frontiers | Research Topics

Advances in the genetic and epigenetic of bone and joint diseases

Edited by
Fang Fang Yu, Lu Qiao and Qiling Yuan

Coordinated by
Minjie Fu

Published in
Frontiers in Genetics

ISSN 1664-8714
ISBN 978-2-8325-3519-6
DOI 10.3389/978-2-8325-3519-6

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Advances in the genetic and epigenetic of bone and joint diseases

Topic editors
Fang Fang Yu — Zhengzhou University, China
Lu Qiao — Seoul National University, Republic of Korea
Qiling Yuan — Xi'an Jiaotong University, China

Topic Coordinator
Minjie Fu — Seoul National University, Republic of Korea

Citation
Yu, F. F., Qiao, L., Yuan, Q., Fu, M., eds. (2023). *Advances in the genetic and epigenetic of bone and joint diseases*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-3519-6

Table of contents

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 06 April 2022
doi: 10.3389/fgene.2022.859513

Characterization of Plasma Extrachromosomal Circular DNA in Gouty Arthritis

Jingyuan Pang[1†], Xiaoguang Pan[2†], Ling Lin[2†], Lei Li[3,4†], Shuai Yuan[5†], Peng Han[2], Xiaopeng Ji[6], Hailong Li[6,7], Can Wang[6], Zhaobin Chu[3,4], Haoru Wu[8], Guangyi Fan[3,9], Xiao Du[3,9] and Aichang Ji[6,7]**

[1]Medical School, Shandong University of Traditional Chinese Medicine, Jinan, China, [2]Lars Bolund Institute of Regenerative Medicine, Qingdao-Europe Advanced Institute for Life Sciences, BGI-Qingdao, Qingdao, China, [3]BGI-Qingdao, BGI-Shenzhen, Qingdao, China, [4]College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China, [5]Emergency Department, Qingdao Third People's Hospital, Qingdao, China, [6]Shandong Provincial Key Laboratory of Metabolic Diseases and Qingdao Key Laboratory of Gout, The Affiliated Hospital of Qingdao University, Qingdao, China, [7]Institute of Metabolic Diseases, Qingdao University, Qingdao, China, [8]State Key Laboratory of Medicinal Chemical Biology and College of Pharmacy, Nankai University, Tianjin, China, [9]BGI-Shenzhen, Shenzhen, China

Objective: Extrachromosomal circular DNA elements (eccDNAs) are known for their broad existence in cells and plasma, which may potentially play important roles in many biological processes. Our aim was to identify potentially functional or marked eccDNAs in gout patients.

Methods: The Circle-Seq approach was applied for eccDNA detection from plasma in acute gout patients and healthy controls. Further analysis was performed on the distribution of genomic elements and eccDNA gene annotations in two groups.

Results: We detected 57,216 and 109,683 eccDNAs from the acute gout and healthy control plasma, respectively. EccDNAs were mapped to the reference genome to identify diverse classes of genomic elements and there was no significant difference of eccDNAs on genomic element annotation between gout and control group. A total of 256 eccDNA-associated genes were detected as gout unique eccDNA genes, including COL1A1 and EPB42, which potentially contribute to hyperuricemia and gout, and a couple of genes involved in inflammation or immune response. Enrichment analysis showed that these eccDNA genes were highly correlated with defense response, stress response, and immune and inflammatory responses, including T cell receptor signaling pathway, Fc epsilon RI signaling pathway, and JAK-STAT signaling pathway.

Conclusion: Our discovery reveals the novel potential biological roles of plasma eccDNAs in gouty arthritis.

Keywords: eccDNA, gouty arthritis, inflammation, plasma DNA topologics, EccDNA-associated genes

OPEN ACCESS

Edited by:
Min Li,
Central South University, China

Reviewed by:
Luca Ermini,
Luxembourg Institute of Health, Luxembourg
Wei Lan,
Guangxi University, China
Xiaoqing Peng,
Central South University, China

****Correspondence:***
Xiao Du
duxiao@genomics.cn
Aichang Ji
jiaich@126.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: *21 January 2022*
Accepted: *22 March 2022*
Published: *06 April 2022*

Citation:
Pang J, Pan X, Lin L, Li L, Yuan S, Han P, Ji X, Li H, Wang C, Chu Z, Wu H, Fan G, Du X and Ji A (2022) Characterization of Plasma Extrachromosomal Circular DNA in Gouty Arthritis. Front. Genet. 13:859513. doi: 10.3389/fgene.2022.859513

INTRODUCTION

Nowadays gout is the most common inflammatory arthritis in males, which is mainly triggered by the deposition of monosodium urate (MSU) crystals in tissues and joints (So and Martinon, 2017). The epidemiology worldwide shows an increase in the incidence and prevalence of gout worldwide (Kuo et al., 2015; Dalbeth et al., 2019). Current studies suggest that innate immunostimulatory activity, especially the MSU crystals-mediated NLRP3 inflammasome activation plays a vital role in gouty arthritis development. Many inflammatory cytokines, such as interleukin-1β (IL-1β) and IL-6, participate in this immune response, ultimately leading to gouty attacks (Dalbeth et al., 2016; Wu et al., 2020).

Genetic variation has an important impact on the risk of gout (Barnett, 2018; Major et al., 2018). In the past decades, several research *via* genome-wide association studies on gout have been reported, revealing some genetic variants or genes related to gout and serum urate levels (Major et al., 2018; Tin et al., 2019). However, their explanation of the pathogenesis of gout is still limited. Therefore, other pathogenic mechanisms of gout need to be explored urgently.

Extrachromosomal circular DNA elements (eccDNAs) are derived from genomic DNAs and range in size from a few hundred bases to megabases (Wang et al., 2021). EccDNAs are known for their broad existence across different species. There are increasing evidence showing that they play important functions in diverse biological processes, such as DNA damage repair (Dillon et al., 2015), hypertranscription (Dillon et al., 2015; Hull et al., 2019), homologous recombination (Gresham et al., 2010), replication stress (Dillon et al., 2015), gene amplification in cancer (Paulsen et al., 2018; Morton et al., 2019), and ageing (Zhu et al., 2017; Hull et al., 2019; Qiu et al., 2021). Recently, Zhang's research team from Boston reported that eccDNA was potent innate immunostimulants (Wang et al., 2021). Their studies showed that eccDNAs could activate IFNα, IFNβ, IL-6 or TNF in bone marrow-derived dendritic cells (BMDCs), and marrow-derived macrophages (BMDMs) (Wang et al., 2021).

Here, we sequenced and investigated the eccDNAs in plasma from patients with acute gout and healthy humans. This is very different from previous studies of gout genetics that focused on chromosomal DNA. We detected a large number of diverse eccDNAs with two distinctive peaks at 201 bp and 339 bp. We found that eccDNAs were common in human plasma and distributed in all genomic structures, including genes, intergenic and repetitive regions. We identified 256 eccDNA-associated genes solely represented in gout patients compared to healthy controls, from which we identified a couple of genes involved in inflammatory response, including TLR6, IL2RA, PTGS1, MAPK13, CHRNA4, GCH1, and IL5. Our discovery reveals the novel potential biological roles of plasma eccDNAs in gouty arthritis.

METHODS

Plasma DNA Extraction, eccDNA Purification, and Amplification

Fresh blood was collected from four acute gout patients and four healthy people. Plasma was acquired by centrifuging the fresh blood at 3000 rpm for 10 min. Cell free DNA (cfDNA) was extracted from 300 μL plasma using a MGIEasy Circulating DNA Extraction Kit (MGI-BGI, China). DNA content was measured by Qubit 3.0. The cfDNA was digested using Plasmid-SafeTM ATP-dependent DNase (Epicentre, E310K) at 37°C for 1.5 h with a final concentration of 0.4 U/μl to remove linear dsDNA. After inactivating the ATP-dependent DNase at 70°C for 30 min, the remaining cfDNA was amplified using the REPLI-g Single Cell Kit (Qiagen, Cat #150343) at 30°C for 8 h. The amplified products were purified using Agencourt AMPure XP (Beckman Coulter, A63881) and quantified using Qubit 3.0.

Library Preparation and Sequencing

The obtained eccDNA-enriched DNAs were fragmented into 100 to 300 bp using sonication (Covaris, Inc. Woburn, MA). We took 30 ng of the fragmented DNA to prepare DNA library using the MGIEasy DNA Library Preparation Kit. The library was sequenced using the MGISEQ-2000.

Preprocess of Sequencing Data

The quality control of sequencing reads was performed by Fastqc-v0.11.3. Fastp-0.19.6 was used for filtering and removing low quality reads with default parameters. In order to identify circular DNA from Circle-seq data, we aligned all reads to the two-plasmid sequences (*P895 and P1035*) and the whole human genome sequence (GRCh38 download from UCSC) using BWA-MEM (v0.7.17) (Li, 2013) with default parameters, separately. To comply with the requirements of the Circle-Map (Prada-Luengo et al., 2019) pipeline (v1.1.4), we added the parameter -q to the BWA-MEM when reads were aligned to the human genome. We got two BAM files via sorting by reads names and coordinates.

Circular DNA Identification

We used the Circle-Map pipeline for eccDNA detection (Prada-Luengo et al., 2019). All the details were the same as the description from GitHub-wiki (https://github.com/iprada/Circle-Map/wiki/Tutorial:-Identification-of-circular-DNA-using-Circle-Map-Realign). Filtering was conducted to improve the accuracy with parameters as followed: 1) split reads >= 2, 2) Circle score >=200, 3) Coverage increments in start coordinate >= 0.33, 4) Coverage increments in end coordinate >= 0.33, 5) Coverage continuity <= 0.1, and 6) Standard deviation <Mean coverage. All eccDNA regions that passed the filtering conditions were used for downstream analysis.

Analysis of eccDNA

Using the Ensembl database assembly GRCh38, we identified gene positions on chromosomes. Genes were annotated using the UCSC table-browser (https://genome.ucsc.edu/cgi-bin/hgTables/). Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed based on the differentially expressed eccDNA-associated genes using Gene Set Enrichment Analysis (GSEA, http://www.gsea-msigdb.org/gsea/msigdb/annotate.jsp). The genes differed between groups were plotted using R package

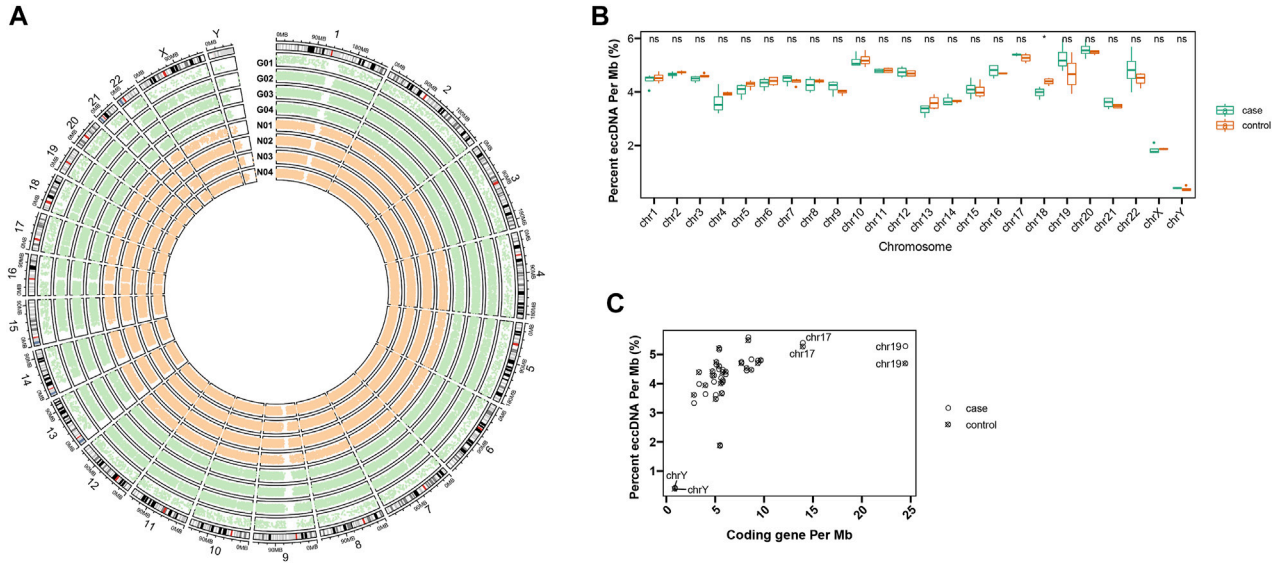


FIGURE 1 | Distribution of plasma eccDNAs across human chromosomes. **(A)** The circos plot of eccDNA distributions from four acute gout patients (G01-G04) and four healthy controls (N01-N04). **(B)** The eccDNA distributions of gout and control on each chromosome. Ns stands for there are no significant differences of eccDNA distribution between control and gout group. * shows there are significant differences of eccDNA distribution between control and gout group on chromosome 18. **(C)** The frequency of eccDNAs per Mb and coding genes per Mb on each chromosome.

ComplexHeatmap (v2.11.1). GSEA annotation results were plotted using R package ggplot2 (v3.3.5).

Motif Signature

We got the 15 bp upstream and downstream sequences flanking the junction of all merged eccDNAs using Circle-Map with bedtools-flank/slop/getfasta (v2.25.0) (Quinlan and Hall, 2010) functions. The start and end sequence motifs in R were generated using the ggseqlogo version 0.1 package (Wagih, 2017).

Statistical Analysis

Statistical analyses were conducted using R version ≥4.1.1. The difference between two groups was checked using Wilcoxon rank-sum test. Correlation analysis were performed by Pearson correlation test. p value <0.05 was considered statistically significant.

RESULTS

Identification and Distribution of eccDNA in Plasma

Circle-Seq method was applied for eccDNA detection from plasma from one group of acute gout (n = 4) and another group of healthy people (n = 4). Acute gout presents as an acutely painful joint inflammation. A total of 166,899 eccDNAs were identified from these eight plasma samples, showing eccDNAs were common in human plasma. We detected 57,216 and 109,683 eccDNAs from acute gout and healthy control plasma, respectively. Plasma eccDNA sequences from gout and healthy controls were distributed across all chromosomes with a combined length of 5.73 and 10.84 Mb of the human genome, respectively (**Figure 1A**). When checking frequency of eccDNAs per Mb on each chromosome, chromosomes 20, 17, 10, and 19 displayed a relatively higher eccDNA frequency compared to the others (**Figure 1B**). Across the chromosomes, chromosome 17 and 19 showed high coding genes per Mb along with the highest eccDNA frequency (**Figures 1B,C**). This pattern was in line with previous observation in human muscle and blood eccDNA study (Møller et al., 2018), suggesting that features of coding genes might play a role on eccDNA formation. Interestingly, we found that chromosome 18 were differentially represented in eccDNA frequency between gout and control groups, while no significant differences were detected on other chromosomes. (**Figure 1B**; chromosome 18: $p < 0.05$).

Genomic Characteristics of Plasma eccDNA

Size distribution of eccDNAs from these eight plasma samples showed comparable size distributions between the gout and healthy control group (**Figure 2A**). The plasma eccDNAs displayed two distinctive peaks at 201 bp and 339 bp, and the 339 bp peak was more predominant than the 201 bp peak (**Figure 2A**). This size signature was in concordance with previous observations studying plasma eccDNA (Kumar et al., 2017; Sin et al., 2020). When comparing the eccDNA size <2 kb between the gout and control group, we detected significantly longer eccDNA sequences in the healthy control by Wilcoxon rank sum test (p-value < 2.2e -16; **Figure 2B**).

Genomic Annotation of eccDNA

The overall detected eccDNAs were mapped to the reference genome to identify diverse classes of genomic elements (**Figure 3**). Firstly, a substantial number of eccDNA sequences were mapped to diverse repetitive regions. The plasma eccDNA molecules were mapped to SINEs (0.45%), telomeres (0.19%),

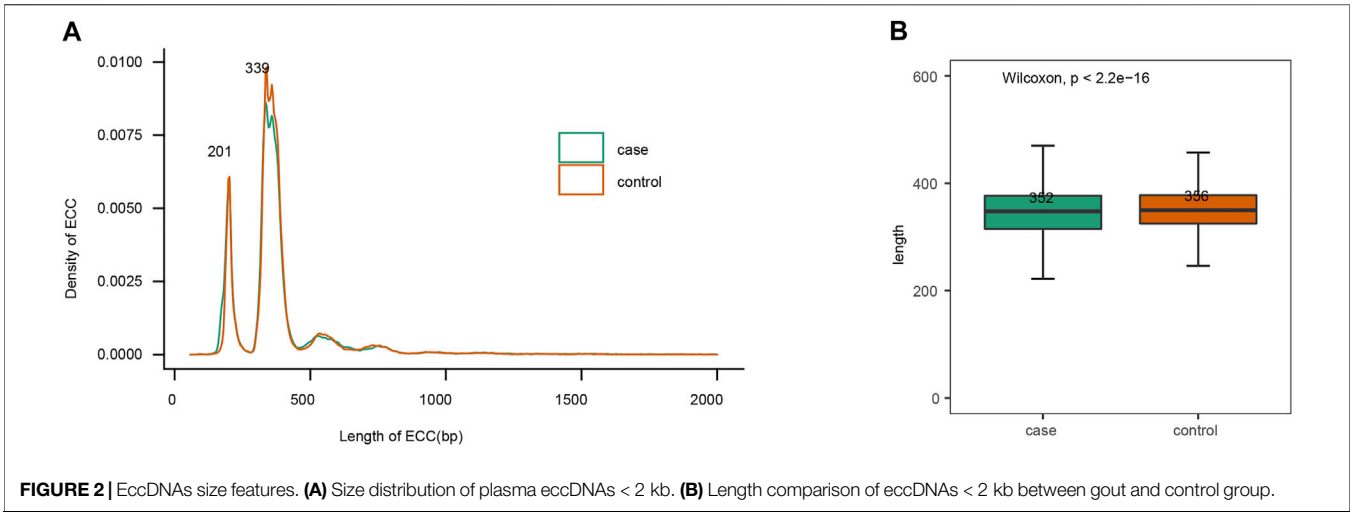


FIGURE 2 | EccDNAs size features. **(A)** Size distribution of plasma eccDNAs < 2 kb. **(B)** Length comparison of eccDNAs < 2 kb between gout and control group.

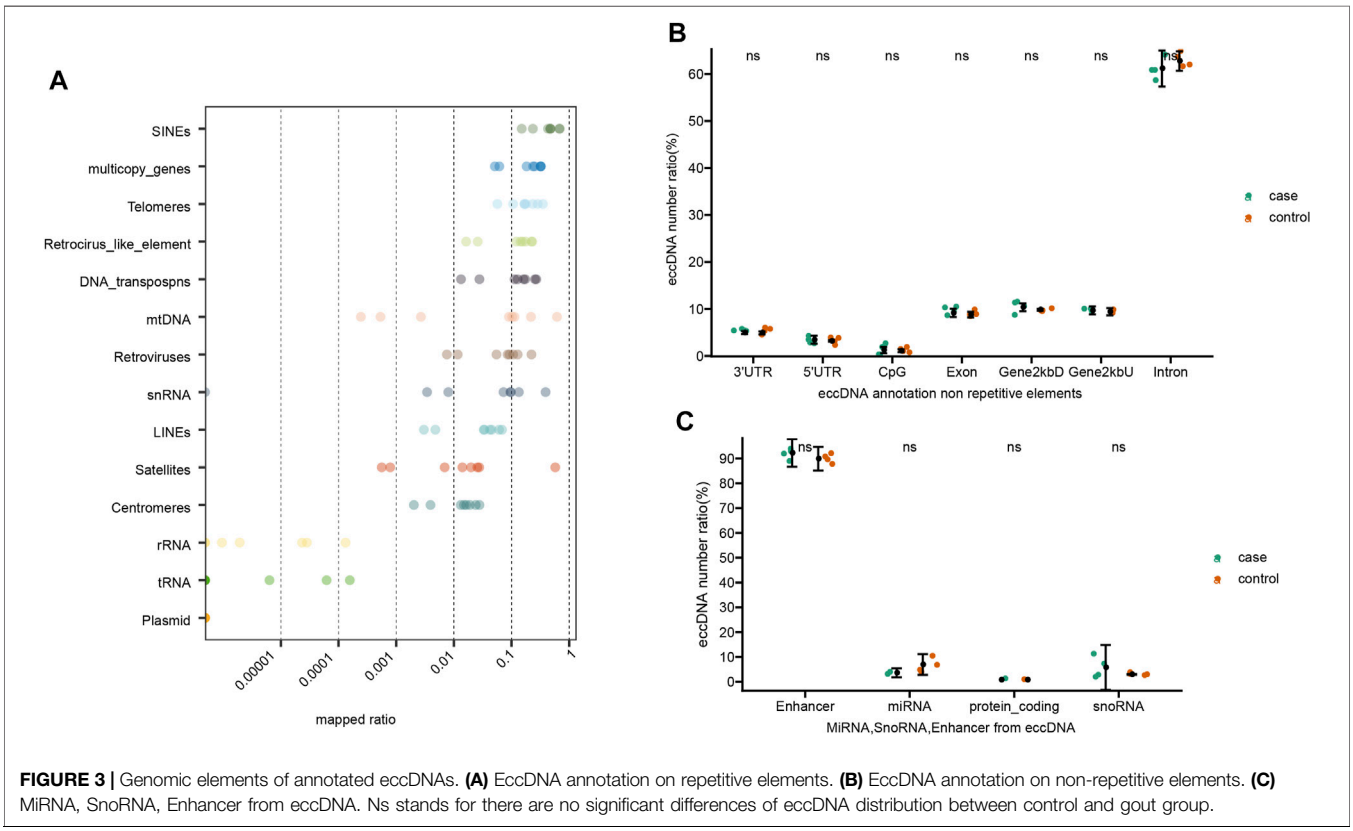
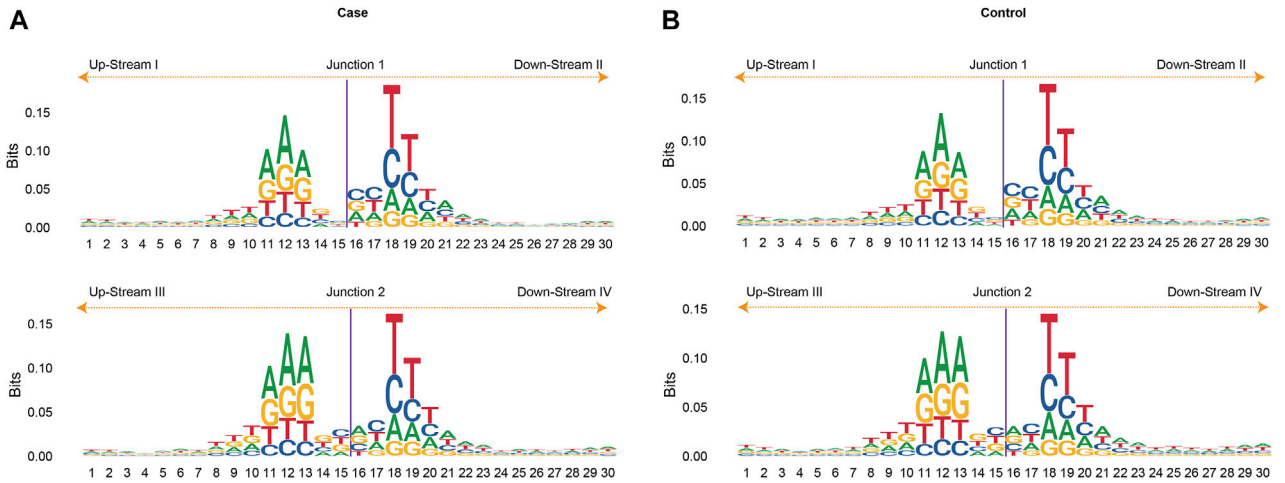


FIGURE 3 | Genomic elements of annotated eccDNAs. **(A)** EccDNA annotation on repetitive elements. **(B)** EccDNA annotation on non-repetitive elements. **(C)** MiRNA, SnoRNA, Enhancer from eccDNA. Ns stands for there are no significant differences of eccDNA distribution between control and gout group.

Retroviruses (0.09%), LINEs (0.04%), satellites (0.08%), and centromeres (0.02%; **Figure 3A**), detected from all types of repetitive regions in human genome. This was in concordance with the knowledge that a large fraction of eccDNAs derive from repetitive sequences (Ling et al., 2021). Apart from that, 70.23% of all eccDNAs were mapped to genic regions. Among that, 9 and 62% were represented in introns and exons, respectively. Five percent and 3% were mapped to 3′UTRs and 5′UTRs, respectively (**Figure 3B**). Some plasma eccDNAs were also detected from enhancers (**Figure 3C**). No significant differences were detected between gout and control group in genomic element annotation (**Figures 3B,C**).

Trinucleotide Motifs Flanking eccDNA Junctions

The eccDNA junction was defined as the site where two ends of a genomic sequence ligated to generate an eccDNA. Therefore, the



C

I	II	III	IV	Frequency
AAA	CTC	AAA	TTT	0.00873897
AAA	TTT	AAA	TTC	0.00873897
AAA	TTT	AAG	TCT	0.00873897
CGG	TCA	TAG	CCG	0.00873897
AAA	TCC	AGA	TTC	0.00699117
AAA	TCC	GAA	TTT	0.00699117
AAA	TTC	GCA	TGC	0.00699117
AAA	TTT	GGA	TTT	0.00699117
AAA	TTT	GTG	TTT	0.00699117
AAG	TTT	AAA	TTT	0.00699117
AGA	TTC	AAA	TTT	0.00699117
GAA	TCA	AAA	TCA	0.00699117
GTG	CTC	AAA	TTT	0.00699117
AAA	AAT	AAA	CTC	0.00524338
AAA	CCA	AAA	ACA	0.00524338

D

I	II	III	IV	Frequency
AAA	TTT	AAA	TTT	0.01094102
AAA	TTT	AAA	TTC	0.00729401
AAA	TTT	AGA	TTT	0.00729401
AAA	TTT	GAA	TTT	0.00729401
AAA	TTA	AAA	TTT	0.00638226
ATA	TTC	GAA	TTT	0.00638226
TAA	TTC	AAA	TTT	0.00638226
AAA	CTT	GAG	TTT	0.00547051
AAA	TAA	AAA	TTT	0.00547051
AAA	TCC	AAA	TTT	0.00547051
AAA	TCT	AAA	TTT	0.00455876
AAA	TTC	GAG	TTT	0.00455876
AAA	TTT	AAA	TCA	0.00455876
AAA	TTT	AAG	CTT	0.00455876
AAA	TTT	GAT	TTT	0.00455876

FIGURE 4 | The motif analysis of eccDNA junctions. The nucleotide composition from 15 bp upstream to 15 bp downstream of the start (Junction 1) and end position (Junction 2) of the eccDNA junction in acute gout **(A)** and healthy control **(B)**. The frequency of base distribution at each position is plotted as a stacked histogram. The size of four bases at each position in the figure reflects their frequency. The 3-bp motifs flanking the start and end site in acute gout **(C)** and healthy control **(D)**. Data were ranked by motif frequencies. I, II, III, IV refers to 3-bp upstream of start site, 3-bp downstream of start site, 3-bp upstream of end site, and 3-bp downstream of end site, respectively.

sequence pattern flanking the eccDNA junctional sites was important as it may shed light on exploring the mechanisms of eccDNA biogenesis. The nucleotide composition from 15 bp upstream to 15 bp downstream of the start and end position of the junction of all merged eccDNAs was analyzed and plotted in **Figures 4A,B**. The frequency of base distribution at each position flanking the start and end site is plotted as a stacked histogram. The relative size of four bases at each position in the figure indicates their frequency. The height of each letter is directly proportional to the frequency of the corresponding base at that position. The bases in each position are arranged from large to small, which can identify the conservative sequence from the bases at the top. As shown in **Figures 4A,B**, both the start and end positions of eccDNA junctions were flanked by a pair of high-frequency trinucleotide segments with 4-bp "spacers" in between. These trinucleotide segments, which were located at 11–13 bp and 18–20 bp from junction 1 and junction 2 (**Figure 4A**), were termed as motifs I, II, III, and IV in **Figure 4C**. The sequences of these four trinucleotide segments with the top 20 frequencies were listed in **Figures 4C,D** for gout and control independently. A recurrent pattern was observed from these sequences: sequence AAA showed high frequency in motifs I and III, whereas sequences TTT and TTC showed dominance in motifs II and IV.

Functional Analysis of eccDNA

We used the overall 166,899 eccDNAs for gene annotation and annotated 13,979 genes in total. Among that, 8,488 (60.7%) genes were shared by all eight samples. 1,544 (11%) and 3,947 (28.2%) genes were only represented in eccDNAs from gout and healthy control plasma, respectively (**Figure 5A**). To investigate the differentially represented eccDNA genes in gout patients, eccDNA genes that were represented in at least two gout patients but absent from the healthy control were defined as gout unique eccDNA genes. We ended up with 256 eccDNA genes only represented in gout, from which 27 genes were identified involved in inflammatory response or potentially gout-related diseases (**Figures 5B,C**). Specifically, Gene TLR6, IL2RA, PTGS1, MAPK13, CHRNA4, GCH1, and IL5 play a role in inflammation or immune response. COL1A1 and EPB42, which potentially contribute to hyperuricemia and gout, were

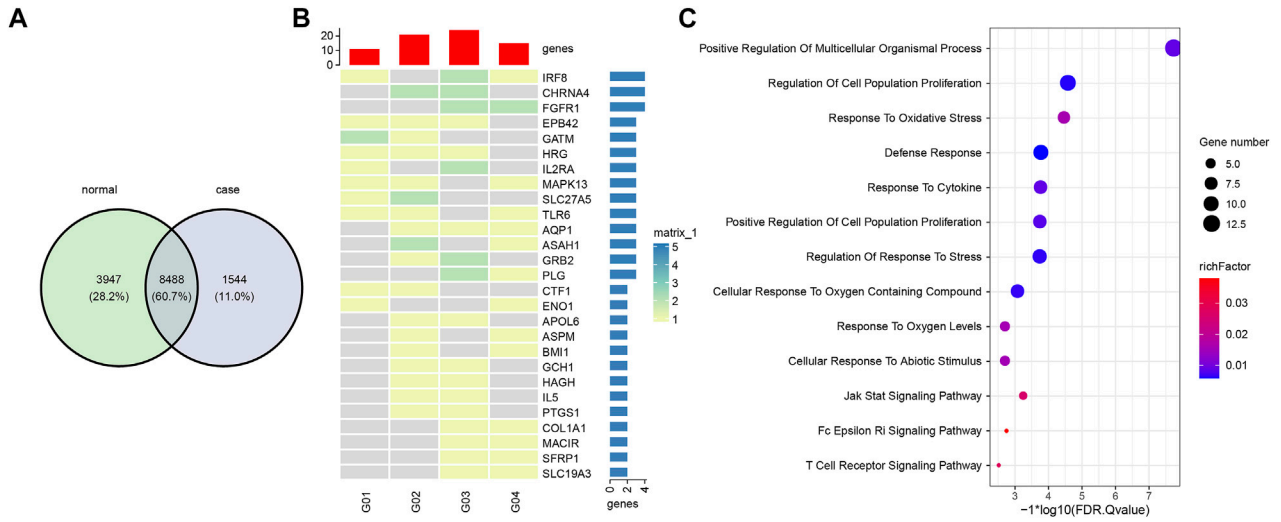


FIGURE 5 | Comparison and functional analysis of eccDNA genes. **(A)** Shared and unique eccDNA genes between gout and healthy controls. **(B)** A heatmap of 27 gout unique eccDNA genes involved in inflammatory response or potentially gout-related diseases. **(C)** Functional enrichment of 27 gout unique eccDNA genes.

also revealed by gout eccDNAs. A list of diverse genes associated with rheumatoid arthritis were detected, including PTGS1, MACIR, APOL6, IRF8, AQP1, ASAH1, ENO1, FGFR1, GRB2, HAGH, HRG, MAPK13, PLG, and SFRP1. Other detected genes including SLC19A3, CTF1, and GATM are related to kidney diseases. GO and KEGG enrichment analysis showed that these 27 gout-unique eccDNA genes were highly correlated with defense response, stress response, and immune and inflammatory responses, including T cell receptor signaling pathway, Fc epsilon RI signaling pathway, and JAK-STAT signaling pathway (**Figure 5C**).

DISCUSSION

The eccDNA is reported to be ubiquitous in eukaryotes (Paulsen et al., 2018; Wang et al., 2021). It is not only present in tumor tissues, but also abundant in plasma and serum of normal tissues (Zhu et al., 2017). The eccDNA can cause gene overexpression due to its high copy number, enhanced transcriptional activity, and enhanced activity of chromosomal and extrachromosomal genes (Nathanson et al., 2014; Zhu et al., 2017; Paulsen et al., 2018; Zhu et al., 2021). The high occurrence of cell-free eccDNAs in the blood system suggest eccDNAs may play vital roles in biological processes, which may potentially act as novel biomarkers for early disease detection, risk assessment, and monitoring of drug treatment response (Paulsen et al., 2018).

The occurrence of gout is presented by the trigger of the inflammatory response, which is induced by the deposition of urate and the overactivation of the innate immunity sequentially (Barnett, 2018; Dalbeth et al., 2019; Dalbeth et al., 2021). A series of signaling pathways are involved in this process. Acute gout attacks depend on the activation of macrophage TLRs/NF-κB signaling pathway (signal 1) and NLRP3 inflammasome complex/IL-1β (signal 2) dual signaling pathways (Wu et al., 2020; Clavijo-Cornejo et al., 2021). As MSU crystals stimulate TLR2 and TLR4 in macrophages, signal 1 is activated, generating the components ASC, Pro-Caspase 1, NLRP3, and Pro-IL-1β. Once MSU crystals enter macrophages by endocytosis, the NLRP3 inflammatory complex is activated. A series of signaling pathways involving proinflammatory cytokines and chemokines are initiated by the mature IL-1β, with neutrophils delivered to the crystal deposition site (So and Martinon, 2017; Dalbeth et al., 2021).

In this study, we report for the first time of the eccDNA characteristics in the acute phase of gout in blood. A couple of eccDNAs of inflammatory factors or genes were different in gout patients and healthy controls. The TLR6, IL2RA, PTGS1, MAPK13, and IL5, which were important in inflammatory response, were only discovered in the acute gout patients in our study. In particular, TLR6 is involved in the key signaling pathway of gout inflammatory response, the NF-kappaB signaling pathway (Hou et al., 2008; Dalbeth et al., 2019). The contents of interleukin 2 (IL-2) were significantly higher than the blank group in acute gouty arthritis rats (Sun et al., 2015). IL2RA is the receptor of IL-2, and IL-2 is the inflammatory markers of gout (Alzahrani et al., 2019). Studies also showed that Th17 was involved in gouty arthritis. The inflammatory Th17 was elevated in MSU crystal-induced arthritis in acute gout rats, which was in concordance with the development of inflammation in gouty arthritis (Dai et al., 2018). The IL5 and MAPK13 discovered in our study are involved in IL-17 signaling pathway (Hashimoto et al., 2005; Yue et al., 2021).

The eccDNA has been reported to activate IFNα, IFNβ, IL-6 or TNF as the potent innate immunostimulants dependent on eccDNA circularity (Wang et al., 2021). In our study, we discovered some eccDNAs associated with inflammatory genes in acute gouty arthritis patients, which might potentially play a role in the activation of the inflammatory response caused by urate crystals and serve as biomarkers for gouty arthritis. Limited by our sample size, this requires further experimental verification.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://db.cngb.org/cnsa, CNP0002557.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the institutional review board on bioethics and biosafety of the Affiliated Hospital of Qingdao University and the IRB-BGI. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AJ, XP, and GF designed the study. LiL, PH, XJ and HL collected the data. JP, XP, LeL, CW, ZC, SY, HW, and XD analyzed the data. JP, XD and AJ wrote the manuscript. All authors had full access to the final version of the manuscript and agreed to its submission.

FUNDING

This work was supported by Shandong University of Traditional Chinese Medicine and grants from The Mechanisms of Hyperuricemia-induced Pancreatic β Cell Impairments (S202010441003).

REFERENCES

Alzahrani, M. I., Alzahrani, M. I., Alsolami, E. A., Al Sulami, A. H., Abdulrahman, A., Alhejaili, M., et al. (2019). Evaluative Role of Family Physician in Diagnosis and Management of Gout in Primary Health Care Centers: A Simple Literature Review. *Arch. Pharm. Pract.* 1, 126.

Barnett, R. (2018). Gout. *The Lancet* 391, 2595. doi:10.1016/s0140-6736(18)31366-7

Clavijo-Cornejo, D., Hernández-González, O., and Gutierrez, M. (2021). The Current Role of NLRP3 Inflammasome Polymorphism in Gout Susceptibility. *Int. J. Rheum. Dis.* 24, 1257–1265. doi:10.1111/1756-185X.14205

Dai, X.-J., Tao, J.-H., Fang, X., Xia, Y., Li, X.-M., Wang, Y.-P., et al. (2018). Changes of Treg/Th17 Ratio in Spleen of Acute Gouty Arthritis Rat Induced by MSU Crystals. *Inflammation* 41, 1955–1964. doi:10.1007/s10753-018-0839-y

Dalbeth, N., Choi, H. K., Joosten, L. A. B., Khanna, P. P., Matsuo, H., Perez-Ruiz, F., et al. (2019). Gout. *Nat. Rev. Dis. Primers* 5, 69. doi:10.1038/s41572-019-0115-y

Dalbeth, N., Gosling, A. L., Gaffo, A., and Abhishek, A. (2021). Gout. *The Lancet* 397, 1843–1855. doi:10.1016/s0140-6736(21)00569-9

Dalbeth, N., Merriman, T. R., and Stamp, L. K. (2016). Gout. *The Lancet* 388, 2039–2052. doi:10.1016/s0140-6736(16)00346-9

Dillon, L. W., Kumar, P., Shibata, Y., Wang, Y.-H., Willcox, S., Griffith, J. D., et al. (2015). Production of Extrachromosomal MicroDNAs Is Linked to Mismatch Repair Pathways and Transcriptional Activity. *Cel Rep.* 11, 1749–1759. doi:10.1016/j.celrep.2015.05.020

Gresham, D., Usaite, R., Germann, S. M., Lisby, M., Botstein, D., and Regenberg, B. (2010). Adaptation to Diverse Nitrogen-Limited Environments by Deletion or Extrachromosomal Element Formation of the GAP1 Locus. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18551–18556. doi:10.1073/pnas.1014023107

Hashimoto, T., Akiyama, K., Kobayashi, N., and Mori, A. (2005). Comparison of IL-17 Production by Helper T Cells Among Atopic and Nonatopic Asthmatics and Control Subjects. *Int. Arch. Allergy Immunol.* 137, 51–54. doi:10.1159/000085432

Hou, B., Reizis, B., and DeFranco, A. L. (2008). Toll-like Receptors Activate Innate and Adaptive Immunity by Using Dendritic Cell-Intrinsic and -extrinsic Mechanisms. *Immunity* 29, 272–282. doi:10.1016/j.immuni.2008.05.016

Hull, R. M., King, M., Pizza, G., Krueger, F., Vergara, X., and Houseley, J. (2019). Transcription-induced Formation of Extrachromosomal DNA during Yeast Ageing. *Plos Biol.* 17, e3000471. doi:10.1371/journal.pbio.3000471

Kumar, P., Dillon, L. W., Shibata, Y., Jazaeri, A. A., Jones, D. R., and Dutta, A. (2017). Normal and Cancerous Tissues Release Extrachromosomal Circular DNA (eccDNA) into the Circulation. *Mol. Cancer Res.* 15, 1197–1205. doi:10.1158/1541-7786.mcr-17-0095

Kuo, C.-F., Grainge, M. J., Zhang, W., and Doherty, M. (2015). Global Epidemiology of Gout: Prevalence, Incidence and Risk Factors. *Nat. Rev. Rheumatol.* 11, 649–662. doi:10.1038/nrrheum.2015.91

Li, H. (2013). Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. *arXiv preprint arXiv:13033997*.

Ling, X., Han, Y., Meng, J., Zhong, B., Chen, J., Zhang, H., et al. (2021). Small Extrachromosomal Circular DNA (eccDNA): Major Functions in Evolution and Cancer. *Mol. Cancer* 20, 113. doi:10.1186/s12943-021-01413-8

Major, T. J., Dalbeth, N., Stahl, E. A., and Merriman, T. R. (2018). An Update on the Genetics of Hyperuricaemia and Gout. *Nat. Rev. Rheumatol.* 14, 341–353. doi:10.1038/s41584-018-0004-x

Møller, H. D., Mohiyuddin, M., Prada-Luengo, I., Sailani, M. R., Halling, J. F., Plomgaard, P., et al. (2018). Circular DNA Elements of Chromosomal Origin Are Common in Healthy Human Somatic Tissue. *Nat. Commun.* 9, 1069. doi:10.1038/s41467-018-03369-8

Morton, A. R., Dogan-Artun, N., Faber, Z. J., MacLeod, G., Bartels, C. F., Piazza, M. S., et al. (2019). Functional Enhancers Shape Extrachromosomal Oncogene Amplifications. *Cell* 179, 1330–1341. doi:10.1016/j.cell.2019.10.039

Nathanson, D. A., Gini, B., Mottahedeh, J., Visnyei, K., Koga, T., Gomez, G., et al. (2014). Targeted Therapy Resistance Mediated by Dynamic Regulation of Extrachromosomal Mutant EGFR DNA. *Science* 343, 72–76. doi:10.1126/science.1241328

Paulsen, T., Kumar, P., Koseoglu, M. M., and Dutta, A. (2018). Discoveries of Extrachromosomal Circles of DNA in Normal and Tumor Cells. *Trends Genet.* 34, 270–278. doi:10.1016/j.tig.2017.12.010

Prada-Luengo, I., Krogh, A., Maretty, L., and Regenberg, B. (2019). Sensitive Detection of Circular DNAs at Single-Nucleotide Resolution Using Guided Realignment of Partially Aligned Reads. *BMC bioinformatics* 20, 663–669. doi:10.1186/s12859-019-3160-3

Qiu, G.-H., Zheng, X., Fu, M., Huang, C., and Yang, X. (2021). The Decreased Exclusion of Nuclear eccDNA: From Molecular and Subcellular Levels to Human Aging and Age-Related Diseases. *Ageing Res. Rev.* 67, 101306. doi:10.1016/j.arr.2021.101306

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics* 26, 841–842. doi:10.1093/bioinformatics/btq033

Sin, S. T. K., Jiang, P., Deng, J., Ji, L., Cheng, S. H., Dutta, A., et al. (2020). Identification and Characterization of Extrachromosomal Circular DNA in Maternal Plasma. *Proc. Natl. Acad. Sci. U.S.A.* 117, 1658–1665. doi:10.1073/pnas.1914949117

So, A. K., and Martinon, F. (2017). Inflammation in Gout: Mechanisms and Therapeutic Targets. *Nat. Rev. Rheumatol.* 13, 639–647. doi:10.1038/nrrheum.2017.155

Sun, X., Zhang, X., Han, J., Liu, C., Qiao, Y., Xie, Y., et al. (2015). Kidney Method for Acute Gouty Arthritis Rats with the Influence of IL-2, IL-16 Expression Level. *Liaoning J. Traditional Chin. Med.* 08.

Tin, A., Marten, J., Halperin Kuhns, V. L., Li, Y., Wuttke, M., Kirsten, H., et al. (2019). Target Genes, Variants, Tissues and Transcriptional Pathways Influencing Human Serum Urate Levels. *Nat. Genet.* 51, 1459.

Wagih, O. (2017). Ggseqlogo: a Versatile R Package for Drawing Sequence Logos. *Bioinformatics* 33, 3645–3647. doi:10.1093/bioinformatics/btx469

Wang, Y., Wang, M., Djekidel, M. N., Chen, H., Liu, D., Alt, F. W., et al. (2021). eccDNAs Are Apoptotic Products with High Innate Immunostimulatory Activity. *Nature* 599, 308–314. doi:10.1038/s41586-021-04009-w

Wu, M., Tian, Y., Wang, Q., and Guo, C. (2020). Gout: a Disease Involved with Complicated Immunoinflammatory Responses: a Narrative Review. *Clin. Rheumatol.* 39, 2849–2859. doi:10.1007/s10067-020-05090-8

Yue, H., Gu, J., Zhao, X., Liang, W., and Wu, Z. (2021). Role of the Interleukin-17 Pathway in the Pathogenesis of Atrial Fibrillation Associated with Inflammation. *Arch. Med. Sci.* 17, 262–265. doi:10.5114/aoms/130392

Zhu, J., Zhang, F., Du, M., Zhang, P., Fu, S., and Wang, L. (2017). Molecular Characterization of Cell-free eccDNAs in Human Plasma. *Sci. Rep.* 7, 10968. doi:10.1038/s41598-017-11368-w

Zhu, Y., Gujar, A. D., Wong, C.-H., Tjong, H., Ngan, C. Y., Gong, L., et al. (2021). Oncogenic Extrachromosomal DNA Functions as mobile Enhancers to Globally Amplify Chromosomal Transcription. *Cancer Cell* 39, 694–707. doi:10.1016/j.ccell.2021.03.006

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 18 November 2022
DOI 10.3389/fgene.2022.1064565

OPEN ACCESS

EDITED BY
Lu Qiao,
Seoul National University, South Korea

REVIEWED BY
Pengfei Qu,
Northwest Women's and Children's Hospital, China
Mingzhu Zhang,
Capital Medical University, China

*CORRESPONDENCE
Hongmou Zhao,
zhao_hongmou@hotmail.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

RECEIVED 08 October 2022
ACCEPTED 02 November 2022
PUBLISHED 18 November 2022

CITATION
Ning Y, Zhang P, Zhang F, Chen S, Liu Y, Chen F, Wu Y, Li S, Wang C, Gong Y, Hu M, Huang R, Zhao H, Guo X, Wang X and Yang L (2022), Abnormal expression of TSG-6 disturbs extracellular matrix homeostasis in chondrocytes from endemic osteoarthritis.
Front. Genet. 13:1064565.
doi: 10.3389/fgene.2022.1064565

Abnormal expression of TSG-6 disturbs extracellular matrix homeostasis in chondrocytes from endemic osteoarthritis

Yujie Ning[1†], Pan Zhang[2†], Feiyu Zhang[1], Sijie Chen[1], Yanli Liu[1], Feihong Chen[1], Yifan Wu[1], Shujin Li[1], Chaowei Wang[1], Yi Gong[1], Minhan Hu[1], Ruitian Huang[1], Hongmou Zhao[3]*, Xiong Guo[1,4], Xi Wang[1] and Lei Yang[5]

[1]Key Laboratory of Trace Elements and Endemic Diseases, School of Public Health, Xi'an Jiaotong University Health Science Center, National Health Commission of the People's Republic of China, Xi'an, China, [2]Sichuan Center for Disease Control and Prevention, Chengdu, China, [3]Foot and Ankle Surgery Department, Honghui Hospital of Xi'an Jiaotong University, Xi'an, China, [4]Clinical Research Center for Endemic Disease of Shaanxi Province, the Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China, [5]School of Nursing, Health Science Center, Xi'an Jiaotong University, Xi'an, China

Background and aims: Kashin-Beck disease (KBD) is a unique endemic osteochondropathy with unclear pathogenesis in China. T-2 toxin exposure has been identified as a significant risk factor of KBD. However, the mechanism of articular cartilage damage induced by T-2 toxin is a conundrum. We explored the role of the extracellular matrix-related gene TSG-6 in the articular chondrocyte damage process under the exposure of HT-2 toxin.

Methods: TSG-6 was identified as a candidate gene by mining our previous gene expression profiling of KBD and verified by qRT-PCR and immunohistochemistry. Then, TSG-6 was silenced by RNA interference technology and overexpressed induction by TNF-α. Gradient concentrations of HT-2 toxin were added to intervene with C28/I2 chondrocytes. MTT was used to observe the proliferation and cell viability of chondrocytes, and qRT-PCR was utilized to detect the expression changes of MMP1, MMP3, MMP13, COL2A1, and proteoglycan before and after treatments for verification.

Results: TSG-6 was upregulated in KBD chondrocytes at the mRNA level and upregulated in the superficial, middle, and deep zones of KBD cartilage. After TSG-6 silencing, the expression of MMP1, MMP3, MMP13, and proteoglycan was significantly decreased while COL2A1 expression was significantly increased, which was reversed after the overexpression of TSG-6 induced by TNF-α ($p < 0.05$). The survival rate of chondrocytes was correspondingly reduced with an increase in the HT-2 toxin concentration. Compared with the blank control group, the expression of MMPs was increased in the intervention group of HT-2 toxin, while the expression of proteoglycan and COL2A1 decreased ($p < 0.05$).

Conclusion: The upregulation of the TSG-6 gene may play a role in promoting the damage and degradation of the extracellular matrix in KBD chondrocytes under the exposure of HT-2 toxin.

KEYWORDS
Kashin-Beck disease, chondrocytes, TSG-6, HT-2 toxin, extracellular matrix

Introduction

Kashin-Beck disease (KBD) is an endemic osteochondropathy found in China, southeastern regions of Transbaikalia in Russia, and North Korea, and sporadic cases are found in Central Asia and the European part of Russia (Ermakov et al., 2019). It is characterized by childhood onset, arthritic pain, enlarged joints, limited mobility, and dwarfism in advanced patients (Han et al., 2018). The clinical manifestations and pathological features of KBD are similar to those of OA such as joint pain, stiffness, limitation of motion, swelling, chondrocyte necrosis, apoptosis, and degradation of the extracellular matrix (ECM) (Wang et al., 2017). Although the etiology and pathogenesis for KBD still need to be elucidated, cereal T-2 toxin exposure has been identified as an important risk factor of KBD (Sun et al., 2019). However, the mechanism of articular cartilage damage induced by the T-2 toxin is unclear.

TSG-6, which is also known as TNFAIP6, encodes a secreted protein that contains a hyaluronic acid-binding domain; therefore, TSG-6 belongs to a member of the hyaluronan-binding protein family. Previous studies determined that the hyaluronic acid domain is primarily involved in the stability of the extracellular matrix and cell migration (Dyer et al., 2016). Moreover, TSG-6 plays an important role in anti-inflammatory and cartilage protection from arthritis by forming a negative feedback loop (Wisniewski et al., 1996). In recent decades, we identified an environment-gene interaction that plays an important role in the cartilage damage process of KBD (Wang et al., 2017; Lin et al., 2019; Yang et al., 2019). For example, T-2 toxin can stimulate IL-1ß production in chondrocytes, which collectively aggravates cell damage by inhibiting the transcription of collagen II and aggrecan, suppresses the release of sulfated glycosaminoglycans (sGAG) and TIMP1, and promotes matrix metalloproteinase production including MMP1, MMP3, and MMP13 (Chang et al., 2017). Moreover, T-2 toxin combined with low-selenium administration in a rat model proved to remarkably upregulate TGF-βR-I-II proteins to activate the TGF-β signaling pathway, which plays an important role in the induction of chondrocyte death in the deep zone of articular cartilage similar to human KBD (Zhang et al., 2018).

Extracellular matrix (ECM) degradation in chondrocytes is a typical pathological abnormality of KBD (He et al., 2018). However, it is unclear whether the T-2 toxin can directly induce this degradation. Specifically, the critical intermediate element and functional pathway remain need to be elucidated. Therefore, we identified candidate T-2 toxin-responsive genes and their effect on the ECM degradation process of KBD. In this study, the comparison between the previous gene expression profiles of chondrocytes in KBD and OA articular cartilage was analyzed, and TSG-6 was identified as a target gene verified by qRT-PCR and IHC. Then, TSG-6 was silenced and overexpressed to observe the effect on the change in the ECM of chondrocytes. Finally, gradients of HT-2 toxin, the main cell-lethal metabolite of T-2 toxin after ingestion, were added to C28/I2 cells for verification.

Materials and methods

All subjects provided informed consent (orally if the subject was illiterate) regarding the sample collection. The study protocol was approved by the ethics committee of Xi'an Jiaotong University (Approval No. 2018–206).

Articular cartilage sample recruitment

Diagnoses and degree classifications of patients were strictly applied according to the national criteria of KBD in China [WS/T 207–2010]. All subjects with alterations, such as defects and sclerosis on the bone end of phalanges combined with compression changes of the calcaneus and talus and enlarged/deformed limb joints (e.g., hand, elbow, knee, and ankle) manifested on X-ray, were diagnosed with KBD. Subjects were excluded if they were suffering or had previously suffered from any other osteoarticular diseases (such as rheumatoid arthritis, gout, or skeletal fluorosis) or any other type of macrosomia, osteochondrodysplasia disorder, or chronic disease (such as hypertension, diabetes, or coronary heart disease) or had received any treatment in the past six months. Clinical information was collected from patient records. Articular cartilage samples from KBD and OA patients were collected from individuals who underwent arthroplasty of the knee. Healthy controls were obtained from patients who suffered trauma or amputation due to an accident.

Donors signed a written informed consent form. Subjects were of Chinese Han lineage. For the verification of TSG-6 by qRT-PCR and immunohistochemistry, specimens of adult articular cartilage were collected from five KBD patients who had arthroplasty in the knee and five healthy subjects who had an amputation due to an accident in Xi'an Red Cross Hospital (Supplementary Table S1). For silencing and overexpression of TSG-6, articular cartilage samples were obtained from five KBD patients and three controls for the same reason described above (Supplementary Table S2).

