-------------------------------|----------|----------------------------------|
| GOTERM_BP_FAT | GO:0022402~cell cycle process | 1.49E-04 | CENPN, NUF2, FBXO5, ECT2, TRIP13 |
| GOTERM_BP_FAT | GO:0007049~cell cycle         | 4.41E-04 | CENPN, NUF2, FBXO5, ECT2, TRIP13 |
| GOTERM_BP_FAT | GO:0000280~nuclear division   | 5.62E-04 | CENPN, NUF2, FBXO5, TRIP13       |



(D)

| Category      | Term  | P-Value  | Genes                     |
|---------------|---|----------|---------------------------|
| GOTERM_BP_FAT | GO:0016032~viral process                              | 4.38E-04 | IFIT3, RSAD2, STAT1, GBP1 |
| GOTERM_BP_FAT | GO:0044764~multi-organism cellular process            | 4.48E-04 | IFIT3, RSAD2, STAT1, GBP1 |
| GOTERM_BP_FAT | GO:0044419~interspecies interaction between organisms | 5.39E-04 | IFIT3, RSAD2, STAT1, GBP1 |
| KEGG_PATHWAY  | pon05164:Influenza A                                  | 2.52E-02 | RSAD2, STAT1              |



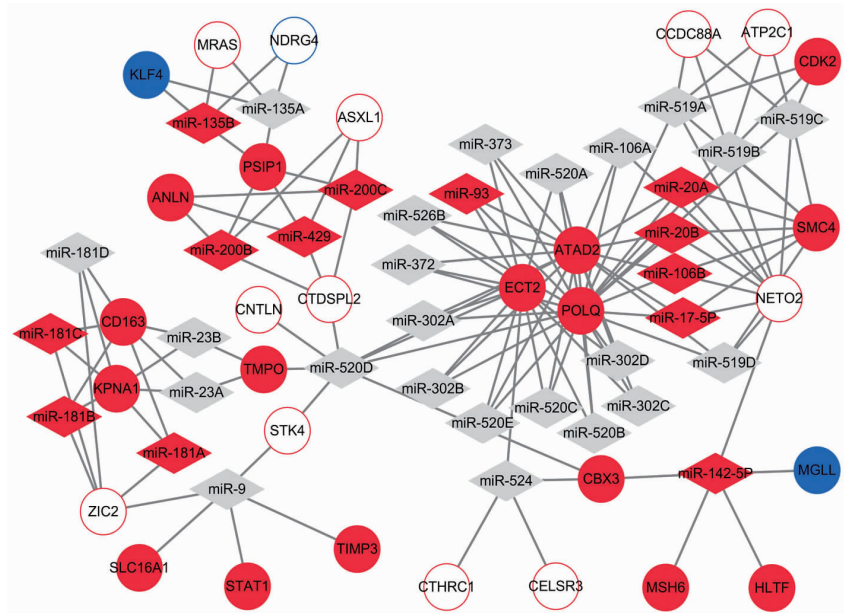
(F)

| Category      | Term  | P-Value  | Genes                                |
|---------------|---|----------|--------------------------------------|
| GOTERM_BP_FAT | GO:0010033~response to organic substance      | 9.68E-04 | AR, CXCR4, ESR1, CXCL11, TIMP1, SPP1 |
| GOTERM_BP_FAT | GO:0033993~response to lipid                  | 1.41E-03 | AR, ESR1, CXCL11, SPP1               |
| GOTERM_BP_FAT | GO:2000241~regulation of reproductive process | 2.06E-03 | AR, CDK2, TIMP1                      |
| KEGG_PATHWAY  | ssc05200:Pathways in cancer                   | 7.53E-03 | AR, CXCR4, GNG2, CDK2                |
| KEGG_PATHWAY  | ssc04062:Chemokine signaling pathway          | 1.53E-02 | CXCR4, GNG2, CXCL11                  |
| KEGG_PATHWAY  | ssc04151:PI3K-Akt signaling pathway           | 4.99E-02 | GNG2, CDK2, SPP1                     |

