Quality Improvement Study

Medicine®

OPEN

# Bioinformatics analysis to identify the critical genes, microRNAs and long noncoding RNAs in melanoma

Qian Zhang, PhD, Yang Wang, MD, Jiulong Liang, MD, Yaguang Tian, MD, Yu Zhang, MM, Kai Tao, MD*

**Abstract**

Melanoma, which is usually induced by ultraviolet light exposure and the following DNA damage, is the most dangerous skin cancer. The purpose of the present study was to screen key molecules involved in melanoma.

Microarray data of E-MTAB-1862 were downloaded from the ArrayExpress database, which included 21 primary melanoma samples and 11 benign nevus samples. In addition, the RNASeq version 2 and microRNA (miRNA) sequencing data of cutaneous melanoma were downloaded from The Cancer Genome Atlas database. After identifying the differentially expressed genes (DEGs) using Limma package, enrichment analysis and protein-protein interaction (PPI) network analysis were performed separately for them using DAVID software and Cytoscape software. In addition, survival analysis and regulatory network analysis were further performed by log-rank test and Cytoscape software, respectively. Moreover, real-time reverse transcription polymerase chain reaction (RT-PCR) was performed to further verify the expression patterns of several selected DEGs.

A total of 382 DEGs were identified in primary melanoma samples, including 206 upregulated genes and 176 downregulated genes. Functional enrichment analysis showed that *COL17A1* was enriched in epidermis development. In the PPI network, CXCL8 (degree = 29) and STAT1 (degree = 28) had higher degrees and could interact with each other. Survival analysis showed that 21 DEGs, 55 long noncoding RNAs (lncRNAs) and 32 miRNAs were found to be associated with prognosis. Furthermore, several regulatory relationships were found in the lncRNA-gene regulatory network (such as *RP11-361L15.4* targeting *COL17A1*) and the miRNA-gene regulatory network (such as *hsa-miR-375* targeting *CCL27* and *hsa-miR-375* targeting insulin-like growth factor 1 receptor [*IGF1R*]). Real-time RT-PCR results showed that the overall direction of differential expression was consistent except *COL17A1*.

*CXCL8* interacted with *STAT1*, *CCL27*, and *IGF1R* targeted by *hsa-miR-375*, and *COL17A1* targeted by *RP11-361L15.4* might function in the development and progression of melanoma, which should be verified by more detailed experiments.

**Abbreviations:** BP = biological process, BRAF = B-Raf proto-oncogene, serine/threonine kinase, CC = cellular component, DEG = differentially expressed gene, GO = Gene Ontology, IGF1R = insulin-like growth factor 1 receptor, lncRNA = long noncoding RNA, MF = molecular function, miRNA = microRNAs, miRNASeq = miRNA-sequencing, NF-κB = nuclear factor kappa B, PI3K = phosphatidylinositol 3-kinase, PPI = protein-protein interaction, RNASeqV2 = RNASeq version 2, SPRY4-IT1 = SPRY4 intronic transcript 1, TCGA = The Cancer Genome Atlas.

**Keywords:** differentially expressed genes, long noncoding RNA, melanoma, microRNA, protein-protein interaction network, regulatory network

## 1. Introduction

Melanoma is the most aggressive type of skin cancer arising from the malignant transformation of melanocytes.[1] Malignant melanoma has a strong tendency to spread to other parts of the body and causes serious illness and death.[2] In 2016, there was an estimation of 76,380 new diagnoses of cutaneous melanoma and 10,130 deaths related to melanoma in the United States.[3] Despite significant breakthroughs and advances in early diagnosis and prevention as well as targeted therapies, the prognosis of melanoma remains unoptimistic. In addition, evidence shows that the 5-year survival rate decreases to 63% in patients with melanoma with regional metastases and it is only 16% in patients with distant metastases.[4] Therefore, identifying molecular pathogenic basis for melanoma and novel treatment strategies remains an active area of research.

Recently, many researches have focused on the roles of mRNAs, microRNAs (miRNAs), and long noncoding RNAs (lncRNAs) in the tumorigenesis of melanoma. For instance, neural precursor cell expressed, developmentally downregulated 9 , which is significantly overexpressed in human metastatic melanoma, promotes invasion ability in vitro and metastasis ability in vivo of melanocytes.[5] Recently, Zeng et al[6] demonstrated that transactivating response region DNA-binding protein (TDP-43) was a novel oncogene in melanoma and it can regulate melanoma

Editor: Mihalis Panagiotidis.

The authors have no conflicts of interest to disclose.

Department of Plastic Surgery, General Hospital of Shenyang Military Area Command, Shenyang, Liaoning, China.

* Correspondence: Kai Tao, Department of Plastic Surgery, General Hospital of Shenyang Military Area Command, PLA. 83. Wenhua Road, Shenhe District, Shenyang, Liaoning, China 110016 (e-mail: taokaitkai@126.com).



Medicine (2017) 96:29(e7497)

Received: 19 January 2017 / Received in final form: 7 June 2017 / Accepted: 15 June 2017

http://dx.doi.org/10.1097/MD.0000000000007497

proliferation and metastasis. *MiR-211* inhibits expressions of the entire melastatin locus insulin-like growth factor type 2 receptor, transforming growth factor β receptor type II, and nuclear factor–activated T cell 5 , suggesting that *miR-211* suppresses melanoma invasion in the progression of human melanoma.[7] A study reported that miR-7-5p could inhibit melanoma cell proliferation and metastasis in part through its direct suppression of nuclear factor kappa B (NF-κB) subunit RelA expression and NF-κB signaling.[8] The lncRNA SPRY4 intronic transcript 1 (*SPRY4-IT1*) is overexpressed in melanoma cells, accumulates in cell cytoplasm, and affects cell dynamics, indicating that *SPRY4-IT1* may be related to the molecular mechanisms of human melanoma.[9] The lncRNA BRAF–activated noncoding RNA is reported to overexpress in human melanoma, and be implicated in the proliferation of melanoma cell in vivo and in vitro through mediating the activation of mitogen-activated protein kinase (MAPK) pathway.[10] Overexpressed lncRNA growth arrest-specific transcript 5 can reduce the invasion and migration of melanoma SK-Mel-110 cells through inhibiting the expression and activity of matrix metalloproteinase 2 to some degree.[11] However, a number of melanoma-associated cellular molecules remain unexploited, and the precise molecular mechanisms of melanoma are not yet fully understood.

Bioinformatic data mining of gene expression microarray data provides a useful tool for revealing many previously unrecognized significant genes implicated in the pathogenesis of diseases.[12,13] For instance, a recent study analyzed previously published melanoma gene expression profiles from available database and identified several lncRNAs and mRNAs that may be associated with melanoma tumorigenesis and metastasis.[14] In 2016, Eriksson et al[15] carried out analysis of gene expression profiles to identify the common changes in metastatic and nonmetastatic primary melanomas and benign nevi, finding that collagen triple helix repeat containing-1 regulated by B-Raf proto-oncogene, serine/threonine kinase (*BRAF*), nuclear factor of activated T cells-c2, and transforming growth factor β played an essential role in migration and invasion of melanoma cells and could be used as therapeutic target for metastatic melanomas. In order to identify more genes, miRNAs and lncRNAs associated with melanoma, the microarray data deposited by Eriksson et al, as well as the RNASeq version 2 (RNASeqV2) and miRNA sequencing (miRNASeq) data of cutaneous melanoma from The Cancer Genome Atlas (TCGA) database were downloaded and reanalyzed. We screened the differentially expressed genes (DEGs) between primary melanoma samples and benign nevus samples. Afterwards, enrichment analysis and protein-protein interaction (PPI) network analysis were conducted for the DEGs. In addition, survival analysis and regulatory network construction were further carried out to select the key molecules in melanoma.