Cartilage tissue collection and chondrocyte isolation

All articular cartilage samples, including all of the cartilage zones (including calcified) and subchondral bone, were harvested

from the lateral tibial plateau and obtained within 1 h after operation. Chondrocytes were isolated as follows: articular cartilage specimens were washed twice with sterile phosphate-buffered saline (PBS) with antibiotics (penicillin and streptomycin), cut into pieces (1 mm^3), and subjected to enzymatic digestion with 0.25% trypsin at 37 °C in an atmosphere of 5% CO_2 for up to 30 min. Cell suspensions were centrifuged at 1,000×g for 5 min, the supernatant was completely aspirated, and the cells were digested in basal media supplemented with 0.2% type II collagenase at 37 °C using an Eppendorf Thermomixer for 12–16 h.

Immunohistochemical verification

Cartilage tissues were fixed with 4% (w/v) paraformaldehyde for 24 h immediately after acquisition and decalcified in 10% (w/v) ethylenediaminetetraacetic acid disodium salt (EDTA-Na2) for 2–3 weeks. Samples were dehydrated in an alcohol series, cleared in xylene, and embedded in paraffin wax. Paraffin sections were cut into 5 μm sections, mounted on slides, and stored at room temperature until ready for staining. The paraffin-embedded sections were baked at 65 °C for 1 h, deparaffinized with xylene and then rehydrated in decreasing concentrations of ethanol. Endogenous peroxidase activity was blocked by 0.3% (w/v) hydrogen peroxide for 10 min at room temperature, and the sections were then washed with 1×PBS. Then, sections were incubated in a 10 mol/L urea solution diluted with water at 37 °C for 20 min and washed with 1×PBS. Sections were incubated in 0.1% trypsin/$CaCl_2$ at 37 °C and digested for another 20 min for antigen retrieval. After blocking with 5% (w/v) goat serum for 20 min at room temperature, anti-TNFAIP6 (1:100 dilution, PA5-76008, Invitrogen) antibodies and IgG as a negative control were applied onto the sections, and the samples were further incubated overnight at 4 °C. After washing with 1×PBS, sections were incubated using a human serum amyloid P (SAP) kit (solution B contains a goat anti-rabbit secondary antibody; Zhongshan, Jinqiao, Guangzhou, China) according to the manufacturer's instructions. The substrate 3,3′-diaminobenzidine was added to stain the sections, and hematoxylin counterstaining was performed. Finally, sections were dehydrated and mounted under alcohol-cleaned coverslips. All IHC staining was assessed under light microscopy by two pathologists who were blinded to the origin of the samples. Articular cartilage was divided into three cell morphologies, namely, the superficial, middle, and deep zones, according to light microscopy observation. Chondrocytes in the superficial zone were relatively small and flat and were oriented with the long axis parallel to the surface; cells in the middle zone were larger and more rounded and were randomly distributed in the matrix with fibers running in oblique directions; and cells in the deep zone were larger in size and were arranged in a columnar manner perpendicular to the surface. An assessment of staining throughout each cartilage zone included systematic counting of positive and negative cells starting from the cartilage surface and progressing down through all layers of cartilage. Five randomly chosen fields in each zone were counted at 50× magnification. The percentage of positive cells was calculated using the number of positively stained cells divided by the total number of cells (positively and negatively stained cells) in the chosen fields of view. Percentages of positive cells in different zones were calculated for each case and for the different groups.

Candidate gene identification and verification

According to the previous gene expression profile in cartilage from KBD (Wang et al., 2009) and OA microarray data GSE57218 from the GEO database in NCBI, TSG-6 (TNF-stimulated gene 6, also known as TNFAIP6) was identified as one of the common differentially expressed genes of importance in ECM reconstruction and by maintaining cellular homeostasis. Differential expression of TSG-6 at the mRNA level in KBD chondrocytes was verified respectively by quantitative real-time polymerase chain reaction (qRT-PCR). TaKaRa Company (Japan) synthesized the TSG-6 primer (forward-GGAGTGAAAGATGGGCATGCC, reverse-CTCATTTGGGAAGCCTGGA). Experiments were conducted according to the manufacturer's instructions.

TSG-6 silencing and overexpression

To clarify the effect of TSG-6 abnormality in the ECM degradation of KBD chondrocytes, we silenced this gene in KBD chondrocytes by RNA interference technology and then separately overexpressed TNF-α induction. A pretest was applied to determine the optimum multiplicity of infection (MOI) and transfection time of the lentivirus. We designed five culture conditions combined with three preset MOI (Supplementary Table S3) and selected the group with an infection efficiency above 80% along with good cell growth status. The lentivirus system (GENECHEM, Shanghai) consisted of the objective vector GV248, pHelper 1.0, and pHelper 2.0. We designed three sections of TSG-6 for construction of the lentivirus vector including TNFAIP6-RNAi(54322–1), TNFAIP6-RNAi(54323–1), and TNFAIP6-RNAi(54324–11) (Supplementary Table S4). Transfected 293T cells were introduced into the lentivirus system for 48–72 h and then the supernatant was collected. Next, viruses were concentrated and purified by 0.45 μm filtration and ultracentrifugation. Finally, the

quality of lentiviruses was determined involving physical form, sterility, and a titer test (*via* fluorescence).

Primarily, chondrocytes were cultured in 96-well plates (5×10^4 cells/well) for 36 h. Then, 200 μl TNF-α (0, 5, 15, 25, 50, 75, 100 ng/ml, 5 repeats for each concentration) were added to further culture chondrocytes with 5% CO_2 for 24 h, 48 h, and 72 h (three repeats for each time point) at 37 °C. TNF-α is not supposed to remarkably affect the proliferation of chondrocytes. We attempted cell culture with serum-free medium as described in previous studies (Zhang et al., 2012; Yoshida et al., 2018). However, a downward trend of cell proliferation was observed when the TNF-α concentration added up to 100 ng/ml by MTT. This might have been due to the serum-free culture condition. Therefore, we added 2% serum into the medium, but the proliferation still presented a downward trend at 100 ng/ml of TNF-α. Hence, 100 ng/ml was set as the upper limited concentration. Both trends emerged at 48 h. No significant difference in the affection of cell proliferation was observed between the previously mentioned concentrations, which suggests that there is no need to try so many sub-dosages. Finally, we selected four concentrations of TNF-α (0, 5, 25, 100 ng/ml) with serum-free medium and one optimistic time point (48 h) for TSG-6 induction.

qRT-PCR was used to detect the expression changes of extracellular matrix MMP1, MMP3, MMP13, COL2A1, and proteoglycan (Abcam, United Kingdom, 1:500 dilution) before and after treatments. Information about the primers (TaKaRa, Bejing) of each gene is listed in Supplementary Table S5. Proliferation and viability of chondrocytes was observed by MTT before and after the gradient concentrations of TNF-α treatment.

HT-2 toxin intervention

To explore whether TSG-6 is the intermediate element in the chondrocyte ECM degradation process caused by T-2 toxin, gradients (0, 12, 50, 100, 150, 200, 250 nM, and 200 μl) of HT-2 toxin were added to intervene with the C28/I2 chondrocytes. Conventional culture conditions and repeat counts were the same as the overexpression section.

The morphological changes of chondrocytes were observed under a light microscope. The MTT test showed a dose-dependent inhibition effect of HT-2 on cell viability at 0, 12, 50, and 100 nM. However, the inhibitory effect did not continue to significantly increase with the increase in intervention concentration from 100, 150, and 200–250 nM. Therefore, we selected 0, 12, 50, and 100 nM to utilize in the following study.

Total RNA was extracted in subsequent experiments. The expression of TSG-6, MMPs, COL2A1, and proteoglycans in the extracellular matrix were detected using qRT-PCR before and after intervention according to previously described methods (Wang et al., 2017).

Statistical analysis

A statistical analysis was performed using the Statistical Package for the Social Sciences for Windows version 18.0 (SPSS, Inc.). Individual samples were divided into triplicate for the study. The differences in means were determined by one-way analysis of variance (ANOVA) for multiple comparisons. Student's t-test was applied to determine the difference between two groups. The normality and homogeneity of variance of the data were tested before any further analyses. A non-parametric test was performed if the conditions for normality and homogeneity properties were not fulfilled, and p values less than 0.05 were considered significant.

Results

Identification of upregulated TSG-6 in KBD

We used immunohistochemistry to detect TSG-6 protein expression in KBD cartilage tissue (Figure 1A). Results showed that TSG-6 was upregulated in the superficial, middle, and deep zones of KBD cartilage compared to cartilage tissue from normal controls. We validated the gene expression of TSG-6 in patients with KBD using qRT-PCR and found that the results were consistent with those of the differentially expressed TSG-6 gene in KBD chondrocytes screened by previous microarray analysis (Figure 1B).

Upregulated TSG-6 was responsible for ECM degradation in KBD chondrocytes

The lentiviral transfection pretest determined that the optimal transfection conditions were MOI = 100, ENi.S + P(E)+virus for culture, and 96 h for the duration. TNFAIP6-RNAi (54322–1) had the best infection efficiency and showed a 72% interference rate (Table 1; Figure 1C). After TSG-6 silencing, the expression levels of MMP1, MMP3, MMP13, and proteoglycan were significantly decreased, while the expression of COL2A1 was significantly increased in KBD chondrocytes ($p < 0.05$, Figure 1D).

Generally, with increasing TNF-α concentration and prolonged time, cell proliferation was faintly affected regardless of whether the chondrocytes were cultured with or without serum. However, cell proliferation showed a decreasing trend at 48 h and a significant difference at a dosage 100 ng/ml without serum compared to the blank

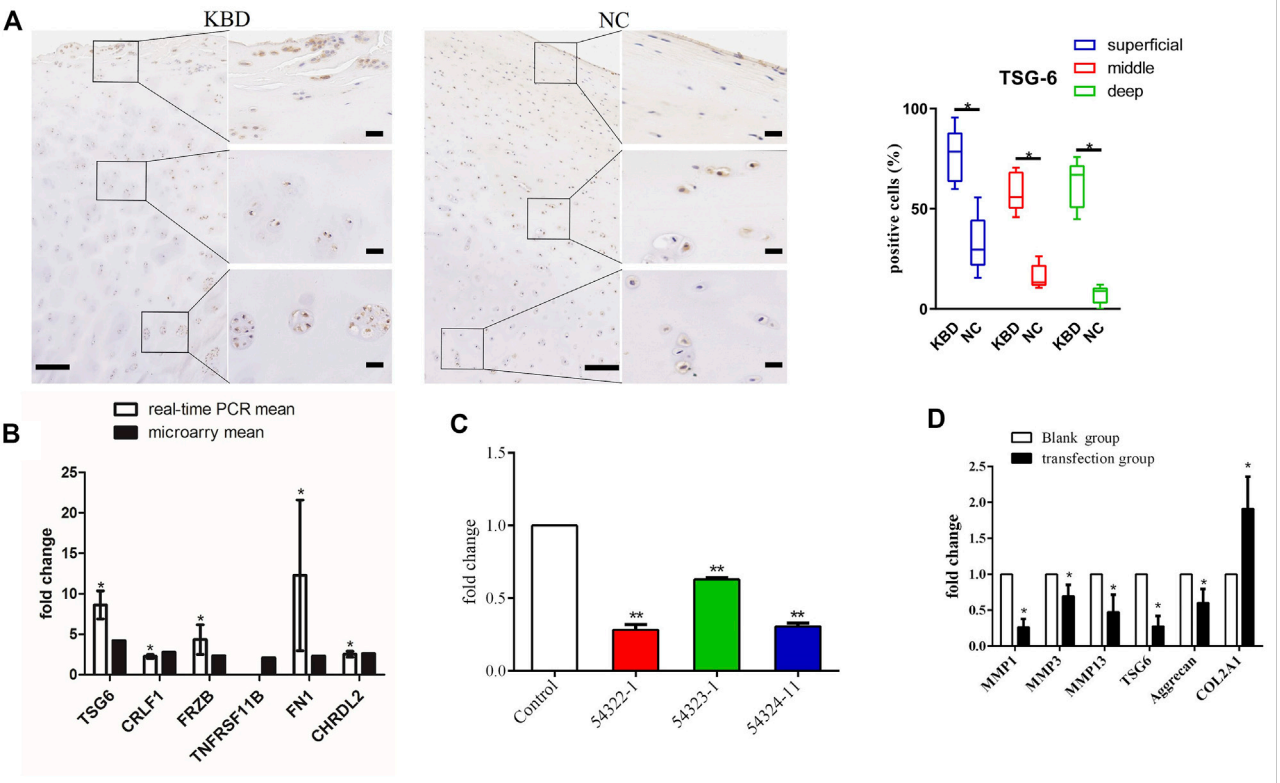


FIGURE 1
(A) Representative immunohistochemistry staining of TSG-6 in KBD and healthy cartilage tissues. Scale bar, left, 500 μm; right, 50 μm, and comparative quantification of positive cells of different areas (superficial, middle, and deep) between KBD and healthy cartilage tissues displayed by box plot (n = 3). *p < 0.05. **(B)** The mRNA levels of the selected genes between the microarray and quantitative RT-PCR in chondrocytes from patients with KBD (n = 3) and normal controls (n = 3). The histogram shows the fold change of selected genes as measured by microarray and quantitative RT-PCR; * indicates p < 0.05. **(C)** The selection of an optimistic transfection target for TSG-6, and the interference effect of TSG-6RNAi target sequence interfered with RNAi and detected by quantitative RT-PCR; ** indicates p < 0.001. **(D)** The mRNA levels of MMP1, MMP3, MMP13, aggrecan, and COL2A1 after silencing TSG-6; * indicates p < 0.05.

TABLE 1 Determination of the interference rate by qRT-PCR.

TSG-6 gene sections	Fold change	Interference rate (%)	p Value
Control	1	—	-
TSG-6-RNAi (54322–1)	0.28	72	<0.001
TSG-6-RNAi (54323–1)	0.63	37	<0.01
TSG-6-RNAi (54324–11)	0.31	69	<0.001

group (p < 0.05) (Figure 2A). And cell proliferation showed no significant difference when cultured with serum (Figure 2B).

Before intervention, the expression levels of MMP1, MMP3, and MMP13 in the KBD group notably increased, while the expression levels of COL2A1 and proteoglycan significantly decreased compared with those in the healthy control group. TNF-α intervention results demonstrated that the expression of TSG-6 increased in KBD and healthy chondrocytes in a dose-dependent manner. Increased MMPs and decreased COL2A1 and proteoglycan were more significant in TNF-α treated KBD chondrocytes than in the blank KBD group (p < 0.05, Figure 2C). Moreover, healthy chondrocytes treated with TNF-α also had significantly increased MMPs and decreased COL2A1 and proteoglycan compared with the blank control group (p < 0.05, Figure 2D).

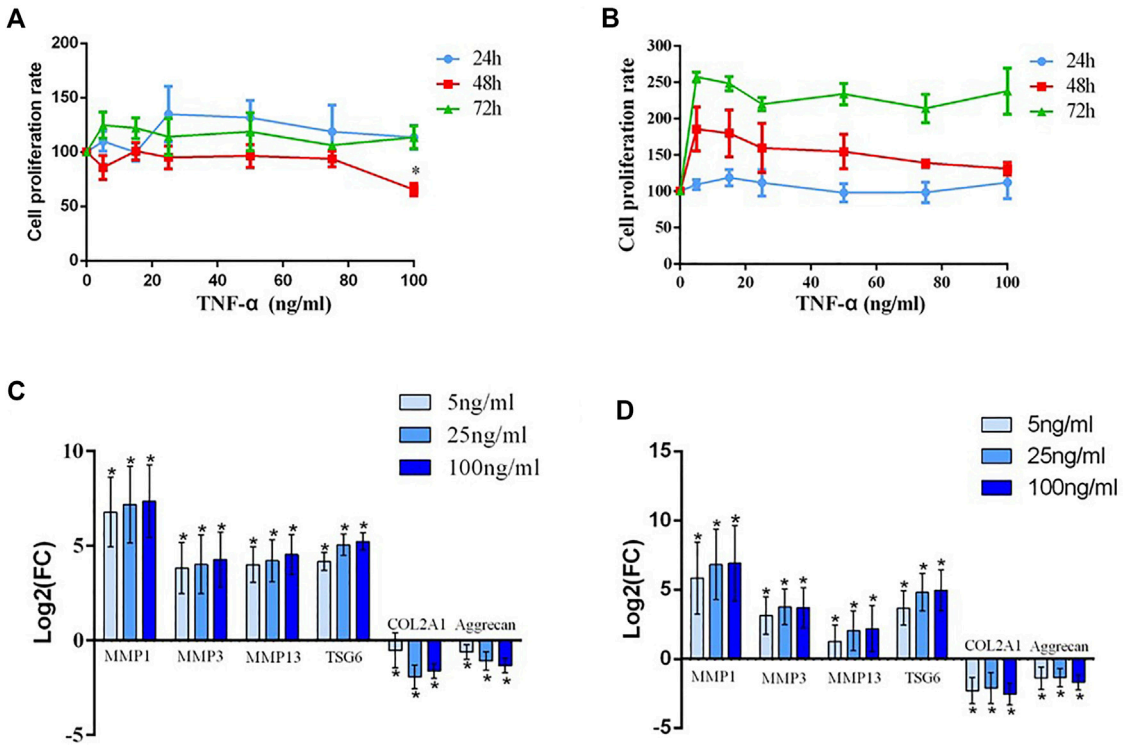


FIGURE 2
(A) The detection of the human chondrocyte proliferation rate after treatment with different concentrations (20, 40, 60, 80, and 100 ng/ml) of TNF-α and without serum for 24 h, 48 h, and 72 h; * indicates $p < 0.05$. **(B)** The detection of the human chondrocyte proliferation rate after treatment with different concentrations (20, 40, 60, 80, and 100 ng/ml) of TNF-α and with serum for 24 h, 48 h, and 72 h. **(C)** The mRNA expression of MMP1, MMP3, MMP13, TSG-6, COL2A1, and aggrecan after chondrocytes from patients with KBD were interfered with TNF-α; * indicates $p < 0.05$. **(D)** The mRNA expression of MMP1, MMP3, MMP13, TSG-6 COL2A1, and aggrecan after chondrocytes from normal controls were interfered with TNF-α; * indicates $p < 0.05$.

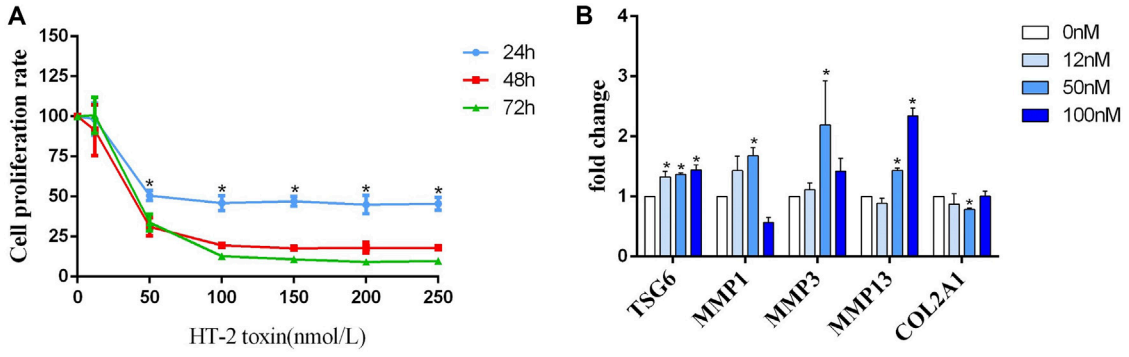


FIGURE 3
(A) The detection of the human C28/I2 cell proliferation rate after treatment with different concentrations (0, 50, 100, 150, 200, and 250 nmol/L) of HT-2 toxin for 24 h, 48 h, and 72 h; * indicates $p < 0.05$. **(B)** The mRNA expression of MMP1, MMP3, MMP13, TSG-6, and COL2A1 after human C28/I2 cells were treated with different concentrations (0, 12, 50, and 100 nmol/L) HT-2 toxin; * indicates $p < 0.05$.

TSG-6 may be an intermediate factor in the chondrocyte ECM degradation process caused by HT-2 toxin

C28/I2 chondrocytes were treated with gradient concentrations of HT-2 toxin to clarify whether TSG-6 mediates environmental exposure and chondrocyte ECM degradation. By using MTT and light microscopy, we found that the survival rate of chondrocytes was correspondingly reduced in a dose-dependent manner (Figure 3A). Compared with the blank control group, the mRNA levels of TSG-6, MMP1, MMP3, and MMP13 were increased in the HT-2 toxin intervention group (0, 12, 50, and 100 nM), while the expression of COL2A1 was significantly decreased ($p < 0.05$, Figure 3B).

Discussion

For the first time, this study found that TSG-6 was upregulated in KBD chondrocytes, which may function as an intermediate factor between HT-2 toxin exposure and cartilage injury by promoting the ECM degradation of chondrocytes.

TSG-6 is a member of the hyaluronan (HA) binding protein family and participates in maintaining ECM stability and cell migration (Wisniewski et al., 2014). This gene is described as a double-edged sword for OA. On one hand, it could combine with aggrecan and chondroitin-4-sulfate to inhibit inflammation and protect against articular cartilage. TSG-6 reduces bone resorption activity and promotes the formation of osteophytes that contain both bone and cartilage (Broeren et al., 2015). Link_TSG6 and rhTSG-6 proteins protect cartilage from degradation by suppressing the response of MSC-derived chondrocytes to inflammatory cytokines (IL-1 and TNF) and inhibiting their expression of ADAMTS4, ADAMTS5, and MMP13 (Day et al., 2016). On the other hand, TSG-6 in intermediate and deep regions of OA cartilage could block matrix assembly directly or indirectly through activation of matrix metalloproteases (MMPs), such as MMP3, MMP9, MMP12, and MMP13, and therefore promote the progression of cartilage damage (Chou et al., 2018). Therefore, TSG-6 activity was proposed as a promising independent biomarker for OA progression (Wisniewski et al., 2014). However, determining whether TSG-6 is a protective factor or a risk enhancer of KBD development needs to be explored.

We observed decreased MMP1, MMP3, and MMP13 along with increased COL2A1 after TSG-6 silencing *via* lentivirus in KBD chondrocytes. In contrast, overexpression of TSG-6 increased MMPs and decreased aggrecan and COL2A1 in a dose-dependent manner. These findings indicate that TSG-6 upregulation in KBD chondrocytes could disturb ECM homeostasis and accelerate the cartilage damage process.

KBD is a degenerative osteochondropathy that manifests with excessive apoptosis and ECM degradation. The association between T-2 toxin and the occurrence of KBD was established by Professor Yang in the 1990s based on an epidemiological investigation and *in vitro* research (Jianbo, 1998). Evidence showed that T-2 toxin aggregated excessively in cereals obtained from endemic areas compared to nonendemic areas (Jianbo et al., 1995). Chondrocytes treated with T-2 toxin upregulated matrix metalloproteinases including MMP1, MMP3, and MMP13 and downregulated aggrecan and COL2A1 at both the mRNA and protein levels (Zhang et al., 2010; Lu et al., 2012). However, we scarcely detected T-2 toxin in patient serum and chondrocytes, which led us to deliberate on the functional form of this pathogenic factor. In a rat model and one hour after consumption, T-2 toxin mainly transformed into HT-2 toxin and was distributed in multiple organs. HPLC-MS/MS determination results showed that the order of concentration of HT-2 was as follows: femur > knee joint > rib cartilage > liver > skeletal muscle > heart > kidney. The conversion ratio occurred in the range of 68.2–90.7% (Yu et al., 2017b). In the T-2 toxin treated chondrocytes, after 48 h, the concentration of T-2 decreased from 20 ng/ml to 6.67 ± 1.02 ng/ml, while the concentration of HT-2 increased from 0 ng/ml to 6.88 ± 1.23 ng/ml (Yu et al., 2017a). This suggests that T-2 toxin was almost completely metabolized into HT-2 at the end of 48 h. Additionally, it is suggested that both T-2 toxin and HT-2 toxin can lead to apoptosis and autophagy in human chondrocytes (Yu et al.,2019b). Hence, this study discovered the effect of HT-2 toxin on the ECM degradation of KBD.

The MTT assay revealed that HT-2 could reduce the proliferation of C28/I2 cells in a dose- and time-dependent manner, which was consistent with results induced by the T-2 toxin. What attracted our attention was that cell viability did not further decrease any time point when the HT-2 concentration reached 100 nM. Therefore, the cytotoxic effect of the HT-2 toxin on chondrocytes may have a threshold.

Compared with the blank control group, the expression of TSG-6 along with MMP1, MMP3, and MMP13 was upregulated in the HT-2 toxin intervention group in a dose-dependent manner, while the expression of proteoglycan and COL2A1 significantly decreased. Therefore, based on the effect of TSG-6 on accelerating ECM degradation in KBD chondrocytes as proven by viral silencing and TNF-α induction, TSG-6 might work as an intermediate factor between HT-2 exposure and cartilage damage in the development of KBD.

During developmental and inflammatory processes, heavy chains (HCs) from the serum-derived proteoglycan inter-α-inhibitor (IαI) are covalently transferred to HA *via* the TSG-6 enzyme to form an HC-HA complex, which is similar to osteoarthritis and rheumatoid arthritis (Lauer et al., 2015; Torihashi et al., 2015). In the transfer action, HA substitutes for the chondroitin sulfate (CS) chain on bikunin to bind to the HCs *via* an ester bond. Although bikunin and its CS chain are not

components of HA-HC complexes, they are integral for HA-HC complex formation (Lord et al., 2013). The organized pericellular HA coat is anchored to its principal cell surface receptor cluster of differentiation 44 (CD44). TSG-6-dependent HC-HA interactions induce the formation of a pericellular HA matrix, which facilitates CD44 relocalization within the cell membrane. This receptor complex then activates the downstream signaling cascade that is responsible for reorganization of the actin cytoskeleton (Martin et al., 2016). Significantly elevated levels of HA and CD44 were observed in the serum of both juvenile and adult KBD patients compared to healthy controls in non-KBD areas (Cao et al., 2008; Yu et al., 2014). T-2 toxin was identified as responsible for increased CD44 and HA in cultured articular chondrocytes (Li et al., 2008). Therefore, the effect of HT-2 toxin in inhibiting aggrecan synthesis, promoting aggrecanases, and consequently inducing ECM degradation in chondrocytes is likely associated with an abnormal HA-HC transferring action, which is interesting to elucidate and promote the study of KBD pathogenesis.

Conclusion

Unlike the chondroprotective function in osteoarthritis, the upregulation of the TSG-6 gene may play an important role in promoting the ECM degradation of KBD chondrocytes. Although it remains unknown whether the pathological HC-HA complex formation process is responsible for TSG-6 induced articular ECM degradation, this gene could be an intermediate factor between HT-2 toxin exposure and chondrocyte injury in KBD. Additionally, an *in vivo* study would make this promising result more convincing.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, and further inquiries can be directed to the corresponding author.

Ethics statement

Studies involving human participants were reviewed and approved by Xi'an Jiaotong University. The patients/participants provided their written informed consent prior to participating in this study.

Author contributions

YN, PZ, and HZ designed the study. YG, MH, RH, YL, FC, FZ, SC, YW, SL, and CW performed the analysis and interpretation of the data. YN, XW, LY, XG, and HZ analyzed the results and drafted the manuscript. All authors critically reviewed and amended the manuscript, and all authors read and approved the final manuscript.

Funding

This study was financially supported by the National Natural Science Foundation of China (82273752, 81903263, and 81620108026) and the China Postdoctoral Science Foundation (2021M692543 and 2022M712526).

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1064565/full#supplementary-material

References

Broeren, M. G., Bennink, M. B., Arntz, J., Blom, A. B., Van Den Berg, W. B., and Van De Loo, F. A. (2015). AB0101 Viral expression of TSG-6 can stimulate osteophyte formation in experimental osteoarthritis. *Ann. Rheum. Dis.* 74, 924.2–924. doi:10.1136/annrheumdis-2015-eular.4767

Cao, J., Li, S., Shi, Z., Yue, Y., Sun, J., Chen, J., et al. (2008). Articular cartilage metabolism in patients with kashin-beck disease: An endemic osteoarthropathy in China. *Osteoarthr. Cartil.* 16, 680–688. doi:10.1016/j.joca.2007.09.002

Chang, Y., Wang, X., Sun, Z., Jin, Z., Chen, M., Lammi, M. J., et al. (2017). Inflammatory cytokine of IL-1β is involved in T-2 toxin-triggered chondrocyte injury and metabolism imbalance by the activation of Wnt/β-catenin signaling. *Mol. Immunol.* 91, 195–201. doi:10.1016/j.molimm.2017.08.019

Chou, C. H., Attarian, D. E., Wisniewski, H. G., Band, P. A., and Kraus, V. B. (2018). TSG-6 - a double-edged sword for osteoarthritis (OA). *Osteoarthr. Cartil.* 26, 245–254. doi:10.1016/j.joca.2017.10.019

Day, A. J., Drummond, S. P., Anand, S., Bartnik, E., and Milner, C. M. (2016). A novel chondroprotective property of tsg-6 has therapeutic potential for oa. *Osteoarthr. Cartil.* 24, S19–S20. doi:10.1016/j.joca.2016.01.064

Dyer, D. P., Salanga, C. L., Johns, S. C., Valdambrini, E., Fuster, M. M., Milner, C. M., et al. (2016). The anti-inflammatory protein TSG-6 regulates chemokine function by inhibiting chemokine/glycosaminoglycan interactions. *J. Biol. Chem.* 291, 12627–12640. doi:10.1074/jbc.M116.720953

Ermakov, V., Bech, J., Gulyaeva, U., Tyutikov, S., Safonov, V., Danilova, V., et al. (2019). Relationship of the mobile forms of calcium and strontium in soils with their accumulation in meadow plants in the area of Kashin-Beck endemia. *Environ. Geochem. Health* 42, 159–171. doi:10.1007/s10653-019-00323-5

Han, L., Yang, X., Sun, W., Li, Z., Ren, H., Li, B., et al. (2018). The study of GPX3 methylation in patients with Kashin-Beck Disease and its mechanism in chondrocyte apoptosis. *Bone* 117, 15–22. doi:10.1016/j.bone.2018.08.017

He, Y., Zhang, Y., Zhang, D., Zhang, M., Wang, M., Jiang, Z., et al. (2018). 3-morpholinosydnonimine (SIN-1)-induced oxidative stress leads to necrosis in hypertrophic chondrocytes *in vitro*. *Biomed. Pharmacother.* 106, 1696–1704. doi:10.1016/j.biopha.2018.07.128

Jianbo, Y., Dianjun, S., and Zhiwu, W. (1995). Determination of T-2 toxin in the staple food from the sick families in Kashin-Beck disease (KBD) areas. *Chin. Jouranl Endemiology* 14, 146–149.

Jianbo, Y. (1998). *Research on the cause of Kaschin-Beck disease*. Harerbin: Heilongjiang Science and Technology Press.

Lauer, M. E., Loftis, J., De La Motte, C., and Hascall, V. C. (2015). Analysis of the heavy-chain modification and TSG-6 activity in pathological hyaluronan matrices. *Methods Mol. Biol.* 1229, 543–548. doi:10.1007/978-1-4939-1714-3_42

Li, S. Y., Cao, J. L., Shi, Z. L., Chen, J. H., Zhang, Z. T., Hughes, C. E., et al. (2008). Promotion of the articular cartilage proteoglycan degradation by T-2 toxin and selenium protective effect. *J. Zhejiang Univ. Sci. B* 9, 22–33. doi:10.1631/jzus.B071322

Lin, X., Shao, W., Yu, F., Xing, K., Liu, H., Zhang, F., et al. (2019). Individual and combined toxicity of T-2 toxin and deoxynivalenol on human C-28/I2 and rat primary chondrocytes. *J. Appl. Toxicol.* 39, 343–353. doi:10.1002/jat.3725

Lord, M. S., Day, A. J., Youssef, P., Zhuo, L., Watanabe, H., Caterson, B., et al. (2013). Sulfation of the bikunin chondroitin sulfate chain determines heavy chain.hyaluronan complex formation. *J. Biol. Chem.* 288, 22930–22941. doi:10.1074/jbc.M112.404186

Lu, M., Cao, J., Liu, F., Li, S., Chen, J., Fu, Q., et al. (2012). The effects of mycotoxins and selenium deficiency on tissue-engineered cartilage. *Cells Tissues Organs* 196, 241–250. doi:10.1159/000335046

Martin, J., Midgley, A., Meran, S., Woods, E., Bowen, T., Phillips, A. O., et al. (2016). Tumor necrosis factor-stimulated gene 6 (TSG-6)-mediated interactions with the inter-α-inhibitor heavy chain 5 facilitate tumor growth factor β1 (TGFβ1)-dependent fibroblast to myofibroblast differentiation. *J. Biol. Chem.* 291, 13789–13801. doi:10.1074/jbc.M115.670521

Sun, L., Cui, S., Deng, Q., Liu, H., Cao, Y., Wang, S., et al. (2019). Selenium content and/or T-2 toxin contamination of cereals, soil, and children's hair in some areas of heilongjiang and gansu provinces, China. *Biol. Trace Elem. Res.* 191, 294–299. doi:10.1007/s12011-018-1620-7

Torihashi, S., Ho, M., Kawakubo, Y., Komatsu, K., Nagai, M., Hirayama, Y., et al. (2015). Acute and temporal expression of tumor necrosis factor (TNF)-alpha-stimulated gene 6 product, TSG6, in mesenchymal stem cells creates microenvironments required for their successful transplantation into muscle tissue. *J. Biol. Chem.* 290, 22771–22781. doi:10.1074/jbc.M114.629774

Wang, W. Z., Guo, X., Duan, C., Ma, W. J., Zhang, Y. G., Xu, P., et al. (2009). Comparative analysis of gene expression profiles between the normal human cartilage and the one with endemic osteoarthritis. *Osteoarthr. Cartil.* 17, 83–90. doi:10.1016/j.joca.2008.05.008

Wang, X., Ning, Y., Zhang, P., Yang, L., Wang, Y., and Guo, X. (2017). Chondrocytes damage induced by T-2 toxin via Wnt/β-catenin signaling pathway is involved in the pathogenesis of an endemic osteochondropathy, Kashin-Beck disease. *Exp. Cell Res.* 361, 141–148. doi:10.1016/j.yexcr.2017.10.012

Wisniewski, H. G., Colon, E., Liublinska, V., Karia, R. J., Stabler, T. V., Attur, M., et al. (2014). TSG-6 activity as a novel biomarker of progression in knee osteoarthritis. *Osteoarthr. Cartil.* 22, 235–241. doi:10.1016/j.joca.2013.12.004

Wisniewski, H. G., Naime, D., Hua, J. C., Vilcek, J., and Cronstein, B. N. (1996). TSG-6, a glycoprotein associated with arthritis, and its ligand hyaluronan exert opposite effects in a murine model of inflammation. *Pflugers Arch.* 431, R225–R226. doi:10.1007/BF02346350

Yang, L., Zhang, J., Li, X., Xu, C., Wang, X., and Guo, X. (2019). Expression profiles of selenium-related genes in human chondrocytes exposed to T-2 toxin and deoxynivalenol. *Biol. Trace Elem. Res.* 190, 295–302. doi:10.1007/s12011-018-1560-2

Yoshida, K., Nakashima, A., Doi, S., Ueno, T., Okubo, T., Kawano, K. I., et al. (2018). Serum-free medium enhances the immunosuppressive and antifibrotic abilities of mesenchymal stem cells utilized in experimental renal fibrosis. *Stem Cells Transl. Med.* 7, 893–905. doi:10.1002/sctm.17-0284

Yu, F. F., Lin, X. L., Wang, X., Liu, H., Yang, L., Goldring, M. B., et al. (2017a). Selenium promotes metabolic conversion of T-2 toxin to HT-2 toxin in cultured human chondrocytes. *J. Trace Elem. Med. Biol.* 44, 218–224. doi:10.1016/j.jtemb.2017.08.009

Yu, F. F., Lin, X. L., Yang, L., Liu, H., Wang, X., Fang, H., et al. (2017b). Comparison of T-2 toxin and HT-2 toxin distributed in the skeletal system with that in other tissues of rats by acute toxicity test. *Biomed. Environ. Sci.* 30, 851–854. doi:10.3967/bes2017.115

Yu, F. F., Qi, X., Shang, Y. N., Ping, Z. G., and Guo, X. (2019b). Prevention and control strategies for children kashin-beck disease in China: A systematic review and meta-analysis. *Med. Baltim.* 98, e16823. doi:10.1097/MD.0000000000016823

Yu, F. F., Xia, C. T., Fang, H., Han, J., Younus, M. I., and Guo, X. (2014). Evaluation of the therapeutic effect of treatment with intra-articular hyaluronic acid in knees for kashin-beck disease: A meta-analysis. *Osteoarthr. Cartil.* 22, 718–725. doi:10.1016/j.joca.2014.04.012

Zhang, A., Cao, J. L., Yang, B., Chen, J. H., Zhang, Z. T., Li, S. Y., et al. (2010). Effects of moniliformin and selenium on human articular cartilage metabolism and their potential relationships to the pathogenesis of Kashin-Beck disease. *J. Zhejiang Univ. Sci. B* 11, 200–208. doi:10.1631/jzus.B0900074

Zhang, S., He, H., Day, A. J., and Tseng, S. C. (2012). Constitutive expression of inter-α-inhibitor (IαI) family proteins and tumor necrosis factor-stimulated gene-6 (TSG-6) by human amniotic membrane epithelial and stromal cells supporting formation of the heavy chain-hyaluronan (HC-HA) complex. *J. Biol. Chem* 287, 12433–12444. doi:10.1074/jbc.M112.342873

Zhang, Y., He, Y., Zhang, D., Zhang, M., Wang, M., Ma, T., et al. (2018). Death of chondrocytes in Kashin-Beck disease: Apoptosis, necrosis or necroptosis? *Int. J. Exp. Pathol.* 99, 312–322. doi:10.1111/iep.12297

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 22 November 2022
DOI 10.3389/fgene.2022.1051047

OPEN ACCESS

EDITED BY
Fang Fang Yu,
Zhengzhou University, China

REVIEWED BY
Mingyi Yang,
Xi'an Honghui Hospital, China
Juan Yue,
Civil Aviation General Hospital, China

*CORRESPONDENCE
Tianwei Sun,
suntianweitj@163.com

SPECIALTY SECTION
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

RECEIVED 22 September 2022
ACCEPTED 20 October 2022
PUBLISHED 22 November 2022

CITATION
Yu H, Luo G, Sun T and Tang Q (2022), Causal effects of homocysteine levels on the components of sarcopenia: A two-sample mendelian randomization study.
Front. Genet. 13:1051047.
doi: 10.3389/fgene.2022.1051047

Causal effects of homocysteine levels on the components of sarcopenia: A two-sample mendelian randomization study

Hongwei Yu[1], Gan Luo[2], Tianwei Sun[3]* and Qiong Tang[4]

[1]School of Medicine, Nankai University, Tianjin, China, [2]Graduate School of Tianjin Medical University, Tianjin, China, [3]Department of Spinal Surgery, Tian-jin Union Medical Centre, Nankai University People's Hospital, Tianjin, China, [4]Department of Respiratory Medicine, Tian-jin Union Medical Centre, Nankai University People's Hospital, Tianjin, China

Background: Currently, it is unclear whether there is a causal association between genetically predicted plasma homocysteine (Hcy) levels and the risk of sarcopenia. We performed a Mendelian randomization (MR) study to assess the association between circulating Hcy levels and the components [grip strength, walking pace, and appendicular lean mass (ALM)] of sarcopenia.

Methods: Independent single nucleotide polymorphisms (SNPs) significantly associated with plasma Hcy levels served as instrumental variables. Summary-level data regarding the components of sarcopenia. Were obtained from the UK Biobank. Inverse variance weighted (IVW) as the primary method was used for Mendelian randomization (MR) analysis. We also use four models, weighted median, MR-Egger regression, Maximum likelihood, and Penalised weighted median, as supplementary methods to IVW. The MR-Egger intercept test, Cochran's Q test, and "leave-one-out" sensitivity analysis were performed to evaluate the horizontal pleiotropy, heterogeneities, and stability of the causal association between Hcy levels and the components of sarcopenia.

Results: The IVW-MR analysis suggested significant negative associations of increased plasma Hcy levels with grip strength (right: effect = −0.036, SE = 0.032, p = 5.53E-4; left: effect = −0.045, SE = 0.010, p = 1.45E-5), walking pace (effect = −0.038, SE = 0.011, p = 3.18E-4), and ALM (effect = −0.058, 0.013, p = 1.03E-5). However, there were no significant associations of decreased plasma Hcy levels with grip strength (right: effect = 0.005, SE = 0.021, p = 0.82; left: effect = −0.006, SE = 0.014, p = 0.64), walking pace (effect = 0.01, 0.020, p = 0.61), or ALM (effect = -0.034, SE = 0.018, p = 0.06).The accuracy and robustness of these findings were confirmed by sensitivity tests.

Conclusion: Increased circulating Hcy levels were associated with lower grip strength, slower walking pace, and decreased ALM.

KEYWORDS

homocysteine, grip strength, walking pace, appendicular lean mass, mendelian randomization

1 Introduction

Sarcopenia is defined as an age-related loss of muscle mass and function (Cruz-Jentoft et al., 2010). In 2018, the European Working Group on Sarcopenia in Older People (EWGSOP) redefined sarcopenia to emphasize that muscle function (muscle strength and physical performance) is the most important component of sarcopenia, as muscle strength is a better predictor of poor outcomes than muscle mass (Cruz-Jentoft et al., 2019). Worldwide, as many as 50 million people have sarcopenia, with sarcopenia most prevalent in elderly individuals. With aging of the global population, the prevalence of sarcopenia is expected to increase (Fernandes et al., 2022); the number of people with sarcopenia is predicted to reach 500 million by 2050 (Bhasin et al., 2020). Sarcopenia leads to decreased muscle strength and flexibility, fatigue, and increased risks of falls, fractures, physical disability, and death (Álvarez-Bustos et al., 2022; Yu et al., 2022); these symptoms and risks severely reduce the quality of life of patients and impose a greater burden on families and society (Álvarez-Bustos et al., 2022). An increasing number of studies have investigated sarcopenia, identifying possible causative factors of nutritional dysfunction (Pagano et al., 2018), altered endocrine function (Fielding et al., 2011), and a systemic inflammatory response (Dolan et al., 2019). However, the causative factors of sarcopenia are complex and diverse; thus, they merit further investigation.

Homocysteine (Hcy) is a sulfur-containing amino acid, that is, an intermediate product of methionine and cysteine metabolism (Selhub, 1999). A variety of factors can lead to increased plasma levels of Hcy, resulting in hyperhomocysteinemia. Increased Hcy levels can damage cells, tissues and organs and affect the body (Yoo et al., 2022). Clinical studies have found an association between plasma Hcy levels and risks of cardiovascular disease (Yuan et al., 2021), cerebrovascular disease (Spence and Hankey, 2022), kidney disease (Samson et al., 2022) and skeletal muscle system disease (Veeranki et al., 2015); thus, Hcy levels merit as much clinical monitoring as blood glucose, blood pressure and blood lipids. Epidemiological studies have shown an association between increased plasma Hcy levels and decreased muscle strength and physical function (Ter Borg et al., 2016; Lee et al., 2020). Observational studies have further expanded these associations of decreased muscle strength to include decreased muscle mass (McDermott et al., 2007; Choi et al., 2022). However, the findings of recent observational studies are somewhat contradictory. Cross-sectional and longitudinal studies by Granic et al. (2017) and Eguchi et al. (2021), respectively, did not find an association between Hcy levels and muscle function; similarly, long-term longitudinal studies by Elshorbagy et al. (2008) and Swart et al. (2016) did not find an association between Hcy levels and muscle mass. These conflicting findings may be because observational studies are inevitably subject to confounding factors and/or reverse causation. Furthermore, even if observational studies reveal consistent associations between Hcy levels and sarcopenia, it remains unclear whether Hcy levels are a causal factor for sarcopenia or a concomitant manifestation of sarcopenia due to a common causal factor. Randomized controlled trials (RCTs) are useful for controlling for confounding factors and directly inferring a causal relationship between variables (e.g., plasma Hcy levels and sarcopenia). However, in reality, RCTs are difficult to perform due to the expenses of recruiting human participants and obtaining materials, lengthy durations, and possible ethical constraints (Davies et al., 2018). Mendelian randomization (MR) studies complement RCTs; their design is based on the principle of random assignment of alleles at gamete formation and the use of the genotype as an instrumental variable to examine the intermediate phenotype and thus infer its causal association with the disease state. Hence, the effect estimates of MR studies are not influenced by confounding factors or reverse causal associations (Davies et al., 2018).

Muscle and bone, two important components of the locomotor system, are jointly regulated by a variety of physiological factors. The combination of sarcopenia and osteoporosis is referred to as dysmobility syndrome (Chen et al., 2021); this syndrome is one of the leading causes of falls, fractures, disability and even death in elderly individuals. Previous MR studies have identified that decreased plasma Hcy levels can causally increase bone mineral density (Wang et al., 2021). Here, we used MR analysis to investigate whether there is a causal association between increased or decreased plasma Hcy levels and the components of sarcopenia (grip strength, walking pace, and ALM), respectively. In addition, we used the MR-Egger intercept test, Cochran's Q test, and"leave-one-out" sensitivity analysis to evaluate the horizontal pleiotropy, heterogeneities, and stability of the causal association between Hcy levels and the components of sarcopenia.

2 Methods

2.1 Study design

To ensure that reliable causal relationships are obtained, MR studies need to satisfy the following three core assumptions (Cruz-Jentoft et al., 2010, Pagoni et al., 2019) relevance, i.e., the instrumental variables are strongly correlated with the exposure factors; (Cruz-Jentoft et al., 2019); independence, i.e., the instrumental variables are not correlated with confounders of exposure and outcome; and (Fernandes et al., 2022) exclusion restriction, i.e., the instrumental variables can only affect the outcome through the exposure factors.

2.2 Data sources

Genetic variants associated with plasma Hcy levels were derived from the genome-wide association study (GWAS) meta-analysis published by van Meurs et al. (2013). This meta-analysis included nine cohorts of 44,147 individuals of European ancestry and identified 18 single nucleotide polymorphisms (SNPs) significantly associated with plasma Hcy levels. For more information on the above nine cohorts, please refer to the Supplementary Presentation. Regarding the sarcopenia components, grip strength is moderately correlated with strength in other parts of the body; thus, it can be used as a reliable substitute for whole-body strength measurements (Cruz-Jentoft et al., 2019). Walking pace is considered a fast, safe and highly reliable test for sarcopenia and is widely used in clinical practice (Cruz-Jentoft et al., 2019). Therefore, we selected grip strength and walking speed as the outcomes. GWAS data for both grip strength and walking pace were obtained from the UK Biobank (Cox, 2018). In brief, the UK Biobank is a large prospective cohort study that recruited over 500,000 participants (age range: 43%–79% and 54% female) in the UK from 2006 to 2010. Grip strength data (n = 461,089) were collected using the Jamar dynamometer, which is a validated and widely used method of measuring grip strength (Cox, 2018). Self-reported walking pace (n = 335,349) was obtained from answers to the question "How would you describe your usual walking speed?" The response options were "slow", "steady/average", and "fast". Slow was defined as less than three miles per hour, steady/average was defined as between 3–4 miles per hour, and fast was defined as more than four miles per hour (Cox, 2018). Because direct measurement of muscle mass is usually not possible, lean body mass is considered a valid measure of muscle mass. ALM, the most commonly used muscle mass approximator in sarcopenia studies, is widely used in the EWGSOP (2) and Asian Working Group for Sarcopenia (AWGS) (Chen et al., 2020) diagnostic criteria for sarcopenia. We chose ALM rather than whole-body lean mass (WBLM) because ALM reduces the influence of the systemic water component, nonadipose soft tissue, cardiac muscle and vascular smooth muscle on the results (Cruz-Jentoft et al., 2019). ALM data were extracted from a GWAS by Pei et al. (2020) (n = 450,243), in which study samples were obtained from the UK Biobank. These ALM data were obtained by using bioelectric impedance analysis (BIA) to determine the sum of the deglutition masses of the arms and legs (Cox, 2018). The data sources for both exposure and outcome are summarized in Table 1.

2.3 Extraction of instrumental variables

To satisfy assumption 1 (i.e., the relevance assumption), we selected SNPs from the initial 18 SNPs that were independently (R^2 < 0.001, window size = 10,000 kb) and significantly associated with plasma Hcy levels at a genome-wide significance level (p < 5E−8) as potential instrumental variables (IVs). In addition, to ensure that the potential IVs had sufficient power to detect the causal influence of exposure on the outcomes, we calculated the F-statistics of the potential IVs using an online tool (https://sb452.shinyapps.io/overlap). IVs with F-statistics >10 were considered to have sufficiently robust estimation power to determine causal effects and were considered candidate IVs. To satisfy assumptions two and 3 (i.e., the independence and exclusion-restriction assumptions, respectively), we first identified confounders associated with plasma Hcy levels, muscle function and mass from three meta-analyses (Han et al., 2017; Shen et al., 2019; Gao et al., 2021), and then the PhenoScanner database (Version 2; http://www.phenoscanner.medschl.cam.ac.uk/) was used to examine the association of each candidate IV with the confounders. Candidate IVs were excluded if they had a significant association with outcomes or confounders (p < 5E-8). Once we identified the included IVs, we performed the first round of MR analysis. As the exact biological functions of many genetic variants remain unknown, we also used Mendelian randomization pleiotropy residual sum and outlier (MR-PRESSO) to identify and remove pleiotropic SNPs (Verbanck et al., 2018). If pleiotropic SNPs were identified by MR-PRESSO in the first round of analysis, we excluded them and before performing the second round of MR analysis.

2.4 Mendelian randomization analysis

The inverse variance-weighted Mendelian randomization (IVW-MR) method was the primary method used to assess causal effects between plasma Hcy levels and grip strength, walking pace, and ALM. IVW-MR uses a meta-analysis to combine Wald estimates for each SNP to obtain an overall estimate of the effect of exposure on outcomes (Burgess et al., 2016). IVW-MR can use both fixed- and random-effects models. In the present study, if significant heterogeneity was observed (p < 0.05), a random-effects IVW-MR model was applied. We also used MR-Egger (Bowden et al., 2015), weighted median (Bowden et al., 2016), maximum likelihood (Milligan, 2003), and penalized weighted median (Bowden et al., 2015) methods to complement and validate the results of the IVW-MR analysis. Detailed information regarding the above MR methods can be found elsewhere. If multiple comparisons were conducted, a Bonferroni-corrected p value < 0.0125 (0.05/4) was considered statistically significant. All analyses were performed using R version 4.1.1 with the two-sample MR package.

TABLE 1 Data sources used for the MR analysis ALM:appendicular lean mass.

	Traits	Sample size	Number of SNPS	Population	Consortium or cohort study
Exposure	Plasma Homocysteine levels	44147	2,090,256	European	Van Meurs, et al./PMID:23824729
Outcomes	Grip strength (right)	461,089	9,851,867	European	UK Biobank (MRC IEU)/(https://doi.org/10.5523/bris.2fahpksont1zi26xosyamqo8rr)
	Grip strength (left)	461,026	9,851,867	European	UK Biobank (MRC IEU)/(https://doi.org/10.5523/bris.2fahpksont1zi26xosyamqo8rr)
	Walking pace	335,349	10,894,596	European	UK Biobank (Neale Lab)/(http://www.nealelab.is/uk-biobank)
	ALM	450243	18,071,518	European	Pei YF, et al./PMID:33097823

TABLE 2 Summary statistics for homocysteine-associated SNPs using as instrumental variables.

SNP	Chrr	Nearby gene	EA	OA	EAF	Beta	SE	*p* Value	F	Homocysteine levels
rs9369898	6	MUT	A	G	0.62	0.045	0.007	2.20E−10	42	unit increase
rs154657	16	DPEP1	A	G	0.47	0.096	0.007	1.70E−43	204	unit increase
rs4660306	1	MMACHC	T	C	0.33	0.043	0.007	2.30E−09	37	unit increase
rs548987	6	SLC17A3	C	G	0.13	0.06	0.01	1.10E−08	36	unit increase
rs1801222	10	CUBN	A	G	0.34	0.045	0.007	8.40E−10	41	unit increase
rs12780845	10	CUBN	A	G	0.65	0.053	0.009	7.80E−10	56	unit increase
rs838133	19	FUT2	A	G	0.45	0.042	0.007	7.50E−09	39	unit increase
rs2275565	1	MTR	G	T	0.21	−0.054	0.009	2.00E−10	43	unit decrease
rs2251468	12	HNF1A	C	A	0.65	−0.051	0.007	1.30E−12	53	unit decrease
rs234709	21	CBS	C	T	0.45	−0.072	0.007	3.90E−24	113	unit decrease
rs7130284	11	NOX4	T	C	0.07	−0.124	0.013	1.90E−20	89	unit decrease
rs42648	7	GTPB10	A	G	0.4	−0.039	0.007	2.00E−08	33	unit decrease

SNP: single-nucleotide polymorphism; Chr: chromosome; EA: effect allele; OA, other allele; EAF: effect allele frequency. F: F statistic; SE: standard error.