注:A、C、E:红色圆圈代表上调基因,蓝色圆圈代表下调基因,圆圈间线段代表蛋白质间存在互相作用;B、D、F:PPI 亚网络中枢纽基因的 GO 和 KEGG 富集分析;GO:基因本体;KEGG:京都基因和基因组百科全书

图 3 PPI 亚网络及 GO 和 KEGG 富集分析

Fig 3 Significant modules of the PPI network and GO and KEGG enrichment analyses



注:红色和蓝色圆形节点分别表示上调和下调的 DEG,其中实心节点表示为 PPI 网络中的 DEG,空心节点表示未出现在 PPI 网络中的 DEG;红色菱形节点表示上调的 DEM,灰色菱形节点表示该网络中非 DEM 的其余 miRNA;DEG:差异表达基因;PPI:蛋白质-蛋白质相互作用网络;DEM:差异表达 miRNA

图 4 miRNA-靶基因调控网络

Fig 4 miRNA-target regulatory network

### 3 讨论

宫颈癌是发展中国家最常见的妇科恶性肿瘤,是导致女性因癌症死亡的第二大原因<sup>[1]</sup>,其癌前病变为 CIN,最常见的亚型为 CSC。为了探索宫颈癌进展的分子机制以及筛选潜在的分子标志物及治疗靶点,基于数据集 GSE63514 和 GSE86100,对正常宫颈上皮样本、CIN 样本和宫颈癌样本进行了生物信息学分析。在本研究中,共筛选出宫颈癌发生、发展相关的 118 个上调 DEG 和 132 个下调 DEG 以及 74 个上调 DEM 和 92 个下调 DEM,并通过生物信息学分析的方法,对其进行了功能和通路富集分析及 PPI 网络和亚网络以及 miRNA-靶基因调控网络的构建和分析,从而进一步筛选出可能在宫颈癌进展过程中起关键作用的 DEG 和 DEM。

本研究发现 miRNA-靶基因调控网络中的部分 miRNA 为本研究筛选得到的 DEM,部分基因为 PPI 网络中枢纽蛋白的编码基因。最值得注意的是 DEM miR-20A、miR-20B、miR-106B 和 miR-17-5P,以及它们共同调控的 ATAD2、POLQ 和 SMC4,以上 3 个 DEG 不仅为 PPI 网络中的枢纽基因和第 1 个亚网络的构成节点,而且三者彼此之间均存在互相作用。因此,受 miR-20A、miR-20B、miR-106B 和 miR-17-5P 调控的基因 ATAD2、SMC4 和 POLQ 极有可能在宫颈癌的进展中发挥着重要的作用。

ATAD2 编码的蛋白为 ATP 酶家族蛋白 2,能够干预基因组的调控,包括细胞增殖、分化和凋亡。在

人类多种癌症中均检测到 ATAD2 的过度表达,包括口腔鳞状细胞癌、食道鳞状细胞癌、卵巢癌、乳腺癌、胃癌以及结直肠癌等;而且 ATAD2 蛋白还被认为是肿瘤的“增殖驱动者”、“转录共调节因子”和“凋亡负性调节因子”,并已通过基因敲低研究证实其为一种相当有前景的抗癌药物。与目前研究结果相吻合,在一项针对宫颈癌 circRNA-miRNA-mRNA 网络的研究中<sup>[2]</sup>,ATAD2 也被确定为枢纽基因;另外,经临床资料分析及体外实验证实,ATAD2 的确在促进宫颈癌细胞增殖、迁移和侵袭中发挥重要作用<sup>[3]</sup>。

