Identification of core genes and pathways in melanoma metastasis via bioinformatics analysis

Yumei Li
Fuzhou University
Bifei Li
Fuzhou University
Fan Chen
Fuzhou University
Weiyu Shen
Fuzhou University
Vladimir L. Katanaev
"University of gevena
lee jia (cmapcjia1234@163.com)
Fuzhou University https://orcid.org/0000-0001-6839-5545

Primary research

Keywords: melanoma metastasis, bioinformatics analysis, differentially expressed genes, hub gene

Posted Date: July 6th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-39969/v1

Abstract

Background Metastasis is the leading cause of melanoma mortality. Current therapies are rarely curative for metastatic melanoma, revealing the urgent need to identify more effective preventive and therapeutic targets. This study aimed to screen for the key core genes and molecular mechanisms related to the metastasis of melanoma. Methods Gene expression profile, GSE8401 including 31 primary melanoma and 52 metastatic melanoma clinical samples, was downloaded from the Gene Expression Omnibus (GEO). Differentially expressed genes (DEGs) between metastatic melanoma and primary melanoma were screened using GEO2R. Assays of gene ontology (GO), Kyoto Encyclopedia of Gene and Genome (KEGG) pathway and protein-protein interaction (PPI) were performed to visualize these DEGs through Database for Annotation, Visualization and Integrated Discovery (DAVID) software and Search Tool for the Retrieval of Interacting Genes (STRING) and Cytoscape with Molecular Complex Detection (MCODE) plug-in tools. Top 10 genes with high degree were defined as hub genes. Furthermore, paired post-metastatic melanoma cells and pre-metastatic melanoma cells were established by experimental mouse model of melanoma metastasis to verify the expression of these hub genes. Results 424 DEGs between the metastatic melanoma and primary melanoma were screened, including 60 upregulated genes enriched in ECM-receptor interaction and progesterone-mediated oocyte maturation and 364 downregulated genes enriched in amoebiasis, melanogenesis, and ECM-receptor interaction. CDH1, EGFR, KRT5, COL17A1, KRT14, IVL, DSP, DSG1, FLG and CDK1 were defined as the hub genes. . In addition, paired post-metastatic melanoma cells (A375M) and pre-metastatic melanoma cells (A375) were established and qRT-PCR analysis confirmed the expression of the hub genes during melanoma metastasis. Conclusion This bioinformatic study has provided a deeper understanding of the molecular mechanisms of melanoma metastasis. KRT5, IVL and COL17A1 have emerged as possible biomarkers and therapeutic targets in metastasis of melanoma.

Background

Grimly, as the most commonly encountered malignant skin cancer, melanoma incidence has been steadily increasing in the last 40 years, and its incidence is ascending faster than that of any other solid tumor[1]. More seriously, because of melanoma’s extraordinary predisposition to spread and its rapid progression towards metastasis, patients who develop a metastasis are almost always incurable, with a median survival time of only 6–9 months, a 15% 3-year survival rate, and a 5-year survival rate of only 4.6%[2].

Metastasis is the most serious event in the clinic, leading to the majority of deaths of melanoma patients[3]. Despite the landscape of genetic alterations and elaborate molecular mechanisms that have been discovered in melanoma, less has been elucidated about the underlying biology that drives its metastasis[4]. Achieving a more in-depth understanding of the metastatic melanoma process is urgent to develop specific therapies for improving current therapy and reducing patient mortality.

Historically, several genes have been identified and were expected to be targets for preventing melanoma metastasis. For instance, PGCLα, a metabolic transcriptional coactivator, suppresses melanoma metastasis through protection against oxidative stress[5]. KiSS-1 overexpression was found to inhibit the metastasis of C8161 melanoma cells, and Survivin can promote melanoma metastasis through Akt-dependent upregulation of α5 integrin[6, 7]. However, these findings were hardly sufficient to develop a complete overview of melanoma metastasis. Furthermore, the peculiarity of the high heterogeneity and clinical phenotype of melanoma cells implies a complicated regulatory mechanism of metastasis[8].

Microarray-based gene expression analysis is a powerful and promising tool for functional genomics research, especially the accurate and comprehensive analysis of complex networks involved in biological processes[9, 10]. In our present study, we chose the GSE8401 profile from the Gene Expression Omnibus (GEO) and utilized the GEO2R tool to excavate the differentially expressed genes (DEGs). Subsequently, gene ontology (GO) and biological function annotation were performed, and a PPI network was constructed. We identified the hub genes and key pathways associated with melanoma metastasis that might provide new insight into clinical melanoma treatment. This bioinformatics method will support advances in the exploration of cancer metastasis.

Materials And Methods

Microarray data

The gene expression profile of GSE8401 (GDS3966), based on the Agilent GPL96 platform ([HG-U133A] Affymetrix Human Genome U133A Array), was obtained from the Gene Expression Omnibus (GEO), a free and publicly available database[11]. The GSE8401 dataset includes 83 clinical samples, 31 primary melanomas and 52 melanoma metastases. The heat map of the full expression range of the genes was acquired by the GEO online data analysis tool (https://www.ncbi.nlm.nih.gov/sites/GDSbrowser).

Identification of the differentially expression genes (DEGs) and hub genes

GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/), an interactive web tool, allows users to compare different groups of cancer clinic samples in a GEO series in order to identify genes that were differentially expressed across different experimental conditions[12]. Herein, we applied this tool to detect the DEGs between primary melanomas and melanoma metastases in GSE8410. The results are exhibited as a table of genes ordered by significance. The adjusted *p*-value was utilized to reduce the false positive rate, applying the Benjamini and Hochberg false discovery rate method by default. A total of 424 DEGs were found, and the top 10 genes with a high degree of connectivity were defined as the hub genes.

Gene ontology and KEGG pathway analysis of DEGs

Gene ontology analysis (GO) is a commonly used productive method for annotating genes and gene products and for identifying the biological characteristics of high-throughput genome or transcriptome

data[13]. The Kyoto Encyclopedia of Genes and Genomes (KEGG) is an open and collective database integrating genomes, biological pathways, diseases, drugs, and chemical substances[14]. Genes within the cut-off criteria of adjusted *P* value < 0.05 and |logFC| ≥ 1.5 were designated as DEGs and subjected to Gene ontology and KEGG pathway analysis. The Database for Annotation, Visualization and Integrated Discovery (DAVID, https://david.ncifcrf.gov/), an online bioinformatics tool, was applied to interpret the GO function and KEGG pathway enrichment and visualize the biological processes (BP), molecular functions (MF), cellular components (CC) and pathways of these DEGs[15]. A *P* value < 0.05 was set as the cut-off criterion.