## 2. Materials and methods

### 2.1. Data source

The flow diagram of bioinformatics analysis is shown in Figure 1. Microarray data under the accession number of E-MTAB-1862 were extracted from the European Bioinformatics Institute ArrayExpress database (http://www.ebi.ac.uk/arrayexpress), which were sequenced on the platform of A-AFFY-44—Affymetrix GeneChip Human Genome U133 Plus 2.0 (HG-U133_Plus_2). The E-MTAB-1862 dataset included 21 primary melanoma samples (8 females and 13 males, mean Breslow thickness = 6.8) and 11 benign nevus samples (6 females and 5 males) isolated from skin. In addition, the RNASeqV2 and miRNASeq data of cutaneous melanoma were downloaded from TCGA database. This study just reanalyzed the microarray data downloaded from public database and performed bioinformatics analysis. The authors declare that no experiments were performed on humans or animals for this investigation. Thus, ethics approval or consent to participate was not applicable.

### 2.2. Data preprocessing and DEGs screening

After E-MTAB-1862 was downloaded, the raw data were converted to expression matrix using the Affy package[16] in R. Combined with the annotation information from the platform, probes were converted into gene symbols. When multiple probes mapped to the same gene symbol, their mean value was considered as the gene expression value. Through aligning to the GENCODE lncRNA data,[17] the RNASeqV2 data were further annotated.

The DEGs between primary melanoma samples and benign nevus samples were screened by the linear models for microarray

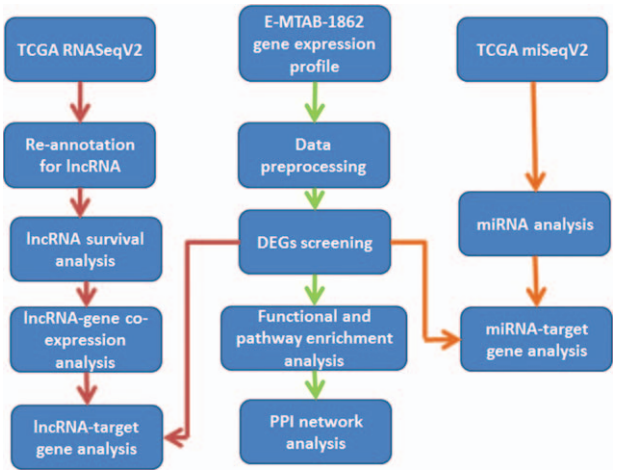


**Figure 1.** The flow diagram of bioinformatics analysis. DEG = differentially expressed gene, lncRNA = long noncoding RNA, PPI = protein-protein interaction, RNASeqV2 = RNASeq version 2, TCGA = The Cancer Genome Atlas.

data (Limma) package[18,19] in R. The adjusted *P* value (all named false discovery rate, FDR) <.05 and $|\log_2$fold change (FC)| >1 were used as the cut-off criteria.

### 2.3. Functional and pathway enrichment analysis

Gene Ontology (GO, http://www.geneontology.org/) describes genes and gene products from molecular function (MF), cellular component (CC), and biological process (BP) aspects.[20] Kyoto Encyclopedia of Genes and Genomes (http://www.genome.jp/kegg/), which integrates information of systemic functional, genomic, and chemical aspects, is known for pathway mapping.[21] The DEGs were performed with functional and pathway enrichment analyses using DAVID software (http://david.abcc.ncifcrf.gov/),[22] with the FDR <0.05.

### 2.4. PPI network construction

The Search Tool for the Retrieval of Interacting Genes (STRING, http://string-db.org/) database contains a large amount of known and predicted interaction information.[23] Cytoscape software (http://www.cytoscape.org/) is developed for constructing a unified framework via integrating biomolecular interaction networks and gene expression data.[24] The interactions among the proteins encoded by the DEGs were predicted by the STRING database,[23] and then the PPI network was visualized using the Cytoscape software.[24] The proteins involved in more interactions were key nodes in the network.

### 2.5. Survival analysis

The median value (each variable value in the statistical population was sorted in the order of size to form a series, and the variable value in middle position was the median value) of each lncRNA, miRNA, and mRNA in tumor samples was calculated. Then, the samples with expression higher than the median value were divided into high expression group, and the samples with expression lower than the median value were divided into low expression group. Using the KMsurv package[25,26] and the survival package[27] in R, the Kaplan-Meier curve was constructed for the high and low expression groups. Afterwards, the significant difference in prognosis between the 2 groups was obtained by log-rank test,[28] with *P* value <.05 as the threshold.

### 2.6. Predicting the target genes of the risk lncRNAs and miRNAs

The target genes of lncRNAs were predicted based on coexpression relationships, and the genes with FDR <0.05 were obtained as the potential targets of the lncRNAs. What's more, miRanda (http://www.mocrorna.org),[29] MirTarget2 (http://mirdb.org/miRDB),[30] PicTar (http://pictar.mdc-berlin.de/),[31] PITA (http://genie.weizmann.ac.il/pubs/mir07/mir07_data.html),[32] TargetScan (http://www.targetscan.org/),[33,34] miRecords (http://www.mirecords.umn.edu),[35] and Mirwalk (http://mirwalk.uni-hd.de/)[36] databases were used to search the target genes of prognostic risk miRNAs. The genes being found in 1 or more databases were considered as the targets of prognostic risk miRNAs. Finally, the regulatory network was constructed using Cytoscape software.[24]

### 2.7. Quantitative real-time polymerase chain reaction validation

In an attempt to validate differential expression of genes obtained by analyzing the microarray data, we selected a total of 5 genes, 2 upregulated and 3 downregulated, and quantified their expression levels by real-time polymerase chain reaction (PCR). Real-time PCR was performed on 5 primary melanoma samples and 10 benign nevus samples. Total RNA from the sample tissue was isolated using Trizol (TAKARA, Japan) protocol. First-strand cDNA was generated from 0.5 μg total RNA using PrimerScript RT Master Mix (Takara Biotechnology Co, Ltd, Dalian, China) according to the manufacturer's instructions and quantitative PCR was performed using SYBR Premix Ex Taq (Perfect Real-Time) Kit (Takara Biotechnology, Dalian, PR China) according to the manufacturer's protocol. PCRs were performed in triplicate in a total reaction volume of 20 μL, including 10 μL SYBR Premix Ex Taq (2×), 1 μL of PCR Forward Primer (10 μM), 1 μL of PCR Reverse Primer (10 μM), and 8 μL of cDNA (diluted in double-distilled water). The quantitative real-time PCR reaction was set at an initial step of 3 minutes at 50°C and 3 minutes at 95°C; and 95°C (10 seconds) and 60°C (30 seconds) in a total of 40 cycles. All experiments were done in triplicate and all samples were normalized to glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Primers used for amplification were as follows: CXCL8 sense, 5′- CCTGATTTCTGCAGCTCTGTG-3′ and CXCL8 antisense, 5′-CCAGACAGAGCTCTC.

TTCCAT-3′; STAT1 sense, 5′-GGTTAACGTTCGCACTCTG TG-3′ and STAT1 antisense, 5′-GCTGCTGAAGTTCGTACC AC-3′; COL17A1 sense, 5′-GCAGCAGCGGCTACATA

AAC-3′ and COL17A1 antisense, 5′-AGAAGAGTTGC-CACTGGAGC-3′; insulin-like growth factor 1 receptor (IGF1R) sense, 5′-CTCTGGCCGACGAGTGGAG-3′ and IGF1R antisense, 5′-CTCGGTAATGACCGT

GAGCTT-3′; CCL27 sense, 5′- AGGCTGAGCAACAT-GAAGGG-3′ and CCL27 antisense, 5′- TGCTGGGTGGCAG-TAGGAAT-3′; GAPDH sense, 5′- CAGTGCCAGCCTCGTCT CAT-3′ and GAPDH antisense, 5′-AGGGGCCATCCACAGT CTTC-3′. Statistical analysis of the PCR data was conducted using the relative expression software tool algorithm, which uses a pairwise-fixed reallocation and randomization test to determine significance (28). Data are expressed as the mean ± SEM from 3 independent experiments. Statistical analysis was carried out with SPSS 22.0 software (SPSS Inc, Chicago, IL) using a Student *t* test. A value of $P < .05$ was considered as statistically significant.

