Cancer Management and Research

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Identification of key candidate genes and small molecule drugs in cervical cancer by bioinformatics strategy

This article was published in the following Dove Press journal:
Cancer Management and Research

Xin Tang[1]
Yicong Xu[2,3]
Lin Lu[2,3]
Yang Jiao[2,3]
Jianjun Liu[2,3]
Linlin Wang[2,3]
Hongbo Zhao[2,3]

[1]School of Rehabilitation, Kunming Medical University, Kunming, China; [2]Institute of Molecular and Clinical Medicine, Kunming Medical University, Kunming, China; [3]Yunnan Key Laboratory of Stem Cell and Regenerative Medicine, Kunming, China

Purpose: Cervical cancer (CC) is one of the most common malignant tumors among women. The present study aimed at integrating two expression profile datasets to identify critical genes and potential drugs in CC.

Materials and methods: Expression profiles, GSE7803 and GSE9750, were integrated using bioinformatics methods, including differentially expressed genes analysis, Kyoto Encyclopedia of Genes and Genomes pathway analysis, and protein–protein interaction (PPI) network construction. Subsequently, survival analysis was performed among the key genes using Gene Expression Profiling Interactive Analysis websites. Connectivity Map (CMap) was used to query potential drugs for CC.

Results: A total of 145 upregulated genes and 135 downregulated genes in CC were identified. The functional changes of these differentially expressed genes related to CC were mainly associated with cell cycle, DNA replication, p53 signaling pathway, and oocyte meiosis. A PPI network was identified by STRING with 220 nodes and 2,111 edges. Thirteen key genes were identified as the intersecting genes of the enrichment pathways and the top 20 nodes in PPI network. Survival analysis revealed that high mRNA expression of *MCM2*, *PCNA*, and *RFC4* was significantly associated with longer overall survival, and the survival was significantly better in the low-expression *RRM2* group. Moreover, CMap predicted nine small molecules as possible adjuvant drugs to treat CC.

Conclusion: Our study found key dysregulated genes involved in CC and potential drugs to combat it, which might provide insights into CC pathogenesis and might shed light on potential CC treatments.

Keywords: cervical cancer, bioinformatics, cell cycle, biomarker, drug

Introduction

Cervical cancer (CC) is the second most common malignant tumor among women, responsible for ~527,600 new cases and >265,700 deaths annually.[1] Despite advances in screening detection and new treatment strategies, CC is one of the leading causes of cancer death among females in many developing countries.[2,3] Although most patients can be cured if diagnosed at an early stage, poor prognosis is observed with secondary metastatic cancer and tumor relapse.

Although human papillomavirus (HPV) is a prerequisite for CC, only a small number of women infected by this virus develop cancer. Thus, other risk factors should be considered as cofactors contributing to the progression of CC.[4] Dysregulated genes play important roles in CC development.[5] Several studies have used gene expression profiling to identify key genes between CC samples and normal cervix.[6–9] Hundreds

Correspondence: Hongbo Zhao
Institute of Molecular and Clinical Medicine, Kunming Medical University, No. 1168, West Chunrong Road, Chenggong District, Kunming 650500, China
Tel/fax +86 871 6592 2699
Email zhao.hongbo@hotmail.com

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/CMAR.S171661

of differentially expressed genes (DEGs) were detected. However, DEGs reported in different studies vary enormously with only some of them consistently detected. Therefore, the discovery of novel effective therapeutic targets against CC is urgently required.

A number of chemotherapeutic agents have shown activity against CC, including cisplatin,[10] bevacizumab,[11] carboplatin,[12] paclitaxel,[13] ifosfamide,[14] and topotecan.[15] Various combinations of these agents are recommended as therapies.[16] A recent systematic literature review found that carboplatin–paclitaxel is equally effective and less toxic than cisplatin–paclitaxel as the first-line therapy for metastatic CC.[17] However, patients overall survival (OS) times remains short, indicating an urgent need to discover some molecular drugs that are more efficient and selective. Based on bioinformatics approaches, several studies found small molecules as potential anticancer agents.[18–20]

In this study, we selected the following microarray datasets GSE7803 and GSE9750 from the Gene Expression Omnibus (GEO) database to identify DEGs. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis using the identified DEGs was investigated. A protein–protein interaction (PPI) network was constructed to elucidate the significant relationships among DEGs and to identify key genes. Furthermore, the Kaplan–Meier estimator was used on the Gene Expression Profiling Interactive Analysis (GEPIA) website. Candidate small molecules were identified for their potential use in the treatment of CC.

Materials and methods

Data collection

Two CC microarray datasets were downloaded from the GEO website (http://www.ncbi.nlm.nih.gov/geo/). GSE7803 microarray data contained 21 CC tissues and 10 normal cervical epithelia tissues.[6] GSE9750 included 33 tumors samples and 24 healthy cervical samples.[7] Both the profile datasets were based on the Affymetrix GPL96 platform (Affymetrix Human Genome U133A Array). Because Connectivity Map (CMap) strictly required all probesets obtained from the Affymetrix Human Genome U133A Array,[21] we predicted the drugs for the DEGs measured only in this platform with high accuracy. GSE63514 data included 28 cancer cases and 24 normal cases[8] and were chosen to validate *RRM2* mRNA expression in our analysis.

Data preprocessing and DEGs screening

The raw data were standardized and transformed into expression values using the affy package of Bioconductor (http://www.bioconductor.org/).[22] DEGs between cancer and normal samples were selected by significance analysis using the empirical Bayes methods within limma package.[23] False discovery rate (FDR) <0.05 and |log2 (fold change)| >1 were set as the cutoff criteria for the identification of DEGs. Common dysregulated probesets between GSE7803 and GSE9750 were selected for subsequent analyzes.

KEGG pathway analysis

Pathway enrichment analysis was performed using the clusterProfiler package and a pathway with an adjusted *P*-value <0.05 was considered significantly enriched.[24] DEGs that we identified could be involved in multiple pathways, Thus, some overlap was observed among the pathways. We identified the significant pathways that shared the same DEGs and used Cytoscape (version 3.5.1) to construct graphical representations of the interactive relationships among the pathways.[25]

PPI network construction and analysis

The PPI pairs of the screened DEGs were analyzed using the online database STRING version 10.5 (https://string-db.org/).[26] The pairs with combined scores >0.4 were used for the PPI network construction, then the Cytoscape software was used to construct the network and analyze the interaction relationship of the candidate DEGs encoding proteins in CC.

Validation of key genes

Key genes were identified as the intersecting genes of the enrichment pathways and top 20 nodes in PPI network. To confirm the reliability of these genes from our detection, we analyzed their prognostic and expression in CC using GEPIA.[27] GEPIA is an interactive web application for gene expression analysis based on 9,736 tumors and 8,587 normal samples from the Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression databases.[28,29] We evaluated the expression of key genes in CC tissues and normal tissues. Then the survival curve and boxplot were performed to visualize the relationships.

Identification of candidate small molecules

The CC gene signature was used to query CMap to find potential drugs for use in patients.[21] CMap is an in silico method to predict potential drugs that could possibly reverse, or induce, the biological state encoded in particular gene expression signatures. The common differently expressed probesets in GSE7803 and GSE9750 between CC samples and healthy controls were divided into upregulated and

downregulated groups. Then, these probesets were used to query the CMap database. Finally, the enrichment score representing similarity was calculated, ranging from –1 to 1. A positive connectivity score indicates that a drug is able to induce the input signature in human cell lines. Conversely, a negative connectivity score indicates that a drug is able to reverse the input signature. Negative connectivity scores were investigated, which indicate potential therapeutic value. After rank ordering all instances, the connectivity score of various instances were filter by the number of instances (N>10) and *P*-value (<0.05).

Results

DEGs identification

The two mRNA expression profiles, including 54 patients with CC and 34 healthy individuals, were included in our study. Using a FDR <0.05 and |logFC| >1 as cutoff criteria, we extracted 443 and 848 differentially expressed probesets from the expression profile datasets GSE7803 and GSE9750, respectively. In GSE7803, 212 unregulated probes and 231 downregulated probes were identified. A total of 376 unregulated probes and 472 downregulated probes were identified in GSE9750. After being overlapped, the common 149 upregulated and 146 downregulated probesets corresponding to 145 upregulated and 135 downregulated genes were identified from the two profile datasets (Table S1).

CC significant pathways evaluation

A total of 16 pathways with adjusted *P*-value <0.05 were found enriched including 10 upregulated and 6 downregulated pathways (Table 1). The most significant upregulated pathway was cell cycle; the other significant pathways included DNA replication, oocyte meiosis, p53 signaling pathway, microRNAs in cancer, and cellular senescence. The downregulated pathways included arachidonic acid metabolism, serotonergic synapse, gap junction, estrogen signaling pathway, signaling pathways regulating pluripotency of stem cells, and proteoglycans in cancer (Figure 1A). In order to consider the potentially biological complexities in which a gene may belong to multiple pathways and provide information of numeric changes, we constructed pathway–gene networks to extract the complex association (Figure 1B, C). Cell cycle pathway contained the most significant genes in the network.