Construction of the PPI network and module analysis

The Search Tool for the Retrieval of Interacting Genes (STRING) is an online tool that assesses protein-protein interaction (PPI) information[16]. The STRING (version 10.5) was used to evaluate the potential PPI relationships among the DEGs. Only experimentally validated interactions with a combined score ≥ 0.4 were selected as significant. The PPI network was constructed and visualized using the Cytoscape software 3.6.0. A plug-in molecular complex detection (MCODE) in Cytoscape was used to screen the modules of the PPI network. Modules inferred used the default settings with the degree cutoff = 2, node score cutoff = 0.2, K-core = 2, and max. Depth = 100. Additionally, the hub genes were mapped into the STRING with confidence score ≥ 0.4 and the maximum number of interactors ≤ 5. The KEGG pathway analysis of the genes in each module was realized by DAVID.

Cell, cell culture and postmetastatic cell line establishment

The human melanoma A375 cell line was purchased from the Cell Resource Center of Shanghai Institute for Biological Sciences (Chinese Academy of Sciences, China) and used as the premetastatic parental cell line. The postmetastatic melanoma cell line, A375M, was derived from the pulmonary metastatic nodules of pre-metastasized parental A375 cells via trypsinization as described previously[17]. In short, A375 cells (3×10^5 cells in 0.1 ml saline solution per mouse) were intravenously injected into the tails of 6–8 weeks old BALB/C nude mice (Slac Animal Inc., Shanghai, China). After 7 weeks, all of the mice were sacrificed, and the pulmonary metastatic nodules were stripped to obtain the monoplasts via trypsinization. Subsequently, these monoplasts were cultured *in vitro* to establish the A375M cell line.

The cell lines mentioned above were cultured in RPMI 1640 medium (HyClone, Logan, UT, USA) supplemented with 10% (v/v) fetal bovine serum (FBS, Gemini, West Sacramento, CA) and 100 U/mL penicillin/streptomycin (Sigma-Aldrich, St. Louis, MO, USA) and maintained in a humidified atmosphere with 5% CO_2 at 37 °C. The cells were harvested by 0.25% trypsin (HyClone) before use.

All of the cell lines were regularly subjected to mycoplasma testing. The A375 and A375M cell lines were characterized by the Genetic Testing Biotechnology Corporation (Suzhou, China) using short tandem repeat (STR) markers. The animal studies were performed following the animal protocol procedures approved by the Institutional Animal Care and Use Committee (IACUC) of Fuzhou University, which are

consistent with the AAALAS guidelines. All possible efforts were made to minimize animal suffering and sacrifice.

Quantification of gene expression by qRT-PCR

Total RNA was acquired using TRIzol (Thermo Fish, USA). Then, 100 ng of total RNA was used for cDNA synthesis in a 10 μl reaction system with oligo (dT), random 6 mers, and primeScript RT Enzyme Mix Ⅰ (Takara, Japan). The Quantitative Real-Time PCR (qRT-PCR) procedure was carried in a 20 μl reaction containing target primers, cDNA template, and SYBR Premix Ex Taq (TaKaRa), and performed by using a CFX96™ real-time PCR detection system (Bio-Rad, USA). Primers corresponding to the PCR targets were synthesized by Shanghai Sangon Biotech (Shanghai, China). Unless otherwise mentioned, each reaction was carried out in triplicate. Relative mRNA expression levels for each gene were analyzed using the $2^{-\triangle\triangle}$Ct method and normalized to the endogenous reference gene β-actin. All of the primer sequences are listed in Table 4.

Table 4
Primer sequences used in this study were listed.

Gene	Forward primer sequence(5' → 3')	Reverse primer sequence(5' → 3')
EGFR	GCCACCTGCGTGAAGAAGTGTC	ACGCCGTCTTCCTCCATCTCATAG
CDH1	ATGAAGAAGGAGGCGGAGAAGAGG	TGCAACGTCGTTACGAGTCACTTC
KRT5	AGAAGCCGAGTCCTGGTATCAGAC	CTTGGTGTTGCGGAGGTCATCG
KRT14	GAAGAACCACGAGGAGGAGATGAATG	CGTTCAGAATGCGGCTCAGGTC
DSP	TTAGTCCGACCTGGTACTGCTCTG	CTTCCTCCACTGGTAACCTCAAGTTG
IVL	CAGGCCAGGTCCAAGACATTCAAC	GCTTCTGCTGCTGGTGCTCTAC
DSG1	CTGCTGGCATTGGACTCCTCATC	AGCACCTCCACAATCACAACAGATC
KRT10	TCGATACAGCTCAAGCAAGCACTAC	CACTCCTCCTCCTCCTCCACATC
KRT1	GAGCTGAATCGTGTGATCCAGAGAC	TCACTGATGGACTGCTGCAAGTTG
FLG	CCATCACAGCCACACCACATCC	GTGCCGTCTCCTGATTGTTCCTC
COL17A1	GCTCTTCGGCCTCATTGCTCTG	CAGTATGCTCCTCCTGATCCTCTCC
CDK1	TGGGGTCAGCTCGTTACTCA	GGAGTGCCCAAAGCTCTGAA
β-actin	AGAAAATCTGGCACCACACC	AGAGGCGTACAGGGATAGCA

Statistical analysis

The data for all experiments were managed using GraphPad Prism software 5.0 and represented as the mean ± s.d. A paired *t*-test was used for two-group comparisons. Statistical significance is indicated by asterisks corresponding to $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$.

Results

Identification of Differentially Expressed Genes (DEGs)

The GSE4801 profile, including 31 primary melanomas samples and 52 melanoma metastases samples, were submitted to the GEO2R online analysis tool. Using adjusted *p* value < 0.01 and |logFC| >1.5 as cut-off criteria, a total of 424 DEGs were detected, among which 60 genes were upregulated and 364 were downregulated. The heat map of the full range of genes is shown in Fig. 1, and the top 15 significantly upregulated and downregulated genes are listed in Table 1. TGM2 was the most significantly upregulated gene; JUP was the most significantly downregulated gene.

Table 1
Differentially expressed genes between primary melanoma and metastatic melanoma.

