RESEARCH

Open Access

Identification of a potential bioinformatics-based biomarker in keloids and its correlation with immune infiltration



Zihan Li^{1,2,3†}, Chuwei Zhang^{1,2†}, Qingrong Zhang^{1,5}, Yipeng Dong^{1,2}, Xinyu Sha⁴, Ming Jiang^{1,2}, Jun Yan^{1,2}, Wenmiao Wang⁴, Houqiang Li⁴, Yi Zhang^{1*} and You Lang Zhou^{1,3*}

Abstract

Keloid formation is a pathological consequence resulting from cutaneous irritation and injury, primarily attributed to excessive collagen matrix deposition and fibrous tissue proliferation. Chronic inflammation, left uncontrolled over an extended period, also stands as a substantial contributing factor. The precise mechanisms underlying keloid formation remain unclear. Therefore, this study aimed to identify key genes for diagnostic purposes. To achieve this, we used two Gene Expression Omnibus (GEO) data sets to identify differentially expressed genes. We identified one particular gene, homeobox *C9 (HOXC9)*, using a thorough strategy involving two algorithms (least absolute shrinkage and selection operator and support vector machine-recursive feature elimination) and weighted gene co-expression network analysis. We then assessed its expression in normal and keloid tissues. In addition, we explored its temporal expression patterns via Mfuzz time clustering analysis. In our comprehensive analysis, we observed that immune infiltration, as well as cell proliferation, are crucial to keloid formation. Thus, we investigated immune cell infiltration in the keloid and normal groups, as well as the correlation between *HOXC9* and these immune cells. It was found that *HOXC9* was closely associated with the immune microenvironment of keloids. This shows that *HOXC9* can serve as a potential biomarker and therapeutic target for keloids.

[†]Zihan Li and Chuwei Zhang have contributed equally to this work.

*Correspondence: Yi Zhang 198zy@163.com

- You Lang Zhou
- youlangzhou@163.com
- ¹ Department of Burn and Plastic Surgery, Affiliated Hospital of Nantong University, Nantong, China
- ² Nantong University, Nantong, China
- ³ The Hand Surgery Research Center, Department of Hand Surgery,
- Affiliated Hospital of Nantong University, Nantong, China

⁴ Department of Thoracic Surgery, Affiliated Hospital of Nantong University, Nantong, China

⁵ Institute of Burn Research, Southwest Hospital, State Key Lab of Trauma, Burn and Combined Injury, Chongqing Key Laboratory for Disease Proteomics, Third Military Medical University (Army Medical University), Chongqing, China

Introduction

Keloid is an aberrant tissue characterized by overgrowth that arises after the healing of skin injuries or wounds, often for unknown reasons. Keloids can form over months or years and are linked to excruciating pain, inflammation, and different physical and psychosocial symptoms, such as pruritus [1]. The back, anterior earlobes, shoulders, and chest are the most prevalent sites for keloid formation, with the latter two being especially frequent due to higher strain. [2].

During the wound-healing process, myofibroblasts and fibroblasts are the primary mediators of collagen matrix remodeling and formation. These cells form an extracellular matrix with increased rigidity [3]. Different types of glycoproteins, collagens, and glycosaminoglycans form the keloid matrix [4]. Over time, the keloid matrix's



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

overproduction of type III collagen is replaced by a significant proportion of type I collagen. Fibers in keloids are typically larger compared with those in normal skin [5]. Due to their morphology and clinically aggressive behavior, keloids caused by dermal trauma are frequently regarded as a form of non-specific aetiologic and benign fibroproliferative reticular skin tumor. The clinical features and morphology of these entities bear resemblance to neoplastic cutaneous tumors, exhibiting uncontrolled growth, excessive vascularization, and tissue invasion [6, 7].

The immune microenvironment is a pivotal factor in keloid formation. Keloids can form as a result of an exaggerated inflammatory response to tissue injury. This pathological state is characterized by an elevated number of inflammatory cells, particularly T cells, macrophages, and mast cells, within keloid tissues to varied extents [8]. These cells secrete a diverse array of cytokines/ chemokines, including interleukins (ILs) and transforming growth factor β 1, which are crucial inflammatory response-related mediators [9].

To date, the exact pathophysiology of keloid formation remains unclear; however, it is thought to be caused by skin tension [10], tissue hypoxia [11], chronic inflammation [12], autoimmunity [13], and genetic factors [14]. No effective treatment is currently available, and keloids often recur even after monotherapies, such as surgical removal [2]. As a result, identifying multiple biomarkers for this illness could potentially enhance patients' prognoses.

Cellular protein expression is critically involved in determining cellular development status [15]. Gene expression encompasses mRNA and protein transcription, translation, and turnover. The link between mRNA and protein levels affects translation and degradation processes, which are important for regulating transcription, mRNA stability, and gene expression [16]. Gene expression is crucial for the development and sustainability of several biological activities, including cell proliferation, cell cycle, and apoptosis, and irregularities in it can lead to disorders, such as cancer [17]. Therefore, understanding these molecules' biological activities in keloids is crucial for identifying potential treatment applications. This study aimed to uncover potential markers to distinguish patients with keloids from those who did not have keloids. The intramodular hub gene and module eigengene (ME) can be used to summarize gene clusters and find highly linked genes using the weighted gene coexpression network analysis (WGCNA) technique. In addition, it calculates module membership (MM) measures and establishes connections between modules and external sample attributes using the eigengene network methodology. Gene screening techniques that use this network-based methodology can identify prospective therapeutic targets and aid in the development of effective therapies [18]. In this study, we used WGCNA as one of the methods to screen key genes for disease.