Table 1 Pathway enrichment analysis of DEGs function in CC

ID	Description	Adjusted *P*-value	Count	Gene symbol
Upregulated				
hsa04110	Cell cycle	2.02E–19	21	*BUB1B*, *CCNB1*, *CCNB2*, *CCNE2*, *CDC7*, *CDK1*, *CDKN2A*, *CDKN2C*, *E2F3*, *MAD2L1*, *MCM2*, *MCM3*, *MCM4*, *MCM5*, *MCM6*, *MCM7*, *ORC6*, *PCNA*, *PTTG1*, *SMC1A*, *TTK*
hsa03030	DNA replication	3.59E–11	10	*FEN1*, *MCM2*, *MCM3*, *MCM4*, *MCM5*, *MCM6*, *MCM7*, *PCNA*, *RFC4*, *RFC5*
hsa04114	Oocyte meiosis	7.63E–04	8	*AURKA*, *CCNB1*, *CCNB2*, *CCNE2*, *CDK1*, *MAD2L1*, *PTTG1*, *SMC1A*
hsa04115	p53 signaling pathway	1.16E–03	6	*CCNB1*, *CCNB2*, *CCNE2*, *CDK1*, *CDKN2A*, *RRM2*
hsa05206	MicroRNAs in cancer	1.07E–02	10	*CCNE2*, *CDKN2A*, *DDIT4*, *DNMT1*, *E2F3*, *EZH2*, *MIR106B*, *MIR25*, *PLAU*, *STMN1*
hsa04218	Cellular senescence	1.44E–02	7	*CCNB1*, *CCNB2*, *CCNE2*, *CDK1*, *CDKN2A*, *CXCL8*, *E2F3*
hsa03430	Mismatch repair	2.09E–02	3	*PCNA*, *RFC4*, *RFC5*
hsa04914	Progesterone-mediated oocyte maturation	3.43E–02	5	*AURKA*, *CCNB1*, *CCNB2*, *CDK1*, *MAD2L1*
hsa05166	HTLV-I infection	3.43E–02	8	*BUB1B*, *CCNB2*, *CDKN2A*, *CDKN2C*, *E2F3*, *MAD2L1*, *PCNA*, *PTTG1*
hsa03410	Base excision repair	4.22E–02	3	*FEN1*, *MBD4*, *PCNA*
Downregulated				
hsa00590	Arachidonic acid metabolism	2.61E–02	5	*ALOX12*, *ALOX12B*, *ALOX15B*, *GPX3*, *PTGDS*
hsa04726	Serotonergic synapse	2.79E–02	6	*ALOX12*, *ALOX12B*, *ALOX15B*, *CYP2C18*, *DUSP1*, *ITPR2*
hsa04540	Gap junction	3.37E–02	5	*GJA1*, *ITPR2*, *PDGFD*, *TUBA1A*, *TUBB2A*
hsa04915	Estrogen signaling pathway	3.37E–02	6	*CALML3*, *ESR1*, *FOS*, *ITPR2*, *KRT10*, *KRT13*
hsa04550	Signaling pathways regulating pluripotency of stem cells	3.37E–02	6	*FGFR2*, *FZD1*, *ID4*, *IGF1*, *ISL1*, *KLF4*
hsa05205	Proteoglycans in cancer	3.70E–02	7	*CCND1*, *DCN*, *ESR1*, *FZD1*, *IGF1*, *ITPR2*, *PDCD4*

Abbreviations: CC, cervical cancer; DEGs, differentially expressed genes; HTLV-I, human T-lymphotropic virus type 1.

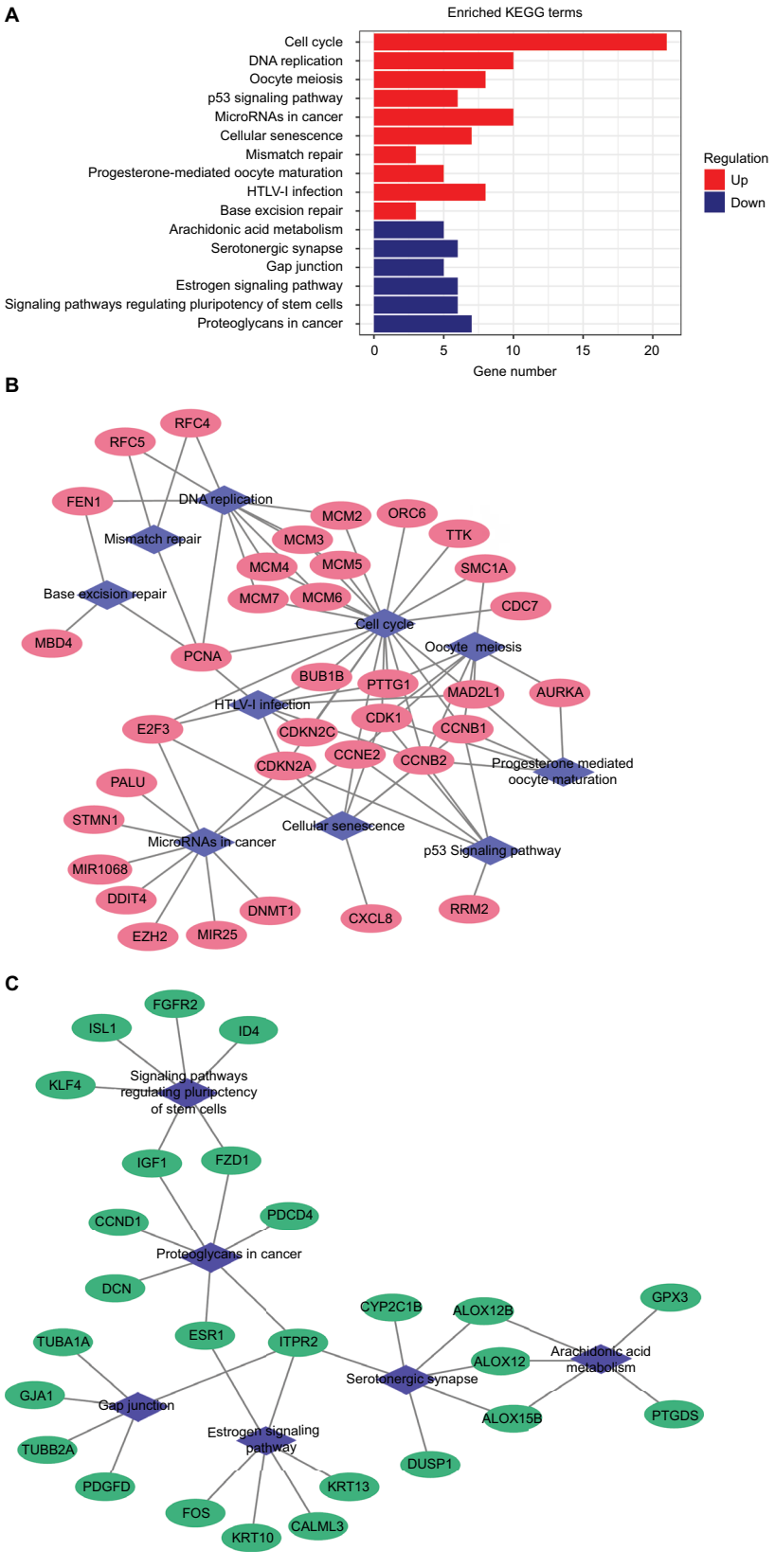

Figure 1 Significantly enriched pathway terms associated to DEGs in CC.
Notes: (**A**) KEGG pathways in CC DEGs enrichment analysis. (**B**) Upregulated pathway–gene network including 35 upregulated genes and 10 pathways. (**C**) Downregulated pathway–gene network including 26 downregulated genes and 6 pathways.
Abbreviations: CC, cervical cancer; DEGs, differentially expressed genes; KEGG, Kyoto Encyclopedia of Genes and Genomes; HTLV-I, human T-lymphotropic virus type 1.

PPI network construction

STRING was used for mining proteins expressed by DEGs which can interact with others. At a combined score >0.4, a total of 222 DEGs (118 upregulated and 104 downregulated genes) among the 280 commonly altered DEGs were filtered into the DEGs PPI network, containing 222 nodes and 2,111 edges (Figure 2A). NetworkAnalyzer app in Cytoscape was used to calculate the node degree.[25] The genes CDK1, PCNA, TOP2A, CCNB1, RFC4, MAD2L1, NDC80, CCNB2, AURKA, TYMS, MCM2, FEN1, RRM2, NCAPG, TTK, PRC1, MCM4, ZWINT, DTL, and MCM6 were the most significant 20 node degree genes and were selected as the hub nodes, since they might play important roles in CC progression (Figure 2B).

Key gene signatures identification in CC

Compared with KEGG enrichment genes, 13 of the top 20 nodes in the PPI network, including AURKA, CCNB1, CCNB2, CDK1, FEN1, MAD2L1, MCM2, MCM4, MCM6, PCNA, RFC4, RRM2, and TTK were found as key genes. Further survival analyses on these key genes were employed to evaluate their effects on CC patients' survival using GEPIA. Expression levels of *MCM2*, *PCNA*, *RFC4*, and *RRM2* were significantly related to the OS of patients with cervical squamous cancer ($P<0.05$). High expression of *MCM2*, *PCNA*, and *RFC4* could result in a high survival rate, and increased *RRM2* expression in CC was significantly associated with shorter patients' survival (Figure 3A–D). The expression of these four genes was significantly higher in CC tissues compared to that of normal tissues ($P<0.01$; Figure 3E–H). Together, the high level of these four genes might represent the important prognostic factor to predict the survival of CC. GSE63514 was used to validate *RRM2* mRNA expression. The results showed that *RRM2* expression was significantly higher in CC compared to that of normal tissues ($P<0.01$; Figure 4A).The PPI network based on RRM2 found that PCNA and RFC4 have a close relationship with RRM2, and most of the proteins in the network were related to cell cycle (Figure 4B).

Related small molecule drugs screening

In order to screen out small molecule drugs, consistent differently expressed probesets between CC samples and healthy controls were analyzed with CMap. The related small molecules with highly significant correlations are listed in Table 2. Among these molecules, trichostatin A (TSA), tanespimycin, vorinostat, trifluoperazine, prochlorperazine, and thioridazine showed higher negative correlation and the potential to treat CC.

Discussion

Driver genes play vital roles during stages of cancer progression. Although many studies on CC development are available, more efforts are needed to identify driver genes and candidate drugs that may shed light on CC treatments. This study integrated two gene profile datasets based on Affymetrix Human Genome U133A Array, utilized bioinformatics methods to analyze these datasets, and identified 280 commonly changed DEGs (145 upregulated and 135 downregulated). Pathway enrichment analysis indicated that cell cycle, DNA replication, oocyte meiosis, p53 signaling pathway, cellular senescence, and DNA repair-relevant biological pathways were overrepresented among the upregulated genes. The PPI network was constructed including 222 nodes/DEGs and 2,111 edges. Thirteen key genes were identified and chosen for survival analysis. *MCM2*, *PCNA*, *RFC4*, and *RRM2* were clearly related to the prognosis of patients. In addition, small molecules that can provide new insights in CC therapeutic studies were identified.

