Identification and analysis of dysregulated lncRNA and associated ceRNA in the pathogenesis of keloid

Xilei Duan, Yuemeng Wu, Zheng Zhang, Zhong Lu

Department of Dermatology, Huashan Hospital, Fudan University, Shanghai 200040, China

Contributions: (I) Conception and design: Z Lu, Z Zhang; (II) Administrative support: Z Lu; (III) Provision of study materials or patients: X Duan, Y Wu; (IV) Collection and assembly of data: X Duan, Y Wu; (V) Data analysis and interpretation: X Duan, Y Wu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

These authors contributed equally to this work.

Correspondence to: Zhong Lu; Zheng Zhang. Department of Dermatology, Huashan Hospital, Fudan University, 12 Middle Wulumuqi Road, Shanghai 200040, China. Email: luzhong20100806@qq.com; lypqh@163.com.

Background: Keloid is an excessive fibrosis disease caused by the abnormal proliferation of collagen fibers following trauma. Previous studies have shown that genetic factors have been considered to play important roles in keloid formation. This study is aimed to investigate the regulatory network of messenger RNAs (mRNAs) microRNAs (miRNAs) and long non-coding RNAs (lncRNAs) in keloid, and identifying its key biomarkers

Methods: We performed RNA-seq and miRNA-seq on keloid and normal skin samples. Sequencing datasets were analyzed by bioinformatics. Gene ontology (GO) and pathway analysis presented the characteristics of associated protein-coding genes. Differentially expressed ceRNAs were validated by quantitative reverse transcriptase-PCR (qRT-PCR).

Results: We identified a total of 319 lncRNAs, 1,533 mRNAs and 40 miRNAs as keloid-specific RNAs. Both the GO biological processes and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were analyzed for 1,219 specific genes with differentially expressed mRNAs. Then, with 509 key lncRNAs, 25 miRNAs, and 94 mRNAs, we constructed a ceRNA network and explored any potential underlying mechanisms. In the regulation of the actin cytokeleton pathway, we validated 2 pairs of ceRNAs EGFR/miR-370-3p/lnc-GLB1L-1 and ITGB5/miR-204/lnc-CASP9-3 in another sample size in keloid.

Conclusions: Through RNA-seq and miRNA-seq, we identified keloid-associated lncRNAs, mRNAs and miRNAs, which can be used as potential therapeutic targets and biomarkers for keloid. Our study may lay a foundation for future pathogenesis studies.

Keywords: Keloid; competing endogenous RNA; long non-coding RNA (lncRNA); RNA sequence

Introduction

Keloid is an excessive fibrosis disease caused by abnormal proliferation of collagen fibers following trauma. None of the many methods that have been used to treat keloid, including triamcinolone injection, radical therapy and surgery, have been found to be efficient. This stresses the need for further insight into the pathogenesis driving this disease. Recently, links have been identified between keloid and several SNPs and loci, including rs873549 at lq41, rs8032158 at 15p21.3, rs940187 and rs1511412 at 3q22.3 (1), and lq41 and 15p21.3 in the Chinese Han population (2). Furthermore, several signaling pathways were focused as plausible pathways involved in keloid formation, such as phosphatidylinositol-3-kinase (PI3K)/Akt (3), mitogen-activated protein kinase (MAPK) (4),...
and transforming growth factor-beta 1 (TGF-β1) (5). However, the genetic and epigenetic basis of keloid is yet to be detected, and a more meticulous understanding of its pathogenesis is required.

Non-coding RNAs (ncRNAs) including long non-coding RNA (lncRNA) and microRNA (miRNA) play important roles in the pathogenesis of various diseases (6). There is an increasing amount of evidence which highlights lncRNA’s key contribution to tumorigenesis through its modulation of oncogenic and tumor suppressor pathways in tumor biology (7), including keloid. For example, miRNA-21 may promote keloid genesis via the P13K/AKT signaling pathway (8). Additionally, nowadays, more and more researchers focus on a gene regulatory network based on ceRNA (competing endogenous RNA) and the interaction of messenger RNA (mRNA), lncRNA, circRNA, pseudogene, etc. Zhu et al. reported that via miR-200c, lncRNA-ATB can inhibit the expression of zinc finger protein 217 (ZNF217), thus regulating the secretion of TGF-β2 in keloid fibroblast. As such, the ZNF217/TGF-β2/miR-200c/lncRNA-ATB signaling pathway may be a potential biomarker for the pathogenesis of keloid (9).