## 3. Results

### 3.1. Data preprocessing and DEGs analysis

The expression values of 20,545 genes were obtained from the 32 samples included in E-MTAB-1862. After further annotating the RNASeqV2 data downloaded from TCGA database, the expression values of 465 lncRNAs were also acquired from 448 samples. Among the 448 samples, 154 samples had annotation information of death time and were used for the subsequent survival analysis.

Under the thresholds of $|\log_2$FC >1 and FDR <0.05, total of 382 DEGs (including 206 upregulated genes and 176 downregulated genes) were screened in primary melanoma samples compared with benign nevus samples. Therefore, the number of upregulated genes was more than that of downregulated genes.

### 3.2. Functional and pathway enrichment analysis

The top 5 GO and pathway terms enriched for the upregulated genes are listed in Table 1, including immune response (GO_BP, $P = 1.70E-14$), extracellular region part (GO_CC, $P = 1.74E-07$), chemokine activity (GO_MF, $P = 1.09E-05$), and toll-like receptor signaling pathway (pathway, $P = 7.68E-05$). In addition, the GO terms significantly enriched for the downregulated genes including epidermis development (GO_BP; $P = 8.98E-07$; which involved collagen, type XVII, alpha 1, *COL17A1*), integral to plasma membrane (GO_CC, $P = 4.19E-04$) and sequence-specific DNA binding (GO_MF, $P = 5.64E-03$) (Table 2). For the downregulated genes, there were no significantly enriched pathways. The enriched GO_BP terms for the upregulated genes and the downregulated genes separately are shown in Figures 2 and 3.

### 3.3. PPI network construction

A total of 226 nodes (including 137 upregulated genes and 89 downregulated genes) and 598 interactions were involved in the PPI network constructed for the DEGs (Fig. 4). Especially, chemokine (C-X-C motif) ligand 8 (CXCL8, degree = 29) and signal transducers and activators of transcription 1 (STAT1, degree = 28) were key nodes in the PPI network for they had higher degrees, and they could interact with each other.

### 3.4. Survival analysis

Survival analysis was performed for the DEGs, lncRNAs, and miRNAs in melanoma samples. As a result, a total of 21 DEGs, 55 lncRNAs, and 32 miRNAs were found to be associated with prognosis.

**Table 1**

**The top 5 functional and pathway terms enriched for the upregulated genes.**

| Category | Term | Gene number | Gene symbol | *P* |
|---|---|---|---|---|
| GO_BP | GO:0006955~immune response | 37 | *GPR183, MICB, S100A7, CCR1, GPR65, CXCL9, CCL8, RSAD2, IFI44L, CCL5, CXCL11, IL7R, C1QC, CCL4, PNP, CXCL10, CXCR4, FCGR1B, FCER1G, CFI, FCGR3B, PTPRC, C5AR1, GBP5, BST2, LY96, GZMA, CTSS, TRIM22, C1QA, C1QB, RGS1, TNFSF13B, CXCL13, SERPINB4, CD14, GBP1* | 1.70E-14 |
| GO_BP | GO:0009611~response to wounding | 28 | *C3AR1, S100A8, TNC, CCR1, S100A9, CXCL9, CCL8, CXCL11, CCL5, CCL4, C1QC, CXCL10, CXCR4, CFI, PAPSS2, SPP1, PLEK, LY96, S100A12, C1QA, CCNB1, C1QB, HIF1A, CD36, CXCL13, VCAN, CD14, IGFBP4* | 8.91E-11 |
| GO_BP | GO:0006954~inflammatory response | 22 | *C3AR1, S100A8, LY96, CCR1, S100A9, CXCL9, CCL8, CCL5, CXCL11, C1QC, CCL4, S100A12, CXCL10, C1QA, C1QB, HIF1A, CXCR4, CXCL13, CFI, CD14, IGFBP4, SPP1* | 1.90E-10 |
| GO_BP | GO:0006952~defense response | 29 | *C3AR1, S100A8, S100A7, CCR1, S100A9, CXCL9, CCL8, RSAD2, CCL5, CXCL11, C1QC, CCL4, CXCL10, CXCR4, CFI, TYROBP, SPP1, PTPRC, C5AR1, LY96, S100A12, C1QA, C1QB, HIF1A, CXCL13, MNDA, PLA2G2A, CD14, IGFBP4* | 5.07E-10 |
| GO_BP | GO:0009615~response to virus | 12 | *PTPRC, BST2, ISG15, CXCR4, CCL8, RSAD2, IFI44, STAT1, CCL5, TRIM22, CCL4, ISG20* | 5.58E-08 |
| GO_CC | GO:0044421~extracellular region part | 31 | *WNT5A, CTHRC1, MMP9, LUM, TNC, CXCL9, CCL8, CCL5, CXCL11, CCL4, MMP1, CXCL10, ISG15, CFI, SRGN, SPP1, COL4A2, COL4A1, LY96, NID2, COL5A2, C1QA, C1QB, TNFSF13B, CXCL13, PI3, SULF1, PLA2G2A, VCAN, IGFBP2, IGFBP4* | 1.74E-07 |
| GO_CC | GO:0005576~extracellular region | 47 | *WNT5A, CTHRC1, S100A7, MMP9, LUM, TNC, CXCL9, CCL8, CCL5, CXCL11, IL7R, C1QC, CCL4, MMP1, CXCL10, C1ORF54, ISG15, CFI, FCGR3B, SRGN, SPP1, CYR61, COL4A2, COL4A1, PLEK, LY96, GZMA, OLFML2B, APOLD1, C10ORF99, NID2, CTSS, TCN1, COL5A2, C1QA, C1QB, TNFSF13B, CXCL13, SULF1, PI3, PLA2G2A, FCRLA, VCAN, IGFBP2, CD14, ADAMDEC1, IGFBP4* | 5.44E-07 |
| GO_CC | GO:0005615~extracellular space | 21 | *WNT5A, LY96, MMP9, CXCL9, CCL8, CCL5, CXCL11, CCL4, CXCL10, C1QA, C1QB, TNFSF13B, ISG15, CXCL13, SULF1, PLA2G2A, CFI, IGFBP2, IGFBP4, SRGN, SPP1* | 6.79E-05 |
| GO_CC | GO:0005578~proteinaceous extracellular matrix | 12 | *WNT5A, CTHRC1, COL4A2, COL4A1, TNC, MMP9, LUM, PI3, VCAN, NID2, COL5A2, MMP1* | 8.82E-04 |
| GO_CC | GO:0031012~extracellular matrix | 12 | *WNT5A, CTHRC1, COL4A2, COL4A1, TNC, MMP9, LUM, PI3, VCAN, NID2, COL5A2, MMP1* | 1.62E-03 |
| GO_MF | GO:0008009~chemokine activity | 7 | *CXCL13, CXCL9, CCL8, CXCL11, CCL5, CCL4, CXCL10* | 1.09E-05 |
| GO_MF | GO:0042379~chemokine receptor binding | 7 | *CXCL13, CXCL9, CCL8, CXCL11, CCL5, CCL4, CXCL10* | 1.58E-05 |
| GO_MF | GO:0005125~cytokine activity | 10 | *NAMPT, TNFSF13B, CXCL13, CXCL9, CCL8, CXCL11, CCL5, CCL4, SPP1, CXCL10* | 3.21E-04 |
| GO_MF | GO:0019864~IgG binding | 3 | *FCGR1B, FCER1G, FCGR3B* | 3.27E-03 |
| GO_MF | GO:0005539~glycosaminoglycan binding | 7 | *PTPRC, SELL, CCL8, VCAN, CD14, CYR61, HMMR* | 4.67E-03 |
| Pathway | hsa04620:Toll-like receptor signaling pathway | 9 | *LY96, CXCL9, CXCL11, STAT1, CCL5, CCL4, CD14, SPP1, CXCL10* | 7.68E-05 |
| Pathway | hsa04060:Cytokine-cytokine receptor interaction | 12 | *TNFRSF21, TNFSF13B, CXCR4, CXCL13, CCR1, CXCL9, CCL8, CXCL11, IL7R, CCL5, CCL4, CXCL10* | 1.01E-03 |
| Pathway | hsa04512:ECM-receptor interaction | 7 | *COL4A2, CD36, COL4A1, TNC, COL5A2, HMMR, SPP1* | 1.11E-03 |
| Pathway | hsa04062:Chemokine signaling pathway | 10 | *CXCR4, CXCL13, CCR1, CXCL9, CCL8, CXCL11, STAT1, CCL5, CCL4, CXCL10* | 1.16E-03 |
| Pathway | hsa04610:Complement and coagulation cascades | 6 | *C1QA, C1QB, C3AR1, C5AR1, CFI, C1QC* | 2.75E-03 |

