Screening and Identification of Novel Biomarkers Associated with Cutaneous Squamous Cell Carcinoma

Wenxing Su
Second Affiliated Hospital of Soochow University
Biao Huang
First Affiliated Hospital of Soochow University
Wei Han
First Affiliated Hospital of Soochow University
Lu An
Second Affiliated Hospital of Soochow University
Yi Guan
Soochow University
Jiang JI (jijiang2222@126.com)
Second Affiliated Hospital of Soochow University
Daojiang Yu
the second affiliated hospital of chengdu medical college

Research

Keywords: cutaneous squamous cell carcinoma, bioinformatic analysis, differentially expressed genes, biomarker

Posted Date: September 28th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-80056/v1

Abstract

Background: Cutaneous squamous cell carcinoma (cSCC) is the leading cause of death in patients with non-melanoma skin cancers (NMSC). However, unclear pathogenesis of cSCC limits the application of molecular targeted therapy.

Results: To identify the hub genes in the pathogenesis and progression of cSCC, we downloaded the microarray data sets GSE2503, GSE45164 and GSE66359 from the Gene Expression Omnibus (GEO) database, and identified differentially expressed genes (DEGs) between tumor and non-tumor tissues. Functional enrichment analysis was performed using DAVID. The STRING online website was used to construct a protein-protein interaction network (PPI), and then Cytoscape performed module analysis and degree calculation. 146 DEGs were identified with significant differences, including 113 up-regulated genes and 33 down-regulated genes. The enriched functions and pathways of the DEGs include microtubule-based movement, ATP binding, cell cycle, p53 signaling pathway, oocyte meiosis and PLK1 signaling events. Nine hub genes were identified, namely CDK1, AURKA, RRM2, CENPE, CCNB1, KIAA0101, ZWINT, TOP2A, ASPM. The differential expression of these genes has been verified in other data sets. In addition, the ROC curve also confirmed their ability to predict disease.

Conclusion: By integrated bioinformatic analysis, the DEGs and hub genes identified in this study elucidated the molecular mechanism of the pathogenesis and progression of cSCC, and are expected to become future biomarkers or therapeutic targets.

Background

Cutaneous squamous cell carcinoma (cSCC) is a type of malignant tumor that originates in the epidermis or appendage keratinocytes, with an incidence only second to basal cell carcinoma (BCC), accounting for about 20% of all non-melanoma skin cancers (NMSC)[1, 2]. Recent studies have shown that somatic mutations in cSCC are much more frequent than in other squamous cell carcinomas (head and neck and lung squamous cell carcinomas)[3], indicating a complex genetic background of cSCC—its pathogenesis involves multiple genes and pathways. The most common altered tumor suppressor genes in cSCC are P53, CDKN2A, NOTCH1 and NOTCH2[4, 5]. It has been confirmed that 54–95% of cSCC contain UV radiation-induced P53 mutations[6]. The IHC score of P53 protein expression is closely related to the histological grade and TNM stage of cSCC. Tumors with high P53 protein expression are more aggressive than tumors with low P53 protein expression[7]. cSCC often finds heterozygous deletions or point mutations in the CDKN2A gene locus, and the deletion of p16INK4a is thought to be related to the progression of AK to cSCC[8]. Notch is a direct target of p53, and more than 75% of cSCCs have NOTCH1 and NOTCH2 mutations[9]. However, driving genes are not yet clear. It is thus significant to understand the exact molecular mechanisms of cSCC occurrence, proliferation, and recurrence in order to develop effective diagnostic and therapeutic strategies.

In recent years, gene expression analysis, such as DNA microarray analysis, has provided an effective and global way to elucidate gene expression changes in many cancer types, including skin cancer. However, it is difficult to obtain reliable results from an independent chip analysis for its false positive rate is high. In this study, therefore, three mRNA microarray datasets downloaded from the Gene Expression Omnibus (GEO) were analyzed to obtain differential expression profiles (DEGs) of cSCC and normal epidermis. In addition, functional enrichment analysis was performed on these DEGs, and a protein-protein interaction (PPI) network between them was constructed to reveal the molecular mechanisms of pathogenesis and progression. In summary, the DEGs and hub genes identified in this study elucidated the molecular mechanism of the pathogenesis and progression of cSCC, and are expected to become future biomarkers or therapeutic targets.

Methods

Raw data collection

GEO (http://www.ncbi.nlm.nih.gov/geo)[10] is a public repository containing microarray-based gene expression profiles. Three gene expression datasets (GSE2503[11], GSE45164[12] and GSE66359[13])were downloaded from it. Besides, the gene expression datasets GSE53462[14], GSE45216[15] and GSE42677[16] were downloaded to verify the data. Table 1 shows the details of the six datasets. The probes were converted into the corresponding gene symbol according to the annotation information on the platform. The GSE2503 dataset contains 5 cSCC tissue samples and 6 normal human epidermis samples. GSE45164 contains 10 cSCC samples and 3 non-cancerous samples. GSE66359 contains 8 cSCC samples and 5 normal human epidermal keratinocytes. The GSE53462 dataset contains 5 cSCC tissue samples and 5 normal human epidermis samples. There are 30 cSCC samples in the GSE45216 dataset and we selected 10 of them which share the same background and testing platform with those from the GSE42677 dataset as controls to draw the ROC curve.

Identification of DEGs

The DEGs between cSCC and non-cancerous samples were screened using GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r). GEO2R is an interactive web tool that allows users to compare two sets of data in the GEO series to determine the DEG between them[17]. The P-values and Benjamini and Hochberg false discovery rate were adopted to indicate statistical significance while controlling false-positives. Probe sets without corresponding gene symbols or genes with more than one probe set were removed or averaged, respectively. logFC (fold change) >1 and P-value < 0.05 were considered statistically significant.

Enrichment analyses of DEGs

DAVID is an online bioinformatics database that integrates biological data and analysis tools, providing users with a comprehensive set of gene and protein function annotation information for extracting biological information (https://david.ncifcrf.gov/, version 6.8)[18]. KEGG is an encyclopedia of genes and

genomes, used for biological interpretation of genome sequences and other high-throughput data. Gene Ontology (GO) is a recognized standardized classification system used to define the unique biological functions of genes and their RNA or protein products obtained from high-throughput genome or transcriptome analysis. The results of GO and KEGG enrichment analysis of DEGs were obtained through DAVID. P-value < 0.05 was considered statistically significant.

PPI network construction and module analysis

The STRING (http://string-db.org) (version 10.0) database is a database that calculates physical interactions and predicts protein (gene) relationships by collecting gene regulation relationships, protein interaction relationships, protein co-expression, etc[19]. In the present study, PPI network of DEGs was constructed based on STRING database, and interaction with a combined score > 0.4 was considered statistically significant. Cytoscape (version 3.6.1) is an open source bioinformatics software platform for visualizing molecular interaction networks[20]. The plug-in Molecular Complex Detection (MCODE) (version 1.4.2) of Cytoscape is an APP for clustering a given network based on the topology to find densely connected regions[21]. The PPI networks were drawn using Cytoscape and the most significant module in the PPI networks was identified using MCODE. The criteria for selection were as follows: MCODE scores >10, degree cut-off=2, node score cut-off=0.2, Max depth=100 and k-score=2. Subsequently, FunRich[22], an open-access software enabling functional enrichment analysis and interaction network analysis of genes and proteins, was used to analyze the biological pathways of the module.

Hub genes selection and analysis

Hub genes pathway analysis was performed and visualized by ClueGO (version 2.5.4) and CluePedia (version 1.5.4), the plug-in of Cytoscape. P-value < 0.05 was considered statistically significant. The hub genes were selected with degrees ≥ 30. A network of the genes and their co-expression genes was analyzed using GeneMANIA (http://www.genemania.org/)[23], which is a convenient web portal for analyzing gene lists and predicting gene function. The Drug-Gene Interaction database (DGIdb) mines existing resources and generates assumptions about how genes are therapeutically targeted or prioritized for drug development[24]. In this study, DGIdb 3.0 (http://www.dgidb.org/) was used to predict drugs based on the hub genes. The parameters were set as: preset filters: FDA approved; antineoplastic; all the default. All the drug-gene relationship pairs related to the module genes were predicted, and Cytoscape was used to construct the network map.

