Identification and Validation of Key Genes in Giant Congenital Melanocytic Nevi by RNA Sequencing Analysis

Xialin Cheng
Xiang'an Hospital of Xiamen University
Tao Dai
Henan University of Science and Technology
Wu Bao
Xiang'an Hospital of Xiamen University
Lingxi Chen
Xiang'an Hospital of Xiamen University
Zexin Zhang
Xiang'an Hospital of Xiamen University
Chiyu Jia (jiachiyu@qq.com)
Xiang'an Hospital of Xiamen University

Research Article

Keywords: Giant congenital melanocytic nevi (GCMNs), pathogenes, RNA-seq, Cytoscape software

Posted Date: December 2nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1120605/v1

Abstract

Background

Giant congenital melanocytic nevi (GCMNs) are melanotic lesion present at birth, which often cause severe psychological and financial burden to patients and their families.However, the pathogenes is still unclear.

Objective

We aim to identify key genes and biological processes that related to the development of GCMN.

Methods

We sequenced ten pairs of GCMN tissues and adjacent normal tissues by high-throughput RNA-seq, then used GO and KEGG analysis to find inportment pathways, and used MCODE,Cluego and Cytohubba plugin of Cytoscape software to identify hub genes.

Results

A total of 1163 differentially expressed genes were identified. 29 BPs, 18 CCs, and 17 MFs were significantly enriched in GO analysis and no pathway was significantly enriched in KEGG analysis. PPI Visual Network consisted of 779 nodes and 2359 edges,which was be divided into 25 functional modules by MCODE. We discovered most of the hub genes were located in module 5,and the top 3 hub genes (*PTGS2,EGF,SOX10*) in module 5 were involved in GO and KEGG enrichment pathways –"Arachidonic acid metabolism","glycosaminoglycan biosynthetic process","developmental pigmentation" respectively.

Conclusion

PTGS2, EGF and *SOX10* are thought to be the three most important hub genes and may play essential roles in the development of GCMN.

1 Introduction

Giant congenital melanocytic nevi (GCMNs) are benign melanocytic skin tumors that originate from the neural crest and present at birth[1]. Generally, GCMN is defined by the maximal diameter of its projected adult size (PAS)>40cm^2. The clinical importance of GCMNs is the increased risk of melanomas[3] and neurocutaneous melanosis (NCM)[4],both of which are associated with extremely poor prognosis. The great impairment on appearance and severe complications often cause serious psychological and economic burden to patients and their families. Currently, the mainstay of GCMN treatment is surgical excision. However, the therapeutic effect is not ideal due to the complexity of the disease. Its root cause lies in the the lack of understanding of the disease,therefore, further exploring the pathogenesis of the disease is essential.

With the development of the sequencing technology, bioinformatics analysis provides a new approach to the deep investigation of complicated diseases. Several scholars studied the molecular characterization of GCMN patients by next-generation sequencing and found approximately 80% of patients harbor somatic mutations. *NRAS,BRAF,KRAS,APC,MET* are common mutated genes and are involved in *RAS/RAF/MEK/ERK, PI3K/Akt* and *JAK/STAT* pathways[5,6]. Rouille, et al. [7]reported that inhibition of *MEK* and/or *PI3K* pathways could significantly reduce the proliferative ability of melanocytes in vivo and in vitro, which provides a new direction for the treatment of GCMN. Other molecular events may also intervene in the occurrence of GCMN[8]. Therefore, it is crucial to further elucidate the underlying mechanisms of GCMNs and pave the way for novel therapeutic targets.

The mRNA contains abundant genetic information and is necessary for protein translation, however the molecular function and clinical value of transcriptomic expression profiles in GCMN remains unclear. RNA sequencing (RNA-seq) presents a convenient method to obtain comprehensive RNA expression information of specific tissues in a certain state[9].In this study, we applied the RNA-Seq technology to identify all differentially expressed genes (DEGs) between GCMN tissues and paired normal tissues.Gene Ontology(GO),Kyoto Gene and Genomic Encyclopedia(KEGG) pathway analyses and protein-protein interaction (PPI) network analysis were performed in order to explore key genes and molecular processes.

2 Materials And Methods

2.1 Sample collection and preparation

Ten pairs of GCMN and adjacent normal tissues for RNA-seq and Hematoxylin-Eosin staining(H&E staining) were collected from January 2020 to June 2020, and paired normal tissues were taken from surgically dissected tissues 5 mm away from the lesion edge. All patients met the diagnostic criteria for GCMN, and none of them had combined melanoma and neurocutaneous melanosis at the time of treatment. The dissected skin samples were rapidly washed with physiological normal saline and divided into two parts, one of which was instantly frozen in liquid nitrogen and thereafter stored at -80C° for RNA extraction and the other was fixed in 10% neutral buffered formalin and paraffin-embedded, and stained with hematoxylin and eosin. Then, the samples were observed under light microscope for picture collection and analysis to ensure the quality of samples and the effectiveness of the results(Fig. 1A.1B). Another 14 GCMN tissues and 14 adjacent normal tissues were obtained from Xiang'an Hospital for use in RT-qPCR.

All subjects/ subjects' guardians gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Xiang'an Hospital of Xiamen University.

2.2 RNA extraction and Illumina sequencing

Total RNA was extracted using RNeasy Fibrous Tissue Mini Kit (Qiagen, Valencia CA,USA) according to the manufacturer's instructions, The purity and integrity of RNA was checked using the NanoPhotometer spectrophotometer (IMPLEN, CA, USA) and the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA) respectively. The RNA integrity number (RIN) values of all samples were above 7.