In this study, we performed RNA-seq and miRNA-seq on keloid and normal skin samples. We systematically identified genome-wide dysregulated mRNA, lncRNA, and miRNA, constructed a ceRNA network and validated 2 pairs of ceRNAs in another sample size. Understanding these novel RNA interactions will lead to a better understanding of the molecular mechanism of keloid pathogenesis, and lay the foundation for further research in this area, especially for the development of novel treatment options.

**Methods**

**Samples and clinical information**

Thirteen patients with keloid tissue and 12 surrounding healthy tissue were recruited from Huashan Hospital at Fudan University. The diagnoses of the patients (mean ± SD age of 28.8±5.67 years) were confirmed by at least two dermatologists and assessed according to the Vancouver scar scale (VSS). The features of keloid tissues are depicted in Figure S1; briefly, these are abnormal accumulation of extracellular matrix (ECM) and excess proliferation of fibroblasts. All participants were required to sign informed consent documents. This research was conducted in accordance with the principles outlined in the Declaration of Helsinki.

**RNA sequencing**

Total RNA was isolated from tissue samples using TRIzol (Invitrogen, Carlsbad, CA, USA) according to the manufacturer’s included protocol. Total RNA was accessed by electrophoresis on a denaturing agarose gel and quantified by NanoDrop spectrophotometer (Thermo Scientific, Wilmington, DL, USA). A TruSeq Small RNA Sample Preparation Kit (Illumina, San Diego, CA, USA) and a TruSeq RNA Sample Preparation Kit (Illumina, San Diego, CA, USA) were used to generate the miRNA and RNA sequencing library. A HiSeq 4000 platform (Illumina, San Diego, CA, USA) was used to sequence the library preparations by 150 bp paired-end reads for RNA-seq and 50 bp of single-end reads for miRNA-seq.

**Sequencing data analysis**

We checked the quality of the data processing of raw reads using Fastqc (v0.10.1, http://www.bioinformatics. babraham.ac.uk/projects/fastqc/) and, where necessary, used TrimGalore (http://www.bioinformatics.babraham.ac.uk/ projects/trim_galore/) to trim low-quality bases (a quality of less than 10) and adaptors. We used HISAT2 (v2.0.4) to map clean reads to the reference genome for RNA-seq, as well as using StringTie (v1.3.1) to assemble the mapped reads of each sample (10). Differential expression analysis of mRNA and lncRNA was conducted using the DESeq2 R package, and for miRNA-seq, Bowtie (v1.1.2) was used to map the clean reads to the hg19 reference genome. Known mature miRNA expression profiles were generated using miRDeep2. For normalization, we used the bioconductor package edgeR (11).

**Functional enrichment analysis**

To analyze the differential mRNAs and lncRNA-associated differential mRNAs, we used gene ontology (GO) analysis, and identified the significant pathway associated with differential mRNAs through Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis. Enrichment analysis of the GO and KEGG pathways was performed using the clusterProfiler R package.

**ceRNA network analysis**

The miRanda was used to predict miRNA/lncRNA interactions. A ceRNA network of mRNA/miRNA/lncRNA
qRT-PCR was performed using a SYBR-Green PCR kit. The primers used for the qRT-PCR are detailed in Table S1.

Statistical analysis

Statistical analysis was performed using Statistical Programme for Social Sciences 20.0 (SPSS, Inc., Chicago, IL, USA), R Studio (R version 3.4.1). Differences in the qRT-PCR results were calculated by paired t-tests. The threshold used for statistically significant difference was P<0.05.

Results

mRNA, lncRNA and miRNA differential expression analysis (Figure 1)

With a threshold of adjusted P value <0.05 and fold change >2 or <0.5, a total of 319 lncRNAs (68 downregulated and 251 upregulated) and 1,533 mRNAs (792 downregulated and 741 upregulated) between keloid samples and normal samples were detected in RNA-seq (Figure 1A,B,D,E). Thirteen downregulated miRNAs and 27 upregulated miRNAs were detected in miRNA-seq (Figure 1C,F). A hierarchical clustering heatmap and volcano plots were used to display the differentially expressed mRNAs, lncRNAs and miRNAs.