BP = biological process, CC = cellular component, GO = Gene Ontology, MF = molecular function.

**Table 2**

**The top 5 GO terms enriched for the downregulated genes.**

| Category | Term | Gene number | Gene symbol | P-value |
|---|---|---|---|---|
| GO_BP | GO:0008544~epidermis development | 12 | *C1ORF68, FOXQ1, COL17A1, IRF6, POU2F3, KRT15, AHNAK2, LCE2B, KRT31, KRT2, SOX9, ALDH3A2* | 8.98E-07 |
| GO_BP | GO:0007398~ectoderm development | 12 | *C1ORF68, FOXQ1, COL17A1, IRF6, POU2F3, KRT15, AHNAK2, LCE2B, KRT31, KRT2, SOX9, ALDH3A2* | 1.94E-06 |
| GO_BP | GO:0007155~cell adhesion | 17 | *CDHR1, BCAN, EFS, PCDH7, SOX9, NRCAM, IGSF11, COL17A1, LY6D, F5, PKP1, FAT2, CLDN1, CNTN1, CNTN3, CHL1, ALX1* | 5.96E-04 |
| GO_BP | GO:0022610~biological adhesion | 17 | *CDHR1, BCAN, EFS, PCDH7, SOX9, NRCAM, IGSF11, COL17A1, LY6D, F5, PKP1, FAT2, CLDN1, CNTN1, CNTN3, CHL1, ALX1* | 6.05E-04 |
| GO_BP | GO:0030216~keratinocyte differentiation | 5 | *IRF6, POU2F3, AHNAK2, LCE2B, KRT2* | 3.04E-03 |
| GO_CC | GO:0005887~integral to plasma membrane | 24 | *FGFR3, ENPP2, PTPRZ1, CDHR1, F2RL1, CSPG4, MAL, CD1A, PCDH7, SDC4, TRPM1, KCNJ13, NRCAM, SLC1A4, EDNRB, IGF1R, COL17A1, EPHB6, GPR37, GRIA1, SLC1A6, CLDN1, RYR1, DEGS1* | 4.19E-04 |
| GO_CC | GO:0031226~intrinsic to plasma membrane | 24 | *FGFR3, ENPP2, PTPRZ1, CDHR1, F2RL1, CSPG4, MAL, CD1A, PCDH7, SDC4, TRPM1, KCNJ13, NRCAM, SLC1A4, EDNRB, IGF1R, COL17A1, EPHB6, GPR37, GRIA1, SLC1A6, CLDN1, RYR1, DEGS1* | 5.75E-04 |
| GO_CC | GO:0000267~cell fraction | 20 | *PTGS1, EPHX2, MAL, GYG2, PCSK2, EDNRB, IGF1R, GSTM3, CYP39A1, LY6D, PYGL, GRIA1, SLC1A6, GPX3, HSPB2, RYR1, CNTN1, LGI3, RAPGEFL1, DEGS1* | 4.31E-03 |
| GO_CC | GO:0005882~intermediate filament | 7 | *SLC1A4, KRT77, PKP1, KRT15, KRT31, KRT2, GAN* | 6.70E-03 |
| GO_CC | GO:0045111~intermediate filament cytoskeleton | 7 | *SLC1A4, KRT77, PKP1, KRT15, KRT31, KRT2, GAN* | 7.41E-03 |
| GO_MF | GO:0043565~sequence-specific DNA binding | 13 | *PKNOX2, HLF, FOXQ1, POU2F3, GATA3, LHX2, TFAP2B, NR3C2, RXRG, SOX9, NR2E1, ETV5, ALX1* | 5.64E-03 |
| GO_MF | GO:0015267~channel activity | 10 | *SLC1A4, MCOLN3, KCNK7, GRIA1, ANO1, RYR1, PLLP, MAL, TRPM1, KCNJ13* | 8.74E-03 |
| GO_MF | GO:0022803~passive transmembrane transporter activity | 10 | *SLC1A4, MCOLN3, KCNK7, GRIA1, ANO1, RYR1, PLLP, MAL, TRPM1, KCNJ13* | 8.88E-03 |
| GO_MF | GO:0030528~transcription regulator activity | 22 | *KLF5, HLF, BMP2, RXRG, NR3C2, SOX9, PPP1R13L, NR2E1, CDKN1C, PKNOX2, FOXQ1, IRF6, GATA3, PIR, LHX2, POU2F3, BCL11A, TFAP2B, ID4, ETV5, CRYM, ALX1* | 1.65E-02 |
| GO_MF | GO:0005496~steroid binding | 4 | *PYGL, NR3C2, RXRG, NR2E1* | 1.74E-02 |

BP = biological process, CC = cellular component, GO = Gene Ontology, MF = molecular function.

