Hereditary gingival fibromatosis (HGF) is an uncommon condition characterised by localised or widespread lesions on the gingiva and its cause is still unclear. It can be inherited as autosomal dominant or recessive mode and cause enlargement of the gingiva.1 An excessive amount of gum tissue can cover the visible part of the teeth, leading to tooth displacement, retention of primary teeth, and difficulties with speech and chewing.2 Rarely present at birth, alveolar ridge thickening often

Hub Genes, Possible Pathways and Predicted Drugs in Hereditary Gingival Fibromatosis by Bioinformatics Analysis

Rong Xia YANG1, Fan SHI1, Shu Ning DU1, Xin Yu LUO1, Wan Qing WANG1, Zhi Lu YUAN1, Dong CHEN1

Objective: To explore potential pathogenic processes and possible treatments using unbiased and reliable bioinformatic tools.

Methods: Gene expression profiles of control and hepatocyte growth factor (HGF) samples were downloaded from CNP0000995. Analysis of differentially expressed genes (DEGs) was conducted using R software (version 4.2.1, R Foundation, Vienna, Austria). Functional enrichment analyses were performed using the Gene Ontology (GO), Kyoto Encyclopaedia of Genes and Genomes (KEGG) and Gene Set Enrichment Analysis (GSEA) databases, then the protein-protein interaction (PPI) network was constructed to screen the top 10 hub genes. Finally, five genes related to cell junctions were selected to build gene-miRNA interactions and predict small-molecule drugs.

Results: A total of 342 downregulated genes and 188 upregulated genes were detected. Candidate pathways include the extracellular matrix (ECM) receptor interaction pathway, the TGF-β signalling pathway and the cell adhesion molecule (CAM) pathway, which were discovered through KEGG and GSEA enrichment studies. GO analyses revealed that these DEGs were significantly enriched in cell adhesion, the adherens junction and focal adhesion. Five hub genes (CDH1, SNAP25, RAC2, APOE and ITGB4) associated with cell adhesion were identified through PPI analysis. Finally, the gene-miRNA regulatory network identified three target miRNAs: hsa-miR-7110-5p, hsa-miR-149-3p and hsa-miR-1207-5p. Based on the gene expression profile, the small-molecule drugs zebularine, ecuronium and prostratin were selected for their demonstrated binding activity when docked with the mentioned molecules.

Conclusion: This study offered some novel insights into molecular pathways and identified five hub genes associated with cell adhesion. Based on these hub genes, three potential therapeutic miRNAs and small-molecule drugs were predicted, which are expected to provide guidance for the treatment of patients with HGF.

Keywords: CDH1, cell-cell junction, hereditary gingival fibromatosis, MiRNA, small-molecule drugs

starts when permanent teeth begin to emerge. It tends to worsen during childhood and may occasionally persist into adolescence. Some drugs, including immunosuppressants, calcium channel blockers and antiseizure medications, are also linked to gingival overgrowth. It is important to establish an antidiastole and develop appropriate treatment plans based on the aetiology.

The pathological manifestations of the gingiva consisted of dense connective tissue that is rich in collagen fibre bundles and infiltrated by inflammatory cells. Researchers have identified several hallmarks of epithelial-to-mesenchymal transition (EMT), including basal lamina rupture and migration of epithelial cells into connective tissue. EMT is the process whereby epithelial cells transform into fibroblast-like cells associated with organ development, oral submucous fibrosis and the invasion of oral squamous cell carcinoma in oral tissue. Recently, EMT has been proposed as a new pathogenic mechanism that contributes to the development of drug-induced gingival fibrosis. Cellular proliferation and the abnormal accumulation of extracellular matrix (ECM) components, including collagen, integrin, fibronectin and glycosaminoglycans, are also the causes of gingival fibrosis. The molecular processes that cause this pathogenic process, however, are not fully understood.

Current treatment options for this condition are limited to surgical procedures, such as gingivectomy and gingivoplasty, which unfortunately have a high recurrence rate. For clinical diagnosis and treatment, it is crucial to understand the molecular and pathophysiological mechanisms of HGF. The purpose of this study is to investigate the molecular pathways associated with the HGF RNA-seq datasets obtained from the China National GeneBank Database (CNGBdb) and to explore potential therapeutic targets.

Materials and methods

Data collection and processing

We downloaded the BAM-formatted human genome RNA-seq datasets of primary fibroblasts from HGFs and normal gingival tissues, which included two HGFs and three normals (the BAM data of CNR0453836 is incomplete), from CNGBdb (https://db.cngb.org/). In our study, the data were first processed through a series of steps: Trim Galore was used for quality control of the CNP000099510 high-throughput data. STAR was employed to compare the reference genome and RSEM software was used to quantify gene expression.