SMC4 是编码染色质结构维持蛋白 4 的基因,该编码蛋白是参与染色体浓缩和分离的复合物的核心亚基。研究发现,SMC4 蛋白的有丝分裂特异性磷酸化可以降低浓缩素在染色体上的动态转换率<sup>[4]</sup>,且与 ATAD2 相似,SMC4 蛋白也含有 ATPase 结构域,其失活可导致细胞的死亡<sup>[5]</sup>。关于 SMC4 在乳腺癌、胰腺导管癌、肝细胞癌、肺腺癌和神经胶质瘤等癌症中过表达的报道有很多,但迄今为止尚无其在宫颈癌中表达情况的报道。有关 SMC4 在宫颈癌中的唯一研究是 SMC4 蛋白可以负性调控主基因组调控蛋白 CTCF(CCCTC-binding factor,CCCTC 结合因子),后者能够介导核糖体 RNA 基因的转录<sup>[6]</sup>。但在其他类型的癌症中,SMC4 已被证明在肿瘤的发生和发展中起着关键作用,例如,敲低 SMC4 将导致前列腺癌细胞系 DU145 和 LNCaP-AR 的迁移和侵袭显著减少<sup>[7]</sup>。



POLQ 是蛋白质 DNA 聚合酶  $\theta$  的编码基因,像 ATAD2 和 SMC4 一样,DNA 聚合酶  $\theta$  也具有 ATPase 活性。其作用主要在于影响 DNA 复制和修复由各种因素(例如核酸内切酶和 Cas9 切口酶)引起的 DNA 双链断裂<sup>[8]</sup>。据研究报道,POLQ 在包括乳腺癌、大肠癌、肺癌在内的多种癌症中表达上调,并与患者的不良预后相关。另有多项研究表明,POLQ 缺陷的人类肿瘤细胞系对放射线和其他双链断裂剂敏感,意味着 POLQ 极有可能被用作治疗癌症的药物靶标。然而在宫颈癌中,本研究是首次发现 POLQ 作为潜在的促癌基因的作用。

此外,由 miRNA-靶基因调控网络的结果可见,ATAD2、SMC4 和 POLQ 三者还同时与 miR-20A、miR-20B、miR-106B 和 miR-17-5P 存在关联,而以上 miRNA 又均受 HPV 蛋白 E6/E7 的调控<sup>[9-11]</sup>,并在宫颈癌中过表达<sup>[12-14]</sup>。据报道,miR-20A 与宫颈癌细胞的增殖、迁移、侵袭、转移和顺铂耐药有关,可以抑制 IFN- $\gamma$  和 TNF- $\alpha$  的分泌,减弱 NK 细胞对宫颈癌细胞的杀伤作用<sup>[15-19]</sup>。有关 miR-106B 的研究表明,其过度表达可显著促进宫颈癌细胞系 HeLa 和 SiHa 的迁移<sup>[12]</sup>。miR-17-5P 也被报道可以增强宫颈癌的增殖和转移<sup>[20]</sup>。

综上所述,ATAD2、SMC4 和 POLQ 都是具有 ATPase 活性的基因,并且都与细胞的基因组调控有关,其失调将导致肿瘤的发生和发展;而且三者作为宫颈癌进展过程中的 DEG,既是 PPI 网络中互相关联的枢纽基因,又同时受 DEM——miR-20A、miR-20B、miR-106B 和 miR-17-5P 的调控。这是首次通过生物信息学分析正常宫颈上皮-CIN-宫颈癌进展过程中的差异表达基因的研究,并通过一系列综合分析得出,miR-20A、miR-20B、miR-106B 和 miR-17-5P 共同调控下的基因 ATAD2、SMC4 和 POLQ,在宫颈癌的发生、发展中很可能发挥着至关重要的作用,并有可能用作宫颈癌的临床分子标志物及治疗靶点。但由于本研究的样本量尤其是 miRNA 样本量相对较小,所以可能具有一定的局限性,且部分研究结果尚需进一步的临床及基础实验证实。