Validation of hub genes in other databases

Finally, Hub genes expression were verified in the dataset GSE53462 by Studen’s t text. P-value < 0.05 was considered statistically significant. To differentiate hub genes in terms of their abilities in recognizing cSCC, the ROC curve based on the comparison of GSE45216 and GSE42677 was drawn via the pROC package[25].

Results

Identification of DEGs in cSCC

After standardization of the microarray results, DEGs (1860 in GSE2503, 1649 in GSE45164 and 1990 in GSE66359) were identified. Their volcanic and heat maps are shown in Figure 1. The overlap among the 3 datasets contained 146 genes as shown in the Venn diagram (Fig. 2A), consisting of 113 upregulated genes and 33 downregulated genes between cSCC tissues and non-cancerous tissues.

Enrichment analyses of DEGs

Use the online tool DAVID for the enrichment analysis of GO and KEGG pathways to classify DEGs and the pathways involved. Through GO function enrichment analysis, 60 main terms were found enriched. The most enriched items in biological processes (BP) (31 items) and the least enriched items in molecular functions (MF) (6 items). GO analysis results show that: For BP, DEGs is significantly enriched in microtubule-based movement, negative regulation of cell growth and positive regulation of apoptotic process (Fig. 3A). Regarding the cell composition (CC), DEGs is mainly concentrated on nucleoplasm, extracellular exosome and cytoplasm (Fig. 3B). In terms of MF, DEGs mainly focuses on ATP binding, ATPase activity and microtubule motor activity (Fig. 3C). KEGG pathway analysis showed that the upregulated DEGs was mainly concentrated in cell cycle, p53 signaling pathway, oocyte meiosis and progesterone-mediated oocyte maturation (Fig. 3D).

PPI network construction and module analysis

The PPI network of DEGs is shown in Figure 2B, consisting of 109 nodes and 617 interaction pairs. The most significant module (score=28.857) was aggregated from the PPI network (Figure 2C), including 29 nodes and 404 interaction pairs. Then it was put into FunRich for further functional analysis of genes. All the genes in this module are up-regulated. Funrich analysis of enriched biological pathway for the module showed that the DEGs were mainly enriched in PLK1 signaling events and polo-like kinase signaling events in the cell cycle, as shown in Figure 4.

Hub gene selection and analysis

A total of 9 genes with degrees ≥ 30 was identified as hub genes. Names, abbreviations, degree and functions of these hub genes are listed in Table 1. The ClueGO revealed that the most involved pathways were p53 signaling pathway, TP53 regulates transcription of cell cycle genes and TP53 regulates transcription of genes involved in G2 cell cycle arrest (Figure 5A). A network of the hub genes and their co-expression genes was analyzed using GeneMANIA online platform. Those 9 genes showed the complex DEGs PPI network with the Co-expression of 72.69%, Predicted of 22.58%, Co-localization of 1.86%, Physical interactions of 1.73%, Pathway of 1.12% and Genetic interactions of 0.02% (Figure 5B). Based on the DGIdb predictions of the module genes, we obtained 30 drug-gene interaction pairs, including four upregulated genes (AURKA, RRM2, CENPE and TOP2A) and 29 drugs (FDA-listed + antitumor drugs), as shown in Figure 5C.

Validation of hub genes in other databases

Finally, the result of independence testing analysis suggested that all hub genes were significantly increased in cSCC tumor cells compared to normal tissues (Figure 6). ROC curve results indicated that AURKA, RRM2, CENPE, CCNB1, KIAA0101, ZWINT and TOP2A can identify CSCC with a reasonable degree of accuracy while CDK1 and ASPM not (Figure 7).

Discussion

The incidence of cSCC is increasing globally. Cutaneous squamous cell carcinoma is the leading cause of NMSC death, as invasive cSCC shows potential for recurrence and metastasis[26]. The most important risk factors for cSCC include UV exposure, older age, fair skin, and immunosuppression[27]. Previous reports have confirmed that P53 mutations, CDKN2A mutations, RAS mutations in cell signaling, and tumor suppressor NOTCH1 and NOTCH2 mutations are closely related to cSCC[6–9, 28], failing to fully unveil the underlying molecular mechanisms behind the aggressive progression of cSCC subpopulations. This might account for the high mortality rate of cSCC in NMSC[29]. Thus, both potential and efficient markers for diagnosis and therapy are in urgent need. Microarray technology, which was proven to be useful for identifying new biomarkers in other diseases enable us to explore the genetic alterations of cSCC.

In the current study, through analysis of a large sample of cSCC and corresponding normal tissues, 146 differential genes were identified, including 113 up-regulated gene and 33 down-regulated genes. The up-regulated DEGs are mainly enriched in the cell cycle, the P53 signaling pathway and oocyte meiosis. According to previous studies, dysregulation of the cell cycle process plays an important role in tumorigenesis or development[30] and the p53 signaling pathway is associated with expression profiling of cSCC during the process and epigenetic abnormalities[31, 32]. In addition, biological pathway for the most significant module showed that the DEGs were mainly enriched in PLK1 signaling events and polo-like kinase signaling events in the cell cycle. Previous studies have confirmed that by inhibiting cSCC keratinocyte PLK1 signaling in vitro, cancer cells can die first, emphasizing the importance of PLK1 signaling pathway in the occurrence and development of cSCC[33]. In this regard, our results are in consistence with all these theories.

A total of 9 genes were identified as hub genes with degrees ≥ 30, namely CDK1, AURKA, RRM2, CENPE, CCNB1, KIAA0101, ZWINT, TOP2A, ASPM. And these genes have been verified in the dataset GSE53462. Among these genes, there are 4 druggable genes, including AURKA, RRM2, CENPE and TOP2A. AURKA is one of three members of the highly conserved mitogen kinase family and plays an important role in regulating cell division, which is necessary for timely access to mitosis, centrosome maturation, and assembly of bipolar spindles[34]. Previous studies have found that the expression rates of AURKA in squamous cell carcinoma and adenocarcinoma are significantly different[35]. In addition, Enrique C et al. established a mouse model of AURKA overexpression, suggesting that AURKA has a clear role in the malignant progression of cSCC[36]. RRM2 encodes one of two non-identical subunits for ribonucleotide

reductase. The overexpression of RRM2 significantly enhances the invasiveness of the cells and plays a key role in determining the degree of tumor malignancy[37, 38]. However, the role RRM2 plays in the development of cSCC is unclear. The protein encoded by CENPE is a forward-directed kinesin, belonging to the kinesin-7 subfamily, which is critical in mitosis[39]. Increasing evidence shows that CENPE may be a useful drug target for several tumors without targeted therapy[40]. Recent studies have confirmed that CENPE is highly expressed in lung adenocarcinoma tissues and promotes lung adenocarcinoma cell proliferation[41]. TOP2A encodes DNA topoisomerase and is involved in important cellular functions such as DNA replication, transcription, recombination, and mitosis. It is a sign of cell proliferation in normal and tumor tissues. High expression of TOP2A occurs most often in breast cancer and affects the patient's disease-free survival period and total survival period, which is regarded as an important prognostic molecular index in the development process of breast cancer[42–44]. In addition, the high expression of TOP2A was related to the cell cycle, and targeting TOP2A was also considered to be an important method for treating human cancer[45]. We have reason to believe that these genes may have potential to play an important role in the occurrence and development of cSCC.

CDK1, a serine/threonine kinase, regulates cell cycle progression by binding to cyclin B to form a complex called cyclin B-CDK1. Studies have shown that CDK1 is overexpressed in breast cancer and liver cancer, causing tumor cell proliferation and development[46, 47]. Besides, CDK1 is a marker of clinical prognosis in colon cancer[48]. CCNB1 is a member of the cyclin family. CCNB1 and CDC2 combine to form an M-phase promoting factor (MPF), which promotes cells from G2 to M phase[49]. Overexpression of CCNB1 damages the cell's G2/M detection point and causes MPF increase. This detection point cannot detect whether the DNA is damaged, whereas the damaged DNA to continue to undergo mitosis; and it would also cause the proteasome to break down and recognize all MPF in the middle stage of division, resulting in continuous proliferation and development of tumor cells[50]. CCNB1 and squamous cells build on each other, allowing cancer cells to proliferate and differentiate, while new cancer cells promote the expression of CCNB1 to further increase[51]. Previous studies have reported that KIAA0101 overexpression in mammalian cells can prevent UV-induced apoptosis, suggesting its essential role in regulating DNA repair, cell proliferation, apoptosis and cell cycle progression[52]. KIAA0101 is closely related to the invasion and metastasis of cancer cells[53]. However, there is no direct evidence for a correlation between this gene and cSCC. ZWINT is a centromere complex component required for mitotic spindle checkpoints and is involved in centromere function and cell growth[54]. Recently, ZWINT overexpression has been reported in ovarian cancer and hepatocellular carcinoma, which is intimately linked to tumor progression and poor prognosis[55, 56]. ASPM, as a cell cycle progression gene, is a key in mitotic spindle regulation[57]. Previous studies have shown that ASPM is highly expressed in ovarian, pancreatic, and prostate cancers and is significantly associated with poor prognosis[58–60]. In addition, according to recent findings, knockout of TPX2 in prostate cancer can induce cell cycle quiescence and apoptosis, reduce the ability of cells to invade, and inhibit cell proliferation[61].