In this study, we integrated two data sets sourced from the GEO database to augment the sample size. We compared gene expression profiles between the normal and disease groups and identified the differentially expressed genes (DEGs). WGCNA was then used to obtain hub genes, which were then intersected with the DEGs. The LASSO technique was used to further screen the genes selected to identify a key gene. We analyzed the expression of this key gene in normal and keloid tissues, examined its temporal changes concerning other genes using Mfuzz time clustering [19], and recognized the most relevant gene cluster. Finally, we examined immune cell infiltration in the keloid and normal groups, as well as the correlation between this key gene and immune cells.

Materials and methods

Data sources and quality control

We retrieved two data sets of original expression profile data from the NCBI's GEO database, GSE7890 [20] and GSE83286 [21], for this investigation. These data sets were associated with the GPL570and GPL19612 platforms, respectively. Smith J. et al. (https://www.ncbi. nlm.nih.gov/geo/query/acc.cgi?acc=GSE7890) collected samples from both normal dermis and keloids of adult males and females in the GSE7890 data set. From these samples, primary cells were collected for further culture and analysis. To ensure data integrity, samples treated with hydrocortisone were excluded from our study, leaving us with a total of ten samples. The GSE83286 data set contained three samples of keloid lesions from human earlobe keloids and three samples of normal skin from earlobe piercing. Guo L. et al. provided (https://www. ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE83286). In addition, we used the GSE145725 [22] (https://www.ncbi. nlm.nih.gov/geo/query/acc.cgi?acc=GSE145725) data set based on GPL16043, which contained ten normal fibroblasts and nine keloid fibroblasts, for validation.

Preparation and identification of DEGs

To annotate two series matrix files, several official gene symbols from the data table of the microarray platform were employed. The ensuing gene expression matrix files were then combined into a single file. The expression data from the three data sets were then batch-normalized using the "sva" R package. The "limma" R package was used to perform DEG analysis on the generated gene expression matrix files. To identify DEGs, we used a cutoff threshold of P < 0.05 and |logFC| > 1.5. Based on

the findings of the DEG analysis, volcano plots and heat maps were generated.

Enrichment analysis

Gene ontology (GO)

The "clusterProfiler" package was used in R to perform GO enrichment analysis. To find enriched GO terms, we set a threshold of *P* value < 0.05 and an adjusted *P* value < 0.05.

protein-protein interaction (PPI) network analysis

Utilizing the Search Tool for the Retrieval of Interacting Genes (STRING) database (http://www.string-db. org), the DEGs identified in this investigation were used to develop a PPI network. The minimal interaction score needed was set at 0.4.

WGCNA network construction and module identification

Using the "WGCNA" R package, we analyzed co-expression networks [23]. Initially, we examined the samples for potential outliers or missing values and processed or removed them as necessary. We then used an automatic network construction function to develop the co-expression network. The soft thresholding power was determined using R's "pickSoftThreshold" function. After that, we performed a hierarchical clustering analysis (min-ModuleSize=60) after converting the adjacency matrix into a topological overlap matrix (TOM). Similar modules were merged based on commonalities discovered using the dynamic tree-cut function. We computed the ME, which reflects the module's first principal component and characterizes the module's expression pattern. In addition, we calculated the MM to determine the relationship between genes and MEs. This metric serves as an indicator of the dependability of genes in modules and supports the identification of modules related to clinical features. Following that, gene significance (GS) was calculated to link modules to clinical features, and the resulting data were shown. Finally, we retrieved pertinent gene information from the corresponding modules for further study.

Screening biomarkers

To identify key genes, we selected appropriate module genes and intersected them with DEGs. We then performed minimum absolute value convergence using the glmnet package and variable selection using LASSO. Using the cv. glmnet function in the glmnet package, a model cross-validation analysis was performed with the following parameters: Family="binomial," alpha=1, and nfolds=10. The "svmRadial" method in R was also used to implement another algorithm, SVM–RFE, and a model

cross-validation analysis was carried out, for further analysis and identification of the key gene.

Verification of biomarkers

The receiver operating characteristic (ROC) approach was used with the R package "pROC" to evaluate the sensitivity and specificity of the genes identified through SVM–RFE and LASSO analyses. We collected skin and keloid tissues from patients having scar excision surgery to isolate fibroblasts and primary keratinocytes for the expression profile data.

Cell- and tissue-specific gene expression analysis

The Human Protein Atlas (HPA) database (https://www. proteinatlas.org/) is a web-based resource that contains data from transcriptomics, proteomics, and system biology. It facilitates the identification of target gene expression in various cells, tissues, and organs and provides detailed maps of these biological components. The database includes information on patient survival curves as well as the protein expression of normal tissues as well as diseased tissues. Another useful online tool for retrieving and comparing gene expression patterns across several animal species is the Bgee database (https://bgee.org/). The gene expression levels of normal tissues can be compared using the expression data provided on this website, which are derived exclusively from healthy tissues utilized in earlier studies.

Mfuzz time clustering analysis

Genes within our samples were clustered using Mfuzz, an R tool created for the soft clustering of microarray data, based on their expression profiles to establish unique gene sets. After removing outliers or missing values from this analysis, we determined that there were 30 gene clusters. Subsequently, we used the "GSVA" method to calculate the enriched scores for these gene clusters. This enabled us to assess the variations in scores across several gene clusters between the control and scar groups. To investigate the expression features of the three genes indicated above further, we calculated and plotted shifts in marker gene expression across multiple gene clusters and performed correlation tests.