A total amount of 1 μg RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext UltraTM RNA Library Prep Kit for Illumina (NEB, USA). mRNA was purified from total RNA using poly-T oligo-attached magnetic beads, then clean data (clean reads) were obtained by removing reads containing adapter, reads containing ploy-N and low quality reads from raw data.

2.3 Data Analysis

Index of the reference genome was built and paired-end clean reads were aligned to the reference genome using Hisat2 v2.0.5[10]. After data homogenization and quality control, FPKM(fragments per kilobase) of each gene was calculated based on the length of the gene and reads count mapped to this gene. Genes with fold change > 2 and adjusted P-value <0.01 (P-values were adjusted using the Benjamini and Hochberg's approach to control the false discovery rate) determined by DESeq2 were considered as differentially expressed[11]

2.4 GO and KEGG enrichment analysis of differentially expressed genes

To understand the biological function of DEGs, GO term and KEGG (http://www.genome.jp/kegg/) pathway analysis was implemented by the clusterProfiler R package[12]. Gene Ontology (GO) can annotate target genes and perform analysis from three levels, biological processes, cellular components and molecular function[13]. KEGG (Kyoto Encyclopedia of Genes and Genomes) is a database resource for understanding high-level functions and utilities of the biological system[14]. A corrected P value less than 0.05(padj<0.05) was considered statistical significantly.

2.5 PPI Network Construction and Hub Gene Identification

A PPI network of DEGs was generated based on the STRING database (https://string-db.org/) with a combined score > 0.4 and further analyzed and visualized by Cytoscape software (version 3.7.1). The Cytoscape app MCODE were used to divide clusters of the above network with the default parameters: Degree Cutoff = 2, Node Score Cutoff = 0.2, K-Core = 2 and Max.Depth = 100.Next ,the ClueGO plugin of Cytoscape was employed for Biological process analysis upon Cluster 5, with a level of significance set for 0.05 and Kappa Score of 0.4.with a score value greater than 5. Then we explored hub genes using 12 scoring methods(MCC, DMNC, MNC, Degree, EPC, BottleNeck, EcCentricity, Closeness, Radiality, Betweenness, Stress, Clustering Coefficient available in another Cytoscape plugin cytoHubba. The top 10 genes from each of these algorithms were choosen. Any gene that appeared more than one time and also presented in top 6 MCODE clusters was considered as a hub gene.

2.6 RT-qPCR analysis

RT-qPCR(Reverse transcription-quantitative polymerase chain reaction) was performed to confirm the expression of hub genes in GCMN tissues and adjacent normal tissues. Total RNA was extracted from each sample using TRIzol reagent (Qiagen,Valencia, CA,U.S.A.) and first-strand cDNA was synthesized with the cDNA Synthesis kit (TOYOBO, Osaka, Japan). Next, amplification was performed using the Power SYBR Green real-time PCR Master Mix (TOYOBO, Osaka, Japan). All procedures were conducted in three technical replicates. The relative expression of each target gene was normalized to GAPDH,and data were analyzed by the $2 - \Delta\Delta Ct$ method.

2.7 Statistical analysis

All statistical analyses were conducted using GraphPad Prism 7.0 and the results were expressed as the means ± standard deviation. Student's t-test was performed to compare differences between two groups,The $\chi 2$ test and Fisher's exact test were used for GO and KEGG analysis. P value less than 0.05 was considered statistically significant.

3 Results

3.1 Identification of DEGs

To systematically analyze that mRNAs may play a role in GCMN, we performed RNA-Seq experiments on 10 pairs of GCMN tissues and their matched adjacent normal tissues. Approximately 45 million reads were sequenced for each sample. In this study, a total of 1163 differentially expressed genes (DEGs) were identified with the cut-off criterion as $p < 0.01$ and $| log2fc | > 1$, among which 519 genes were up-regulated and 644 genes were down-regulated(Fig. 1C).

3.2 GO and KEGG enrichment analysis

To discern the potential functions of 1163 DEGs, Gene Ontology(GO) and Kyoto Encyclopaedia of Genes and Genomes(KEGG) pathway analysis were carried out. Through GO analysis, we identified 29 BPs, 18 CCs, and 17 MFs as being significantly enriched. For the biological process(BP) category, the DEGs were mainly involved in epidermis development, keratinocyte differentiation, developmental pigmentation, glycosaminoglycan biosynthetic process. For the cell component (CC) category, the DEGs were markedly located in intermediate filament, melanosome, collagen trimer, extracellular matrix. For the molecular function (MF) category, the DEGs were primarily related to structural constituent of epidermis, metal ion transmembrane transporter activity, peroxidase activity, actin binding. The top 10 significantly enriched terms of BP,CC,MF are shown in Fig. 2A.

The result of KEGG analysis suggested that no pathway was significantly enriched and the overall pathways obtained (padj>0.05)mainly associated with cancer, metabolism and infection, such as JAK-STAT signaling pathway, Small cell lung cancer, Arachidonic acid metabolism, Tyrosine metabolism, Malaria and Human papillomavirus infection. Scatter plot for top 20 pathways in KEGG enrichment of DEGs are shown in Fig. 2B.

3.3 PPI Network and Hub Gene Identification

We used Cytoscape software to visualize the PPI network, which consists of 779 nodes and 2359 edges. Cytoscape plug-in MCODE was applied for the module analysis, and a total of 25 functional modules were obtained from the entire network. Detailed informations of top six modules are listed in Table1.

Table 1
The DEGs in the top sixc module clusters sorted by MCODE score.