We would like to acknowledge the limitations of this study. First, this is a retrospective study and all the data of this study were obtained from publicly available databases. Second, further studies, including in

vivo and in vitro experiments, are required to confirm these results. Third, the underlying mechanisms of signaling pathways in cSCC has to be studied further.

In summary, the purpose of this study is to identify DEGs that may be related to the occurrence and development of cSCC. A total of 9 Hub genes were identified and could be used as diagnostic biomarkers or drug treatment targets for cSCC. Moreover, the high expression of hub genes in cSCC has been verified in other data sets, which makes the results more reliable. However, the biological function of these genes in cSCC needs further research.

Abbreviations

cSCC: Cutaneous squamous cell carcinoma; GEO: Gene Expression Omnibus; DEGs: differentially expressed genes; PPI: protein-protein interaction network; BCC: basal cell carcinoma; NMSC: non-melanoma skin cancers; FC: fold change; GO: Gene Ontology; MCODE: Molecular Complex Detection; BP: biological processes; MF: molecular functions; CC: cell composition; MPF: M-phase promoting factor.

Declarations

Acknowledgements

Not applicable.

Authors' contributions

The work presented here was carried out in collaboration among all authors. JJ and YDJ defined the research theme, discussed analysis, interpretation, and presentation. SWX, HB, HW drafted the manuscript, analyzed the data, developed the algorithm and interpreted the results. GY polished up the writing and refined the language. AL helped collect data and references. All authors read and approved the final manuscript.

Funding

This research was supported by the National Natural Science Foundation of China (81703157).

Availability of data and materials

All data generated or analyzed during this study are included in this article.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors consent to the publication of this study.

Competing interests

The authors declare no conflict of interests.

References

1. Karia PS, Han J, Schmults CD: **Cutaneous squamous cell carcinoma: estimated incidence of disease, nodal metastasis, and deaths from disease in the United States, 2012**. *J Am Acad Dermatol* 2013, **68**(6):957-966.
2. Rogers HW, Weinstock MA, Feldman SR, Coldiron BM: **Incidence Estimate of Nonmelanoma Skin Cancer (Keratinocyte Carcinomas) in the U.S. Population, 2012**. *JAMA dermatology* 2015, **151**(10):1081-1086.
3. Pickering CR, Zhou JH, Lee JJ, Drummond JA, Peng SA, Saade RE, Tsai KY, Curry JL, Tetzlaff MT, Lai SY *et al*: **Mutational landscape of aggressive cutaneous squamous cell carcinoma**. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2014, **20**(24):6582-6592.
4. South AP, Purdie KJ, Watt SA, Haldenby S, den Breems N, Dimon M, Arron ST, Kluk MJ, Aster JC, McHugh A *et al*: **NOTCH1 mutations occur early during cutaneous squamous cell carcinogenesis**. *J Invest Dermatol* 2014, **134**(10):2630-2638.
5. Inman GJ, Wang J, Nagano A, Alexandrov LB, Purdie KJ, Taylor RG, Sherwood V, Thomson J, Hogan S, Spender LC *et al*: **The genomic landscape of cutaneous SCC reveals drivers and a novel azathioprine associated mutational signature**. *Nature communications* 2018, **9**(1):3667.
6. Boukamp P: **Non-melanoma skin cancer: what drives tumor development and progression?** *Carcinogenesis* 2005, **26**(10):1657-1667.
7. Bukhari MH, Niazi S, Chaudhry NA: **Relationship of immunohistochemistry scores of altered p53 protein expression in relation to patient's habits and histological grades and stages of squamous cell carcinoma**. *Journal of cutaneous pathology* 2009, **36**(3):342-349.
8. Gray SE, Kay E, Leader M, Mabruk M: **Analysis of p16 expression and allelic imbalance / loss of heterozygosity of 9p21 in cutaneous squamous cell carcinomas**. *Journal of cellular and molecular medicine* 2006, **10**(3):778-788.
9. Okuyama R, Tagami H, Aiba S: **Notch signaling: its role in epidermal homeostasis and in the pathogenesis of skin diseases**. *J Dermatol Sci* 2008, **49**(3):187-194.
10. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Marshall KA *et al*: **NCBI GEO: archive for high-throughput functional genomic data**. *Nucleic acids research* 2009, **37**(Database issue):D885-890.
11. Padilla RS, Sebastian S, Jiang Z, Nindl I, Larson R: **Gene expression patterns of normal human skin, actinic keratosis, and squamous cell carcinoma: a spectrum of disease progression**. *Archives of*

dermatology 2010, **146**(3):288-293.

12. Brooks YS, Ostano P, Jo SH, Dai J, Getsios S, Dziunycz P, Hofbauer GF, Cerveny K, Chiorino G, Lefort K *et al*: **Multifactorial ERbeta and NOTCH1 control of squamous differentiation and cancer**. *The Journal of clinical investigation* 2014, **124**(5):2260-2276.
13. Farshchian M, Nissinen L, Siljamaki E, Riihila P, Toriseva M, Kivisaari A, Ala-Aho R, Kallajoki M, Verajankorva E, Honkanen HK *et al*: **EphB2 Promotes Progression of Cutaneous Squamous Cell Carcinoma**. *J Invest Dermatol* 2015, **135**(7):1882-1892.
14. Jee BA, Lim H, Kwon SM, Jo Y, Park MC, Lee IJ, Woo HG: **Molecular classification of basal cell carcinoma of skin by gene expression profiling**. *Molecular carcinogenesis* 2015, **54**(12):1605-1612.
15. Lambert SR, Mladkova N, Gulati A, Hamoudi R, Purdie K, Cerio R, Leigh I, Proby C, Harwood CA: **Key differences identified between actinic keratosis and cutaneous squamous cell carcinoma by transcriptome profiling**. *British journal of cancer* 2014, **110**(2):520-529.
16. Mitsui H, Suarez-Farinas M, Gulati N, Shah KR, Cannizzaro MV, Coats I, Felsen D, Krueger JG, Carucci JA: **Gene expression profiling of the leading edge of cutaneous squamous cell carcinoma: IL-24-driven MMP-7**. *J Invest Dermatol* 2014, **134**(5):1418-1427.
17. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M *et al*: **NCBI GEO: archive for functional genomics data sets--update**. *Nucleic acids research* 2013, **41**(Database issue):D991-995.
18. Huang da W, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources**. *Nature protocols* 2009, **4**(1):44-57.
19. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C *et al*: **STRING v9.1: protein-protein interaction networks, with increased coverage and integration**. *Nucleic acids research* 2013, **41**(Database issue):D808-815.
20. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T: **Cytoscape 2.8: new features for data integration and network visualization**. *Bioinformatics (Oxford, England)* 2011, **27**(3):431-432.
21. Bandettini WP, Kellman P, Mancini C, Booker OJ, Vasu S, Leung SW, Wilson JR, Shanbhag SM, Chen MY, Arai AE: **MultiContrast Delayed Enhancement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study**. *Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance* 2012, **14**:83.
22. Pathan M, Keerthikumar S, Ang C-S, Gangoda L, Quek CYJ, Williamson NA, Mouradov D, Sieber OM, Simpson RJ, Salim A *et al*: **FunRich: An open access standalone functional enrichment and interaction network analysis tool**. *Proteomics* 2015, **15**(15):2597-2601.
23. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes CT *et al*: **The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function**. *Nucleic acids research* 2010, **38**(Web Server issue):W214-220.