Functional enrichment analyses of keloid

GO and KEGG enrichment analysis was performed on the 1,533 differentially expressed mRNAs to investigate their biological functions (Figures 2,3). The GO functional enrichment analysis explored BP, CC and MF. In BP analysis, microtubule binding, molecular adaptor activity and collagen binding were the three most enriched terms. In CC analysis, cell cortex, synaptic membrane and cell-substrate adherens junction were the three most enriched terms. In MF analysis, regulation of neuron projection development, adherens junction organization and the epidermal growth factor receptor signaling pathway were the three most enriched terms. The KEGG pathway enrichment analysis showed that differentially expressed mRNA were mainly enriched in 14 pathways, the top 3 of which were regulation of actin cytoskeleton, yersinia infection and proteoglycans in cancer (Table 1).

Co-expression analysis

A correlation coefficient was calculated to identify the co-expression of lncRNAs with mRNAs, after filtering out the genes that were not significantly differentially expressed from all of the samples. Only the pairs of lncRNA and mRNA with adjusted P<0.05 were regarded as having a true correlation. Finally, 860 mRNAs and 835 lncRNAs with 6,728 correlations were retained to construct a co-expression network between lncRNA and mRNA (Figure 4). We also found that 207 differentially expressed lncRNAs interacted with 268 mRNAs in the co-expression network mainly enriched in GO term ECM and axonogenesis (Figure 5).

cRNA network construction and analysis

A cRNA network was established by applying the theory that the binding of miRNA and target mRNA could lead to gene silencing, while cRNA can play as a sponge to competitively bind with the same miRNA. Target mRNAs of miRNA were predicted by miRand, and we found 1,615 mRNAs and 49 miRNAs can be matched. Of these mRNAs, a negative correlation with miRNA expression levels was chosen. Target lncRNAs were selected using the method mentioned above. Based on the interaction of mRNA-miRNA, lncRNA-miRNA with shared miRNA, we constructed a general mRNA-miRNA-lncRNA network. Totally, we formed a ceRNA network consisting of 94 mRNAs, 509 lncRNA and 25 miRNAs (1,016 edges and 628 nodes, Figure 6). The intersection of the target mRNAs and differentially expressed mRNAs in keloid was identified.

qRT-PCR validation of transcriptome data of selected cRNA networks

Among the 94 mRNAs in ceRNA, 5 genes—the epidermal growth factor receptor (EGFR), integrin-β5 (ITGB5), the S100 calcium binding protein A4 (S100A4), the molecule CD68, and galanin (GAL) (12-14)—have been reported to be associated with keloids. ITGB5 and EGFR were involved in the top enriched pathways of regulation of the actin cytoskeleton. Here, we selected EGFR and ITGB5 genes as candidate genes for validation of cRNAs in keloid. For each mRNA, we found unique mRNA/miRNA
interactions and multiple IncRNA/miRNA interactions. Since most of the IncRNAs in the selected ceRNA networks have low expression, IncRNAs were selected based on their high expression in RNA-seq data (reads >30), and with unique isoform. We chose 2 groups of ceRNA network: EGFR/miR-370-3p/Inc-GLB1L-1 and ITGB5/miR-204/Inc-CASP9-3 to perform qRT-PCR validation using samples from another 10 patients with keloid and 9 of nearby normal tissue. As shown in Figure 7, the expression of all of the selected ceRNAs was coincident with the result of RNA-seq, which confirmed the accuracy of sequencing. In the first ceRNA network: EGFR and Inc-GLB1L-1 was
microtubule binding, molecular adaptor activity, collagen binding, GTPase activator activity, protein binding, bridging, Rho guanyl-nucleotide exchange factor activity, integrin binding, guanyl-nucleotide exchange factor activity, actin binding, Ras guanyl-nucleotide exchange factor activity, cell cortex, synaptic membrane, cell-substrate adherens junction, site of polarized growth, focal adhesion, cytoplasmic region, cell-substrate junction, cell leading edge, lamellipodium, regulation of neuron projection development, adherens junction organization, epidermal growth factor receptor signaling pathway, integrin-mediated signaling pathway, cell-matrix adhesion, regulation of cell morphogenesis, axonogenesis, cell junction organization, axon development, cell-substrate adhesion.

**Figure 2** KEGG enrichment scatter plot of the keloid DE mRNA. KEGG, Kyoto encyclopedia of genes and genomes; GO, gene ontology; DE mRNA, different expressed messenger RNA.