Many researchers have found that four key genes were involved in cell cycle, participating in tumorigenesis and tumor proliferation. MCM2 has been studied in a wide range of human malignancies and is associated with tumor histopathological grade in several malignancies, including colon, oral cavity, ovarian, urothelial, and non-small cell lung carcinoma.[30–34] In cervical carcinoma and precancerous lesions, MCM2 is overexpressed and positively correlated with high risk types of HPV.[35] Amaro Filho et al also reported an increasing expression of MCM2 in invasive CC compared to control, but they suggested that MCM2 is not a good biomarker when comparing the different clinical stages of CC.[36] PCNA acts as a central coordinator of DNA transactions by providing a multivalent interaction surface for factors involved in DNA replication and cell cycle regulation. Owing to its function, PCNA has been widely used as a tumor marker for cancer cell progression and patient prognosis.[37–39] A recent systematic literature review found that the expression of PCNA is significantly associated with poor 5-year survival, International Federation of Gynecology and Obstetrics stage, or WHO grade, suggesting its use as a valuable prognostic and diagnostic biomarker in CC and gliomas.[40] RFC4 is involved in cancer. Knockdown of *RFC4* in HepG2 cells induces apoptosis.[41] Similar results were discovered in breast carcinoma.[42] In colorectal cancer,

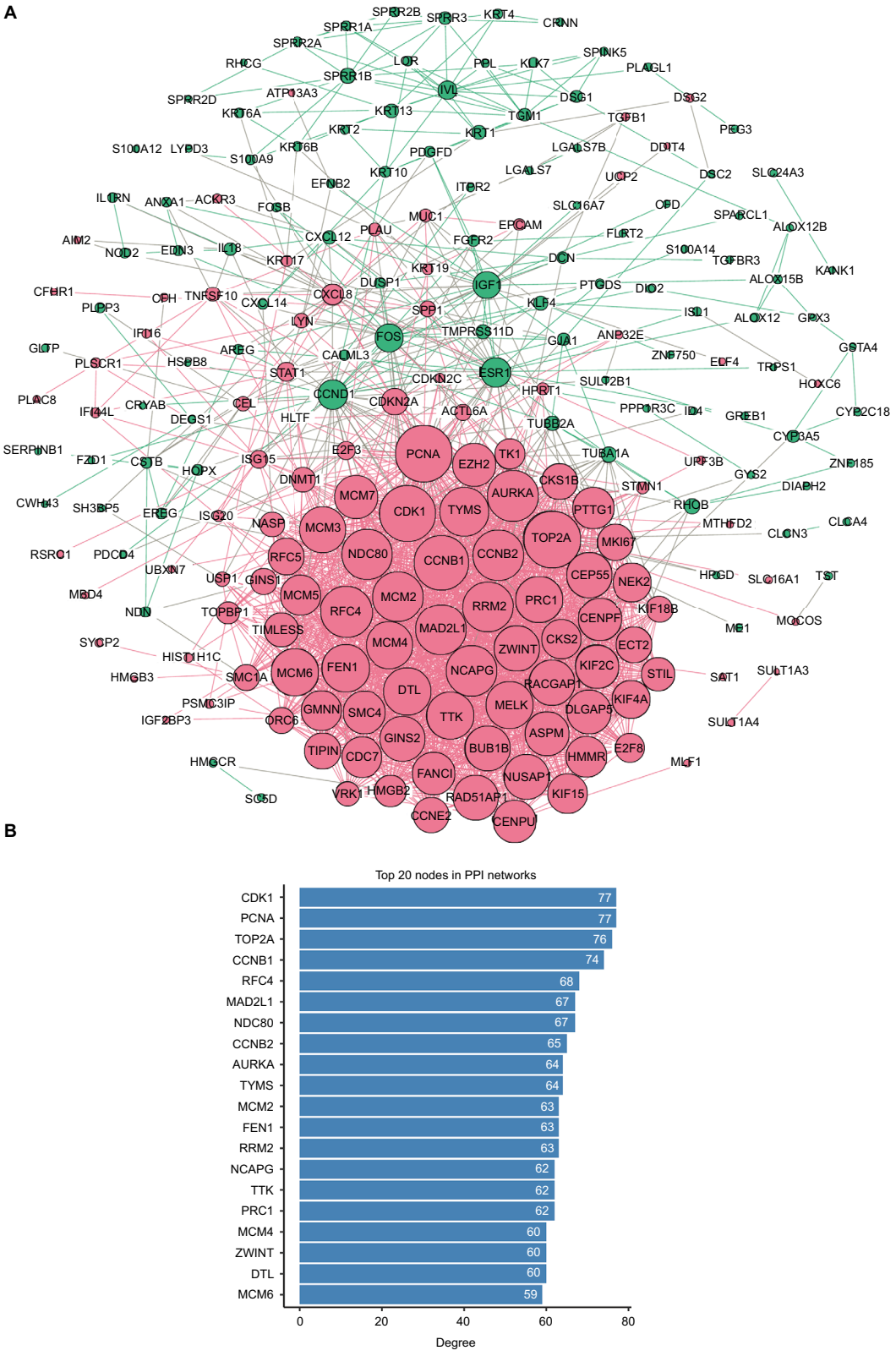

Figure 2 PPI network analysis.

Notes: (**A**) Using the STRING online database, a total of 222 DEGs (118 upregulated in red standing for upregulation and 104 downregulated genes in green standing for downregulation) were filtered into the DEGs PPI network. Bigger nodes represent genes with more links. (**B**) Degree of the top 20 nodes in the PPI network. All these nodes are upregulated genes.

Abbreviations: DEGS, differentially expressed genes; PPI, protein–protein interaction.

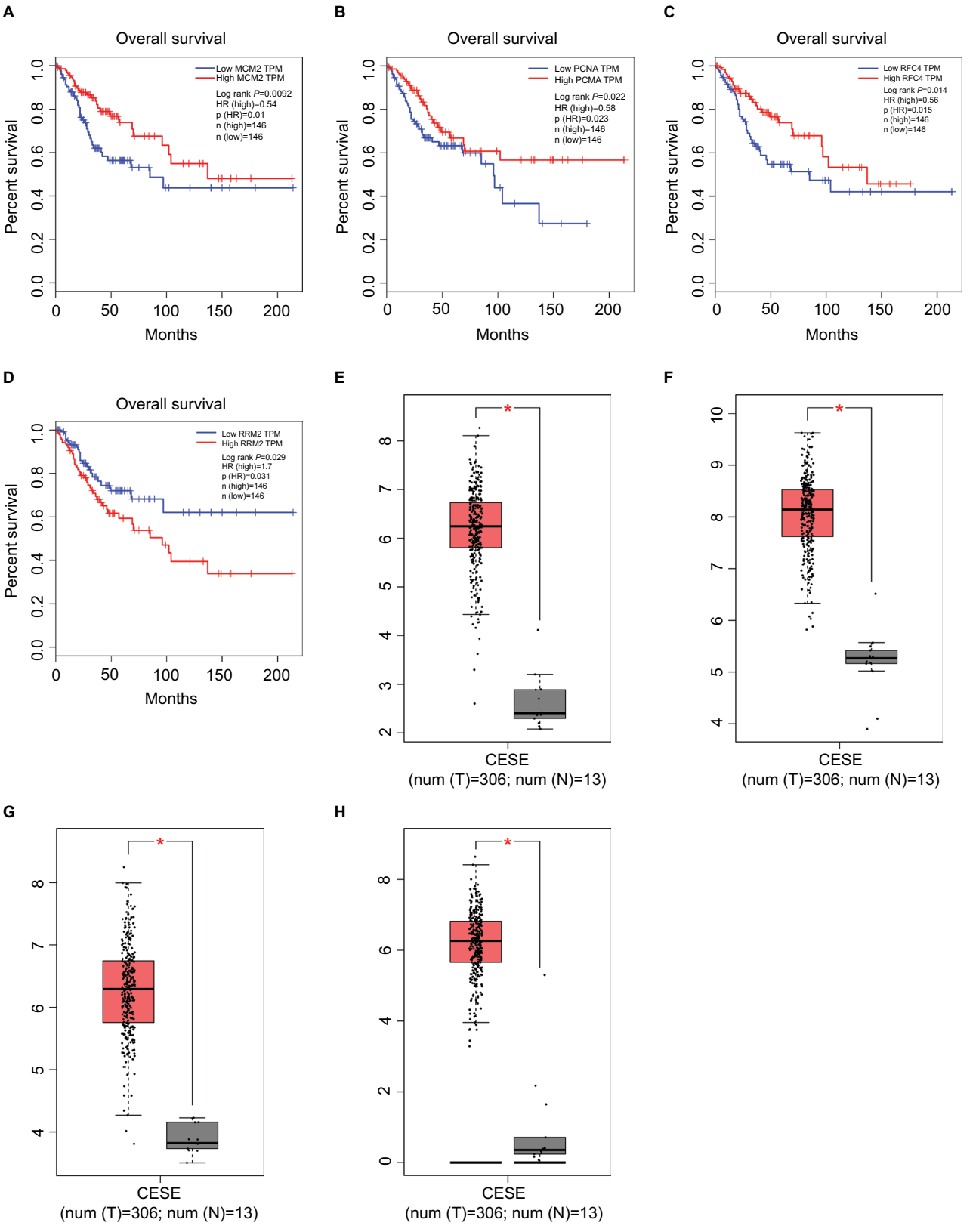

Figure 3 Survival curves and expression boxplots of key genes using GEPIA website.

Notes: (**A–D**) Expression level of *MCM2, PCNA, RFC4, and RRM2* was significantly related to the overall survival of patients with cervical squamous cancer ($P<0.05$). (**E–H**) *MCM2*, *PCNA*, *RFC4*, and *RRM2* were significantly upregulated in cervical squamous cancer compared with normal tissues ($P<0.01$).

Abbreviations: CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; GEPIA, Gene Expression Profiling Interactive Analysis; TPM, transcripts per million.