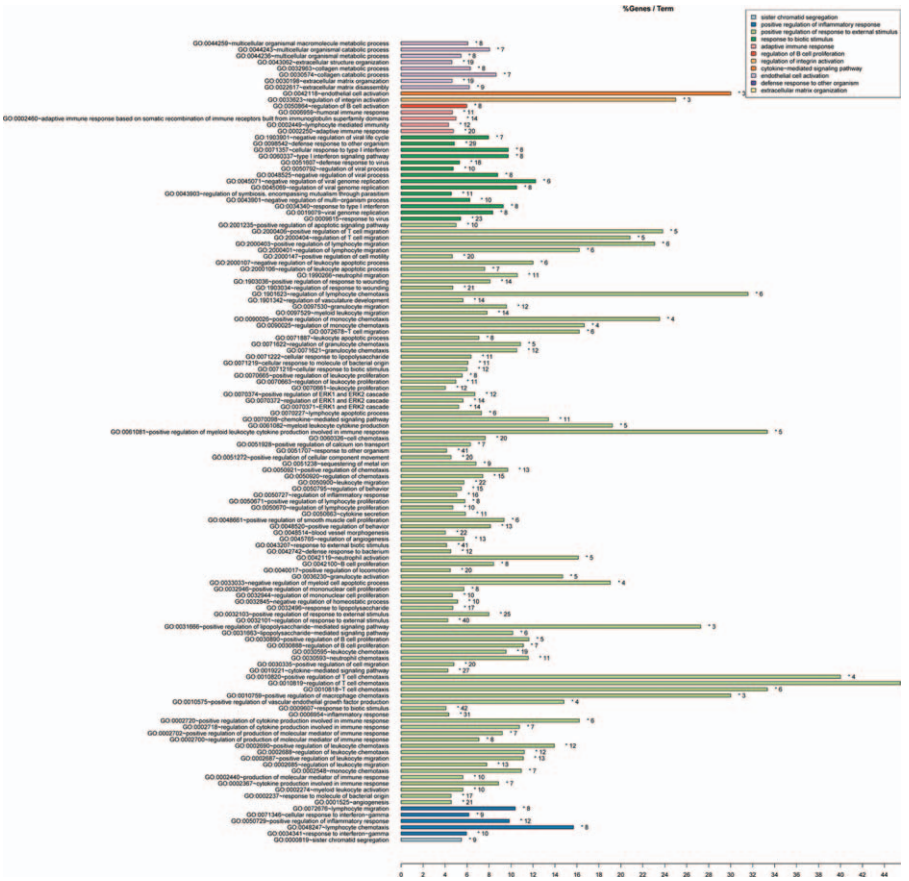


**Figure 2.** The Gene Ontology_biological process (GO_BP) terms enriched for the upregulated genes.

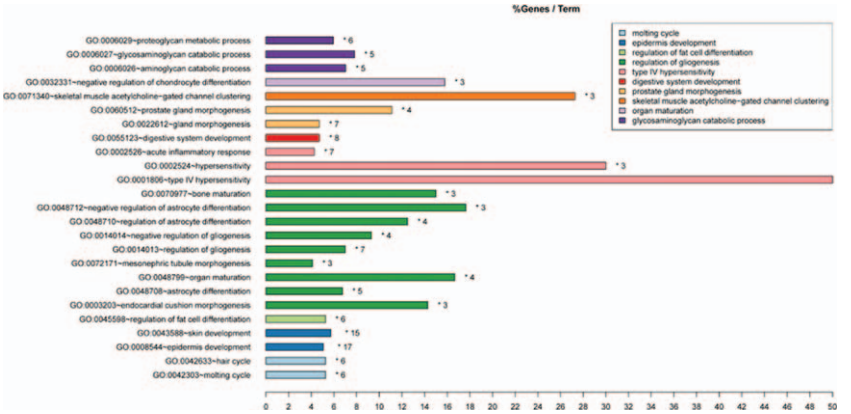


**Figure 3.** The Gene Ontology_biological process (GO_BP) terms enriched for the downregulated genes.

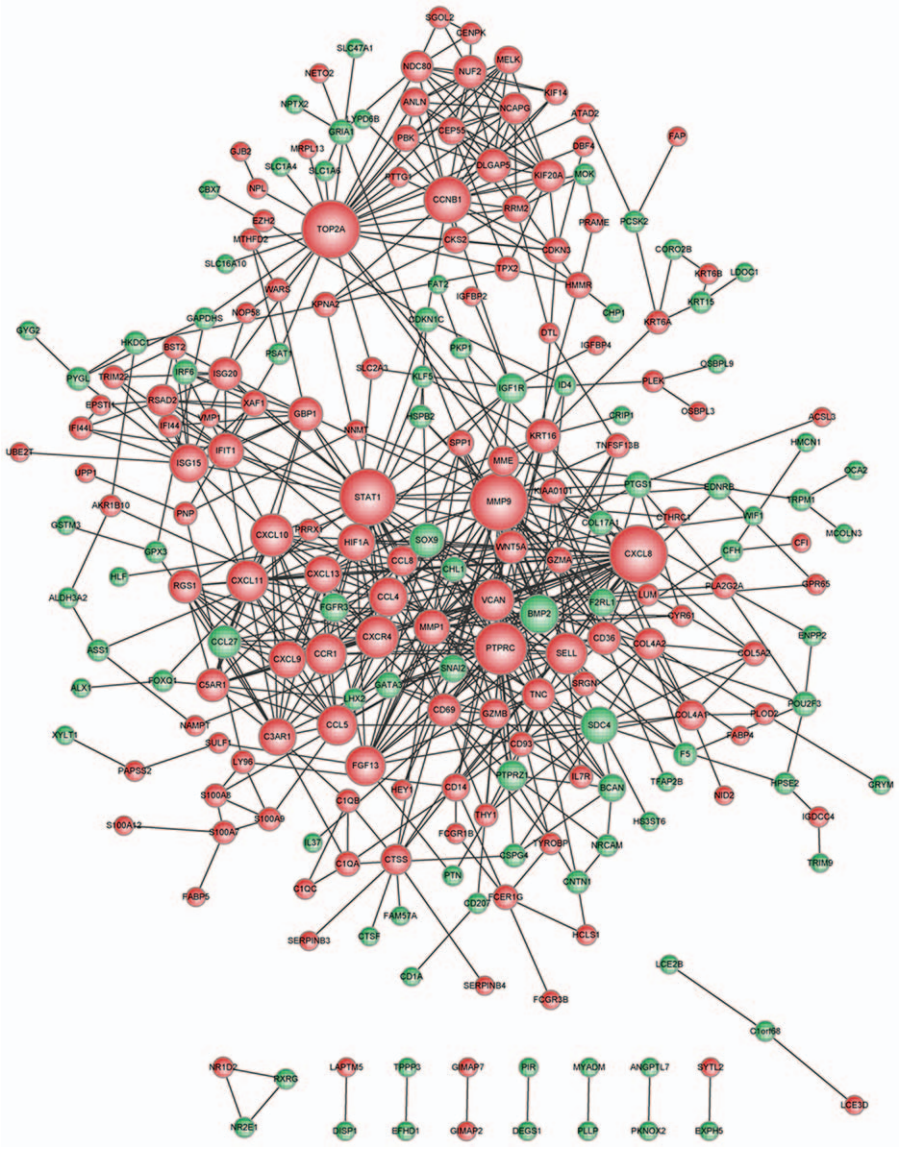


**Figure 4.** The protein-protein interaction network constructed for the differentially expressed genes.

### 3.5. Regulatory network construction

Based on coexpression relationships, the target genes of all prognostic risk lncRNAs were predicted (such as *RP11–361L15.4* targeting *COL17A1*) and the lncRNA-gene regulatory network is shown in Figure 5. Meanwhile, 78 target genes were obtained for 6 (*hsa-miR-100*, *hsa-miR-10a*, *hsa-miR-183*, *hsa-miR-184*, *hsa-miR-22*, and *hsa-miR-375*) of the prognostic risk miRNAs (such as *hsa-miR-375* targeting chemokine (C-C motif) ligand 27, *CCL27*; and *hsa-miR-375* targeting *IGF1R*). In addition, the miRNA gene regulatory network is shown in Figure 6.

### 3.6. Validation of several DEGs by real-time RT-PCR

Aiming to further verify the expression patterns of selected DEGs, real-time PCR, which allows quantitative analysis of mRNA expression, was applied. A total of 5 primary melanoma samples and 10 benign nevus samples (control) were obtained from 15 additional persons who were not part of the microarray analysis. As shown in Figure 7, the overall direction of differential expression was consistent. Among 5 genes assayed, 2 genes (*STAT1* and *IGF1R*) yielded concordance with the results of microarray analysis (*STAT1* was significantly upregulated and *IGF1R* was significantly downregulated). The other 2 genes (*CXCL8* and *CCL27*) showed no significant changes and *COL17A1* showed a slight change in the opposite direction, which may be affected by the small sample size.