24. Cotto K, Wagner A, Feng Y-Y, Kiwala S, Coffman A, Spies G, Wollam A, Spies N, Griffith O, Griffith M: **DGIdb 3.0: a redesign and expansion of the drug-gene interaction database**. In.: bioRxiv; 2017.
25. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M: **pROC: an open-source package for R and S+ to analyze and compare ROC curves**. *BMC bioinformatics* 2011, **12**:77.
26. Weinberg AS, Ogle CA, Shim EK: **Metastatic cutaneous squamous cell carcinoma: an update**. *Dermatologic surgery : official publication for American Society for Dermatologic Surgery [et al]* 2007, **33**(8):885-899.
27. Green AC, Olsen CM: **Cutaneous squamous cell carcinoma: an epidemiological review**. *Br J Dermatol* 2017, **177**(2):373-381.
28. Li L, Li M, Xu S, Bu W, Zhang M, Gu H, Chen X: **Is Ras a potential target in treatment against cutaneous squamous cell carcinoma?** *Journal of Cancer* 2018, **9**(18):3373-3381.
29. Kivisaari A, Kahari VM: **Squamous cell carcinoma of the skin: Emerging need for novel biomarkers**. *World journal of clinical oncology* 2013, **4**(4):85-90.
30. Tripathi V, Shen Z, Chakraborty A, Giri S, Freier SM, Wu X, Zhang Y, Gorospe M, Prasanth SG, Lal A *et al*: **Long noncoding RNA MALAT1 controls cell cycle progression by regulating the expression of oncogenic transcription factor B-MYB**. *PLoS genetics* 2013, **9**(3):e1003368.
31. Warren TA, Broit N, Simmons JL, Pierce CJ, Chawla S, Lambie DL, Quagliotto G, Brown IS, Parsons PG, Panizza BJ *et al*: **Expression profiling of cutaneous squamous cell carcinoma with perineural invasion implicates the p53 pathway in the process**. *Scientific reports* 2016, **6**:34081.
32. Chen HQ, Gao D: **Inhibitory effect of microRNA-154 targeting WHSC1 on cell proliferation of human skin squamous cell carcinoma through mediating the P53 signaling pathway**. *The international journal of biochemistry & cell biology* 2018, **100**:22-29.
33. Watt SA, Pourreyron C, Purdie K, Hogan C, Cole CL, Foster N, Pratt N, Bourdon JC, Appleyard V, Murray K *et al*: **Integrative mRNA profiling comparing cultured primary cells with clinical samples reveals PLK1 and C20orf20 as therapeutic targets in cutaneous squamous cell carcinoma**. *Oncogene* 2011, **30**(46):4666-4677.
34. Vader G, Lens SM: **The Aurora kinase family in cell division and cancer**. *Biochimica et biophysica acta* 2008, **1786**(1):60-72.
35. Twu NF, Yuan CC, Yen MS, Lai CR, Chao KC, Wang PH, Wu HH, Chen YJ: **Expression of Aurora kinase A and B in normal and malignant cervical tissue: high Aurora A kinase expression in squamous cervical cancer**. *European journal of obstetrics, gynecology, and reproductive biology* 2009, **142**(1):57-63.
36. Torchia EC, Chen Y, Sheng H, Katayama H, Fitzpatrick J, Brinkley WR, Caulin C, Sen S, Roop DR: **A genetic variant of Aurora kinase A promotes genomic instability leading to highly malignant skin tumors**. *Cancer research* 2009, **69**(18):7207-7215.
37. Duxbury MS, Whang EE: **RRM2 induces NF-kappaB-dependent MMP-9 activation and enhances cellular invasiveness**. *Biochemical and biophysical research communications* 2007, **354**(1):190-196.

38. Zhang K, Hu S, Wu J, Chen L, Lu J, Wang X, Liu X, Zhou B, Yen Y: **Overexpression of RRM2 decreases thrombspondin-1 and increases VEGF production in human cancer cells in vitro and in vivo: implication of RRM2 in angiogenesis**. *Molecular cancer* 2009, **8**:11.
39. Hou S, Li N, Zhang Q, Li H, Wei X, Hao T, Li Y, Azam S, Liu C, Cheng W *et al*: **XAB2 functions in mitotic cell cycle progression via transcriptional regulation of CENPE**. *Cell death & disease* 2016, **7**(10):e2409.
40. El-Arabey AA, Salama SA, Abd-Allah AR: **CENP-E as a target for cancer therapy: Where are we now?** *Life sciences* 2018, **208**:192-200.
41. Shan L, Zhao M, Lu Y, Ning H, Yang S, Song Y, Shi X, Chai W: **[Corrigendum] CENPE promotes lung adenocarcinoma proliferation and is directly regulated by FOXM1**. *International journal of oncology* 2019, **55**(6):1397.
42. Tokiniwa H, Horiguchi J, Takata D, Kikuchi M, Rokutanda N, Nagaoka R, Sato A, Odawara H, Tozuka K, Oyama T *et al*: **Topoisomerase II alpha expression and the Ki-67 labeling index correlate with prognostic factors in estrogen receptor-positive and human epidermal growth factor type-2-negative breast cancer**. *Breast cancer (Tokyo, Japan)* 2012, **19**(4):309-314.
43. Olszewski W, Pienkowski T, Olszewski W, Mrozkowiak A, Bauer-Kosinska B, Piascik A, Olszewska K, Wojnowska A, Michalski W: **Topoisomerase 2alpha status in invasive breast carcinoma - comparison of its clinical value according to immunohistochemical and fluorescence in situ hybridization methods of evaluation**. *Polish journal of pathology : official journal of the Polish Society of Pathologists* 2014, **65**(4):283-290.
44. Jun KY, Park SE, Liang JL, Jahng Y, Kwon Y: **Benzo[b]tryptanthrin inhibits MDR1, topoisomerase activity, and reverses adriamycin resistance in breast cancer cells**. *ChemMedChem* 2015, **10**(5):827-835.
45. Delgado JL, Hsieh CM, Chan NL, Hiasa H: **Topoisomerases as anticancer targets**. *The Biochemical journal* 2018, **475**(2):373-398.
46. Reese JM, Bruinsma ES, Monroe DG, Negron V, Suman VJ, Ingle JN, Goetz MP, Hawse JR: **ERbeta inhibits cyclin dependent kinases 1 and 7 in triple negative breast cancer**. *Oncotarget* 2017, **8**(57):96506-96521.
47. Zhang Y, Huang W, Ran Y, Xiong Y, Zhong Z, Fan X, Wang Z, Ye Q: **miR-582-5p inhibits proliferation of hepatocellular carcinoma by targeting CDK1 and AKT3**. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* 2015, **36**(11):8309-8316.
48. Sung WW, Lin YM, Wu PR, Yen HH, Lai HW, Su TC, Huang RH, Wen CK, Chen CY, Chen CJ *et al*: **High nuclear/cytoplasmic ratio of Cdk1 expression predicts poor prognosis in colorectal cancer patients**. *BMC cancer* 2014, **14**:951.
49. Chevaleyre C, Benhamouda N, Favry E, Fabre E, Mhoumadi A, Nozach H, Marcon E, Cosler G, Vinatier E, Oudard S *et al*: **The Tumor Antigen Cyclin B1 Hosts Multiple CD4 T Cell Epitopes Differently Recognized by Pre-Existing Naive and Memory Cells in Both Healthy and Cancer Donors**. *Journal of immunology (Baltimore, Md : 1950)* 2015, **195**(4):1891-1901.