**Figure 3** GO enrichment summary of the keloid DE mRNA. GO, gene ontology; DE mRNA, different expressed messenger RNA.
downregulated, miR-370-3p was upregulated in keloid tissue. In the second ceRNA network: ITGB5 and Inc-CASP9-3 was upregulated, miR-204 was downregulated in keloid tissue.

**Discussion**

Currently, large-scale cDNA sequencing and RNA sequencing already determines more than 10,000 lncRNAs in the mammalian genome, including humankind. The miRNAs and lncRNAs discussed make an important impact in the progression of multifarious diseases including keloid. For example, research conducted about lncRNA-ATB suggested a signaling axis composition of ZNF217/TGF-β2/miR-200c/lncRNA-ATB, and there has also been an association reported between keloids and CACNA1G-AS1, HOXA11-AS, LINC00312 and other lncRNAs (9,15). Nevertheless, the expression profiles of mRNA, lncRNA and miRNA remain elusive. In our study, we compared genome-wide expression profiles between the lesioned skin and normal skin of keloid patients. The differentially expressed mRNAs, lncRNAs and miRNAs were identified and analyzed to elucidate the pathogenesis of keloid.

It has been proven that genetics and epigenetics are clearly related to the pathogenesis of keloid. Until now, only one GWAS carried out on keloid analysis, and 4 SNPs (rs873549, rs1511412, rs940187 and rs8032158) were proven to be associated with keloid (1). We suspect that this...
Figure 5 GO enrichment summary of the keloid DE IncRNA. GO, gene ontology; DE IncRNA, different expressed long non-coding RNA.

Figure 6 The competing endogenous RNA regulatory network in keloid.
relative lack of research into the genetic basis of keloid is at least part of the reason that we are unable to find significant associations between known SNPs and our sequencing data.

Our research found the most significant pathway to be the regulation of the actin cytoskeleton. A vast array of core eukaryotic phenotypes is governed by the actin cytoskeleton, including cell movement, endocytosis, vesicular trafficking, and cytokinesis (16). Consequently, cellular motility and the invasion of cancer also participated with the dynamic remodeling of actin filaments (17). This pathway can explain the invasive nature of keloids compared with hypertrophic scars. Regulation of the actin cytoskeleton pathway further relates to the adherens junction pathway, the MAPK signaling pathway and the focal adhesion pathway, which also participates in the generation and development of many human cancers. In this pathway, EGFR, ITGB5 were also reported to be associated with keloid. Based on the above analysis, we further selected 2 pairs of ceRNAs: EGFR/miR-370-3p/Inc-GLB1L-1 and ITGB5/miR-204/Inc-CASP9-3 for relative expression qRT-PCR verification. The final quantitative results were consistent with the sequencing results, and successfully validate that IncRNA GLB1L-1 serving as a competitive endogenous RNA, downregulates EGFR expression via sponging miR-370-3p.

Our results further validate that IncRNA CASP9-3 serving as a competitive endogenous RNA, upregulates ITGB5 expression via sponging miR-204.

EGFR belongs to the tyrosine kinase receptor family and has tyrosine kinase activity. When combined with the corresponding ligand EGF, it can activate the tyrosine protein kinase binding domain of the receptor, while further activating the downstream signal transduction pathway, which causes cell division and proliferation and increased DNA synthesis (18). EGFR is also a key gene in the mechanism of keloid development. Previous studies indicate that EGFR is less expressed and that there is a decrease of activation in keloid tissue (4). Besides, EGF-induced keloid tissue fibroblasts have less migration than normal tissue fibroblasts, and mitosis is reduced (19). This indicates that the low expression of EGFR promotes the development of keloid. MiR-370-3p is located on human chromosome 14 within the DLK1/DIO3 domain, which also considered to be a tumor-associated genomic region (20). Previous studies have shown that downregulation of miR-370-3p expression

---

**Figure 7** qRT-PCR analysis of expression levels of 2 groups of ceRNA networks: EGFR (A)/miR-370-3p (B)/Inc-GLB1L-1 (C) and ITGB5 (D)/miR-204 (E)/Inc-CASP9-3 (F). *, P<0.05; ***, P<0.001. ceRNA, competing endogenous RNA.
is relevant to ovarian cancer, cholangiocarcinoma, hepatocellular carcinoma, bladder cancer, oral squamous cell carcinoma, neuroblastoma cells and acute myeloid leukemia. In contrast, another study found that miR-370-3p is upregulated in gastric cancer, nephroblastoma, and prostate cancer, promoting the progression. In keloids, the biological and functional mechanisms involved in miR-370-3p are unknown. IncRNA GLB1L-1 is located on chromosome 2 (chr2: 220101439-220118807) and has a long-chain non-coding RNA of 17,369 bp in length. It has not yet been reported on. Its low expression in keloid tissue suggests that it may be associated with the occurrence of keloids. The mechanism is related, and further study into its specific regulatory mechanism is required.