Identification of differentially expressed genes (DEGs)

The original datasets were converted into a recognisable format using R software (version 4.2.1). The “DESeq2” R package was then used to normalise and identify differentially expressed genes (DEGs). The threshold values for DEG detection were determined to be $|\log_2(\text{fold change [FC]})| > 1$ and $P < 0.05$. A volcano plot was created using http://www.sangerbox.com.  

Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was conducted in R software using the package “clusterProfiler”. Gene Ontology (GO) analysis, which includes enrichment for biological processes (BPs), cellular components (CCs) and molecular functions (MFs), was conducted via http://metascape.org. $P < 0.05$ was considered statistically significant.

Further Gene Set Enrichment Analysis (GSEA) analysis

All genomic data from both HGFs and normal gingival samples were uploaded to for Gene Set Enrichment Analysis (GSEA). Enrichments were considered significant when the NES > 1.0, the normalised $P < 0.05$ and the FDR < 0.25.

Construction of protein-protein interaction (PPI) network

The Search Tool for the Retrieval of Interacting Genes (STRING, https://string-db.org) was used to identify protein-protein interactions (PPI) among DEGs. The interaction score was set at 0.4. The result of the PPI network was processed and analysed using Cytoscape version 3.9.1. The key genes in the PPI network were then screened using the CytoHubb plugin.

Further screening of hub genes

Hub gene–associated cell junctions according to the STRING database were further screened and the evidence was listed.

Gene-miRNA network construction

Targeted miRNAs were predicted using miRWalk 3.0 (http://mirwalk.umm.uni-heidelberg.de/) and the miRDB database (http://www.mirdb.org/). The minimum seed sequence length was set to 6 mer and the
Fig 1 Volcano plot of 530 DEGs. The black dots represent genes with no appreciable variation in expression, whereas the red and green dots represent genes that are upregulated and downregulated genes, respectively (a). Heatmap of all DEGs in the HGF group and normal group. N represents gingival fibroblasts from normal and H represents gingival fibroblasts from HGF (b). Upregulated genes as identified by KEGG, ECM receptor interaction and the TGF-β signalling pathway are significantly enriched (c). Downregulated genes as identified by KEGG and CAMs are significantly enriched (d). Upregulated genes as identified by GO analysis and extracellular matrix are significantly enriched (e). Downregulated genes as identified by GO analysis and cell-cell adhesion are significantly enriched (f). Pathway and CC enrichment plots of GSEA showed CAMs pathways, ECM receptor interaction pathway, TGF-β signalling pathway and cell-cell adhesion in HGF are more active than normal (g).
selection criteria were set at $P < 0.95$. MiRNAs that target more than two genes were selected as a result.

Small-molecule drug prediction
Connectivity Map (https://clue.io/) is an online database of chemical reagent action expression profiles. We inputted the top 150 upregulated genes and the top 150 downregulated DEGs into the Query tool. From the results, we selected the top three small-molecule chemical medicines with the highest scores to be molecularly docked with hub genes.

Molecular docking
The PDB database (http://www.rcsb.org/) was used to obtain the 3D structures of the protein molecules, and the PubChem database (https://pubchem.ncbi.nlm.nih.gov/) was employed to obtain the structures of the small-molecule medications. The 3D structure was examined using the molecular docking program AutoDock Vina 1.2.0 and 3D docking pictures were generated using PyMOL 2.5.2. It is widely acknowledged that a specific binding activity between the small-molecule medication and the protein when the binding energy is less than $-4.25$ kcal/mol. Additionally, when the binding energy is less than $-5.0$ kcal/mol, it is deemed that the two have a strong binding activity.12

Statistical analysis
All analyses were performed using R software. The “DESeq2” R package used negative binomial distribution and functional enrichment analyses used a hypergeometric test. Significance was defined as $P < 0.05$.

Results

DEG analysis
The CNP0000995 dataset contained 530 DEGs, including 188 upregulated genes and 342 downregulated genes. These DEGs were presented in a volcano plot (Fig 1a) and a heatmap (Fig 1b).

Enriched GO, KEGG pathway analysis of the DEGs
For the analysis of GO and KEGG pathway enrichment, the ECM receptor interaction pathway, TGF-β signalling pathway and cell adhesion molecule (CAM) were significantly enriched in the KEGG analysis (Fig 1c and d). GO analysis results showed that changes in BPs of DEGs were significantly enriched in several categories, including cell-cell adhesion, cell junction organisation, extracellular matrix, regulation of transforming growth factor beta production and collagen fibril organisation. Changes in cell-substrate junctions and the extracellular matrix were primarily enriched in CCs (Fig 1e and f).

GSEA analysis
GSEA analysis found that cell-cell junctions, CAMs pathways, the ECM receptor interaction pathway and TGF-β signalling pathway were significantly enriched in the HGF group (Fig 1g).