PPI network

To investigate the gene clusters selected earlier, we conducted protein interaction network analyses. Using the STRING database, we set the confidence levels for clusters 1 and 5 at ≥ 0.07 and ≥ 0.09 , respectively. We used Cytoscape (3.7.2) to visualize the results.

Analysis of immune cell infiltration and differences

Each tissue's immune infiltration was examined using CIBERSORT, a program that analyzes gene expression profiles to determine the cell composition of complex tissues. In addition, we assessed the relative number of immune cells in our samples using the ssGSEA algorithm. We previously acquired comprehensive gene sets related to immune cells for this study. The enrichment scores we calculated for the samples served as an estimate of the immune cell infiltration level. Using these results, we constructed heat maps of immune cell distributions. Subsequently, we conducted comparisons to identify variations in the expression patterns of distinct immune cells between the keloid and normal groups.

Analysis of correlation between biomarkers and immune cells

Using the R packages "tidyverse," "reshape2," and "limma," we conducted a correlation analysis between immune cells and marker genes. We formatted the data using the "ggplot2" package to effectively present the results.

Results

Identification of DEGs

We obtained the GSE7890 and GSE83286 data sets from the GEO database, which have been previously studied. After correcting for batch effects, we combined them. We identified 40 DEGs in total using the "limma" R package, with 22 upregulated and 18 downregulated genes. The heatmap shows all of the DEGs (Fig. 1A). In addition, Fig. 1B shows DEG distribution, with red and blue dots showing the same expression characteristics as previously.

DEG PPI and function prediction

We conducted GO analyses comprising biological processes (BPs), cellular components (CCs), and molecular functions (MFs) to acquire insights into the probable functionalities of these 40 DEGs. These DEGs were shown to be enriched in BPs such as cell proliferation implicated in heart morphogenesis, positive control of mesenchymal cell proliferation, and stem cell population maintenance, according to the findings. In addition, they were linked to CCs, such as adherens junctions, the apical plasma membrane, and the extracellular matrix,

(See figure on next page.)

which contains collagen. Using MF analysis, it was discovered that DEGs had a strong correlation with collagen binding, RNA polymerase II-specific activity, DNA-binding transcription activator activity, and extracellular matrix structural constituents (Fig. 1C, D). GOcircus (Fig. 1E) additionally demonstrated that seven genes were enriched in GO terms.

We developed a PPI network using the STRING database (http://www.string-db.org), an online database, based on the interactions between these genes. These interactions are presented in Fig. 1F. For upcoming research, genes that do not interact with others were eliminated.

WGCNA and key module identification

Using WGCNA, we identified co-expression modules of genes with high degrees of co-expression and topological overlap resemblance Initially, the samples were clustered, and a sample clustering tree was generated (Fig. 2A). A scale-free network was successfully designed using the R function "pickSoftThreshold" by establishing a soft thresholding power of 7 according to the scale-free topology criterion, resulting in an R2 value of 0.83. The adjacency matrix was then converted into a TOM to display node similarity while accounting for weighted correlations (Fig. 2B, C). We merged similar modules after locating them using hierarchical clustering and dynamic tree-cut analysis. More ME, MM, and GS calculations were performed to link modules to clinical characteristics. The generated data were presented using heatmaps (Fig. 2D) and bar charts to underline the importance of each module. These results (Fig. 2E) revealed that the turquoise module gene, which was considered a key module, had the lowest P value.

Screening biomarkers

We chose the turquoise module genes and intersected them with the DEGs based on earlier calculations before further screening the resulting genes (Fig. 2F). We identified three high-expressed and four low-expressed DEGs in the turquoise module. LASSO and SVM–RFE were then used. Cross-validation analysis in the LASSO regression model selected the optimal λ value, minimizing error and finding three keloid-related genes: *HSPA2*, *HECW2*, and *homeobox C9 (HOXC9)* (Fig. 2G). Using the

Fig. 1 Identification of differentially expressed genes (DEGs) and their enrichment. A Heatmap visualizes the DEGs, with red indicating high expression and blue indicating low expression. Deeper colors represent greater differences. B Volcano plot shows DEG distribution. Blue and red dots show low- and high-expressed DEGs, respectively. C Gene ontology (GO) enrichment analysis of DEGs. Every GO term's number of genes is shown on the horizontal axis. D, E Circlize and circos diagrams of the GO enrichment analysis of DEGs. F Protein–protein interaction network of DEGs

Page 4 of 16



Fig. 1 (See legend on previous page.)

SVM–RFE technique, two genes, *HOXC9* and *F2R*, were discovered similarly (Fig. 2H). Finally, using the intersection of the genes found by the two algorithms, we discovered HOXC9 to be the biomarker (Fig. 2I).

Biomarker verification

We used ROC analysis to determine *HOXC9*'s diagnostic value in the merged GSE7890 and GSE83286 samples. The high diagnostic value of *HOXC9* was indicated by its area under the curve value of 1.000 (Fig. 3A). To validate the reliability of these results, we selected the GSE145725 data set for verification, yielding a result of 0.844 for the gene (Fig. 3C). This finding further confirms the validity of our previous analysis.