## 4. Discussion

In the current study, a total of 382 DEGs were identified in primary melanoma samples, including 206 upregulated genes and 176 downregulated genes. In the PPI network, CXCL8 (degree = 29) and STAT1 (degree = 28) had higher degrees and could interact with each other. Previous study demonstrates that *CXCL8* expression acts in mediating multiple cellular phenotypes related to growth and metastasis of melanoma.[37] *CXCL8* can cause cell proliferation and angiogenesis through both receptors (chemokine [C-X-C motif] receptor 1, *CXCR1*; and chemokine [C-X-C motif] receptor 2, *CXCR2*) and mediate invasive and migration of melanoma cells via *CXCR2*; hence, blocking the CXCL8 signaling axis may be a promising therapeutic method for metastatic melanoma.[38–40] The STAT1/STAT3 balance in tumor cells and host lymphocytes is significantly regulated by interferonα2b, and the ratio of pSTAT1/pSTAT3 in tumor cells at baseline may be a promising predictor of therapeutic effect in patients with cutaneous melanoma.[41] *STAT1* is speculated to be involved in the progression of late-stage melanoma; thus, interference with the *STAT1* pathway may have curative effects for patients with late-

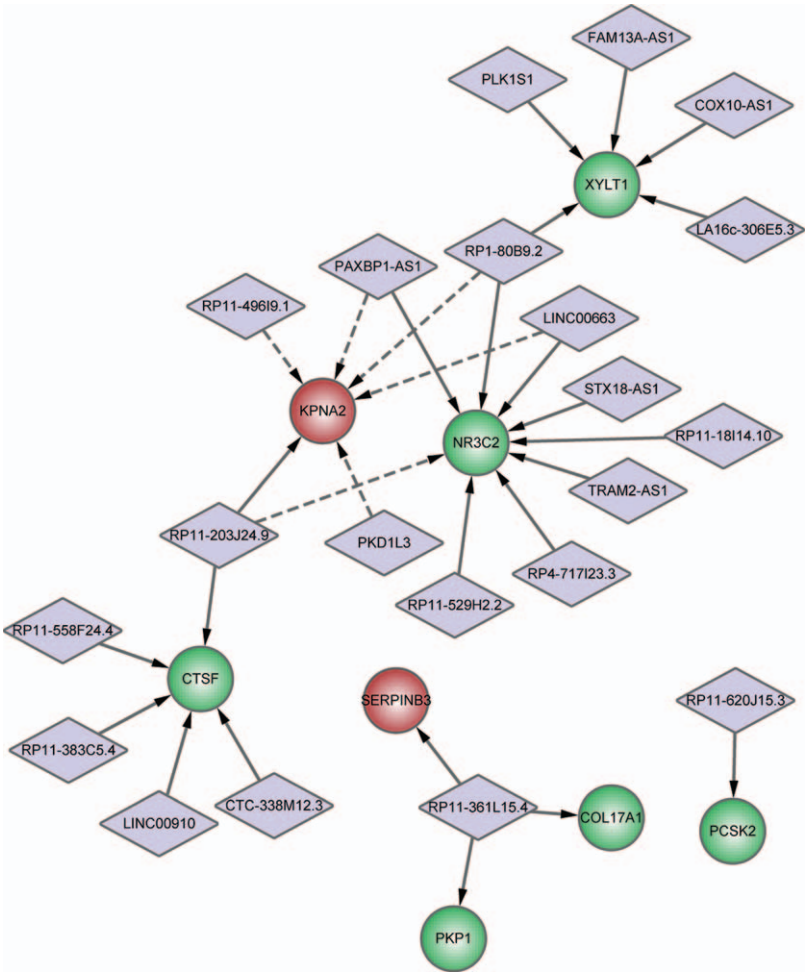


**Figure 5.** The regulatory network of prognostic risk long noncoding RNAs (lncRNAs) and their target genes. Purple diamonds represent lncRNAs. Red and green circles stand for upregulated and downregulated genes, respectively. Dotted lines indicate negative correlation between lncRNAs and target genes.

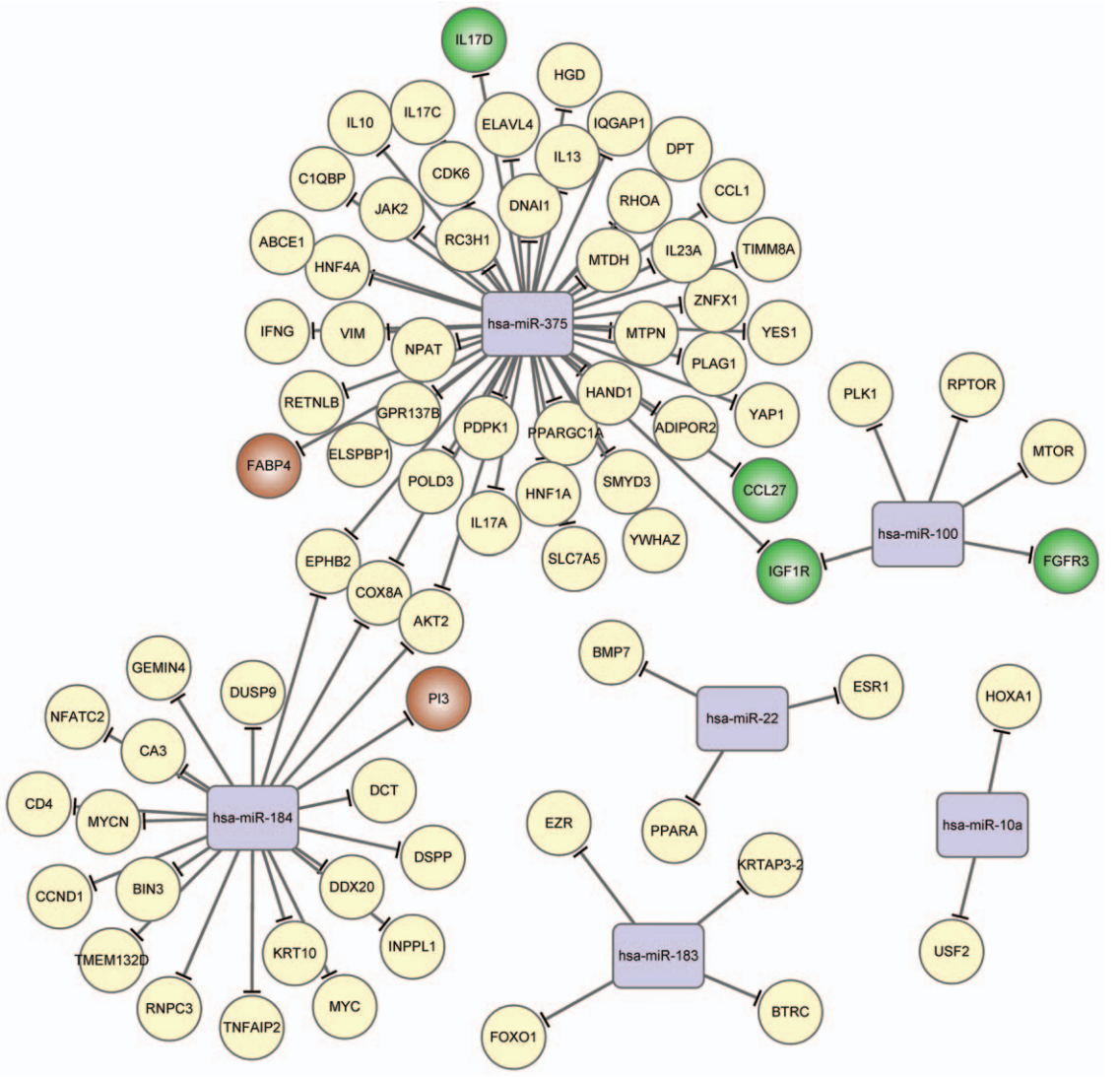


**Figure 6.** The regulatory network of prognostic risk microRNAs (miRNAs) and their target genes. Purple rectangles represent miRNAs. Red and green circles stand for upregulated and downregulated genes, respectively.