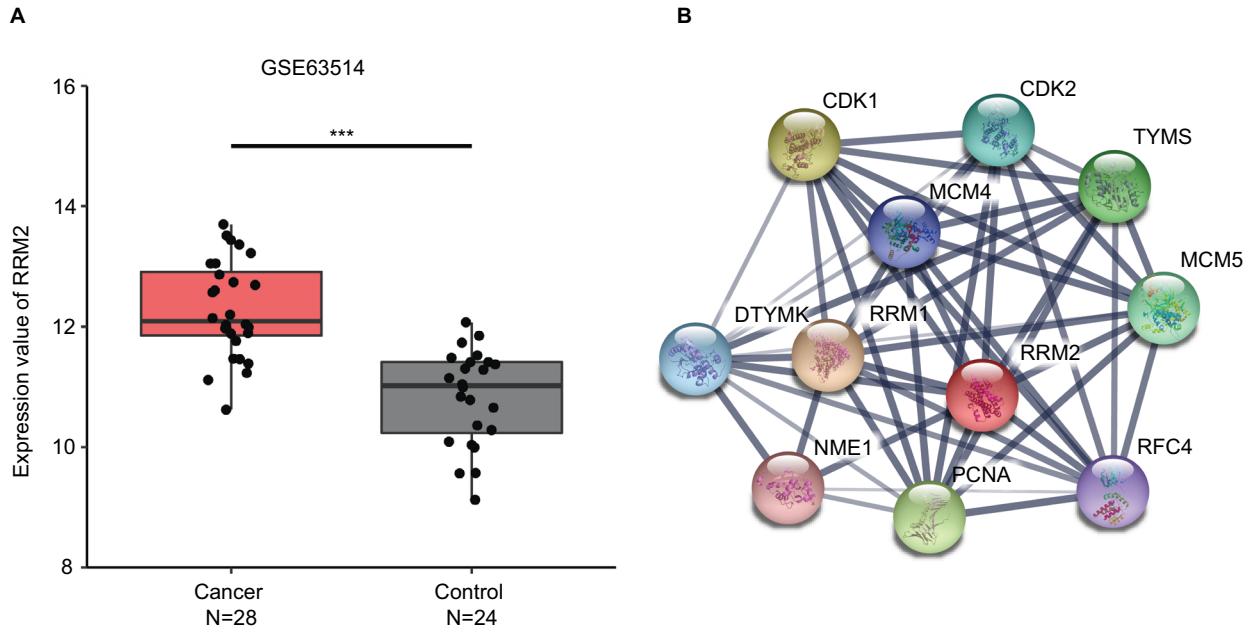

Figure 4 RRM2 validation using GSE63514 and PPI network.
Notes: (**A**) GSE63514 showed higher expression of *RRM2* in CC tissues compared with normal cervical tissues (P<0.01). (**B**) RRM2 PPI network based on STRING.
Abbreviations: CC, cervical cancer; PPI, protein–protein interaction.

Table 2 Results of CMap analysis

Rank	CMap name	Mean	N	Enrichment	*P*-value
1	Trichostatin A	–0.480	182	–0.419	0
2	Tanespimycin	–0.372	62	–0.301	0.00002
3	Vorinostat	–0.551	12	–0.571	0.00034
4	Trifluoperazine	–0.511	16	–0.488	0.00054
5	Prochlorperazine	–0.461	16	–0.436	0.00277
6	Thioridazine	–0.407	20	–0.375	0.00526
7	Alpha-estradiol	–0.367	16	–0.365	0.02104
8	Fluphenazine	–0.403	18	–0.326	0.03608
9	Chlorpromazine	–0.366	19	–0.310	0.04109

Abbreviation: CMap, Connectivity Map.

overexpression of RFC4 is associated with tumor progression and poor survival outcome.[43] Additionally, with gene network reconstruction, RFC4 is regarded as one of the main drivers in cell cycle network in CC.[44] Together with our results, *MCM2*, *PCNA*, and *RFC4* were significantly upregulated in CC compared with normal samples, and in CC patients, the survival rate was positively correlated with the high expression of these genes.

RRM2 is markedly upregulated in many patients' cancer types and indeed acts as an oncogene.[45] *RRM2* knockdown reduces cell proliferation and invasive ability in gastric cancer and pancreatic adenocarcinoma.[46,47] Wang et al reported that RRM2 expression inhibition significantly increases apoptosis, promotes cell cycle arrest at the G1 phase, and inhibits tumor formation in CC nude mice transplant models.[48] Several studies showed that RRM2 is an independent prognostic factor and may predict poor survival in ovarian cancer, bladder cancer, breast cancer, and CC.[49–52] In this study, according to the PPI network, RRM2 closely interacts with PCNA and RFC4 involved in CC progression. Therefore, a further exploration of cell cycle and related genes was of enormous significance.

Consistent with our results, recent studies have also reported the identification of DEGs in CC. van Dam et al used three publicly available Affymetrix gene expression datasets (GSE5787, GSE7803, and GSE9750) and identified five cancer hallmarks enriched pathways in CC, showing that cell cycle deregulation is the major component of CC biology. They also identified seven probesets that were highly expressed in both CIN3 samples compared to normal samples and in cancer samples compared to CIN3 samples. From these probesets, six genes (*AURKA*, *DTL*, *HMGB3*, *KIF2C*, *NEK2*, and *RFC4*) were overexpressed in CC cell lines compared to cancer samples, suggesting their potential role as biomarkers in CC early diagnosis.[53] One of these genes, such as *RFC4*, was also identified in our study. Furthermore, our conclusion generated from both expression and survival analysis suggested that *RFC4* might have a prognostic value. Another report from Li et al was based on TCGA data.[54] They found that *MCM2*, *MCM4*, *MCM5*, *PCNA*, and *RNASEH2A* participating in DNA replication pathway might be prognostic biomarkers in CC patients. *MCM2* and *PCNA* were also found in our results.

Several small molecules with potential therapeutic efficacy against CC were identified. The most significant

small molecules in our result have been reported to display anticancer activity. TSA, as a histone deacetylase (HDAC) inhibitor, shows a potential therapeutic effect in various types of cancer cells, when combined with radiotherapy or chemotherapy.[55,56] In particular, TSA and its hydroxamate analogs can effectively and selectively induce tumor growth arrest at very low concentrations.[57] Additionally, TSA can inhibit HeLa cells growth via Bcl-2-mediated and caspase-dependent apoptosis.[58] Vorinostat is a hydroxamate-based pan-HDAC inhibitor also known as suberoylanilide hydroxamic acid used for the treatment of cutaneous T-cell lymphoma.[59] In HeLa cell, both mRNA and protein levels of HPV18 E6 and E7 were reduced after vorinostat treatment.[60] Furthermore, vorinostat promotes SiHa apoptosis through upregulation of p21 and Bax mRNA and protein, leading to cell cycle arrest in G0/G1 phase.[61] Thioridazine, a derivative of phenothiazine, displays anticancer abilities in a variety of cancer types and can reverse multidrug resistance.[62–64] Kang et al found that thioridazine can inhibit the PI3K/Akt/mTOR/p70S6K signaling pathway and exert cytotoxic effect on CC cells by inducing cell cycle arrest and apoptosis.[65] Thus, we might suppose that these identified drugs could play certain roles to combat CC.

Conclusion

Using bioinformatics analysis, 280 DEGs were identified, which were significantly enriched in several pathways, mainly associated with cell cycle, DNA replication, oocyte meiosis, p53 signaling pathway, and cellular senescence. We also identified key genes including *MCM2*, *PCNA*, *RFC4*, and *RRM2* that might play important roles in CC and that might represent novel biomarkers in CC diagnosis, prognosis, and therapy. Additionally, a group of small molecules was identified that might be exploited as adjuvant drugs for improved therapeutics for CC. However, further investigations are required to validate the predicted drugs.

Acknowledgment

This work was supported by grants from the National Natural Science Foundation of China (grant no 81360336), the Joint Special Funds for the Department of Science and Technology of Yunnan Province – Kunming Medical University (grant no 2015FB017) and One Hundred Young and Middle - Aged Academic and Technical Backbone Project of Kunming Medical University (grant no. 60117190449).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015;65(2):87–108.
2. Forouzanfar MH, Foreman KJ, Delossantos AM, et al. Breast and cervical cancer in 187 countries between 1980 and 2010: a systematic analysis. *Lancet*. 2011;378(9801):1461–1484.
3. Maguire R, Kotronoulas G, Simpson M, Paterson C. A systematic review of the supportive care needs of women living with and beyond cervical cancer. *Gynecol Oncol*. 2015;136(3):478–490.
4. zur Hausen H. Human papillomaviruses in the pathogenesis of anogenital cancer. *Virology*. 1991;184(1):9–13.
5. Ojesina AI, Lichtenstein L, Freeman SS, et al. Landscape of genomic alterations in cervical carcinomas. *Nature*. 2014;506(7488):371–375.
6. Zhai Y, Kuick R, Nan B, et al. Gene expression analysis of preinvasive and invasive cervical squamous cell carcinomas identifies HOXC10 as a key mediator of invasion. *Cancer Res*. 2007;67(21):10163–10172.
7. Scotto L, Narayan G, Nandula SV, et al. Identification of copy number gain and overexpressed genes on chromosome arm 20q by an integrative genomic approach in cervical cancer: potential role in progression. *Genes Chromosomes Cancer*. 2008;47(9):755–765.
8. den Boon JA, Pyeon D, Wang SS, et al. Molecular transitions from papillomavirus infection to cervical precancer and cancer: role of stromal estrogen receptor signaling. *Proc Natl Acad Sci U S A*. 2015;112(25):E3255–E3264.
9. Pappa KI, Polyzos A, Jacob-Hirsch J, et al. Profiling of discrete gynecological cancers reveals novel transcriptional modules and common features shared by other cancer types and embryonic stem cells. *PLoS One*. 2015;10(11):e0142229.
10. Leisching GR, Loos B, Botha MH, Engelbrecht AM. The role of mTOR during cisplatin treatment in an in vitro and ex vivo model of cervical cancer. *Toxicology*. 2015;335:72–78.
11. Tewari KS, Sill MW, Long HJ, et al. Improved survival with bevacizumab in advanced cervical cancer. *N Engl J Med*. 2014;370(8):734–743.
12. Katanyoo K, Tangjitgamol S, Chongthanakorn M, et al. Treatment outcomes of concurrent weekly carboplatin with radiation therapy in locally advanced cervical cancer patients. *Gynecol Oncol*. 2011;123(3):571–576.
13. Wang X, Shen Y, Zhao Y, et al. Adjuvant intensity-modulated radiotherapy (IMRT) with concurrent paclitaxel and cisplatin in cervical cancer patients with high risk factors: a phase II trial. *Eur J Surg Oncol*. 2015;41(8):1082–1088.
14. Downs LS, Chura JC, Argenta PA, et al. Ifosfamide, paclitaxel, and carboplatin, a novel triplet regimen for advanced, recurrent, or persistent carcinoma of the cervix: a phase II trial. *Gynecol Oncol*. 2011;120(2):265–269.
15. Muderspach LI, Blessing JA, Levenback C, Moore JL Jr. A phase II study of topotecan in patients with squamous cell carcinoma of the cervix: a gynecologic oncology group study. *Gynecol Oncol*. 2001;81(2):213–215.
16. Eskander RN, Tewari KS. Chemotherapy in the treatment of metastatic, persistent, and recurrent cervical cancer. *Curr Opin Obstet Gynecol*. 2014;26(4):314–321.
17. Lorusso D, Petrelli F, Coinu A, Raspagliesi F, Barni S. A systematic review comparing cisplatin and carboplatin plus paclitaxel-based chemotherapy for recurrent or metastatic cervical cancer. *Gynecol Oncol*. 2014;133(1):117–123.