Expression of the biomarker in normal tissues

To gain a comprehensive understanding of this biomarker, we examined its expression across the entire body (Fig. 4A) using data from the HPA database. We also investigated its enrichment profile in diverse skin tissue cell types. According to our findings, the following cell types expressed *HOXC9*: endothelial cells, smooth muscle cells, eccrine sweat gland cells, adipocytes, and fibroblast_2 (mesenchymal cells) (Fig. 4B). In addition, using the Bgee database, we evaluated its expression in multiple skin locations and the adjacent tissues of healthy persons, and a histogram of the results is shown in Fig. 4C. These results revealed that the differences in gene expression between the subcutaneous adipose and skin tissue were not particularly pronounced, suggesting its possible role in keloid formation.

Temporal biomarker expression clustering and correlation analysis

Based on the expression pattern, we organized the genes into 30 distinct clusters. Subsequently, we explored the changes in *HOXC9* expression across these clusters, elucidating how the expression of other genes correlated with the biomarker levels (Fig. 5A). To delve deeper into the expression characteristics of the biomarker, we calculated the enrichment scores for each gene set in which the sample genes were located. We compared the scores and visualized the outcomes between the keloid and normal groups (Fig. 5B). The gene clusters with P < 0.05 are highlighted in this figure. Following this, correlation tests were conducted, which indicated that *HOXC9* displayed the highest level of positive association with gene cluster 5 and the most substantial negative correlation with cluster 1 (Fig. 5C). The relationships were subsequently shown (Fig. 5D, E), and two gene clusters were chosen for further research.

Gene cluster PPI and core gene role

Using online tools, protein–protein interaction network analysis was performed on the two gene clusters stated above. Using Cytoscape software, we further constructed PPI networks for clusters 1 and 5. Cluster 1 had a total of 409 nodes and 439 identified edges, according to STRING, whereas Cluster 5 had 440 nodes and 770 identified edges. The final results are presented in Fig. 5F, G. Based on previous studies, *MDM2* and *CDK1*, the cluster's core genes, are linked to *HOXC9*. These key genes affect cell growth and immune infiltration. Due to the importance of immune infiltration in keloid formation, the immune microenvironment must be studied.

Correlation between biomarker and immune cell infiltration

The surrounding microenvironment, in particular the immune microenvironment, plays a crucial role in the formation of keloid lesions. Uncertainty persists regarding the precise relationship between immune cells and keloid formation. We investigated the immune infiltration in several tissues using CIBERSORT, and the outcomes are shown in Fig. 6A. We generated a correlation heatmap of immune cells (Fig. 6B). In addition, we utilized the ssGSEA algorithm to investigate immune cell infiltration in the tissues. We made a heatmap of immune cells (Fig. 6C) to better visualize our results. Red denotes a positive correlation, blue denotes a negative correlation, and the intensity of the color denotes the strength of the association. The expression of certain immune cells was then investigated in keloid and normal tissues. Significant variations were found in the expression of type 2 T helper cells, gamma delta T cells, plasmacytoid dendritic cells, natural killer T cells, neutrophils, natural killer cells, and

(See figure on next page.)

Fig. 2 Screening biomarkers. **A** Weighted gene co-expression network analysis (WGCNA) of genes. Branches are distinguished by different colors, representing seven distinct modules. **B**, **C** Evaluation of the scale-free topology criterion with R2=0.83 and the soft thresholding power=7. **D** Correlation plot of WGCNA modules with trait attributes. Rows and columns represent modules and trait attributes, respectively. Red denotes a positive correlation, whereas blue denotes a negative correlation. **E** Significance of each module's genes. **F** Venn diagram showing the intersection genes of DEGs and the turquoise module. **G** Cross-validation for choosing LASSO regression model tuning parameters. Three genes were examined, and dotted vertical lines were created at the optimal levels. **H** Screening genes using the SVM–RFE algorithm. The number of genes with the optimal accuracy is represented by the blue dots. **I** Venn diagram showing the discovery of significant genes using LASSO and SVM analyses



Fig. 2 (See legend on previous page.)



Fig. 2 continued



Fig. 3 Biomarker verification. A HOXC9-related ROC analysis. B, C GSE145725 data set-based HOXC9 validation



Fig. 4 The biomarker's expression in normal tissues. A Human protein Atlas HOXC9 expression. B HOXC9 expression in different skin cell types. C Human tissue HOXC9 expression

(See figure on next page.)

Fig. 5 Time clustering of expression and correlation tests of biomarkers. A Changes in gene expression in 30 different gene clusters as marker gene expression increases. B Differences in enrichment scores between the keloid and normal groups. C Correlation of HOXC9 with different gene clusters. D, E Correlation of HOXC9 with clusters 1 and 5. F, G Protein–protein interaction network of gene clusters 1 and 5



Fig. 5 (See legend on previous page.)



Fig. 5 continued

(Th2) cells (Fig. 6D). We conducted a correlation analysis to determine the association between immune cells and HOXC9, which demonstrated a positive relationship between HOXC9 and macrophages (Fig. 6E).

Discussion

Keloids are characterized by the formation of scar tissue that extends beyond the original skin injury site, and they commonly occur in persons who are predisposed to this condition [24]. The etiology of these benign fibroproliferative disorders is uncertain. They are further distinguished by blood vessel hyperproliferation [25, 26], epidermal thickening, increased mesenchymal cell density [27], thick and dense hyalinized collagen fibers [28], and excessive fibronectin proliferation [29]. Keloid tissue has an extensive extracellular matrix meshwork [27] due to excessive collagen deposition and growth [30] and irregular fiber distribution. In contrast, normal skin possesses collagen bundles parallel to the epidermis [31].