	MCODE score	Composition	Genes
Module 1	13.000	13nodes 78edges	KRT74,KRT10,KRT83,KRT2,KRT26,KRT77,KRT33B,KRT27,KRT23,KRT36,KRT71,KRT32,KRT25
Module 2	10.762	22nodes 113edges	COL23A1,COL22A1,LCE6A,LCE2D,COLGALT2, LCE2C,LOR,EVPL,COL28A1,CSTA,COL19A1,LCE1A,COL13A1,COL4A5,LCE1F,COL17A1,COL11A2, COL6A5,COL7A1,RPTN,TCHH,COL2A1
Module 3	8.875	17nodes 71edges	RBBP6,KBTBD8,FBXO7,TRIM63,ASB9,NPFFR1, GPR143,LMO7,ASB14,LNX1,EDNRB,TRIM9, CYSLTR2,GAN,LTB4R,P2RY1,FBXW7
Module 4	6.000	7nodes 18edges	KRTAP19-5,KRTAP2-2,KRTAP19-1, KRTAP9-3,KRTAP4-4,KRTAP19-3,KRTAP9-4
Module 5	5.625	17nodes 45edges	SOX10,GPR37,PAX3,EGF,GPR183,HEBP1,SLC24A5,DCT,PMEL,MITF,CCR8,NPY4R,MLPH,TYR,IL18, PTGS2,C3
Module 6	5.500	13nodes 33edges	DNAH11,DNAAF3,SLC39A10,MT1G,RSPH4A, MT1E,NME8,SOD3,OFD1,MT2A,MT1X,SPAG1, SERPINE2

Subsequently, hub genes were identified using 12 types of calculation methods in cytoHubba(Fig. 3A). detailed information about these hub genes is presented in Table2. We discovered most of the hub genes were located in module 5, thus Module5 was chosen for further evaluation through hierarchical clustering analysis (Fig. 3B)and Cluego analysis(Fig. 3C). According to the result, Module5 was primarily involved in melanin biosynthetic process,and most genes were upregulated. The top 3 hub genes (*PTGS2,EGF,SOX10*) with the most frequent occurrences were selected for further experimental verification.

Table 2
Six hub genes identified by MCODE and Cytohubba

Gene symbol	Gene description	log2FC	occurrence*
PTGS2	prostaglandin-endoperoxide synthase 2	-2.348798977	9
EGF	epidermal growth factor	-2.549945398	8
SOX10	SRY-box 10	3.102842853	8
C3	complement C3	1.527467154	6
KRT27	keratin 27	5.710775079	3
EVPL	envoplakin	-2.260789619	3
EDNRB	endothelin receptor type B	3.198479721	3
TYR	tyrosinase	3.644471127	3
MITF	melanogenesis associated transcription factor	2.285262595	3
KRT23	keratin 23	-2.875258674	2
occurrence*: the times of ranking in the top 10 in 12 types of calculation methods			

3.4 Validation of hub genes using RT-qPCR

To validate the results observed in RNA-seq, we undertook RT-qPCR to quantify the relative expression of 3 hub genes selected above. Results show that compared with adjacent normal tissues, *SOX10* was significantly up-regulated and *PTGS2,EGF* were significantly down-regulated in GCMN tissues, which is consistent with the RNA-Seq data (Figure.4).The expression patterns of selected genes in qRT-PCR were consist with RNA-seq.

4 Discussion

The lesion area of GCMN becomes larger and the number of satellite nevi increases as the child grows. In addition,relevant complications may seriously affect the quality of life and even longevity[15]. These problems cannot be solved completely by surgical removal and new medical options based on the key disease-associated genes seem to be promising. Previously published studies mainly focused on the association between somatic mutations and clinical characteristics. In this study, we employed RNA-Seq to screen out 1163 DEGs, and found several novel biological processes and key genes that have not been investigated in GCMN.

GO classification and enrichment analysis revealed that several biological processes related to melanin synthesis were significantly enriched and the genes involved were up-regulated, which is consistent with the increased number of melanocytes and darker skin color of lesions[16].In addition, primary biological processes of the DEGs were involved in the regulation of epidermis development, skin development, and keratinocyte differentiation. The thickened cuticle was also observed in diseased tissues with the HE staining results obtained, indicating that keratinocytes are strikingly different between pathological and normal tissues in terms of quantity and gene expression. A previous study suggested that melanocytes transfered melanin to keratinocytes through melanosomes, and the number of melanosomes in keratinocytes determined the degree of skin pigmentation[17]. We hypothesized that the two types of cells interact through direct contacting and cytokines secreting, and jointly participate in the progression of disease. Extracellular matrix (ECM) can provide the tissue with the mechanical support and mediate the cell-microenvironment interactions. Meanwhile, the extracellular structural changes may play facilitating roles in tumor progression[18]. Our results showed that GO analysis were linked to extracellular matrix synthesis, including collagen trimer and glycosaminoglycan biosynthetic process. ECM not only directly regulates cellular processes, such as adhesion, migration, proliferation and cell differentiation[19], but also acts in binding, guiding and regulating signal molecules[20]. Tannous, et al. [15]demonstrated that the structure of ECM in the lesion site was disordered and melanocytes were extensively infiltrated. This consistency implies that ECM is involved in the formation and development of GCMN by affecting the behavior of melanocytes.