2.5 Sensitivity analysis

To further ensure the robustness of our MR estimates, we performed the following sensitivity analyses. First, Cochran's Q test was used to quantify the heterogeneity among the genetic instruments (Pierce and Burgess, 2013). Second, the MR-Egger intercept test (Burgess and Thompson, 2017) and MR-PRESSO global test (Verbanck et al., 2018)were used to examine whether our MR analysis was affected by horizontal pleiotropy. Third, we performed a leave-one-out analysis to examine whether the overall estimates were disproportionately affected by specific SNPs.

3 Results

3.1 Single nucleotide polymorphisms selection and validation

Of the initial 18 SNPs, we excluded five (rs12134663, rs12921383, rs7422339, rs2851391, and rs957140) due to linkage disequilibrium. Additionally, rs1801133 was excluded due to its significant association with folic acid (P = 1E-28). The remaining 12 SNPs were categorized as IVs in our study (Table 2). Seven of these SNPs (rs9369898, rs154657, rs4660306, rs548987, rs1801222, rs12780845, and rs838133) were associated with increases in genetically predicted plasma levels of Hcy, and five SNPs (rs2275565, rs2251468, rs234709, rs7130284, and rs42648) were associated with decreases in genetically predicted plasma levels of Hcy. All 12 SNPs were valid (F > 10).

3.2 Associations of increased plasma hcy levels with grip strength, walking pace, and appendicular lean mass

The IVW-MR analysis suggested significant negative associations between increased plasma Hcy levels and grip strength (right: effect = −0.036, SE = 0.032, p = 5.53E-4; left: effect = −0.045, SE = 0.010, p = 1.45E-5), walking pace

10.3389/fgene.2022.1051047

TABLE 3 MR analysis of association between homocysteine levels and the three components of sarcopenia.

Traits	SNP(n)	Methods	Beta	SE	p Value
Hcy levels (unit increase)					
Grip strength (right)	7	MR-Egger	−0.013	0.032	6.90E−01
	7	Weighted_median	−0.036	0.013	6.02E−03
	7	IVW	−0.036	0.010	5.53E−04
	7	Maximum likelihood	−0.036	0.011	6.40E−04
	7	Penalised weighted median	−0.036	0.013	7.04E−03
Grip strength (left)	7	MR-Egger	−0.019	0.032	5.79E−01
	7	Weighted_median	−0.045	0.013	6.80E−04
	7	IVW	−0.045	0.010	1.45E−05
	7	Maximum likelihood	−0.046	0.011	2.05E−05
	7	Penalised weighted median	−0.045	0.014	1.24E−03
Walking pace	7	MR-Egger	−0.072	0.032	7.47E−02
	7	Weighted_median	−0.036	0.013	5.66E−03
	7	IVW	−0.038	0.011	3.18E−04
	7	Maximum likelihood	−0.038	0.011	3.60E−04
	7	Penalised weighted median	−0.036	0.013	6.36E−03
ALM	7	MR-Egger	−0.056	0.041	2.26E−01
	7	Weighted_median	−0.057	0.017	1.01E−03
	7	IVW	−0.058	0.013	1.03E−05
	7	Maximum likelihood	−0.059	0.014	1.55E−05
	7	Penalised weighted median	−0.057	0.017	8.39E−04
	Hcy levels (unit decrease)				
Grip strength (right)*	4	MR-Egger	0.008	0.071	0.93
	4	Weighted_median	0.013	0.016	0.42
	4	IVW	0.005	0.021	0.82
	4	Maximum likelihood	0.005	0.013	0.71
	4	Penalised weighted median	0.014	0.016	0.38
Grip strength (left)*	3	MR-Egger	0.064	0.052	0.43
	3	Weighted_median	0.004	0.016	0.81
	3	IVW	−0.006	0.014	0.64
	3	Maximum likelihood	−0.007	0.014	0.64
	3	Penalised weighted median	0.004	0.016	0.81
Walking pace	5	MR-Egger	0.031	0.066	0.67
	5	Weighted_median	0.014	0.016	0.39
	5	IVW	0.010	0.020	0.61
	5	Maximum likelihood	0.011	0.012	0.38
	5	Penalised weighted median	0.019	0.016	0.25
ALM*	4	MR-Egger	1.80E−4	0.080	0.99
	4	Weighted_median	−0.032	0.021	0.14
	4	IVW	−0.034	0.018	0.06
	4	Maximum likelihood	−0.034	0.018	0.06
	4	Penalised weighted median	−0.032	0.021	0.13

*: Second round of MR, analysis after identification and removal of pleiotropic SNPs, by MR-PRESSO; Of these, rs2251468 was excluded in the MR, analysis of grip strength (right), rs2251468 and rs42648 were excluded in the MR, analysis of grip strength (left), and rs7130284 was excluded in the MR, analysis of ALM.

(effect = −0.038, SE = 0.011, p = 3.18E-4), and ALM (effect = −0.058, 0.013, p = 1.03E-5), as shown in Table 3. Similar results were observed regarding the MR-Egger, weighted median, maximum likelihood, and penalized weighted median methods (Table 3). The direction of the results of all sensitive analysis methods is consistent with the direction of the IVW-MR results, suggesting that the results are stable. No significant heterogeneity was detected among

TABLE 4 Heterogeneity and pleiotropy analysis between homocysteine levels and the three components of sarcopenia.

	Heterogeneity test			MR-egger intercept test					MR-PRESSO global test			
	Q	Q_df	P value	Q	Q_df	P value	Intercept	SE	P value	Pleiotropy variant #	P value	Pleiotropy variant *	P value
Hcy levels (unit increase)													
Grip strength (right)	3.60	6	0.73	3.05	5	0.69	−0.001	0.002	0.49	NA	0.81		
Grip strength (left)	5.76	6	0.45	5.03	5	0.41	−0.002	0.002	0.43	NA	0.57		
Walking pace	6.30	6	0.39	5.00	5	0.42	0.002	0.002	0.31	NA	0.44		
ALM	4.13	6	0.66	4.13	5	0.53	< 0.001	0.002	0.95	NA	0.74		
Hcy levels (unit decrease)													
Grip strength (right)	7.93	3	0.05	7.92	2	0.02	< 0.001	0.005	0.97	rs2251468	0.02	NA	0.13
Grip strength (left)	3.55	2	0.17	1.18	1	0.28	−0.006	0.004	0.39	rs2251468/ rs42648	< 0.01	NA	NA
Walking pace	11.20	4	0.02	10.80	3	0.01	−0.001	0.004	0.76	NA	0.07		
ALM	0.55	3	0.91	0.36	2	0.84	−0.002	0.005	0.71	rs7130284	0.06	NA	0.89

#: The first round MR, analysis; *: The second round of MR, analysis after identification and removal of pleiotropic SNPs, by MR-PRESSO. Q: Cochran Q statistics; SE, standard error.

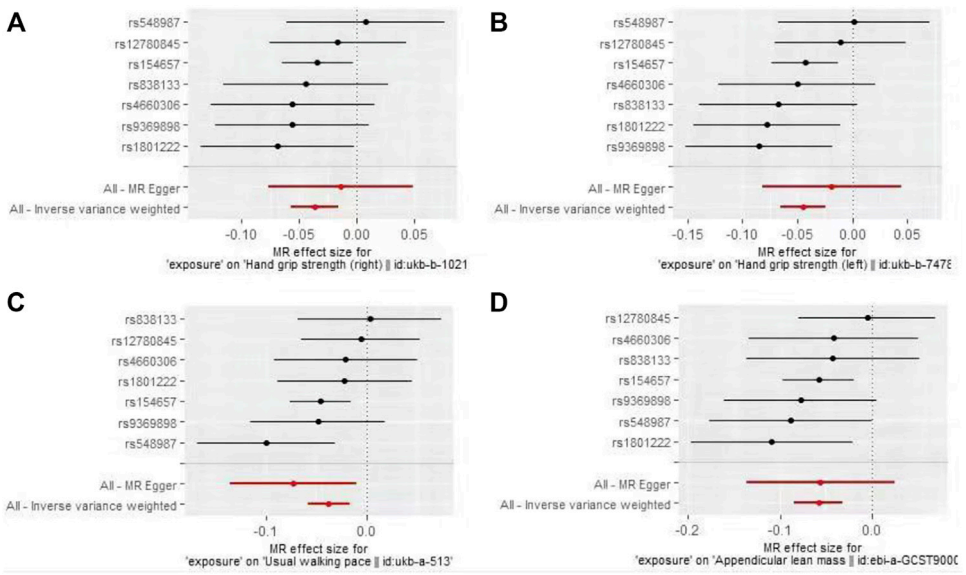


FIGURE 1
Forest plots for MR analyses of the causal effect of increased plasma homocysteine levels on Grip strength (right, **(A)**, Grip strength (left, **(B)**, Walking pace **(C)**, and ALM **(D)**.

the 7 IVs according to Cochran's Q-test (Table 4); thus, all IVW-MR analyses used fixed-effects models. The MR–Egger intercept test and MR-PRESSO global test indicated that the MR results were unlikely to be confounded by horizontal pleiotropy and reverse causality (Table 4). Therefore, all MR analyses included only the first round. The effect of individual SNPs on causal estimation was demonstrated by forest plots (Figure 1). The leave-one-out analysis showed that the MR analysis results were stable and unaffected by any single SNP (Figure 2).

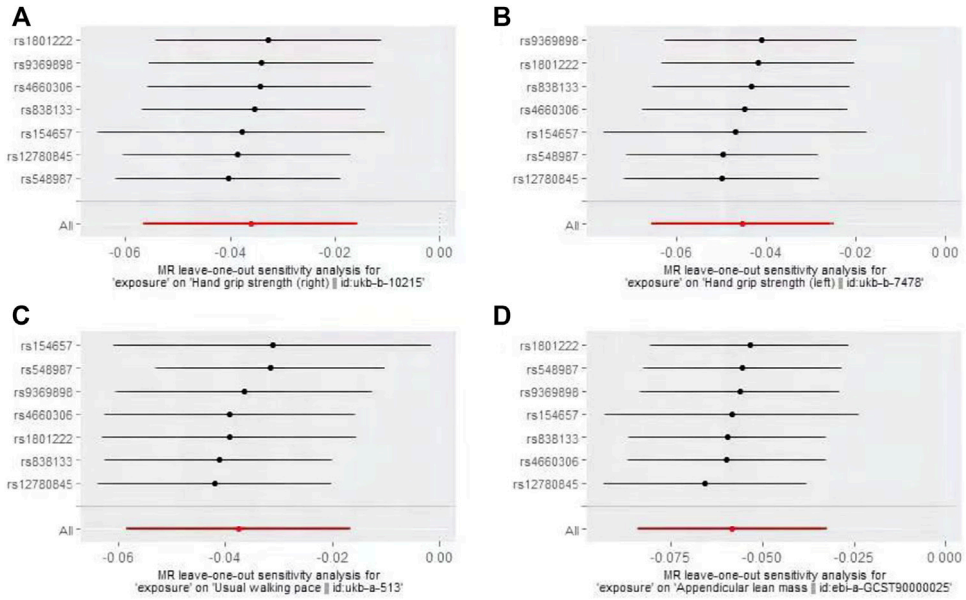


FIGURE 2
Plots of "leave-one-out"analyses MR analyses of the causal effect of increased plasma homocysteine levels on Grip strength (right, **(A)**, Grip strength (left, **(B)**, Walking pace **(C)**, and ALM **(D)**.

3.3 Associations of decreased plasma hcy levels with grip strength, walking pace, and appendicular lean mass

In the first round of analysis, the MR-PRESSO outlier test identified four pleiotropic SNPs associated with grip strength (right: rs2251468; left: rs2251468 and rs42648) and ALM (rs7130284). Therefore, a second round of analysis was performed after excluding the pleiotropic SNPs. The IVW-MR analysis found no evidence of associations between decreased plasma Hcy levels and grip strength (right: effect = 0.005, SE = 0.021, $p = 0.82$; left: effect = −0.006, SE = 0.014, $p = 0.64$), walking pace (effect = 0.01, 0.020, $p = 0.61$), or ALM (effect = −0.034, SE = 0.018, $p = 0.06$), (Table 3; Figure 3). The results of all MR methods used for sensitivity analysis are shown in Table 3. In the MR analysis of grip strength (left), the results of MR-Egger, Weighted median and Penalised weighted median were opposite to those of IVW-MR; In the MR analysis of ALM, the result of MR-Egger was opposite to that of IVW-MR, but the results of all MR methods in grip strength (left) and ALM were not significant, suggesting that the results were also stable. In the MR analyses of grip strength (right) and walking pace, the results of the MR methods used for sensitivity analysis were in the same direction as the IVW-MR results. Cochran's Q-test detected slight heterogeneity in IVs associated with grip strength (right) and walking pace (Table 4); therefore, the IVW-MR analysis utilized random-effects models. No heterogeneity was observed in IVs associated with grip strength (left) or ALM (Table 4); thus, the IVW-MR analysis utilized fixed-effects models. The MR–Egger intercept test and MR-PRESSO global test indicated that the MR results were unlikely to be confounded by horizontal pleiotropy and reverse causality (Table 4). The effect of individual SNPs on causal estimation was demonstrated by forest plots (Figure 3). Additionally, the leave-one-out analysis showed that the MR analysis results were stable and unaffected by any single SNP (Figure 4).

4 Discussion

To our knowledge, this is the first study to use large-scale GWAS data and MR analysis to explore the causal relationship between Hcy levels and the three components (grip strength, walking pace and ALM) of sarcopenia. This MR study showed that increased plasma Hcy levels were significantly and negatively associated with grip strength, walking pace, and ALM, and similar results were obtained in all sensitivity analyses. However, decreased plasma Hcy levels were not associated with grip strength, walking pace, or ALM. Although inconsistent directions emerged in the sensitivity analysis of grip strength (left) and ALM, but since the results of all MR methods are not significant, we consider that this no-causality association is still the stable. Based on this, our findings provide new insights into the effects of Hcy on muscle function and mass.

Consensus regarding the relationship between plasma Hcy levels and muscle strength and physical performance in

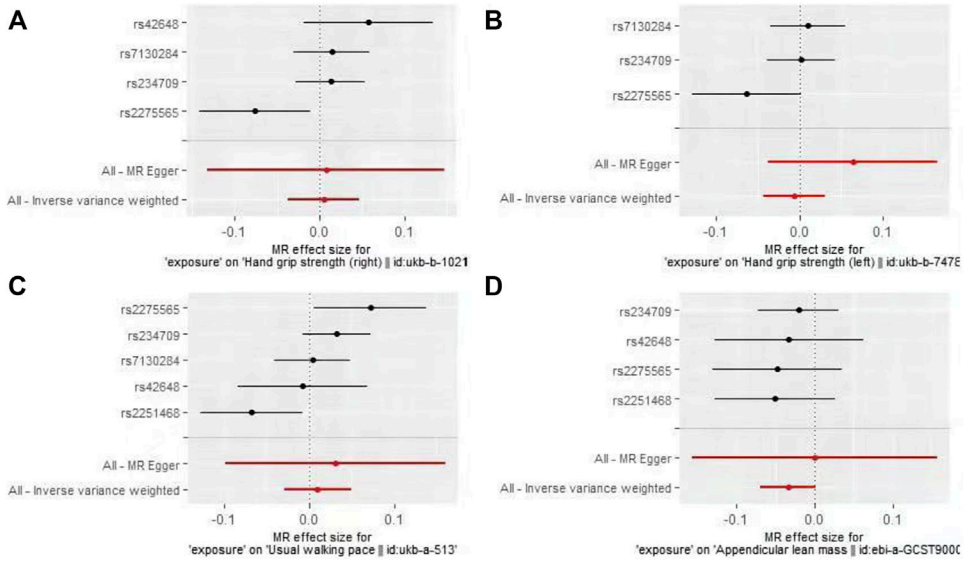


FIGURE 3
Forest plots for MR analyses of the causal effect of decreased plasma homocysteine levels on Grip strength (right, **(A)**, Grip strength (left, **(B)**, Walking pace **(C)**, and ALM **(D)**.

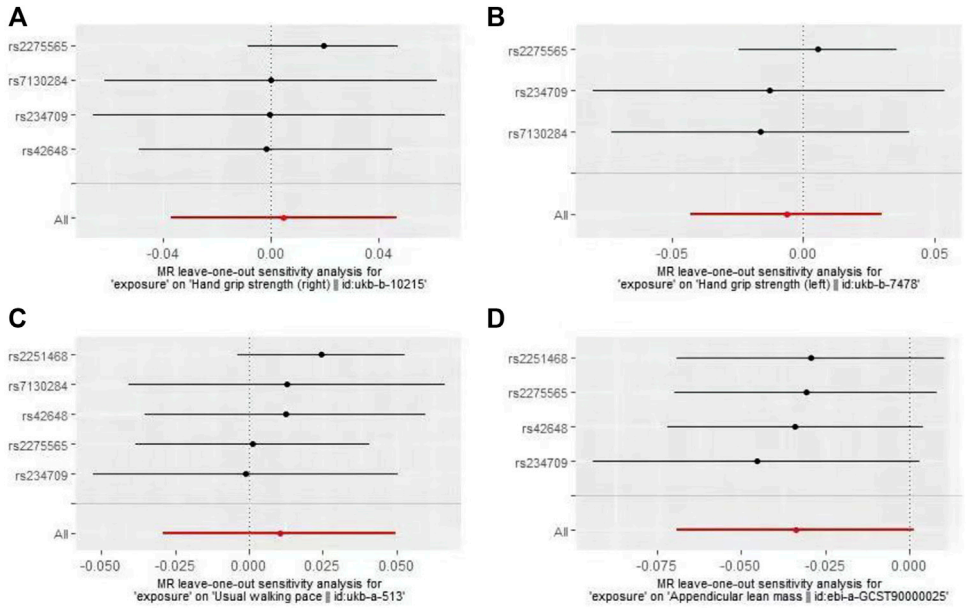


FIGURE 4
Plots of "leave-one-out"analyses MR analyses of the causal effect of decreased plasma homocysteine levels on Grip strength (right, **(A)**, Grip strength (left, **(B)**, Walking pace **(C)**, and ALM **(D)**.

observational studies is lacking. Granic et al. (2017) did not find an association between Hcy levels and grip strength in a longitudinal study that included 845 participants aged 85 years or older, either in the cross-sectional study (wave 1) or in the subsequent 5-year follow-up (waves 2–4); however, the use of this elderly population (>85 years old) may reduce its generalizability to older individuals of less advanced age. In addition, Hcy levels in Granic et al.'s study were measured by detecting Hcy

biomarkers and further determined by principal component analysis (PCA); thus, the results may have been influenced by the selection and number of biomarkers included in the PCA(20). In contrast, a cross-sectional study by Lee et al. (2020) that included 1,582 community members >50 years of age reported significant negative associations of increased Hcy levels with grip strength and walking pace. These strong associations persisted after adjusting for confounding factors such as sex, age, smoking status and underlying disease. Conflicting results between observational studies may be explained by interference from confounding factors. In this MR study, to eliminate the interference of confounding factors (pleiotropic SNPs) on the results, we took a series of quality control steps for instrumental variable selection (PhenoScanner database checking and MR-PRESSO outlier test), then we found a significant negative correlation between high levels of Hcy with grip strength and walking pace, which may resolve the current controversy to some extent. Furthermore, several animal experiments have reported that increased Hcy levels influence skeletal muscle function, consistent with our findings. Using a mouse model of hyperhomocysteine (HHcy) constructed using cysteine beta synthase deficient (CBS-) mice, Veeranki et al. (2015) concluded that HHcy leads to reduced proliferation of skeletal muscle satellite cells by inducing excessive oxidative stress and p38 MAPK signaling. Another experiment utilizing a CBS- mouse model found that skeletal muscle weakness associated with HHcy involves mitochondrial dysfunction and epigenetic modifications (Majumder et al., 2018). Thus, some of the mechanisms by which increased plasma Hcy levels lead to skeletal muscle dysfunction have been identified by *in vivo* experiments. Physical performance is a multidimensional concept involving not only muscles but also central and peripheral nerve function. The pathogenesis of HHcy in various vascular conditions, such as induction of oxidative stress (Behera et al., 2021), promotion of apoptosis (Pai et al., 2021), and stimulation of inflammatory factor release (Liang et al., 2021), has been extensively investigated; similarly, studies are exploring the pathogenesis of HHcy in various neurological diseases (Lipton et al., 1997; Moretti and Caruso, 2019). Although it is unclear whether HHcy reduces physical performance through muscle damage or neurological damage, there is no doubt that Hcy impacts muscles, nerves, and even blood vessels and thereby leads accelerates declines in physical performance.

However, it remains controversial whether Hcy reduces muscle mass. Several previous studies (Elshorbagy et al., 2008; Atkins et al., 2014) did not find a significant association between HHcy and muscle mass. However, these studies were either biased in terms of sample selection (e.g., hospitalized participants rather than community-dwelling older adults) or did not control for confounding factors associated with sarcopenia or Hcy levels (e.g., lifestyle, underlying medical conditions, and mental health) to a great extent. Choi et al. (2022) conducted a cross-sectional study of 114,583 community-dwelling adults and found significant associations of Hcy levels with mild low muscle mass (LMM) and severe LMM. These associations persisted after adjusting for factors such as age, sex, lifestyle, physical activity, and glycated hemoglobin. In another study of 423 patients with peripheral arterial disease, a significant inverse association was found between plasma Hcy levels and calf muscle area (McDermott et al., 2007). Similarly, in the present MR study, we found a significant negative correlation between increased Hcy levels and ALM. This etiologically verified negative correlation suggests that the abovementioned pathophysiological mechanism by which Hcy damages muscle not only impairs muscle function but also reduces muscle mass. In the future, well-designed *in vivo* or animal experiments are needed to elucidate the exact mechanisms underlying Hcy-induced muscle damage.

The use of MR analysis is a major strength of this study, as it allowed the use of large-scale genome-wide summary data, thus improving the reliability of our findings. To minimize potential bias due to population stratification, the GWAS data used in this study were all sourced from European populations. We also performed a linkage disequilibrium test to exclude the effects of nonindependent SNPs on the results. To prevent the influences of pleiotropy and reverse causation, we first identified and excluded pleiotropic SNPs that were significantly associated with outcomes or confounders with the PhenoScanner database before the first round of analysis. As the exact biological functions of many genetic variants remain unknown, we also identified and removed potentially pleiotropic SNPs with the MR-PRESSO outlier test before the second round of analysis. Our MR study showed a causal correlation between Hcy levels and the major components of sarcopenia, so Hcy levels can be used as an early predictor of sarcopenia. Also, because of the causal association, reducing Hcy levels could also be used as a treatment for sarcopenia.

This study has some limitations. First, this study did not determine the relationship between Hcy levels and sarcopenia based on the cutoff values required for the diagnosis of sarcopenia. However, since grip strength, walking pace and ALM are valid predictors of sarcopenia (Cruz-Jentoft et al., 2019), we believe that this study supports a negative association between increased plasma Hcy levels and sarcopenia. Second, in our study, ALM was measured using BIA, an indirect measure that may be less accurate than other methods that directly quantify muscle mass, such as magnetic resonance imaging (MRI), computed tomography (CT) or dual-energy x-ray absorptiometry (DXA). However, the EWGSOP2 consensus repor (Cruz-Jentoft et al., 2019) states that, considering affordability and portability, BIA-based muscle mass measurements may be preferable to other measures in large populations. Third, physical performance can be measured by

different tests such as walking pace, short distance physical performance test (SPPB) and timed up and go test (TUG). We only use walking pace as an assessment of physical performance, because it is considered a rapid, safe and highly reliable test for sarcopenia and has been shown to predict adverse outcomes associated with sarcopenia—disability, falls and death (Cruz-Jentoft et al., 2019). In addition, sex differences in skeletal muscle kinetics and fiber composition necessitate sex-specific analyses of muscle function and mass (De Giuseppe et al., 2022). Because we were unable to obtain sex-stratified GWAS data on plasma Hcy levels, we did not investigate the effect of Hcy levels on sarcopenia in a sex-stratified manner. Finally, these results may not be generalizable to populations of other ethnicities, as this study was primarily limited to individuals of European ancestry.

5 Conclusion

In conclusion, our MR analysis supported a causal association between increased plasma Hcy levels and lower grip strength, slower walking pace, and decreased ALM, while decreased plasma Hcy levels were not associated with all three components of sarcopenia. Prevention of increased plasma Hcy levels may avoid the occurrence of sarcopenia.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: See Table 1 for data sources and direct links.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

HY and TS conducted research design and data collection. HY and GL conducted data analysis. TS and QT interpreted the results. The first draft of the manuscript was written by HY. All authors read and approved the final manuscript.

Funding

This work was funded by Tianjin Key Medical Discipline (Specialty) Construction Project (TJXZDXK-064B) and Tian-Jin union medical centre Science and Technology Foundation (2020YJ007).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1051047/full#supplementary-material

References

Álvarez-Bustos, A., Carnicero-Carreño, J. A., Davies, B., Garcia-Garcia, F. J., Rodríguez-Artalejo, F., Rodríguez-Mañas, L., et al. (2022). Role of sarcopenia in the frailty transitions in older adults: a population-based cohort study. *J. Cachexia Sarcopenia Muscle* 13, 2352–2360. doi:10.1002/jcsm.13055

Atkins, J. L., Whincup, P. H., Morris, R. W., and Wannamethee, S. G. (2014). Low muscle mass in older men: the role of lifestyle, diet and cardiovascular risk factors. *J. Nutr. Health Aging* 18 (1), 26–33. doi:10.1007/s12603-013-0336-9

Behera, J., Kelly, K. E., and Tyagi, N. (2021). Hydrogen sulfide prevents ethanol-induced ZO-1 CpG promoter hypermethylation-dependent vascular permeability via miR-218/DNMT3a axis. *J. Cell. Physiol.* 236 (10), 6852–6867. doi:10.1002/jcp.30382

Bhasin, S., Travison, T. G., Manini, T. M., Patel, S., Pencina, K. M., Fielding, R. A., et al. (2020). Sarcopenia definition: The position statements of the sarcopenia definition and outcomes consortium. *J. Am. Geriatr. Soc.* 68 (7), 1410–1418. doi:10.1111/jgs.16372

Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through egger regression. *Int. J. Epidemiol.* 44 (2), 512–525. doi:10.1093/ije/dyv080

Bowden, J., Davey Smith, G., Haycock, P. C., and Burgess, S. (2016). Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.* 40 (4), 304–314. doi:10.1002/gepi.21965

Burgess, S., and Thompson, S. G. (2017). Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur. J. Epidemiol.* 32 (5), 377–389. doi:10.1007/s10654-017-0255-x

Burgess, S., Dudbridge, F., and Thompson, S. G. (2016). Combining information on multiple instrumental variables in mendelian randomization: Comparison of allele score and summarized data methods. *Stat. Med.* 35 (11), 1880–1906. doi:10.1002/sim.6835

Chen, L. K., Woo, J., Assantachai, P., Auyeung, T. W., Chou, M. Y., Iijima, K., et al. (2020). Asian working group for sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. *J. Am. Med. Dir. Assoc.* 21 (3), 300–307. doi:10.1016/j.jamda.2019.12.012

Chen, F. P., Lin, Y. J., Chao, A. S., Lin, Y. C., Sung, C. M., Chen, J. F., et al. (2021). Utilizing nomograms to predict prevalent vertebral fracture risk: an analysis of dysmobility syndrome in a community-dwelling population. *Biomed. J*, S2319-4170(21)00168-2. doi:10.1016/j.bj.2021.11.008

Choi, J. H., Seo, J. W., Lee, M. Y., Lee, Y. T., Yoon, K. J., and Park, C. H. (2022). Association between elevated plasma homocysteine and low skeletal muscle mass in asymptomatic adults. *Endocrinol. Metab.* 37 (2), 333–343. doi:10.3803/EnM.2021.1202

Cox, N. (2018). UK Biobank shares the promise of big data. *Nature* 562 (7726), 194–195. doi:10.1038/d41586-018-06948-3

Cruz-Jentoft, A. J., Baeyens, J. P., Bauer, J. M., Boirie, Y., Cederholm, T., Landi, F., et al. (2010). Sarcopenia: European consensus on definition and diagnosis: Report of the European working group on sarcopenia in older people. *Age Ageing* 39 (4), 412–423. doi:10.1093/ageing/afq034

Cruz-Jentoft, A. J., Bahat, G., Bauer, J., Boirie, Y., Bruyère, O., Cederholm, T., et al. (2019). Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing* 48 (4), 601. doi:10.1093/ageing/afz046

Davies, N. M., Holmes, M. V., and Davey Smith, G. (2018). Reading mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ Clin. Res. ed.* 362, k601. doi:10.1136/bmj.k601

De Giuseppe, R., Tomasinelli, C. E., Vincenti, A., Di Napoli, I., Negro, M., and Cena, H. (2022). Sarcopenia and homocysteine: is there a possible association in the elderly? A narrative review. *Nutr. Res. Rev.* 35 (1), 98–111. doi:10.1017/S095442242100010X

Dolan, R. D., Almasaudi, A. S., Dieu, L. B., Horgan, P. G., McSorley, S. T., and McMillan, D. C. (2019). The relationship between computed tomography-derived body composition, systemic inflammatory response, and survival in patients undergoing surgery for colorectal cancer. *J. Cachexia Sarcopenia Muscle* 10 (1), 111–122. doi:10.1002/jcsm.12357

Eguchi, Y., Toyoguchi, T., Inage, K., Fujimoto, K., Orita, S., Suzuki, M., et al. (2021). Advanced glycation end products are associated with sarcopenia in older women: aging marker dynamics. *J. Women Aging* 33 (3), 328–340. doi:10.1080/08952841.2019.1697161

Elshorbagy, A. K., Nurk, E., Gjesdal, C. G., Tell, G. S., Ueland, P. M., Nygård, O., et al. (2008). Homocysteine, cysteine, and body composition in the hordaland homocysteine study: does cysteine link amino acid and lipid metabolism? *Am. J. Clin. Nutr.* 88 (3), 738–746. doi:10.1093/ajcn/88.3.738

Fernandes, L. V., Paiva, A. E. G., Silva, A. C. B., de Castro, I. C., Santiago, A. F., de Oliveira, E. P., et al. (2022). Prevalence of sarcopenia according to EWGSOP1 and EWGSOP2 in older adults and their associations with unfavorable health outcomes: a systematic review. *Aging Clin. Exp. Res.* 34 (3), 505–514. doi:10.1007/s40520-021-01951-7

Fielding, R. A., Vellas, B., Evans, W. J., Bhasin, S., Morley, J. E., Newman, A. B., et al. (2011). Sarcopenia: an undiagnosed condition in older adults. Current consensus definition: prevalence, etiology, and consequences. International working group on sarcopenia. *J. Am. Med. Dir. Assoc.* 12 (4), 249–256. doi:10.1016/j.jamda.2011.01.003

Gao, Q., Hu, K., Yan, C., Zhao, B., Mei, F., Chen, F., et al. (2021). Associated factors of sarcopenia in community-dwelling older adults: A systematic review and meta-analysis. *Nutrients* 13 (12), 4291. doi:10.3390/nu13124291

Granic, A., Davies, K., Martin-Ruiz, C., Jagger, C., Kirkwood, T. B. L., von Zglinicki, T., et al. (2017). Grip strength and inflammatory biomarker profiles in very old adults. *Age Ageing* 46 (6), 976–982. doi:10.1093/ageing/afx088

Han, L., Liu, Y., Wang, C., Tang, L., Feng, X., Astell-Burt, T., et al. (2017). Determinants of hyperhomocysteinemia in healthy and hypertensive subjects: A population-based study and systematic review. *Clin. Nutr.* 36 (5), 1215–1230. doi:10.1016/j.clnu.2016.11.011

Lee, W. J., Peng, L. N., Loh, C. H., and Chen, L. K. (2020). Sex-different associations between serum homocysteine, high-sensitivity C-reactive protein and sarcopenia: Results from I-Lan Longitudinal Aging Study. *Exp. Gerontol.* 132, 110832. doi:10.1016/j.exger.2020.110832

Liang, C., Wang, Q. S., Yang, X., Zhu, D., Sun, Y., Niu, N., et al. (2021). Homocysteine causes endothelial dysfunction via inflammatory factor-mediated activation of epithelial sodium channel (ENaC). *Front. Cell Dev. Biol.* 9, 672335. doi:10.3389/fcell.2021.672335

Lipton, S. A., Kim, W. K., Choi, Y. B., Kumar, S., D'Emilia, D. M., Rayudu, P. V., et al. (1997). Neurotoxicity associated with dual actions of homocysteine at the N-methyl-D-aspartate receptor. *Proc. Natl. Acad. Sci. U. S. A.* 94 (11), 5923–5928. doi:10.1073/pnas.94.11.5923

Majumder, A., Singh, M., Behera, J., Theilen, N. T., George, A. K., Tyagi, N., et al. (2018). Hydrogen sulfide alleviates hyperhomocysteinemia-mediated skeletal muscle atrophy via mitigation of oxidative and endoplasmic reticulum stress injury. *Am. J. Physiol. Cell Physiol.* 315 (5), C609–c22. doi:10.1152/ajpcell.00147.2018

McDermott, M. M., Ferrucci, L., Guralnik, J. M., Tian, L., Green, D., Liu, K., et al. (2007). Elevated levels of inflammation, d-dimer, and homocysteine are associated with adverse calf muscle characteristics and reduced calf strength in peripheral arterial disease. *J. Am. Coll. Cardiol.* 50 (9), 897–905. doi:10.1016/j.jacc.2007.05.017

Milligan, B. G. (2003). Maximum-likelihood estimation of relatedness. *Genetics* 163 (3), 1153–1167. doi:10.1093/genetics/163.3.1153

Moretti, R., and Caruso, P. (2019). The controversial role of homocysteine in neurology: From labs to clinical practice. *Int. J. Mol. Sci.* 20 (1), E231. doi:10.3390/ijms20010231

Pagano, A. F., Brioche, T., Arc-Chagnaud, C., Demangel, R., Chopard, A., and Py, G. (2018). Short-term disuse promotes fatty acid infiltration into skeletal muscle. *J. Cachexia Sarcopenia Muscle* 9 (2), 335–347. doi:10.1002/jcsm.12259

Pagoni, P., Dimou, N. L., Murphy, N., and Stergiakouli, E. (2019). Using Mendelian randomisation to assess causality in observational studies. *Evid. Based. Ment. Health* 22 (2), 67–71. doi:10.1136/ebmental-2019-300085

Pai, P. Y., Chou, W. C., Chan, S. H., Wu, S. Y., Chen, H. I., Li, C. W., et al. (2021). Epigallocatechin gallate reduces homocysteine-caused oxidative damages through modulation SIRT1/AMPK pathway in endothelial cells. *Am. J. Chin. Med.* 49 (1), 113–129. doi:10.1142/S0192415X21500063

Pei, Y. F., Liu, Y. Z., Yang, X. L., Zhang, H., Feng, G. J., Wei, X. T., et al. (2020). The genetic architecture of appendicular lean mass characterized by association analysis in the UK Biobank study. *Commun. Biol.* 3 (1), 608. doi:10.1038/s42003-020-01334-0

Pierce, B. L., and Burgess, S. (2013). Efficient design for mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am. J. Epidemiol.* 178 (7), 1177–1184. doi:10.1093/aje/kwt084

Samson, M. E., Yeung, L. F., Rose, C. E., Qi, Y. P., Taylor, C. A., and Crider, K. S. (2022). Vitamin B-12 malabsorption and renal function are critical considerations in studies of folate and vitamin B-12 interactions in cognitive performance: Nhanes 2011-2014. *Am. J. Clin. Nutr.* 116 (1), 74–85. doi:10.1093/ajcn/nqac065

Selhub, J. (1999). Homocysteine metabolism. *Annu. Rev. Nutr.* 19, 217–246. doi:10.1146/annurev.nutr.19.1.212

Shen, Y., Chen, J., Chen, X., Hou, L., Lin, X., and Yang, M. (2019). Prevalence and associated factors of sarcopenia in nursing home residents: A systematic review and meta-analysis. *J. Am. Med. Dir. Assoc.* 20 (1), 5–13. doi:10.1016/j.jamda.2018.09.012

Spence, J. D., and Hankey, G. J. (2022). Problem in the recent American Heart association guideline on secondary stroke prevention: B vitamins to lower homocysteine do prevent stroke. *Stroke* 53 (8), 2702–2708. doi:10.1161/STROKEAHA.122.038640

Swart, K. M., Ham, A. C., van Wijngaarden, J. P., Enneman, A. W., van Dijk, S. C., Sohl, E., et al. (2016). A randomized controlled trial to examine the effect of 2-year vitamin B12 and folic acid supplementation on physical performance, strength, and falling: Additional findings from the B-PROOF study. *Calcif. Tissue Int.* 98 (1), 18–27. doi:10.1007/s00223-015-0059-5

Ter Borg, S., de Groot, L. C., Mijnarends, D. M., de Vries, J. H., Verlaan, S., Meijboom, S., et al. (2016). Differences in nutrient intake and biochemical nutrient status between sarcopenic and nonsarcopenic older adults-results from the maastricht sarcopenia study. *J. Am. Med. Dir. Assoc.* 17 (5), 393–401. doi:10.1016/j.jamda.2015.12.015

van Meurs, J. B., Pare, G., Schwartz, S. M., Hazra, A., Tanaka, T., Vermeulen, S. H., et al. (2013). Common genetic loci influencing plasma homocysteine concentrations and their effect on risk of coronary artery disease. *Am. J. Clin. Nutr.* 98 (3), 668–676. doi:10.3945/ajcn.112.044545

Veeranki, S., Lominadze, D., and Tyagi, S. C. (2015). Hyperhomocysteinemia inhibits satellite cell regenerative capacity through p38 alpha/beta MAPK signaling. *Am. J. Physiol. Heart Circ. Physiol.* 309 (2), H325–H334. doi:10.1152/ajpheart.00099.2015

Verbanck, M., Chen, C. Y., Neale, B., and Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* 50 (5), 693–698. doi:10.1038/s41588-018-0099-7

Wang, P., Liu, L., and Lei, S. F. (2021). Causal effects of homocysteine levels on the changes of bone mineral density and risk for bone fracture: A two-sample mendelian randomization study. *Clin. Nutr.* 40 (4), 1588–1595. doi:10.1016/j.clnu.2021.02.045

Yoo, T. K., Rhim, H. C., Lee, Y. T., Yoon, K. J., and Park, C. H. (2022). Relationship between hyperhomocysteinemia and coexisting obesity with low skeletal muscle mass in asymptomatic adult population. *Sci. Rep.* 12 (1), 12439. doi:10.1038/s41598-022-16401-1

Yu, H., Luo, G., Yu, B., Sun, T., Tang, Q., and Jia, Y. (2022). Robot-assisted kyphoplasty improves clinical and radiological features better than fluoroscopy-assisted kyphoplasty in the treatment of vertebral compression fractures: A meta-analysis. *Front. Surg.* 9, 955966. doi:10.3389/fsurg.2022.955966

Yuan, S., Mason, A. M., Carter, P., Burgess, S., and Larsson, S. C. (2021). Homocysteine, B vitamins, and cardiovascular disease: a mendelian randomization study. *BMC Med.* 19 (1), 97. doi:10.1186/s12916-021-01977-8

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 05 December 2022
DOI 10.3389/fgene.2022.1044264

OPEN ACCESS

EDITED BY
Qiling Yuan,
Xi'an Jiaotong University, China

REVIEWED BY
Yiyi Zhou,
Second Military Medical University, China
Peng Zhu,
Huazhong University of Science and Technology, China

*CORRESPONDENCE
Yongqing Xu,
xuyongqingkm@163.net

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

RECEIVED 14 September 2022
ACCEPTED 18 November 2022
PUBLISHED 05 December 2022

CITATION
Shi X, Ni H, Wu Y, Guo M, Wang B, Zhang Y, Zhang B and Xu Y (2022), Diagnostic signature, subtype classification, and immune infiltration of key m6A regulators in osteomyelitis patients.
Front. Genet. 13:1044264.
doi: 10.3389/fgene.2022.1044264

Diagnostic signature, subtype classification, and immune infiltration of key m6A regulators in osteomyelitis patients

Xiangwen Shi[1†], Haonan Ni[1†], Yipeng Wu[1,2,3†], Minzheng Guo[1], Bin Wang[1], Yue Zhang[1], Bihuan Zhang[1] and Yongqing Xu[2,3*]

[1]School of Medicine, Kunming Medical University, Kunming, China, [2]Department of Orthopedic Surgery, 920th Hospital of Joint Logistics Support Force, Kunming, China, [3]Laboratory of Clinical Medical Center, Yunnan Traumatology and Orthopedics, Kunming, China

Background: As a recurrent inflammatory bone disease, the treatment of osteomyelitis is always a tricky problem in orthopaedics. N6-methyladenosine (m6A) regulators play significant roles in immune and inflammatory responses. Nevertheless, the function of m6A modification in osteomyelitis remains unclear.

Methods: Based on the key m6A regulators selected by the GSE16129 dataset, a nomogram model was established to predict the incidence of osteomyelitis by using the random forest (RF) method. Through unsupervised clustering, osteomyelitis patients were divided into two m6A subtypes, and the immune infiltration of these subtypes was further evaluated. Validating the accuracy of the diagnostic model for osteomyelitis and the consistency of clustering based on the GSE30119 dataset.

Results: 3 writers of Methyltransferase-like 3 (METTL3), RNA-binding motif protein 15B (RBM15B) and Casitas B-lineage proto-oncogene like 1 (CBLL1) and three readers of YT521-B homology domain-containing protein 1 (YTHDC1), YT521-B homology domain-containing family 3 (YTHDF2) and Leucine-rich PPR motif-containing protein (LRPPRC) were identified by difference analysis, and their Mean Decrease Gini (MDG) scores were all greater than 10. Based on these 6 significant m6A regulators, a nomogram model was developed to predict the incidence of osteomyelitis, and the fitting curve indicated a high degree of fit in both the test and validation groups. Two m6A subtypes (cluster A and cluster B) were identified by the unsupervised clustering method, and there were significant differences in m6A scores and the abundance of immune infiltration between the two m6A subtypes. Among them, two m6A regulators (METTL3 and LRPPRC) were closely related to immune infiltration in patients with osteomyelitis.

Conclusion: m6A regulators play key roles in the molecular subtypes and immune response of osteomyelitis, which may provide assistance for personalized immunotherapy in patients with osteomyelitis.

KEYWORDS

N6-methyladenosine (m6A), osteomyelitis, diagnosis, immune infiltration, subtype

Introduction

Osteomyelitis is a devastating bone disease usually caused by bacterial, fungal and mycobacterial infections. Acute and chronic osteomyelitis both cause gradual deterioration of bone, as well as a severe inflammatory response, even invading the surrounding tissues outside the bone marrow cavity (Ma et al., 2021a). The incidence of posttraumatic osteomyelitis varies from 2% to 50% (Zalavras, 2017). In addition, diabetic ulcers due to vascular insufficiency are a common risk factor for osteomyelitis. The incidence of osteomyelitis is on the rise due to the increasing prevalence of diabetic foot infection. More than one-third of patients with diabetic foot infections are reported to be accompanied by osteomyelitis of the foot (Berendt et al., 2008). It is difficult to cure osteomyelitis, and even surgical intervention has an 8% mortality rate (Noskin et al., 2005).

The key to successful treatment of osteomyelitis is early diagnosis. However, the wide variation in clinical presentations of osteomyelitis makes accurate diagnosis of osteomyelitis challenging. The gold standard for the diagnosis of osteomyelitis includes bone histopathological examination and bacterial culture, which should be combined with imaging examination, laboratory test and physical examination (Hogan et al., 2013). The indicators of preliminary evaluation included erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), and blood culture. Although CRP is elevated earlier than ESR and can be used to detect acute osteomyelitis, it is nonspecific for osteomyelitis (Unkila-Kallio et al., 1994; Peltola and Pääkkönen, 2014). Even when bacteria are cultured, 40% of osteomyelitis patients show negative results (Scott et al., 1990). In addition, early osteomyelitis is also difficult to diagnose with imaging, and the positive X-ray rate of patients with osteomyelitis for 2 weeks has been reported to be less than 20% (Pasquet et al., 2015). Therefore, no specific markers or methods have been identified for the diagnosis of osteomyelitis.

N6-methyladenosine (m6A) is the most prevalent chemical modification throughout the eukaryotic population, occurring in messenger RNAs (mRNAs), transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs) (Linder et al., 2015; Xiao et al., 2016; Wang et al., 2020). The m6A modification process is mainly regulated by a combination of methylation recognition proteins (readers), m6A methyltransferases (writers), and m6A demethylases (erasers), which perform the functions of recognition, installation, and removal of methylation, respectively (Shi et al., 2019; Huang et al., 2021). Due to its reversible modifications that can affect the transcription and translation of mRNA profiles, which in turn regulate various fundamental cellular processes such as differentiation, metabolism, inflammation and immunity (Zhao et al., 2020; Zhu et al., 2020; Liu et al., 2022), it has begun to receive widespread attention from researchers.

Recently, m6A modifications have been found to be involved in the development of various inflammatory diseases, including skeletal diseases, neurodegenerative diseases, cardiac diseases, metabolic diseases, and cancer (Chokkalla et al., 2019; Zong et al., 2019; Zhang et al., 2020a). Zhang et al. (2019) found that METTL3 knockdown inhibited osteoblast differentiation and mineralization *via* the Smad signaling pathway under LPS-stimulated inflammatory conditions. In addition, METTL3 depletion also mediates the inflammatory response of osteoblasts through the MAPK signaling pathway, as evidenced by the production of inflammatory cytokines such as IL6 and IL12. However, the relationship and regulatory mechanisms between osteomyelitis and m6A modifications remain unclear. In this study, we used a random forest (RF) approach to establish a diagnostic model of osteomyelitis, and further two subtypes (m6A clusters and gene clusters) were identified using unsupervised clustering in patients with osteomyelitis. This study is designed to assist in the diagnosis of osteomyelitis and personalized immunotherapy based on different molecular subtypes.

Material and methods

Data collection and processing

Two gene expression datasets (GSE16129 and GSE30119) containing osteomyelitis and healthy samples were selected for this study and both were downloaded from the GEO database website (https://www.ncbi.nlm.nih.gov/geo/). GSE16129 datasets came from three platforms (GPL96, GPL97 and GPL6106), while GSE30119 datasets came from the GPL6947 platform. The gene expression data were merged using the "Perl" language (http://www.perl.org/) and normalized using the "sva" and "limma" packages of R software (Ritchie et al., 2015). As the merged microarray data come from three different platforms, the ComBat method was applied to correct the data in batches to eliminate the batch effects caused by multiple factors (Larsen et al., 2014; Müller et al., 2016; Zhang et al., 2020b). 97 osteomyelitis samples and 29 healthy samples from the GSE16129 dataset served as the test group, and 39 osteomyelitis samples and 44 healthy samples from the GSE30119 dataset served as the validation group. Data with large gene expression were changed with log2.

Differential analysis of m6A regulators expression in osteomyelitis patients and healthy samples

26 m6A regulators were selected as references, including 15 readers, 9 writers and 2 erasers (Table 1) (Boulias et al., 2019;

TABLE 1 The characteristics of 26 m6A regulators.

Type	Gene	Number
Writers	METTL3/METTL14/METTL16/WTAP/VIRMA/ZC3H13/RBM15/RBM15B/CBLL1	9
Readers	YTHDC1/YTHDC2/YTHDF1/YTHDF2/YTHDF3/HNRNPC/FMR1/LRPPRC/HNRNPA2B1/IGFBP1/IGFBP2/IGFBP3/RBMX/ELAVL1/IGF2BP1	15
Erasers	FTO/ALKBH5	2

Zhang et al., 2020a). Based on these genes, the expression data of m6A regulators from 97 patients with osteomyelitis and 29 healthy samples in the text group were extracted by "limma" package, and the DEGs were screened by differential analysis between two groups. The results were visualized in the form of heatmap and boxplot with "pheatmap" and "ggpubr" packages (Zhao et al., 2014). The filter criteria were set to |logFC| value >1 and p-value < 0.05. In addition, the corresponding chromosome characteristics of m6A regulators, such as name, starting position and ending position, were obtained by consulting the literature (Dai et al., 2021). The locations of m6A regulators on chromosomes were extracted using the "Perl" program and were presented as a circle diagram using "RCirco" package (Zhang et al., 2013).

Diagnostic signature of osteomyelitis based on m6A regulators

Machine learning algorithms such as random forest (RF) and support vector machine (SVM) have been used for identifying and classifying key features of diseases, detecting hard-to-discern features in complex datasets, and analyzing gene expression data (Huang et al., 2018). RF is an integrated classifier consisting of many decision trees that exhibits robust performance by injecting randomness into the training of the trees to prevent overfitting of the data (Polan et al., 2016). SVM is also very effective in identifying complex datasets and is often used to identify cancer subtypes (Abinash and Vasudevan, 2022).

To better identify and screen important m6A regulators, two models were built by RF and SVM algorithms based on the text group. The boxplot of residual error, receiver operating characteristic (ROC) and reverse cumulative distribution plot of residual error were constructed to compare the accuracy of the two models. The smaller the residual error, the higher the accuracy of the models. After determining the model, the m6A regulators were scored according to their importance distribution, expressed as the Mean Decrease Gini (MDG). A higher value of a gene means that the gene is more important in this model. Genes with MDG values >10 were ultimately screened out. The "rms" package was then used to construct a nomogram model to predict the incidence of osteomyelitis. Calibration curves and decision curve analysis (DCA) were used in the analysis to assess the fit of the model, and a clinical impact curve was drawn to evaluate whether the model was beneficial to patients (Iasonos et al., 2008). Finally, the diagnostic model for osteomyelitis was validated based on the GSE30119 dataset.

Identification of molecular subtypes based on significant m6A regulators

In consensus clustering, the K-means algorithm is run multiple times to obtain the input partition, and the common matrix is calculated based on the partition result. The ultimate goal is to detect sample subtypes with similar characteristics (Brière et al., 2021). The "ConsensusClusterPlus" package of R software was used to cluster osteomyelitis samples with m6A regulators to create subtypes with differential characteristics (Seiler et al., 2010; Yu et al., 2012). The most appropriate number of clusters was screened based on the consistency heatmap, consistency scores, cumulative distribution function (CDF) and CDF delta area curve. Specifically, the value of K corresponding to when the CDF reaches its approximate maximum value is the best grouping result (Wilkerson and Hayes, 2010). The maximum value of the number of clusters K was set to 10. To test the accuracy of clusters, the results of consensus clustering were validated based on the GSE30119 dataset and principal component analysis (PCA) was performed on the classified samples to determine whether the groupings accurately reflected the characteristics of osteomyelitis patients (David and Jacobs, 2014). Furthermore, the expression of m6A regulators was further analyzed to determine whether it differed significantly between molecular subtypes.

Correlation analysis between m6A subtypes and immune cell infiltration

We further used single-sample gene set enrichment analysis (ssGSEA) in the "gsva" package to combine the 23 immune gene datasets with "high-low discriminant analysis" to calculate the immune infiltration score for each sample (Xiao et al., 2020). The boxplot was used to indicate whether the abundance of immune cells differed between m6A subtypes. The correlation between m6A regulators and immune cell infiltration was represented by a heatmap, and the top 2 m6A regulators with the strongest correlation were

screened. The samples were divided into low and high expression groups based on the median expression value of each gene, and boxplots were used to observe whether there was a significant difference in the abundance of immune cells between the high and low expression groups.

Correlation analysis between m6A subtypes and clinical characteristics

To test the relationship between the two m6A subtypes of osteomyelitis and clinical characteristics, information on the age and sex of osteomyelitis patients in the test and validation groups was extracted to further compare whether there were significant differences in age and sex between the two subtypes, represented in the form of boxplots and histograms. Furthermore, the results of clinical correlation were validated based on the GSE30119 dataset.

GO functional annotation and KEGG pathway analysis based on DEGs between the two m6A subgroups

To explore the molecular mechanisms of m6A regulators, GO and KEGG enrichment analyses were performed using the "ClusterProfile" and "enrichplot" packages (The Gene Ontology, 2006; Kanehisa et al., 2012). In the GO enrichment analysis, the top 10 molecular functions (MF), cellular components (CC) and biological processes (BP) related to DEGs were screened, while the KEGG analysis focused on the biological pathways involved in DEGs.

Identification of molecular subtypes based on the DEGs of two m6A subgroups

Based on the DEGs of the two m6A subtypes, consensus clustering analysis was performed using the "ClusterProfile" package. A consistency heatmap and CDF were used to determine the optimal number of clusters. Clustering conditions were set as follows: maxK = 8, reps = 50, pFeature = 1, pItem = 0.8, clusterAlg = "km", distance = "euclidean", and seed = 123,456.

Correlation analysis between gene subgroups and immune cell infiltration

To explore the relationship between the 23 immune cells and the gene cluster, a boxplot was used to visualize whether the abundance of immune cells was significantly different in each molecular subtype. The expression of m6A regulators was also visualized using boxplots to determine whether it was different between gene subtypes.