18. Yeh CT, Wu AT, Chang PM, et al. Trifluoperazine, an antipsychotic agent, inhibits cancer stem cell growth and overcomes drug resistance of lung cancer. *Am J Respir Crit Care Med*. 2012;186(11):1180–1188.
19. Chen MH, Lin KJ, Yang WL, et al. Gene expression-based chemical genomics identifies heat-shock protein 90 inhibitors as potential therapeutic drugs in cholangiocarcinoma. *Cancer*. 2013;119(2):293–303.
20. Hassane DC, Guzman ML, Corbett C, et al. Discovery of agents that eradicate leukemia stem cells using an in silico screen of public gene expression data. *Blood*. 2008;111(12):5654–5662.
21. Lamb J, Crawford ED, Peck D, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science*. 2006;313(5795):1929–1935.
22. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy – analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*. 2004;20(3):307–315.
23. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
24. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–287.
25. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–2504.
26. Szklarczyk D, Morris JH, Cook H, et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res*. 2017;45(D1):D362–D368.
27. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. 2017;45(W1):W98–W102.
28. Cancer Genome Atlas Research Network; Weinstein JN, Collisson EA, Mills GB, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet*. 2013;45(10):1113–1120.
29. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet*. 2013;45(6):580–585.
30. Wang Y, Li Y, Zhang WY, et al. mRNA expression of minichromosome maintenance 2 in colonic adenoma and adenocarcinoma. *Eur J Cancer Prev*. 2009;18(1):40–45.
31. Szelachowska J, Dziegiel P, Jelen-Krzeszewska J, et al. Correlation of metallothionein expression with clinical progression of cancer in the oral cavity. *Anticancer Res*. 2009;29(2):589–595.
32. Gakiopoulou H, Korkolopoulou P, Levidou G, et al. Minichromosome maintenance proteins 2 and 5 in non-benign epithelial ovarian tumours: relationship with cell cycle regulators and prognostic implications. *Br J Cancer*. 2007;97(8):1124–1134.
33. Burger M, Denzinger S, Hartmann A, Wieland WF, Stoehr R, Obermann EC. Mcm2 predicts recurrence hazard in stage Ta/T1 bladder cancer more accurately than CK20, Ki67 and histological grade. *Br J Cancer*. 2007;96(11):1711–1715.
34. Zhang X, Teng Y, Yang F, et al. MCM2 is a therapeutic target of lovastatin in human non-small cell lung carcinomas. *Oncol Rep*. 2015;33(5):2599–2605.
35. Zheng J. Diagnostic value of MCM2 immunocytochemical staining in cervical lesions and its relationship with HPV infection. *Int J Clin Exp Pathol*. 2015;8(1):875–880.
36. Amaro Filho SM, Nuovo GJ, Cunha CB, et al. Correlation of MCM2 detection with stage and virology of cervical cancer. *Int J Biol Markers*. 2014;29(4):e363–e371.
37. Tachibana KE, Gonzalez MA, Coleman N. Cell-cycle-dependent regulation of DNA replication and its relevance to cancer pathology. *J Pathol*. 2005;205(2):123–129.
38. Zhao H, Lo YH, Ma L, et al. Targeting tyrosine phosphorylation of PCNA inhibits prostate cancer growth. *Mol Cancer Ther*. 2011;10(1):29–36.
39. Wang LF, Chai CY, Kuo WR, Tai CF, Lee KW, Ho KY. Correlation between proliferating cell nuclear antigen and p53 protein expression and 5-year survival rate in nasopharyngeal carcinoma. *Am J Otolaryngol*. 2006;27(2):101–105.
40. Lv Q, Zhang J, Yi Y, et al. Proliferating cell nuclear antigen has an association with prognosis and risks factors of cancer patients: a systematic review. *Mol Neurobiol*. 2016;53(9):6209–6217.
41. Arai M, Kondoh N, Imazeki N, et al. The knockdown of endogenous replication factor C4 decreases the growth and enhances the chemosensitivity of hepatocellular carcinoma cells. *Liver Int*. 2009;29(1):55–62.
42. Srihari S, Kalimutho M, Lal S, et al. Understanding the functional impact of copy number alterations in breast cancer using a network modeling approach. *Mol Biosyst*. 2016;12(3):963–972.
43. Xiang J, Fang L, Luo Y, et al. Levels of human replication factor C4, a clamp loader, correlate with tumor progression and predict the prognosis for colorectal cancer. *J Transl Med*. 2014;12:320.
44. Mine KL, Shulzhenko N, Yambartsev A, et al. Gene network reconstruction reveals cell cycle and antiviral genes as major drivers of cervical cancer. *Nat Commun*. 2013;4:1806.
45. Xu X, Page JL, Surtees JA, et al. Broad overexpression of ribonucleotide reductase genes in mice specifically induces lung neoplasms. *Cancer Res*. 2008;68(8):2652–2660.
46. Kang W, Tong JH, Chan AW, et al. Targeting ribonucleotide reductase M2 subunit by small interfering RNA exerts anti-oncogenic effects in gastric adenocarcinoma. *Oncol Rep*. 2014;31(6):2579–2586.
47. Duxbury MS, Whang EE. RRM2 induces NF-kappaB-dependent MMP-9 activation and enhances cellular invasiveness. *Biochem Biophys Res Commun*. 2007;354(1):190–196.
48. Wang N, Li Y, Zhou J. Downregulation of ribonucleotide reductase subunits M2 induces apoptosis and G1 arrest of cervical cancer cells. *Oncol Lett*. 2018;15(3):3719–3725.
49. Ferrandina G, Mey V, Nannizzi S, et al. Expression of nucleoside transporters, deoxycitidine kinase, ribonucleotide reductase regulatory subunits, and gemcitabine catabolic enzymes in primary ovarian cancer. *Cancer Chemother Pharmacol*. 2010;65(4):679–686.
50. Morikawa T, Maeda D, Kume H, Homma Y, Fukayama M. Ribonucleotide reductase M2 subunit is a novel diagnostic marker and a potential therapeutic target in bladder cancer. *Histopathology*. 2010;57(6):885–892.
51. Kretschmer C, Sterner-Kock A, Siedentopf F, Schoenegg W, Schlag PM, Kemmner W. Identification of early molecular markers for breast cancer. *Mol Cancer*. 2011;10(1):15.
52. Su YF, Wu TF, Ko JL, et al. The expression of ribonucleotide reductase M2 in the carcinogenesis of uterine cervix and its relationship with clinicopathological characteristics and prognosis of cancer patients. *PLoS One*. 2014;9(3):e91644.
53. van Dam PA, van Dam PJ, Rolfo C, et al. In silico pathway analysis in cervical carcinoma reveals potential new targets for treatment. *Oncotarget*. 2016;7(3):2780–2795.
54. Li X, Tian R, Gao H, et al. Identification of significant gene signatures and prognostic biomarkers for patients with cervical cancer by integrated bioinformatic methods. *Technol Cancer Res Treat*. 2018;17:1–12.
55. Ranganathan P, Rangnekar VM. Exploiting the TSA connections to overcome apoptosis-resistance. *Cancer Biol Ther*. 2005;4(4):391–392.
56. Hajji N, Wallenborg K, Vlachos P, Nyman U, Hermanson O, Joseph B. Combinatorial action of the HDAC inhibitor trichostatin A and etoposide induces caspase-mediated AIF-dependent apoptotic cell death in non-small cell lung carcinoma cells. *Oncogene*. 2008;27(22):3134–3144.
57. Vanhaecke T, Papeleu P, Elaut G, Rogiers V. Trichostatin A-like hydroxamate histone deacetylase inhibitors as therapeutic agents: toxicological point of view. *Curr Med Chem*. 2004;11(12):1629–1643.
58. You BR, Park WH. Trichostatin A induces apoptotic cell death of HeLa cells in a Bcl-2 and oxidative stress-dependent manner. *Int J Oncol*. 2013;42(1):359–366.
59. Duvic M, Vu J. Vorinostat: a new oral histone deacetylase inhibitor approved for cutaneous T-cell lymphoma. *Expert Opin Investig Drugs*. 2007;16(7):1111–1120.

60. He H, Liu X, Wang D, et al. SAHA inhibits the transcription initiation of HPV18 E6/E7 genes in HeLa cervical cancer cells. *Gene*. 2014;553(2):98–104.
61. Xing J, Wang H, Xu S, Han P, Xin M, Zhou JL. Sensitization of suberoylanilide hydroxamic acid (SAHA) on chemoradiation for human cervical cancer cells and its mechanism. *Eur J Gynaecol Oncol*. 2015;36(2):117–122.
62. Gil-Ad I, Shtaif B, Levkovitz Y, et al. Phenothiazines induce apoptosis in a B16 mouse melanoma cell line and attenuate in vivo melanoma tumor growth. *Oncol Rep*. 2006;15(1):107–112.
63. Li J, Yao QY, Xue JS, et al. Dopamine D2 receptor antagonist sulpiride enhances dexamethasone responses in the treatment of drug-resistant and metastatic breast cancer. *Acta Pharmacol Sin*. 2017;38(9):1282–1296.
64. Seo SU, Cho HK, Min KJ, et al. Thioridazine enhances sensitivity to carboplatin in human head and neck cancer cells through downregulation of c-FLIP and Mcl-1 expression. *Cell Death Dis*. 2017;8(2):e2599.
65. Kang S, Dong SM, Kim BR, et al. Thioridazine induces apoptosis by targeting the PI3K/Akt/mTOR pathway in cervical and endometrial cancer cells. *Apoptosis*. 2012;17(9):989–997.