50. Lee MH, Cho Y, Jung BC, Kim SH, Kang YW, Pan CH, Rhee KJ, Kim YS: **Parkin induces G2/M cell cycle arrest in TNF-alpha-treated HeLa cells**. *Biochemical and biophysical research communications* 2015, **464**(1):63-69.
51. Patil GB, Hallikeri KS, Balappanavar AY, Hongal SG, Sanjaya PR, Sagari SG: **Cyclin B1 overexpression in conventional oral squamous cell carcinoma and verrucous carcinoma- A correlation with clinicopathological features**. *Medicina oral, patologia oral y cirugia bucal* 2013, **18**(4):e585-590.
52. Simpson F, Lammerts van Bueren K, Butterfield N, Bennetts JS, Bowles J, Adolphe C, Simms LA, Young J, Walsh MD, Leggett B *et al*: **The PCNA-associated factor KIAA0101/p15(PAF) binds the potential tumor suppressor product p33ING1b**. *Experimental cell research* 2006, **312**(1):73-85.
53. Jain M, Zhang L, Patterson EE, Kebebew E: **KIAA0101 is overexpressed, and promotes growth and invasion in adrenal cancer**. *PloS one* 2011, **6**(11):e26866.
54. Starr DA, Saffery R, Li Z, Simpson AE, Choo KH, Yen TJ, Goldberg ML: **HZwint-1, a novel human kinetochore component that interacts with HZW10**. *Journal of cell science* 2000, **113 (Pt 11)**:1939-1950.
55. Xu Z, Zhou Y, Cao Y, Dinh TL, Wan J, Zhao M: **Identification of candidate biomarkers and analysis of prognostic values in ovarian cancer by integrated bioinformatics analysis**. *Medical oncology (Northwood, London, England)* 2016, **33**(11):130.
56. Ying H, Xu Z, Chen M, Zhou S, Liang X, Cai X: **Overexpression of Zwint predicts poor prognosis and promotes the proliferation of hepatocellular carcinoma by regulating cell-cycle-related proteins**. *OncoTargets and therapy* 2018, **11**:689-702.
57. Higgins J, Midgley C, Bergh AM, Bell SM, Askham JM, Roberts E, Binns RK, Sharif SM, Bennett C, Glover DM *et al*: **Human ASPM participates in spindle organisation, spindle orientation and cytokinesis**. *BMC cell biology* 2010, **11**:85.
58. Alsiary R, Bruning-Richardson A, Bond J, Morrison EE, Wilkinson N, Bell SM: **Deregulation of microcephalin and ASPM expression are correlated with epithelial ovarian cancer progression**. *PloS one* 2014, **9**(5):e97059.
59. Wang WY, Hsu CC, Wang TY, Li CR, Hou YC, Chu JM, Lee CT, Liu MS, Su JJ, Jian KY *et al*: **A gene expression signature of epithelial tubulogenesis and a role for ASPM in pancreatic tumor progression**. *Gastroenterology* 2013, **145**(5):1110-1120.
60. Xie JJ, Zhuo YJ, Zheng Y, Mo RJ, Liu ZZ, Li BW, Cai ZD, Zhu XJ, Liang YX, He HC *et al*: **High expression of ASPM correlates with tumor progression and predicts poor outcome in patients with prostate cancer**. *International urology and nephrology* 2017, **49**(5):817-823.
61. Pai VC, Hsu CC, Chan TS, Liao WY, Chuu CP, Chen WY, Li CR, Lin CY, Huang SP, Chen LT *et al*: **ASPM promotes prostate cancer stemness and progression by augmenting Wnt-Dvl-3-beta-catenin signaling**. *Oncogene* 2019, **38**(8):1340-1353.

Tables

Table 1.

Details of the GEO datasets. Cutaneous squamous cell carcinoma (cSCC), health control (HC).

Dataset	Platform	No. of Samples (cSCC vs HC)
GSE2503	GPL96[HG-U133A] Affymetrix Human Genome U133A Array	5, 6
GSE45164	GPL571 [HG-U133A_2] Affymetrix Human Genome U133A 2.0 Array	10, 3
GSE66359	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	8, 5
GSE53462	GPL10558 Illumina HumanHT-12 V4.0 expression beadchip	5, 5
GSE45216	GPL570[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	30, 0
GSE42677	GPL571[HG-U133A_2] Affymetrix Human Genome U133A 2.0 Array	0, 10

Table 2.

Functional roles of 9 hub genes with degree ≥ 30.

Gene symbol	Degree	Full name	Function
CDK1	35	cyclin dependent kinase 1	CDK1 can regulate the cell cycle progression, apoptosis and carcinogenesis of tumor cells
AURKA	32	aurora kinase A	AURKA is one of three members of the highly conserved mitogen kinase family and plays an important role in regulating cell division.
RRM2	32	ribonucleotide reductase regulatory subunit M2	RRM2 can enhanced the invasiveness of the cells and played a key role in determining the degree of tumor malignancy.
CENPE	31	centromere protein E	CENPE may be a useful drug target for several tumors without targeted therapy.
CCNB1	31	cyclin B1	CCNB1 and squamous cells play a complementary role, allowing cancer cells to further proliferate and differentiate.
KIAA0101	30	KIAA0101	KIAA0101 is closely related to the invasion and metastasis of cancer cells.
ZWINT	30	ZW10 interacting kinetochore protein	ZWINT is a centromere complex component required for mitotic spindle checkpoints and is involved in centromere function and cell growth.
TOP2A	30	topoisomerase (DNA) II alpha	This gene encodes a DNA topoisomerase, an enzyme that controls and alters the topologic states of DNA during transcription. TOP2A acts as a target for several anticancer agents and mutations of this gene have been associated with drug resistance
ASPM	30	abnormal spindle microtubule assembly	ASPM, as a cell cycle progression gene, plays a critical role in mitotic spindle regulation.

Figures

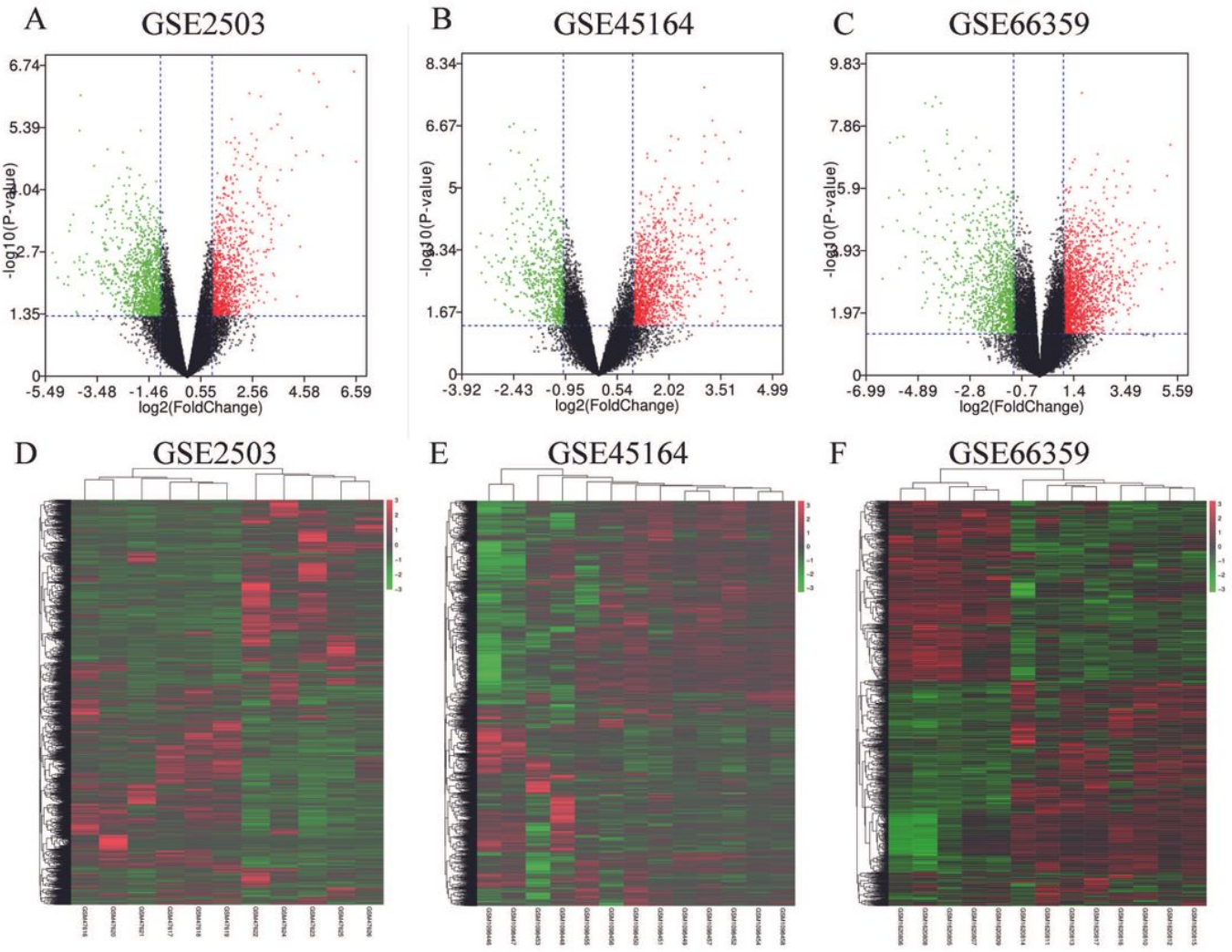

Figure 1

Differentially expressed analysis of GSE2503, GSE45164 and GSE66359 datasets. A-C was the differentially expressed volcano figure. D-F were heat maps of DEGs. Among them, red indicates up-regulated genes and green indicates down-regulated genes. Black indicates no differential expression.