Differential analysis of m6A scores between the m6A subgroups or gene subgroups

Based on the expression of m6A regulators in each sample, the m6A score of each sample was obtained using PCA. Next, the "ggpubr" package was used to assess whether there were significant differences in m6A scores between the two m6A molecular subtypes or gene molecular subtypes. A Sankey diagram was drawn using the "ggalluvial" package to compare the similarity between m6A molecular subtypes and gene molecular subtypes. Finally, differences in m6A scores were validated based on the GSE30119 dataset.

Differential analysis of collagen and interleukin-related genes between the m6A subgroups or gene subgroups

To further understand the relationship between m6A modification and the occurrence and development of osteomyelitis, we screened collagen and interleukin-related genes, which are involved in osteogenic differentiation and osteomyelitis inflammation, respectively. Boxplots were used to observe whether there were significant differences in the expression of these genes between the two m6A subtypes or gene subtypes. The expression of differences in collagen and interleukin-related genes were validated based on the GSE30119 dataset.

Establishment of osteomyelitis model in rats

Twenty 8-week-old male SD rats weighing approximately 300 g were purchased from Kunming Animal Institute (Kunming, China), of which 10 were used as the experimental group and 10 as the control group. After anesthesia with intravenous sodium pentobarbital, the proximal third of the right tibia was exposed, and the medullary cavity was drilled approximately 1.5 mm and injected with 5 μl of 1×10^6 *S. aureus*. The control group was injected with sterile phosphate buffer solution. The specimen was sealed with bone wax after implantation of the 5 mm × 1 mm steel needle. The specific procedure of the experimental model was based on the method of Vergidis et al. (2015). All experimental procedures were reviewed and approved by the 920th Hospital of Joint Logistics Support Force Ethics Review Committee (2022–025–01).

Quantitative real-time PCR

Tibial tissues from the 10 rat models of osteomyelitis and 10 control groups were collected, and total RNA was extracted from tibial tissues using TRIzol reagent (Ambion Inc., Austin, TX), followed by reverse transcription of RNA to cDNA using SweScript RT I First Strand cDNA Synthesis Kit (Service Bio, Guangzhou, China) to reverse transcribe RNA to cDNA. Quantitative real-time PCR was performed using Universal Blue SYBR Green qPCR Master Mix (Service Bio, Guangzhou, China), and the level of GAPDH was used as an internal control. Relative expression was calculated according to the comparative Ct ($2^{-\Delta\Delta CT}$) method (Livak and Schmittgen, 2001). The sequences of the primers used are shown in Supplementary Table S1.

Immunohistochemistry

Right tibial tissues from the experimental group and the control group in the same location were taken and immunohistochemically stained using 4% paraformaldehyde fixation and paraffin embedding. Following paraffin removal, experimental tissues were incubated with primary antibodies against METTL3, YTHDC1, YTHDF2, RBM15B, LRPPRC and CBLL1 overnight at 4°C and control tissues were incubated with sterile phosphate-buffered saline (PBS; Service Bio, Guangzhou, China). Representative slices were incubated with biotinylated secondary antibodies (Service Bio, Guangzhou, China) at 22°C (Ramos-Vara, 2005). After the samples were sectioned, the staining differences of 6 significant m6A regulators in tibial tissues were evaluated under the microscope.

Statistical analysis

The Kruskal–Wallis test was used to compare the differences between groups, with $p < 0.05$ indicating statistical significance. All statistical analyses were performed using R software (version 4.1.3) and the corresponding program package (http://www.bioconductor.org/).

Results

Expression and specific characteristics of m6A regulators in osteomyelitis patients and healthy samples

The test group consisted of 97 osteomyelitis samples and 29 healthy samples, including 71 males and 55 females, with an average age of 7.4 ± 4.7 years, and the validation group consisted of 39 osteomyelitis samples and 44 healthy samples, including 41 males and 42 females, with an average age of 7.2 ± 4.3 years. Based on the 26 m6A regulators used as reference, 10 m6A regulators were identified in all samples and then screened for 6 m6A regulators after differential expression analysis (Figures 1A,B). The 6 m6A regulators include 3 writers of Methyltransferase-like 3 (METTL3), RNA-binding motif protein 15B (RBM15B) and Casitas B-lineage proto-oncogene like 1 (CBLL1) and three readers of YT521-B homology domain-containing protein 1 (YTHDC1), YT521-B homology domain-containing family 3 (YTHDF2) and Leucine-rich PPR motif-containing protein (LRPPRC). In addition, we observed the chromosome characteristics of m6A regulators in the form of a circle diagram and found that m6A regulators associated with osteomyelitis were mainly distributed on chromosomes 1, 2, 3, 4, 6, 7, 14 and 19 (Figure 1C). Specific characteristics of chromosomes, including starting position and ending position, are shown in Table 2.

Diagnostic signature for osteomyelitis based on m6A regulators

RF and SVM were used to construct diagnostic signatures. The boxplot and line plot revealed that the residual of RF was smaller than that of SVM (Figure 2A), and the AUC value of the model built by RF was larger than that of SVM (RF: 1.000 vs. SVM: 0.961) (Figure 2B). Therefore, RF model was selected as the best model to predict the incidence of osteomyelitis. The RF model was constructed based on the trees (Figure 2C), of which 6 significant m6A regulators (METTL3, YTHDC1, YTHDF2, RBM15B, LRPPRC and CBLL1) had MDG values greater than 10, and all of them were included in this RF model (Figure 2D). Furthermore, a nomogram model was established to predict the incidence of osteomyelitis by summing the scores for each gene (Figure 2E). In general, a higher total score represented a higher risk of osteomyelitis. The calibration curve and DCA curve suggested that the RF model fit well (Figure 2F). In the validation group, a nomogram model (Figure 2G) was established to predict the incidence of osteomyelitis and the calibration curve and DCA curve also suggested that the RF model fit well (Figure 2H).

Results of molecular subtypes based on m6A regulators

Based on CDF, consistency score, CDF delta area curve and number of samples, two molecular subtypes of m6A were identified (cluster A and cluster B) and consistency score (k = 2) > 0.8 (Figures 3A,B). There were 45 samples in cluster A osteomyelitis and 52 samples in cluster B osteomyelitis. In the validation group, cluster A osteomyelitis included 6 samples and cluster B osteomyelitis included 33 samples (Figures 3C,D). The results of PCA suggested that cluster A and cluster B

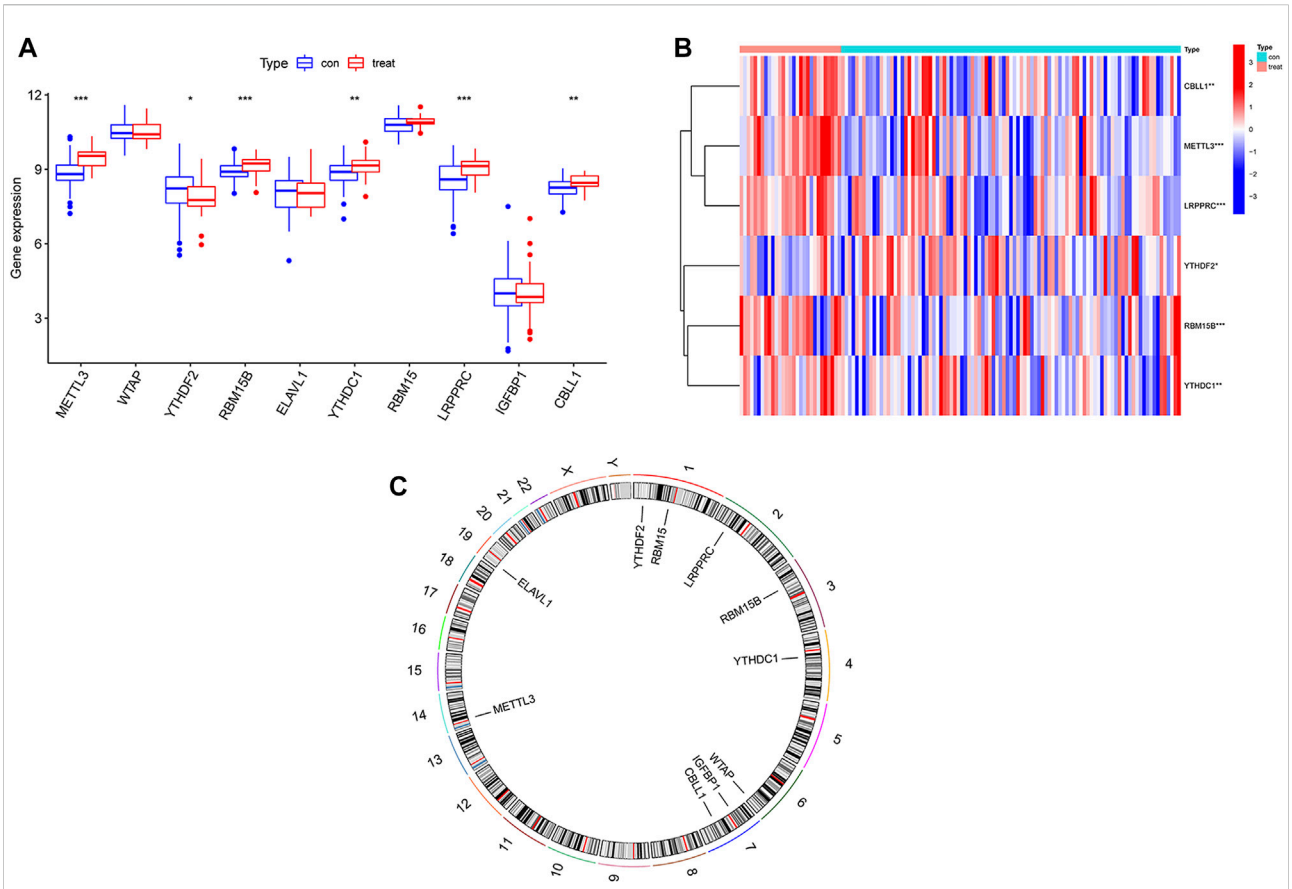


FIGURE 1
Expression of m6A regulators in osteomyelitis samples and healthy samples. **(A)** Expression of all 10 m6A regulators in osteomyelitis samples and normal samples. **(B)** Heatmap of 6 significant m6A regulators between the two groups. **(C)** Chromosomal characteristics of all 10 m6A regulators.

TABLE 2 Specific characteristics of m6A regulators on chromosomes.

Gene	Chromosome	Starting position	Ending position
METTL3	Chromosome 14	21498133	21511375
WTAP	Chromosome 6	159725585	159756319
RBM15	Chromosome 1	110338506	110346681
RBM15B	Chromosome 3	51391268	51397908
CBLL1	Chromosome 7	107743697	107761667
YTHDC1	Chromosome 4	68310387	68350089
YTHDF2	Chromosome 1	28736621	28769775
LRPPRC	Chromosome 2	43886508	43996005
IGFBP1	Chromosome 7	45888357	45893668
ELAVL1	Chromosome 19	7958579	8445041

osteomyelitis were well distinguished by the expression of m6A regulators in both the text and validation groups (Figures 3E,F). Visualization of the results by boxplot and heatmap suggested that the 5 regulators were differentially expressed between the two m6A subtypes, with METTL3, CBLL1 and LRPPRC highly expressed in cluster A osteomyelitis, RBM15B and YTHDC1 highly expressed in cluster B osteomyelitis (Figures 3G,H). The different m6A clusters of the test (k = 1–10) and validation groups (k = 1–9) can be found in Supplementary Figure S1.

Immune infiltration analysis of m6A subtypes

The results of ssGSEA analysis revealed that a total of 7 immune cells were significantly different in two m6A subtypes, among which activated B cells, activated CD8 T cells, CD56 bright natural killer cells, and eosinophils were highly abundant in cluster A osteomyelitis, while activated dendritic cells, natural killer cells, and type 2 T helper cells were highly abundant in cluster B osteomyelitis (Figure 4A). The correlation heatmap suggested that the METTL3 gene showed a strong positive correlation with activated B cells (r = 0.49), and

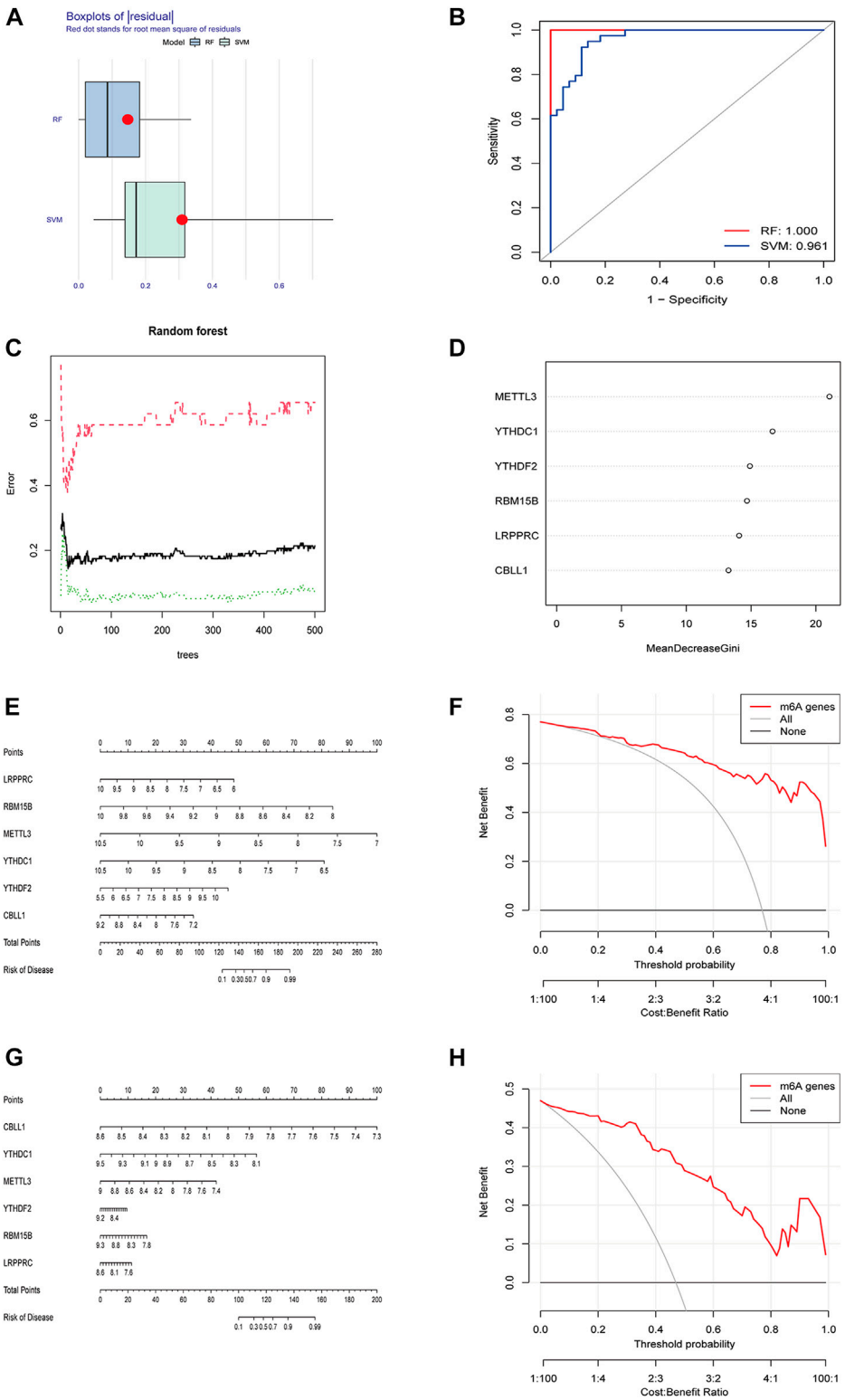


FIGURE 2
Building and validating of diagnostic models based on RF and SVM machine learning methods. **(A)** Boxplot of residuals. **(B)** ROC curves for both models. **(C)** Construction of random forest model. **(D)** MDG value of 6 significant m6A-related DEGs. **(E)** Nomogram of the model based on the GSE16129 dataset. **(F)** Calibration curve of the text model. **(G)** Nomogram of the model based on the GSE30119 dataset. **(H)** Calibration curve of the validation model. RF, random forest; SVM, support vector machine; ROC, receiver operating characteristic; MDG, Mean Decrease Gini; DCA, Decision Curve Analysis.

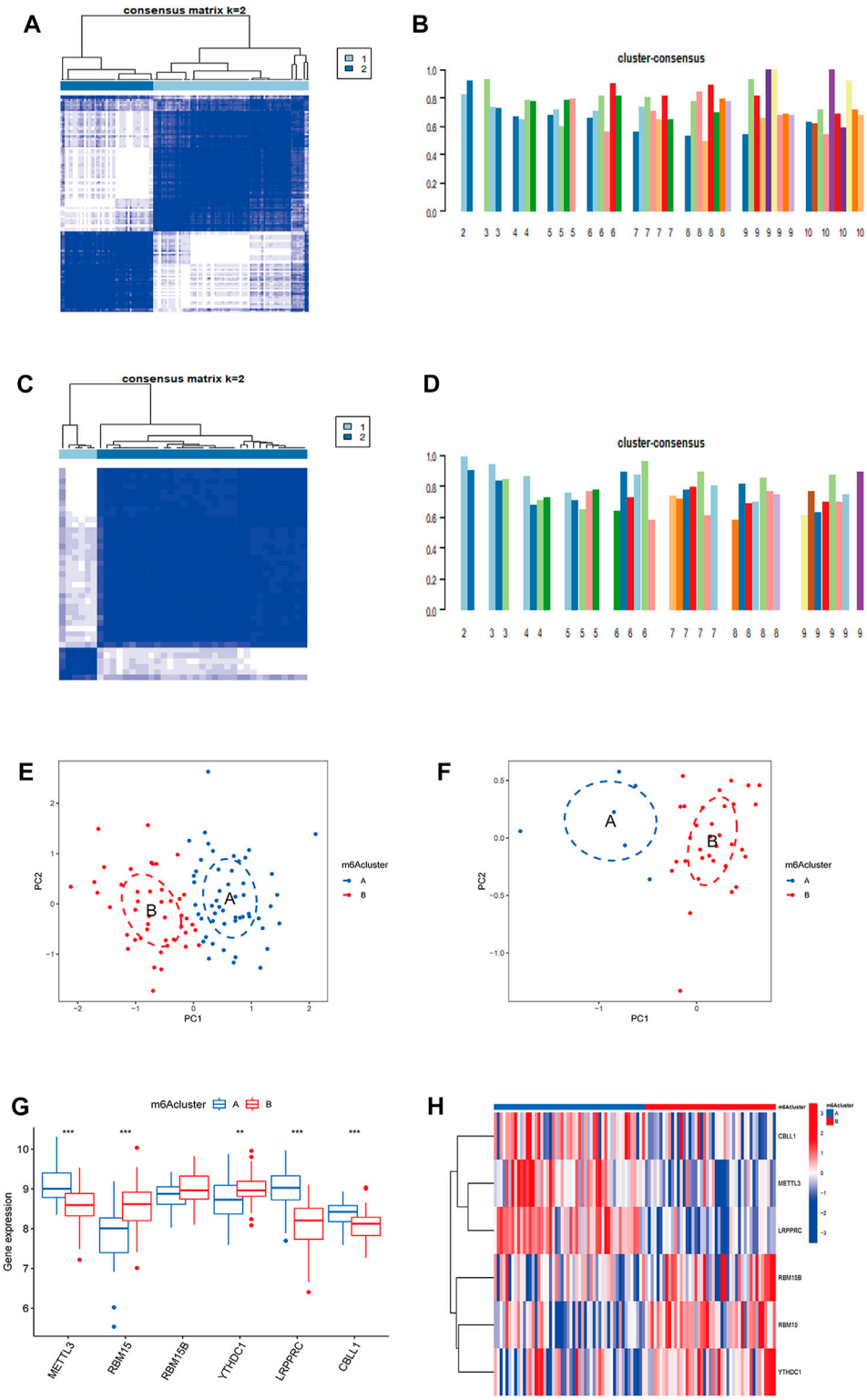


FIGURE 3
Molecular subgroups based on clustering analysis of m6A regulators. **(A)** Heatmap of 2 clusters (k = 2) based on m6A regulators from the GSE16129 dataset. **(B)** Consistency scores of all the text clusters (k = 2–10). **(C)** Heatmap of 2 clusters (k = 2) based on m6A regulators from the GSE30119 dataset. **(D)** Consistency scores of all the validation clusters (k = 2–9). **(E)** PCA analysis of the 2 clusters in the text group. **(F)** PCA analysis of the 2 clusters in the validation group. **(G)** Box diagram of differential expression of 6 significant m6A regulators in two m6A clusters from the text group. **(H)** Heatmap of differential expression of 6 significant m6A regulators in two m6A clusters from the text group.

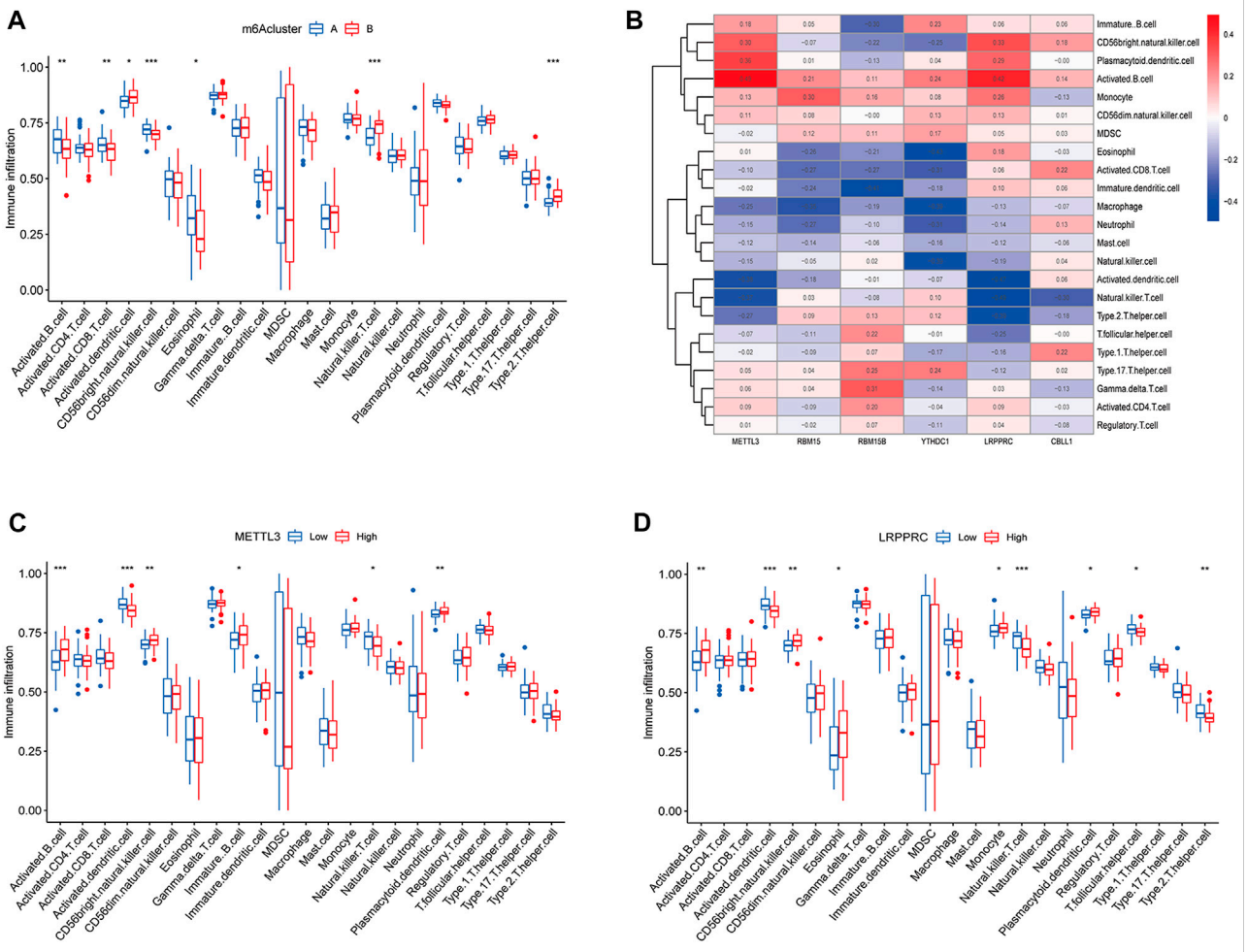


FIGURE 4
Immune cells infiltration analysis of m6A regulators. **(A)** Box diagram of immunocyte abundance in molecular subgroups of m6A regulators. **(B)** Heatmap of correlation between immunocyte abundance and m6A regulators expression. **(C)** Box diagram of comparison of immunocyte abundance in the METTL3 high- and low-expression groups. **(D)** Box diagram of comparison of immunocyte abundance in the LRPPRC high- and low-expression groups.

a strong negative correlation with activated dendritic cells (r = −0.38) and natural killer T cells (r = −0.37). The LRPPRC gene also showed a strong negative correlation with natural killer T cells (r = −0.49) (Figure 4B). Therefore, METTL3 and LRPPRC were chosen to study their abundance of immune infiltration in the high and low expression groups (Figures 4C,D). In the METTL3 high-expressed group, activated B cells, CD56 bright natural killer cells, immature dendritic cells, and plasmacytoid dendritic cells were highly infiltrated, whereas in the low-expressed group, activated dendritic cells, and natural killer T cells were highly infiltrated in the low-expressed group. In the high LRPPRC expression group, activated B cells, CD56 bright natural killer cells, eosinophils, monocytes, and plasmacytoid dendritic cells were highly infiltrated, whereas in the low expression group, activated dendritic cells, natural killer T cells, T follicular helper cells, and type 2 T helper cells were highly infiltrated in the low-expressed group.

Clinical correlation analysis of m6A subtypes

To test the clinical correlation of m6A subtypes, we analyzed the relationship between the two m6A subtypes and clinical characteristics. In the m6A subtypes of the test group, age and proportion of male were significantly lower in subgroup A than in subgroup B (Figures 5A,B). However, there were no significant difference in age and proportion of male in the two m6A subtypes of the validation group (Figures 5C,D).

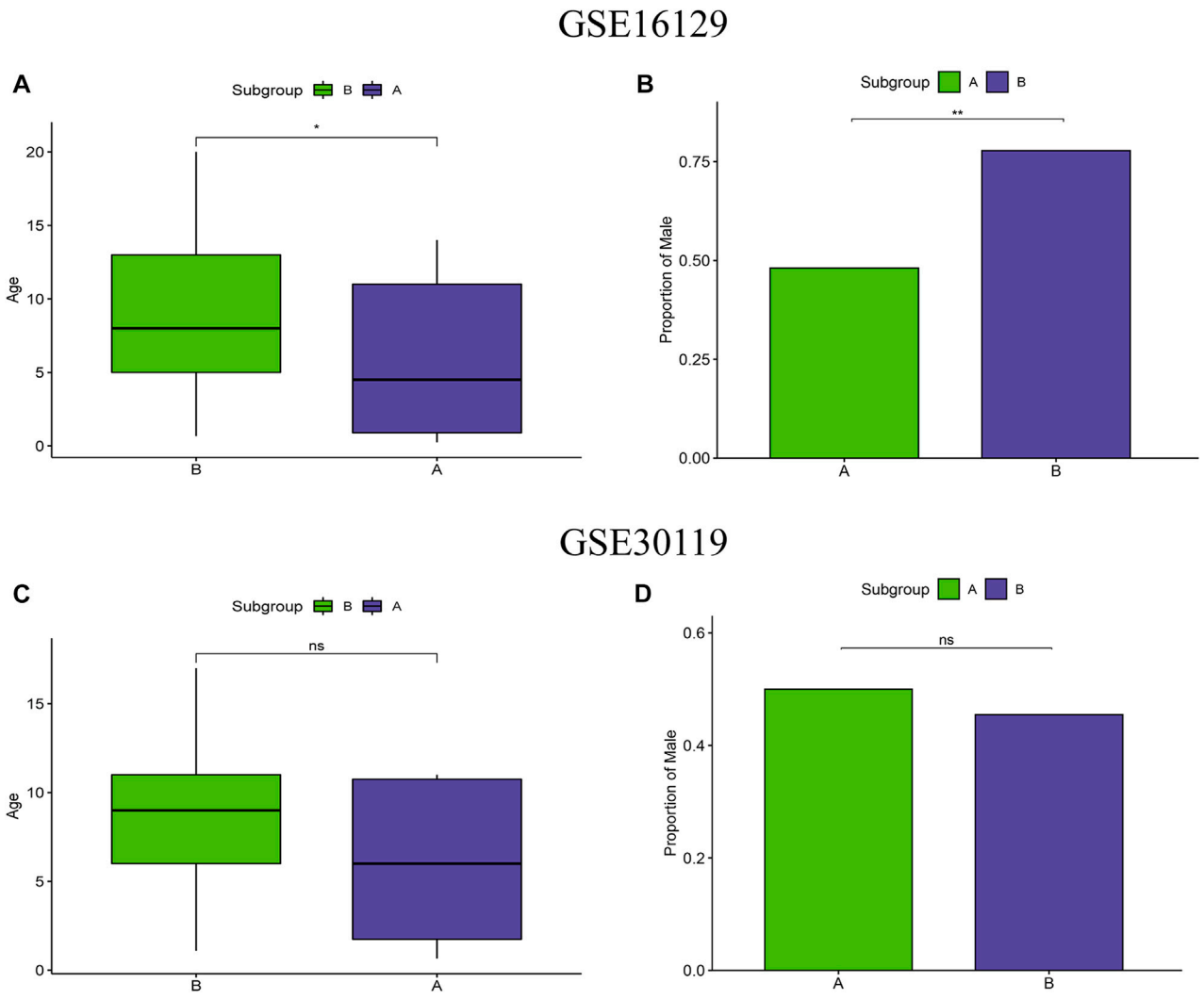


FIGURE 5
Clinical correlation analysis of two m6A clusters based on the GSE16129 and GSE30119 datasets. Differences in age **(A)** and proportion of male **(B)** between two m6A clusters based on the GSE16129 dataset. Differences in age **(C)** and proportion of male **(D)** between two m6A clusters based on the GSE30119 dataset.

Effects of GO functional annotation and KEGG pathway analysis

A total of 167 m6A-related DEGs were obtained from the two m6A subtypes using the "limma" and "VennDiagram" packages (Figure 6A). To understand the molecular functions and biological processes of the DEGs, GO annotation and KEGG pathway analysis were performed on DEGs from cluster A osteomyelitis and cluster B osteomyelitis (Figures 6B–D). In GO analysis, DEGs were mainly involved in biological processes (BP) such as T cell activation, Ras protein signal transduction, and regulation of protein stability; cellular components (CC) such as nuclear envelope, nuclear speck, and transcription regulator complex; and molecular functions (MF) such as RNA polymerase II–specific, DNA–binding transcription activator activity, and ubiquitin–like protein ligase binding. KEGG pathway analysis revealed that DEGs were mainly involved in T cell activation and proteasomal protein catabolic processes.

Immune infiltration of gene subtypes

To validate the m6A subtypes, two molecular subtypes were further confirmed based on 167 DEGs of the m6A subtypes (gene cluster A and gene cluster B) (Figures 7A,B). There were 54 cases of gene cluster A osteomyelitis and 43 cases of gene cluster B osteomyelitis. Among the 6 m6A regulators, METTL3 and

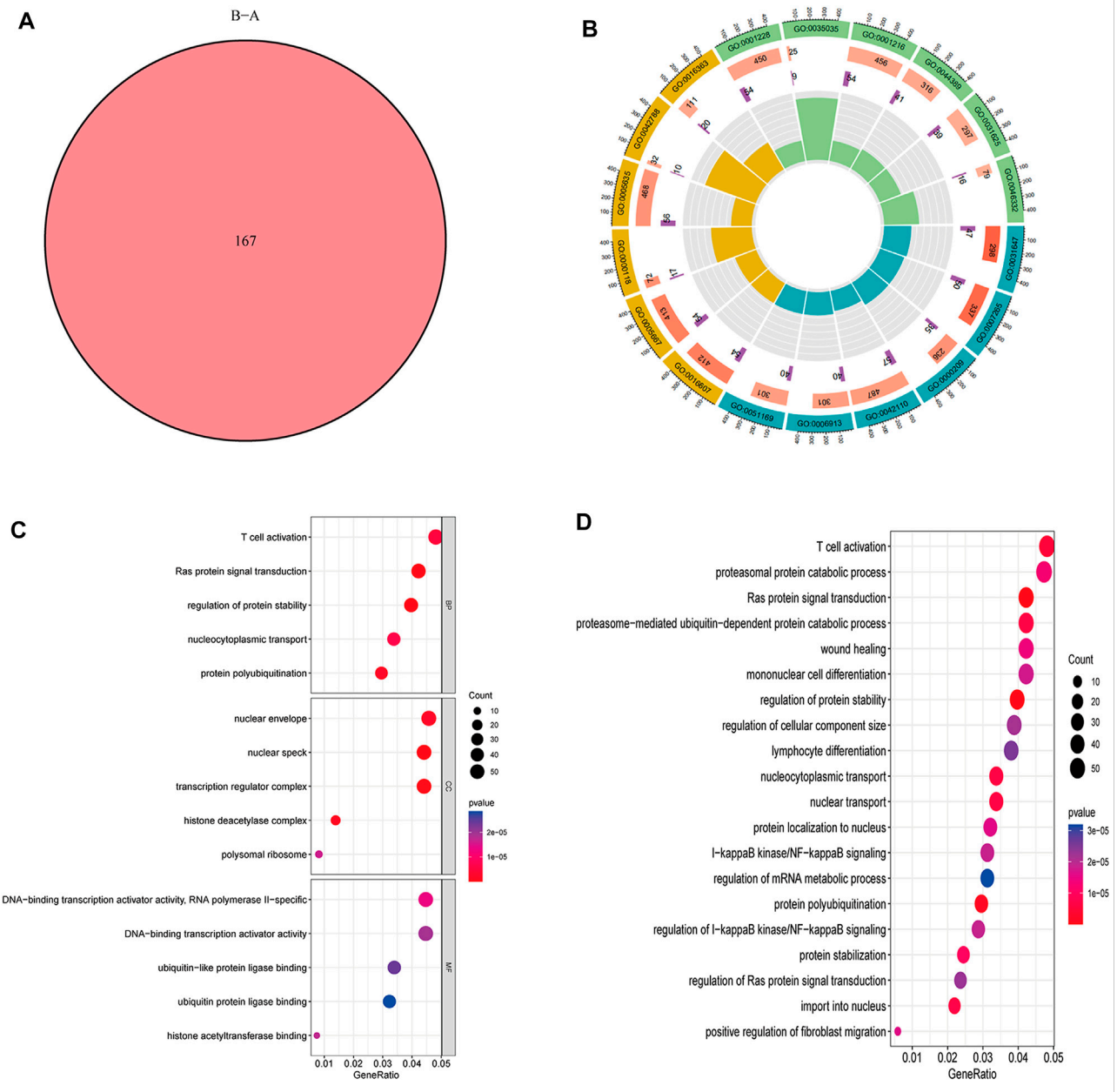


FIGURE 6
GO functional annotation and KEGG pathway analysis of DEGs based two m6A subgroups. **(A)** Venn diagram of DEGs. Circle **(B)** and bubble diagram **(C)** of GO functional annotation. **(D)** Top 20 KEGG pathway analysis of DEGs.

LRPPRC were highly expressed in subgroup A osteomyelitis, whereas YTHDC1 was highly expressed in gene cluster B osteomyelitis (Figure 7C). In addition, the results of immune infiltration showed that 9 immune cells were significantly different between the two gene subtypes, in which activated B cells, activated CD8 T cells, CD56 bright natural killer cells, eosinophils, monocytes, and regulatory T cells were highly infiltrated in gene cluster A, while activated dendritic cells, natural killer T cells, and type 2 T helper cells were highly infiltrated in gene cluster B (Figure 7D).

Differential analysis of m6A scores, collagen- and interleukin-related genes in the two typing methods

Based on the expression of m6A regulators, 97 samples from the test group and 83 samples from the validation group were scored using PCA analysis, and the score of each sample is shown in Supplementary Table S2. We then compared m6A scores in the two typing approaches, and the results showed that m6A

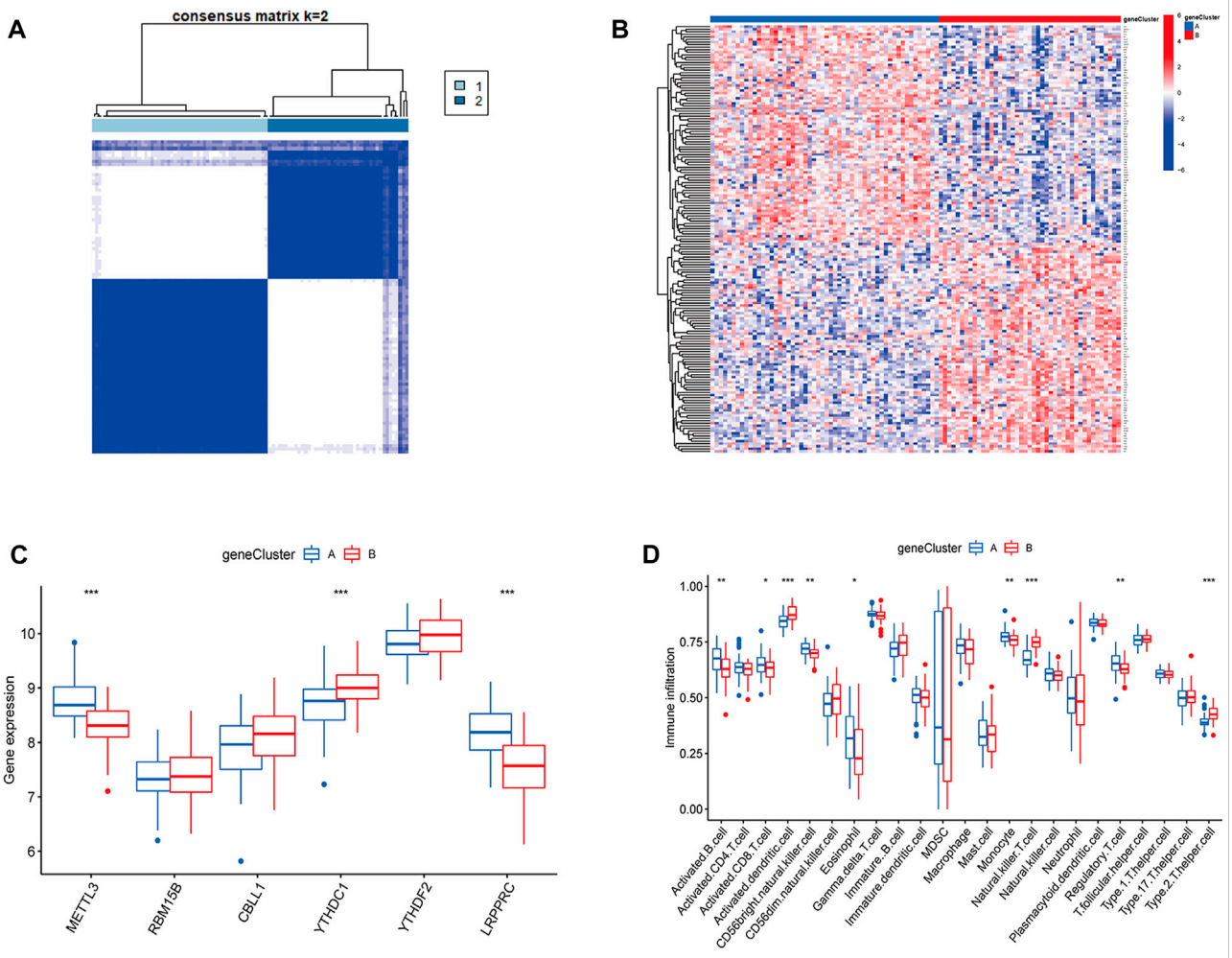


FIGURE 7
Molecular subgroups and immune cells infiltration analysis based on m6A-associated DEGs. **(A)** Clustered heatmap of 2 clusters (k = 2) based on m6A-associated DEGs. **(B)** Heatmap of typing for 167 DEGs. **(C)** Box diagram of m6A regulators between the two gene subgroups. **(D)** Box diagram of immunocyte abundance between the two gene subgroups.

scores in cluster B or gene cluster B were significantly higher than those in cluster A or gene cluster A (Figures 8A,B). The Sankey diagram suggested a high similarity between m6A subtypes and gene subtypes (Figure 8C). In the validation group, the m6A scores in cluster B or gene cluster B were also significantly higher than those in cluster A or gene cluster A (Figures 8D,E) and the Sankey diagram suggested a similarity between m6A subtypes and gene subtypes (Figure 8F).

To further reveal the relationship between m6A subtypes and the development of osteomyelitis, we found that between the two m6A subtypes from the test group, the expression of 3 collagen-related genes was significantly different, with COL4A1, COL8A2 and COL18A1 both highly expressed in cluster A osteomyelitis. The expression of 2 interleukin-related genes was significantly different, with IL11RA and IL17RA highly expressed in cluster B osteomyelitis (Figures 9A,B). Between the two gene subtypes, the expression of 3 collagen-related genes was significantly different, among which COL4A1, COL8A2 and COL18A1 were all highly expressed in gene cluster A osteomyelitis. The expression of 2 interleukin-related genes were significantly different, among which IL11RA and IL17RA were highly expressed in gene cluster B osteomyelitis (Figures 9C, D). Similar to the results of the test group, the validation group exhibited a weak correlation with collagen- and interleukin-related genes. Between the two m6A subtypes from the validation group, the expression of 2 collagen-related genes was significantly different, with COL8A2 and COL18A1 both highly expressed in cluster A osteomyelitis. The expression of 1 interleukin-related gene was significantly different, with IL11RA being highly expressed in cluster A osteomyelitis (Figures 9E,F). Between the two gene subtypes, the expression of 2 collagen-related genes was significantly different, among

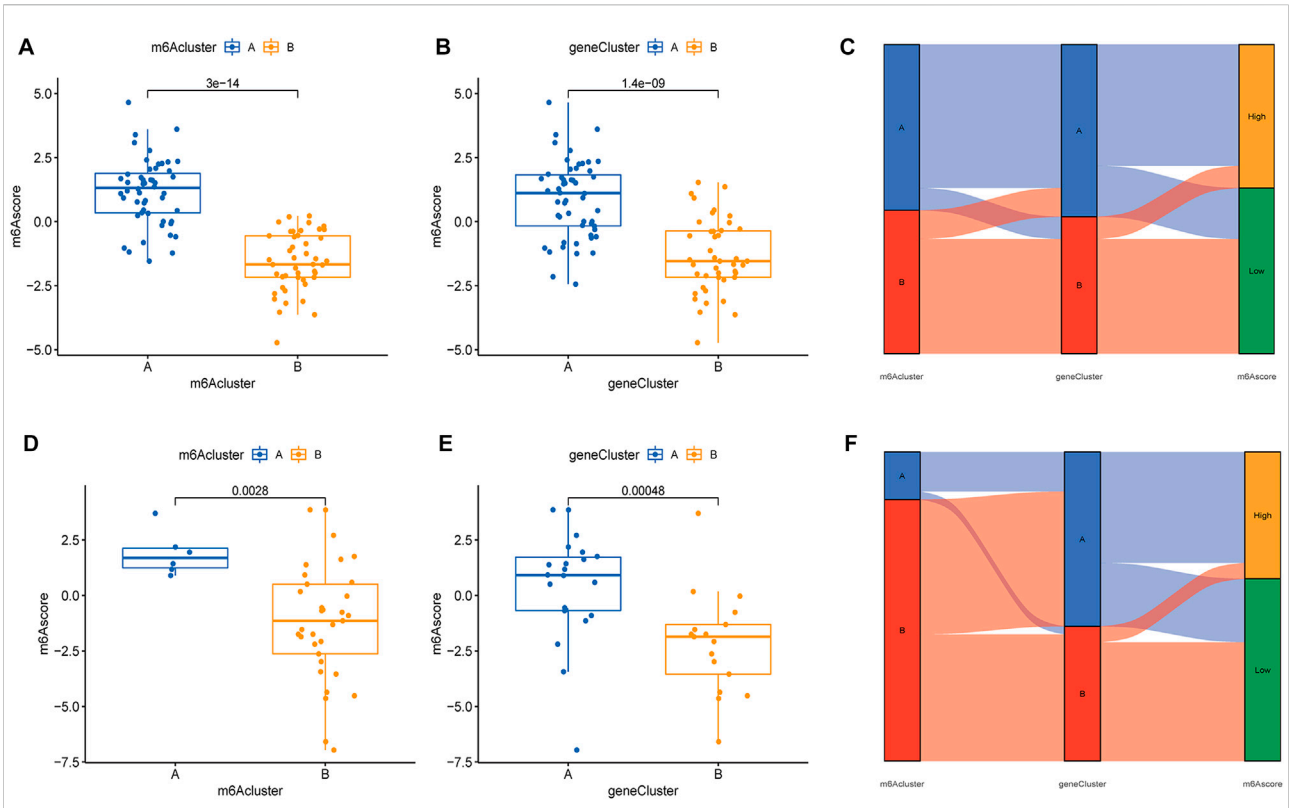


FIGURE 8
Differential analysis of m6A scores in the two typing methods based on the GSE16129 and GSE30119 datasets. Differential analysis of m6A scores in the two m6A subgroups **(A)** and two gene subgroups **(B)** based on the GSE16129. **(C)** A Sankey diagram was showed the relationship between m6A subgroups, m6A gene subgroups, and m6A scores based on the GSE16129. Differential analysis of m6A scores in the two m6A subgroups **(D)** and two gene subgroups **(E)** based on the GSE30119. **(F)** A Sankey diagram was showed the relationship between m6A subgroups, m6A gene subgroups, and m6A scores based on the GSE30119.

which COL8A2 and COL18A1 were all highly expressed in gene cluster A osteomyelitis. The expression of 2 interleukin-related genes was significantly different, among which IL11RA was highly expressed in gene cluster A osteomyelitis (Figures 9G,H).

Validation of m6A regulators in rat models of osteomyelitis

Based on the 6 significant m6A regulators with differential expression in the dataset, we further validated the expression of the 6 m6A regulators (METTL3, YTHDC1, YTHDF2, RBM15B, LRPPRC and CBLL1) in 10 osteomyelitis samples and 10 control samples by qRT–PCR analysis. The qRT–PCR results showed that the mRNA expression of METTL3, YTHDC1, RBM15B, LRPPRC and CBLL1 in osteomyelitis tissues was significantly increased, while that of YTHDF2 was significantly decreased ($p < 0.05$), which was consistent with the results of bioinformatics analysis (Figures 10A–F). To verify the effectiveness of our analysis, immunohistochemical staining was further performed on the focal bone tissue and healthy bone tissue of rat model of osteomyelitis (Figures 11A–F). The results suggested that METTL3, YTHDC1, RBM15B, LRPPRC and CBLL1 were highly expressed in the focal bone tissue of rats, while YTHDF2 was expressed at low levels. Compared with the healthy bone tissue, the positive rates of METTL3, YTHDC1, RBM15B, LRPPRC and CBLL1 in the focus bone tissue were higher, while YTHDF2 was lower ($p < 0.05$), which is consistent with the results of bioinformatics analysis and qRT–PCR.

Discussion

In this study, two datasets (GSE16129 and GSE30119) were collected, the GSE16129 dataset as the test group and the GSE30119 dataset as the validation group. A diagnostic model for osteomyelitis was established based on m6A regulators using a RF approach, and calibration curves and DCA plots demonstrated that the model fit well. In addition, we classified 97 osteomyelitis samples into two molecular subtypes based on the expression of 6 significant m6A regulators and 167 m6A-associated DEGs. The abundance of immune cell infiltration and

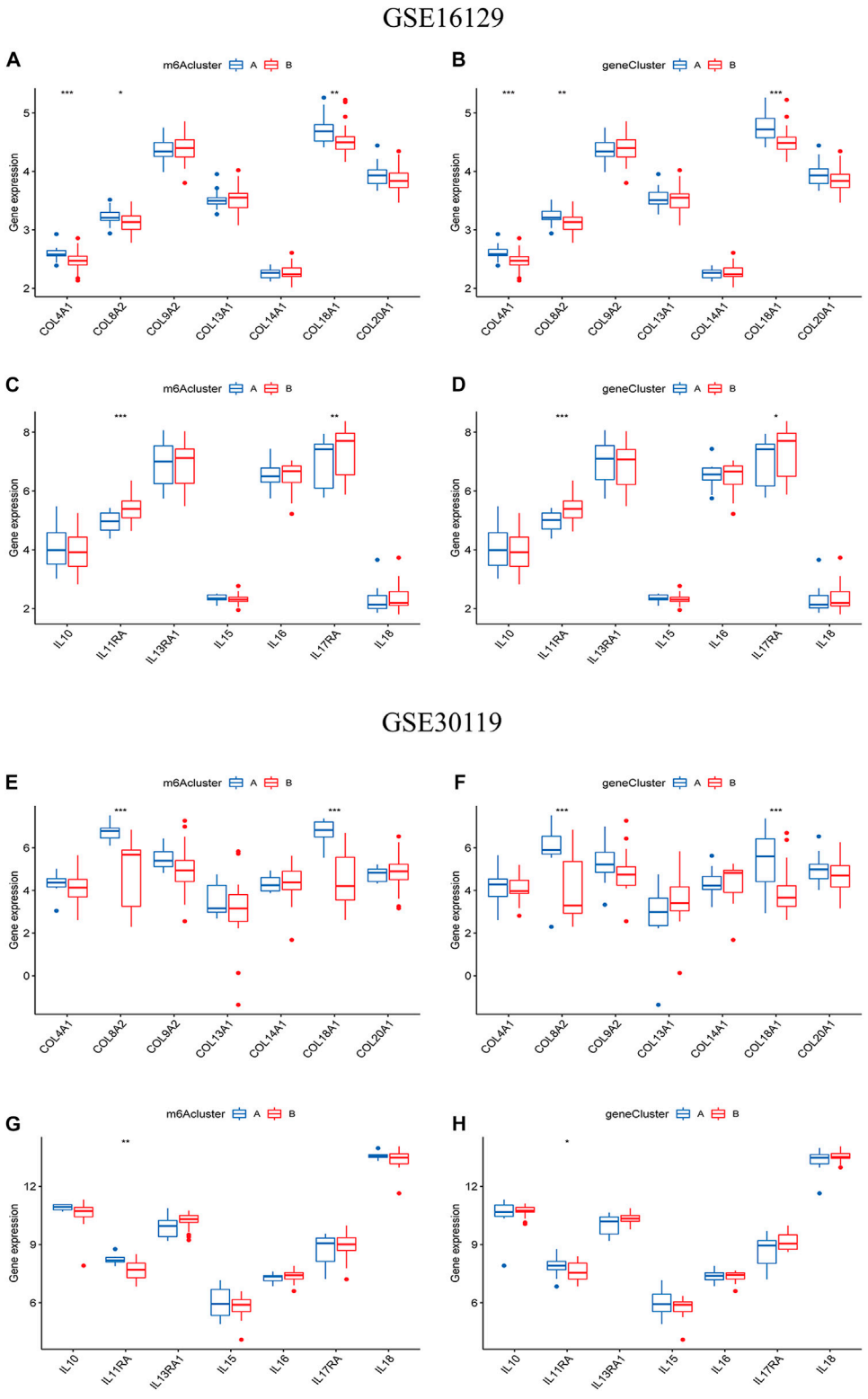


FIGURE 9
Differential analysis of collagen-related and interleukin-related genes in the two typing methods. Differential analysis of collagen-related genes between the two m6A **(A)** and two gene subgroups **(B)** based on the GSE16129. Differential analysis of interleukin-related genes between the two m6A **(C)** and two gene subgroups **(D)** based on the GSE16129. Differential analysis of collagen-related genes between the two m6A **(E)** and two gene subgroups **(F)** based on the GSE30119. Differential analysis of interleukin-related genes between the two m6A **(G)** and two gene subgroups **(H)** based on the GSE30119.