Currently, a variety of methods, including compression therapy, corticosteroid injections, and surgical procedures, are used in clinical practice for the treatment of keloid scars. However, the recurrence rates of keloids remain high [12]. Keloids can potentially impair limb function and affect patients' quality of life in some cases. Our study aimed to discover prospective targets for keloid therapy. We found that DEGs were abundant in BPs, CCs, and MFs associated with keloid fibroblasts by means of differential expression analysis. In terms of MF, many DEGs were involved in DNA replication activities and the construction of the extracellular matrix. In BPs, we observed enhanced cell proliferation abilities. In addition, their positive regulation of mesenchymal cells and the maintenance of stem cells were emphasized. These findings suggest that these DEGs may lead to the highly activated state of cells in keloid tissue, resulting in excessive proliferation and increased deposition of the extracellular matrix. Furthermore, some skin cells, such as dermal fibroblasts, can secrete extracellular matrix components, such as collagen. As the proliferation ability and quantity of cells increase, so does the extracellular matrix increase, including collagen. Moreover, the cells become tightly connected, contributing to the tough texture of keloids. Cellular component analysis further supports these findings.

We used the LASSO algorithm and WGCNA to explore potential biomarkers. Using this comprehensive approach, we discovered a key gene, *HOXC9*, that was

(See figure on next page.)

Fig. 6 Immune cell infiltration and correlation between biomarkers and immune cells. A Infiltration of different immune cells in tissues. B Heatmap of immune cell infiltration. C Correlation heatmap of immune cells. Red indicates a positive correlation, blue indicates a negative correlation and darker colors denote stronger correlations. D Differences in different types of immune cells between the two groups of samples. E The correlation of HOXC9 with different immune cells







A

action

dia.





D



Е



Fig. 6 (See legend on previous page.)

related to keloids in the GSE7890 and GSE83286 data sets. In normal tissues, our data demonstrated that the *HOXC9* expression pattern was similar in the skin and its adjacent tissues. However, *HOXC9*'s specific expression pattern in keloid tissues indicated its crucial role in keloid formation regulation.

The homeobox gene family contains the protein-coding gene HOXC9. Homeobox genes encode a highly conserved group of transcription factors known for their involvement in morphogenesis during the development of multicellular organisms. HOXC9 has been linked to different disorders, despite not previously being connected to keloids. It is well-known that HOXC9 plays a crucial function in regulating vascular morphogenesis and preserving endothelial cell quiescence[32]. HOXC9 may have an impact on angiogenesis, which is crucial for the formation of keloids. In addition, HOXC9 can modulate these processes via inhibiting interleukin (IL)-8 [33]. Studies have linked low expression of *HOXC9* to various conditions, including lymph node metastasis, papillary thyroid cancer, and Hashimoto's thyroiditis [34]. Conversely, high HOXC9 expression has been observed in human gastric cancer [35] and non-small cell lung cancer (NSCLC) [33]. It might be linked to a poor tumor prognosis depending on the situation. Overexpression of this gene promotes the invasion and development of NSCLC cells. It has also been determined that the circCENPF/ has-miR184 axis is a potential target for the development of cancer cells [36]. HOXC9 overexpression can also promote the development of stem cell- and metastasis-like cells in gastric cancer cells [37]. The inhibition of the DAPK1/RIG1/STAT1 axis can lead to interferongamma resistance [38]. HOXC9 overexpression enhances invasiveness but decreases proliferation in various breast cancer cell lines. This is suggestive of a switch from the proliferative to the invasive phase of cancer cell development [39]. Consequently, *HOXC*9 may play a crucial role in cell proliferation, cell migration, angiogenesis, and other activities significantly linked to keloid formation.

Using a soft cluster analysis, we identified two gene sets closely associated with changes in *HOXC9* expression. The key genes in these two sets were *CDK1* and *MDM2*. According to reports, *MDM2* can activate fibroblasts, resulting in fibrosis [40]. It may also contribute to the emergence of inflammation [41]. *CDK1* is necessary for controlling cell division during the G2/M phase and mitosis. Its overexpression has been linked to the development and progression of various cancers [42]. When *THRIL* is overexpressed, *CDK1*, which is involved in cellular motility and proliferation, is downregulated [43]. Several studies and analyses have connected *CDK1* to immune infiltration in lung adenocarcinoma [44], colorectal cancer [45], and hepatocellular carcinoma [46,

47]. These findings illustrate that these two key genes, which have a high correlation with *HOXC9*, are involved in cell proliferation and are intimately related to immune infiltration. Keloids exhibit a close relationship with the immune microenvironment, and genes related to immune infiltration and immune cell abnormalities may underlie keloid formation. Therefore, it is imperative to investigate the immune microenvironment within keloid tissues.