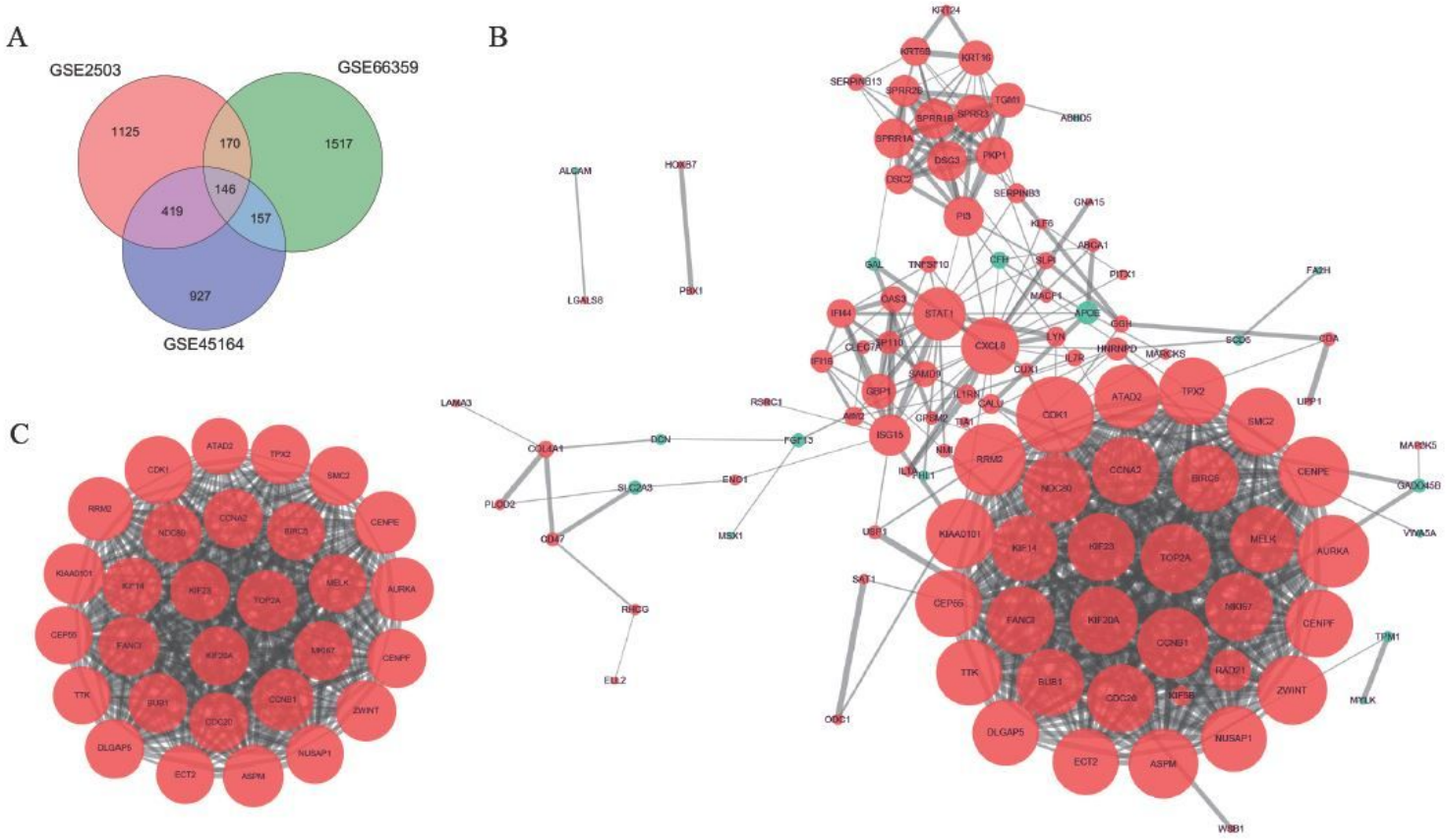

Figure 2

Venn diagram, PPI network and the most significant module of DEGs. (A) DEGs were selected with a fold change >1 and P-value < 0.05 among the mRNA expression profiling sets GSE2503, GSE45164 and GSE66359. The 3 datasets showed an overlap of 146 genes. (B) The PPI network of DEGs was constructed using Cytoscape. (C) The most significant module was obtained from PPI network with 29 nodes and 404 edges. Upregulated genes are marked in light red; downregulated genes are marked in light green. The node size is based on the degree value.

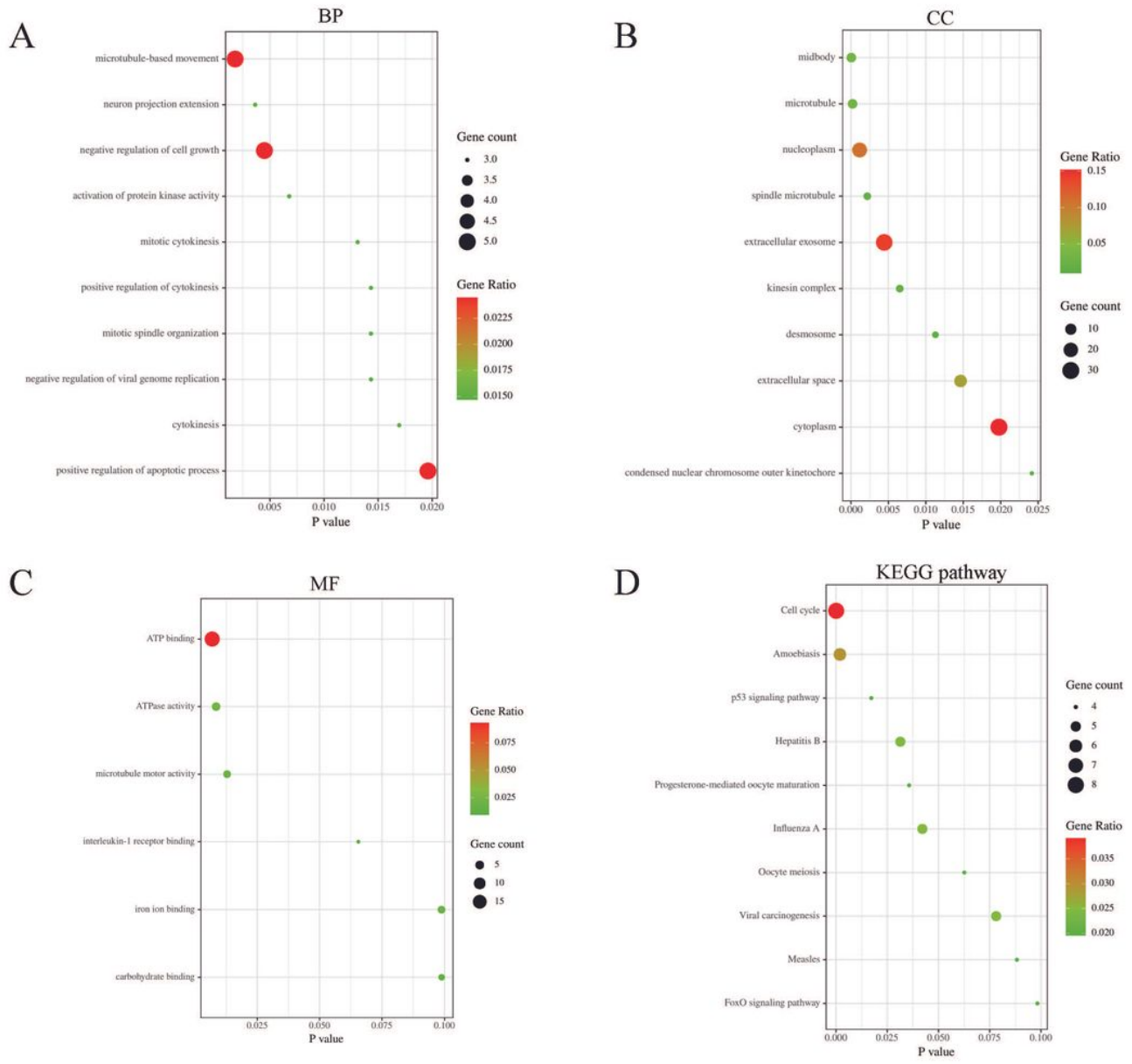

Figure 3

GO and KEGG Enrichment Analysis of the DEGs. (A) The top 10 enriched GO categories of biological process (BP). (B) The top 10 enriched GO categories of cellular component (CC). (C) The top 6 enriched GO categories of molecular function (MF). (D) A total of 10 signaling pathways in the KEGG enrichment. The abscissa represents the P-value, and the ordinate represents the terms. The size of the circle represents the number of genes involved, and the color represents the frequency of the genes involved in the term total genes.