Although no pathway was significantly enriched in KEGG analysis, pathway“Arachidonic acid metabolism”(padj=0.277) attracted our attention. Gledhill, et al. [21]reported that arachidonic acid metabolism pathway exists in both keratinocytes and melanocytes, which can stimulate melanocyte proliferation and melanin production by secreting PGE2 and other regulatory factors. Furthermore, the arachidonic acid metabolic pathway has been proved to be activated or inhibited in some pigmented diseases, such as vitiligo[22] and solar lentigo[23]. The role of over-expression of this pathway in GCMN lesions need to be further explored.

From the PPI network analysis, 10 genes were recognized as hub genes. Among them, *PTGS2, EGF, SOX10* appeared most frequently in the top ten rankings and were involved in the mentioned biological processes. *PTGS2*, also known as *COX-2*, serves as an important role in prostaglandin production from arachidonic acid, and is highly expressed in inflammatory sites and melanoma[24]. Both animal experiments and cell experiments have confirmed that the growth, migration and invasion ability of melanoma cells were inhibited after the knockout of *PTGS2* gene[24,25]. However ,it should be noted that compared with the normal group, *PTGS2* gene was down-regulated in the diseased group in our study.

Epidermal growth factor (*EGF*) is commonly used to treat chronic ulcers of skin, such as diabetic foot ulcers[26] and severe burn injuries[27],as it can promote wound healing by inducing granulation tissue growth, angiogenesis and epidermal remodeling. In recent years, some studies have demonstrated that *EGF* has the inhibitory effects of skin pigmentation by blocking the activity of tyrosinase[28]. In addition, *EGF* is a typical anti-inflammatory and antioxidant cytokine, which could also reduce melanin production[26,29].Kim, et al. [30]proposed *EGF*-containing ointment had anti-pigmentation effects through performing a randomized controlled trial. In our study, it has been identified that the expression level of *EGF* was dramatically decreased in the lesions and may be related to local over-pigmentation.

As an essential transcription factor, *Sox10* was the first gene in *SOX* family which was found to be involved in neural crest development[31]. By synergising with *PAX3*, *SOX10* is capable of up-regulating the expression of the *MITF*(melanocyte master transcription factor)[32] and further activating dopachrome tautomerase and tyrosinase[33],the chain reaction can vastly promote melanocyte proliferation and synthesis of melanin. Numerous studies have been reported that *SOX10* is highly expressed in tumors, especially melanoma, and plays an essential role in the proliferation, migration, and invasion of melanoma cells [34].Kaufman, et al. [35]also found that overexpression of *SOX10* in melanocytes significantly accelerated the formation of melanoma. Therefore, the overexpression of *SOX10* in GCMN tissues is very likely to be linked to abnormal proliferation and high canceration of melanocytes and thus exploring the mechanisms of *SOX10* in the occurrence and development of GCMN may be of great significance.

Several limitations exist in the present study because the sample size is not large enough. further experiments in vivo and in vitro are required to validate our findings and to explore the role of the hub genes and biological processes.

5 Conclusions

In summary, we identified some key genes and biological processes that may be associated with the occurrence and development of GCMN through RNA-seq analysis, among which *PTGS2, EGF* and *SOX10* are thought to be the three most important hub genes. Besides relating to melanin synthesis, these three genes are involved in “Arachidonic acid metabolism”,”glycosaminoglycan biosynthetic process”,”developmental pigmentation” respectively. These new findings may help to elucidate the pathophysiology of GCMN.

Declarations

Data availability

The RNA sequencing data presented in this study can be found in SRA database,The accession numbers are: PRJNA730598.

We create a temporary link for journal reviewers to access the SRA data: https://dataview.ncbi.nlm.nih.gov/object/PRJNA730598?reviewer=p93eh1e1arfq0ccbkbc8fmmdnk.

Acknowledgments

We would like to thank Huile Pei,Mingming Zhang and Jilong Guo for their help in collecting skin samples.

Funding

This work was supported by the Natural Science Foundation of Fujian Province, China(2019J01011).

Conflict of interest: The authors have no conflict of interest to declare.

References

1. Mort, R.L., Jackson, I.J. & Patton, E.E. The melanocyte lineage in development and disease (vol 142, pg 620, 2015). *Development* **142**, 1387-1387 (2015).
2. Krengel, S., Scope, A., Dusza, S.W., Vonthein, R. & Marghoob, A.A. New recommendations for the categorization of cutaneous features of congenital melanocytic nevi. *Journal of the American Academy of Dermatology* **68**, 441-451 (2013).
3. Viana, A.C.L., Goulart, E.M.A., Gontijo, B. & Bittencourt, F.V. A prospective study of patients with large congenital melanocytic nevi and the risk of melanoma. *An Bras Dermatol* **92**, 199-204 (2017).
4. Jakchairoongruang, K., Khakoo, Y., Beckwith, M. & Barkovich, A.J. New insights into neurocutaneous melanosis. *Pediatr Radiol* **48**, 1786-1796 (2018).
5. da Silva, V.M.*, et al.* Genetic Abnormalities in Large to Giant Congenital Nevi: Beyond NRAS Mutations. *J Invest Dermatol* **139**, 900-908 (2019).
6. Salgado, C.M.*, et al.* BRAF Mutations Are Also Associated with Neurocutaneous Melanocytosis and Large/Giant Congenital Melanocytic Nevi. *Pediatr Devel Pathol* **18**, 1-9 (2015).
7. Rouille, T.*, et al.* Local Inhibition of MEK/Akt Prevents Cellular Growth in Human Congenital Melanocytic Nevi. *J Invest Dermatol* **139**, 2004-+ (2019).
8. Stark, M.S. Large-Giant Congenital Melanocytic Nevi: Moving Beyond NRAS Mutations. *J Invest Dermatol* **139**, 756-759 (2019).
9. Haas, B.J.*, et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* **8**, 1494-1512 (2013).
10. Kim, D., Landmead, B. & Salzberg, S.L. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**, 357-U121 (2015).
11. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol* **11**(2010).
12. Yu, G.C., Wang, L.G., Han, Y.Y. & He, Q.Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omics* **16**, 284-287 (2012).
13. Ashburner, M.*, et al.* Gene Ontology: tool for the unification of biology. *Nat Genet* **25**, 25-29 (2000).
14. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**, D353-D361 (2017).