As inflammatory proliferative fibrous diseases, immune infiltration is critically involved in keloid formation. Previous studies have underscored the substantial role of various inflammatory cells in the wound-healing process. For instance, the equilibrium between the M1 and M2 phenotypes is critical for controlling inflammation and tissue healing, which are regulated by macrophages. High macrophage infiltration, mostly of the M2 type, has been found in the keloid specimens [48]. Similarly, T cells are vital contributors. T-cell subset diversity contributes to these cells' complex role in the keloid formation process. For instance, it is known that Th2 cells release IL-4 and IL-13, which can both boost collagen synthesis and metabolism and cause the accumulation of reticular fibrin [49]. Th1 cells can attenuate tissue fibrosis by secreting IFN-y, which suppresses fibroblast proliferation and decreases the expression of type I and III collagen genes [50]. Treg cells, known for regulating other effector T cells, may directly influence collagen deposition in keloids. Furthermore, the substantial increase in mast cell population in keloids implies their integral role in keloid formation [51], with their numbers and activation status being positively y associated with the degree of keloid formation [52]. By secreting IL-4, VEGF, and basic fibroblast growth factors via the PI-3 K/Akt signaling pathway, mast cells can enhance fibroblast proliferation [52, 53]. Furthermore, it is worth noting that inflammatory chemicals play an important role as key mediators in the execution of immune cell activities. TGF- β levels are increased by Th2 cytokines, such as IL-4 and IL-13, leading to fibrosis. Hence, blocking Th2 cytokines represents a potential treatment avenue for keloids [54]. According to a previous study, keloid tissue's growing margin (perilesional region) also contains IL-17. It can encourage the production of SDF-1 in fibroblasts, which in turn facilitates the recruitment of Th17 cells that produce IL-17, establishing a positive feedback loop [55]. In addition, IL-17 greatly increases STAT3 and HIF-1 expression in healthy fibroblasts, impairing autophagy, which is linked to an increase in necroptosis and fibrosis [56]. This finding shows that suppressing STAT3 might be a feasible keloid therapy strategy [55].

We investigated the degree of immune cell infiltration in the samples from the GEO database, since immune infiltration contributes significantly to the formation of scarring. According to our findings, there were significant variations in the levels of neutrophils, natural killer T cells, gamma delta T cells, plasmacytoid dendritic cells, Th2 cells, and natural killer T cells between the two groups. A previous study has shown that an abnormal T-cell response triggers keloid progression [57]. In our research, we delved into the link between HOXC9 and immune cells, and the findings are presented in the preceding section. HOXC9 has been shown to exhibit a certain relationship with T-cell activity [58], which is a factor that has previously been discussed in the context of keloids. Collectively, our findings underscore the critical function of HOXC9 in keloids. However, its functional mechanism and association with immune infiltration remains unclear, warranting further investigation.

In our study, we not only compared the differences in HOXC9 expression between normal and keloid tissues but also highlighted the aberrant HOXC9 expression in keloids by evaluating its expression in normal tissues. Given the findings of differential and enrichment analyses, we hypothesize that HOXC9 may promote keloid formation by promoting cell proliferation and migration, increasing extracellular matrix deposition, and enhancing angiogenesis in keloid tissues. Furthermore, we identified key gene sets associated with the upregulation of HOXC9, further highlighting its connection to keloid formation and the immune microenvironment. Nevertheless, this study has some limitations. First, some differences were present between the samples. In this investigation, we chose GSE7890 and GSE83286 data sets and evaluated their combined results with GSE145725. These data sets, on the other hand, contain samples obtained from tissue samples or primary cells obtained from tissues. Consequently, variations between them could lead to differences in gene expression. Second, because the sample size is still too small, more microarray data must be evaluated to increase the results' reliability. Finally, further experimental validation, including cellular and animal experiments, is required to substantiate these findings. Our study findings suggest that HOXC9 may influence the immune microenvironment of keloids, thereby promoting their growth. Subsequent experiments may include the establishment of animal models of keloid. These models can be used to observe tissue sizes and histological changes after reducing HOXC9 expression to verify the effect of HOXC9 on keloid growth. Furthermore, in cellular experiments, primary cells, such as fibroblasts and keratinocytes, can be cultured to investigate the effects of HOXC9 reduction on cellular activities, including cell proliferation, migration, and apoptosis. In addition, keloid formation is influenced by environmental factors, such as inflammatory cell infiltration. The association between *HOXC9* and immune infiltration remains unclear [36]. Therefore, investigating whether reducing *HOXC9* expression can lead to improved immune cell infiltration. Gene therapy methods are gaining prominence, but their application in keloid treatment remains relatively rare. Validating the role and expression of *HOXC9* in keloids through various methods can pave the way for its clinical use. This approach offers a feasible method for keloid treatment.

Conclusion

In this study, we performed bioinformatics analyses to detect key genes, such as *HOXC9*, which exhibited differential expression between normal and keloid tissues. Our findings demonstrate that *HOXC9* is critical in keloid formation and intricately linked to the keloid immune microenvironment. Therefore, it may serve as a valuable biomarker for keloid diagnosis and treatment. Our study findings not only provide valuable insights for future experimental investigations but also offer a novel avenue for clinical intervention.

Acknowledgements

We would like to thank the Clinical Experimental Center of the Affiliated Hospital of Nantong University for providing the necessary experimental equipment for this study.

Author contributions

ZL: methodology, software, formal analysis, investigation, writing—original draft, and visualization. CZ: methodology, software, formal analysis, investigation, and writing—original draft. QZ: validation, writing—review and editing, visualization, and funding acquisition. YD and XS: project administration and investigation. MJ and JY: project administration and investigation. WW and HL: validation and supervision. YOZ: resources, conceptualization, and writing—review and editing, review and editing. VZ: conceptualization, resources, writing—review and editing, and funding acquisition.

Funding

This study was funded by the Technology Projects of Nantong (No. HS2020006), China University Industry-Academia-Research Innovation (No. 2021JH033), and the Project of Nantong Science and Technology Bureau (JC2021178).