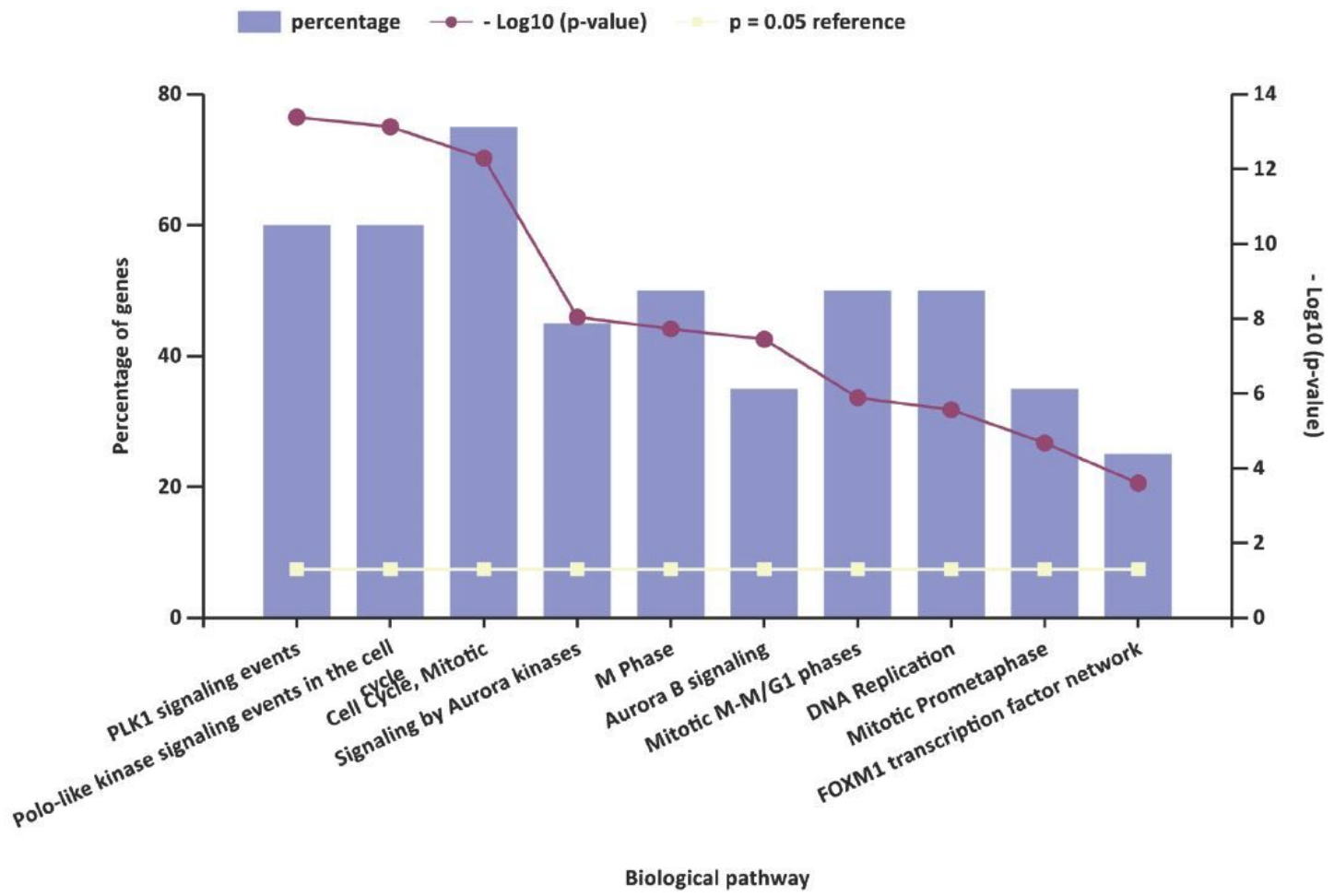

Figure 4

The Funrich software drew a bar chart of five biological pathways based on the P-value and the percentage of genes, among which biological pathways with P-value < 0.05 are statistically significant.

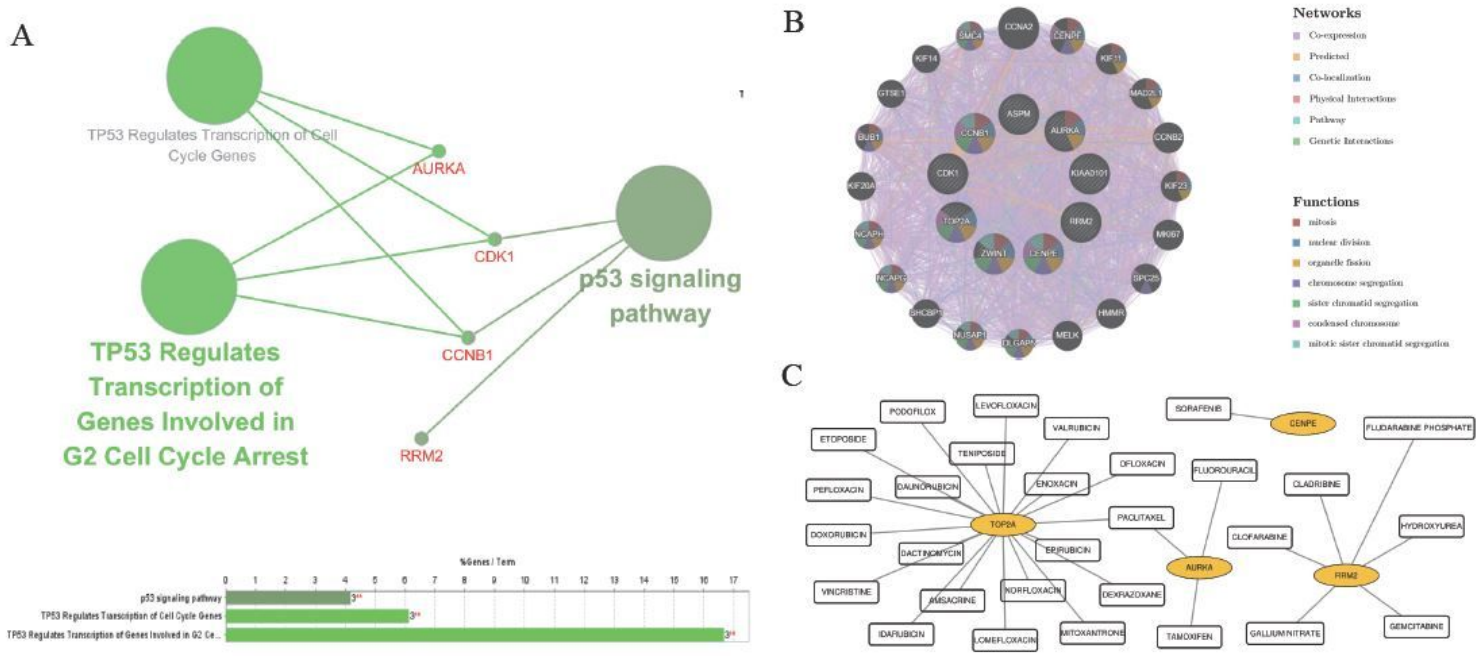

Figure 5

KEGG pathway, co-expression network and drug-gene interaction analysis of the hub genes. (A) The most significant pathway and related genes. The results show that these hub genes are mainly involved in p53 signaling pathway, TP53 regulates transcription of cell cycle genes and TP53 regulates transcription of genes involved in G2 cell cycle arrest. (B) Hub genes and their co-expression genes were analyzed using Gene MANIA. (C) Drug-gene interaction diagram, yellow circle indicates the differentially expressed gene and blank square indicates the drug.