15. Tannous, Z.S., Mihm, M.C., Sober, A.J. & Duncan, L.M. Congenital melanocytic nevi: Clinical and histopathologic features, risk of melanoma, and clinical management. *Journal of the American Academy of Dermatology* **52**, 197-203 (2005).
16. Viana, A.C.L., Gontijo, B. & Bittencourt, F.V. Giant congenital melanocytic nevus. *An Bras Dermatol* **88**, 863-878 (2013).
17. Delevoye, C. Melanin Transfer: The Keratinocytes Are More than Gluttons. *J Invest Dermatol* **134**, 877-879 (2014).
18. Chacon-Solano, E*., et al.* Fibroblast activation and abnormal extracellular matrix remodelling as common hallmarks in three cancer-prone genodermatoses. *Brit J Dermatol* **181**, 512-522 (2019).
19. Belvedere, R., Bizzarro, V., Parente, L., Petrella, F. & Petrella, A. Effects of Prisma (R) Skin dermal regeneration device containing glycosaminoglycans on human keratinocytes and fibroblasts. *Cell Adhes Migr* **12**, 168-183 (2018).
20. Sarrazin, S., Lamanna, W.C. & Esko, J.D. Heparan Sulfate Proteoglycans. *Csh Perspect Biol* **3**(2011).
21. Gledhill, K*., et al.* Prostaglandin-E-2 is produced by adult human epidermal melanocytes in response to UVB in a melanogenesis-independent manner. *Pigm Cell Melanoma R* **23**, 394-403 (2010).
22. Wang, J.Y*., et al.* Network pharmacological mechanisms of Vernonia anthelmintica (L.) in the treatment of vitiligo: Isorhamnetin induction of melanogenesis via up-regulation of melanin-biosynthetic genes. *Bmc Syst Biol* **11**(2017).
23. Aoki, H., Moro, O., Tagami, H. & Kishimoto, J. Gene expression profiling analysis of solar lentigo in relation to immunohistochemical characteristics. *Brit J Dermatol* **156**, 1214-1223 (2007).
24. Ercolano, G*., et al.* Knockdown of PTGS2 by CRISPR/CAS9 System Designates a New Potential Gene Target for Melanoma Treatment. *Front Pharmacol* **10**(2019).
25. Panza, E*., et al.* Differential expression of cyclooxygenase-2 in metastatic melanoma affects progression free survival. *Oncotarget* **7**, 57077-57085 (2016).
26. Garcia-Ojalvo, A*., et al.* Systemic translation of locally infiltrated epidermal growth factor in diabetic lower extremity wounds. *Int Wound J* **16**, 1294-1303 (2019).
27. Williams, F.N. & Herndon, D.N. Metabolic and Endocrine Considerations After Burn Injury. *Clin Plast Surg* **44**, 541-+ (2017).
28. Yun, W.J*., et al.* Epidermal Growth Factor and Epidermal Growth Factor Signaling Attenuate Laser-Induced Melanogenesis. *Dermatol Surg* **39**, 1903-1911 (2013).
29. Techapichetvanich, T*., et al.* The effects of recombinant human epidermal growth factor containing ointment on wound healing and post inflammatory hyperpigmentation prevention after fractional ablative skin resurfacing: A split-face randomized controlled study. *J Cosmet Dermatol-Us* **17**, 756-761 (2018).
30. Kim, H.O*., et al.* A Randomized Controlled Trial on the Effectiveness of Epidermal Growth Factor-Containing Ointment on the Treatment of Solar Lentigines as Adjuvant Therapy. *Medicina-Lithuania* **57**(2021).
31. Castillo, S.D. & Sanchez-Cespedes, M. The SOX family of genes in cancer development: biological relevance and opportunities for therapy. *Expert Opin Ther Tar* **16**, 903-919 (2012).
32. Seberg, H.E., Van Otterloo, E. & Cornell, R.A. Beyond MITF: Multiple transcription factors directly regulate the cellular phenotype in melanocytes and melanoma. *Pigm Cell Melanoma R* **30**, 454-466 (2017).
33. Rooper, L.M., Huang, S.C., Antonescu, C.R., Westra, W.H. & Bishop, J.A. Biphenotypic sinonasal sarcoma: an expanded immunoprofile including consistent nuclear beta-catenin positivity and absence of SOX10 expression. *Hum Pathol* **55**, 44-50 (2016).
34. Yu, L.M*., et al.* Sex-Determining Region Y Chromosome-Related High-Mobility-Group Box 10 in Cancer: A Potential Therapeutic Target. *Front Cell Dev Biol* **8**(2020).
35. Kaufman, C.K*., et al.* A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation. *Science* **351**(2016).

Tables

Table 1 The DEGs in the top sixc module clusters sorted by MCODE score.