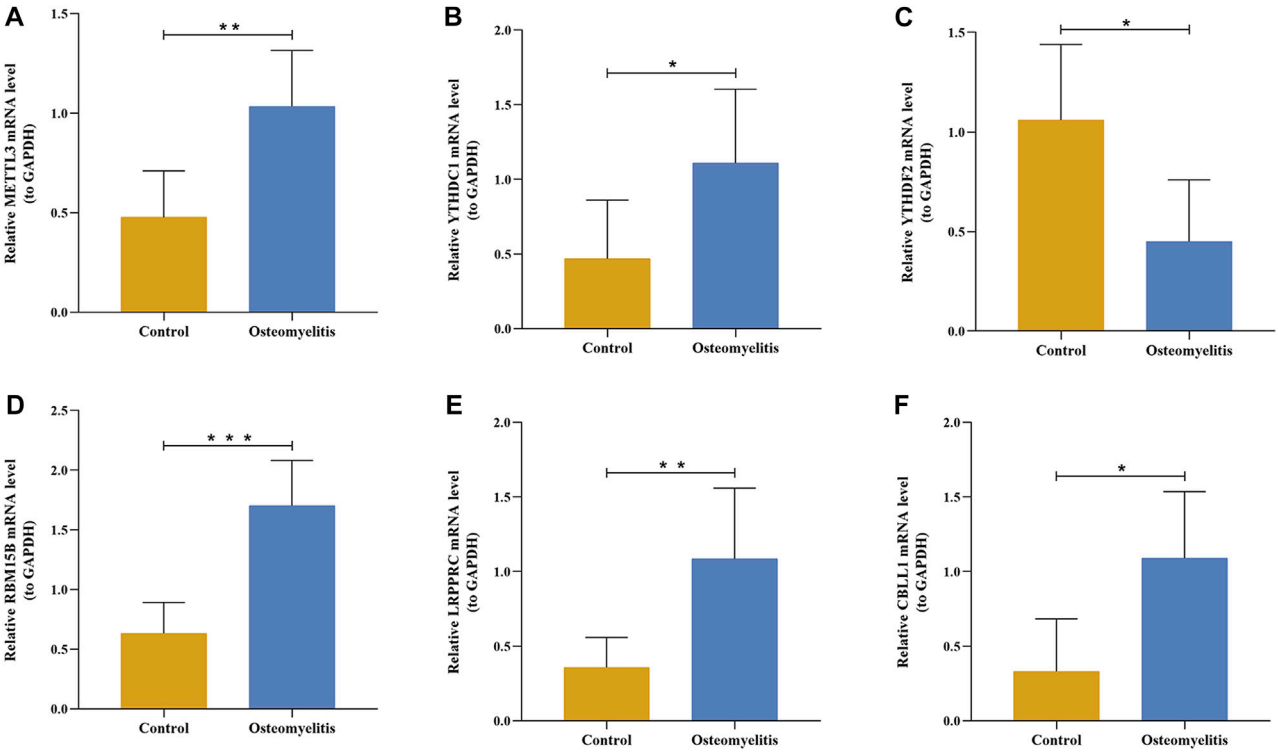


FIGURE 10
The mRNA expression levels of 6 significant m6A regulators in the focal bone tissue and healthy bone tissue of rat model of osteomyelitis. The mRNA expression levels of METTL3 **(A)**, YTHDC1 **(B)**, YTHDF2 **(C)**, RBM15B **(D)**, LRPPRC **(E)**, and CBLL1 **(F)** in focal bone tissues were significantly higher than that of healthy bone tissues.

genes involved in osteomyelitis were significantly different between the two m6A subtypes or gene subtypes. The accuracy of both the osteomyelitis diagnostic model and the m6A molecular subtypes were validated using the GSE30119 dataset. All these findings suggest that the diagnostic model and molecular subtypes were crucial to differentiate osteomyelitis. To our knowledge, this is the first study to comprehensively evaluate the functions of m6A regulators in the diagnosis and molecular subtypes of osteomyelitis.

The diagnostic model of osteomyelitis included 6 significant m6A-related DEGs, namely METTL3, YTHDC1, YTHDF2, RBM15B, LRPPRC and CBLL1, which all had importance scores greater than 10. A previous study found that METTL3 knockdown promoted the expression levels of the inflammatory cytokines IL6 and IL12, as well as the expression of osteogenic-related genes and osteoblast differentiation in an LPS-mediated inflammatory environment (Zhang et al., 2019). *In vitro*, overexpression of METTL3 can attenuate LPS-induced inflammatory responses in macrophages *via* the NF-κB pathway (Wang et al., 2019a). Silencing YTHDF2 exacerbated LPS-induced inflammatory response in RAW264.7 cells through activation of the NF-κB pathway (Yu et al., 2019). A recent study also demonstrated that METTL3 knockdown promoted proliferation and differentiation of osteoblasts under LPS-mediated inflammatory conditions and similar results were observed after YTHDF2 knockdown (Kong et al., 2022). These findings were similar to our results. We found that METTL3 was highly expressed and YTHDF2 was expressed at low levels in osteomyelitis samples and validated this finding in a rat model of osteomyelitis. It is suggested that METTL3 and YTHDF2 may be involved in promoting osteogenic differentiation and attenuating the inflammatory response in the inflammatory environment of osteomyelitis, making them potential therapeutic targets. 15 m6A regulators, including METTL3, YTHDC1, YTHDF2, RBM15B, LRPPRC and CBLL1, are involved in the pathological process of periodontitis (Zhang et al., 2021). Our study also came to similar conclusions, suggesting that these 6 m6A regulators may play a key regulatory role in the development of inflammatory diseases.

To investigate the role of m6A modification in disease, researchers were the first to conduct studies in the field of oncology (Liu et al., 2020; Zhao et al., 2021a; Li and Zhang, 2021). In the non-oncology field, studies based on consensus

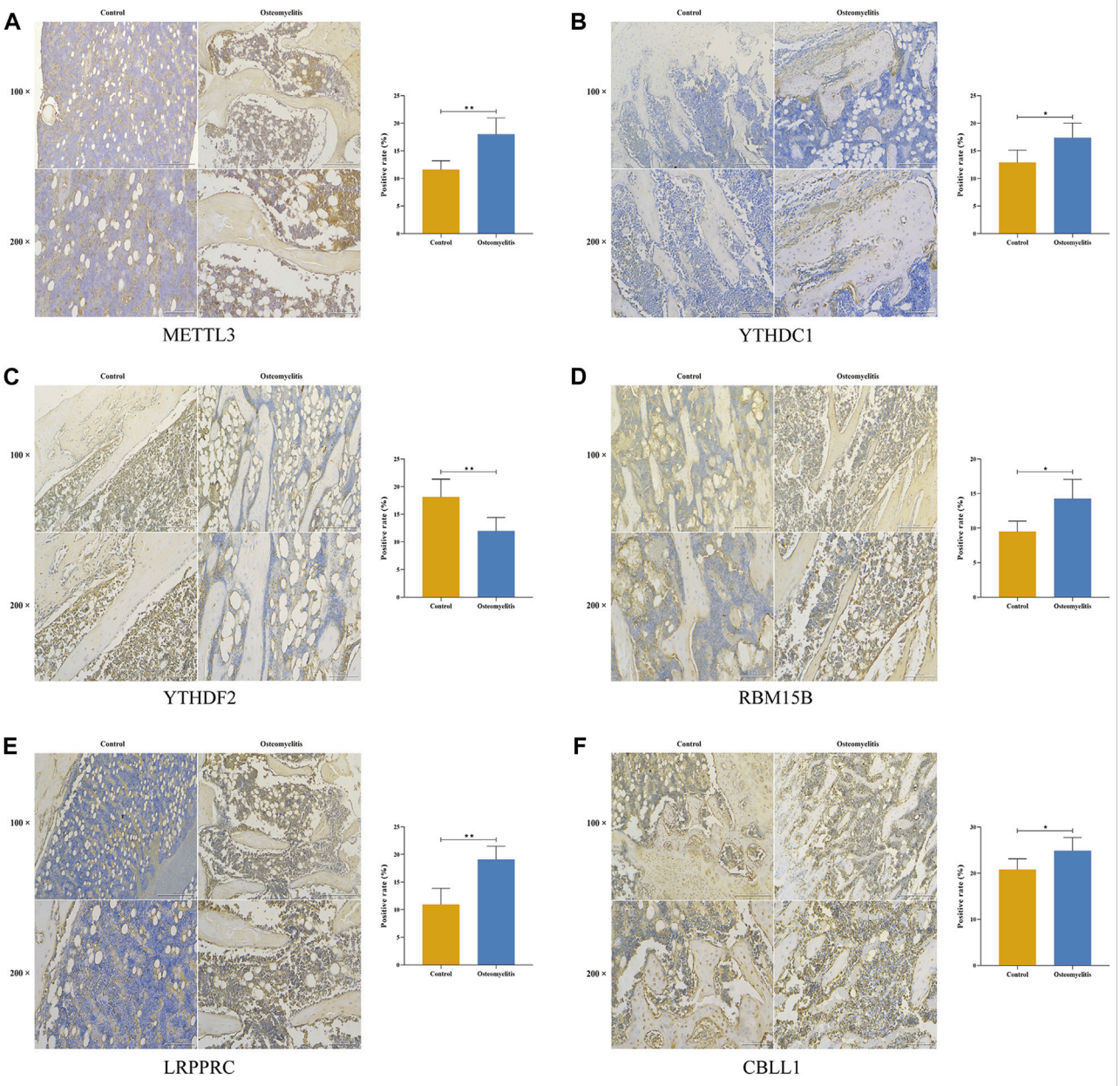


FIGURE 11
Immunohistochemistry of 6 significant m6A regulators in the focal bone tissue and healthy bone tissue of rat model of osteomyelitis. METTL3 **(A)**, YTHDC1 **(B)**, YTHDF2 **(C)**, RBM15B **(D)**, LRPPRC **(E)**, and CBLL1 **(F)** were abundantly expressed in focal bone tissues, and the positive rates were significantly higher than that of healthy bone tissues.

clustering analysis have been conducted to explore their regulatory mechanisms in periodontitis and acute myocardial infarction (Zhao et al., 2021b; Zhang et al., 2021). In the clustering analysis, two kinds of molecular subtypes were identified based on m6A regulators and gene clusters. In the m6A regulators-associated molecular subtypes, METTL3, CBLL1 and LRPPRC were significantly more highly expressed in subgroup A osteomyelitis than in subgroup B, while RBM15B and YTHDC1 expression were higher in subgroup A osteomyelitis. We obtained consistent findings in the gene subtypes. These results suggested that above 5 m6A regulators were important for distinguishing between subgroup A osteomyelitis and subgroup B osteomyelitis.

In addition, m6A modification has been proven to be involved in the regulation of the cellular immune response (Shriwas et al., 2020; Ma et al., 2021b; Liu et al., 2021). We

found significant differences in the abundance of numerous immune cells in molecular subtypes based on either m6A clusters or gene clusters. In these subtypes, 7 immune cells were different in m6A subtypes, and 9 immune cells were different in gene subtypes. Specifically, there were significant differences between activated B cells, activated CD8 T cells, activated dendritic cells, CD56 bright natural killer cells, eosinophil, natural killer T cells, and type 2 T helper cells. We further identified METTL3 and LRPPRC as two m6A regulators with strong relevance to immune cell infiltration. 9 immune cell infiltration abundances were significantly different in METTL3, and 8 immune cells were significantly different in LRPPRC. Liu et al. found that METTL3 knockdown inhibited M1-type macrophages but promoted the polarization of M2-type macrophages (Liu et al., 2019). In our study, the proportion of monocyte infiltration was higher in the group with low METTL3 expression. On the other hand, METTL3 mediates the activation and maturation of dendritic cells. MTTTL3 knockout can reduce the production of viruses in B cells infected by viruses (Hesser et al., 2018; Wang et al., 2019b). In this study, we indicated that activated B cells had higher infiltration abundance in the group with high METTL3 expression, suggesting that METTL3 may play an important role in the immune response of B cells and be closely related to the immune regulation of osteomyelitis. However, there are still few studies on the regulatory role of LRPPRC in immune cells. In addition, the inflammatory response is mainly caused by immune cells and cytokines (Luo et al., 2021). To explore the role of inflammatory cytokines in osteomyelitis, we explored the expression of interleukin-related genes in molecular subtypes. The differences between IL11RA and IL17RA in the m6A cluster and gene cluster suggested that the two genes may be involved in osteomyelitis-mediated inflammation. Additionally, collagens and collagen-related genes have been shown to be involved in a variety of bone diseases (Myllyharju and Kivirikko, 2001; Marini and Blissett, 2013; Forlino and Marini, 2016), and we also analyzed the differential expression of collagen-related genes among molecular subtypes of osteomyelitis. The results suggested that both COL4A1 and COL18A1 were highly expressed in subgroup A osteomyelitis.

Our study has some certain limitations. Firstly, we tried to find additional datasets to validate the consistency of the osteomyelitis diagnostic model and clustering analysis, but no eligible datasets were found outside the GEO database. The GSE30119, as the only validation data set, leads to clustering results with unbalanced sample allocation and future large sample datasets for verification are necessary. Secondly, this study analyzed the relationship between m6A regulators and immune cell infiltration and briefly validated the expression of key m6A regulators in rat model of osteomyelitis, but its specific regulatory mechanism in osteomyelitis-mediated inflammatory conditions requires more *in vitro* and clinical experiments to demonstrate.

Conclusion

In this study, a well-fitting nomogram model was constructed to predict the incidence of osteomyelitis based on six significant m6A regulators using a random forest approach, while two m6A isoforms were obtained using unsupervised clustering analysis. Furthermore, METTL3 and LRPPRC were determined to be closely associated with immune infiltration by immune infiltration analysis, which will provide guidance for personalized immunotherapy of osteomyelitis patients in the future.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Ethics statement

The animal study was reviewed and approved by all experimental procedures were the 920th Hospital of Joint Logistics Support Force Ethics Review Committee (2022-025-01).

Author contributions

XS and YX conceived the study and wrote the manuscript. HN and YW carried out the data collection and data analysis. MLi contributed to the data curation, methodology, and validation. All authors reviewed the results and approved the final version of the manuscript. XS and HN contributed equally to this work.

Funding

This study was funded by National Natural Science Foundation of China (Grant No.81772367, 82072392); the Yunnan Traumatology and Orthopedics Clinical Medical Center (Grant No. ZX20191001); the Grants from Yunnan Orthopedics and Sports Rehabilitation Clinical Medicine Research Center (Grant No. 202102AA310068).

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1044264/full#supplementary-material

References

Abinash, M. J., and Vasudevan, V. (2022). Boundaries tuned support vector machine (BT-SVM) classifier for cancer prediction from gene selection. *Comput. Methods Biomech. Biomed. Engin.* 25 (7), 794–807. doi:10.1080/10255842.2021.1981300

Berendt, A. R., Peters, E. J., Bakker, K., Embil, J. M., Eneroth, M., Hinchliffe, R. J., et al. (2008). Diabetic foot osteomyelitis: A progress report on diagnosis and a systematic review of treatment. *Diabetes. Metab. Res. Rev.* 24 (1), S145–S161. doi:10.1002/dmrr.836

Boulias, K., Toczydłowska-Socha, D., Hawley, B. R., Liberman, N., Takashima, K., Zaccara, S., et al. (2019). Identification of the m(6)Am methyltransferase PCIF1 reveals the location and functions of m(6)Am in the transcriptome. *Mol. Cell* 75 (3), 631–643. e8. doi:10.1016/j.molcel.2019.06.006

Brière, G., Darbo, É., Thébault, P., and Uricaru, R. (2021). Consensus clustering applied to multi-omics disease subtyping. *BMC Bioinforma.* 22 (1), 361. doi:10.1186/s12859-021-04279-1

Chokkalla, A. K., Mehta, S. L., Kim, T., Chelluboina, B., Kim, J., and Vemuganti, R. (2019). Transient focal ischemia significantly alters the m(6)A epitranscriptomic tagging of RNAs in the brain. *Stroke* 50 (10), 2912–2921. doi:10.1161/STROKEAHA.119.026433

Dai, B., Sun, F., Cai, X., Li, C., Liu, H., and Shang, Y. (2021). Significance of RNA N6-methyladenosine regulators in the diagnosis and subtype classification of childhood asthma using the gene expression omnibus database. *Front. Genet.* 12, 634162. doi:10.3389/fgene.2021.634162

David, C. C., and Jacobs, D. J. (2014). Principal component analysis: A method for determining the essential dynamics of proteins. *Methods Mol. Biol.* 1084, 193–226. doi:10.1007/978-1-62703-658-0_11

Forlino, A., and Marini, J. C. (2016). Osteogenesis imperfecta. *Lancet* 387 (10028), 1657–1671. doi:10.1016/S0140-6736(15)00728-X

Hesser, C. R., Karijolich, J., Dominissini, D., He, C., and Glaunsinger, B. A. (2018). N6-methyladenosine modification and the YTHDF2 reader protein play cell type specific roles in lytic viral gene expression during Kaposi's sarcoma-associated herpesvirus infection. *PLoS Pathog.* 14 (4), e1006995. doi:10.1371/journal.ppat.1006995

Hogan, A., Heppert, V. G., and Suda, A. J. (2013). *Arch. Orthop. Trauma Surg.* 133 (9), 1183–1196. doi:10.1007/s00402-013-1785-7

Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., and Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer Genomics. *Cancer Genomics Proteomics* 15 (1), 41–51. doi:10.21873/cgp.20063

Huang, W., Chen, T. Q., Fang, K., Zeng, Z. C., Ye, H., Chen, Y. Q., et al. (2021). N6-methyladenosine methyltransferases: Functions, regulation, and clinical potential. *J. Hematol. Oncol.* 14 (1), 103. doi:10.1186/s13045-019-0800-z

Iasonos, A., Schrag, D., Raj, G. V., and Panageas, K. S. (2008). How to build and interpret a nomogram for cancer prognosis. *J. Clin. Oncol.* 26 (8), 1364–1370. doi:10.1200/JCO.2007.12.9791

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 40, D109–D114. doi:10.1093/nar/gkr988

Kong, Y., Zhang, Y., Cai, Y., Li, D., Yi, B., and Xu, Q. (2022). METTL3 mediates osteoblast apoptosis by regulating endoplasmic reticulum stress during LPS-induced inflammation. *Cell. Signal.* 95, 110335. doi:10.1016/j.cellsig.2022.110335

Larsen, M. J., Thomassen, M., Tan, Q., Sørensen, K. P., and Kruse, T. A. (2014). Microarray-based RNA profiling of breast cancer: Batch effect removal improves cross-platform consistency. *Biomed. Res. Int.* 2014, 651751. doi:10.1155/2014/651751

Li, F., and Zhang, P. (2021). The N6-methyladenosine- (m6A-) associated genes act as strong key biomarkers for the prognosis of pancreatic adenocarcinoma. *Comput. Math. Methods Med.* 2021, 8715823. doi:10.1155/2021/8715823

Linder, B., Grozhik, A. V., Olarerin-George, A. O., Meydan, C., Mason, C. E., and Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods* 12 (8), 767–772. doi:10.1038/nmeth.3453

Liu, C., Yang, Z., Li, R., Wu, Y., Chi, M., Gao, S., et al. (2021). Potential roles of N6-methyladenosine (m6A) in immune cells. *J. Transl. Med.* 19 (1), 251. doi:10.1186/s12967-021-02918-y

Liu, L., Li, H., Hu, D., Wang, Y., Shao, W., Zhong, J., et al. (2022). Insights into N6-methyladenosine and programmed cell death in cancer. *Mol. Cancer* 21 (1), 32. doi:10.1186/s12943-022-01508-w

Liu, T., Wei, Q., Jin, J., Luo, Q., Liu, Y., Yang, Y., et al. (2020). The m6A reader YTHDF1 promotes ovarian cancer progression via augmenting EIF3C translation. *Nucleic Acids Res.* 48 (7), 3816–3831. doi:10.1093/nar/gkaa048

Liu, Y., Liu, Z., Tang, H., Shen, Y., Gong, Z., Xie, N., et al. (2019). The N(6)-methyladenosine (m(6)A)-forming enzyme METTL3 facilitates M1 macrophage polarization through the methylation of STAT1 mRNA. *Am. J. Physiol. Cell Physiol.* 317 (4), C762–c75. doi:10.1152/ajpcell.00212.2019

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25 (4), 402–408. doi:10.1006/meth.2001.1262

Luo, J., Xu, T., and Sun, K. (2021). N6-Methyladenosine RNA modification in inflammation: Roles, mechanisms, and applications. *Front. Cell Dev. Biol.* 9, 670711. doi:10.3389/fcell.2021.670711

Ma, X., Xia, W., Zong, Y., Jiang, C., Shan, H., Lin, Y., et al. (2021). Tumor necrosis factor-α promotes Staphylococcus aureus-induced osteomyelitis through downregulating endothelial nitric oxide synthase. *J. Microbiol. Immunol. Infect.* 54 (6), 1018–1027. doi:10.1016/j.jmii.2020.08.002

Ma, Z., Gao, X., Shuai, Y., Xing, X., and Ji, J. (2021). The m6A epitranscriptome opens a new charter in immune system logic. *Epigenetics* 16 (8), 819–837. doi:10.1080/15592294.2020.1827722

Marini, J. C., and Blissett, A. R. (2013). New genes in bone development: what's new in osteogenesis imperfecta. *J. Clin. Endocrinol. Metab.* 98 (8), 3095–3103. doi:10.1210/jc.2013-1505

Müller, C., Schillert, A., Röthemeier, C., Trégouët, D. A., Proust, C., Binder, H., et al. (2016). Removing batch effects from longitudinal gene expression - quantile normalization plus ComBat as best approach for microarray transcriptome data. *PLoS One* 11 (6), e0156594. doi:10.1371/journal.pone.0156594

Myllyharju, J., and Kivirikko, K. I. (2001). Collagens and collagen-related diseases. *Ann. Med.* 33 (1), 7–21. doi:10.3109/07853890109002055

Noskin, G. A., Rubin, R. J., Schentag, J. J., Kluytmans, J., Hedblom, E. C., Smulders, M., et al. (2005). The burden of *Staphylococcus aureus* infections on hospitals in the United States: An analysis of the 2000 and 2001 nationwide inpatient sample database. *Arch. Intern. Med.* 165 (15), 1756–1761. doi:10.1001/archinte.165.15.1756

Pasquet, J., Chevalier, Y., Couval, E., Bouvier, D., and Bolzinger, M. A. (2015). Zinc oxide as a new antimicrobial preservative of topical products: Interactions with common formulation ingredients. *Int. J. Pharm.* 479 (1), 88–95. doi:10.1016/j.ijpharm.2014.12.031

Peltola, H., and Pääkkönen, M. (2014). Acute osteomyelitis in children. *N. Engl. J. Med.* 370 (4), 1365–1366. doi:10.1056/NEJMc1402234

Polan, D. F., Brady, S. L., and Kaufman, R. A. (2016). Tissue segmentation of computed tomography images using a random forest algorithm: A feasibility study. *Phys. Med. Biol.* 61 (17), 6553–6569. doi:10.1088/0031-9155/61/17/6553

Ramos-Vara, J. A. (2005). Technical aspects of immunohistochemistry. *Vet. Pathol.* 42 (4), 405–426. doi:10.1354/vp.42-4-405

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007

Scott, R. J., Christofersen, M. R., Robertson, W. W., Jr., Davidson, R. S., Rankin, L., and Drummond, D. S. (1990). Acute osteomyelitis in children: A review of 116 cases. *J. Pediatr. Orthop.* 10 (5), 649–652. doi:10.1097/01241398-199009000-00015

Seiler, M., Huang, C. C., Szalma, S., and Bhanot, G. (2010). ConsensusCluster: A software tool for unsupervised cluster discovery in numerical data. *Omics* 14 (1), 109–113. doi:10.1089/omi.2009.0083

Shi, H., Wei, J., and He, C. (2019). Where, when, and how: Context-dependent functions of RNA methylation writers, readers, and erasers. *Mol. Cell* 74 (4), 640–650. doi:10.1016/j.molcel.2019.04.025

Shriwas, O., Mohapatra, P., Mohanty, S., and Dash, R. (2020). The impact of m6A RNA modification in therapy resistance of cancer: Implication in chemotherapy, radiotherapy, and immunotherapy. *Front. Oncol.* 10, 612337. doi:10.3389/fonc.2020.612337

The gene Ontology (GO) project in 2006. *Nucleic Acids Res.* 2006;34:D322–D326. doi:10.1093/nar/gkj021

Unkila-Kallio, L., Kallio, M. J., Eskola, J., and Peltola, H. (1994). Serum C-reactive protein, erythrocyte sedimentation rate, and white blood cell count in acute hematogenous osteomyelitis of children. *Pediatrics* 93 (1), 59–62. doi:10.1542/peds.93.1.59

Vergidis, P., Schmidt-Malan, S. M., Mandrekar, J. N., Steckelberg, J. M., and Patel, R. (2015). Comparative activities of vancomycin, tigecycline and rifampin in a rat model of methicillin-resistant *Staphylococcus aureus* osteomyelitis. *J. Infect.* 70 (6), 609–615. doi:10.1016/j.jinf.2014.12.016

Wang, H., Hu, X., Huang, M., Liu, J., Gu, Y., Ma, L., et al. (2019). Mettl3-mediated mRNA m(6)A methylation promotes dendritic cell activation. *Nat. Commun.* 10 (1), 1898. doi:10.1038/s41467-019-09903-6

Wang, J., Yan, S., Lu, H., Wang, S., and Xu, D. (2019). METTL3 attenuates LPS-induced inflammatory response in macrophages via NF-κB signaling pathway. *Mediat. Inflamm.* 2019, 3120391. doi:10.1155/2019/3120391

Wang, T., Kong, S., Tao, M., and Ju, S. (2020). The potential role of RNA N6-methyladenosine in Cancer progression. *Mol. Cancer* 19 (1), 88. doi:10.1186/s12943-020-01204-7

Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26 (12), 1572–1573. doi:10.1093/bioinformatics/btq170

Xiao, B., Liu, L., Li, A., Xiang, C., Wang, P., Li, H., et al. (2020). Identification and verification of immune-related gene prognostic signature based on ssGSEA for osteosarcoma. *Front. Oncol.* 10, 607622. doi:10.3389/fonc.2020.607622

Xiao, W., Adhikari, S., Dahal, U., Chen, Y. S., Hao, Y. J., Sun, B. F., et al. (2016). Nuclear m(6)A reader YTHDC1 regulates mRNA splicing. *Mol. Cell* 61 (4), 507–519. doi:10.1016/j.molcel.2016.01.012

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Yu, R., Li, Q., Feng, Z., Cai, L., and Xu, Q. (2019). m6A reader YTHDF2 regulates LPS-induced inflammatory response. *Int. J. Mol. Sci.* 20 (6), E1323. doi:10.3390/ijms20061323

Zalavras, C. G. (2017). Prevention of infection in open fractures. *Infect. Dis. Clin. North Am.* 31 (2), 339–352. doi:10.1016/j.idc.2017.01.005

Zhang, B., Wu, Q., Li, B., Wang, D., Wang, L., and Zhou, Y. L. (2020). m(6 A regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. *Mol. Cancer* 19 (1), 53. doi:10.1186/s12943-020-01170-0

Zhang, H., Meltzer, P., and Davis, S. (2013). RCircos: an R package for Circos 2D track plots. *BMC Bioinforma.* 14, 244. doi:10.1186/1471-2105-14-244

Zhang, X., Zhang, S., Yan, X., Shan, Y., Liu, L., Zhou, J., et al. (2021). m6A regulator-mediated RNA methylation modification patterns are involved in immune microenvironment regulation of periodontitis. *J. Cell. Mol. Med.* 25 (7), 3634–3645. doi:10.1111/jcmm.16469

Zhang, Y., Gu, X., Li, D., Cai, L., and Xu, Q. (2019). METTL3 regulates osteoblast differentiation and inflammatory response via Smad signaling and MAPK signaling. *Int. J. Mol. Sci.* 21 (1), E199. doi:10.3390/ijms21010199

Zhang, Y., Parmigiani, G., and Johnson, W. E. (2020). ComBat-seq: Batch effect adjustment for RNA-seq count data. *Nar. Genom. Bioinform.* 2 (3), lqaa078. doi:10.1093/nargab/lqaa078

Zhao, H., Xu, Y., Xie, Y., Zhang, L., Gao, M., Li, S., et al. (2021). m6A regulators is differently expressed and correlated with immune response of esophageal cancer. *Front. Cell Dev. Biol.* 9, 650023. doi:10.3389/fcell.2021.650023

Zhao, S., Guo, Y., Sheng, Q., and Shyr, Y. (2014). Advanced heat map and clustering analysis using heatmap3. *Biomed. Res. Int.* 2014, 986048. doi:10.1155/2014/986048

Zhao, W., Qi, X., Liu, L., Ma, S., Liu, J., and Wu, J. (2020). Epigenetic regulation of m(6)A modifications in human cancer. *Mol. Ther. Nucleic Acids* 19, 405–412. doi:10.1016/j.omtn.2019.11.022

Zhao, X., Ge, L., Wang, J., Song, Z., Ni, B., He, X., et al. (2021). Exploration of potential integrated models of N6-methyladenosine immunity in systemic lupus erythematosus by bioinformatic analyses. *Front. Immunol.* 12, 752736. doi:10.3389/fimmu.2021.752736

Zhu, Z. M., Huo, F. C., and Pei, D. S. (2020). Function and evolution of RNA N6-methyladenosine modification. *Int. J. Biol. Sci.* 16 (11), 1929–1940. doi:10.7150/ijbs.45231

Zong, X., Zhao, J., Wang, H., Lu, Z., Wang, F., Du, H., et al. (2019). Mettl3 deficiency sustains long-chain fatty acid absorption through suppressing traf6-dependent inflammation response. *J. Immunol.* 202 (2), 567–578. doi:10.4049/jimmunol.1801151

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 05 December 2022
DOI 10.3389/fgene.2022.1079739

OPEN ACCESS

EDITED BY
Lu Qiao,
Seoul National University, South Korea

REVIEWED BY
Jin Dai,
Children's Hospital of Soochow University, China
Xi Wang,
Xi'an Jiaotong University, China

*CORRESPONDENCE
Jiaxiang Wang,
wjiaxiang@zzu.edu.cn
Yingzhong Fan,
zdyfyxew@126.com

SPECIALTY SECTION
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

RECEIVED 25 October 2022
ACCEPTED 18 November 2022
PUBLISHED 05 December 2022

CITATION
Shi L, Liu Q, Yang H, Wang Q, Wang J and Fan Y (2022), Inflammation-related pathways involved in damaged articular cartilage of rats exposed to T-2 toxin based on RNA-sequencing analysis. *Front. Genet.* 13:1079739. doi: 10.3389/fgene.2022.1079739

Inflammation-related pathways involved in damaged articular cartilage of rats exposed to T-2 toxin based on RNA-sequencing analysis

Longyan Shi, Qiuliang Liu, Heying Yang, Qi Wang, Jiaxiang Wang* and Yingzhong Fan*

Department of Pediatric Surgery, First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China

Many studies have shown that ingestion of the T-2 toxin is harmful to articular cartilage. However, the mechanisms underlying damaged articular cartilage induced by T-2 toxin have not been elucidated. Twenty-four SD rats were randomly divided into T-2 toxin and control groups. In the control group, the 12 rats were administered 4% absolute ethanol by gavage, and in the T-2 toxin group, the 12 rats were administered T-2 toxin (100 ng/g, BW/day) by gavage. After the rats were sacrificed, the knee joints were collected, and RNA was extracted using TRIzol reagent for RNA sequencing (RNA-seq). Differentially expressed mRNA was identified based on $p < 0.05$ and $|\log_2(\text{fold change})| > 1$. The T-2 toxin-related genes were obtained from the GeneCards database. An online tool (https://www.bioinformatics.com.cn) was used for enrichment analysis. Hematoxylin and eosin (H&E) staining was used to observe damaged articular cartilage, and immunohistochemical (IHC) staining was used to validate differentially expressed proteins. The H&E staining shows the number of cells decreased significantly, and the arrangement of chondrocytes became disordered in the T-2 toxin group. RNA-seq analysis identified 195 upregulated and 89 downregulated mRNAs in the T-2 toxin group. The top immune-related biological processes (Gene Ontology) were regulation of hormone secretion, regulation of peptide hormone secretion, and regulation of transcription involved in cell fate commitment. KEGG pathway enrichment analysis revealed that the IL-17 and tumor necrosis factor signaling pathways were significantly expressed, and the IL-17 signaling pathway was also identified in the enrichment analysis of T-2 toxin-related genes. Also, Mmp3, Tnf, Mapk10, Ccl11, Creb5, Cxcl2, and Cebpb were significantly enriched in the two pathways. The immunohistochemical staining showed that the levels of Mmp3 and Tnf proteins were significantly increased in the T-2 toxin group, which was consistent with the RNA-seq results. This study revealed the critical

roles of IL-17 and TNF signaling pathways in damaged cartilage induced by T-2 toxin.

KEYWORDS

T-2 toxin, articular cartilage, inflammatory response, RNA-seq, IL-17 signaling pathway, TNF signaling pathway

Introduction

The T-2 toxin is the most toxic class A trichothecene mycotoxin produced by *Fusarium*. It is widely distributed in nature and mainly contaminates cereals and grains, such as wheat, barley, oats, rice, maize, and sorghum, and is harmful to humans and poultry. At present, it is believed that the T-2 toxin can cause a variety of human and animal diseases such as alimentary toxic aleukia and Kashin–Beck disease (KBD) (Lutsky et al., 1978; Li et al., 2016). The Soviet Union scholar proposed the hypothesis of food mycotoxin contamination in 1964, and the causative organism was considered to be a toxigenic strain of *Fusarium sporotrichiella. var. poae Bilai* (Sokoloff, 1988). The rats fed with *Fusarium* showed epiphyseal lesions and femur shortening, while the rats in the control group developed normally (Guo et al., 2014). In 1990s, Yang Jianbo proposed a relationship between the T-2 toxin and KBD according to epidemiological investigations (Dianjun Sun; Wang et al., 2020). In 1995, he investigated edible grains of severe KBD areas in Shaanxi and Sichuan provinces and the commercially available grains in non-KBD areas from Henan, Hubei, Hunan, and Guangxi provinces. The results showed the concentration of T-2 toxin (ranging from 2.0 to 1,549.4 ng/g) in 10 out of 15 family-grown flour samples in the KBD areas was positive, which was significantly higher than that of the commercial flour samples in non-KBD areas (ranging from 2.0 to 885.3 ng/g). Meanwhile, the study also confirmed the threshold concentration of T-2 toxin in flour should not exceed 300 ng/g (Yang et al., 1995). Therefore, the concentration of T-2 toxin in the flour and corn meal of families affected by KBD resulted in an abnormal aggregation phenomenon.

According to the food mycotoxin contamination hypothesis, the grain harvest season is rainy, and thus, fungi contaminate the grain during harvesting, drying, and storage. As healthy children in the KBD areas consumed fungus-contaminated grains, the T-2 toxin selectively damaged the developing articular cartilage, resulting in damage to the membrane system of chondrocytes, inhibition of DNA synthesis, and induction of chondrocytes apoptosis (Guo et al., 2014). When chondrocytes were incubated with T-2 toxin, the growth rate decreased, the protein and DNA contents were reduced, levels of collagen II and proteoglycan decreased, expression of matrix metalloproteinase 13 increased, and chondrocyte apoptosis was induced (Liu et al., 2021; Li H. N. et al., 2022). When the chicks were fed diets containing T-2 toxin for 5 weeks, their articular cartilage underwent degenerative changes, which caused deep necrosis of the epiphyseal plate cartilage (Yu et al., 2022). Wistar rats were treated with T-2 toxin (100 ng/kg) for 10 months, and the results showed chondrocyte degeneration/necrosis and loss, chondrocyte clones, and loss of proteoglycan staining of articular cartilage in the femorotibial cartilage. The rats exposed to T-2 toxin showed degenerative lesions in articular cartilage similar to spontaneous osteoarthritis (Wang et al., 2011; Yu et al., 2022). However, the deep necrosis of articular cartilage in human KBD has not been replicated.

A previous study examined the DNA methylation dynamics of human chondrocytes treated with T-2 toxin and found that the dysfunction of the MAPK pathway was involved in the damaged human chondrocyte (Yang et al., 2022). However, previous studies have only been performed at the cellular level and have not been validated in cartilage tissues. In this study, SD rats were treated with T-2 toxin (100 ng/g) for 4 weeks, and the articular cartilage tissues were collected. RNA transcription sequencing technology was used to explore the mechanism of damaged articular cartilage induced by the T-2 toxin.

Methods and materials

Establishment and treatment of animal model

A total of 24 SPF-grade SD rats (aged 4 weeks, 60–80 g) were purchased from the Henan Provincial Laboratory Animal Center. They were maintained in a spacious animal house and fed a standard mouse diet and water *ad libitum*. The feed and ultrapure water were kept for 1 week, and after weighing, the rats were randomly divided into the control and T-2 toxin groups. The 12 rats were administered 4% absolute ethanol by gavage in the control group, and the 12 rats were administered T-2 toxin (100 ng/g; BW/day; J&K Scientific LLC, USA) by gavage in the T-2 toxin group (Yu et al., 2022). During this period, the weight and other indicators of their growth and development were recorded. After treatment for 1 month, three rats in each group were anesthetized with urethane, and three articular cartilages were used to extract RNA. All animal studies were reviewed and approved by the Animal Ethics Research Committee of Zhengzhou University (ZZUIRB 2021-69).

Hematoxylin–eosin (H&E) staining

After the rats were sacrificed, the knee joints were collected, fixed in 4% paraformaldehyde for 24 h, and decalcified with 10% EDTA for 1 month. The decalcified knee joints were embedded in paraffin

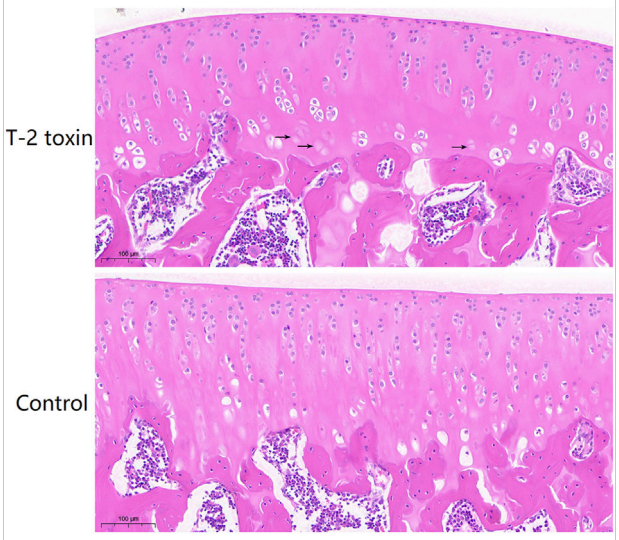


FIGURE 1
Pathological changes of rat cartilage tissues in the T-2 toxin and control groups using H&E staining. The black arrows indicated chondrocytes without the nucleus.

and sectioned. The tissue sections were stained with H&E. Finally, the stained tissue sections were dehydrated and sealed with neutral gum, and the collected images were observed under a microscope.

Total RNA extraction

RNA was extracted using the TRIzol reagent (Thermo Fisher Scientific, USA) for RNA sequencing (RNA-seq). After collecting the RNA samples, the concentration was accurately quantified using Qubit, and the ratios of $OD_{260/280}$ and $OD_{260/230}$ were detected using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, USA). Then, the degree of RNA degradation was detected by agarose gel electrophoresis and an Agilent 2100 Bioanalyzer (Agilent Technologies, USA).

Construction of the DNA-specific library

An appropriate amount of total RNA was used to remove rRNA, and the RNA was purified and fragmented. Reverse transcription was used to synthesize the first strand cDNA and second strand cDNA. Next, a single adenylate "A" was added to the 3′ ends on both sides of the modified double-stranded cDNA to prevent self-ligation of the flat ends between the cDNA fragments. Ligation buffer and a double-stranded sequencing linker were added to the aforementioned reaction system, and the Illumina sequencing linker was connected to both ends of the DNA library using T4 DNA ligase (NEB, UK). The Agencourt SPRI select Nucleic Acid Fragment Screening kit was used to screen the fragment size while the library was being purified, and a two-

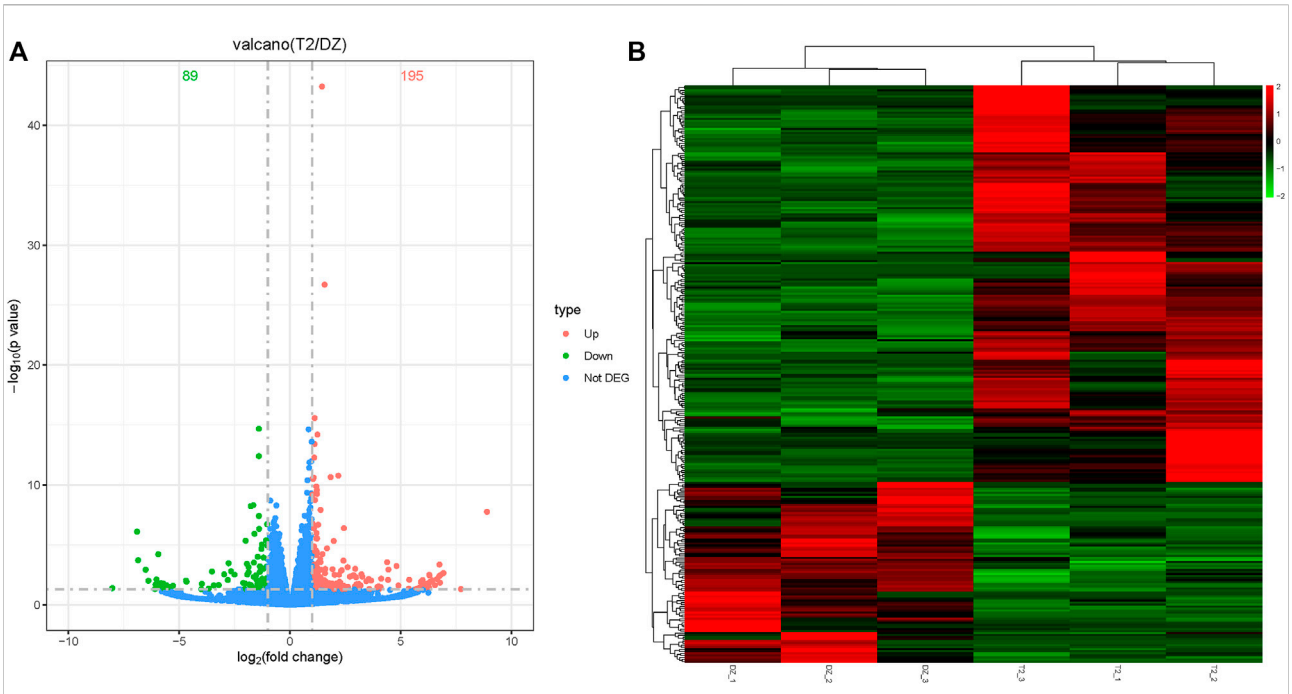


FIGURE 2
Differentially expressed mRNA analysis and cluster analysis. **(A)** Volcano of differentially expressed mRNA in the T-2 toxin group compared with the control group. Green represents downregulated mRNA, and orange represents upregulated mRNA. **(B)** Hierarchical clustering of differentially expressed mRNA in the T-2 toxin group compared with the control group. Green represents downregulated mRNA, and red represents upregulated mRNA.

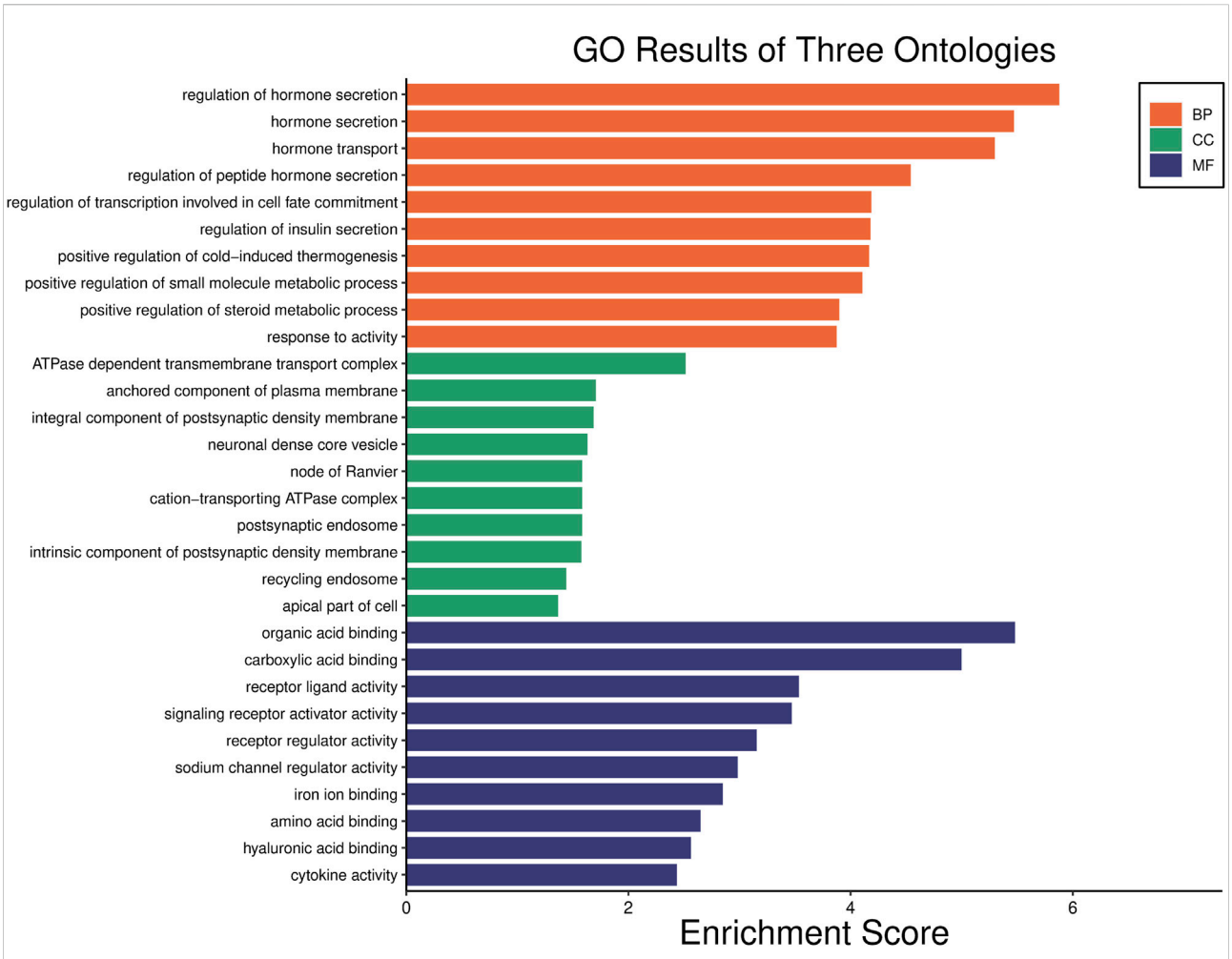


FIGURE 3
Gene Ontology (GO) functional classification. The significant GO terms of differentially expressed genes in the T-2 toxin group compared with the control group. The orange, green, and blue bars represent the biological process, cellular component, and molecular function, respectively.

step screening method was used to select the original library with a fragment peak of 300 bp. In a 50-μl reaction system, high fidelity polymerase was used to amplify the original library to ensure a sufficient total number of libraries. The constructed library was qualified using Qubit and an Agilent 2100 Bioanalyzer (Agilent Technologies, USA) and then sequenced using an Illumina HiSeq 2500 System (Illumina, USA).

Differential expression mRNA analysis

Deseq2 software was used to analyze differentially expressed mRNA in the T-2 toxin and control groups. p-values <0.05 and $|\log_2$ (fold change)$| > 1$ were considered differentially expressed mRNAs, $\log_2$ (fold change) > 1 was labeled as an upregulated mRNA, and $\log_2$ (fold change) < −1 was labeled as a downregulated mRNA.

Enrichment analysis

Gene Ontology (GO) is an international standard classification system for gene functions. Through GO enrichment analysis, significantly differentially expressed genes were classified according to cellular components, molecular functions, and biological processes. KEGG (Kyoto Encyclopedia of Genes and Genomes) is a vital public database of classical pathways for experimental verification. Pathway enrichment analysis takes the KEGG pathway as the unit and applies a hypergeometric test to determine which pathways are significantly enriched in a specific gene set compared to the whole genome background. The analysis was based on differentially expressed genes marked as up and down, and an online tool (https://www.bioinformatics.com.cn) was used to perform GO and KEGG enrichment analyses (Luo & Brouwer, 2013).

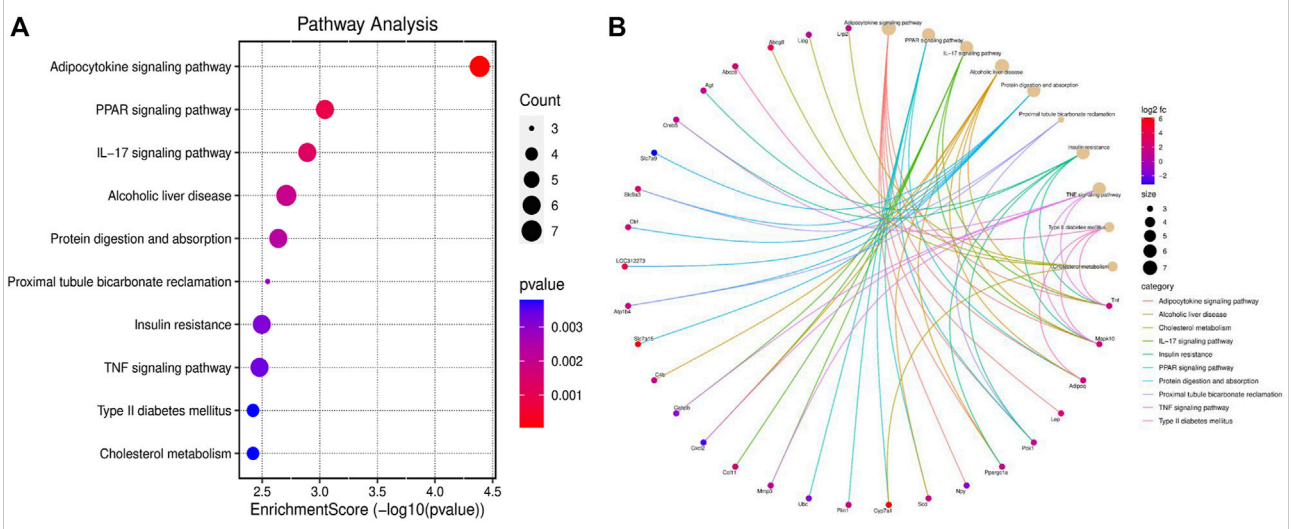


FIGURE 4
KEGG pathway enrichment analysis (top 10). **(A)** Top 10 significant KEGG pathways of differentially expressed genes in the T-2 toxin group compared with the control group. **(B)** Relationship between significant pathways and enriched genes, and different colors represent different pathways. *p*-value <0.05 was considered significant pathways.

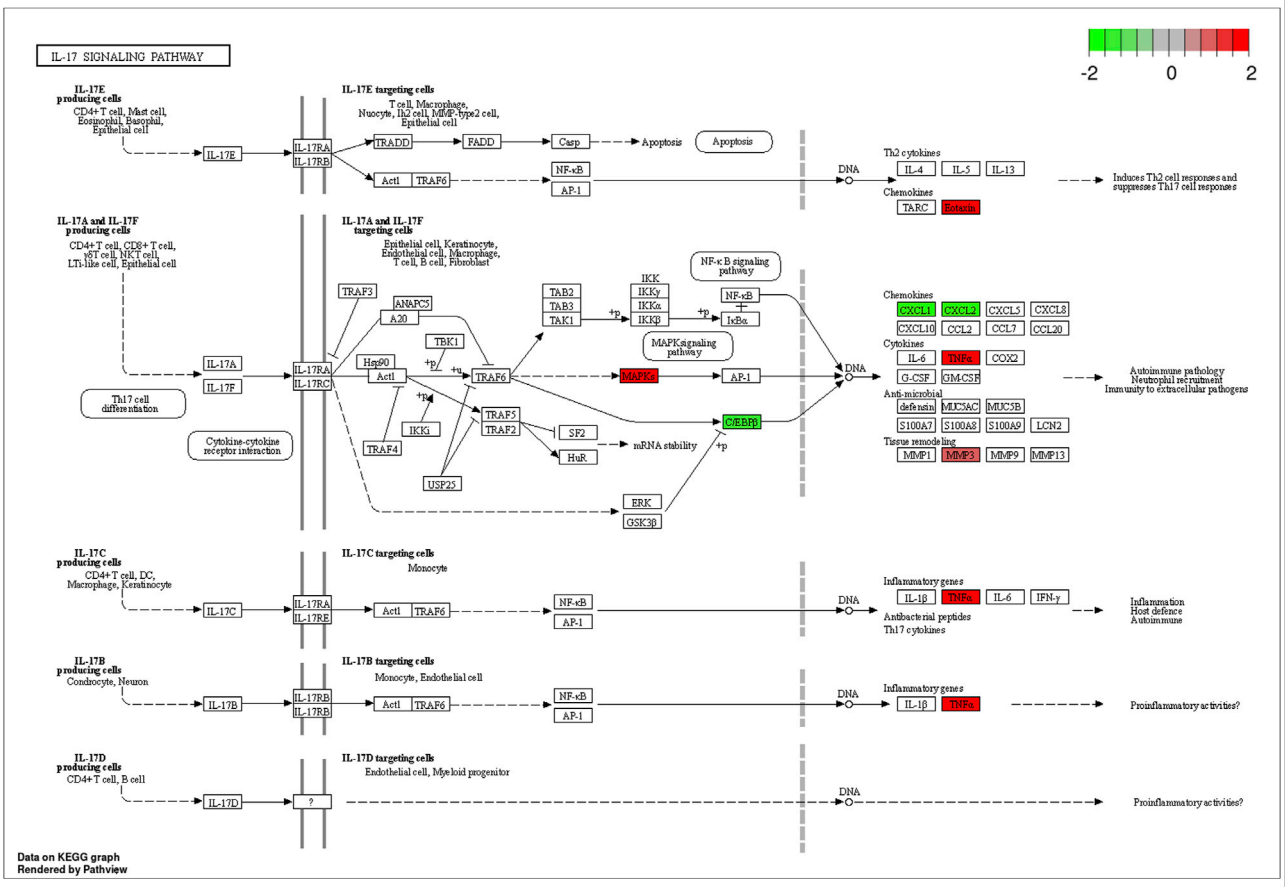


FIGURE 5
Differentially expressed genes enriched in the path diagram of the IL-17 signaling pathway.