ITGB5, which is a member of the integrin family, has been found to promote tumorigenesis in many cancers through its modulation of interaction between cells and the ECM. There is growing evidence which highlights the contribution of ITGB5 in cancer cell migration, invasion and angiogenesis (21,22). It also can be an attractive target for tumor therapy because of its correlation with oncogenes (23). The cooperation between ITGB5 and the endothelial growth factor receptor 2 (VEGF2) inhibits extrinsic apoptosis to enhance cell survival (24,25). ITGB5 also plays a pivotal role in epithelial-mesenchymal transition (EMT) through the transforming growth factor β (TGF-β) (26). Studies have shown that ITGB5 is significantly upregulated both in keloid and hypertrophic scar (27). MiR-204 is located on human chromosome 9. It specifically binds with target genes to regulate both the proliferation and differentiation of cells. Previous studies have shown that the expression of miR-204 was downregulated in gastric and thyroid cancer tissues (28). MiR-204 acts while the autophagy-related miRNAs participate in the growth of keloid fibroblast (29). IncRNA CASP9-3 is located on chromosome 1 (chr1: 15326680-15343876) and has a long-chain non-coding RNA of 3,105 bp in length. Furthermore, more experiments are required into Inc-CASP9-3 as a ceRNA of ITGB5 via miR-204.

In conclusion, our study analyzed the IncRNA, mRNA and miRNA expression profiles of keloid based on RNA-seq and miRNA-seq. The key IncRNA and miRNA which were identified in our study may serve as valuable diagnostic and therapeutic targets which could lay a foundation for further genetic research. Through co-expression analysis and ceRNA network construction, it was revealed that 2 groups of ceRNA networks: EGFR/miR-370-3p/Inc-GLB1L-1 and ITGB5/miR-204/Inc-CASP9-3 may participate in the pathogenesis of keloid.

Acknowledgments

Funding: This research is supported by the Natural Science Foundation of China (Grant No. 81402589) and the Key Laboratory of Dermatology of the Education Ministry of China, Anhui Medical University (Grant No. AY2017-1-008).

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was approved by Medical Ethics Committee of Huashan Hospital Affiliated to Fudan University (No. 2018-091). All participants were required to sign informed consent documents.

References


Figure S1 Keloid-histologic features: strikingly thick, glassy, homogeneous collagen bundles oriented haphazardly throughout the dermis (A, hematoxylin-eosin stain, 100×; B, hematoxylin-eosin stain, 400×).

Table S1 PCR primer sequences

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFR-forward</td>
<td>CAGTGGAGCGAATCCTTTGG</td>
</tr>
<tr>
<td>EGFR-reverse</td>
<td>GGTTCAGCAACACCCCTGCC</td>
</tr>
<tr>
<td>hsa-miR-370-3p</td>
<td>GCTGGGGTGGAACCTGGTA</td>
</tr>
<tr>
<td>Inc-GLB1L-1-forward</td>
<td>AGCCCTGAAATTCCTAGCC</td>
</tr>
<tr>
<td>Inc-GLB1L-1-reverse</td>
<td>GTATCTCAGGGCTGGCAGGAG</td>
</tr>
<tr>
<td>ITGB5-forward</td>
<td>ATGTGCCTGGTGCTCAAAAGA</td>
</tr>
<tr>
<td>ITGB5-reverse</td>
<td>TCAAAAATGGCTGTGGAGGTGA</td>
</tr>
<tr>
<td>HSA-miR-204</td>
<td>TTCCCTTTGTCTATCTGCTTCTGCTAGA</td>
</tr>
<tr>
<td>Inc-CASP9-3-forward</td>
<td>GCCATTTGGGATGCCGATTT</td>
</tr>
<tr>
<td>Inc-CASP9-3-reverse</td>
<td>AGAACCCTGCTTTCTGCTAGA</td>
</tr>
</tbody>
</table>