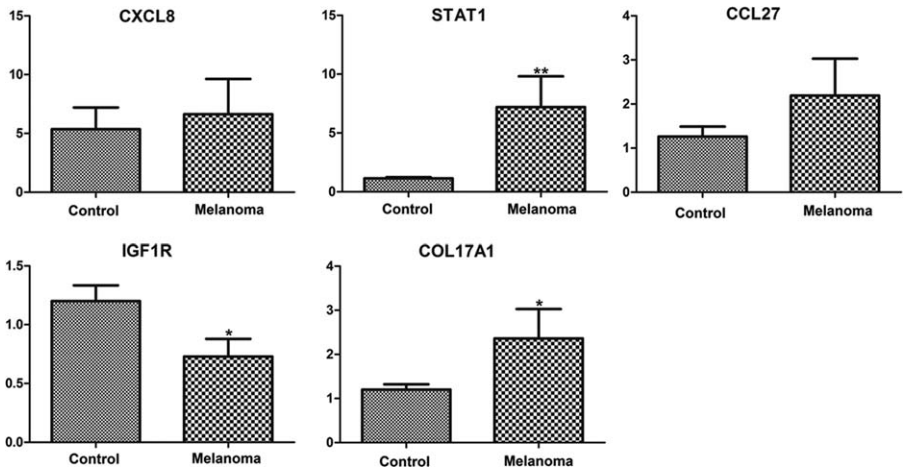


**Figure 7.** Real-time reverse transcription polymerase chain reaction (RT-PCR) validation of several identified differentially expressed genes. Control indicates benign nevus samples. Melanoma indicates primary melanoma samples. $^{*}$, $P < .05$; $^{**}$, $P < .01$.

stage melanoma.[42] Real-time PCR in the present study had further confirmed that *STAT1* was overexpressed in melanoma. Collectively, we suggested that *CXCL8* and *STAT1* might function in the development and progression of melanoma via interacting with each other. More experimental investigations are required to verify the role of *CXCL8* and *STAT1* and their interaction.

Survival analysis showed that 21 DEGs, 55 lncRNAs, and 32 miRNAs were found to be associated with prognosis. Furthermore, several regulatory relationships were found in the miRNA gene regulatory network (such as *hsa-miR-375* targeting *CCL27* and *hsa-miR-375* targeting *IGF1R*). Ectopic expression of *miR-375* inhibits motility, proliferation, and invasion, and causes shape changes of melanoma cells, indicating that *miR-375* may be implicated in the development and progression of melanomas.[43]*CCL27* and chemokine (C-C motif) receptor 10 (*CCR10*) in human melanomas may help tumor cells to grow, invade tissue, evade immune response, and spread to lymph nodes.[44] Villanueva et al[45] found that IGF1R/phosphatidylinositol 3-kinase (PI3K) signaling is boosted in resistant melanomas, and combined treatment with mitogen-activated protein kinase and *IGF1R/PI3K* inhibitors can overcome acquired resistance to *BRAF*.[46] Uveal melanomas express *MET*, epidermal growth factor receptor, hepatocyte growth factor, and *IGF-1R*, in addition, epidermal growth factor and *IGF1* may have correlation with the pathogenesis of uveal melanoma.[47] In human melanoma, *IGF1* is found to be a novel growth factor for autocrine-driven proliferation in vitro, and IGF1-IGF1R autocrine pathway may be a therapeutic target for the disease.[48] Therefore, *CCL27* and *IGF1R* targeted by *hsa-miR-375* might also associate with the mechanisms of melanoma. The validation of real-time PCR strongly supported the downregulation of *IGF1R*; the expression of *CCL27* and the role of *hsa-miR-375* in melanoma should be further verified by multiple approaches.

Functional enrichment analysis showed that *COL17A1* was enriched in epidermis development. Fibrillar collagen can act as a powerful barrier to viral distribution and matrix-modifying treatments significantly promote the treatment response of human melanoma.[49] Van Kempen et al[50] firstly report that the proangiogenic effect of type I collagen in a de novo tumor model keeps ahead of the development of invasive melanomas. Collagen types XVII, which is a type of epithelial adhesion protein, has correlation with the epidermal basement membrane and plays a key functional role.[51] Moreover, collagen XVII endodomain accumulated in melanocytic tumors is related to malignant transformation and regulates antibody-induced melanoma apoptosis.[52] The majority of evidence has demonstrated that microarrays are invaluable discovery tools with acceptable reliability for genome-wide gene expression screening.[53] However, microarray data alone often lack the degree of specificity, which was needed to make accurate biological conclusions.[54] Genes found to be differentially expressed by microarray analysis should be validated with other well-established technologies for analyzing gene expression, such as real-time reverse transcription (RT)-PCR. Dallas et al[55] reported a 13% to 16% nonconcordance between quantitative reverse transcription PCR (qRT-PCR) and microarray data, which may emphasize the continuing requirement for caution in validation of gene expression data. In this study, the qRT-PCR results showed a slight change in the opposite direction of *COL17A1*. However, considering the role of collagen XVII demonstrated by the previous study,[52] we still suggested the significant role of *COL17A1* in melanoma. On the contrary, in the lncRNA-gene regulatory network, *COL17A1* was targeted by the lncRNA *RP11–361L15.4*, indicating that *COL17A1* targeted by *RP11-361L15.4* might function in the pathogenesis of melanoma. More experiments should be performed to explore the expression of *COL17A1* and the association of *RP11-361L15.4* and *COL17A1*.

In conclusion, we have conducted an in-depth bioinformatics analysis and identified a total of 382 DEGs in primary melanoma samples. *CXCL8* and *STAT1* might function in the development and progression of melanoma via interacting with each other. *CCL27* and *IGF1R* targeted by *hsa-miR-375* might also associate with the mechanisms of melanoma. Furthermore, *COL17A1* targeted by *RP11-361L15.4* might function in the pathogenesis of melanoma. A subset of selected genes was validated by real-time RT-PCR and several genes were confirmed their differential expression in small size of melanoma samples. More in vivo or in vitro experiments with larger sample size were required to further explore the exact roles of the identified genes, miRNAs and lncRNAs in melanoma, which may represent potential tumor markers or drug targets.

## References


[1] Queirolo P, Pfeffer U. Metastatic melanoma: how research can modify the course of a disease. Cancer Metastasis Rev 2017;36:3–5.

[2] Agaimy A, Specht K, Stoehr R, et al. Metastatic malignant melanoma with complete loss of differentiation markers (undifferentiated/dedifferentiated melanoma): analysis of 14 patients emphasizing phenotypic plasticity and the value of molecular testing as surrogate diagnostic marker. Am J Surg Pathol 2016;40:181–91.

[3] Reddy BY, Miller DM, Tsao H. Somatic driver mutations in melanoma. Cancer 2017;123:2104–17.

[4] Weinstein D, Leininger J, Hamby C, et al. Diagnostic and prognostic biomarkers in melanoma. J Clin Aesthet Dermatol 2014;7:13–24.

[5] Kim M, Gans JD, Nogueira C, et al. Comparative oncogenomics identifies NEDD9 as a melanoma metastasis gene. Cell 2006;125:1269–81.

[6] Zeng Q, Cao K, Liu R, et al. Identification of TDP-43 as an oncogene in melanoma and its function during melanoma pathogenesis. Cancer Biol Ther 2017;18:8–15.

[7] Levy C, Khaled M, Iliopoulos D, et al. Intronic miR-211 assumes the tumor suppressive function of its host gene in melanoma. Mol Cell 2010;40:841–9.

[8] Giles KM, Brown R, Ganda C, et al. microRNA-7-5p inhibits melanoma cell proliferation and metastasis by suppressing RelA/NF-(B. Oncotarget 2016;7:31663–80.

[9] Khaitan D, Dinger ME, Mazar J, et al. The melanoma- upregulated long noncoding RNA SPRY4-IT1 modulates apoptosis and invasion. Cancer Res 2011;71:3852–62.

[10] Li R, Zhang L, Jia L, et al. Long non-coding RNA BANCR promotes proliferation in malignant melanoma by regulating MAPK pathway activation. PLoS One 2014;9:e100893.