Supplementary material

Table S1 Common dysregulated probes identified in GSE7803 and GSE9750

Number	Probe name	Gene symbol	logFC		Adjusted *P*-value	
			GSE7803	GSE9750	GSE7803	GSE9750
Upregulated						
1	200783_s_at	*STMN1*	1.0912	1.0985	4.11E–05	1.25E–04
2	201202_at	*PCNA*	1.8714	1.3752	4.41E–09	2.87E–05
3	201291_s_at	*TOP2A*	2.4680	2.4862	1.05E–08	5.16E–06
4	201292_at	*TOP2A*	1.2403	2.0680	1.71E–06	2.30E–04
5	201506_at	*TGFBI*	1.1157	1.1297	2.21E–02	1.60E–02
6	201555_at	*MCM3*	1.0145	1.2759	3.14E–07	5.41E–07
7	201589_at	*SMC1A*	1.3907	1.0445	5.43E–06	8.52E–05
8	201650_at	*KRT19*	1.5763	2.3797	2.89E–02	8.95E–05
9	201663_s_at	*SMC4*	1.4923	1.5673	1.60E–06	4.31E–05
10	201664_at	*SMC4*	1.7038	1.8408	9.71E–06	7.08E–06
11	201697_s_at	*DNMT1*	1.0219	1.2128	1.05E–08	4.10E–09
12	201761_at	*MTHFD2*	1.6111	1.3783	4.73E–05	4.37E–05
13	201839_s_at	*EPCAM*	1.6182	2.2101	3.36E–04	1.82E–04
14	201890_at	*RRM2*	1.7698	2.3342	1.05E–05	1.44E–06
15	201897_s_at	*CKS1B*	1.3398	1.2246	1.04E–06	4.26E–05
16	201930_at	*MCM6*	1.5487	1.4628	6.91E–09	1.25E–08
17	201970_s_at	*NASP*	1.3464	1.1522	1.38E–08	9.44E–09
18	202107_s_at	*MCM2*	1.7296	2.2864	1.91E–08	7.29E–10
19	202219_at	*SLC6A8*	1.4325	1.8698	5.38E–03	2.75E–05
20	202234_s_at	*SLC16A1*	1.3561	1.2683	4.70E–03	7.50E–04
21	202338_at	*TK1*	1.1038	1.2846	2.40E–06	2.91E–06
22	202412_s_at	*USP1*	1.0775	1.2056	3.24E–04	3.54E–04
23	202430_s_at	*PLSCR1*	1.2889	1.1148	1.29E–04	7.10E–05
24	202446_s_at	*PLSCR1*	1.6150	1.5934	1.16E–06	1.41E–06
25	202503_s_at	*PCLAF*	2.0118	1.9842	2.33E–10	1.18E–04
26	202589_at	*TYMS*	1.5263	2.2375	1.95E–04	8.52E–08
27	202619_s_at	*PLOD2*	1.8339	2.0812	1.44E–06	1.41E–07
28	202620_s_at	*PLOD2*	2.8767	2.2084	3.54E–09	5.09E–07
29	202625_at	*LYN*	1.1767	1.4027	1.20E–03	2.09E–03
30	202626_s_at	*LYN*	1.5737	1.7521	1.07E–03	8.51E–05
31	202633_at	*TOPBP1*	1.6677	1.4670	3.67E–07	7.76E–07
32	202666_s_at	*ACTL6A*	1.5525	1.1292	9.46E–07	1.46E–03
33	202688_at	*TNFSF10*	1.0878	1.1141	3.63E–02	3.00E–02
34	202705_at	*CCNB2*	1.0405	1.7114	7.17E–05	8.58E–06
35	202854_at	*HPRT1*	1.0911	1.2618	2.75E–04	7.59E–05
36	202859_x_at	*CXCL8*	1.5100	2.7989	1.66E–02	3.86E–04
37	202887_s_at	*DDIT4*	1.0791	2.1221	4.05E–02	3.21E–04
38	202983_at	*HLTF*	2.0242	1.2271	4.08E–07	2.36E–03
39	203046_s_at	*TIMELESS*	1.1641	1.4554	1.96E–08	1.55E–07
40	203209_at	*RFC5*	1.5612	1.2914	1.93E–07	7.95E–06
41	203213_at	*CDK1*	1.5427	1.9989	1.11E–05	9.35E–06
42	203358_s_at	*EZH2*	1.6144	1.6656	3.31E–07	3.91E–05
43	203362_s_at	*MAD2L1*	1.0749	1.2051	4.94E–03	1.40E–02
44	203554_x_at	*PTTG1*	1.2746	1.3424	1.32E–05	3.11E–03
45	203693_s_at	*E2F3*	1.2182	1.2321	1.16E–06	1.28E–04
46	203744_at	*HMGB3*	1.1938	1.6479	3.01E–04	1.41E–07
47	203755_at	*BUB1B*	1.0860	1.8703	3.34E–07	1.96E–06
48	203764_at	*DLGAP5*	1.4536	2.2334	3.73E–05	9.69E–05

(*Continued*)

Table S1 (*Continued*)

Number	Probe name	Gene symbol	logFC		Adjusted *P*-value	
			GSE7803	GSE9750	GSE7803	GSE9750
49	203819_s_at	*IGF2BP3*	1.2319	1.4909	2.75E–02	2.22E–02
50	203856_at	*VRK1*	1.1485	1.1641	4.00E–06	2.15E–04
51	204023_at	*RFC4*	1.5116	2.1996	7.47E–08	2.54E–07
52	204026_s_at	*ZWINT*	1.5906	1.8649	2.05E–04	2.07E–05
53	204092_s_at	*AURKA*	1.1672	1.0470	1.39E–07	2.36E–05
54	204146_at	*RAD51AP1*	1.6216	1.6478	4.26E–07	8.74E–05
55	204159_at	*CDKN2C*	1.5231	1.2248	2.07E–03	3.73E–03
56	204162_at	*NDC80*	1.5684	1.3280	9.05E–05	1.03E–03
57	204170_s_at	*CKS2*	1.4786	1.5687	6.24E–05	5.32E–03
58	204416_x_at	*APOC1*	1.0620	1.3204	8.58E–03	9.07E–04
59	204439_at	*IFI44L*	1.5993	1.4595	4.00E–02	5.25E–02
60	204510_at	*CDC7*	1.3374	1.6750	1.88E–06	2.38E–06
61	204580_at	*MMP12*	1.6409	2.9620	2.22E–03	2.97E–05
62	204641_at	*NEK2*	1.3957	2.1694	6.79E–08	2.33E–07
63	204698_at	*ISG20*	1.0250	1.3923	6.87E–04	1.78E–05
64	204767_s_at	*FEN1*	1.4080	1.7083	1.31E–09	3.37E–07
65	204784_s_at	*MLF1*	1.6420	1.6891	6.24E–05	3.01E–04
66	204822_at	*TTK*	1.4460	1.4235	7.52E–08	2.45E–03
67	204825_at	*MELK*	1.8745	1.9957	2.90E–07	3.60E–07
68	205034_at	*CCNE2*	1.3408	1.7091	4.18E–04	5.08E–04
69	205157_s_at	*KRT17*	1.5725	3.3509	2.32E–02	1.54E–05
70	205339_at	*STIL*	1.0641	1.5078	1.16E–06	9.67E–08
71	205449_at	*SAC3D1*	1.3483	1.1408	4.86E–05	2.80E–04
72	205479_s_at	*PLAU*	1.2262	1.4554	1.65E–03	4.38E–04
73	205483_s_at	*ISG15*	1.3717	2.0505	1.65E–02	5.73E–04
74	205569_at	*LAMP3*	1.7550	1.2873	4.09E–04	2.57E–02
75	205691_at	*SYNGR3*	1.2198	1.4950	5.72E–04	3.29E–04
76	205910_s_at	*CEL*	1.8737	1.2510	1.62E–02	3.08E–02
77	206102_at	*GINS1*	1.6544	1.9896	6.95E–05	4.57E–06
78	206332_s_at	*IFI16*	1.5947	1.2876	8.99E–07	2.56E–04
79	206513_at	*AIM2*	2.0306	2.3769	1.22E–03	2.74E–03
80	206546_at	*SYCP2*	1.3491	2.3512	2.17E–03	5.99E–05
81	206632_s_at	*APOBEC3A*, *APOBEC3B*	2.9688	1.9572	1.68E–08	2.46E–04
82	206858_s_at	*HOXC6*	2.1365	1.4749	1.66E–05	1.02E–03
83	207039_at	*CDKN2A*	4.6085	4.0377	3.50E–14	1.62E–14
84	207165_at	*HMMR*	1.4593	1.0523	2.62E–06	1.30E–02
85	207332_s_at	*TFRC*	1.2833	1.3954	4.66E–04	5.87E–03
86	207828_s_at	*CENPF*	1.4100	1.9877	1.36E–07	9.05E–08
87	208079_s_at	*AURKA*	2.2857	2.0803	1.85E–09	5.77E–08
88	208691_at	*TFRC*	1.5192	1.4901	6.53E–06	5.22E–04
89	208795_s_at	*MCM7*, *MIR25*, *MIR93*, *MIR106B*	1.1151	1.3206	1.55E–07	2.44E–05
90	208808_s_at	*HMGB2*	1.5215	1.1577	8.56E–07	7.41E–04
91	208965_s_at	*IFI16*	1.4542	1.3939	2.56E–05	6.32E–04
92	208966_x_at	*IFI16*	1.7717	1.3388	4.13E–07	6.92E–05
93	208998_at	*UCP2*	1.9521	1.1262	2.31E–05	2.83E–03
94	209398_at	*HIST1H1C*	1.1786	1.2497	5.37E–03	7.31E–03
95	209408_at	*KIF2C*	1.4768	1.5638	1.85E–09	4.49E–09
96	209579_s_at	*MBD4*	1.2877	1.1884	6.26E–07	6.34E–05
97	209773_s_at	*RRM2*	1.2805	1.9504	2.40E–03	6.29E–05
98	209875_s_at	*SPP1*	2.5457	3.4037	3.80E–04	3.18E–06
99	209900_s_at	*SLC16A1*	1.5149	1.2504	1.45E–03	1.73E–03
100	209969_s_at	*STAT1*	1.8886	2.1349	6.82E–04	4.07E–04

(*Continued*)

Table S1 (*Continued*)