Availability of data and materials

The genetic/RNA sequencing data for this study were obtained from the NCBI-Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/ geo/).

Declarations

Ethics approval and consent to participate

The authors are fully accountable for ensuring the accuracy and integrity of this work. The study adhered to the Helsinki Declaration (revised in 2013).

Consent for publication

Not applicable.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Received: 28 February 2023 Accepted: 1 October 2023 Published online: 02 November 2023

- References
- 1. Ud-Din S, Bayat A. New insights on keloids, hypertrophic scars, and striae. Dermatol Clin. 2014;32(2):193–209.
- Mustoe TA, et al. International clinical recommendations on scar management. Plast Reconstr Surg. 2002;110(2):560–71.
- Ashcroft KJ, Syed F, Bayat A. Site-specific keloid fibroblasts alter the behaviour of normal skin and normal scar fibroblasts through paracrine signalling. PLoS ONE. 2013;8(12): e75600.
- Satish L, et al. Gene expression patterns in isolated keloid fibroblasts. Wound Repair Regen. 2006;14(4):463–70.
- Verhaegen PD, et al. Differences in collagen architecture between keloid, hypertrophic scar, normotrophic scar, and normal skin: an objective histopathological analysis. Wound Repair Regen. 2009;17(5):649–56.
- 6. Ud-Din S, Bayat A. Keloid scarring or disease: unresolved quasi-neoplastic tendencies in the human skin. Wound Repair Regen. 2020;28(3):422–6.
- Tan S, Khumalo N, Bayat A. Understanding keloid pathobiology from a quasi-neoplastic perspective: less of a scar and more of a chronic inflammatory disease with cancer-like tendencies. Front Immunol. 2019;10:1810.
- Shaker SA, Ayuob NN, Hajrah NH. Cell talk: a phenomenon observed in the keloid scar by immunohistochemical study. Appl Immunohistochem Mol Morphol. 2011;19(2):153–9.
- Nangole FW, Agak GW. Keloid pathophysiology: fibroblast or inflammatory disorders? JPRAS Open. 2019;22:44–54.
- Song H, et al. Tension enhances cell proliferation and collagen synthesis by upregulating expressions of integrin alphavbeta3 in human keloid-derived mesenchymal stem cells. Life Sci. 2019;219:272–82.
- Wolfram D, et al. Hypertrophic scars and keloids—a review of their pathophysiology, risk factors, and therapeutic management. Dermatol Surg. 2009;35(2):171–81.
- Ogawa R. Keloid and hypertrophic scars are the result of chronic inflammation in the reticular dermis. Int J Mol Sci. 2017. https://doi.org/10.3390/ijms18030606.
- 13. Kazeem AA. The immunological aspects of keloid tumor formation. J Surg Oncol. 1988;38(1):16–8.
- 14. Bayat A, et al. Keloid disease: clinical relevance of single versus multiple site scars. Br J Plast Surg. 2005;58(1):28–37.
- Buccitelli C, Selbach M. mRNAs, proteins and the emerging principles of gene expression control. Nat Rev Genet. 2020;21(10):630–44.
- de Sousa Abreu R, et al. Global signatures of protein and mRNA expression levels. Mol Biosyst. 2009;5(12):1512–26.
- 17. Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet. 2010;11(9):597–610.
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinform. 2008;9:559.
- Kumar L, Matthias EF. Mfuzz: a software package for soft clustering of microarray data. Bioinformation. 2007;2(1):5–7.
- Smith JC, et al. Gene profiling of keloid fibroblasts shows altered expression in multiple fibrosis-associated pathways. J Invest Dermatol. 2008;128(5):1298–310.
- 21. Guo L, et al. Expression profile of long noncoding RNAs in human earlobe keloids: a microarray analysis. Biomed Res Int. 2016;2016;5893481.
- Kang Y, et al. Hypoxia and HIF-1α regulate collagen production in keloids. J Invest Dermatol. 2020;140(11):2157–65.
- 23. Tian Z, et al. Identification of important modules and biomarkers in breast cancer based on WGCNA. Onco Targets Ther. 2020;13:6805–17.
- Troiano M, et al. Giant keloid of left buttock treated with post-excisional radiotherapy. J Radiol Case Rep. 2011;5(9):8–15.
- Syed F, Bayat A. Notch signaling pathway in keloid disease: enhanced fibroblast activity in a Jagged-1 peptide-dependent manner in lesional vs extralesional fibroblasts. Wound Repair Regen. 2012;20(5):688–706.
- Syed F, et al. Keloid disease can be inhibited by antagonizing excessive mTOR signaling with a novel dual TORC1/2 inhibitor. Am J Pathol. 2012;181(5):1642–58.
- Friedman DW, et al. Regulation of collagen gene expression in keloids and hypertrophic scars. J Surg Res. 1993;55(2):214–22.
- Lee JY, et al. Histopathological differential diagnosis of keloid and hypertrophic scar. Am J Dermatopathol. 2004;26(5):379–84.