Construction and analysis of the PPI network
A PPI network was established that included upregulated and downregulated genes based on a combined score $> 0.4$, which was considered statistically significant. Cytoscape software (version 3.9.1) was used to visualise the integrated regulatory network and identify the high (marked in red) and low (marked in green) levels of gene expression (Fig 2a). The top 10 genes, ranked by betweenness level using the CytoHubba plugin in Cytoscape, were chosen (Fig 2b). These genes included CDH1, SNAP25, RAC2, APOE, ITGB4, SFN, AHR, CXCL12, SCN2A and CD24 (Table 1).

Further screening of hub genes
Five hub genes associated with cell adhesion according to the String database were screened out, namely CDH1, SNAP25, RAC2, APOE and ITGB4 (Fig 2c), and the evidence was listed (knowledge and text mining) (Table 2).

Further miRNA mining and interaction network analysis
Predicted miRNAs for five genes associated with cell adhesion and gene-miRNA interactions were visualised in Cytoscape 3.9.1 (Fig 3). Target miRNAs with a high number of gene crosslinks $\geq 2$ were selected, including hsa-miR-7110-5p, hsa-miR-149-3p and hsa-miR-1207-5p.

Small-molecule drug screening
The gene expression profile was compared and examined using the Connectivity Map database. We selected three small-molecule drugs with the potential to modify the expression profile and ranked them in order of effectiveness, from highest to lowest. First was zebular-
Fig 2  PPI network of DEGs formed using Cytoscape, with red and green dots representing genes that are upregulated and downregulated, respectively (a). Ten hub genes (CDH1, SNAP25, RAC2, APOE, ITGB4, SFN, AHR, CXCL12, SCN2A and CD24) with the greatest significance were chosen in the PPI network and the colour from red to yellow indicates high to low scores (b). Five hub genes (CDH1, SNAP25, RAC2, APOE and ITGB4) associated with cell adhesion were chosen (marked in green) (c).

Table 1  Top 10 genes in the network ranked by betweenness method.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Symbol</th>
<th>Up/down</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DH1</td>
<td>Down</td>
<td>84</td>
</tr>
<tr>
<td>2</td>
<td>SNAP25</td>
<td>Down</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>RAC2</td>
<td>Down</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>APOE</td>
<td>Down</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>ITGB4</td>
<td>Down</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>SFN</td>
<td>Down</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>AHR</td>
<td>Up</td>
<td>9</td>
</tr>
<tr>
<td>8</td>
<td>CXCL12</td>
<td>Down</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>SCN2A</td>
<td>Down</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>CD24</td>
<td>Down</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 2  Evidence for five genes.

<table>
<thead>
<tr>
<th>Knowledge</th>
<th>Name</th>
<th>Source</th>
<th>Evidence</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text mining</td>
<td>Name</td>
<td>Z-score</td>
<td>Confidence</td>
<td></td>
</tr>
<tr>
<td>CDH1</td>
<td>7.3</td>
<td>★★★★</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAC2</td>
<td>5.2</td>
<td>★★★</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNAP25</td>
<td>6.2</td>
<td>★★★</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ine, a DNA methyltransferase inhibitor, with an enrichment score of −85.32. Second was vecuronium, which acts as an antagonist to the acetylcholine receptor, with an enrichment score of −82.22, and third was prostratin, a PKC activator, with an enrichment score of −81.29 (Table 3).
Molecular docking

We docked zebularine, vecuronium and prostratin with the five molecules (APOE, CDH1, RAC2, SNAP25 and ITGB4). Table 4 displays the data for molecular docking and binding energy. The binding site map for vecuronium and APOE is shown in Fig 4a. Through ARG-32, APOE interacts with ligands to form hydrogen bonds. The binding energy of Selumetinib with DCT is −6.18 kcal/mol, indicating high binding activity. The binding site map for vecuronium and CDH1 is presented in Fig 4b. Through ILE-146, CDH1 interacts with ligands to form hydrogen bonds. The binding energy of Selumetinib with DCT is −4.52 kcal/mol, indicating strong binding activity. A diagram of the prostratin and CDH1 binding site is shown in Fig 4c. Through ALA-132, GLU-129 and TYR-148, CDH1 interacts with ligands via hydrogen bonds. The binding energy of Selumetinib with DCT is 4.52 kcal/mol, suggesting some binding activity. Fig 4d shows the binding site map of zebularine and RAC2. RAC2 interacts with ligands through GLN-180 and TYR-154, forming a hydrogen bond. Zebularine and RAC2 have a binding energy of −4.44 kcal/mol, suggesting that they exhibit binding activity.
Discussion

The present authors identified 530 DEGs, including 188 upregulated genes and 342 downregulated genes. The ECM receptor interaction pathway, TGF-β signalling pathway and CAMs pathway were all involved in HGF, according to KEGG and GSEA enrichment studies. GO functional analysis showed that DEGs were primarily enriched in cell adhesion, the adherens junction and focal adhesion. In addition, five genes (CDH1, SNAP25, RAC2, APOE and ITGB4) associated with cell adhesion were screened. Finally, the authors made predictions of miRNA molecules and small-molecule drugs based on these five hub genes.