A.Top 15 up-regulated DEGs			
Gene symbol	**Gene ID**	**Fold change**	**Adjust P value**
PSPH	5723	2.8486465	2.98E-10
SPP1	6696	2.767764	1.99E-07
IGF2BP3	10643	2.5198993	1.44E-08
DNAJB9	4189	2.2987063	4.92E-08
MAGEA6///MAGEA3	4105///4102	2.1640936	2.14E-03
ADAM12	8038	2.1087045	2.97E-09
IGF2BP3	10643	2.0515876	1.22E-08
RRM2	6241	1.9764716	1.95E-06
ITGB3	3690	1.9427608	2.67E-05
DHFR	1719	1.8997797	1.24E-06
MAGEA6	4105	1.8867916	1.09E-02
PAEP	5047	1.8672555	6.61E-03
CDK1	983	1.8363552	1.61E-06
UGT8	7368	1.8051681	6.63E-06
EXOC5	10640	1.7962246	3.13E-06
B.Top 15 down-regulated DEGs			
Gene symbol	**Gene ID**	**Fold change**	**Adjust P value**
S100A7	6278	-8.6805712	9.06E-20
KRT14	3861	-8.6624659	6.11E-17
KRT16	3868	-8.0258657	1.74E-20
SPRR1A	6698	-7.8002823	2.57E-19
KRT6A	3853	-6.8262439	1.06E-18
KRT17///JUP	3872///3728	-6.7016996	1.41E-19
KRT5	3852	-6.6593224	1.79E-17
KRT6C///KRT6B///KRT6A	286887///3854///3853	-6.4903185	1.36E-18
LOR	4014	-6.4902197	1.94E-16

A.Top 15 up-regulated DEGs			
SFN	2810	-6.3583246	6.30E-19
LGALS7B///LGALS7	653499///3963	-6.2382627	4.02E-15
KRT6B	3854	-6.0953977	8.68E-19
PKP1	5317	-6.0861137	1.14E-17
S100A8	6279	-6.0784181	2.57E-19
FLG	2312	-5.9593454	4.24E-14

GO function and KEGG pathway enrichment analysis of the DEGs

The top 5 enrichment analyses are shown in Table 2 for each part of the GO analysis. The upregulated genes significantly participated in the formation of cellular components, including condensed chromosome kinetochore, chromosome and centromeric region, condensed chromosome and centromeric region, and condensed chromosome, chromosomal region; and for the biological processes included sister chromatid segregation, mitotic cell cycle process, nuclear chromosome segregation, mitotic nuclear division, and cell cycle process; and for molecular function included small molecule binding, adenyl ribonucleotide binding, adenyl nucleotide binding, ATP binding, and nucleotide binding. In addition, the downregulated genes were mainly involved in cellular components, including extracellular region, extracellular region part, extracellular exosome, extracellular vesicle, and extracellular organelle; and for biological processes including epidermis development, skin development, keratinocyte differentiation, epidermal cell differentiation, and epithelial cell differentiation; and for molecular functions including structural molecule activity, endo-peptidase inhibitor activity, structural constituent of cytoskeleton, endopeptidase regulator activity, and peptidase inhibitor activity.

Table 2

Gene ontology analysis of differentially expressed genes associated with melanoma metastasis. (FDR: false discovery rate)

A. Up-regulated DEGs				
Category	**Term**	**Count**	P-**Value**	**FDR**
GOTERM_BP_FAT	GO:0000819 ~ sister chromatid segregation	7	2.91E-05	0.049010848
GOTERM_BP_FAT	GO:1903047 ~ mitotic cell cycle process	11	1.04E-04	0.17563807
GOTERM_BP_FAT	GO:0098813 ~ nuclear chromosome segregation	7	1.15E-04	0.193715552
GOTERM_BP_FAT	GO:0007067 ~ mitotic nuclear division	8	1.44E-04	0.241700828
GOTERM_BP_FAT	GO:0022402 ~ cell cycle process	13	1.87E-04	0.313364264
GOTERM_CC_FAT	GO:0000777 ~ condensed chromosome kinetochore	5	3.09E-04	0.381145565
GOTERM_CC_FAT	GO:0000775 ~ chromosome, centromeric region	6	3.14E-04	0.387290385
GOTERM_CC_FAT	GO:0000779 ~ condensed chromosome, centromeric region	5	4.26E-04	0.524074041
GOTERM_CC_FAT	GO:0000793 ~ condensed chromosome	6	5.10E-04	0.626972395
GOTERM_CC_FAT	GO:0098687 ~ chromosomal region	7	7.33E-04	0.9006285
GOTERM_MF_FAT	GO:0036094 ~ small molecule binding	16	0.001218847	1.59877789
GOTERM_MF_FAT	GO:0032559 ~ adenyl ribonucleotide binding	12	0.001322438	1.733564241
GOTERM_MF_FAT	GO:0030554 ~ adenyl nucleotide binding	12	0.001402442	1.837543422
GOTERM_MF_FAT	GO:0005524 ~ ATP binding	11	0.003900765	5.033846973
GOTERM_MF_FAT	GO:0000166 ~ nucleotide binding	14	0.005471503	6.993868115
B. **Down-regulated DEGs**				
Category	**Term**	**Count**	P-**Value**	**FDR**
GOTERM_BP_FAT	GO:0008544 ~ epidermis development	55	7.24E-42	1.36E-38
GOTERM_BP_FAT	GO:0043588 ~ skin development	43	2.57E-32	4.82E-29

A. Up-regulated DEGs				
GOTERM_BP_FAT	GO:0030216 ~ keratinocyte differentiation	30	1.12E-26	2.09E-23
GOTERM_BP_FAT	GO:0009913 ~ epidermal cell differentiation	33	6.61E-25	1.24E-21
GOTERM_BP_FAT	GO:0030855 ~ epithelial cell differentiation	50	4.45E-23	8.35E-20
GOTERM_CC_FAT	GO:0005576 ~ extracellular region	165	2.67E-27	3.67E-24
GOTERM_CC_FAT	GO:0044421 ~ extracellular region part	149	9.38E-27	1.29E-23
GOTERM_CC_FAT	GO:0070062 ~ extracellular exosome	110	2.61E-18	3.57E-15
GOTERM_CC_FAT	GO:1903561 ~ extracellular vesicle	110	3.80E-18	5.21E-15
GOTERM_CC_FAT	GO:0043230 ~ extracellular organelle	110	3.90E-18	5.35E-15
GOTERM_MF_FAT	GO:0005198 ~ structural molecule activity	46	6.04E-14	9.02E-11
GOTERM_MF_FAT	GO:0004866 ~ endopeptidase inhibitor activity	16	1.41E-07	2.10E-04
GOTERM_MF_FAT	GO:0005200 ~ structural constituent of cytoskeleton	13	2.01E-07	3.00E-04
GOTERM_MF_FAT	GO:0061135 ~ endopeptidase regulator activity	16	2.20E-07	3.28E-04
GOTERM_MF_FAT	GO:0030414 ~ peptidase inhibitor activity	16	2.72E-07	4.07E-04