Number	Probe name	Gene symbol	logFC		Adjusted *P*-value	
			GSE7803	GSE9750	GSE7803	GSE9750
101	210580_x_at	*SLX1A-SULT1A3*, *SLX1B-SULT1A4*, *SULT1A3*, *SULT1A4*	1.0598	1.0491	1.80E–03	1.19E–03
102	212022_s_at	*MKI67*	1.4831	1.5781	8.11E–07	1.81E–06
103	212236_x_at	*KRT17*	1.3508	2.7588	3.84E–02	7.95E–06
104	212255_s_at	*ATP2C1*	1.0824	1.0588	1.60E–04	8.54E–05
105	212297_at	*ATP13A3*	1.4423	1.1315	1.69E–06	6.46E–05
106	212621_at	*NEMP1*	1.4685	1.2100	1.63E–07	2.29E–07
107	212840_at	*UBXN7*	1.0252	1.0213	1.25E–04	1.36E–03
108	212977_at	*ACKR3*	2.0327	1.2298	4.91E–03	5.31E–02
109	213007_at	*FANCI*	1.3983	1.6034	2.79E–07	2.80E–07
110	213008_at	*FANCI*	1.0861	1.6205	4.34E–05	4.91E–08
111	213164_at	*SLC5A3*	1.0329	1.0596	2.97E–05	4.04E–04
112	213457_at	*MFHAS1*	1.0606	1.0186	3.06E–02	2.25E–03
113	213693_s_at	*MUC1*	1.2200	1.9901	3.72E–02	4.37E–04
114	213951_s_at	*PSMC3IP*	1.2274	1.2423	1.92E–07	2.52E–07
115	213988_s_at	*SAT1*	1.1293	1.0600	2.52E–03	7.87E–04
116	214329_x_at	*TNFSF10*	2.3354	1.2571	6.72E–06	1.12E–02
117	214710_s_at	*CCNB1*	1.2879	1.7956	3.73E–05	1.97E–04
118	215388_s_at	*CFH*, *CFHR1*	1.4348	1.0577	1.61E–02	4.50E–02
119	216237_s_at	*MCM5*	1.5683	2.0746	1.77E–08	8.15E–11
120	217885_at	*IPO9*	1.1126	1.0339	1.76E–07	1.89E–05
121	217901_at	*DSG2*	1.3065	2.5311	6.24E–05	3.60E–07
122	218009_s_at	*PRC1*	1.6401	2.1259	5.88E–08	1.78E–06
123	218039_at	*NUSAP1*	2.1401	2.3735	7.34E–09	5.69E–06
124	218350_s_at	*GMNN*	1.7878	1.8225	3.13E–07	2.66E–06
125	218355_at	*KIF4A*	1.0153	1.7434	2.38E–06	2.89E–07
126	218542_at	*CEP55*	1.3865	2.4903	2.46E–06	2.29E–07
127	218585_s_at	*DTL*	1.3800	2.8428	2.32E–06	4.81E–09
128	218662_s_at	*NCAPG*	1.6736	1.5539	4.30E–06	1.30E–04
129	218757_s_at	*UPF3B*	1.3448	1.0318	6.72E–05	2.68E–05
130	218883_s_at	*CENPU*	1.2494	1.5454	7.55E–04	3.91E–03
131	219014_at	*PLAC8*	1.2330	1.4057	2.84E–02	3.55E–02
132	219105_x_at	*ORC6*	1.0780	1.0908	2.10E–04	5.03E–07
133	219258_at	*TIPIN*	1.1365	1.4832	3.91E–07	1.27E–06
134	219306_at	*KIF15*	1.0990	1.0348	1.84E–04	1.03E–03
135	219507_at	*RSRC1*	1.3864	1.2573	3.27E–05	3.86E–04
136	219787_s_at	*ECT2*	2.8139	2.5551	1.00E–08	1.37E–06
137	219918_s_at	*ASPM*	1.2168	1.9490	4.36E–05	2.25E–04
138	219959_at	*MOCOS*	1.0344	1.9971	3.59E–03	1.67E–05
139	219978_s_at	*NUSAP1*	1.1780	1.6455	8.94E–04	6.82E–05
140	219990_at	*E2F8*	1.4188	1.1301	1.17E–04	1.10E–03
141	220239_at	*KLHL7*	1.0503	1.0053	3.05E–03	1.40E–03
142	221046_s_at	*GTPBP8*	1.0722	1.0602	1.44E–06	4.04E–04
143	221521_s_at	*GINS2*	1.4631	1.8407	8.85E–06	3.27E–06
144	222036_s_at	*MCM4*	1.0134	1.8445	1.71E–06	1.67E–06
145	222039_at	*KIF18B*	1.5126	1.0619	7.17E–09	2.91E–06
146	222077_s_at	*RACGAP1*	1.5482	1.6939	2.69E–06	5.07E–05
147	222380_s_at	*PDCD6*	1.0922	1.0579	3.96E–04	1.18E–02
148	31845_at	*ELF4*	1.2212	1.0448	2.96E–05	1.00E–05
149	33304_at	*ISG20*	1.1281	1.1279	1.62E–04	7.59E–05
Downregulated						
1	200795_at	*SPARCL1*	−2.6933	−2.5139	3.39E–04	2.05E–04
2	201012_at	*ANXA1*	−1.6637	−1.2447	1.15E–03	1.28E–03

(*Continued*)

Table S1 (*Continued*)

Number	Probe name	Gene symbol	logFC		Adjusted *P*-value	
			GSE7803	GSE9750	GSE7803	GSE9750
3	201041_s_at	*DUSP1*	−1.7177	−1.2448	7.16E−03	3.26E−02
4	201201_at	*CSTB*	−1.6745	−1.0817	6.75E−04	2.05E−04
5	201312_s_at	*SH3BGRL*	−1.0735	−2.0334	1.43E−02	3.28E−04
6	201324_at	*EMP1*	−2.5006	−2.0144	1.69E−05	7.39E−05
7	201325_s_at	*EMP1*	−2.8729	−2.3264	2.29E−08	3.96E−06
8	201348_at	*GPX3*	−1.6408	−2.9297	1.32E−05	9.66E−08
9	201667_at	*GJA1*	−2.1421	−2.0359	3.83E−03	6.21E−03
10	201735_s_at	*CLCN3*	−1.1179	−1.0850	1.09E−03	1.89E−03
11	201811_x_at	*SH3BP5*	−1.2423	−1.3642	3.25E−03	5.04E−03
12	201893_x_at	*DCN*	−1.2951	−2.2495	1.17E−03	7.34E−04
13	202539_s_at	*HMGCR*	−1.0673	−1.3111	2.12E−03	1.05E−02
14	202575_at	*CRABP2*	−1.1350	−1.9855	2.03E−07	1.63E−03
15	202660_at	*ITPR2*	−1.2894	−1.0723	6.71E−07	2.03E−04
16	202668_at	*EFNB2*	−1.0648	−1.0430	2.83E−03	1.00E−02
17	202768_at	*FOSB*	−1.6730	−2.2309	6.70E−03	5.21E−03
18	202967_at	*GSTA4*	−1.4779	−1.8389	4.26E−07	2.57E−04
19	203407_at	*PPL*	−1.5681	−1.8813	3.47E−05	3.27E−06
20	203535_at	*S100A9*	−1.9766	−1.2926	8.14E−03	1.86E−02
21	203585_at	*ZNF185*	−1.3599	−1.3991	1.52E−03	4.20E−05
22	203638_s_at	*FGFR2*	−1.3240	−1.5564	3.23E−04	7.27E−04
23	203700_s_at	*DIO2*	−1.3828	−1.1342	1.94E−03	4.10E−02
24	203913_s_at	*HPGD*	−1.7678	−2.7646	5.41E−05	3.98E−05
25	203914_x_at	*HPGD*	−2.4427	−2.7244	1.11E−05	2.65E−05
26	203961_at	*NEBL*	−1.4367	−1.5881	2.34E−03	5.38E−03
27	204141_at	*TUBB2A*	−1.7240	−1.5928	1.18E−03	1.89E−04
28	204256_at	*ELOVL6*	−1.3219	−1.1355	4.57E−03	4.75E−02
29	204284_at	*PPP1R3C*	−2.6784	−3.7692	4.26E−07	2.33E−09
30	204359_at	*FLRT2*	−1.1097	−2.2141	5.71E−03	4.36E−05
31	204451_at	*FZD1*	−1.1135	−1.0997	1.15E−05	1.01E−04
32	204731_at	*TGFBR3*	−1.3493	−1.8325	7.17E−03	4.04E−04
33	204750_s_at	*DSC2*	−2.0225	−1.1059	2.83E−04	4.70E−02
34	204751_x_at	*DSC2*	−1.9548	−2.2472	2.98E−04	1.15E−04
35	204777_s_at	*MAL*	−4.8179	−5.7789	9.50E−07	1.62E−14
36	204952_at	*LYPD3*	−1.5886	−1.7448	1.08E−04	4.43E−04
37	205064_at	*SPRR1B*	−2.2769	−2.7744	1.39E−03	1.20E−02
38	205185_at	*SPINK5*	−3.8683	−3.6665	3.46E−07	3.15E−05
39	205225_at	*ESR1*	−3.0458	−2.7160	9.37E−06	1.54E−05
40	205239_at	*AREG*	−1.8099	−1.5361	3.96E−02	1.14E−03
41	205363_at	*BBOX1*	−1.8640	−2.8822	2.54E−09	1.27E−05
42	205382_s_at	*CFD*	−2.1856	−2.5747	9.39E−07	1.20E−06
43	205470_s_at	*KLK11*	−1.9196	−2.3007	8.84E−08	1.78E−03
44	205726_at	*DIAPH2*	−1.0814	−1.4450	1.17E−03	1.34E−04
45	205759_s_at	*SULT2B1*	−1.2047	−1.2198	1.04E−06	5.73E−03
46	205765_at	*CYP3A5*	−1.8262	−1.1731	3.95E−06	5.05E−03
47	205767_at	*EREG*	−1.6854	−2.1525	1.38E−04	8.11E−05
48	205778_at	*KLK7*	−1.3998	−1.7611	4.35E−03	1.13E−02
49	205862_at	*GREB1*	−1.5579	−1.8147	1.16E−03	1.94E−06
50	205863_at	*S100A12*	−1.2720	−2.0771	4.51E−03	6.18E−04
51	205900_at	*KRT1*	−4.8450	−5.1604	9.91E−10	1.22E−06
52	206008_at	*TGM1*	−1.3988	−1.4672	3.53E−03	2.12E−02
53	206104_at	*ISL1*	−1.8146	−1.8069	4.46E−05	4.23E−06
54	206295_at	*IL18*	−1.8318	−1.1817	9.99E−08	1.78E−02