The aetiology of HGF is most significant histologically in terms of increased ECM accumulation and degeneration. TGF-β, a multifunctional cytokine that plays a key role in inflammation, wound healing and fibrosis, is an important contributor to ECM deposition and EMT processes in tissue fibrosis. The findings of our bioinformatics analytics fully align with these conclusions.

In this study, we also discovered that the CAMs pathway was significantly enriched in the KEGG and GSEA.
analyses. Furthermore, the GO analysis focused on cell adhesion, adherens junction and focal adhesion. Cell adhesion refers to the ability of a single cell to adhere to the surface of another cell or another inanimate object, such as ECM. To establish connections between cells or between cells and the extracellular matrix, the process of cell adhesion relies on a wide range of proteins known as CAMs, including integrins, selectins, cadherins, members of the immunoglobulin superfamily (IgSF) and mucins. CAMs have mostly been studied in relation to tumours and can mediate the interaction between tumour cells and host cells and ECM, playing a crucial role in the multi-stage sequential process of tumour metastasis. However, some studies also suggest that CAMs can contribute to fibrosis in different organs, such as the liver and lungs.

CDH1, a transmembrane adhesion receptor involved in the regulation of cell adhesion, migration and epithelial cell proliferation, scored highest among the five hub genes. Downregulation of CDH1 is considered a crucial event in the progression of EMT. Iwano et al. found that type II EMT is the primary source of fibroblasts in connective tissue fibrosis. Another study suggests that cyclosporin A can induce type II EMT in gingiva by modifying the morphology of epithelial cells and changing the EMT markers, such as CDH1. The pathogenic function of type II EMT in HGF, however, remains unclear. Roman-Malo et al. observed some histologic features that have been described for EMT in gingival lesions of HGF patients. Based on the above information, we hypothesise that the downregulation of CDH1 leads to the loss of intercellular adhesion, which in turn induces type II EMT and ultimately results in HGF fibrosis. Nevertheless, this is only an initial hypothesis, and more trials are required to confirm it.

The laminin receptor ITGB4, a vital structural component of the hemidesmosome in epithelial cells, was found to be downregulated in our study. It plays a crucial role in tissue morphogenesis by regulating cell adhesion, morphology, polarity and differentiation. According to studies, renal fibrosis and herpes epidermolysis bullosa are linked to the ITGB4 gene. Nevertheless, there is insufficient data regarding the connection between IGTB4 and gingival fibrosis.

miRNAs are a type of traditional non-coding RNA that play a role in regulating various physiological and pathological functions. Current studies have shown that microRNA may participate in the fibrosis of various organs. In HGF, the expression of miR-355-3p is downregulated in gingival fibroblasts as well as in normally occurring human gingival fibroblasts stimulated by TGF-β. In this study, we predicted the miRNAs through the database, resulting in three: hsa-miR-7110-5p, hsa-miR-149-3p and hsa-miR-1207-5p.

Hsa-miR-1207-5p has been reported to be closely associated with cancer development and metastasis. It may function as a novel EMT-negative regulator inhibiting the invasion and metastasis of cancer cells in nasopharyngeal, gastric, lung and oral cancers. MiR-149-3p has been proven to be closely associated with liver and lung fibrosis and may be involved in the EMT process in tumour cells. However, no study has shown that the aforementioned miRNA molecules are associated with therapeutic treatments for gingival fibrosis. This association needs to be confirmed through further trials.

The present study has certain limitations because the sample size is small, which restricts the scope of analysis. In the absence of experimental confirmation, further fundamental studies are required to validate the results and identify the crucial genes and pathways. Additionally, further research must be conducted to investigate whether small-molecule medications and the miRNAs discovered can serve as potential therapeutic targets for HGF.

**Conclusion**

The present authors found potential pathogenic pathways, such as the CAMs pathway and ECM receptor interaction pathway. GO analysis suggested that cell adhesion may potentially play a role in HGF. The authors also screened hub genes involved in cell adhesion, including CDH1, SNAP25, RAC2, APOE and ITGB4.

**Conflicts of interest**

The authors declare no conflicts of interest related to this study.

**Author contribution**

Drs Rong Xia YANG, Fan SHI, Shu Ning DU, Xin Yu LUO, Wan Qing WANG and Zhi Lu YUAN contributed to the data analysis; Dr Rong Xia YANG contributed to the manuscript draft; Dr Dong CHEN contributed to the study conception and manuscript revision.

(Received Jun 30, 2023; accepted Nov 23, 2023)
References