#### 参考文献:

- [1] COHEN P A, JHINGRAN A, OAKNIN A, et al. Cervical cancer[J]. *Lancet*, 2019, 393(10167): 169-182.
- [2] YI Y, LIU Y, WU W, et al. Reconstruction and analysis of circRNA miRNA mRNA network in the pathology of cervical cancer[J]. *Oncol Rep*, 2019, 41(4): 2209-2225.
- [3] ZHENG L, LI T, ZHANG Y, et al. Oncogene ATAD2 promotes cell proliferation, invasion and migration in cervical cancer[J]. *Oncol Rep*, 2015, 33(5): 2337-2344.
- [4] THADANI R, KAMENZ J, HEEGER S, et al. Cell-cycle regulation of dynamic chromosome association of the condensin complex[J]. *Cell Rep*, 2018, 23(8): 2308-2317.
- [5] PALOU R, DHANARAMAN T, MARRAKCHI R, et al. Condensin ATPase motifs contribute differentially to the maintenance of chromosome morphology and genome stability[J]. *PLoS Biol*, 2018, 16(6): e2003980.
- [6] HUANG K, JIA J, WU C, et al. Ribosomal RNA gene transcription mediated by the master genome regulator protein CCCTC-binding factor (CTCF) is negatively regulated by the condensin complex[J]. *J Biol Chem*, 2013, 288(36): 26067-26077.
- [7] ZHAO SG, EVANS JR, KOTHARI V, et al. The landscape of prognostic outlier genes in high-risk prostate cancer[J]. *Clin Cancer Res*, 2016, 22(7): 1777-1786.
- [8] WANG Z, SONG Y, LI S, et al. DNA polymerase theta (POLQ) is important for repair of DNA double-strand breaks caused by fork collapse[J]. *J Biol Chem*, 2019, 294(11): 3909-3919.
- [9] HONEGGER A, SCHILLING D, BASTIAN S, et al. Dependence of intracellular and exosomal microRNAs on viral E6/E7 oncogene expression in HPV-positive tumor cells[J]. *PLoS Pathog*, 2015, 1(3): e1004712.
- [10] LIU X. Up-regulation of miR-20a by HPV16 E6 exerts growth-promoting effects by targeting PDCD6 in cervical carcinoma cells[J]. *Biomed Pharmacother*, 2018, 102: 996-1002.
- [11] ZHENG Z M, WANG X. Regulation of cellular miRNA expression by human papillomaviruses[J]. *Biochim Biophys Acta*, 2011, 1809(11/12): 668-677.
- [12] CHENG Y, GUO Y, ZHANG Y, et al. MicroRNA-106b is involved in transforming growth factor beta1-induced cell migration by targeting disabled homolog 2 in cervical carcinoma[J]. *J Exp Clin Cancer Res*, 2016, 35: 11.
- [13] LI M Y, HU X X. Meta-analysis of microRNA expression profiling studies in human cervical cancer[J]. *Med Oncol*, 2015, 32(6): 510.
- [14] MA D, ZHANG Y Y, GUO Y L, et al. Profiling of microRNA-mRNA reveals roles of microRNAs in cervical cancer [J]. *Chin Med J(Engl)*, 2012, 125(23): 4270-4276.
- [15] KANG H W, WANG F, WEI Q, et al. miR-20a promotes migration and invasion by regulating TNKS2 in human cervical cancer cells[J]. *FEBS Lett*, 2012, 586(6): 897-904.
- [16] XIONG Y, SUN F, DONG P, et al. iASPP induces EMT and cisplatin resistance in human cervical cancer through miR-20a-FBXL5/BTG3 signaling[J]. *J Exp Clin Cancer Res*, 2017, 36(1): 48.
- [17] ZHAO S, YAO D, CHEN J, et al. MiR-20a promotes cervical cancer proliferation and metastasis *in vitro* and *in vivo*[J]. *PLoS One*, 2015, 10(3): e0120905.
- [18] ZHOU Q, DONG J, LUO R, et al. MicroRNA-20a regulates cell proliferation, apoptosis and autophagy by targeting thrombospondin 2 in cervical cancer[J]. *Eur J Pharmacol*, 2019, 844: 102-109.
- [19] ZHU S Y, WU Q Y, ZHANG C X, et al. miR-20a inhibits the killing effect of natural killer cells to cervical cancer cells by downregulating RUNX1[J]. *Biochem Biophys Res Commun*, 2018, 505(1): 309-316.
- [20] CAI N, HU L, XIE Y, et al. MiR-17-5p promotes cervical cancer cell proliferation and metastasis by targeting transforming growth factor-beta receptor 2[J]. *Eur Rev Med Pharmacol Sci*, 2018, 22(7): 1899-1906.