KEGG pathways analysis was then conducted. As shown in Table 3, the upregulated genes were enriched in ECM-receptor interaction, progesterone-mediated oocyte maturation, regulation of actin cytoskeleton, oocyte meiosis, metabolic pathways, and cell cycle, while the downregulated genes were enriched in amoebiasis, melanogenesis, ECM-receptor interaction, pathways in cancer, histidine metabolism, protein digestion and absorption, pancreatic secretion, arachidonic acid metabolism, tyrosine metabolism, bladder cancer, and arrhythmogenic right ventricular cardiomyopathy (ARVC).

Table 3

KEGG pathway analysis of differentially expressed genes associated with melanoma metastasis.

Category	Term	Count	P-Value	Genes	FDR
A. Up-regulated DEGs					
KEGG_PATHWAY	hsa04512:ECM-receptor interaction	3	3.90E-02	ITGA4, ITGB3, SPP1	34.04
KEGG_PATHWAY	hsa04914:Progesterone-mediated oocyte maturation	3	3.90E-02	CDK1, MAD2L1, BUB1	34.04
KEGG_PATHWAY	hsa04810:Regulation of actin cytoskeleton	4	3.94E-02	LIMK1, PIP5K1A, ITGA4, ITGB3	34.33
KEGG_PATHWAY	hsa01100:Metabolic pathways	9	6.17E-02	DHFR, SLC33A1, GLUD2, RRM2, UGT8, PIP5K1A, PSPH, ACSL3, PYGB	48.64
KEGG_PATHWAY	hsa04110:Cell cycle	3	7.33E-02	CDK1, MAD2L1, BUB1	54.92
B.Down-regulated DEGs					
KEGG_PATHWAY	hsa05146:Amoebiasis	13	2.00E-07	IL1R2, GNA15, ARG1, GNAL, LAMB3, LAMA3, LAMC3, SERPINB2,LAMC2, SERPINB4, SERPINB3, SERPINB13, CTSG	2.45E-04
KEGG_PATHWAY	hsa04916:Melanogenesis	9	2.86E-04	DCT, WNT5A, TYRP1, WNT4, FZD10, ADCY2, CALML3, EDN1, CALML5	0.35
KEGG_PATHWAY	hsa04512:ECM-receptor interaction	8	6.70E-04	SDC1, LAMB3, LAMA3, LAMC3, COMP, COL6A2, ITGB4, LAMC2	0.82

A. Up-regulated DEGs					
KEGG_PATHWAY	hsa05200:Pathways in cancer	16	2.67E-03	WNT5A, FGFR2, EGFR, FGFR3, ADCY2, CDH1, BDKRB2, MMP1, JUP, CBLC,WNT4, LAMB3, FZD10, LAMA3, LAMC3, LAMC2	3.24
KEGG_PATHWAY	hsa00340:Histidine metabolism	4	6.70E-03	ALDH2, HAL, ALDH3B2, ALDH3A1	7.93
KEGG_PATHWAY	hsa04974:Protein digestion and absorption	6	1.70E-02	COL17A1, SLC15A1, COL7A1, PRSS2, PRSS3, COL6A2	18.98
KEGG_PATHWAY	hsa04972:Pancreatic secretion	6	2.11E-02	CLCA2, ADCY2, PRSS2, PRSS3, CA2, PLA2G3	23.03
KEGG_PATHWAY	hsa00590:Arachidonic acid metabolism	5	2.11E-02	GPX2, PTGDS, GPX3, ALOX12B, PLA2G3	23.05
KEGG_PATHWAY	hsa00350:Tyrosine metabolism	4	2.14E-02	DCT, TYRP1, ALDH3B2, ALDH3A1	23.31
KEGG_PATHWAY	hsa05219:Bladder cancer	4	3.23E-02	EGFR, FGFR3, CDH1, MMP1	33.24
KEGG_PATHWAY	hsa05412:Arrhythmogenic right ventricular cardiomyopathy	5	3.27E-02	JUP, ITGB4, DSC2, GJA1, DSP	33.55

Hub genes and module screening from the PPI network

Based on the information in the STRING protein query from the public databases of the DEGs, the top 10 genes with a high degree of connectivity were selected and defined as hub genes (Fig. 2). Additionally, we made a PPI network of the hub genes via Cytoscape software. Therein, 250 nodes and 751 edges were analyzed via the plug-in MCODE in Cytoscape, and the top 2 significant modules were filtered (Fig. 3). Based on the GO and KEGG pathway analysis, these two modules were principally associated with the cell cycle, ECM-receptor interaction, focal adhesion, and the PI3K-Akt signaling pathway.

qRT-PCR verification of the hub genes

The expression levels of the 10 hub genes were validated in the paired premetastatic melanoma cells (A375) and the postmetastatic melanoma cells (A375M) through qRTPCR. There were no significant differences in the expression of DSG1, KRT14 and FLG between the metastatic and primary melanoma cells. The expression of CDK1, COL17A1 and EGFR were significantly upregulated in metastatic melanoma ($p < 0.01$; Fig. 4). The expression of CDH1, DSP, IVL and KRT5 were significantly downregulated in metastatic melanoma cells relative to the primary cells ($p < 0.01$ Fig. 4).

Discussion

Melanoma is the most malignant skin cancer type and its metastasis remains essentially incurable because the mutated genes and underlying molecular mechanisms are unknown[18]. In the present study, the gene expression profile of GSE8401 was analyzed, and a total of 424 DEGs, including 60 upregulated genes and 364 downregulated genes, were identified between the 32 primary melanoma samples and the 51 metastatic melanoma samples obtained from patients.

The GO analysis revealed that these upregulated DEGs were mainly participating in small molecule binding, nucleotide binding, and cell cycle processes, while the downregulated DEGs were involved in the extracellular region, extracellular region part, extracellular organelle, exosome, and vesicles. These results illustrated that cell mitosis and malignant proliferation are activated, whereas the interaction with the extracellular environment was suppressed during the metastatic process.