	MCODE score	Composition	Genes
Module 1	13.000	13nodes 78edges	KRT74,KRT10,KRT83,KRT2,KRT26,KRT77,KRT33B,KRT27,KRT23,KRT36,KRT71,KRT32,KRT25
Module 2	10.762	22nodes 113edges	COL23A1,COL22A1,LCE6A,LCE2D,COLGALT2, LCE2C,LOR,EVPL,COL28A1,CSTA,COL19A1,LCE1A,COL13A1,COL4A5,LCE1F,COL17A1,COL11A2, COL6A5,COL7A1,RPTN,TCHH,COL2A1
Module 3	8.875	17nodes 71edges	RBBP6,KBTBD8,FBXO7,TRIM63,ASB9,NPFFR1, GPR143,LMO7,ASB14,LNX1,EDNRB,TRIM9, CYSLTR2,GAN,LTB4R,P2RY1,FBXW7
Module 4	6.000	7nodes 18edges	KRTAP19-5,KRTAP2-2,KRTAP19-1, KRTAP9-3,KRTAP4-4,KRTAP19-3,KRTAP9-4
Module 5	5.625	17nodes 45edges	SOX10,GPR37,PAX3,EGF,GPR183,HEBP1,SLC24A5,DCT,PMEL,MITF,CCR8,NPY4R,MLPH,TYR,IL18, PTGS2,C3
Module 6	5.500	13nodes 33edges	DNAH11,DNAAF3,SLC39A10,MT1G,RSPH4A□MT1E,NME8,SOD3,OFD1,MT2A,MT1X,SPAG1, SERPINE2

Table 2 Six hub genes identified by MCODE and Cytohubba

Gene symbol	Gene description	log2FC	occurrence*
PTGS2	prostaglandin-endoperoxide synthase 2	-2.348798977	9
EGF	epidermal growth factor	-2.549945398	8
SOX10	SRY-box 10	3.102842853	8
C3	complement C3	1.527467154	6
KRT27	keratin 27	5.710775079	3
EVPL	envoplakin	-2.260789619	3
EDNRB	endothelin receptor type B	3.198479721	3
TYR	tyrosinase	3.644471127	3
MITF	melanogenesis associated transcription factor	2.285262595	3
KRT23	keratin 23	-2.875258674	2

occurrence*□the times of ranking in the top 10 in 12 types of calculation methods