(*Continued*)

Table S1 (*Continued*)

Number	Probe name	Gene symbol	logFC		Adjusted *P*-value	
			GSE7803	GSE9750	GSE7803	GSE9750
55	206400_at	*LGALS7, LGALS7B*	−1.2508	−1.7496	1.50E−02	2.32E−02
56	206605_at	*ENDOU*	−2.0623	−3.5113	1.38E−10	3.61E−09
57	206642_at	*DSG1*	−3.6072	−4.3758	2.39E−09	6.70E−07
58	206714_at	*ALOX15B*	−1.2685	−1.0664	8.35E−04	2.03E−02
59	206884_s_at	*SCEL*	−2.4970	−3.2369	4.84E−05	3.50E−06
60	207002_s_at	*PLAGL1*	−1.1346	−1.2887	9.02E−03	5.44E−04
61	207023_x_at	*KRT10*	−1.6054	−1.6701	5.50E−03	1.32E−03
62	207057_at	*SLC16A7*	−1.5167	−1.0148	2.07E−06	2.86E−02
63	207206_s_at	*ALOX12*	−2.4692	−2.9129	2.60E−07	5.89E−06
64	207381_at	*ALOX12B*	−1.8250	−1.5189	8.81E−07	1.92E−02
65	207463_x_at	*PRSS3*	−1.6595	−2.2675	1.69E−05	7.14E−04
66	207480_s_at	*MEIS2*	−1.1053	−1.4587	8.12E−03	3.86E−03
67	207602_at	*TMPRSS11D*	−1.7796	−2.2185	1.94E−04	1.24E−03
68	207720_at	*LOR*	−1.5659	−1.7321	9.03E−03	5.54E−03
69	207761_s_at	*METTL7A*	−1.4377	−1.7274	1.64E−02	1.54E−03
70	207802_at	*CRISP3*	−3.5353	−4.9186	8.56E−07	2.03E−08
71	207908_at	*KRT2*	−1.0700	−1.7438	7.77E−06	2.44E−04
72	207935_s_at	*KRT13*	−3.3723	−3.4606	6.80E−04	5.75E−03
73	208126_s_at	*CYP2C18*	−1.0287	−1.2034	3.17E−04	2.53E−02
74	208228_s_at	*FGFR2*	−1.1180	−1.5103	5.59E−03	2.23E−03
75	208399_s_at	*EDN3*	−1.7159	−2.7146	2.90E−07	1.83E−07
76	208539_x_at	*SPRR2A, SPRR2B, SPRR2D*	−1.0258	−3.4094	4.88E−03	2.74E−04
77	208650_s_at	*CD24*	−1.6380	−1.0461	5.52E−03	9.70E−03
78	208712_at	*CCND1*	−1.7584	−1.3400	1.85E−09	1.70E−04
79	209118_s_at	*TUBA1A*	−1.1814	−1.6618	3.70E−03	8.97E−04
80	209126_x_at	*KRT6A, KRT6B*	−1.0109	−1.7619	7.83E−03	3.16E−03
81	209189_at	*FOS*	−1.2550	−1.7443	6.40E−03	1.23E−02
82	209242_at	*PEG3*	−1.1185	−1.7051	1.60E−04	4.77E−05
83	209250_at	*DEGS1*	−1.6475	−1.0716	1.68E−04	4.35E−03
84	209283_at	*CRYAB*	−1.4519	−2.9331	4.29E−07	1.48E−09
85	209291_at	*ID4*	−2.2617	−1.7203	2.32E−06	1.04E−03
86	209318_x_at	*PLAGL1*	−1.0631	−1.5734	2.80E−02	1.87E−03
87	209335_at	*DCN*	−1.6677	−2.5536	2.57E−03	3.04E−04
88	209540_at	*IGF1*	−1.1470	−2.0332	1.72E−02	5.42E−03
89	209541_at	*IGF1*	−1.4871	−2.5855	1.91E−03	2.83E−03
90	209550_at	*NDN*	−1.1674	−1.7010	4.80E−05	3.81E−04
91	209569_x_at	*NSG1*	−1.3910	−1.9295	8.88E−08	1.28E−04
92	209570_s_at	*NSG1*	−1.6985	−1.3952	1.95E−04	2.86E−05
93	209605_at	*TST*	−1.5512	−1.0193	1.42E−07	1.02E−02
94	209687_at	*CXCL12*	−1.6878	−3.5983	1.05E−03	1.57E−05
95	210020_x_at	*CALML3*	−1.2936	−1.5506	2.22E−03	2.33E−02
96	211423_s_at	*SC5D*	−1.0341	−1.1407	2.63E−04	2.14E−02
97	211548_s_at	*HPGD*	−2.2811	−2.9565	9.06E−06	1.77E−05
98	211549_s_at	*HPGD*	−1.5563	−1.6731	1.97E−05	2.01E−05
99	211597_s_at	*HOPX*	−3.4727	−3.8543	1.85E−09	6.16E−10
100	211748_x_at	*PTGDS*	−1.1759	−2.8281	6.75E−04	3.00E−05
101	211813_x_at	*DCN*	−1.0371	−2.4546	5.80E−03	8.88E−05
102	211896_s_at	*DCN*	−1.7222	−2.8559	4.39E−05	6.65E−04
103	212099_at	*RHOB*	−1.2852	−1.4439	1.59E−02	9.09E−03
104	212187_x_at	*PTGDS*	−1.0762	−2.8519	1.63E−03	2.66E−05
105	212230_at	*PLPP3*	−1.2267	−1.8056	4.59E−03	2.70E−03
106	212268_at	*SERPINB1*	−2.0739	−1.0263	1.49E−06	1.91E−02

(*Continued*)

Table S1 (*Continued*)

Number	Probe name	Gene symbol	logFC		Adjusted *P*-value	
			GSE7803	GSE9750	GSE7803	GSE9750
107	212593_s_at	*PDCD4, MIR4680*	−1.0203	−1.0249	1.71E−06	6.53E−04
108	213005_s_at	*KANK1*	−1.3237	−1.4490	2.32E−06	1.01E−04
109	213240_s_at	*KRT4*	−4.3954	−3.6606	1.22E−05	4.69E−03
110	213287_s_at	*KRT10*	−1.4603	−1.5958	6.51E−03	7.27E−04
111	213421_x_at	*PRSS3*	−1.3548	−1.8542	1.74E−05	2.02E−03
112	213680_at	*KRT6B*	−1.9134	−1.6133	1.08E−02	4.60E−02
113	213796_at	*SPRR1A*	−2.5895	−3.5348	1.67E−02	2.67E−03
114	213895_at	*EMP1*	−1.5169	−1.8942	7.08E−07	1.20E−06
115	214091_s_at	*GPX3*	−1.6787	−3.0400	6.10E−06	7.67E−08
116	214247_s_at	*DKK3*	−1.7619	−1.4292	4.50E−04	1.56E−02
117	214549_x_at	*SPRR1A*	−2.7653	−3.2505	1.35E−04	6.67E−04
118	214599_at	*IVL*	−2.6075	−2.2661	3.08E−05	4.70E−05
119	214621_at	*GYS2*	−1.6288	−1.5121	1.62E−04	9.55E−06
120	214624_at	*UPK1A*	−3.4453	−2.3695	1.95E−11	4.10E−09
121	214696_at	*MIR22, MIR22HG*	−1.2503	−1.0092	9.91E−04	6.55E−04
122	217845_x_at	*HIGD1A*	−1.1969	−1.1395	4.79E−04	1.22E−03
123	218002_s_at	*CXCL14*	−2.3615	−2.3685	5.25E−03	2.29E−03
124	218312_s_at	*ZSCAN18*	−1.5418	−1.8495	2.07E−06	4.65E−06
125	218502_s_at	*TRPS1*	−1.0972	−1.8370	2.27E−04	1.42E−06
126	218677_at	*S100A14*	−1.2332	−1.0900	4.46E−05	2.11E−03
127	218990_s_at	*SPRR3*	−3.8102	−4.0176	5.84E−05	5.09E−04
128	219090_at	*SLC24A3*	−1.4945	−1.8096	2.33E−05	3.51E−05
129	219267_at	*GLTP*	−1.9362	−1.2742	5.88E−07	2.51E−03
130	219304_s_at	*PDGFD*	−1.2399	−2.5374	1.71E−05	1.89E−09
131	219554_at	*RHCG*	−2.3009	−3.4518	1.07E−05	3.81E−05
132	219648_at	*MREG*	−1.1721	−1.0151	3.34E−06	2.70E−02
133	219836_at	*ZBED2*	−1.8152	−1.6920	4.01E−07	4.37E−04
134	219995_s_at	*ZNF750*	−1.3202	−1.4628	3.99E−03	6.24E−03
135	220026_at	*CLCA4*	−1.7727	−2.3287	2.21E−02	2.67E−03
136	220066_at	*NOD2*	−1.2609	−1.3621	1.70E−05	8.88E−05
137	220090_at	*CRNN*	−4.8078	−6.4682	1.02E−11	7.05E−15
138	220266_s_at	*KLF4*	−1.1657	−2.2531	4.46E−05	1.27E−04
139	220403_s_at	*TP53AIP1*	−1.2074	−1.2307	1.17E−03	1.28E−03
140	220431_at	*TMPRSS11E*	−1.1214	−2.9572	3.59E−04	5.14E−04
141	220620_at	*CRCT1*	−2.6450	−4.1521	1.60E−06	6.43E−05
142	220723_s_at	*CWH43*	−1.7748	−2.6534	4.41E−09	7.02E−07
143	221667_s_at	*HSPB8*	−1.6989	−1.8139	7.42E−06	5.09E−07
144	221841_s_at	*KLF4*	−2.1774	−2.2756	3.11E−05	2.90E−04
145	221896_s_at	*HIGD1A*	−1.1638	−1.1978	7.06E−04	9.97E−04
146	57588_at	*SLC24A3*	−1.1619	−1.5407	2.29E−05	1.32E−05

Abbreviation: FC, fold change.