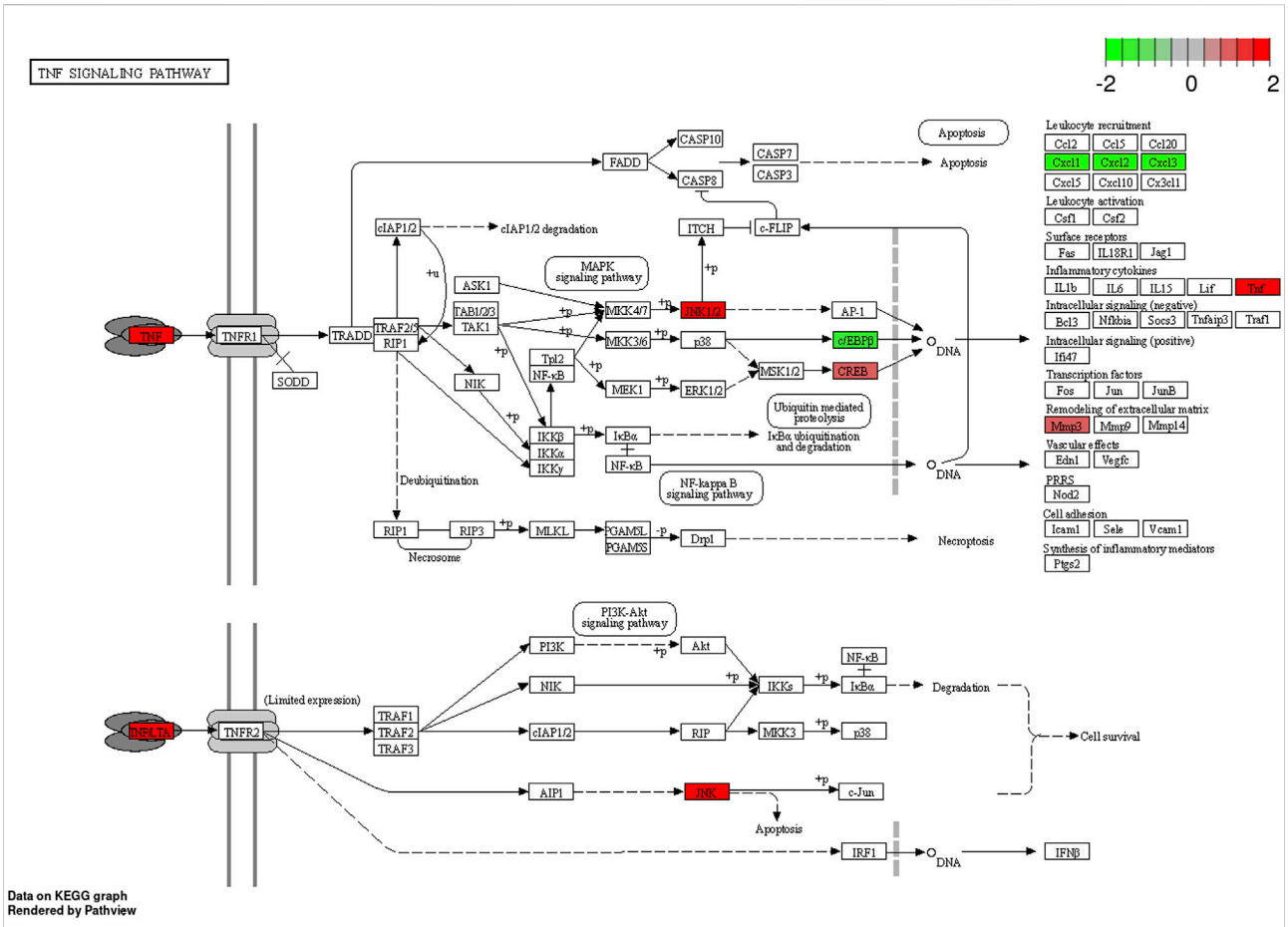


FIGURE 6
Differentially expressed genes enriched in the path diagram of the TNF signaling pathway.

TABLE 1 Differentially expressed genes enriched in the IL-17 and TNF and signaling pathways.

Gene symbol	Gene name	log_2 fold change	*p*-value	Type
Mmp3	Matrix metallopeptidase 3	1.10	5.38×10^{-5}	Up
Tnf	Tumor necrosis factor	2.03	2.08×10^{-2}	Up
Mapk10	Mitogen-activated protein kinase 10	2.23	1.21×10^{-2}	Up
Ccl11	C-C motif chemokine ligand 11	2.26	1.34×10^{-2}	Up
Creb5	cAMP-responsive element-binding protein 5	1.09	5.28×10^{-13}	Up
Cxcl2	C-X-C motif chemokine ligand 2	−2.77	3.34×10^{-4}	Down
Cebpb	CCAAT/enhancer-binding protein-beta	−1.26	6.55×10^{-3}	Down

T-2 toxin-related genes from GeneCards

The T-2 toxin-related genes were obtained from the GeneCards (https://www.genecards.org/) database. The relevance scores of >5 were used as the screening criteria (Safran et al., 2022). KEGG pathways were plotted using a free online platform for data analysis and visualization (https://www.bioinformatics.com.cn).

Immunohistochemical (IHC) staining

The dewaxed and rehydrated tissue sections were placed in a repair box filled with citric acid (pH 6.0) antigen repair buffer. Next, 3% hydrogen peroxide was used to inhibit the endogenous peroxidase activity. After blocking with 3% bovine serum albumin for 30 min at room temperature, the slices were incubated with primary anti-Mmp and anti-Tnf antibodies (1:500 dilution, Wuhan Servicebio

TABLE 2 T-2 toxin-related genes from GeneCards based on the score more than 5.

Gene symbol	Gene name	GIFtS	GC ID	Score
TNF	Tumor necrosis factor	54	GC06P087731	11.21
IL1B	Interleukin 1 beta	50	GC02M112829	10.09
ALB	Albumin	52	GC04P073397	8.36
IL2	Interleukin 2	51	GC04M122451	7.99
FURIN	Furin, paired basic amino acid cleaving enzyme	50	GC15P090868	7.68
IL6	Interleukin 6	53	GC07P022725	7.56
CXCL8	C-X-C motif chemokine ligand 8	44	GC04P073740	7.29
PARP1	Poly (ADP-ribose) polymerase 1	53	GC01M226360	6.45
BCL2	BCL2 apoptosis regulator	53	GC18M063123	6.15
EGF	Epidermal growth factor	54	GC04P109912	5.82
MAPK1	Mitogen-activated protein kinase 1	54	GC22M021759	5.67
CASP3	Caspase 3	52	GC04M184627	5.48
CASP9	Caspase 9	49	GC01M015491	5.43

Abbreviation: GIFtS, GeneCards Inferred Functionality Score; GC ID, GeneCards identifiers.

Biotechnology) overnight at 4°C and then incubated with horseradish peroxidase–conjugated secondary antibodies for 50 min at room temperature. Next, the slices were washed with phosphate buffered saline, developed, and sealed after re-staining the nuclei with hematoxylin. Semi-quantitative analysis of the acquired images was performed using Image Pro Plus 6.0.

Results

Influence of cartilage tissues treated with T-2 toxin

The H&E staining results show the histopathologic changes of rat cartilage tissues in each group (Figure 1). In the control group, chondrocytes were abundant and arranged in an orderly manner from top to bottom, and the morphology of the cells and nuclei was normal. Abnormalities were not observed in the control group. In the T-2 toxin group, the number of cells decreased significantly, and the arrangement of chondrocytes became disordered. The hypertrophic chondrocytes were clustered, and the physalide phenomenon caused by the loss of the nucleus also appeared.

Identification of differentially expressed mRNA

Analyses of differentially expressed mRNAs showed that 284 significant differentially expressed mRNAs were identified in the T-2 group compared to the control group. Of these, 195 mRNAs were significantly upregulated, and 89 mRNAs were downregulated (Figure 2A). Cluster analysis was performed to identify differentially expressed mRNAs between the T-2 and control groups. In Figure 2B, each column represents a different sample, and each row represents the expression levels of mRNA in different colors: up (red) or down (green).

GO analysis

The GO results were divided into three categories (namely, biological process, cellular component, and molecular function). As shown in Figure 3, the top immune-related biological processes were the regulation of hormone secretion, regulation of peptide hormone secretion, regulation of transcription involved in cell fate commitment, positive regulation of the small molecule metabolic process, and response to activity. The top immune-related cellular components were the ATPase-dependent transmembrane transport complex, an anchored component of the plasma membrane, the cation-transporting ATPase complex, the recycling endosome, and the apical part of the cell. The top immune-related molecular functions were receptor–ligand activity, signaling receptor activator activity, receptor regulator activity, sodium channel regulator activity, iron ion binding, and cytokine activity.

KEGG pathway analysis

The KEGG pathway enrichment analysis identified significant pathways. According to the enrichment score, the top 10 significant pathways were the adipocytokine signaling pathway, the PPAR signaling pathway, the IL-17 signaling pathway, alcoholic liver disease, protein digestion and absorption, proximal tubule bicarbonate reclamation, insulin resistance, the TNF signaling pathway, type II diabetes mellitus, and cholesterol metabolism (Figure 4A). Figure 4B shows the relationship between significant pathways and enriched genes, and different colors represent different pathways. The path diagrams of

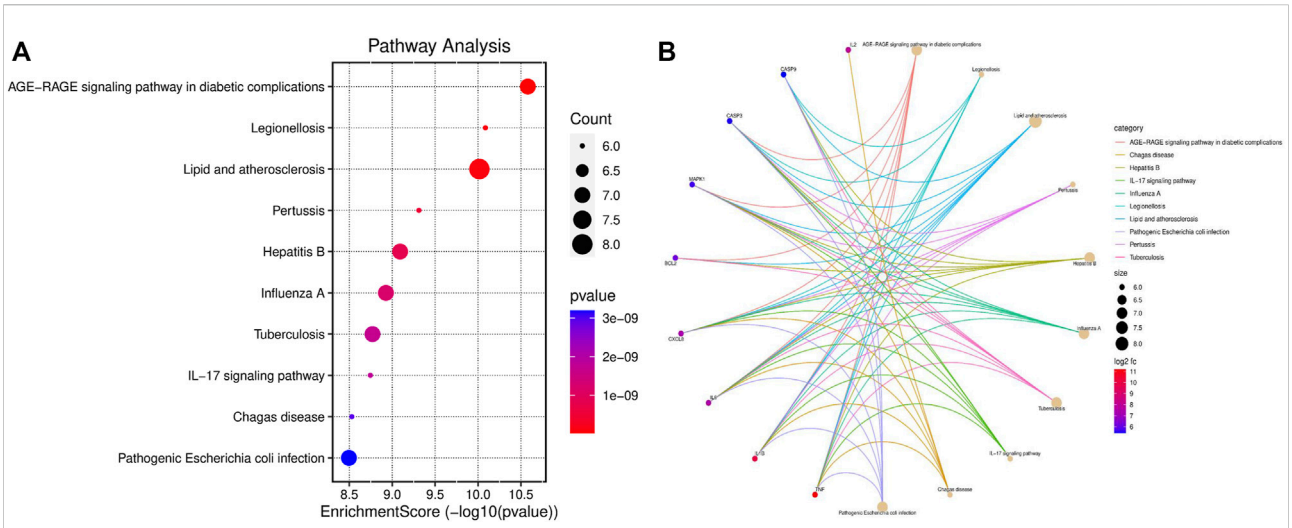


FIGURE 7
KEGG pathway enrichment analysis (top 10). **(A)** Top 10 significant KEGG pathways of T-2 toxin-related genes. **(B)** Relationship between significant pathways and enriched genes, and different colors represent different pathways. *p*-value <0.05 was considered significant pathways.

the IL-17 signaling pathway (Figure 5) and the TNF signaling pathway (Figure 6) were significantly correlated with the damaged articular cartilage exposed to T-2 toxin. Seven identified genes were enriched in the IL-17 and TNF signaling pathways (Table 1).

T-2 toxin-related gene enrichment analysis

In the GeneCards database, the 13 identified T-2 toxin-related genes (Table 2) were screened with a score greater than 5. KEGG pathways were enriched using an online tool (https://www.bioinformatics.com.cn). The IL-17 signaling pathway was significantly correlated with T-2 toxin (Figure 7A). The enriched genes (TNF, IL1B, IL2, IL6, and CXCL8) are presented in Figure 7B.

IHC staining

The expression levels of Mmp3 and Tnf proteins were determined by IHC staining. As shown in Figure 8, the positive expression of Mmp3 and Tnf proteins in the T-2 toxin group was significantly higher than that in the control group, and the expression levels of Mmp3 and Tnf proteins were increased in the T-2 toxin group ($p < 0.05$).

Discussion

It is well known that T-2 toxin is harmful to human health. T-2 toxins can cause cytotoxicity, genotoxicity, immunotoxicity, hepatotoxicity, gastrointestinal toxicity, skeletal toxicity, nephrotoxicity, reproductive toxicity, and neurotoxicity. It damages the immune organs, such as the liver, digestive tract, bones, and kidneys, leading to pathological changes and impaired physiological functions of these organs (Li S. J. et al., 2022). Furthermore, the T-2 toxin has been reported to induce oxidative stress and inflammation, thereby causing organ toxicity. Studies have confirmed that T-2 toxin can increase the excessive production of nitric oxide by increasing various inflammatory cytokines, thus leading to mitochondrial damage and redox imbalance, further inducing apoptosis (Liu et al., 2017). Although T-2 toxin is an environmental risk factor for KBD, the potential mechanism by which T-2 toxin damages articular cartilage remains unknown. In the current study, a series of experiments was carried out to clarify the potential mechanism of damaged cartilage induced by T-2 toxin.

Our results demonstrated T-2 toxin-induced pathological changes in the articular cartilage of rats. The KEGG pathway enrichment results showed that IL-17 and TNF signaling pathways were related to inflammation. In addition, T-2 toxin-related genes obtained from GeneCards were used for enrichment analysis, and it was found that the IL-17 signaling pathway was also significantly expressed. These results indicated that the inflammatory reaction may play an important role in the damaged joint cartilage of rats induced by T-2 toxin. In addition, Mmp3, Tnf, Mapk10, Ccl11, Creb5, Cxcl2, and Cebpb were significantly enriched in the IL-17 and TNF signaling pathways.

Interleukin-17 (IL-17) is a member of a new family of inflammatory cytokines (Amatya et al., 2017). IL-17 is associated with many inflammatory diseases, such as rheumatoid arthritis, asthma, lupus, and allograft rejection (Aggarwal & Gurney, 2002; Moseley et al., 2003; Kolls &

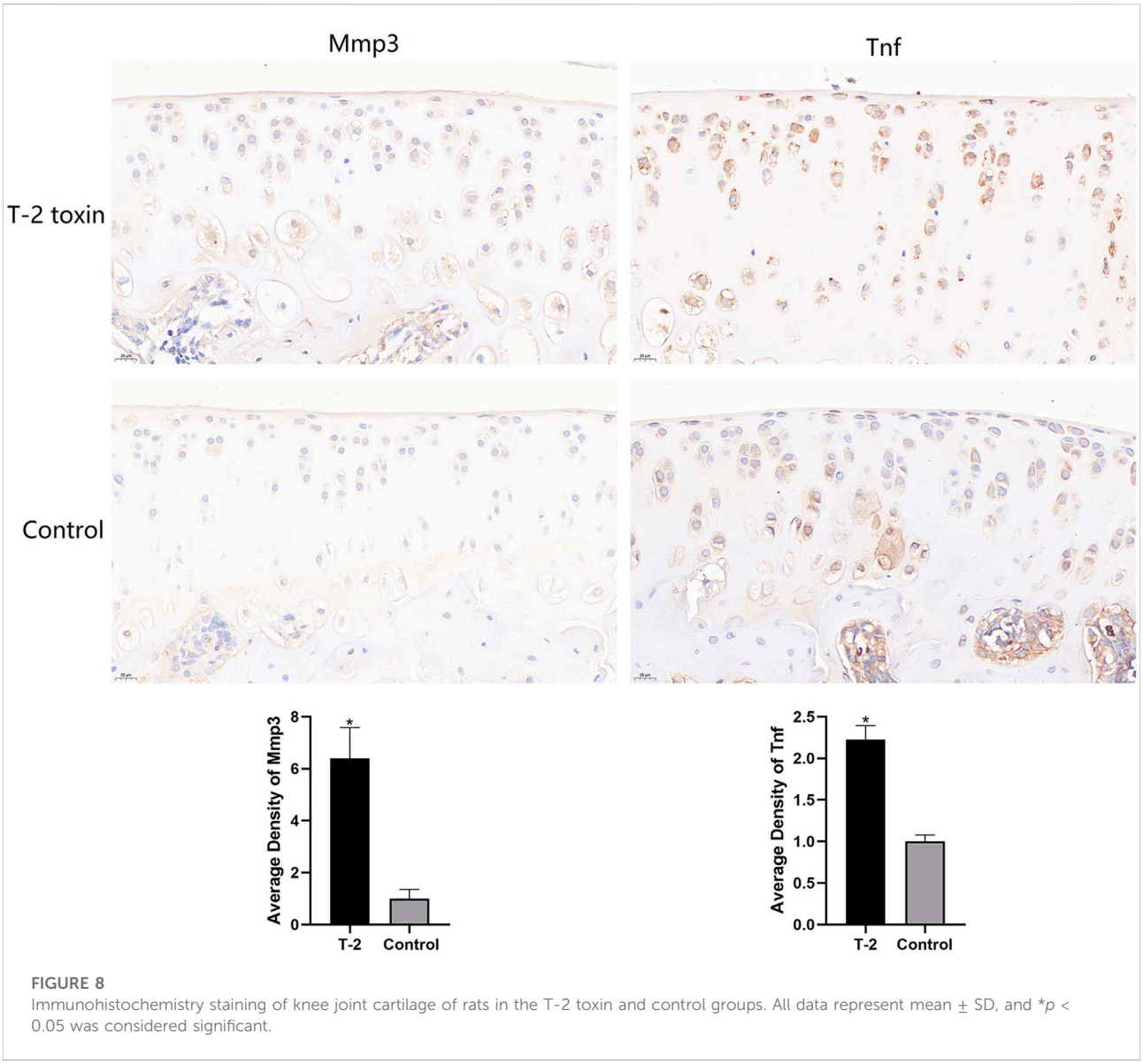


FIGURE 8
Immunohistochemistry staining of knee joint cartilage of rats in the T-2 toxin and control groups. All data represent mean ± SD, and *p < 0.05 was considered significant.

Lindén, 2004). IL-17 receptors are widely distributed in various tissues and participate in the activation of transcription factor NF-κB and the kinase JNK pathway, which are related to inflammatory and immune diseases (Schwandner et al., 2000; Moseley et al., 2003). Tumor necrosis factor (TNF) is mainly produced by activated macrophages and T lymphocytes, and proto-TNF is expressed on the plasma membrane, where it can be cleaved by matrix metalloproteinases (MMPs) in the extracellular region, thus releasing its soluble form (Bradley, 2008). It is well known that pro-inflammatory cytokines including IL-1b, TNF, and IL-6 promote cartilage degradation by stimulating the production of MMPs (Zhou et al., 2014). A proteinase activation cascade is likely initiated by plasmin, the product of plasminogen activator activity, which in turn activates latent stromelysin (MMP3), an activator of latent collagenases (Goldring, 2000). Activation of the NLRP3 inflammasome leads to a wide range of immune responses, including the production of pro-inflammatory cytokines and chemokines and cell death (Wree et al., 2013; Broderick et al., 2015). Studies have shown the liver inflammation and fibrosis induced by the NLRP3 inflammasome were affected by IL-17 and TNF, and the effect of TNF was more significant (Wree et al., 2018). Therefore, some inflammasomes in the articular cartilage may also be affected by IL-17 and TNF, which leads to the enhancement of the inflammatory response. MAPK10 is a member of the Jun N-terminal kinase subgroup in the mitogen-activated protein kinase superfamily and is involved in multiple signaling pathways in key physiological processes, such as apoptosis, differentiation, and proliferation (Davis, 2000; Yan et al., 2018). Some studies have shown that MAPK10 plays a

central role in atrial fibrillation in mice, and MAPK10 knockout can significantly reduce atrial inflammation (Liu et al., 2022). The chemokine C-C motif ligand 11 (CCL11) is a member of the CC chemokine family and is produced by connective tissue cells and leukocytes in the human body (Wakabayashi et al., 2021). CCL11 can induce the migration of many types of leukocytes, including eosinophils, basophils, macrophages, and dendritic cells (Elsner et al., 1996; Peled et al., 1998; Menzies-Gow et al., 2002). The chemokine CXCL2 is induced by macrophage toll-like receptor signal transduction and is responsible for the recruitment of neutrophils, which is the most important step in response to pathogens (De Filippo et al., 2008). Our transcriptome sequencing results showed that the expression of Cxcl2 was downregulated. Therefore, it can be speculated that T-2 toxin may reduce the ability of rats to resist pathogens by inhibiting the expression of Cxcl2.

In addition, the IHC analysis results showed that the expression levels of Mmp3 and Tnf proteins in the T-2 toxin group were significantly higher than those in the control group. This indicated that T-2 toxin can activate the expression of Mmp3 and Tnf, which is consistent with the RNA-seq results. Generally, TNF first binds to its receptors and then transmits molecular signals for biological functions such as inflammation and cell death. It has also been shown that T-2 toxin can increase the level of TNF (Lin et al., 2019). TNF can activate the immune system, but inappropriate or excessive production of TNF may be harmful and lead to rheumatoid arthritis, inflammatory bowel disease, psoriasis arthritis, psoriasis, and non-infectious uveitis (Jang et al., 2021). TNF inhibitors have been successfully developed and applied in the clinical treatment of autoimmune diseases such as Crohn's disease and rheumatoid arthritis (Raychaudhuri & Raychaudhuri, 2009; Lis et al., 2014). MMP3 is a protease that is synthesized and secreted by synovial fibroblasts and chondrocytes, and it can be used as a reliable marker of rheumatoid arthritis (Lerner et al., 2018). The MMP3 enzyme can degrade a variety of collagen, proteoglycans, fibronectin, laminin, and elastin. In addition, it can activate other MMPs, such as MMP-1, MMP-7, and MMP-9. Therefore, MMP3 plays a key role in connective tissue remodeling (Lerner et al., 2018).

In this study, RNA-seq was used to explore the mechanism of damaged cartilage induced by T-2 toxin, which may enhance the inflammatory response by activating the IL-17 and TNF signaling pathways, resulting in an abnormal inflammatory response in chondrocytes and cartilage injury. However, several limitations need to be addressed. First, an animal model was used to study the inflammatory response involved in damaged articular cartilage, and this needs to be validated in the cartilage of KBD patients. Second, further molecular experiments should be performed to validate the IL-17 and TNF signaling pathways. Despite these limitations, our results provided new insight for further understanding the mechanism of damaged articular cartilage induced by T-2 toxin.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

Ethics statement

All animal studies were reviewed and approved by the Animal Ethics Research Committee of Zhengzhou University (ZZUIRB 2021-69).

Author contributions

LS, JW, and YF contributed to design the study. LS, QL, HY, and QW conducted the experiments. LS and QL drafted the manuscript and visualized the results. HY and QW conducted bioinformatics analysis. JW and YF critically reviewed the manuscript. All authors contributed to the manuscript revision and approved the submitted version.

Funding

This work was funded by the Foreign Experts Program of Henan Province (HNGD2021032).

Conflict of interest

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

Publisher's note

Supplementary Material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1079739/full#supplementary-material

References

Aggarwal, S., and Gurney, A. L. (2002). IL-17: Prototype member of an emerging cytokine family. *J. Leukoc. Biol.* 71 (1), 1–8. doi:10.1189/jlb.71.1.1

Amatya, N., Garg, A. V., and Gaffen, S. L. (2017). IL-17 signaling: The yin and the Yang. *Trends Immunol.* 38 (5), 310–322. doi:10.1016/j.it.2017.01.006

Bradley, J. R. (2008). TNF-mediated inflammatory disease. *J. Pathol.* 214 (2), 149–160. doi:10.1002/path.2287

Broderick, L., De Nardo, D., Franklin, B. S., Hoffman, H. M., and Latz, E. (2015). The inflammasomes and autoinflammatory syndromes. *Annu. Rev. Pathol.* 10, 395–424. doi:10.1146/annurev-pathol-012414-040431

Davis, R. J. (2000). Signal transduction by the JNK group of MAP kinases. *Cell* 103 (2), 239–252. doi:10.1016/s0092-8674(00)00116-1

De Filippo, K., Henderson, R. B., Laschinger, M., and Hogg, N. (2008). Neutrophil chemokines KC and macrophage-inflammatory protein-2 are newly synthesized by tissue macrophages using distinct TLR signaling pathways. *J. Immunol.* 180 (6), 4308–4315. doi:10.4049/jimmunol.180.6.4308

Elsner, J., Höchstetter, R., Kimmig, D., and Kapp, A. (1996). Human eotaxin represents a potent activator of the respiratory burst of human eosinophils. *Eur. J. Immunol.* 26 (8), 1919–1925. doi:10.1002/eji.1830260837

Goldring, M. B. (2000). Osteoarthritis and cartilage: The role of cytokines. *Curr. Rheumatol. Rep.* 2 (6), 459–465. doi:10.1007/s11926-000-0021-y

Guo, X., Ma, W. J., Zhang, F., Ren, F. L., Qu, C. J., and Lammi, M. J. (2014). Recent advances in the research of an endemic osteochondropathy in China: Kashin-Beck disease. *Osteoarthr. Cartil.* 22 (11), 1774–1783. doi:10.1016/j.joca.2014.07.023

Jang, D. I., Lee, A. H., Shin, H. Y., Song, H. R., Park, J. H., Kang, T. B., et al. (2021). The role of tumor necrosis factor Alpha (TNF-α) in autoimmune disease and current TNF-α inhibitors in therapeutics. *Int. J. Mol. Sci.* 22 (5), 2719. doi:10.3390/ijms22052719

Kolls, J. K., and Lindén, A. (2004). Interleukin-17 family members and inflammation. *Immunity* 21 (4), 467–476. doi:10.1016/j.immuni.2004.08.018

Lerner, A., Neidhöfer, S., Reuter, S., and Matthias, T. (2018). MMP3 is a reliable marker for disease activity, radiological monitoring, disease outcome predictability, and therapeutic response in rheumatoid arthritis. *Best. Pract. Res. Clin. Rheumatol.* 32 (4), 550–562. doi:10.1016/j.berh.2019.01.006

Li, D., Han, J., Guo, X., Qu, C., Yu, F., and Wu, X. (2016). The effects of T-2 toxin on the prevalence and development of kashin-beck disease in China: A meta-analysis and systematic review. *Toxicol. Res.* 5 (3), 731–751. doi:10.1039/c5tx00377f

Li, H. N., Jin, B. M., Hua, Z., Le-Le, L., Meng-Yuan, L., Xiu-Juan, Z., et al. (2022a). YAP plays a protective role in T-2 toxin-induced inhibition of chondrocyte proliferation and matrix degradation. *Toxicon.* 215, 49–56. doi:10.1016/j.toxicon.2022.06.005

Li, S. J., Zhang, G., Xue, B., Ding, Q., Han, L., Huang, J. C., et al. (2022b). Toxicity and detoxification of T-2 toxin in poultry. *Food Chem. Toxicol.* 169, 113392. doi:10.1016/j.fct.2022.113392

Lin, R., Sun, Y., Ye, W., Zheng, T., Wen, J., and Deng, Y. (2019). T-2 toxin inhibits the production of mucin via activating the IRE1/XBP1 pathway. *Toxicology* 424, 152230. doi:10.1016/j.tox.2019.06.001

Lis, K., Kuzawińska, O., and Bałkowiec-Iskra, E. (2014). Tumor necrosis factor inhibitors - state of knowledge. *Arch. Med. Sci.* 10 (6), 1175–1185. doi:10.5114/aoms.2014.47827

Liu, F., Deng, Y., Zhao, Y., Li, Z., Gao, J., Zhang, Y., et al. (2022). Time series RNA-seq analysis identifies MAPK10 as a critical gene in diabetes mellitus-induced atrial fibrillation in mice. *J. Mol. Cell. Cardiol.* 168, 70–82. doi:10.1016/j.yjmcc.2022.04.013

Liu, X., Guo, P., Liu, A., Wu, Q., Xue, X., Dai, M., et al. (2017). Nitric oxide (NO)-mediated mitochondrial damage plays a critical role in T-2 toxin-induced apoptosis and growth hormone deficiency in rat anterior pituitary GH3 cells. *Food Chem. Toxicol.* 102, 11–23. doi:10.1016/j.fct.2017.01.017

Liu, Y. N., Mu, Y. D., Wang, H., Zhang, M., Shi, Y. W., Mi, G., et al. (2021). Endoplasmic reticulum stress pathway mediates T-2 toxin-induced chondrocyte apoptosis. *Toxicology* 464, 152989. doi:10.1016/j.tox.2021.152989

Luo, W., and Brouwer, C. (2013). Pathview: An R/bioconductor package for pathway-based data integration and visualization. *Bioinformatics* 29 (14), 1830–1831. doi:10.1093/bioinformatics/btt285

Lutsky, I., Mor, N., Yagen, B., and Joffe, A. Z. (1978). The role of T-2 toxin in experimental alimentary toxic aleukia: A toxicity study in cats. *Toxicol. Appl. Pharmacol.* 43 (1), 111–124. doi:10.1016/s0041-008x(78)80036-2

Menzies-Gow, A., Ying, S., Sabroe, I., Stubbs, V. L., Soler, D., Williams, T. J., et al. (2002). Eotaxin (CCL11) and eotaxin-2 (CCL24) induce recruitment of eosinophils, basophils, neutrophils, and macrophages as well as features of early- and late-phase allergic reactions following cutaneous injection in human atopic and nonatopic volunteers. *J. Immunol.* 169 (5), 2712–2718. doi:10.4049/jimmunol.169.5.2712

Moseley, T. A., Haudenschild, D. R., Rose, L., and Reddi, A. H. (2003). Interleukin-17 family and IL-17 receptors. *Cytokine Growth Factor Rev.* 14 (2), 155–174. doi:10.1016/s1359-6101(03)00002-9

Peled, A., Gonzalo, J. A., Lloyd, C., and Gutierrez-Ramos, J. C. (1998). The chemotactic cytokine eotaxin acts as a granulocyte-macrophage colony-stimulating factor during lung inflammation. *Blood* 91 (6), 1909–1916. doi:10.1182/blood.v91.6.1909.1909_1909_1916

Raychaudhuri, S. P., and Raychaudhuri, S. K. (2009). Biologics: Target-specific treatment of systemic and cutaneous autoimmune diseases. *Indian J. dermatol.* 54 (2), 100–109. doi:10.4103/0019-5154.53175

Safran, M. R. N., Twik, M., BarShir, R., Iny Stein, T., Dahary, D., Fishilevich, S., et al. (2022). The GeneCards suite chapter. *Pract. Guide Life Sci. Databases*, 27–56.

Schwandner, R., Yamaguchi, K., and Cao, Z. (2000). Requirement of tumor necrosis factor receptor-associated factor (TRAF)6 in interleukin 17 signal transduction. *J. Exp. Med.* 191 (7), 1233–1240. doi:10.1084/jem.191.7.1233

Sokoloff, L. (1988). Kashin-beck disease: Current status. *Nutr. Rev.* 46 (3), 113–119. doi:10.1111/j.1753-4887.1988.tb05395.x

Wakabayashi, K., Isozaki, T., Tsubokura, Y., Fukuse, S., and Kasama, T. (2021). Eotaxin-1/CCL11 is involved in cell migration in rheumatoid arthritis. *Sci. Rep.* 11 (1), 7937. doi:10.1038/s41598-021-87199-7

Wang, K., Yu, J., Liu, H., Liu, Y., Liu, N., Cao, Y., et al. (2020). Endemic kashin-beck disease: A food-sourced osteoarthropathy. *Semin. Arthritis Rheum.* 50 (2), 366–372. doi:10.1016/j.semarthrit.2019.07.014

Wang, L. H., Fu, Y., Shi, Y. X., and Wang, W. G. (2011). T-2 toxin induces degenerative articular changes in rodents: Link to kaschin-beck disease. *Toxicol. Pathol.* 39 (3), 502–507. doi:10.1177/0192623310396902

Wree, A., Broderick, L., Canbay, A., Hoffman, H. M., and Feldstein, A. E. (2013). From NAFLD to NASH to cirrhosis-new insights into disease mechanisms. *Nat. Rev. Gastroenterol. Hepatol.* 10 (11), 627–636. doi:10.1038/nrgastro.2013.149

Wree, A., McGeough, M. D., Inzaugarat, M. E., Eguchi, A., Schuster, S., Johnson, C. D., et al. (2018). NLRP3 inflammasome driven liver injury and fibrosis: Roles of IL-17 and TNF in mice. *Hepatology* 67 (2), 736–749. doi:10.1002/hep.29523

Yan, J., Thomson, J. K., Zhao, W., Gao, X., Huang, F., Chen, B., et al. (2018). Role of stress kinase JNK in binge alcohol-evoked atrial arrhythmia. *J. Am. Coll. Cardiol.* 71 (13), 1459–1470. doi:10.1016/j.jacc.2018.01.060

Yang, J. B., Sun, D. J., and Wang, Z. W. (1995). Detection report of T-2 toxin in staple food of patients in Kashin Beck disease area. *Chin. J. Endemic Dis.* (03), 146–149.

Yang, X., Xiao, X., Zhang, L., Wang, B., Li, P., Cheng, B., et al. (2022). An integrative analysis of DNA methylation and transcriptome showed the dysfunction of MAPK pathway was involved in the damage of human chondrocyte induced by T-2 toxin. *BMC Mol. Cell Biol.* 23 (1), 4. doi:10.1186/s12860-021-00404-3

Yu, F. F., Zuo, J., Sun, L., Yu, S. Y., Lei, X. L., Zhu, J. H., et al. (2022). Animal models of Kashin-Beck disease exposed to environmental risk factors: Methods and comparisons. *Ecotoxicol. Environ. Saf.* 234, 113419. doi:10.1016/j.ecoenv.2022.113419

Zhou, X., Wang, Z., Chen, J., Wang, W., Song, D., Li, S., et al. (2014). Increased levels of IL-6, IL-1β, and TNF-α in Kashin-Beck disease and rats induced by T-2 toxin and selenium deficiency. *Rheumatol. Int.* 34 (7), 995–1004. doi:10.1007/s00296-013-2862-5

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 24 February 2023
DOI 10.3389/fgene.2023.1130411

OPEN ACCESS

EDITED BY
Qiling Yuan,
Xi'an Jiaotong University, China

REVIEWED BY
Francisco J. Blanco Garcia,
University of A Coruña, Spain
Hanlin Zhang,
College of Letters and Science, University of California, Berkeley, United States
Korosh Morshedi,
Kashan University of Medical Sciences, Iran

*CORRESPONDENCE
Min Yu,
mycdc1234@163.com
Yingjun Li,
2016034036@hmc.edu.cn

SPECIALTY SECTION
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

RECEIVED 23 December 2022
ACCEPTED 13 February 2023
PUBLISHED 24 February 2023

CITATION
Liu H, Cai B, Gong R, Yang Y, Wang J, Zhou D, Yu M and Li Y (2023), Impact of genetically predicted characterization of mitochondrial DNA quantity and quality on osteoarthritis.
Front. Genet. 14:1130411.
doi: 10.3389/fgene.2023.1130411

Impact of genetically predicted characterization of mitochondrial DNA quantity and quality on osteoarthritis

Houpu Liu[1], Bingyue Cai[1], Ruicheng Gong[2], Ye Yang[1], Jing Wang[1], Dan Zhou[3,4], Min Yu[5]* and Yingjun Li[1]*

[1]Department of Epidemiology and Health Statistics, School of Public Health, Hangzhou Medical College, Hangzhou, China, [2]Starbody Clinic, Hangzhou, China, [3]Department of Big Data in Health Science, School of Public Health, Zhejiang University School of Medicine, Hangzhou, China, [4]Vanderbit Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, United States, [5]Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou, China

Background: Existing studies have indicated that mitochondrial dysfunction may contribute to osteoarthritis (OA) development. However, the causal association between mitochondrial DNA (mtDNA) characterization and OA has not been extensively explored.

Methods: Two-sample Mendelian randomization was performed to calculate the impact of mitochondrial genomic variations on overall OA as well as site-specific OA, with multiple analytical methods inverse variance weighted (IVW), weighted median (WM), MR-Egger and MR-robust adjusted profile score (MR-RAPS).

Results: Genetically determined mitochondrial heteroplasmy (MtHz) and mtDNA abundance were not causally associated with overall OA. In site-specific OA analyses, the causal effect of mtDNA abundance on other OA sites, including hip, knee, thumb, hand, and finger, had not been discovered. There was a suggestively protective effect of MtHz on knee OA IVW OR = 0.632, 95% CI: 0.425–0.939, *p*-value = 0.023. No causal association between MtHz and other different OA phenotypes was found.

Conclusion: MtHz shows potential to be a novel therapeutic target and biomarker on knee OA development. However, the variation of mtDNA abundance was measured from leukocyte in blood and the levels of MtHz were from saliva samples rather than cartilage or synovial tissues. Genotyping samples from synovial and cartilage can be a focus to further exploration.

KEYWORDS

mitochondrial heteroplasmy, mitochondrial abundance, osteoarthritis, mendelian randomization, knee OA (KOA)

1 Introduction

Osteoarthritis (OA) is a common joint disease with degenerative changes in articular cartilage and remains major impact on public health worldwide (Nelson, 2018). According to the estimation of the Global Burden of Disease Study (GBD), nearly 300 million individuals throughout the world are suffering from OA, and the number is still increasing. OA is one of the leading causes of pain, disability, and great socioeconomic burden in developed countries (Glyn-Jones et al., 2015). However, the mechanism of OA remains unclear. Identifying risk

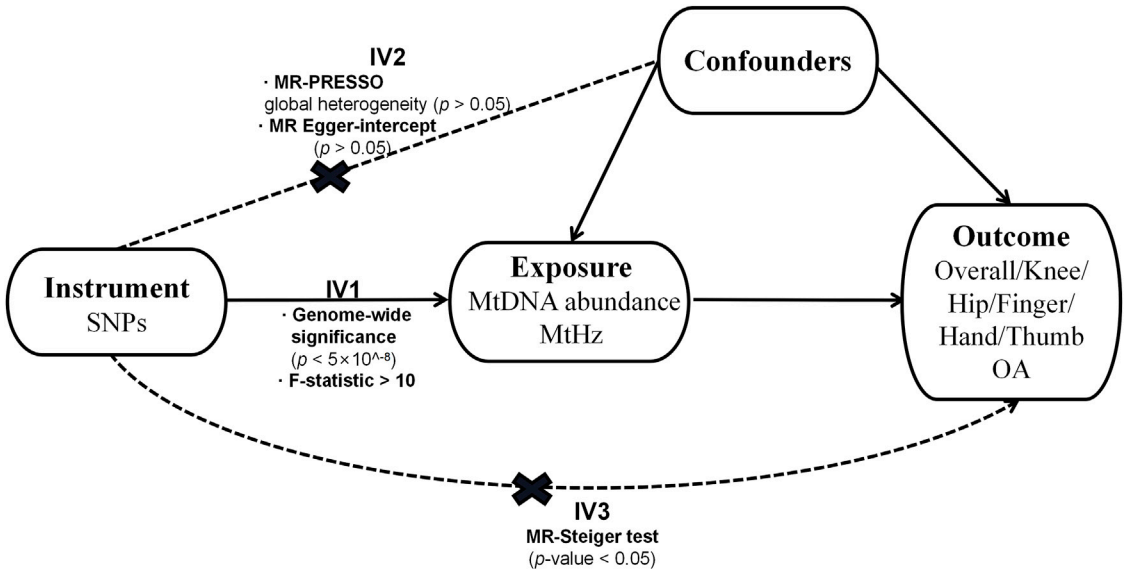


FIGURE 1
The assumptions of MR and how we tested these assumptions in our analyses.

factors such as a disorder-related biomarker is essential to understand OA pathogenesis, decrease the incidence of the disease and develop efficiency prevention strategies.

Mitochondria are complex multifunctional organelles involved in various cell functions, e.g., heat regulation, calcium homeostasis, reactive oxygen species (ROS) production, and proinflammatory cytokines production, the damage of which may affect chondrocyte health at an extent (Blanco et al., 2011; Fernandez-Moreno et al., 2017; Blanco et al., 2018). They have their own genome (mitochondrial DNA, mtDNA) which contain 37 genes and encode 13 proteins, 22 transfer RNAs (tRNAs) and two ribosomal RNAs (rRNAs) (Anderson et al., 1981). As multicopy genome, sequence mutation and copy number variation of mtDNA are prevalent in population (McArdle et al., 2013; McGuire, 2019). As one of the mtDNA quantity characteristics, mtDNA abundance is recognized as a rough estimation for the number of mitochondria (Castellani et al., 2020). Previous studies provided supportive evidence for extensive pathogenicity of mtDNA abundance (Frahm et al., 2005; Hu et al., 2016). As one of the mtDNA quality characteristics, mitochondrial herteroplasmy (MtHz), where mtDNA with distinct sequences coexist, has been found in a large spectrum of human disease, including classical mitochondrial diseases and complex disorders (Ye et al., 2014). MtDNA abundance and MtHz can integrate many aspects of mitochondrial function and serve as promising biomarkers in probing interactions between mitochondria and disorders (Castellani et al., 2020; Tian et al., 2021).

In a recent review, OA was recognized as a potential mitochondrial disease considering the impact of mitochondrial dysfunction on cartilage degradation (Fernandez-Moreno et al., 2022). However, the causality between mitochondria and OA has rarely been investigated. Therefore, the goal of our study was to probe causal effects of characterization of mtDNA quantity and quality on OA development which is crucial to understand the role of mitochondria in OA etiology. A two-sample Mendelian randomization (MR) analysis was performed to investigate potential causal associations between mtDNA abundance and MtHz with OA using summary statistics from large-scale genome-wide association study (GWAS). MR-Steiger test was applied to ascertain whether variation of mtDNA characterization is a cause or a consequence of OA development.

2 Materials and methods

2.1 Study design overview

The design of our research is displayed in Figure 1. We adopted a two-sample MR design to compute the causal effect of the characterization of mtDNA quantity and quality on overall OA and site-specific OA separately (Boer et al., 2021). Genetic association estimates for mtDNA abundance were derived from the United Kingdom biobank study (N = 291,950) (Hagg et al., 2021). Single Nucleotide Polymorphisms (SNPs) associated with MtHz were obtained from the 23andME research program (N = 982,072), a personal genomics and biotechnology company (Nandakumar et al., 2021). The participants included in both GWASs were of European ancestry.

2.2 Selection of genetic instruments

SNPs serving as instrumental variables (IVs) for MtHz and mtDNA abundance, all reach genome-wide significance ($p < 5 \times$ 10–8) (Hagg et al., 2021; Nandakumar et al., 2021). Twenty SNPs were related to MtHz and accounted for 32% observed SNP-heritability, 64 SNPs were included to estimate the genetic liability of mtDNA abundance and explained approximately 8.3% SNP-heritability. We performed linkage disequilibrium (LD)

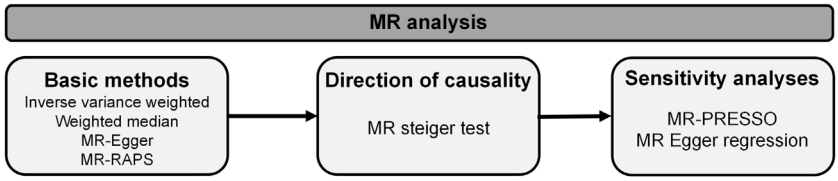


FIGURE 2
Flowchart of Mendelian randomization framework in this study.

pruning in PLINK with 1,000 Genomes Europeans as the reference panel to ascertain whether these genetic variants are independent of each other (r^2 < 0.001) (Burgess et al., 2017). After LD test, five genetic variants linked to MtHz were removed, and 17 IVs related to mtDNA abundance were excluded. These SNPs were then matched in summary GWASs of OA in subsequent analysis and those not available in the outcome GWAS were removed or replaced by the proxy SNPs. F-statistic was calculated to filter the weak instruments. The threshold of F-statistic that was sufficient for identifying causal effect was 10 (Burgess et al., 2011). The selected genetic variants are demonstrated in Supplementary Table S1, S2.

2.3 Osteoarthritis and data sources

Summary statistics data on overall OA and its specific sites, comprising knee, hip, spine, thumb, and hand OA, were obtained from the latest publicly available GWAS of OA. This GWAS covered nine populations with up to 826,690 participants (177,517 OA patients) (Boer et al., 2021). The definition of OA satisfied the criteria of Genetics of Osteoarthritis (GO) including self-reported status, hospital diagnosed, ICD10 codes or radiographic as defined by the TREAT-OA consortium (Boer et al., 2021). All studies contributing data to our analyses were approved by the relevant ethics committees, and all study participants in these studies provided written, informed consent.

2.4 Statistical analysis

Two-sample MR analysis was executed in R software (R version 4.3.0) with the TwoSampleMR (version 0.5.6), MR-PRESSO (version 1.0) and mr. raps (version 0.2) R packages (Hemani et al., 2018; Verbanck et al., 2018). The methods applied in our research are presented in Figure 2. MR-steiger test was applied to infer causal direction between traits under investigation (Hemani et al., 2017). Four methodologies including inverse variance weighted (IVW), weighted median (WM), MR-Egger and MR-robust adjusted profile score (MR-RAPS) were employed to estimate causality between mtDNA abundance and MtHz and OA. IVW was taken as a primary MR analysis method in our research, which is based on the hypothesis that all selected genetic instruments are valid and give an overall causal estimate strengthening causal inference (Lawlor et al., 2008; Burgess et al., 2013). MR-Egger regression takes presence of directional pleiotropy into consideration and measures horizontal pleiotropy with regression intercept, whereas the results of this method are susceptible to outlying genetic variants (Bowden et al., 2015). Weighted Median based on hypothesis that at least 50% of the variants are valid, improves power of causal effect detection but reduces precision (Bowden et al., 2016). MR-RAPS applies robust adjusted profile scores to correct for pleiotropy and makes our results more reliable (Hartwig et al., 2017).

Several sensitivity analyses were applied to detect and correct for heterogeneity and pleiotropy. MR-PRESSO was conducted to detect the existence of horizontal pleiotropy and correct the causal estimate affected by possible pleiotropic outliers (Verbanck et al., 2018). Corresponding p-values were derived based on 1,000 simulations (Verbanck et al., 2018). And the estimation of MR Egger regression intercept was also employed to reflect presence of pleiotropy (p < 0.05) (Burgess and Thompson, 2017). IVW method was utilized to investigate heterogeneity. The level of heterogeneity was quantified by Cochran Q statistics (Carnegie et al., 2020). Leave-one-out sensitivity analysis was performed to identify possibly influential SNPs, which repeated MR analysis with each SNP excluded in turn (Carnegie et al., 2020). As five separate outcomes were tested in our study, main results had statistical significance at p-value<0.01 (0.05/5) after Bonferroni correction.

3 Results

3.1 Causal effect of mitochondrial heteroplasmy on OA

Fifteen LD-independent genetic variants were taken for repeated MR analysis (Supplementary Table S1). We extracted data of above-mentioned SNPs from summary GWAS of outcome traits (overall OA and specific-site OA) and one SNP (rs2286639) was removed in all outcomes except finger OA due to the effect of non-concordant alleles (e.g., A/G vs A/C). In finger OA, two SNP (rs2286639, rs3702096) were excluded for absence of proxy SNPs on the online platform SNiPA (https://snipa.helmholtzmuenchen.de/snipa3/). The mean F-statistic was 36.357 which was above the threshold F value of 10. In total, there were 14 IVs for overall, knee, hip, thumb, and hand OA and 13 IVs for finger OA.

Results of the casual association between MtHz and OA are summarized in Table 1. The primary IVW analysis provided no evidence for the casual association between MtHz and overall OA [odds ratio (OR) = 0.852, 95% confidence interval (CI): 0.664–1.093, p= 0.208]. The results of WM method, MR-RAPS and MR-Egger were consistent with the result of IVW. Both MR-Egger and MR-

TABLE 1 MR results of causal effect of mitochondrial heteroplasmy on OA and its other phenotypes.

Outcome traits	MR methods	Mitochondrial heteroplasmy					
		OR (95%CI)	SE	MR *p*-value	Heterogeneity test	Pleiotropy test	MR-stegier test
					Cochran's Q (*p*)	*p* intercept	
Overall OA	MR-Egger	0.922 (0.619,1.373)	0.203	0.697	12.407 (0.495)	0.625	Direction: TRUE
	Inverse variance weighted	0.852 (0.664,1.093)	0.127	0.208			*p*-value < 0.0001
	MR-RAPS	0.850 (0.662,1.093)	0.128	0.206			
	Weighted median	0.867 (0.640,1.174)	0.155	0.355			
	MRPRESSO	0.851 (0.670,1.082)	0.123	0.209			
Knee OA	MR-Egger	0.667 (0.358,1.243)	0.318	0.226	11.272 (0.588)	0.831	Direction: TRUE
	Inverse variance weighted	**0.632 (0.425,0.939)**	0.202	**0.023**			*p*-value < 0.0001
	MR-RAPS	0.630 (0.422,0.939)	0.204	**0.023**			
	Weighted median	0.637 (0.402,1.009)	0.235	0.055			
	MRPRESSO	**0.631 (0.442,0.903)**	0.182	**0.024**			
Hip OA	MR-Egger	0.708 (0.320,1.568)	0.406	0.412	10.208 (0.676)	0.897	Direction: TRUE
	Inverse variance weighted	0.738 (0.447,1.221)	0.257	0.237			*p*-value < 0.0001
	MR-RAPS	0.734 (0.442,1.220)	0.259	0.234			
	Weighted median	0.711 (0.391,1.293)	0.310	0.272			
	MRPRESSO	0.736 (0.466,1.161)	0.233	0.208			
Thumb OA	MR-Egger	0.708 (0.320,1.568)	0.406	0.412	10.161 (0.681)	0.288	Direction: TRUE
	Inverse variance weighted	1.270 (0.506,3.188)	0.469	0.611			*p*-value < 0.0001
	MR-RAPS	1.270 (0.502,3.216)	0.474	0.615			
	Weighted median	0.711 (0.391,1.293)	0.305	0.264			
	MRPRESSO	1.267 (0.576,2.791)	0.403	0.566			
Hand OA	MR-Egger	0.967 (0.333,2.803)	0.543	0.951	7.958 (0.846)	0.432	Direction: TRUE
	Inverse variance weighted	1.357 (0.687,2.681)	0.347	0.379			*p*-value < 0.0001
	MR-RAPS	1.350 (0.678,2.686)	0.351	0.392			
	Weighted median	1.132 (0.503,2.546)	0.413	0.764			
	MRPRESSO	1.348 (0.769,2.362)	0.286	0.315			
Finger OA	MR-Egger	0.139 (0.010,1.909)	1.336	0.168	8.157 (0.773)	0.251	Direction: TRUE
	Inverse variance weighted	0.483 (0.090,2.588)	0.857	0.395			*p*-value < 0.0001
	MR-RAPS	0.481 (0.088,2.621)	0.865	0.398			
	Weighted median	0.364 (0.049,2.688)	1.021	0.322			
	MRPRESSO	0.483 (0.121,1.927)	0.707	0.323			

MR, mendelian randomization; SNP, single nucleotide polymorphism; OA, osteoarthritis; OR, odds ratio; CI, confidence interval; SE, standard error (the standard error is an estimate of the standard deviation (SD) of the coefficient); *p*, *p*-value.

TABLE 2 MR results of causal effect of mtDNA abundance on OA and its other phenotypes.