Figures

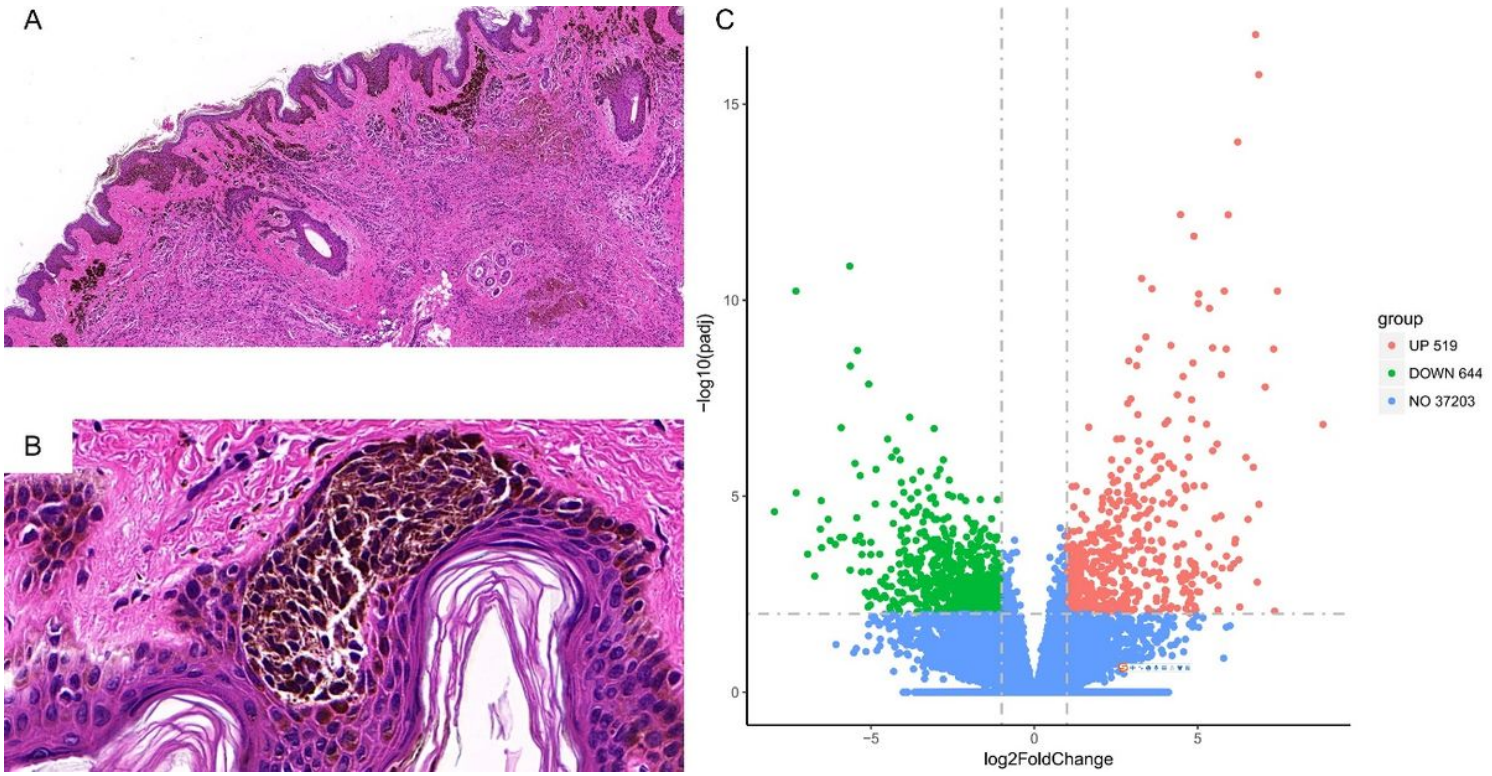

Figure 1

Giant congenital melanocytic nevi(GCMNs) are different from adjacent normal tissues in tissue structure and gene expression.(A). Multiple melanocytes infiltrate the elongated spikes,and surround sweat glands and nerve fibers (B). Melanocytes infiltrate into the epidermis （A and B, Hematoxylin-eosin stain; original magnification: A,×5; B, ×63.）(C). Volcano plot shows all DEGs (| log2fc | > 1.0 and padj < 0.01).Green indicates downregulated genes; red indicates upregulated genes; grey indicates genes without significant differential expression. DEGs,differentially expressed genes.FC, fold change，padj，p-adjusted;

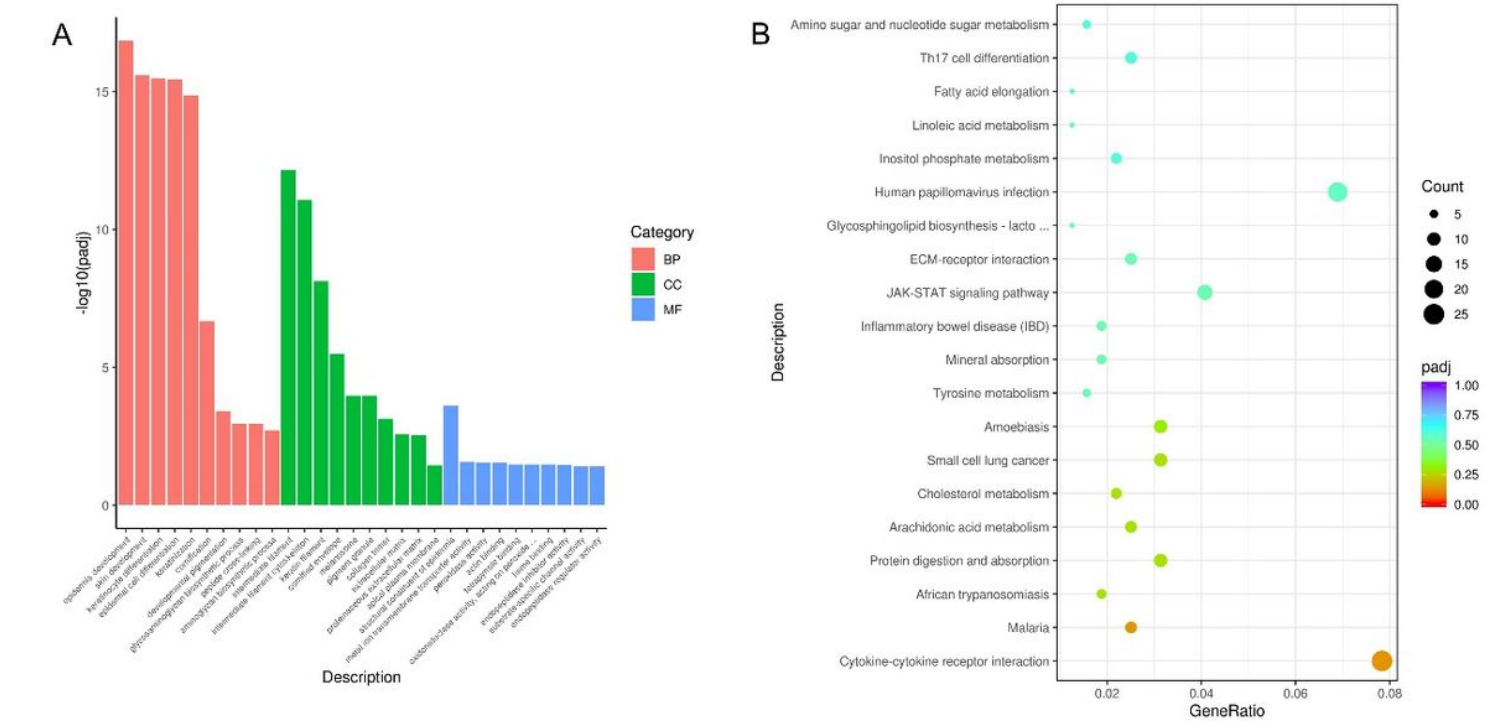

Figure 2

Analysis of DEGs.(A)GO classification. Red column means the biological process,whereas the green column means cellular component,and the blue column indicates molecular function (B)KEGG enrichment. The x-axis is the number of annotated genes in this pathway / the total annotated gene number; the dot size represents annotated gene number; the color from red to purple represents the significance of enrichment