The KEGG analysis demonstrated that the upregulated DEGs were mainly enriched in metabolic pathways and regulation of the actin cytoskeleton, and the downregulated DEGs were enriched in pathways in cancer and chemokine signaling. In addition, some of the upregulated DEGs were enriched in ECM-receptor interactions, as were some of the downregulated DEGs. Cross-talk between the ECM and melanoma cells plays a critical role in melanoma metastasis. Cancer metastasis involves multiple complex processes, which are critically influenced by ECM components[19]. Various ECM-related proteins are significantly deregulated during the progression of cancer, causing both biochemical and biomechanical changes that together promote cancer metastasis[20]. In this study, ITGA4, SPP1, and ITGB3 were increased, while COL6A2, LAMC2, LAMB3, SDC1, ITGB4, LAMC3, COMP, and LAMA3 were decreased. They regulate the interactions between melanoma cells and the ECM. These pathways are promising targets for new drug intervention. Thus, the GO term and KEGG pathway analyses are highlighting possible directions of experimental validation.

Among these DEGs, 10 hub genes with a high degree of connectivity were selected by evaluating the PPI network. According to qRT-PCR analysis, DSG1, FLG and KRT14 showed no significant differences between premetastatic and postmetastatic melanoma cells. The expression levels of EGFR and COL17A1 were inconsistent, with a tendency towards upregulated expression in the GSE8401 profile. In accordance with the literature, the expression of CDH1 was decreased, whereas EGFR and CDK1 were increased in metastatic melanoma compared to the primary site[21, 22, 23]. CDH1, EGFR and CDK1 were well-known

to be commonly dysregulated in malignancies and have been implicated in metastasis, especially in melanoma metastasis[24, 25, 26].

CDH1 encodes the E-cadherin protein, a key biomarker of epithelial-mesenchymal transition (EMT) in cancer cells and a mediator of cell-cell adhesion in epithelial tissues[27]. Loss of E-cadherin promotes the aggressive behaviors of melanoma cells via constitutively active Snail expression during the metastasis process[28, 29]. Mutations that lead to EGFR overexpression and amplification have been found to be associated with the transformation of cancer cells[30, 31]. CDK1 is a master regulator of the cell cycle, and overexpression of CDK1 in melanoma cells is associated with increased malignant proliferation and metastatic potential[32, 33].

Additionally, KRT5, IVL and COL17A1 have not been explored regarding their functional characteristics in melanoma, especially in melanoma metastasis. KRT5 encodes the protein keratin 5 that belongs to the keratin family, which are intermediate filament proteins. It is responsible for the structural integrity of epithelial cells and contributes to cell polarization, cytoskeleton regulation and protein translation[34]. Previous research has confirmed keratin 5 as a stem cell marker in breast cancer and it is associated with cancer recurrence and chemotherapy resistance in ovarian cancer[34, 35, 36]. In the present study, KRT5 was downregulated in postmetastatic melanoma relative to primary melanoma. No similar observations have been made before.

IVL, encoding the protein involucrin, is a marker of keratinocyte terminal differentiation, and it maintains the morphological characteristics of the epidermis[37]. Some findings have suggested that involucrin is a biomarker of EMT and is regulated by AP-1 in squamous cell carcinoma[38]. Another group also demonstrated that involucrin might be involved in breast cancer, cervical cancer and oral cancer[39, 40]. COL17A1 encodes the protein collagen alpha-1(XVII) chain that is responsible for adhesion of cell and matrix. Matsuda K. *et al.* showed COL17A1 was modulated by p53, and it inhibited breast cancer cell migration and invasion[41]. Furthermore, Pascal H. G. Duijf and his colleagues reported that COL17A1 was overexpressed in cervical and other epithelial cancers and predicted an increased metastatic signature[42]. Based on the research conducted so far, no report has attempted to discover abnormalities or functional features about KRT5, IVL and COL17A1 in melanoma. Through a metastatic mouse model of melanoma cells, we further verified the correlation between these hub genes and melanoma metastasis. Our results implicate CDK1, COL17A1, EGFR, CDH1, DSP, IVL and KRT5 in melanoma metastasis.

Conclusion

Taken together, our bioinformatics analysis presented here identified the DEGs involved in melanoma metastasis, which might have central roles in melanoma metastasis. A total of 424 DEGs and 10 hub genes were selected, and their enrichment analysis suggested that an interaction with ECM may play a dominant role in melanoma metastasis. Above all, we found that KRT5, IVL and COL1A1 might serve as biomarkers and therapeutic targets for the prevention and treatment of melanoma metastasis. The

findings of this study may contribute to the more profound elucidation of melanoma metastasis. However, additional verification experiments are necessary to confirm the results of these analyses.

Abbreviations

GEO
Gene Expression Omnibus; DEGs:Differentially expressed genes; DAVID:Database for Annotation, Visualization and Integrated Discovery; STRING:Search Tool for the Retrieval of Interacting Genes; MCODE:Cytoscape with Molecular Complex Detection; GO:gene ontology; KEGG:Kyoto Encyclopedia of Gene and Genome; ECM:Extracellular matrix; ARVC:Arrhythmogenic right ventricular cardiomyopathy; BP:biological processes; MF:molecular functions; CC:cellular components; PPI:protein-protein interaction; STR:short tandem repeat.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and material

Please contact author for data requests.

Conflict of Interests statement

Authors declare no conflicts of interests.

Funding

This work was supported by the National Natural Science Foundation of China (NSFC) (U1505225, 81773063, 81273548); Ministry of Science and Technology of China (2015CB931804).