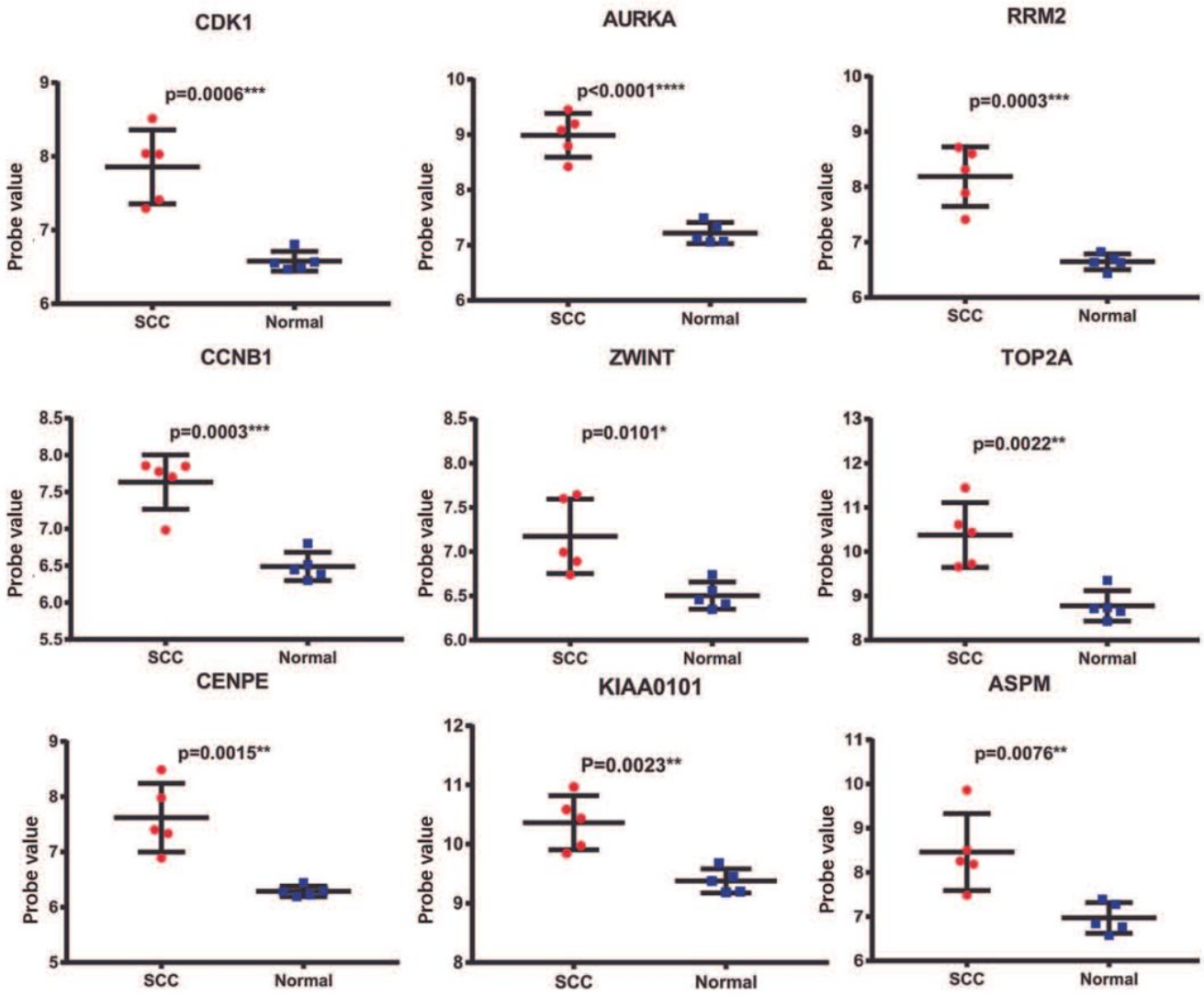

Figure 6

Hub genes expression in the GSE53462 dataset. SCC stands for cutaneous squamous cell carcinoma tumor tissue and Normal represented corresponding normal tissue. P-value < 0.05 was considered statistically significant.

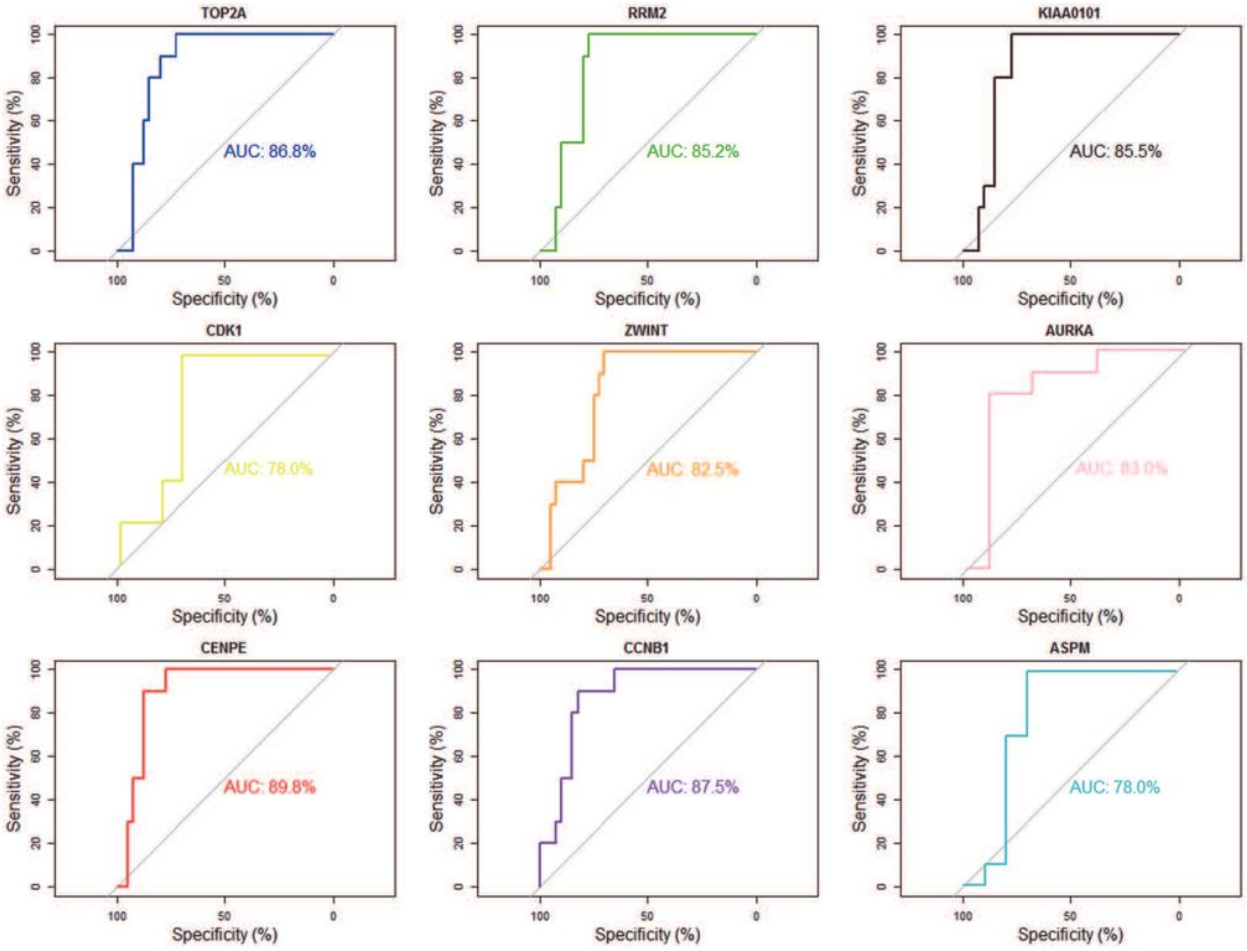

Figure 7

ROC curve of hub genes. Area under the curve (AUC) varies from 0.5 to 1.0. AUC of 0.5 is the chance diagonal, which fails to tell the discrimination between patients with and without the condition. The closer the AUC gets to 1, the more accurate the diagnostic test is (0.5~0.7: low-moderate accuracy, 0.7~0.9: moderate accuracy, 0.9~1: high accuracy).