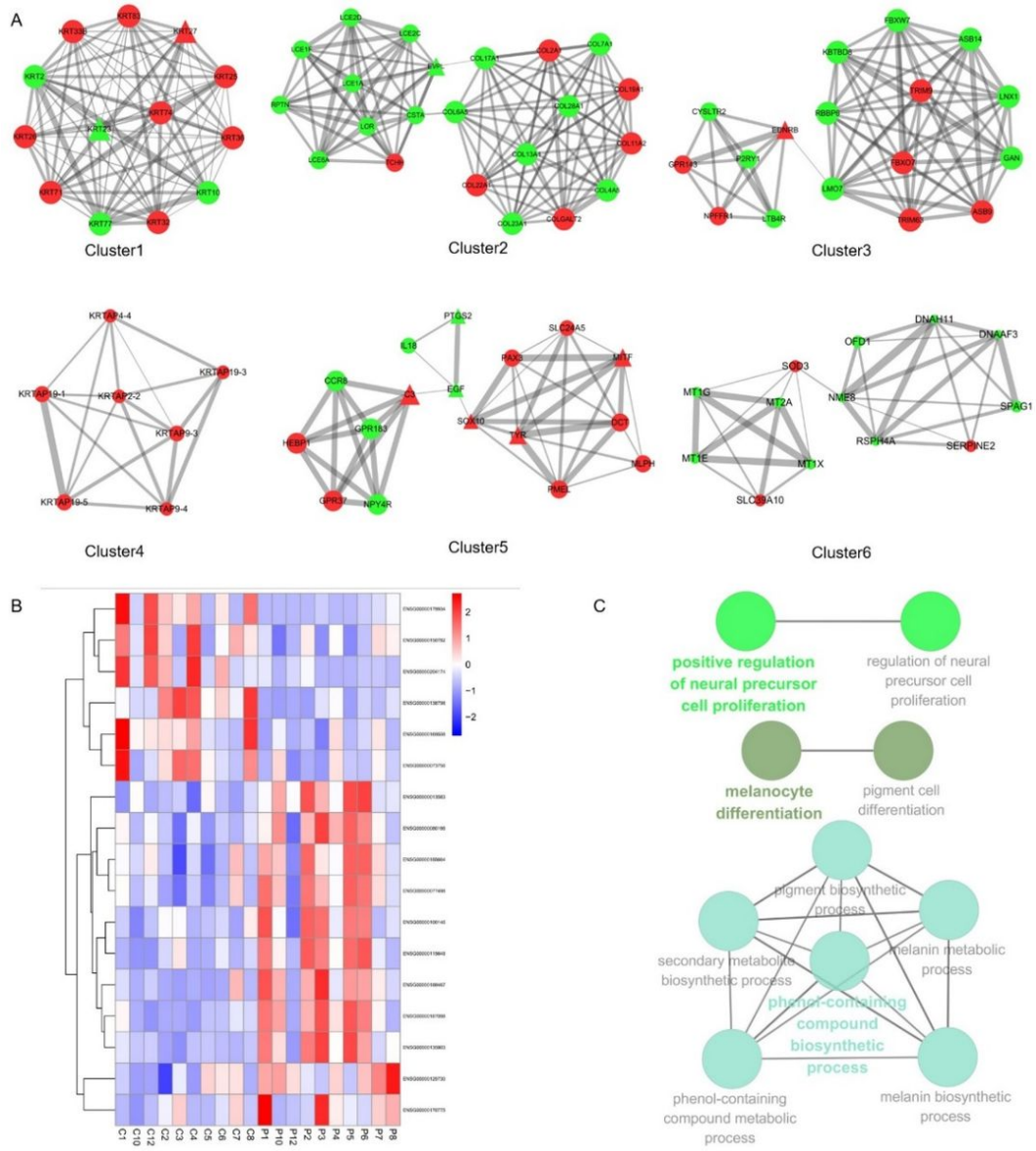

Figure 3

Subnetwork of the top 6 significant MCODE clusters and analysis of Cluster 5.(A). Red nodes and green nodes represent upregulated genes and downregulated genes, respectively. The triangle represents the hub gene.(B). Cluster heatmap of Cluster 5. The red-to-blue gradient bar reflects logarithmic gene expression of each sample, with red and blue representing high and low expression levels, respectively.(C) The representative enriched biological processes of Cluster 5 generated by ClueGO

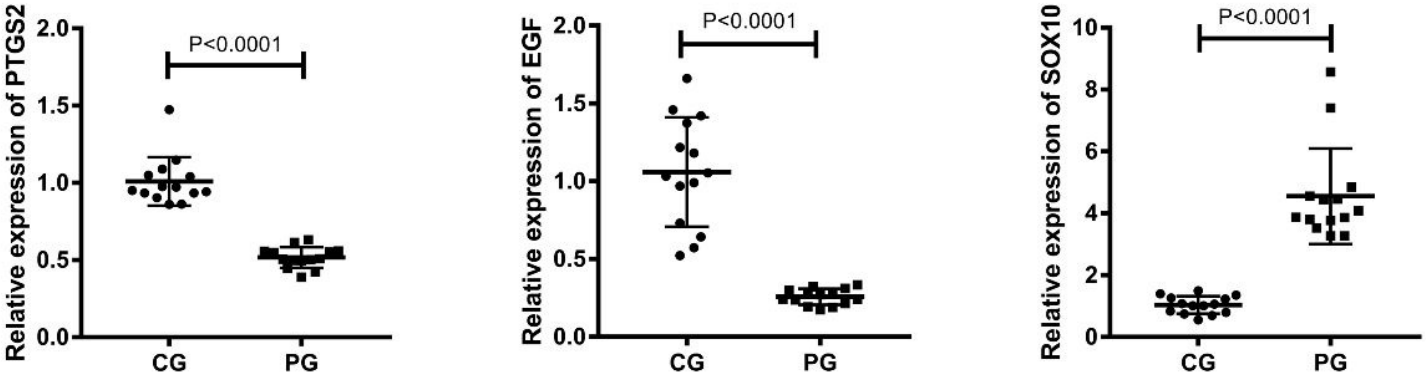

Figure 4

Validation of expression levels of PTGS2, EGF, and SOX10 by RT-qPCR in GCMN and adjacent normal tissues.CG, adjacent normal tissues group;PG,GCMN group