- Babu M, Diegelmann R, Oliver N. Fibronectin is overproduced by keloid fibroblasts during abnormal wound healing. Mol Cell Biol. 1989;9(4):1642–50.
- da Silva IR, et al. In situ cytokine expression and morphometric evaluation of total collagen and collagens Type I and Type III in keloid scars. Mediators Inflamm. 2017;2017:6573802.
- Robles DT, Berg D. Abnormal wound healing: keloids. Clin Dermatol. 2007;25(1):26–32.
- 32. Stoll SJ, Kroll J. HOXC9: a key regulator of endothelial cell quiescence and vascular morphogenesis. Trends Cardiovasc Med. 2012;22(1):7–11.
- Stoll SJ, et al. The transcription factor HOXC9 regulates endothelial cell quiescence and vascular morphogenesis in zebrafish via inhibition of interleukin 8. Circ Res. 2011;108(11):1367–77.
- 34. Cao YM, et al. Prognostic and clinical significance of HOXC9 and HOXD10 in papillary thyroid cancer. Transl Cancer Res. 2021;10(7):3317–25.
- Zhao XF, Yang YS, Park YK. HOXC9 overexpression is associated with gastric cancer progression and a prognostic marker for poor survival in gastric cancer patients. Int J Clin Oncol. 2020;25(12):2044–54.
- Liu Y, Jing L, Zhang J. circRNA-mediated upregulation of HOXC9 is correlated with poor outcome and immune microenvironment infiltrates in LUAD. Biochem Biophys Res Commun. 2022;635:128–35.
- Peng X, et al. miR-26a/HOXC9 dysregulation promotes metastasis and stem cell-like phenotype of gastric cancer. Cell Physiol Biochem. 2018;49(4):1659–76.
- Tang Y, et al. Upregulation of HOXC9 generates interferon-gamma resistance in gastric cancer by inhibiting the DAPK1/RIG1/STAT1 axis. Cancer Sci. 2021;112(9):3455–68.
- Hur H, et al. HOXC9 induces phenotypic switching between proliferation and invasion in breast cancer cells. J Cancer. 2016;7(7):768–73.
- Ye C, et al. MDM2 mediates fibroblast activation and renal tubulointerstitial fibrosis via a p53-independent pathway. Am J Physiol Renal Physiol. 2017;312(4):F760–8.
- Zeng Y, et al. MDM2-mediated ubiquitination of RXRbeta contributes to mitochondrial damage and related inflammation in atherosclerosis. Int J Mol Sci. 2022. https://doi.org/10.3390/ijms23105766.
- 42. Liu X, Wu H, Liu Z. An integrative human pan-cancer analysis of cyclindependent kinase 1 (CDK1). Cancers (Basel). 2022. https://doi.org/10.3390/ cancers14112658.
- Zou Y, et al. LncRNA THRIL is involved in the proliferation, migration, and invasion of rheumatoid fibroblast-like synoviocytes. Ann Transl Med. 2021;9(17):1368.
- Xue J, et al. The CDK1-related IncRNA and CXCL8 mediated immune resistance in lung adenocarcinoma. Cells. 2022. https://doi.org/10.3390/cells 11172688.
- Fadaka AO, et al. Expression of cyclin-dependent kinases and their clinical significance with immune infiltrates could predict prognosis in colorectal cancer. Biotechnol Rep (Amst). 2021;29: e00602.
- Zou Y, et al. CDK1, CCNB1, and CCNB2 are prognostic biomarkers and correlated with immune infiltration in hepatocellular carcinoma. Med Sci Monit. 2020;26: e925289.
- Chen X, et al. Identification of aging-related genes associated with clinical and prognostic features of hepatocellular carcinoma. Front Genet. 2021;12: 661988.
- Xu X, et al. The role of macrophages in the formation of hypertrophic scars and keloids. Burns Trauma. 2020;8:tkaa006.
- Trace AP, et al. Keloids and hypertrophic scars: a spectrum of clinical challenges. Am J Clin Dermatol. 2016;17(3):201–23.
- Harrop AR, et al. Regulation of collagen synthesis and mRNA expression in normal and hypertrophic scar fibroblasts in vitro by interferon-gamma. J Surg Res. 1995;58(5):471–7.
- Bagabir R, et al. Site-specific immunophenotyping of keloid disease demonstrates immune upregulation and the presence of lymphoid aggregates. Br J Dermatol. 2012;167(5):1053–66.
- Komi DEA, Khomtchouk K, Santa Maria PL. A review of the contribution of mast cells in wound healing: involved molecular and cellular mechanisms. Clin Rev Allergy Immunol. 2020;58(3):298–312.
- Zhang Q, et al. Green tea extract and (-)-epigallocatechin-3-gallate inhibit mast cell-stimulated type I collagen expression in keloid fibroblasts via blocking PI-3K/AkT signaling pathways. J Invest Dermatol. 2006;126(12):2607–13.

- 54. Luk K, Fakhoury J, Ozog D. Nonresponse and progression of diffuse keloids to dupilumab therapy. J Drugs Dermatol. 2022;21(2):197–9.
- Lee SY, et al. IL-17 induced stromal cell-derived factor-1 and profibrotic factor in keloid-derived skin fibroblasts via the STAT3 pathway. Inflammation. 2020;43(2):664–72.
- Lee SY, et al. IL-17 induces autophagy dysfunction to promote inflammatory cell death and fibrosis in keloid fibroblasts via the STAT3 and HIF-1alpha dependent signaling pathways. Front Immunol. 2022;13: 888719.
- Chen Z, et al. Characterization of CD45RO(+) memory T lymphocytes in keloid disease. Br J Dermatol. 2018;178(4):940–50.
- Xu J, et al. Molecular profiling of TOX-deficient neoplastic cells in cutaneous T cell lymphoma. Arch Dermatol Res. 2020;312(7):513–25.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