Authors’ contributions

Yumei Li, Bifei Li and Lee Jia conceived and designed the project. Yumei Li and Fan Chen performed the experiments. Weiyu Shen analyzed and interpreted the data. Yumei Li and V.L katanaev wrote and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements
Not applicable

References

1. Cancer, Statistics. 2018.
2. Eggermont AM, Spatz A, Robert C. Cutaneous melanoma. The Lancet. 2014;383:816–27.
3. Braeuer RR, Watson IR, Wu CJ, Mobley AK, Kamiya T, Shoshan E, et al. Why is melanoma so metastatic? Pigment Cell Melanoma Res. 2014;27:19–36.
4. Melanoma.
5. Luo C, Lim J, Lee Y, Granter SR, Thomas A, Vazquez F, et al. A PGC1α-mediated transcriptional axis suppresses melanoma metastasis. Nature. 2016;537:422–6.
6. Lee JH, Miele ME, Hicks DJ, Phillips KK, Trent JM, Weissman BE, et al. KiSS-1, a novel human malignant melanoma metastasis-suppressor gene. J Natl Cancer Inst. 1996;88:1731–7.
7. McKenzie JA, Liu T, Jung JY, Jones BB, Ekiz HA, Welm AL, et al. Survivin promotion of melanoma metastasis requires upregulation of alpha5 integrin. Carcinogenesis. 2013;34:2137–44.
8. Krepler C, Sproesser K, Brafford P, Beqiri M, Garman B, Xiao M, et al. A Comprehensive Patient-Derived Xenograft Collection Representing the Heterogeneity of Melanoma. Cell Rep. 2017;21:1953–67.
9. Sun C, Yuan Q, Wu D, Meng X, Wang B. Identification of core genes and outcome in gastric cancer using bioinformatics analysis. Oncotarget. 2017;8:70271–80.
10. Yan J, Huang Q. Genomics screens for metastasis genes. Cancer Metastasis Rev. 2012;31:419–28.
11. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res. 2013;41:D991-5.
12. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics. 2007;23:1846–7.
13. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25:25–9.
14. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28:27–30.
15. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4:44–57.
16. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015;43:D447-52.
17. Li B, Shen W, Peng H, Li Y, Chen F, Zheng L, et al. Fibronectin 1 promotes melanoma proliferation and metastasis by inhibiting apoptosis and regulating EMT. Onco Targets Ther. 2019;12:3207–21.

18. Miller KD, Siegel RL, Lin CC, Mariotto AB, Kramer JL, Rowland JH, et al. Cancer treatment and survivorship statistics, 2016. CA: A Cancer. Journal for Clinicians. 2016;66:271–89.
19. Barney LE, Dandley EC, Jansen LE, Reich NG, Mercurio AM, Peyton SR. A cell–ECM screening method to predict breast cancer metastasis. Integr Biol-Uk. 2015;7:198–212.
20. Venning FA, Wullkopf L, Erler JT. Targeting ECM Disrupts Cancer Progression. Front Oncol. 2015;5:224.
21. Boone B, Jacobs K, Ferdinande L, Taildeman J, Lambert J, Peeters M, et al. EGFR in melanoma: clinical significance and potential therapeutic target. J Cutan Pathol. 2011;38:492–502.
22. Ravindran Menon D, Luo Y, Arcaroli JJ, Liu S, KrishnanKutty LN, Osborne DG, et al. CDK1 Interacts with Sox2 and Promotes Tumor Initiation in Human Melanoma. Cancer Res. 2018;78:6561–74.
23. Perrot CY, Gilbert C, Marsaud V, Postigo A, Javelaud D, Mauviel A. GLI2 cooperates with ZEB1 for transcriptional repression of CDH1 expression in human melanoma cells. Pigment Cell Melanoma Res. 2013;26:861–73.
24. Ma F, Li W, Liu C, Li W, Yu H, Lei B, et al. MiR-23a promotes TGF-beta1-induced EMT and tumor metastasis in breast cancer cells by directly targeting CDH1 and activating Wnt/beta-catenin signaling. Oncotarget. 2017;8:69538–50.
25. Li WY, Zhao TT, Xu HM, Wang ZN, Xu YY, Han Y, et al. The role of EGFR mutation as a prognostic factor in survival after diagnosis of brain metastasis in non-small cell lung cancer: a systematic review and meta-analysis. Bmc Cancer. 2019;19:145.
26. 10.1002/jcp.26387
Retraction: Pu, S, Zhao Y, Zhou G, Zhu H, Gong L, Zhang W, Huang G, Wang D, Liu D. Effect of CDK1 shRNA on proliferation, migration, cell cycle and apoptosis in non-small cell lung cancer. J. Cell. Physiol. Accepted Author Manuscript. doi: 10.1002/jcp.26387. J Cell Physiol 233 (2018) 7514.
27. Nawijn MC, Hackett TL, Postma DS, van Oosterhout AJ, Heijink IH. E-cadherin: gatekeeper of airway mucosa and allergic sensitization. Trends Immunol. 2011;32:248–55.
28. Kuphal S, Bosserhoff AK. E-cadherin cell-cell communication in melanogenesis and during development of malignant melanoma. Arch Biochem Biophys. 2012;524:43–7.
29. Canel M, Serrels A, Frame MC, Brunton VG. E-cadherin-integrin crosstalk in cancer invasion and metastasis. J Cell Sci. 2013;126:393–401.
30. von Felbert V, Bauerschlag D, Maass N, Brautigam K, Meinhold-Heerlein I, Woitok M, et al. A specific photoimmunotheranostics agent to detect and eliminate skin cancer cells expressing EGFR. J Cancer Res Clin Oncol. 2016;142:1003–11.
31. Katunaric M, Zamolo G, Jonjic N. EGFR activated cell mobility - A link to melanoma ulceration. Med Hypotheses. 2015;85:498–9.
32. Herrera MC, Chymkowitch P, Robertson JM, Eriksson J, Boe SO, Alseth I, et al. Cdk1 gates cell cycle-dependent tRNA synthesis by regulating RNA polymerase III activity. Nucleic Acids Res. 2018;46:12188–9.