[11] Chen L, Yang H, Xiao Y, et al. Lentiviral-mediated overexpression of long non-coding RNA GAS5 reduces invasion by mediating MMP2 expression and activity in human melanoma cells. Int J Oncol 2016;48:1509–18.

[12] He F, Yoo S, Wang D, et al. Large- scale atlas of microarray data reveals the distinct expression landscape of different tissues in Arabidopsis. Plant J 2016;86:472–80.

[13] Zhang H, Yu Z, He J, et al. Identification of the molecular mechanisms underlying dilated cardiomyopathy via bioinformatic analysis of gene expression profiles. Exp Ther Med 2017;13:273–9.

[14] Wang S, Fan W, Wan B, et al. Characterization of long noncoding RNA and messenger RNA signatures in melanoma tumorigenesis and metastasis. PLoS One 2017;12:e0172498.

[15] Eriksson J, Le Joncour V, Nummela P, et al. Gene expression analyses of primary melanomas reveal CTHRC1 as an important player in melanoma progression. Oncotarget 2016;7:15065–92.

[16] Gautier L, Cope L, Bolstad BM, et al. Affy—analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 2004;20:307–15.

[17] Harrow J, Frankish A, Gonzalez JM, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res 2012;22:1760–74.

[18] Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.
[19] Smyth GK. Limma: Linear Models for Microarray Data. Bioinformatics and Computational Biology Solutions Using R and Bioconductor. 2005; New York, NY: Springer; 397–420.
[20] Carbon S, Ireland A, Mungall CJ, et al. AmiGO: online access to ontology and annotation data. Bioinformatics 2009;25:288–9.
[21] Kanehisa M, Araki M, Goto S, et al. KEGG for linking genomes to life and the environment. Nucleic Acids Res 2008;36(suppl 1):D480–4.
[22] Huang DW, Sherman BT, Tan Q, et al. DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res 2007;35(suppl 2): W169–75.
[23] Szklarczyk D, Franceschini A, Kuhn M, et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res 2011;39(suppl 1):D561–8.
[24] Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498–504.
[25] Klein JP, Gerster M, Andersen PK, et al. SAS and R functions to compute pseudo-values for censored data regression. Comput Methods Programs Biomed 2008;89:289–300.
[26] Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. New York, NY: Springer Science & Business Media; 2000.
[27] Therneau T. A Package for Survival Analysis in S, Version 2.38. 2015. Available at: http://cranr-projectorg/web/packages/snow/indexhtml. Accessed April 19, 2016.
[28] Akazawa K, Nakamura T, Palesch Y. Power of logrank test and Cox regression model in clinical trials with heterogeneous samples. Stat Med 1997;16:583–97.
[29] John B, Enright AJ, Aravin A, et al. Human microRNA targets. PLoS Biol 2004;2:e363.
[30] Wang X, El Naqa IM. Prediction of both conserved and nonconserved microRNA targets in animals. Bioinformatics 2008;24:325–32.
[31] Krek A, Grün D, Poy MN, et al. Combinatorial microRNA target predictions. Nat Genet 2005;37:495–500.
[32] Kertesz M, Iovino N, Unnerstall U, et al. The role of site accessibility in microRNA target recognition. Nat Genet 2007;39:1278–84.
[33] Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 2005;120:15–20.
[34] Lewis BP, Shih I-H, Jones-Rhoades MW, et al. Prediction of mammalian microRNA targets. Cell 2003;115:787–98.
[35] Xiao F, Zuo Z, Cai G, et al. miRecords: an integrated resource for microRNA-target interactions. Nucleic Acids Res 2009;37(suppl 1): D105–10.
[36] Dweep H, Sticht C, Pandey P, et al. miRWalk-database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform 2011;44:839–47.
[37] Wu S, Singh S, Varney ML, et al. Modulation of CXCL-8 expression in human melanoma cells regulates tumor growth, angiogenesis, invasion, and metastasis. Cancer Med 2012;1:306–17.
[38] Gabellini C, Trisciuoglio D, Desideri M, et al. Functional activity of CXCL8 receptors, CXCR1 and CXCR2, on human malignant melanoma progression. Eur J Cancer 2009;45:2618–27.
[39] Varney ML, Johansson SL, Singh RK. Distinct expression of CXCL8 and its receptors CXCR1 and CXCR2 and their association with vessel density and aggressiveness in malignant melanoma. Am J Clin Pathol 2006;125:209–16.
[40] Varney ML, Li A, Dave BJ, et al. Expression of CXCR1 and CXCR2 receptors in malignant melanoma with different metastatic potential and their role in interleukin-8 (CXCL-8)-mediated modulation of metastatic phenotype. Clin Exp Metastasis 2003;20:723–31.
[41] Wang W, Edington HD, Rao UN, et al. Modulation of signal transducers and activators of transcription 1 and 3 signaling in melanoma by high-dose IFNα2b. Clin Cancer Res 2007;13:1523–31.
[42] Schultz J, Koczan D, Schmitz U, et al. Tumor-promoting role of signal transducer and activator of transcription (Stat) 1 in late-stage melanoma growth. Clin Exp Metastasis 2010;27:133–40.
[43] Mazar J, DeBlasio D, Govindarajan SS, et al. Epigenetic regulation of microRNA-375 and its role in melanoma development in humans. FEBS Lett 2011;585:2467–76.
[44] Simonetti O, Goteri G, Lucarini G, et al. Potential role of CCL27 and CCR10 expression in melanoma progression and immune escape. Eur J Cancer 2006;42:1181–7.
[45] Villanueva J, Vultur A, Lee JT, et al. Acquired resistance to BRAF inhibitors mediated by a RAF kinase switch in melanoma can be overcome by cotargeting MEK and IGF-1R/PI3K. Cancer Cell 2010;18: 683–95.
[46] Yeh A, Bohula E, Macaulay V. Human melanoma cells expressing V600E B-RAF are susceptible to IGF1R targeting by small interfering RNAs. Oncogene 2006;25:6574–81.
[47] Topcu-Yilmaz P, Kiratli H, Saglam A, et al. Correlation of clinicopathological parameters with HGF, c-Met, EGFR, and IGF-1R expression in uveal melanoma. Melanoma Res 2010;20:126–32.
[48] Molhoek KR, Shada AL, Smolkin M, et al. Comprehensive analysis of RTK activation in human melanomas reveals autocrine signaling through IGF-1R. Melanoma Res 2011;21:274.
[49] McKee TD, Grandi P, Mok W, et al. Degradation of fibrillar collagen in a human melanoma xenograft improves the efficacy of an oncolytic herpes simplex virus vector. Cancer Res 2006;66:2509–13.
[50] Van Kempen LC, Rijntjes J, Mamor- Cornelissen I, et al. Type I collagen expression contributes to angiogenesis and the development of deeply invasive cutaneous melanoma. Int J Cancer 2008;122: 1019–29.
[51] Bechetoille N, Haftek M, Staquet M-J, et al. Penetration of human metastatic melanoma cells through an authentic dermal-epidermal junction is associated with dissolution of native collagen types IV and VII. Melanoma Res 2000;10:427–34.
[52] Krenács T, Kiszner G, Stelkovics E, et al. Collagen XVII is expressed in malignant but not in benign melanocytic tumors and it can mediate antibody induced melanoma apoptosis. Histochem Cell Biol 2012;138: 653–67.
[53] Wang Y, Barbacioru C, Hyland F, et al. Large scale real-time PCR validation on gene expression measurements from two commercial long-oligonucleotide microarrays. BMC Genomics 2006;7:59.
[54] Troyanskaya OG. Putting microarrays in a context: integrated analysis of diverse biological data. Brief Bioinform 2005;6:34–43.
[55] Dallas PB, Gottardo NG, Firth MJ, et al. Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR-how well do they correlate? BMC Genomics 2005;6:59.