Outcome traits	MR methods	MtDNA abundance					
		OR (95%CI)	SE	MR *p*-value	Heterogeneity test	Pleiotropy test	MR-stegier test
					Cochran's Q (*p*)	*p* intercept	
Overall OA	MR-Egger	1.091 (0.856,1.391)	0.124	0.485	**62.748 (0.033)**	0.687	Direction: TRUE
	Inverse variance weighted	1.040 (0.963,1.125)	0.040	0.318			
	MR-RAPS	1.041 (0.974,1.112)	0.034	0.229			*p*-value < 0.0001
	Weighted median	**1.108 (1.001,1,226)**	0.052	**0.048**			
	MRPRESSO	1.040 (0.963,1.125)	0.040	0.323			
Knee OA	MR-Egger	0.823 (0.562,1.206)	0.195	0.324	65.050 (0.027)	0.225	Direction: TRUE
	Inverse variance weighted	1.034 (0.915,1.168)	0.062	0.592			
	MR-RAPS	1.035 (0.934,1.145)	0.052	0.516			*p*-value < 0.0001
	Weighted median	1.056 (0.903,1.236)	0.080	0.494			
	MRPRESSO	1.034 (0.915,1.168)	0.062	0.594			
Hip OA	MR-Egger	**0.616 (0.399,0.950)**	0.221	**0.034**	60.396 (0.062)	**0.027**	Direction: TRUE
	Inverse variance weighted	0.992 (0.857,1.152)	0.076	0.931			
	MR-RAPS	0.993 (0.873,1.130)	0.066	0.919			*p*-value < 0.0001
	Weighted median	0.921 (0.763,1.113)	0.096	0.397			
	MRPRESSO	0.993 (0.857,1.152)	0.076	0.931			
Thumb OA	MR-Egger	0.819 (0.364,1.844)	0.414	0.633	61.709 (0.061)	0.392	Direction: TRUE
	Inverse variance weighted	1.148 (0.880,1.499)	0.136	0.309			
	MR-RAPS	1.151 (0.912,1.454)	0.119	0.235			*p*-value <0.0001
	Weighted median	1.135 (0.790,1.631)	0.185	0.494			
	MRPRESSO	1.148 (0.880,1.499)	0.136	0.314			
Hand OA	MR-Egger	1.028 (0.601,1.759)	0.274	0.920	40.089 (0.680)	0.663	Direction: TRUE
	Inverse variance weighted	1.152 (0.969,1.369)	0.088	0.108			
	MR-RAPS	1.154 (0.967,1.376)	0.090	0.109			*p*-value < 0.0001
	Weighted median	1.144 (0.887,1.476)	0.130	0.299			
	MRPRESSO	1.152 (0.979,1.356)	0.083	0.096			
Finger OA	MR-Egger	0.876 (0.223,3.442)	0.698	0.850	51.688 (0.229)	0.837	Direction: TRUE
	Inverse variance weighted	1.004 (0.647,1.556)	0.224	0.987			
	MR-RAPS	1.003 (0.662,1.520)	0.212	0.986			*p*-value < 0.0001
	Weighted median	1.193 (0.641,2.221)	0.317	0.577			
	MRPRESSO	1.004 (0.647,1.556)	0.224	0.987			

MR, mendelian randomization; SNP, single nucleotide polymorphism; OA, osteoarthritis; mtDNA, mitochondrial DNA; OR, odds ratio; CI, confidence interval; SE, standard error (the standard error is an estimate of the standard deviation (SD) of the coefficient); *p*, *p*-value.

PRESSO reported no existence of horizontal pleiotropy (Egger-intercept: $p > 0.05$; MR-PRESSO global heterogeneity: $p > 0.05$). Cochran's Q test did not detect heterogeneity in overall OA (p= 0.495 > 0.05).

In site-specific OA analyses, the nominally significant results of IVW analysis suggested that MtHz was a potentially protective factor for knee OA (OR = 0.632, 95% CI: 0.425–0.939, p= 0.023, Table 1). MR-RAPS showed similar results with IVW analysis (OR = 0.629, 95% CI: 0.422–0.939, p= 0.023, Table 1). However, no relationship between MtHz and other OA sites, including hip OA (IVW OR = 0.738, 95% CI: 0.447–1.221, p= 0.237), hand OA (IVW OR = 1.357, 95% CI: 0.687–2.681, p= 0.379), thumb OA (IVW OR = 1.270, 95% CI: 0.506–3.188, p= 0.611) or finger OA (IVW OR = 0.483, 95% CI: 0.090–2.588, p= 0.395) was observed in IVW model (Table 1). MR-PRESSO was performed to test for horizontal pleiotropy, and no outliers was identified. The results of MR-Egger intercept were consistent with the MR-PRESSO results (intercept $p > 0.05$). The heterogeneity was tested by Cochran's Q test and MR-PRESSO global heterogeneity test, providing no evidence about the existence of heterogeneity. The MR-Steiger results supported that these SNPs were more predictive of the exposure than of the outcome ($p < 0.05$, Supplementary Table S3). The results of leave-one-out sensitivity analysis and forest plots demonstrated that our study in genetically prediction was robust (Supplementary Figure S1–S3).

3.2 Causal effect of mitochondrial abundance on OA

After LD test, 47 independent genetic variants were chosen for two-sample MR analysis. Several genetic variants could not be matched in summary statistic GWASs of knee (rs35734242), finger (rs1065853), and hand OA (rs12451555). We searched their proxy SNPs on SNiPA and those whose proxy SNPs were absent on SNiPA were excluded. The mean F-statistic for IVs of mtDNA abundance was 93.380, which was above the threshold 10 (Supplementary Table S2). Finally, the SNPS were selected as IVs for thumb, hip, overall OA was 47, and 46 SNPS for hand, finger, knee. Regarding overall OA, MR-PRESSO identified two outliers s (rs59488041, rs12924138) and therefore removed them for repeated MR analysis. In the presence of heterogeneity (Cochran's Q: p= 0.033) and absence of pleiotropy (Egger-intercept: p= 0.687 > 0.05), the results of WM analysis on overall OA were nominally significant (OR = 1.108, 95% CI: 1.001–1,226, p= 0.048, Table 2). IVW method in random effects model was utilized to correct results potentially impacted by heterogeneity, and the results suggested that genetically elevated mtDNA abundance was not casually associated with overall OA (OR = 1.040, 95% CI: 0.963–1.125, p= 0.318). MR-RAPS analysis agreed with the results of IVW analysis (OR = 1.041, 95% CI: 0.974–1.112, p= 0.228, Table 2).

With respect to site-specific OA, we observed that genetically determined MtHz was not causally associated with knee OA (OR = 1.034, 95% CI: 0.915–1.168, p= 0.592), thumb OA (OR = 1.148, 95% CI: 0.880–1.499, p= 0.309), hand OA (OR = 1.152, 95% CI: 0.969–1.369, p = 0.108) and finger OA (OR = 1.004, 95% CI: 0.647–1.556, p= 0.987) in IVW model. The results are presented in Table 2. Both MR-Egger and MR-PRESSO reported no existence of horizontal pleiotropy (Egger-intercept: $p > 0.05$; MR-PRESSO global heterogeneity: $p > 0.05$). Cochran's Q test did not detect heterogeneity in above-mentioned outcome traits ($p > 0.05$). For the analysis of hip OA, rs16978036 were excluded from MR-PRESSO analysis. In presence of horizontal pleiotropy (p= 0.027) and absence of heterogeneity (p= 0.062), the nominally significant result of MR-Egger method was found (OR = 0.616, 95% CI: 0.399–0.950, p= 0.034). However, no evidence about causal relationship between the exposure and hip OA was observed with IVW method and MR-RAPS method (IVW: OR = 0.992, 95% CI: 0.857–1.152, p= 0.931; MR-RAPS: OR = 0.993, 95% CI: 0.873–1.130, p= 0.919) (Table 2). The results of MR-Stegier test were in Supplementary Table S3.

Scatter plots, forest plots and leave-on-out sensitivity analysis plots were displayed in Supplementary Figure S4–6. The results of leave-one-out analysis implicated that the selected genetic variants potentially impacted the pooled results, which suggested that careful interpretations for the results was crucial.

4 Discussion

The causal roles of MtHz and mtDNA abundance in OA pathogenesis were poorly studied in previous works. To our best knowledge, this is the first MR study to evaluate the causal association between mitochondrial genome traits and OA. Four MR methods were employed to estimate the causal association between mtDNA characterization and OA. The results of MR steiger test verified the causal direction of our research (MtHz and mtDNA abundance were exposure and OA were outcome). No effect of mtDNA abundance and MtHz on overall OA was detected. In subgroup analysis, a suggestively protective role of MtHz on knee OA was observed, but not on other sites. And we did not find that mtDNA abundance was causally associated with any site-specific OA.

Although mitochondria is widely recognized as an important factor of OA development (Blanco et al., 2011), the role of MtHz has not been thoroughly investigated (Suliman and Piantadosi, 2016). Indeed, heteroplasmic mutations in mtDNA are often pathogenetic (Ye et al., 2014). An animal study has indicated that the state of heteroplasmy itself was deleterious when the two mtDNA sequences contain no pathogenic variants (Sharpley et al., 2012). However, there is potentially a particular level threshold for MtHz (McCormick et al., 2020). For instance, when the A3243G mutation in mitochondrial DNA is present in more than 10%, patients can manifest Type 2 diabetes (Wallace, 2005). And low-frequency mtDNA variants (0.2%–2% heteroplasmy) are extensively presented in healthy subjects (Payne et al., 2013). Furthermore, MtHz can be beneficial in health promotion as an intermediate state in emergence of novel mtDNA haplogroups (Wallace, 2016). In previous study, MtHz was found to be significantly relevant to several haplogroups (haplogroup H, J, K, T, U and X) with different characterizations among haplogroups (Nandakumar et al., 2021). Mitochondrial haplogroup J has been extensively found to mediate the development of OA (5). A Spanish cohort-study (n_{case} = 457; $n_{control}$ = 262) had reported that haplogroup J was associated with a decreased risk of knee OA (OR = 0.460, 95% CI: 0.282–0.748, p= 0.002) (Rego-Perez et al., 2008). Furthermore, a meta-analysis in European cohorts also suggested that haplogroup J was associated with a lower risk of knee OA (HR = 0.702, 95% CI: 0.541–0.912, p= 0.008) (Fernandez-Moreno et al., 2017). However, the association between mitochondrial DNA variants and OA has not been verified in a large sample observational study (Hudson et al., 2013). The study has only explored the causal relationship in terms of single mutation, but we take MtHz (variation at a whole mtDNA level) as an exposure which contribute to understand the causal role of mitochondria in OA development. In terms of biological mechanism, MtHz has been reported a regulation role of metabolic and epigenomic changes. Kopinski PK et al. had found the levels of mitochondrially drove acetyl-

CoA decreased at high heteroplasmy (Kopinski et al., 2019). And the reduction of acetyl-CoA levels might influence histone acetylation and activate AMP-activated protein kinase (AMPK) to protect from OA development and progression (Chen et al., 2018). Besides, MtHz could play a role in other processes involved in OA pathogenesis to reduce the risk of OA, including regulating energy production, maintaining mitochondrial proteostasis, suppressing matrix metalloproteinase expression, reducing ROS generation, and promoting mitophagy (Blanco et al., 2018). In conjunction with prior observational findings and mechanism studies, our epidemiologic and genetic findings provided supportive evidence that MtHz may have therapeutic value in knee OA and can be regarded as a candidate biomarker after precisely identifying the relationship between MtHz and each haplogroup (Blanco et al., 2018).

However, the reason that MtHz showed diverse effects on different sites of OA remains unclear. So far, we did not find relevant research that investigated the relationship between MtHz and hand, finger, and thumb OA. A case-control study suggested that mitochondrial haplogroups was associated with hip OA (Rego et al., 2010) (OR = 0.661, 95% CI: 0.440–0.993, p= 0.045), however, with no evidence on the causality. In addition, the data sources and sample size used in MR analysis varied in site-specific OA, which mainly included more knee OA cases and less other skeletal joints cases. More MR studies and additional GWAS covering more other site-specific OAs are needed to determine a role of MtHz in the risk of different OA sites.

MtDNA abundance has been regarded as a potential biomarker of mitochondrial function and plays a role in several human diseases (Castellani et al., 2020; Clyde, 2022). Existing literature has suggested that mtDNA abundance may involve in production of inflammatory mediators and regulation of immune function that can influence OA development (Blanco et al., 2018; Zhan et al., 2020). In addition, mtDNA abundance is also associated with sex, advanced age, and elevated BMI, which are also risk factors for OA (16). However, there are few studies to estimate the effects of mtDNA abundance on OA. Only two case-control studies in Asian population were found that reported an association between mtDNA copy number and OA, but their findings were inconsistent (Fang et al., 2014; Zhan et al., 2020). One in Thailand (n_{case} = 204; $n_{control}$ = 169) had found mtDNA abundance in the OA group was significantly lower than that in the control group ($p < 0.0001$), whereas the results were not adjusted by sex and age (Zhan et al., 2020). Another case-control study carried out in southern Chinese (n_{case} = 187; $n_{control}$ = 420) observed a general increase of mtDNA abundance in OA patients (p= 0.019), but obesity was not adjusted in the analysis (Fang et al., 2014). These findings from case-control studies could not determine causal relationships between mtDNA abundance and OA, and the potential confounding variables were not comprehensively considered. Our MR analysis overcame these shortcomings and suggested that genetically determined mtDNA abundance was unrelated to OA. Furthermore, previous study has implicated that the relative contribution of mtDNA abundance might differ between different ethnic groups (Ruiz-Pesini et al., 2004). Considering that our studies have been carried out in European population, the comparisons with results from other ethnic groups such as Asian ancestry require careful considerations.

Our research applies MR methods to investigate causal relationships between mitochondrial genome characterization and OA. However, there are some limitations in our analysis. Firstly, we did not estimate causal effects of mitochondrial genome traits on OA stratified by gender. Mitochondrial genome are maternally inherited, and females may have lower MtHz than males from the perspective of mitochondrial inheritance (Nandakumar et al., 2021). Therefore, the effects of MtHz and mtDNA abundance on OA may differ in gender. Secondly, only participants of European descent are included in the study, but the impact of specific mtDNA variants on diseases could vary in different ethnic groups (Ruiz-Pesini et al., 2004). Additional MR studies on other ethnic groups are needed to probe a causal association between mtDNA characterization and OA. Thirdly, the genetic instruments for mtDNA abundance explain a relatively small amount of phenotypic variance (8.3%), IVs that can account for more variance of mtDNA abundance are warranted to draw robust conclusions. Besides, considering the Bonferroni correction of multiple independent tests, our findings about the association between MtHz and knee OA are deemed suggestive evidence of possible associations ($0.01 < p < 0.05$). Furthermore, the variation of mtDNA abundance and the levels of MtHz varied in different tissues and cell types. And in original GWASs, mtDNA abundance was measured from leukocyte in blood and MtHz from saliva samples rather than cartilage or synovial tissues. Considering that OA is mainly characterized by progressive loss of cartilage and synovial hyperplasia, the application of data from blood samples and saliva samples would limit the explanation ability of our study to a certain extent (Sellam and Berenbaum, 2010). Observational epidemiology studies exploring an association between concrete levels of MtHz and a risk of OA are needed to improve causal inference.

5 Conclusion

In conclusion, our MR analyses elucidated that MtHz is a suggestively protective factor of knee OA, implying that MtHz could be a genetically prediction factor and a therapeutic target in the development of knee OA. No causal association was found between mtDNA abundance and OA. Additional MR analyses are warranted to probe causal relationships between mitochondrial genome traits and OA stratified by gender. Moreover, GWAS covering more than one ethnic population are needed to detect the effect of characterization of mtDNA quantity and quality in different ethnic groups.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

HL mainly designed analysis and wrote the manuscript; BC, RG, DZ performed experiments and analysis; YY verified data and analysis results, JW, MY, YL supervised the entire project. All authors have read, provided critical feedback on intellectual content, and approved the final manuscript.

Acknowledgments

The authors thank all the participants and researchers for their participation in this MR study. All authors interpreted the data, critically revised the manuscript for important intellectual content, approved the final version of the manuscript, and agreed to be responsible for all aspects of the work. The corresponding author attests that all listed co-authors meet authorship criteria and that no others meeting the criteria have been omitted.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1130411/full#supplementary-material

References

Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H., Coulson, A. R., Drouin, J., et al. (1981). Sequence and organization of the human mitochondrial genome. *Nature* 290 (5806), 457–465. doi:10.1038/290457a0

Blanco, F. J., Rego, I., and Ruiz-Romero, C. (2011). The role of mitochondria in osteoarthritis. *Nat. Rev. Rheumatol.* 7 (3), 161–169. doi:10.1038/nrrheum.2010.213

Blanco, F. J., Valdes, A. M., and Rego-Perez, I. (2018). Mitochondrial DNA variation and the pathogenesis of osteoarthritis phenotypes. *Nat. Rev. Rheumatol.* 14 (6), 327–340. doi:10.1038/s41584-018-0001-0

Boer, C. G., Hatzikotoulas, K., Southam, L., Stefansdottir, L., Zhang, Y., Coutinho de Almeida, R., et al. (2021). Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations. *Cell* 184 (18), 6003–6005. doi:10.1016/j.cell.2021.11.003

Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: Effect estimation and bias detection through egger regression. *Int. J. Epidemiol.* 44 (2), 512–525. doi:10.1093/ije/dyv080

Bowden, J., Davey Smith, G., Haycock, P. C., and Burgess, S. (2016). Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.* 40 (4), 304–314. doi:10.1002/gepi.21965

Burgess, S., Butterworth, A., and Thompson, S. G. (2013). Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet. Epidemiol.* 37 (7), 658–665. doi:10.1002/gepi.21758

Burgess, S., Small, D. S., and Thompson, S. G. (2017). A review of instrumental variable estimators for Mendelian randomization. *Stat. Methods Med. Res.* 26 (5), 2333–2355. doi:10.1177/0962280215597579

Burgess, S., Thompson, S. G., and Collaboration, C. C. G. (2011). Avoiding bias from weak instruments in Mendelian randomization studies. *Int. J. Epidemiol.* 40 (3), 755–764. doi:10.1093/ije/dyr036

Burgess, S., and Thompson, S. G. (2017). Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur. J. Epidemiol.* 32 (5), 377–389. doi:10.1007/s10654-017-0255-x

Carnegie, R., Zheng, J., Sallis, H. M., Jones, H. J., Wade, K. H., Evans, J., et al. (2020). Mendelian randomisation for nutritional psychiatry. *Lancet Psychiatry* 7 (2), 208–216. doi:10.1016/S2215-0366(19)30293-7

Castellani, C. A., Longchamps, R. J., Sun, J., Guallar, E., and Arking, D. E. (2020). Thinking outside the nucleus: Mitochondrial DNA copy number in health and disease. *Mitochondrion* 53, 214–223. doi:10.1016/j.mito.2020.06.004

Chen, L. Y., Lotz, M., Terkeltaub, R., and Liu-Bryan, R. (2018). Modulation of matrix metabolism by ATP-citrate lyase in articular chondrocytes. *J. Biol. Chem.* 293 (31), 12259–12270. doi:10.1074/jbc.RA118.002261

Clyde, D. (2022). Genomic insights into conifer evolution. *Nat. Rev. Genet.* 23 (3), 136. doi:10.1038/s41576-022-00452-1

Fang, H., Liu, X., Shen, L., Li, F., Liu, Y., Chi, H., et al. (2014). Role of mtDNA haplogroups in the prevalence of knee osteoarthritis in a southern Chinese population. *Int. J. Mol. Sci.* 15 (2), 2646–2659. doi:10.3390/ijms15022646

Fernandez-Moreno, M., Rego-Perez, I., and Blanco, F. J. (2022). Is osteoarthritis a mitochondrial disease? What is the evidence. *Curr. Opin. Rheumatol.* 34 (1), 46–53. doi:10.1097/BOR.0000000000000855

Fernandez-Moreno, M., Soto-Hermida, A., Vazquez-Mosquera, M. E., Cortes-Pereira, E., Relano, S., Hermida-Gomez, T., et al. (2017). Mitochondrial DNA haplogroups influence the risk of incident knee osteoarthritis in OAI and CHECK cohorts. A meta-analysis and functional study. *Ann. Rheum. Dis.* 76 (6), 1114–1122. doi:10.1136/annrheumdis-2016-210131

Frahm, T., Mohamed, S. A., Bruse, P., Gemund, C., Oehmichen, M., and Meissner, C. (2005). Lack of age-related increase of mitochondrial DNA amount in brain, skeletal muscle and human heart. *Mech. Ageing Dev.* 126 (11), 1192–1200. doi:10.1016/j.mad.2005.06.008

Glyn-Jones, S., Palmer, A. J., Agricola, R., Price, A. J., Vincent, T. L., Weinans, H., et al. (2015). *Osteoarthr. Lancet.* 386 (9991), 376–387. doi:10.1016/S0140-6736(14)60802-3

Hagg, S., Jylhava, J., Wang, Y., Czene, K., and Grassmann, F. (2021). Deciphering the genetic and epidemiological landscape of mitochondrial DNA abundance. *Hum. Genet.* 140 (6), 849–861. doi:10.1007/s00439-020-02249-w

Hartwig, F. P., Davey Smith, G., and Bowden, J. (2017). Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiol.* 46 (6), 1985–1998. doi:10.1093/ije/dyx102

Hemani, G., Tilling, K., and Davey Smith, G. (2017). Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* 13 (11), e1007081. doi:10.1371/journal.pgen.1007081

Hemani, G., Zheng, J., Elsworth, B., Wade, K. H., Haberland, V., Baird, D., et al. (2018). The MR-Base platform supports systematic causal inference across the human phenome. *Elife* 7, e34408. doi:10.7554/eLife.34408

Hu, L., Yao, X., and Shen, Y. (2016). Altered mitochondrial DNA copy number contributes to human cancer risk: Evidence from an updated meta-analysis. *Sci. Rep.* 6, 35859. doi:10.1038/srep35859

Hudson, G., Panoutsopoulou, K., Wilson, I., Southam, L., Rayner, N. W., Arden, N., et al. (2013). No evidence of an association between mitochondrial DNA variants and osteoarthritis in 7393 cases and 5122 controls. *Ann. Rheum. Dis.* 72 (1), 136–139. doi:10.1136/annrheumdis-2012-201932

Kopinski, P. K., Janssen, K. A., Schaefer, P. M., Trefely, S., Perry, C. E., Potluri, P., et al. (2019). Regulation of nuclear epigenome by mitochondrial DNA heteroplasmy. *Proc. Natl. Acad. Sci. U. S. A.* 116 (32), 16028–16035. doi:10.1073/pnas.1906896116

Lawlor, D. A., Harbord, R. M., Sterne, J. A., Timpson, N., and Davey Smith, G. (2008). Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat. Med.* 27 (8), 1133–1163. doi:10.1002/sim.3034

McArdle, M. A., Finucane, O. M., Connaughton, R. M., McMorrow, A. M., and Roche, H. M. (2013). Mechanisms of obesity-induced inflammation and insulin resistance: Insights into the emerging role of nutritional strategies. *Front. Endocrinol. (Lausanne)* 4, 52. doi:10.3389/fendo.2013.00052

McCormick, E. M., Lott, M. T., Dulik, M. C., Shen, L., Attimonelli, M., Vitale, O., et al. (2020). Specifications of the ACMG/AMP standards and guidelines for mitochondrial DNA variant interpretation. *Hum. Mutat.* 41 (12), 2028–2057. doi:10.1002/humu.24107

McGuire, P. J. (2019). Mitochondrial dysfunction and the aging immune system. *Biol. (Basel)* 8 (2), 26. doi:10.3390/biology8020026

Nandakumar, P., Tian, C., O'Connell, J., andMe Research, T., Hinds, D., Paterson, A. D., et al. (2021). Nuclear genome-wide associations with mitochondrial heteroplasmy. *Sci. Adv.* 7 (12), eabe7520. doi:10.1126/sciadv.abe7520

Nelson, A. E. (2018). Osteoarthritis year in review 2017: Clinical. *Osteoarthr. Cartil.* 26 (3), 319–325. doi:10.1016/j.joca.2017.11.014

Payne, B. A., Wilson, I. J., Yu-Wai-Man, P., Coxhead, J., Deehan, D., Horvath, R., et al. (2013). Universal heteroplasmy of human mitochondrial DNA. *Hum. Mol. Genet.* 22 (2), 384–390. doi:10.1093/hmg/dds435

Rego, I., Fernandez-Moreno, M., Fernandez-Lopez, C., Gomez-Reino, J. J., Gonzalez, A., Arenas, J., et al. (2010). Role of European mitochondrial DNA haplogroups in the prevalence of hip osteoarthritis in Galicia, Northern Spain. *Ann. Rheum. Dis.* 69 (1), 210–213. doi:10.1136/ard.2008.105254

Rego-Perez, I., Fernandez-Moreno, M., Fernandez-Lopez, C., Arenas, J., and Blanco, F. J. (2008). Mitochondrial DNA haplogroups: Role in the prevalence and severity of knee osteoarthritis. *Arthritis Rheum.* 58 (8), 2387–2396. doi:10.1002/art.23659

Ruiz-Pesini, E., Mishmar, D., Brandon, M., Procaccio, V., and Wallace, D. C. (2004). Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science* 303 (5655), 223–226. doi:10.1126/science.1088434

Sellam, J., and Berenbaum, F. (2010). The role of synovitis in pathophysiology and clinical symptoms of osteoarthritis. *Nat. Rev. Rheumatol.* 6 (11), 625–635. doi:10.1038/nrrheum.2010.159

Sharpley, M. S., Marciniak, C., Eckel-Mahan, K., McManus, M., Crimi, M., Waymire, K., et al. (2012). Heteroplasmy of mouse mtDNA is genetically unstable and results in altered behavior and cognition. *Cell* 151 (2), 333–343. doi:10.1016/j.cell.2012.09.004

Suliman, H. B., and Piantadosi, C. A. (2016). Mitochondrial quality control as a therapeutic target. *Pharmacol. Rev.* 68 (1), 20–48. doi:10.1124/pr.115.011502

Tian, Q., Moore, A. Z., Oppong, R., Ding, J., Zampino, M., Fishbein, K. W., et al. (2021). Mitochondrial DNA copy number and heteroplasmy load correlate with skeletal muscle oxidative capacity by P31 MR spectroscopy. *Aging Cell* 20 (11), e13487. doi:10.1111/acel.13487

Verbanck, M., Chen, C. Y., Neale, B., and Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* 50 (5), 693–698. doi:10.1038/s41588-018-0099-7

Wallace, D. C. (2005). A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: A dawn for evolutionary medicine. *Annu. Rev. Genet.* 39, 359–407. doi:10.1146/annurev.genet.39.110304.095751

Wallace, D. C. (2016). Genetics: Mitochondrial DNA in evolution and disease. *Nature* 535 (7613), 498–500. doi:10.1038/nature18902

Ye, K., Lu, J., Ma, F., Keinan, A., and Gu, Z. (2014). Extensive pathogenicity of mitochondrial heteroplasmy in healthy human individuals. *Proc. Natl. Acad. Sci. U. S. A.* 111 (29), 10654–10659. doi:10.1073/pnas.1403521111

Zhan, D., Tanavalee, A., Tantavisut, S., Ngarmukos, S., Edwards, S. W., and Honsawek, S. (2020). Relationships between blood leukocyte mitochondrial DNA copy number and inflammatory cytokines in knee osteoarthritis. *J. Zhejiang Univ. Sci. B* 21 (1), 42–52. doi:10.1631/jzus.B1900352

frontiers | Frontiers in Genetics

TYPE Case Report
PUBLISHED 21 March 2023
DOI 10.3389/fgene.2023.1101695

OPEN ACCESS

EDITED BY
Fang Fang Yu,
Zhengzhou University, China

REVIEWED BY
Xuyun Hu,
Beijing Children's Hospital, Capital Medical University, China
Jose Bernardo Quintos,
Rhode Island Hospital, United States

*CORRESPONDENCE
Rong Xiang,
shirlesmile@csu.edu.cn
Xia Wang,
xiawang@csu.edu.cn

SPECIALTY SECTION
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

RECEIVED 18 November 2022
ACCEPTED 10 March 2023
PUBLISHED 21 March 2023

CITATION
Huang H, Jin J, Xiang R and Wang X (2023), Case report: A novel heterozygous frameshift mutation of *ACAN* in a Chinese family with short stature and advanced bone age. *Front. Genet.* 14:1101695. doi: 10.3389/fgene.2023.1101695

Case report: A novel heterozygous frameshift mutation of *ACAN* in a Chinese family with short stature and advanced bone age

Hao Huang[1,2,3], Jieyuan Jin[3,4,5], Rong Xiang[2,3]* and Xia Wang[3,4]*

[1]Department of Nephrology, Xiangya Hospital, Central South University, Changsha, China, [2]Hunan Key Laboratory of Organ Fibrosis, Central South University, Changsha, China, [3]National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China, [4]Department of Pediatrics, Xiangya Hospital, Central South University, Changsha, China, [5]Department of Orthopedics, Xiangya Hospital, Central South University, Changsha, China

Short stature (OMIM: 165800) is a common pediatric disorder. Any abnormality in the cartilage formation of the growth plate can cause short stature. Aggrecan, encoded by *ACAN*, is an important component of the extracellular matrix. Mutations in *ACAN* have been reported to cause short stature. In the present study, we enrolled a Chinese family with short stature and advanced bone age across three generations. Whole-exome sequencing (WES) was performed on the proband to detect the candidate genes causing short stature in family. A novel heterozygous frameshift mutation (NM_013227.3:c.7230delT; NP_001356197.1: p. Phe2410Leufs*9) of the *ACAN* gene was confirmed to be a genetic lesion in this family. This variant, which was located in a functional site globular 3 (G3) domain of ACAN and predicted to be deleterious by informatics programs, was co-segregated with the affected family members by performing Sanger sequencing. Literatures review of growth hormone (GH) treatment outcome of all previously reported *ACAN* patients suggesting that the G3 domain of ACAN may be critical in the development of short stature and growth hormone treatment. These findings not only contribute to the genetic diagnosis and counseling of the family, but will also expand the mutation spectrum of *ACAN*.

KEYWORDS

short stature, novel mutation, acan, whole-exome sequencing, frameshift, case report

1 Introduction

Short stature (OMIM: 165800) is a type of condition that is associated with a child whose height is 2 standard deviations (SD) or more below the mean for children of the same chronological age and sex (Baron et al., 2015; Finken et al., 2018; Grunauer and Jorge, 2018). The growth plate is the target organ of linear growth and height gain (Baron et al., 2015); it is composed of chondrocytes embedded in a collagen matrix that show high metabolic activity. Any abnormality in the cartilage formation of the growth plate can cause linear growth disturbance, which in turn, may lead to short stature.

Existing research suggests that defects in the genes encoding components of either paracrine regulatory systems or the extracellular matrix (ECM), and those encoding intracellular or endocrine proteins result in short stature (Jee et al., 2017). The gene encoding aggrecan (*ACAN*, OMIM: 155760) is located on chromosome 15q26, *ACAN* encodes the production of the proteoglycan aggrecan, which is an important component of the ECM (Stavber et al., 2020). Aggrecan is highly expressed in proliferating chondrocytes of

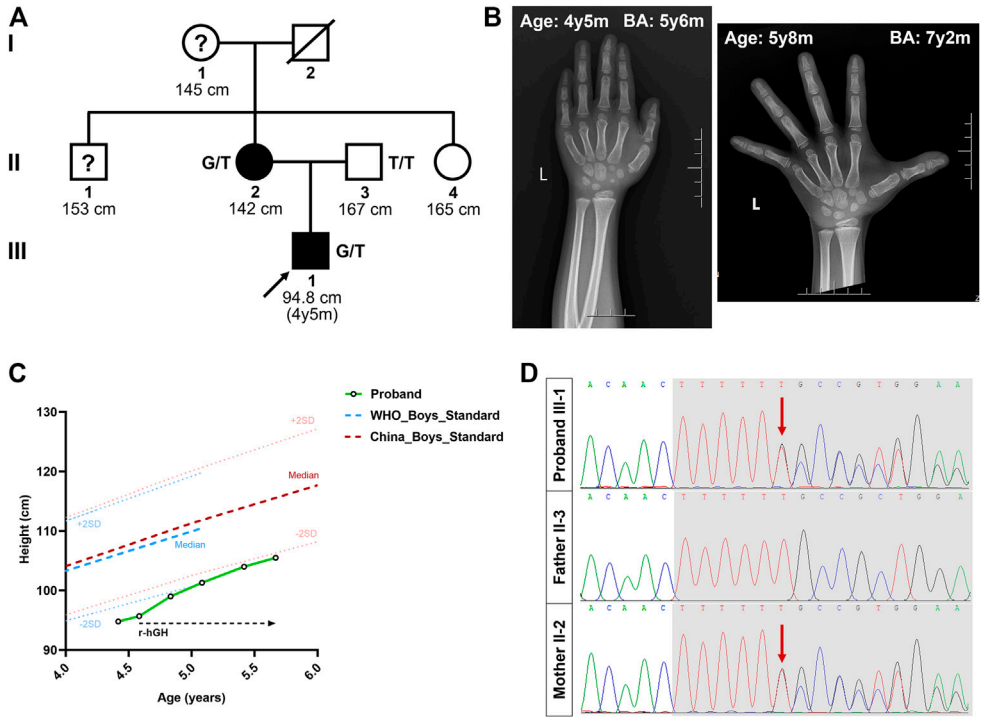


FIGURE 1
Pedigree of the Chinese family with short stature. **(A)** The black circles/squares represent the affected patients diagnosed with short stature diagnosis; the white circles/squares represent the unaffected subjects. The arrow indicates the proband. The question mark indicates subject with unknown genotype but known phenotype. **(B)** X-ray imaging of the proband's left hand at the age of 4 years 5 months and 5 years 8 months. **(C)** The height growth curve of the proband during the one-year follow-up. **(D)** The sequencing results of the ACAN mutations. Sequence chromatograms indicate the heterozygosity for a ACAN frameshift mutation (NM_013227.3:c.7230delT; NP_001356197.1:p.Phe2410Leufs*9) in the proband and his mother. The red arrow indicates the mutation. SDS, standard deviation scores; r-hGH, Recombinant human growth hormone; y, year; m, month; BA, bone age.

the growth plate. The structural or functional disturbances in aggrecan production could significantly affect cartilage, resulting in changes in linear growth. Statistically, in 1%–5% of idiopathic short stature patients, the short stature can be explained by *ACAN* mutations (Jee et al., 2018).

Herein, we enrolled a Han Chinese family with short stature and advanced bone age (BA). Using whole-exome sequencing (WES) and Sanger sequencing, a novel heterozygous frameshift mutation (NM_013227.3:c.7230delT; NP_001356197.1: p. Phe2410Leufs*9) of the *ACAN* gene was confirmed to be a genetic lesion in this family.

2 Case description

The proband, a 4-year-old boy, came to the Xiangya Hospital Central South University with his parents due to growth retardation (Figure 1A). Physical examination of the proband showed a height of 94.8 cm (−2.02 SDS), and his weight was 14.6 kg (−1.29 SDS). According to his parents' recollection, the height of the proband was 70 cm at 12 months of age (−2.40 SDS) and 85 cm at 30 months old (−2.03 SDS). Magnetic resonance imaging revealed a normal size and shape of the pituitary gland. The BA was found to be 5 years and 6 months, as revealed by the X-ray imaging of the left hand (Figure 1B).

We then investigated the proband's family history. His father (II-3) was 167 cm tall (≈−1.31 SDS), and his mother (II-2) was 142 cm tall (≈−3.23 SDS) (the target height of the proband was 161 ± 5 cm). His mother's age at menarche was approximately 12 years. Meanwhile, the heights of his grandmother (I-1), maternal uncle (II-1), and maternal aunt (II-4) were 145 cm (≈−2.78 SDS), 153 cm (≈−3.23 SDS), and 165 cm (≈0.28 SDS), respectively (Figure 1A). No other malformations, including intervertebral disc disease, osteochodritis dissecans or back pain, were observed in the affected members of the family.

At the request of his parents, recombinant human growth hormone (GH) treatment was initiated after the proband's second visit, at a dose of 50 μg/kg/day. During this time, the height of the proband increased to 99.0 cm at 4 years 10 months (−2.17 SDS), 101.3 cm at 5 years 1 month (−1.95 SDS), 104.0 cm at 5 years and 5 months (−1.78 SDS), and 105.5 cm at 5 years and 8 months (−1.76 SDS) (Figure 1C). The growth velocity was also changed from 5.42 cm/year to 6.00 cm/year, after 13 months of GH treatment. Meanwhile, his weight increased to 18.4 kg (−0.52 SDS) and the BA was 8 years during his latest visit, at the age of 5 years 8 months (Figure 1B).

To determine the genetic cause of the short stature in these patients, we extracted the genomic DNA of the proband (III-1) and his mother (II-2). Next, WES was performed to detect the candidate genes associated with short stature and advanced BA. The filtering strategies conformed to those described in our previous study (Huang et al., 2022). Briefly, (a) Variants within intronic, intergenic, and untranslated regions (UTRs), as well as synonymous single nucleotide variants (SNVs), were excluded from later analysis; (b) Common variants were excluded with minor allele frequency ≥0.01 by the 1000 Genomes database, Genome Aggregation Database (gnomAD), version 2.1.1 and an in-house exome database of

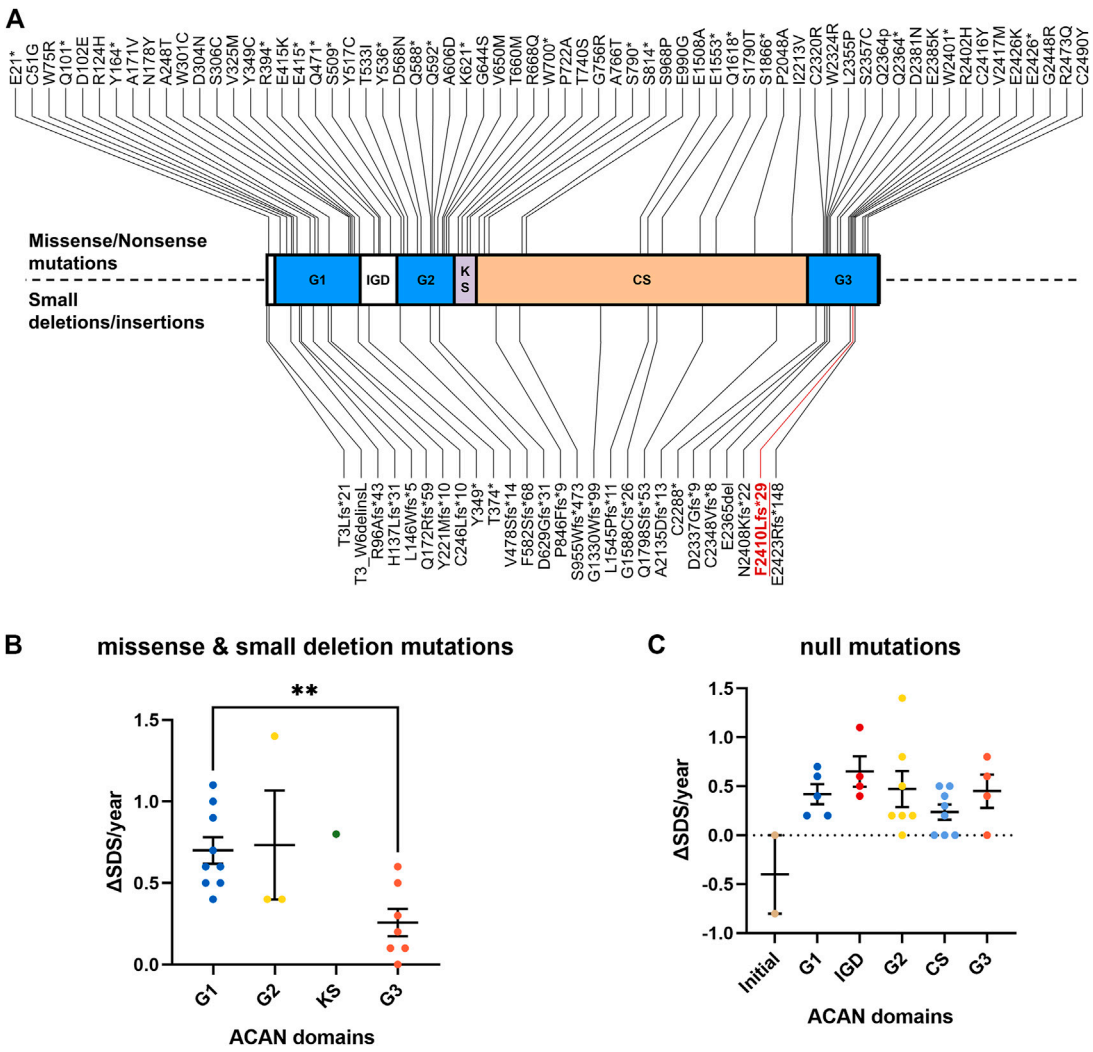


FIGURE 2
Overview of all known ACAN mutations. **(A)** The red font symbolizes the mutation identified in this study. **(B)** The increasing SDS for each year (ΔSDS/year) of patients with ACAN missense and small deletion mutations during GH treatment. **(C)** The increasing SDS for each year (ΔSDS/year) of patients with ACAN null mutations during GH treatment. G1, globular domain 1; IGD, interglobular domain; G2, globular domain 2; KS, keratin sulfate attachment domain; CS, chondroitin sulfate attachment domain; G3, globular domain 3; SDS, standard deviation scores. **p<0.01.

Berry Genomics; (c) Variants classified as Pathogenic/Likely Pathogenic variants, as well as variant of unknown significance (VUS) by American College of Medical Genetics (ACMG) classification (Richards et al., 2015), could be included; (d) Potential causative variants were then screened by a list of growth-associated genes (Supplementary Table S1) (Hauer et al., 2018); (e) Variants that were not existed in the affected family member (II-2) were excluded; (f) Bioinformatics analysis, including Polyphen-2, MutationTaster, SIFT, and CADD, were also carried out to predict the possible impacts of variants; (g) Accordingly, co-segregation analysis was conducted in the family (Supplementary Figure S1).

After data filtering of (a), (b), (c), (d), and (e), only a set of 7 variants in seven genes in the proband were identified (Supplementary Figure S1, Supplementary Tables S2, S3). Moreover, during further filtering of bioinformatic prediction and co-segregation analysis, a novel heterozygous frameshift mutation (NM_013227.3:c.7230delT; NP_001356197.1: p. Phe2410Leufs*9) in exon 16 of *ACAN* was identified in the proband (Figure 1D). No other significant variant related to short stature or advanced BA was identified. This variant of c.7230delT/p.Phe2410Leufs*9 in the *ACAN* gene was also predicted to be deleterious by multiple bioinformatics software (MutationTaster and CADD).

Sanger sequencing revealed that the c.7230delT/p.Phe2410Leufs*9 variant of *ACAN* co-segregated with the affected family members (I-1, II-1 and II-2). In addition, this novel *ACAN* mutation was not found in the 200 local control subjects from the control cohorts. Based on the ACMG guidelines, the variant of *ACAN* (c.7230delT; p. Phe2410Leufs*9) can be classified as likely pathogenic (PVS1_Moderate + PM1+PM2+PP3).

3 Discussion

Short stature is a common clinical manifestation in children. Genetic inheritance is the main determinant of stature (Grunauer and Jorge, 2018). The mutations in *ACAN* gene are a major cause of idiopathic short stature. Aggrecan is a chondroitin sulfate proteoglycan, that is also an important

TABLE 1 Summarized of reported *ACAN* patients and their prognosis data during GH treatment.

ACAN domain	AA change	Sex	Age at Start GH	Height SDS at Start GH	Height SDS at End GH	ΔSDS	ΔSDS/ year	Treat	PMID
—	(c.2T>C)	M	14.7	−4.1	−4.1	0.0	0.0	GH (49.5 μg/kg/d, 4.8 m)	36387899
—	(c.70 + 1G>A)	M	14.3	−4.9	−5.0	−0.1	−0.8	GH (43 μg/kg/d, 1.5 m)	35620465
G1	p.Ile73Serfs*12	F	7.0	−2.5	−2.1	0.4	0.4	GH (32.9 μg/kg/d, 12 m)	34653508
G1	p.Trp75Arg	M	5.5	−2.3	−1.3	1.0	1.0	GH (NA, 12 m)	27870580
G1	p.Trp75Arg	F	8.2	−1.6	−1.1	0.5	0.5	GH (NA, 12 m)	27870580
G1	p.Arg93Alafs*	M	8.6	−2.8	−2.6	0.2	0.2	GH (NA, 11 m) + AI (11 m)	27870580
G1	p.Arg93Alafs*	M	6.2	−2.0	−1.6	0.4	0.2	GH (NA, 30 m)	27870580
G1	p.Val94_Ile112del	M	3.4	−2.2	−3.3	1.1	1.1	GH (50–48 μg/kg/d, 12 m)	34922359
G1	p.Val94_Ile112del	M	5.0	−1.1	−1.8	0.7	0.7	GH (50–48 μg/kg/d, 12 m)	34922359
G1	p.Ala144Pro	M	8.2	−3.4	−2.9	0.5	0.6	GH (36.3 μg/kg/d, 9.6 m)	36387899
G1	p.Tyr164*	F	7.4	−3.1	−3.7	0.6	0.6	GH (50–48 μg/kg/d, 12 m)	34922359
G1	p.Ala180Thr	M	4.5	−2.0	−0.9	1.1	0.9	GH (53 μg/kg/d, 14 m)	35620465
G1	p.Trp203Cys	M	6.0	−4.3	−4.7	0.4	0.4	GH (50–48 μg/kg/d, 12 m)	34922359
G1	p.Val255_Glu352del	F	8.1	−3.0	−3.6	0.6	0.6	GH (50–48 μg/kg/d, 12 m)	34922359
G1	p.Thr282Ile	F	8.2	−3.8	−3.2	0.6	0.5	GH (33–39.6 μg/kg/d, 13.2 m)	36387899
G1	p.Trp301*	M	6.3	−2.9	−2.4	0.5	0.7	GH (33 μg/kg/d, 8.4 m)	36387899
IGD	p.Gly391Valfs*7	F	2.4	−2.6	−3.7	1.1	1.1	GH (50–48 μg/kg/d, 12 m)	34922359
IGD	p.Glu415*	F	8.5	−1.7	−0.8	0.9	0.4	GH (NA, 29 m) + GnRHa (13 m)	27870580
IGD	p.Val478Serfs*	M	7.4	−2.9	−1.4	1.5	0.5	GH (NA, 36 m)	27870580
IGD	(c.1429 + 1G>T)	F	6.0	−2.0	−1.4	0.6	0.6	GH (53 μg/kg/d, 12 m)	35620465
G2	p.Tyr517*	M	10.0	−3.0	−2.8	0.2	0.2	GH (28.6 μg/kg/d, 12 m)	34653508
G2	p.Tyr517*	F	7.2	−2.7	−2.5	0.2	0.2	GH (32.9 μg/kg/d, 12 m)	34653508
G2	p.Tyr536*	F	4.9	−3.7	−3.9	−0.2	0.0	GH (33 μg/kg/d, 97 m) + GnRHa (22 m) + GH (66 μg/kg/ d, 12 m)	27710243
G2	p.Tyr536*	M	11.9	−2.4	−1.6	0.8	0.2	GH (66 μg/kg/d, 42 m) + GnRHa (20 m) + AI (10 m)	27710243
G2	p.Ala606Asp	F	7.0	−0.7	−0.1	0.6	0.4	GH (56.7 μg/kg/d, 18 m)	35620465
G2	p.Ala606Asp	M	2.9	−2.1	−1.4	0.7	1.4	GH (40 μg/kg/d, 6 m)	35620465

(Continued on following page)

10.3389/fgene.2023.1101695

TABLE 1 (*Continued*) Summarized of reported *ACAN* patients and their prognosis data during GH treatment.

ACAN domain	AA change	Sex	Age at Start GH	Height SDS at Start GH	Height SDS at End GH	ΔSDS	ΔSDS/ year	Treat	PMID
G2	p.Ala606Asp	M	9.4	−2.9	−1.9	1.0	0.4	GH (NA, 30 m)	33471655
G2	p.Arg675*	M	3.6	−2.5	−3.9	1.4	1.4	GH (50–48 μg/kg/d, 12 m)	34922359
G2	p.Arg675*	F	9.7	−3.5	−4.3	0.8	0.8	GH (50–48 μg/kg/d, 12 m)	34922359
G2	p.Arg675*	M	6.0	−3.8	−3.0	0.8	0.5	GH (56.1–62.7 μg/kg/d, 18 m)	36387899
KS	p.Gly756Arg	M	3.9	−2.9	−2.2	0.7	0.8	GH (NA, 11 m)	33471655
CS	p.Val1464*	F	10.8	−2.7	−2.6	0.1	0.0	GH (1.0–1.6 mg/d, 46.8 m) + GnRHa (14.4 m)	34456977
CS	p.Val1464*	F	8.2	−2.6	−2.6	0.0	0.0	GH (0.7–1.0 mg/d, 40.8 m)	34456977
CS	p.Gly1588fs	M	12.3	−2.7	−2.5	0.2	0.0	GH (33 μg/kg/d, 76 m) + GnRHa (26 m)	27710243
CS	p.Ile1686Metfs*13	M	12.8	−2.7	−2.6	0.1	0.2	GH (53 μg/kg/d, 5 m) + letrozole (2.5 mg qd)	35620465
CS	p.Leu 1815fs	M	4.8	−4.4	−3.1	1.3	0.5	GH (48 μg/kg/d, 30 m)	33606014
CS	p.Gly 1861fs	F	9.5	−2.9	−2.3	0.6	0.4	GH (48 μg/kg/d, 19 m)	33606014
CS	p.Pro2059Leufs*2	M	6.6	−3.3	−3.0	0.3	0.5	GH (56.1–62.7 μg/kg/d, 7.92 m)	36387899
CS			11.8	−2.5	−2.4	0.1	0.3	GH (49.5 μg/kg/d, 3.6 m)	36387899
G3	p.Leu2355Pro	F	12.0	−3.1	−3.0	0.1	0.1	AI (24 m) + GH (NA, 18 m)	27870580
G3	p.Gln2364Pro	M	6.4	−3.7	−3.8	0.0	0.0	GH (60 μg/kg/d, 91 m)	29769040
G3	p.Gln2364Pro	F	5.6	−2.1	−1.1	1.0	0.3	GH (40–50 μg/kg/d, 48 m) + GnRHa (30 m)	29769040
G3	p.Gln2364Pro	M	7.8	−0.9	−0.3	0.6	0.1	GH (50 μg/kg/d, 60 m) + GnRHa (8 m)	29769040
G3	p.Gln2364*	M	11.7	−2.7	−2.8	−0.1	0.0	GH (66 μg/kg/d, 66 m) + GnRHa (24 m) + AI (13 m) + AI (7 m)	27710243
G3	p.Trp2401*	F	5.2	−1.1	−1.5	0.4	0.4	GH (50–48 μg/kg/d, 12 m)	34922359
G3	p.Trp2401*	F	6.9	−2.3	−2.9	0.6	0.6	GH (50–48 μg/kg/d, 12 m)	34922359
G3	p.Phe2410Leufs*9	M	4.6	−2.6	−1.8	0.8	0.8	GH (50 μg/kg/d, 13 m)	This study
G3	p.Val2417Met	F	5.5	−1.2	−0.6	0.6	0.2	GH (NA, 44 m)	27870580
G3	p.Arg2489Gly	M	7.4	−2.1	−1.7	0.4	0.5	GH (43 μg/kg/d, 9 m)	35620465
G3	p.Cys2490Tyr	M	15.1	−4.5	−4.4	0.1	0.6	GH (NA, 2 m) + GnRHa (2 m)	33471655

G1, globular domain 1; IGD, interglobular domain; G2, globular domain 2; KS, keratin sulfate attachment domain; CS, chondroitin sulfate attachment domain; G3, globular domain 3; M, male; F, female; GH, growth hormone; SDS, standard deviation score; d, day; m, month(s); NA, not applicable; GnRHa, gonadotropin-releasing hormone analog; AI, aromatase inhibitor.

10.3389/fgene.2023.1101695

component of the ECM. Aggrecan is encoded by *ACAN*. Animal experiments shown that dysfunction of the chick and mouse aggrecan can lead to severe chondrodystrophies with perinatal lethality (Li et al., 1993). Moreover, the expression levels of important regulatory genes, including *Col10a1*, *Sox9*, and *Ihh*, were sharply reduced in *ACAN* knockout mice (Lauing et al., 2014). In humans, mutations in the *ACAN* gene are known to cause chondrodysplasia and inherited short stature syndromes associated with accelerated bone maturation (Gibson and Briggs, 2016; Kim et al., 2020; Uchida et al., 2020; Wei et al., 2021). At present, only 65 missense/non-sense and 26 small deletion/insertion variants in the *ACAN* gene that are associated with short stature have been recorded in the Human Gene Mutation Database (HGMD; data retrieved in November 2021).

Aggrecan contains three globular domains (G1 and, G2 at the N-terminus and G3 at the C-terminal), one keratan sulfate domain, and one chondroitin sulfate domain. It also contains two epidermal growth factor-like motifs, one C-type lectin (CLD), and one complement regulatory protein-like motif in the G3 domain (Tompson et al., 2009; Liang et al., 2020). Previous studies have shown that CLD may mediate the interaction between aggrecan and tenascin in coordination with Ca^{2+} (Lundell et al., 2004; Diao and Tajkhorshid, 2008). Mutations of the *ACAN* gene (p.Phe2410Leufs*9), which was also located in the CLD domain of G3 (Figure 2A), may lead to the dysfunction of these interactions. Noticeably, Gkourogianni et al. also described an *ACAN* non-sense mutation (p.Trp2401*) that located in the CLD domain and led to short stature and early-onset osteoarthritis (Gkourogianni et al., 2017). Our study may provide additional evidence for the role of CLD in cartilage development.