33. Santamaria D, Barriere C, Cerqueira A, Hunt S, Tardy C, Newton K, et al. Cdk1 is sufficient to drive the mammalian cell cycle. Nature. 2007;448:811–5.
34. Ricciardelli C, Lokman NA, Pyragius CE, Ween MP, Macpherson AM, Ruszkiewicz A, et al. Keratin 5 overexpression is associated with serous ovarian cancer recurrence and chemotherapy resistance. Oncotarget. 2017;8:17819–32.
35. Vasca V, Vasca E, Freiman P, Marian D, Luce A, Mesolella M, et al. Keratin 5 expression in squamocellular carcinoma of the head and neck. Oncol Lett. 2014;8:2501–4.
36. Ricciardelli C, Lokman NA, Pyragius CE, Ween MP, Macpherson AM, Ruszkiewicz A, et al. Keratin 5 overexpression is associated with serous ovarian cancer recurrence and chemotherapy resistance. Oncotarget. 2017;8:17819–32.
37. Involucrin and Other Markers of Keratinocyte Terminal. Differentiation Author links open overlay panelFiona M.WattD.Phil..
38. Azuma Y, Chou S, Lininger RA, Murphy BJ, Varia MA, Raleigh JA. Hypoxia and differentiation in squamous cell carcinomas of the uterine cervix: pimonidazole and involucrin. Clinical cancer research: an official journal of the American Association for Cancer Research. 2003;9:4944.
39. Lan Y, Chen H, Chen J, Lei Q, Zheng M, Shao Z. Immunolocalization of Vimentin, Keratin 17, Ki-67, Involucrin, β-Catenin and E-Cadherin in Cutaneous Squamous Cell Carcinoma. Pathol Oncol Res. 2014;20:263–6.
40. Tsuda H, Sakamaki C, Fukutomi T, Hirohashi S. Squamoid features and expression of involucrin in primary breast carcinoma associated with high histological grade, tumour cell necrosis and recurrence sites. Br J Cancer. 1997;75:1519–24.
41. Yodsurang V, Tanikawa C, Miyamoto T, Lo P, Hirata M, Matsuda K. Identification of a novel p53 target, COL17A1, that inhibits breast cancer cell migration and invasion. Oncotarget. 2017;8:55790–803.
42. Thangavelu PU, Krenács T, Dray E, Duijf PHG. In epithelial cancers, aberrant COL17A1 promoter methylation predicts its misexpression and increased invasion. Clin Epigenetics 8 (2016).

Figures

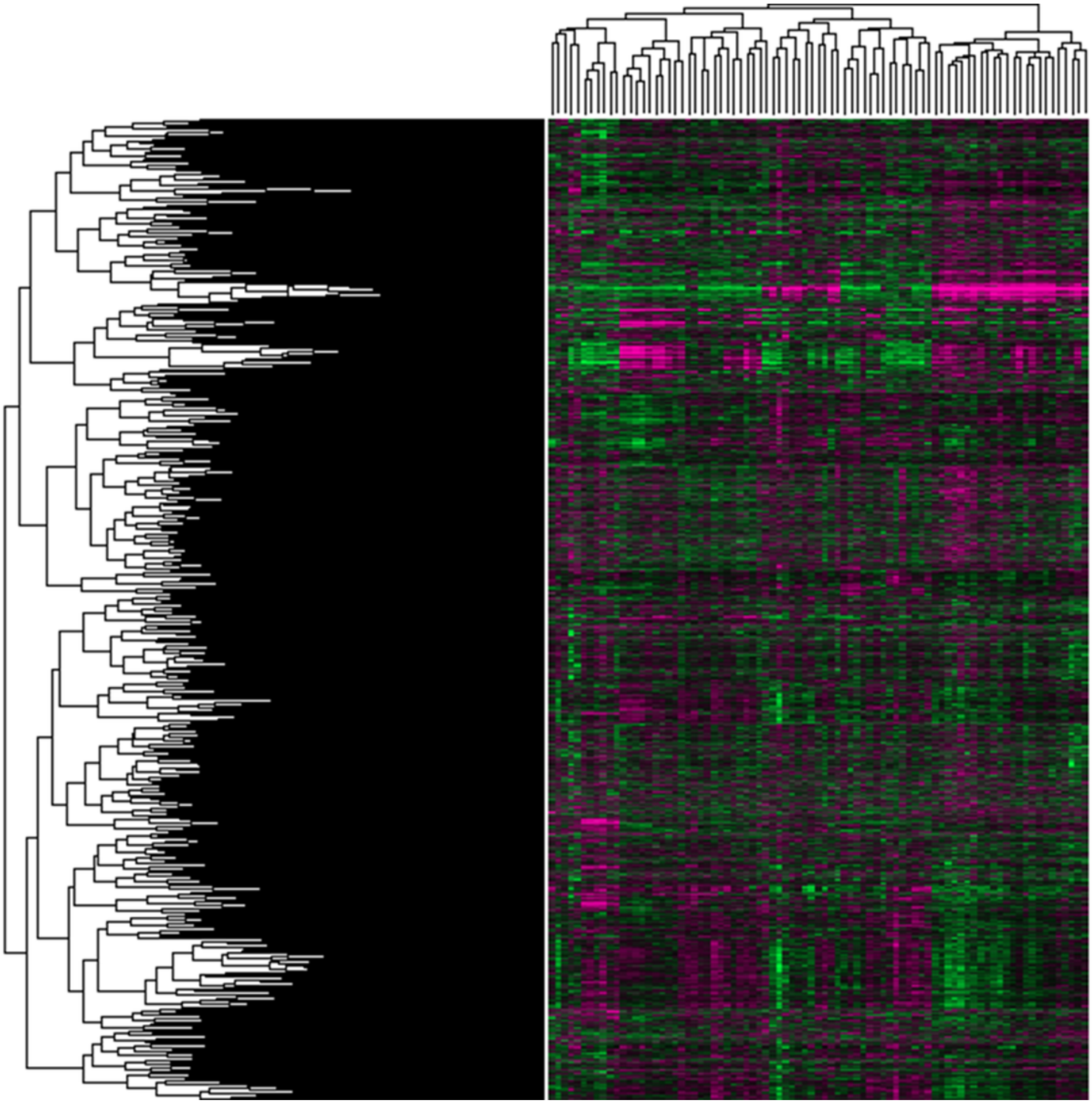


Figure 1

Heat map visualization of alterations of the full range of gene expression patterns between primary melanoma and metastatic melanoma. Purple: upregulation; Green: downregulation.

A

Gene	Degree of connectivity	Adjusted *P* value	Fold Change
CDH1	37	1.30E-03	-2.5
EGFR	36	5.16E-12	-2.86
KRT5	28	1.61E-15	-6.66
COL17A1	25	3.86E-14	-3.96
KRT14	23	6.11E-17	-8.66
IVL	23	3.17E-11	-3.01
DSP	21	5.15E-15	-4.04
DSG1	20	1.02E-14	-4.98
FLG	20	4.24E-14	-5.96
CDK1	19	2.64E-11	1.52

B

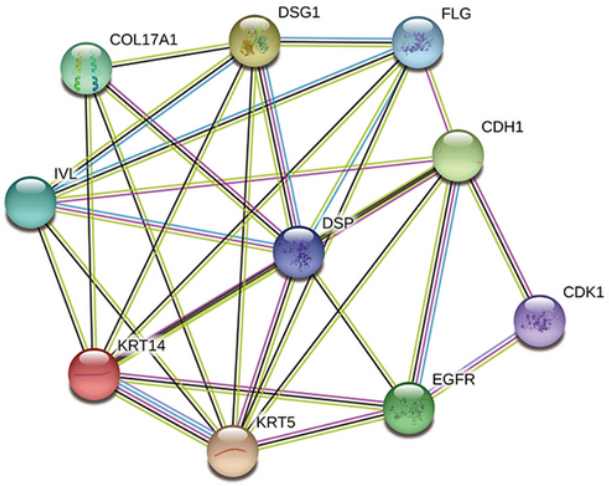
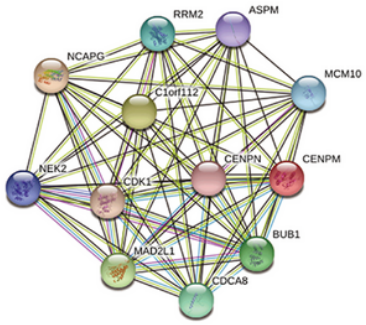


Figure 2

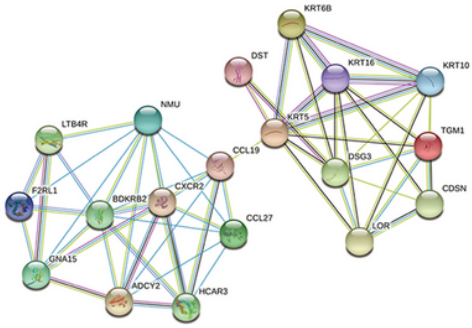
Top 10 hub genes with a higher degree of connectivity. (A) The hub genes are listed. (B) The PPI network of the top 10 hub genes.

A



Term	*P*-Value	Genes
cfa04914:Progesterone-mediated oocyte maturation	4.73E-04	CDK1, MAD2L1, BUB1
cfa04114:Oocyte meiosis	7.32E-04	CDK1, MAD2L1, BUB1
cfa04110:Cell cycle	9.68E-04	CDK1, MAD2L1, BUB1
cfa04115:p53 signaling pathway	0.0284862	CDK1, RRM2

B



Term	*P*-Value	Genes
bta04512:ECM-receptor interaction	3.88E-10	LAMB3, LAMA3, LAMC3, COL6A2, ITGB4, LAMC2, ITGA4
bta04510:Focal adhesion	7.60E-08	LAMB3, LAMA3, LAMC3, COL6A2, ITGB4, LAMC2, ITGA4
bta04151:PI3K-Akt signaling pathway	1.58E-06	LAMB3, LAMA3, LAMC3, COL6A2, ITGB4, LAMC2, ITGA4
bta05412:Arrhythmogenic right ventricular cardiomyopathy	8.27E-05	JUP, ITGB4, DSP, ITGA4
bta05222:Small cell lung cancer	1.71E-04	LAMB3, LAMA3, LAMC3, LAMC2
bta05146:Amoebiasis	3.49E-04	LAMB3, LAMA3, LAMC3, LAMC2
bta05145:Toxoplasmosis	4.38E-04	LAMB3, LAMA3, LAMC3, LAMC2
bta05200:Pathways in cancer	0.001242039	JUP, LAMB3, LAMA3, LAMC3, LAMC2

Figure 3

Top 2 modules from the protein-protein interaction network. (A) Module 1 and the enriched pathways of module 1. (B): Module 2 and the enriched pathways of module 2.

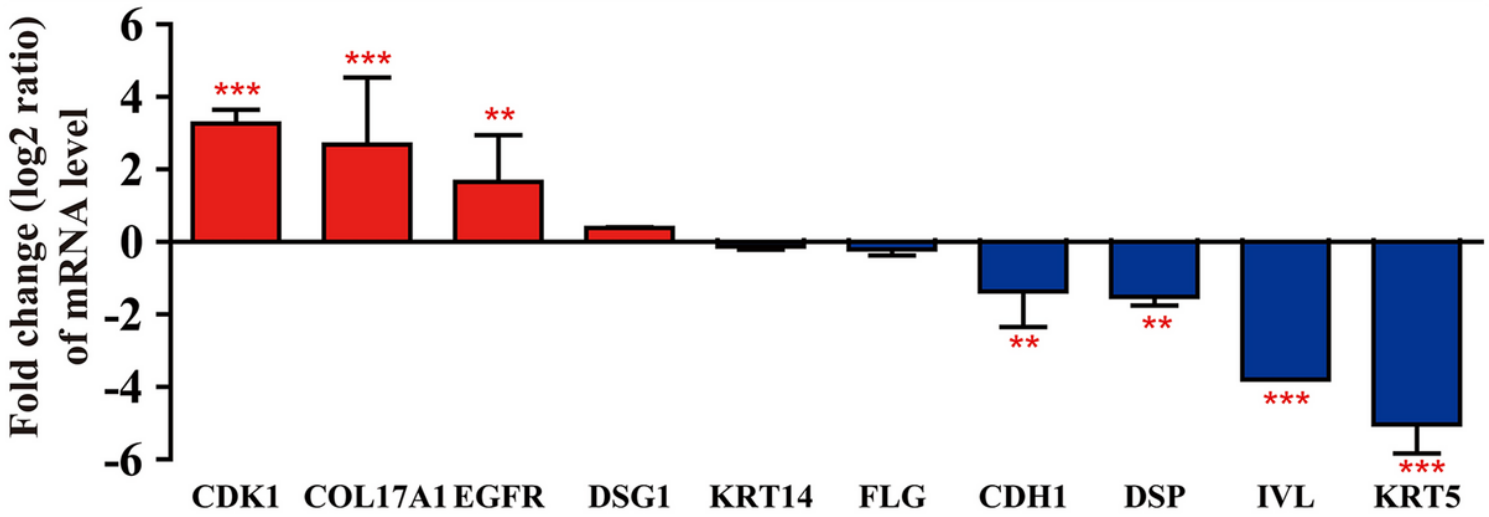


Figure 4

qRT-PCR validation of differentially expressed genes between primary melanoma and metastatic melanoma. CDK1, cyclin-dependent kinase 1; COL17A1, collagen alpha-1(XVII) chain; EGFR, epidermal growth factor receptor; DSG1, desmoglein-1; KRT14, keratin, type I cytoskeletal 14; FLG, filaggrin; CDH1, E-cadherin; DSP, desmoplakin; IVL, involucrin; KRT5, keratin, type II cytoskeletal 5. mRNA expression greater than or equal to one-fold difference.