In this study, we used WES to identify a novel p. Phe2410Leufs*9 frameshift mutation in exon 16 of *ACAN* in a family with short stature and advanced BA. Treatment with GH was performed since his age of 4 years and 7 months. However, the effectiveness of this treatment was limited during the first year of therapy (Figure 1C). Since GH facilitates growth by promoting the release of insulin-like growth factor 1, thus inducing chondrocyte proliferation and hypertrophy (Xu et al., 2018), treatment for a longer time and/or combination therapy may need to restore the normal ECM assembly and improve the height of the patient (van der Steen et al., 2017; Lin et al., 2021).

To clarify the relationship between *ACAN* mutations and the outcome of GH treatment, we reviewed all previously reported *ACAN* patients and their prognosis data during GH treatment (Table 1). A total of 49 *ACAN* patients was included. Most of these *ACAN* patients were firstly cued with different dose of GH (from 30 to 66 μg/kg/day) at age of 5–9. The increasing SDS for each year (ΔSDS/year) were calculated after their GH treatment. The range of ΔSDS/year was from −0.8 to 1.4 and the duration of GH treatment was from 1.5 months to 97 months. As we summarized, for missense and small deletion mutations, patients with *ACAN* mutations in G3 domain showed significant lower ΔSDS/year compared with other ACAN domains (Figure 2B). Especially the cases in Xu et al. (Xu et al., 2018) (Gln2364Pro) and Gkourogianni et al. (Gkourogianni et al., 2017) (Leu2355Pro), the ΔSDS/year were 0.0 or 0.3 even with a relative high dose GH (>40 μg/kg/day) and long duration time of treatment (91 or 18 months, respectively) (Table 1). When we compared the therapeutic differences between non-sense and frameshift mutations in each domain, most of these patients showed a similar GH sensitivity (Figure 2C). We also compared the ΔSDS/year data from who received GH combined with or without gonadotropin-releasing hormone analog (GnRHa). However, due to the limited cases, we found no difference. An example from Xu et al. (Xu et al., 2018) showed the ΔSDS/year only have a slight changed in the same *ACAN* mutation (Gln2364Pro) patients that treatment with or without GnRHa (Table 1). Which may suggest the little effect of GnRHa treatment in *ACAN* patients. However, much more clinical evidence is still demanded. All combined, the above information of point and null variants both suggesting G3, as the last functional domain of ACAN, may be critical in the development of short stature and GH treatment.

In summary, we report a novel mutation in *ACAN* (c.7230delT; p. Phe2410Leufs*9) in a Chinese family with short stature and advanced BA. Our analyses not only further confirmed the clinical diagnosis of short stature in the patients, but also expanded the mutation spectrum of *ACAN* and contributed to the genetic diagnosis and counseling of patients with short stature.

Data availability statement

The data that support the findings of this study are not publicly available because they contain information that could compromise the privacy of research participants; however, these data are available from the corresponding author upon reasonable request. The variation information has been submitted to ClinVar (SCV002819138).

Ethics statement

The studies involving human participants were reviewed and approved by the The Review Board of the Xiangya Hospital of Central South University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the individual (s), and minor (s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

Author contributions

XW and RX conceived and directed the project. HH collected the data and information. HH and JJ analyzed and interpreted the data. HH wrote the manuscript with the help of all the other authors. All authors contributed to the manuscript and approved the submitted version.

Funding

This study was supported by the National Natural Science Foundation of China (81970403); China Postdoctoral Science Foundation (2020TQ0363 and 2020M682598); the Natural Science Foundation of Hunan Province (2021JJ40992); the Emergency Project of Prevention and Control for COVID-19 of

Central South University (160260003), and the Youth Science Foundation of Xiangya Hospital (2021Q11).

Acknowledgments

We thank the patients and their family members for their participation in this study.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1101695/full#supplementary-material

References

Baron, J., Savendahl, L., De Luca, F., Dauber, A., Phillip, M., Wit, J. M., et al. (2015). Short and tall stature: A new paradigm emerges. *Nat. Rev. Endocrinol.* 11, 735–746. doi:10.1038/nrendo.2015.165

Diao, J., and Tajkhorshid, E. (2008). Indirect role of Ca2+ in the assembly of extracellular matrix proteins. *Biophys. J.* 95, 120–127. doi:10.1529/biophysj.108.129106

Finken, M., van der Steen, M., Smeets, C., Walenkamp, M., de Bruin, C., Hokken-Koelega, A., et al. (2018). Children born small for gestational age: Differential diagnosis, molecular genetic evaluation, and implications. *Endocr. Rev.* 39, 851–894. doi:10.1210/er.2018-00083

Gibson, B. G., and Briggs, M. D. (2016). The aggrecanopathies; an evolving phenotypic spectrum of human genetic skeletal diseases. *ORPHANET J. RARE Dis.* 11, 86. doi:10.1186/s13023-016-0459-2

Gkourogianni, A., Andrew, M., Tyzinski, L., Crocker, M., Douglas, J., Dunbar, N., et al. (2017). Clinical characterization of patients with autosomal dominant short stature due to aggrecan mutations. *J. Clin. Endocr. Metab.* 102, 460–469. doi:10.1210/jc.2016-3313

Grunauer, M., and Jorge, A. (2018). Genetic short stature. *GROWTH Horm. IGF Res.* 38, 29–33. doi:10.1016/j.ghir.2017.12.003

Hauer, N. N., Popp, B., Schoeller, E., Schuhmann, S., Heath, K. E., Hisado-Oliva, A., et al. (2018). Clinical relevance of systematic phenotyping and exome sequencing in patients with short stature. *Genet. Med.* 20, 630–638. doi:10.1038/gim.2017.159

Huang, H., Chen, Y., Jin, J., Du, R., Tang, K., Fan, L., et al. (2022). CSRP3, p.Arg122*, is responsible for hypertrophic cardiomyopathy in a Chinese family. *J. GENE Med.* 24, e3390. doi:10.1002/jgm.3390

Jee, Y. H., Andrade, A. C., Baron, J., and Nilsson, O. (2017). Genetics of short stature. *ENDOCRIN Metab. Clin.* 46, 259–281. doi:10.1016/j.ecl.2017.01.001

Jee, Y. H., Baron, J., and Nilsson, O. (2018). New developments in the genetic diagnosis of short stature. *Curr. Opin. Pediatr.* 30, 541–547. doi:10.1097/MOP.0000000000000653

Kim, T. Y., Jang, K. M., Keum, C. W., Oh, S. H., and Chung, W. Y. (2020). Identification of a heterozygous ACAN mutation in a 15-year-old boy with short stature who presented with advanced bone age: A case report and review of the literature. *Ann. Pediatr. ENDOCRIN* 25, 272–276. doi:10.6065/apem.1938198.099

Lauing, K. L., Cortes, M., Domowicz, M. S., Henry, J. G., Baria, A. T., and Schwartz, N. B. (2014). Aggrecan is required for growth plate cytoarchitecture and differentiation. *Dev. Biol.* 396, 224–236. doi:10.1016/j.ydbio.2014.10.005

Li, H., Schwartz, N. B., and Vertel, B. M. (1993). cDNA cloning of chick cartilage chondroitin sulfate (aggrecan) core protein and identification of a stop codon in the aggrecan gene associated with the chondrodystrophy, nanomelia. *J. Biol. Chem.* 268, 23504–23511. doi:10.1016/s0021-9258(19)49491-x

Liang, H., Miao, H., Pan, H., Yang, H., Gong, F., Duan, L., et al. (2020). Growth-promoting therapies may Be useful in short stature patients with nonspecific skeletal abnormalities caused by acan heterozygous mutations: Six Chinese cases and literature review. *Endocr. Pract.* 26, 1255–1268. doi:10.4158/EP-2019-0518

Lin, L., Li, M., Luo, J., Li, P., Zhou, S., Yang, Y., et al. (2021). A high proportion of novel ACAN mutations and their prevalence in a large cohort of Chinese short stature children. *J. Clin. Endocr. Metab.* 106, e2711–e2719. doi:10.1210/clinem/dgab088

Lundell, A., Olin, A. I., Morgelin, M., Al-Karadaghi, S., Aspberg, A., and Logan, D. T. (2004). Structural basis for interactions between tenascins and lectican C-type lectin domains: Evidence for a crosslinking role for tenascins. *STRUCTURE* 12, 1495–1506. doi:10.1016/j.str.2004.05.021

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of medical genetics and genomics and the association for molecular pathology. *Genet. Med.* 17, 405–424. doi:10.1038/gim.2015.30

Stavber, L., Hovnik, T., Kotnik, P., Lovrecic, L., Kovac, J., Tesovnik, T., et al. (2020). High frequency of pathogenic ACAN variants including an intragenic deletion in selected individuals with short stature. *Eur. J. Endocrinol.* 182, 243–253. doi:10.1530/EJE-19-0771

Tompson, S. W., Merriman, B., Funari, V. A., Fresquet, M., Lachman, R. S., Rimoin, D. L., et al. (2009). A recessive skeletal dysplasia, SEMD aggrecan type, results from a missense mutation affecting the C-type lectin domain of aggrecan. *Am. J. Hum. Genet.* 84, 72–79. doi:10.1016/j.ajhg.2008.12.001

Uchida, N., Shibata, H., Nishimura, G., and Hasegawa, T. (2020). A novel mutation in the ACAN gene in a family with autosomal dominant short stature and intervertebral disc disease. *Hum. GENOME Var.* 7, 44. doi:10.1038/s41439-020-00132-8

van der Steen, M., Pfundt, R., Maas, S., Bakker-van, W. W., Odink, R. J., and Hokken-Koelega, A. (2017). ACAN gene mutations in short children born SGA and response to growth hormone treatment. *J. Clin. Endocr. Metab.* 102, 1458–1467. doi:10.1210/jc.2016-2941

Wei, M., Ying, Y., Li, Z., Weng, Y., and Luo, X. (2021). Identification of novel ACAN mutations in two Chinese families and genotype-phenotype correlation in patients with 74 pathogenic ACAN variations. *Mol. Genet. GENOM Med.* 9, e1823. doi:10.1002/mgg3.1823

Xu, D., Sun, C., Zhou, Z., Wu, B., Yang, L., Chang, Z., et al. (2018). Novel aggrecan variant, p. Gln2364Pro, causes severe familial nonsyndromic adult short stature and poor growth hormone response in Chinese children. *BMC Med. Genet.* 19, 79. doi:10.1186/s12881-018-0591-z

frontiers | Frontiers in Genetics

TYPE Brief Research Report
PUBLISHED 23 March 2023
DOI 10.3389/fgene.2023.1116284

OPEN ACCESS

EDITED BY
Fang Fang Yu,
Zhengzhou University, China

REVIEWED BY
Yongyi Zou,
Jiangxi Maternal and Child Health Hospital, China
Jiangong Niu,
University of Texas MD Anderson Cancer Center, United States
Haitao Zhang,
National Institutes of Health (NIH), United States
Shuo Yang,
Zhengzhou University, China

*CORRESPONDENCE
Qianghua Xia,
qhxia@tmu.edu.cn

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

RECEIVED 05 December 2022
ACCEPTED 13 February 2023
PUBLISHED 23 March 2023

CITATION
Zhou W, Jia J, Qu H-Q, Ma F, Li J, Qi X, Meng X, Ding Z, Zheng G, Hakonarson H, Zeng X, Li J and Xia Q (2023), Identification of copy number variants contributing to hallux valgus. *Front. Genet.* 14:1116284. doi: 10.3389/fgene.2023.1116284

Identification of copy number variants contributing to hallux valgus

Wentao Zhou[1†], Jun Jia[2†], Hui-Qi Qu[3†], Feier Ma[1], Junyi Li[1], Xiaohui Qi[1], Xinyi Meng[1], Zhiyong Ding[4], Gang Zheng[5], Hakon Hakonarson[3,6,7], Xiantie Zeng[2], Jin Li[1] and Qianghua Xia[8,1]*

[1]Department of Cell Biology, The Province and Ministry Co-sponsored Collaborative Innovation Center for Medical Epigenetics, Key Laboratory of Immune Microenvironment and Disease (Ministry of Education), School of Basic Medical Sciences, Tianjin Medical University, Tianjin, China, [2]Department of Surgery of Foot and Ankle, Tianjin Hospital, Tianjin, China, [3]Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, United States, [4]Mills Institute for Personalized Cancer Care, Fynn Biotechnologies Ltd., Jinan, China, [5]National Supercomputer Center in Tianjin (NSCC-TJ), Tianjin, China, [6]Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA, United States, [7]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States, [8]Department of Bioinformatics, School of Basic Medical Sciences, Tianjin Medical University, Tianjin, China

Hallux valgus is a common form of foot deformity, and genetic factors contribute substantially to the pathogenesis of hallux valgus deformity. We conducted a genetic study on the structural variants underlying familial hallux valgus using whole exome sequencing approach. Twenty individuals from five hallux valgus families and two sporadic cases were included in this study. A total of 372 copy number variations were found and passed quality control filtering. Among them, 43 were only present in cases but not in controls or healthy individuals in the database of genomic variants. The genes covered by these copy number variations were enriched in gene sets related to immune signaling pathway, and cytochrome P450 metabolism. The hereditary CNVs demonstrate a dominant inheritance pattern. Two candidate pathogenic CNVs were further validated by quantitative-PCR. This study suggests that hallux valgus is a degenerative joint disease involving the dysregulation of immune and metabolism signaling pathways.

KEYWORDS

hallux valgus, whole exome sequencing, copy number variation, immune dysregulation, cytochrome p450 metabolism

Introduction

Hallux valgus (HV) deformity refers to the lateral deflection of the great toe at the first metatarsophalangeal joint (Perera et al., 2011; Hecht and Lin, 2014), which is the most common forefoot deformity often requires surgery with a prevalence rate of 23% in people aged 18–65 (CI: 16.3%–29.6%) (Nix et al., 2010). African Americans are more likely to have HV than people of European ancestry [adjusted odds ratio (aOR) = 2.01, 95% confidence interval [CI] = 1.39–2.92] (Golightly et al., 2012), and no significant difference has been found in the incidence of HV in other populations. Conservative treatment is feasible for patients with deformity but without symptoms or mild symptoms (Bayar et al., 2011). Although the symptoms can be alleviated, they can not be completely reversed. If the patient's pain persists, surgery is necessary. However, many surgical complications and poor

prognosis have brought great burden to individuals and their families (Sammarco and Idusuyi, 2001). Prevention and early intervention are therefore important.

The pathogenesis of HV deformity is complex (Mann and Coughlin, 1981). HV may be associated with inappropriate footwear. HV is 15 times more common in people who wear shoes than in those who don't (Perera et al., 2011), and shoes that tighten the front foot appear to be one of the leading causes of HV. Heredity is another important risk factor for HV development, especially in adolescents. A positive family history of HV has been reported in many studies (Lieberson and Mendes, 1991; Pique-Vidal et al., 2007), and study has found that HV susceptibility is related to genetic polymorphisms associated with arthritis (Perera et al., 2011; Hecht and Lin, 2014). Linear bone arrangement or static stabilizer relaxation due to heredity may also lead to HV deformity (Perera et al., 2011).

Genetic factors make substantial contribution to the pathogenesis of HV deformity. In a study of 350 patients with three generations family trees, 90% had at least one affected relative, which is consistent with autosomal dominant inheritance pattern (Pique-Vidal et al., 2007). The heritability of HV in European ancestry populations is between 0.29 and 0.89 (Hannan et al., 2013), while the rate of HV in Korea is ~0.51 (Lee et al., 2014). In a genome-wide association study (GWAS) on European ancestry population including 1786 cases of HV deformity and 2623 controls, genome-wide SNPs accounted for 50% of the phenotypic variance in males and 48% of the phenotypic variance in females (Hsu et al., 2015). The missing heritability may lie in the contribution of rare variants and structural variants, which have been underexplored in previous studies.

Structural variation is generally defined as a region of DNA of approximately 1 kb or larger that includes changes in copy number, chromosomal position, or orientation between individuals (Escaramis et al., 2015). A major class of genomic structural variation is copy number variation (CNV), which includes deletion and duplication of sequences (Ionita-Laza et al., 2009). In addition to single nucleotide polymorphisms (SNP) (Sachidanandam et al., 2001), CNVs are a major source of variation in the human genome, with significant effects on evolution and disease susceptibility (Conrad and Hurles, 2007). For example, at least 15% of neurodevelopmental diseases are caused by local dose imbalances in dozens of genes due to CNVs (Girirajan et al., 2011). Furthermore, CNVs make significant contributions to development of bone disorders and CNV analysis increases the diagnostic yield for these diseases. CNVs are significantly associated with osteoporosis (Yang et al., 2008; Costantini et al., 2018) and CNV is an important genetic factor for the etiology of fetal skeletal dysplasia (Wit et al., 2014; Bai et al., 2022). In addition, studies have shown that many CNVs confer greater disease risk than SNPs (Ionita-Laza et al., 2009). However, the contribution of CNVs to the pathogenesis and development of HV has not been investigated.

Aiming to explore the contribution of potential pathogenic CNVs to the development of HV deformity, we conducted the first study on the structural variants underlying familial HV using whole exome sequencing (WES) approach, which may be helpful for future risk prediction of HV.

Materials and methods

Samples and ethics statement

In this study, we recruited a total of 22 Chinese participants from 5 families (including 17 cases and 3 controls without any foot deformity) and two sporadic cases (Figure 1). The recruitment and the WES study were approved by Tianjin Hospital, and all participants provided written informed consent. The diagnosis of HV was made by clinical experts of the Foot and Ankle Surgery Group of Orthopaedic Branch of Chinese Medical Association and Foot and Ankle Surgery Professional Committee of Orthopaedic Physician Branch of Chinese Medical Association, according to the expert consensus (Foot and Ankle Working Committee et al., 2015). HV is diagnosed by combining the evaluation of clinical presentations, physical examination, auxiliary imaging examination and medical history. The severity of HV was determined by hallux valgus angle (HVA) and intermetatarsal angle (IMA) (normal: HVA < 16°, IMA < 10°; mild: HVA < 20°, IMA < 13°; moderate: 20° < HVA ≤ 40°, 13° < IMA ≤ 16°; severe: HVA > 40°,IMA > 16°). The single nucleotide variant analysis of the WES data of three families has been reported in our previous publication (Jia et al., 2021). With the addition of data from HV families and sporadic cases, we performed the current CNV study. The genetic genealogy of the 5 families shows that each proband has at least one first-degree relative as a HV patient.

Genomic DNA extraction and whole-exome sequencing

Genomic DNA was extracted from peripheral blood sample of each subject following standard procedures. The TargetSeqTM Enrichment Kit (iGeneTechTM) Human Exome Capture Kit was used for library construction. The Illumina sequencing platform was used for paired next-generation sequencing.

Quality control of sequencing result files

The Trim Galore software (Martin, 2011) (https://github.com/FelixKrueger/TrimGalore) was used to remove low-quality base (−q 25), limit maximum allowable error rate (default −e 0.1), remove reads < 36 nt (-length 36), remove double-ended overlap>3, and remove reads as a unit (-paired) from the raw fastQ files. Burrows-Wheeler Aligner (BWA)-MEM (version 0.7.17) (Li, 2013) was used to align the reads with the reference genome to obtain the SAM file.

After reordering SAM files were converted to BAM (Raw BAM) files and polymerase chain reaction (PCR) duplicates were marked using Picard (v1.91). SamTools (version 1.58) (Li et al., 2009) was used for quality control of the binary alignment graph files generated between them. Then, the Genome Analysis ToolKit (GATK) (version 3.8) (McKenna et al., 2010) was used to re-align indel regions and correct the base mass fraction. And after Base quality

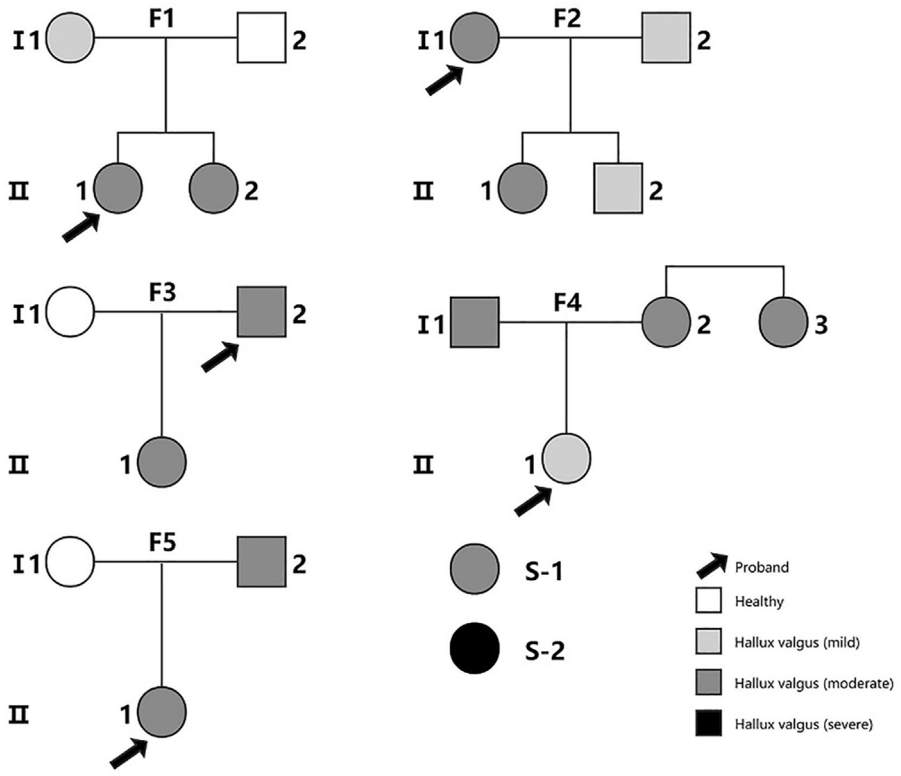


FIGURE 1
The family pedigree diagrams for all the subjects in the study. S-1 and S-2 are two sporadic cases. Families F1, F3 and F5 were included in our previous study of single nucleotide variants (Jia et al., 2021).The color indicates the severity of the patient's illness from mild to severe.

score recalibrations (BQSRs), we used verifyBamID (version 1.1.3) (Jun et al., 2012) to confirm that there was no contamination of the cross samples. The resulting BAM files were used for CNV calling.

CNVs calling by GATK and XHMM

The eXome-Hidden Markov Model (XHMM) (Fromer and Purcell, 2014) was used to call CNVs from filtered normalized targeted exome sequence data. The XHMM divided the chromosomal regions into three types: diploid, deletion, and duplication. The entire pipeline is plotted as a flowchart and shown in Figure 2.

CNVs annotation

We used "scan_regoin.pl" program in GenGen(Wang et al., 2007) to scan genomic features and to find the CNVs that overlapped with those in the database of genomic variants (DGV) (MacDonald et al., 2014) and filtered them out. The DGV database contains genomic structural variants over 50 bp in healthy individuals.

We used "scan_regoin.pl" program in GenGen (Wang et al., 2007) with the hg19_refGene and hg19_refLink files to annotate the CNV regions against the RefGene annotation to find the genes overlapped the CNV regions.

Then, we filtered out CNVs present among controls, and retained case-only CNVs that were repeated among the HV cases but not carried by any of the controls.

Quality control filtering of CNVs

We performed quality control filtering based on "Q_EXACT" and "Q_SOME" scores in the XCNV file according to the instruction of XHMM software. "Q_EXACT" and "Q_SOME" represent the phred-scaled quality of a CNV event along the entire interval and the same CNV event in the interval respectively.

Pathway enrichment of CNVs

The 67 genes overlapped with the 43 CNVs were used as input to the web-portal of OBAS (http://kobas.cbi.pku.edu.cn/) (Bu et al., 2021) and STRING (version 11.0) (https://string-db.org/) (Szklarczyk et al., 2021) for pathway enrichment and protein-protein interaction (PPI) analysis. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database and Search Tool for the Retrieval of Interacting Genes/Proteins database were used as the reference databases.

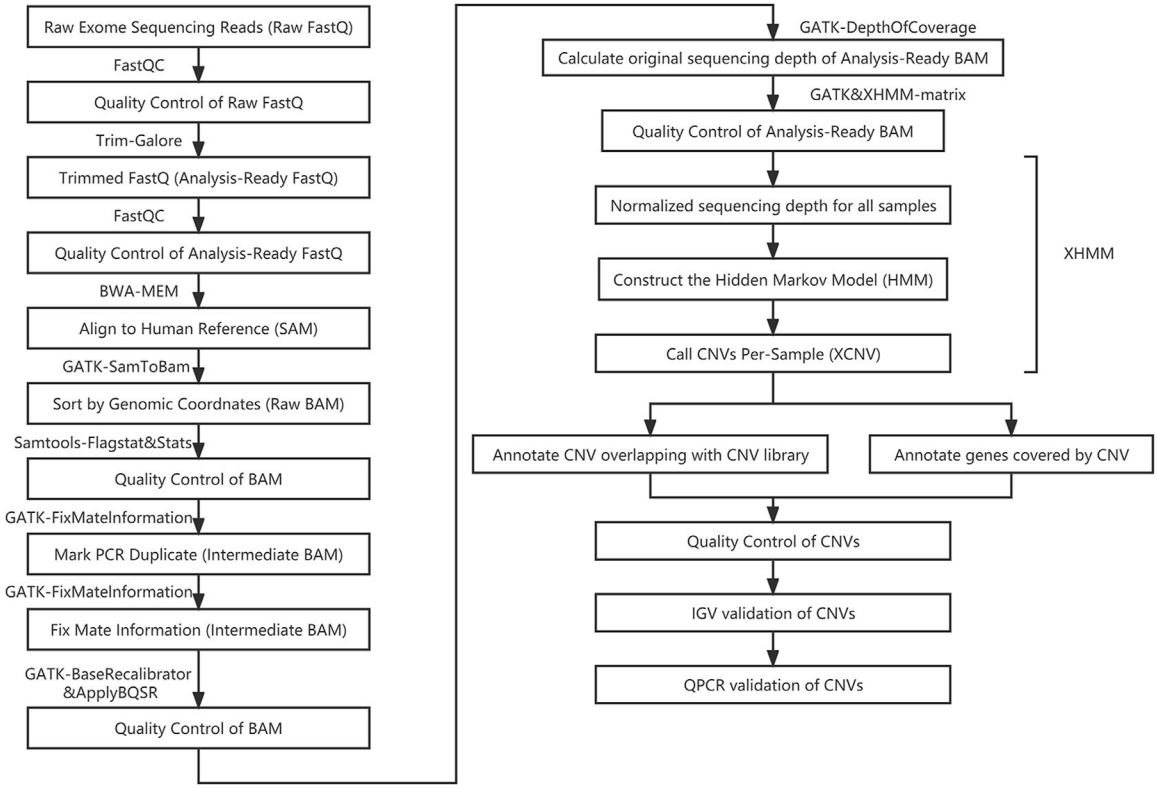


FIGURE 2
The pipeline of copy number variation analysis based on whole exome sequencing data.

Examination of CNVs using integrative genomics viewer

For candidate CNVs of interest, we further conducted visual inspection using the Integrative Genomics Viewer, based on the aligned bam files of the WES data of each CNV carrier and non-carriers within the same family.

qPCR validation of CNVs

For the potential pathogenic CNVs, we chose a ~100 bp fragment in each CNV for qPCR validation. The qPCR experiment was carried out with the Sybr Green I system. Each reaction in a 10 μL system contains 4 ng of genomic DNA, 10 μm per-primer, 2× ChamQ universal SYBR qPCR Master Mix, and DD-H_2O. Each sample was repeated for three times. The geometric mean values of the CT values of the control sequence GAPDH and the sample sequence were calculated to obtain △CT values for each sample. We calculated the 2^{-}–△△CT of each sample by taking the 2^{-}–△ CT of members without CNV in the family as the relative reference value. Finally, the existence of CNV was examined by the relative value of 2^{-}–△△CT of each sample in the pedigree.

Results

The identification of 43 CNVs related to HV

To assess the potential contribution of structural variants to HV, we carried out a CNV study based on the WES data of five HV families including 17 cases and 3 family members without HV and two sporadic cases (Supplementary Table S1). A total of 965 CNVs were detected based on the WES data and 372 passed quality control filtering (Supplementary Table S2).

Among these 372 CNVs, 325 CNVs were present only in cases, including 142 deletions and 183 duplications. They were not detected in the any of the three control samples. Furthermore, 51 CNVs (26 deletions and 25 duplications) were carried by more than one samples. In order to screen for potential pathogenic CNVs, we further filtered out any CNVs with 70% overlap with those carried by healthy individuals in the DGV. Then, 43 CNVs were retained for further analysis, covering the exons of the 67 genes, with 1 CNV co-occurring in four cases, 5 CNVs co-occurring in three cases, and 37 CNVs co-occurring in two cases (Supplementary Table S3).

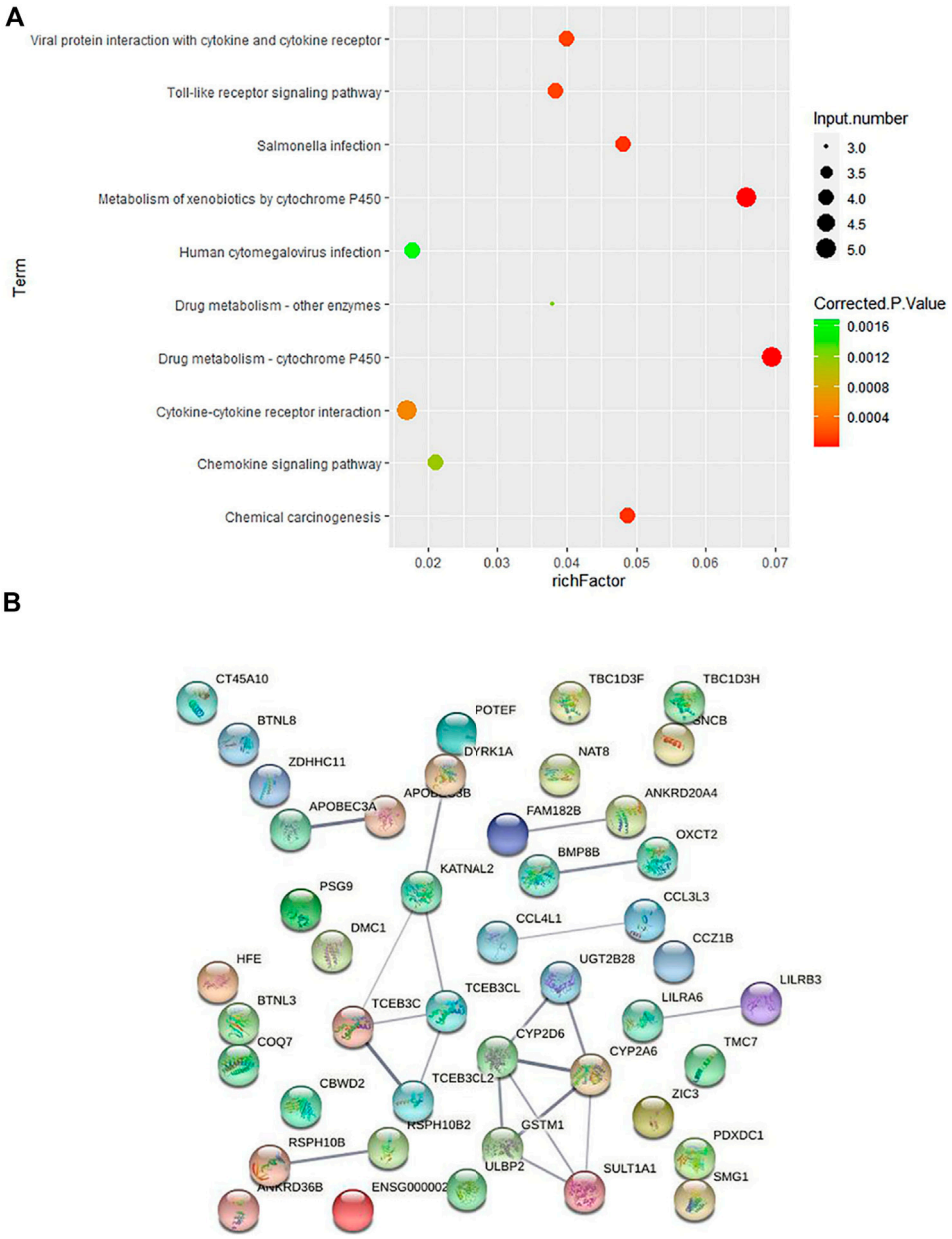


FIGURE 3
Pathway and protein-protein interaction analyses on the genes covered by 43 case-only CNVs **(A)** KEGG pathway enrichment analysis on the genes covered by the 43 case-only CNVs. The color of the dots represents the size of the corrected *p*-value, and the size of the dots represents the number of input genes contained in the corresponding pathway. **(B)** Protein-protein interaction analysis on the genes covered by 43 case-only CNVs. The nodes represent the proteins, and the lines between the nodes indicate the interactions between the proteins.

The involvement of immune and metabolism genes

In order to further explore how these CNVs may contribute to the pathogenesis of HV, we carried out pathway enrichment analysis among the 67 genes covered by case-only CNVs through KEGG PATHWAY database. Ten pathways were statistically significantly enriched ($p < 0.001$) for cytokine genes and genes involved in cytochrome P450 related metabolism and immunity/inflammation (Figure 3A; Supplementary Table S4), such as KEGG pathways "Drug metabolism—cytochrome P450", "Toll-like receptor signaling pathway" and "Cytokine-cytokine receptor interaction", suggesting their contribution to the pathogenesis of HV.

We then conducted the PPI analysis on these genes (Figure 3B). The PPI network also supports the consistent involvement of a group of immune genes and genes functioning in drug metabolism. In the PPI network, we observed the cluster formed by proteins involved in drug metabolism (UGT2B28, CYP2D6, CYP2A6, GSTM1, SULT1A1). In addition to the well-known involvement of cytokine genes CCL4L1 and CCL3L3 in immunity, several other genes also function in regulation of the immune system. HLA-H

TABLE 1 Candidate pathogenic CNVs.

SAMPLE	CNV	INTERVAL	Q_NON_DIPLOID	Q_SOME	GENE
F1-I-1,F1-II-1,F1-II-2	DEL	chr22: 42522498–42536739	99	99	CYP2D6, CYP2D7
F2-I-1,F2-II-1	DEL	chr6: 29855550–29895036	99	93	HCG4B, HLA-H

SAMPLE, sample ID; CNV, type of copy number variation (DEL or DUP); INTERVAL, genomic range of the called CNV; Q_NON_DIPLOID, Phred-scaled quality of not being diploid, i.e., DEL or DUP event in the interval; Q_SOME, Phred-scaled quality of some CNV event in the interval; GENE, gene name.

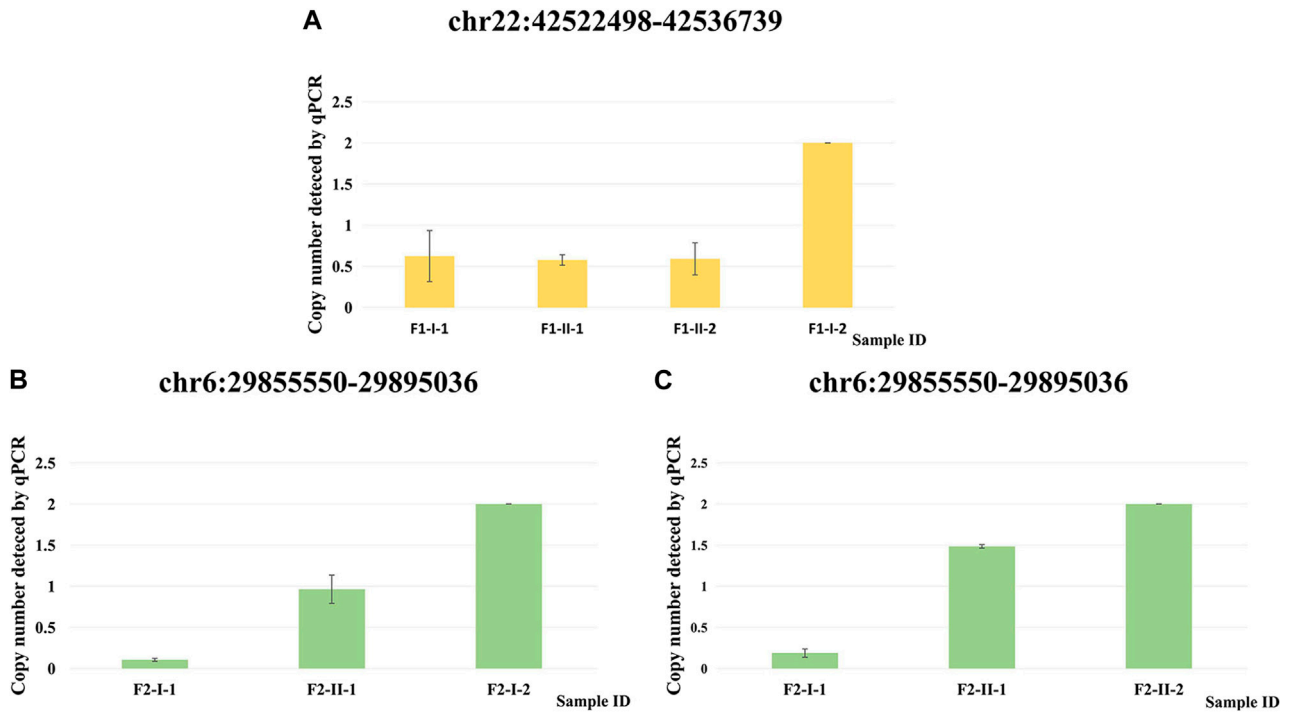


FIGURE 4
qPCR validation of two CNVs. **(A)** Copy number detected by qPCR at chr22: 42522498–42536739 for F1-I-1, F1-II-1, F1-II-2 and F1-I-2. **(B)** Copy number detected by qPCR at chr6: 29855550–29895036 for F2-I-1, F2-II-1 and F2-I-2. **(C)** Copy number detected by qPCR at chr6: 29855550–29895036 for F2-I-1, F2-II-1 and F2-II-2.

(HFE) is an important link between iron homeostasis and immune regulation (Bahram et al., 1999; Lio et al., 2002); DYRK1A gene plays a key role in regulating the differentiation of Th17 and regulatory T cells (Khor et al., 2015); and evidence suggests that genetic polymorphisms of CYP2D6 is associated with autoimmune bullous diseases induction (Rychlik-Sych et al., 2013).

Inheritance pattern of HV-related CNVs

To identify the potential disease-contributing CNVs, we examined the inheritance pattern of the CNVs which passed quality control filtering. We did not find any CNV that fits the recessive pattern. Multiple case-only CNVs consistent with dominant inheritance pattern were found in each family. For family 1, 3 and 5, dominant CNVs were referred to those present in cases (child and parent with HV) but absent in controls (parent without HV). For family 2 and family 4 in which all family members were affected individuals with HV, the CNVs consistent with dominant inheritance were those carried by the child and one of the parents (Supplementary Table S5). Our results are consistent with the hypothesis that HV has a dominant inheritance (Pique-Vidal et al., 2007), but certainly we can not exclude the possibility of recessive inheritance model as a large proportion of the genome were not assayed in our analysis by WES.

Validation of disease-contributing CNVs

Two inherited CNVs were of particular interest and they have the highest quality (Phred-scaled quality of Non-Diploidy = 99) in our analyses (Table 1). Chr22: 42522498–42536739 deletion occurs to all patients in family 1 and absent in the healthy member of the family. This segment of CNV covered genes CYP2D7 and CYP2D6 which are involved in bone excitation effects (Jayaraman et al., 2021) and metabolic pathways related to

human immunity (Effner et al., 2017). Chr6: 29855550–29895036 deletion appears to F2-I-1 and F2-II-1 the two individuals with moderate HV but not in family members F2-I-2 and F2-II-2 who have only mild HV condition. HCG4B and HLA-H genes are in this segment of CNV. HCG4B is a proinflammatory gene whose expression is positively correlated with the expression of HLA-A and may regulate the expression of HLA-A (Chen et al., 2017). HLA-A and HLA-H function in immune homeostasis (Jordier et al., 2019). Therefore evidence from literature suggests that these two CNVs may contribute to the pathogenesis of HV.

We performed qPCR (Supplementary Table S6) as an independent experimental approach to validate these two CNVs. The presence of these two CNVs in family 1 and family 2 respectively was consistent with the results from the CNV calling of WES data (Figure 4).

Discussion

This is the first study that identified the structural variants underlying familial HV using WES approach, The identification provided us new knowledge about the genetics basis of HV, which highlighted the potential contribution of immunity/inflammation, and cytochrome P450 metabolism to the etiology of HV.

Previous studies indicated that HV malformations are likely to be caused by a variety of contributing factors, including genetics, dorsiflexion of the first metatarsal, gastrocnemius equinus, abnormal foot mechanics, and joint hypermobility (Coughlin and Jones, 2007). Interestingly, patients with autoimmune arthritis conditions such as rheumatoid arthritis and psoriatic arthritis are more likely to develop HV malformations (Rojas-Villarraga et al., 2009; Hyslop et al., 2010).

The CNV chr6: 29855550–29895036 observed in both F2-I-1 and F2-II-1 covers the HCG4B and HLA-H genes. It has been found that HCG4B is a pro-inflammatory factor and may regulate the expression of HLA-A by acting on competing endogenous RNAs spongingmiR-122 and miR-1352 (Chen et al., 2017). HLA-A and HLA-H play essential roles in the immune system, including regulation of the innate immune system, class I MHC-mediated antigen processing and presentation, interferon gamma signaling, and antigen processing-cross-presentation (Allan et al., 2002; Nitschke et al., 2016; Jordier et al., 2019). In psoriatic arthritis and rheumatoid arthritis, HLA alleles have been shown to influence susceptibility and severity of these autoimmune joint disease (Mc et al., 2015).

The CNV chr 22: 42529569–42529669 seen in F1-I-1, F1-II-1 and F1-II-2 covers CYP2D6 and CYP2D7.These genes have roles in the cytochrome P450 related xenobiotic metabolic process and oxidation-reduction process (Lessard et al., 1997). Cytochrome P450 changes the structure of cell antigens by metabolizing foreign organisms to produce reactive oxygen species, and then initiates and/or amplifies the autoimmune phenomenon through molecular simulation of the autoimmune response to the original antigen (Namazi, 2009). In addition, studies have shown that secondary metabolites produced by CYP2D6-dependent biotransformation have bone excitation effects (Jayaraman et al., 2021).

These two CNVs may have pathological effects and lead to bone structure deformities through the dysregulation of the immune system and the process of reactive oxygen species production.

Previous genetic studies on HV, including our WES study (Jia et al., 2021), revealed the association of a few genetic variants with HV. Genome-wide association study in population of European ancestry found the sex-specific association of SNVs close to genes AXIN2, ESD, ANXA1 and MRGPRX3 at the marginal significance level (Hsu et al., 2015). A later GWAS with increased sample size identified the genome-wide significant association with SNP in gene CLCA2, being an expression quantitative trait locus for COL24A1 (Arbeeva et al., 2020). Candidate gene study in Chinese population suggested variants in VDR gene and SNP rs1800629 at the 5' of TNF gene are related to HV (Tao et al., 2018; Arbeeva et al., 2020). In our study, we did not find outstanding CNVs in these genes, which could be due to differences in study population, variant types. We also focused on different regions of the human genome from the previous GWAS (the coding exons *versus* the non-coding regions). However, similar underlying genetic mechanisms were revealed. As discussed by Hsu and colleagues, AXIN2 and ANXA1 have important functions in regulation of innate and adaptive immunity, suggesting the potential role of inflammation in HV (Hsu et al., 2015). Hsu and colleagues also discussed the involvement of the serine hydrolase gene ESD gene in the recycling of sialic acids which is related to decreased antioxidant levels. Interestingly, CYP2D6 and CYP2D7 also have roles in oxidation-reduction process. The SNP rs1800629 associated with HV in Chinese population is an eQTL SNP for genes including complement gene C4A and CYP21A1P a member of the Cytochrome P450 Family. Therefore, these identifications and results from our study converge on the similar immune and drug metabolism signaling.

Because of the importance of CNV in the etiology of complex human diseases, various CNV detection methods and computational algorithms have been developed. The traditional array-based approach including array-based comparative hybridization and SNP-array approaches can efficiently detect large CNVs with a relatively accurate rate (Gabrielaite et al., 2021). CNV calling from high-throughput sequencing data has become more and more commonly utilized in research and clinical diagnosis of complex human disorders. Though still being limited by short read length, high-throughput sequencing offers the potential to identify novel CNVs in the genomic regions with sparse coverage or lack of coverage by array probes. As the most frequently used sequencing approach, WES focuses on the exomes, the copy number alterations of which are more likely to have a pathogenic effect than the other regions of the human genome. Calling CNVs from WES data also has the advantages of detecting small CNVs ranging from 1–100 kb (Gordeeva et al., 2021). This approach has the drawback due to non-uniform read depth distribution between exons which makes it prone to false positive detection (Ozden et al., 2022). It is also subjected to missing CNVs in regions with few exons. As another sequencing-based CNV detection approach, CNV-seq technology can potentially provide a unbiased coverage of the human genome and more precise estimation of CNV breakpoints (Liang et al., 2014), however it also has the limitation of relative low read depth and resolution of 0.1 Mb in size which render it more suitable to detect large CNVs. Owing to the low cost, popularity in research use, abundant existing data, simultaneous usage for single nucleotide variant detection, and a number of well-tested CNV-calling algorithms, WES data are still a common choice of CNV analysis. With the gradually reduced sequencing cost and further improved CNV-calling algorithms, WES and CNV-seq are complimentary to each other and be combined in research

and clinical diagnosis to improve clinical efficiency and diagnostic yield of multisystem anomalies (Chen et al., 2022).

There are a few limitations in our study. There are certain technical drawbacks for CNV calling based on WES as aforementioned. Small sample size is another major limitation of our study. To optimize study power and increase the possibility of identifying genetic components of HV, we adopted a family-based approach (Borecki and Province, 2008) focusing on familial Hallux Valgus. To find the potentially pathogenic CNVs related to HV, we also filtered the case-only CNVs against all the CNVs carried by healthy individuals in the DGV database. Larger sample size will be needed to further investigate the substantial contribution of various CNV regions in the human genome (Zarrei et al., 2015) to the etiology of HV, and particularly will gain statistical evidence for CNVs of uncertain clinical significance.

In summary, from an unbiased genomic approach, this first study on the structural variants in familial HV gained us new insights into the HV pathogenesis, mediated by genetic variation of immunity and inflammation, and abnormal cytochrome P450 related metabolism. It provides evidence supporting the dominant inheritance mode of HV. Our results suggest that HV is a degenerative joint disease involving the dysregulation of immune system and metabolism system.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: https://db.cngb.org/ [Accession number CNP0004036].

Ethical statement

The study involving human participants was approved by the institutional review board of Tianjin Hospital. All the participants provided the informed consents.

Author contributions

Conceptualization: XZ, QX, and JL; methodology: QX, HH, and JL; data analysis: WZ, FM, JYL, XM, XQ, ZD, and GZ; resources, JJ and XZ; interpretation: WZ, JJ, and HQ; writing—original draft preparation, WZ, JJ, and FM; writing—review and editing, HQ, XM, JYL, HH, XZ, JL, and QX; supervision, HH, JL, and QX; All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by National Natural Science Foundation of China (No.81771769) and Natural Science Foundation of Guangdong Province (No. 2021A1515012392).

Acknowledgments

We thank all the participants enrolled in the study and the High-performance Computing Platform of Tianjin Medical University for computing support.

Conflict of interest

ZD is an employee of Fynn Biotechnologies Ltd.

The remaining authors declare no conflicts of interests and no competing interests of this work. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1116284/full#supplementary-material

References

Allan, D. S., Lepin, E. J., Braud, V. M., O'Callaghan, C. A., and Mcmichael, A. J. (2002). Tetrameric complexes of HLA-E, HLA-F, and HLA-G. *J. Immunol. Methods* 268, 43–50. doi:10.1016/s0022-1759(02)00199-0

Arbeeva, L., Yau, M., Mitchell, B. D., Jackson, R. D., Ryan, K., Golightly, Y. M., et al. (2020). Genome-wide meta-analysis identified novel variant associated with hallux valgus in Caucasians. *J. Foot Ankle Res.* 13 (1), 11. doi:10.1186/s13047-020-0379-1

Bahram, S., Gilfillan, S., Kuhn, L. C., Moret, R., Schulze, J. B., Lebeau, A., et al. (1999). Experimental hemochromatosis due to MHC class I HFE deficiency: Immune status and iron metabolism. *Proc. Natl. Acad. Sci. U. S. A.* 96 (23), 13312–13317. doi:10.1073/pnas.96.23.13312

Bai, Y., Sun, Y., Liu, N., Wang, L., Jiao, Z., Hou, Y., et al. (2022). Genetic analysis of 55 cases with fetal skeletal dysplasia. *Orphanet J. Rare Dis.* 17 (1), 410. doi:10.1186/s13023-022-02559-4

Bayar, B., Erel, S., Şimşek, İ. E., SüMER, E., and Bayar, K. (2011). The effects of taping and foot exercises on patients with hallux valgus: A preliminary study. *Turkish J. Med. Sci.* 41, 403–409. doi:10.3906/sag-0912-499

Borecki, I. B., and Province, M. A. (2008). Genetic and genomic discovery using family studies. *Circulation* 118 (10), 1057–1063. doi:10.1161/CIRCULATIONAHA.107.714592

Bu, D., Luo, H., Huo, P., Wang, Z., Zhang, S., He, Z., et al. (2021). KOBAS-I: Intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Res.* 49, W317–W325. doi:10.1093/nar/gkab447

Chen, X., Jiang, Y., Chen, R., Qi, Q., Zhang, X., Zhao, S., et al. (2022). Clinical efficiency of simultaneous CNV-seq and whole-exome sequencing for testing fetal structural anomalies. *J. Transl. Med.* 20 (1), 10. doi:10.1186/s12967-021-03202-9

Chen, X., Lu, X., Chen, J., Wu, D., Qiu, F., Xiong, H., et al. (2017). Association of nsv823469 copy number loss with decreased risk of chronic obstructive pulmonary disease and pulmonary function in Chinese. *Sci. Rep.* 7, 40060. doi:10.1038/srep40060

Clinical, G. G. O. M., Professional, C. F. P. D., and Group, O. G. D. P. (2019). Expert consensus on the application of low-depth whole genome sequencing in prenatal diagnosis.

Zhonghua Yi Xue Yi Chuan Xue Za Zhi 36 (4), 293–296. doi:10.3760/cma.j.issn.1003-9406.2019.04.001

Foot and Ankle Working Committee, Chinese Association of Orthopaedic Surgeons Orthopaedic Branch, Chinese Association of Orthopaedic Surgeons, Gao, P., Gui, J. C., Hong, J-J., Hong, J-S., Hu, Y., et al. (2015). Consensus on surgical management of hallux valgus from China. *Orthop. Surg.* 7, 291–296. doi:10.1111/os.12207

Conrad, D. F., and Hurles, M. E. (2007). The population genetics of structural variation. *Nat. Genet.* 39, S30–S36. doi:10.1038/ng2042

Costantini, A., Skarp, S., Kampe, A., Makitie, R. E., Pettersson, M., Mannikko, M., et al. (2018). Rare copy number variants in array-based comparative genomic hybridization in early-onset skeletal fragility. *Front. Endocrinol. (Lausanne)* 9, 380. doi:10.3389/fendo.2018.00380

Coughlin, M. J., and Jones, C. P. (2007). Hallux valgus: Demographics, etiology, and radiographic assessment. *Foot Ankle Int.* 28, 759–777. doi:10.3113/FAI.2007.0759

Effner, R., Hiller, J., Eyerich, S., Traidl-Hoffmann, C., Brockow, K., Triggiani, M., et al. (2017). Cytochrome P450s in human immune cells regulate IL-22 and c-Kit via an AHR feedback loop. *Sci. Rep.* 7, 44005. doi:10.1038/srep44005

Escaramis, G., Docampo, E., and Rabionet, R. (2015). A decade of structural variants: Description, history and methods to detect structural variation. *Brief. Funct. Genomics* 14, 305–314. doi:10.1093/bfgp/elv014

Fromer, M., and Purcell, S. M. (2014). Using XHMM software to detect copy number variation in whole-exome sequencing data. *Curr. Protoc. Hum. Genet.* 81, 7 23 1–21. doi:10.1002/0471142905.hg0723s81

Gabrielaite, M., Torp, M. H., Rasmussen, M. S., Andreu-Sanchez, S., Vieira, F. G., Pedersen, C. B., et al. (2021). A comparison of tools for copy-number variation detection in germline whole exome and whole genome sequencing data. *Cancers (Basel)* 13 (24), 6283. doi:10.3390/cancers13246283

Girirajan, S., Campbell, C. D., and Eichler, E. E. (2011). Human copy number variation and complex genetic disease. *Annu. Rev. Genet.* 45, 203–226. doi:10.1146/annurev-genet-102209-163544

Golightly, Y. M., Hannan, M. T., Dufour, A. B., and Jordan, J. M. (2012). Racial differences in foot disorders and foot type. *Arthritis Care Res. Hob.* 64 (11), 1756–1759. doi:10.1002/acr.21752

Gordeeva, V., Sharova, E., Babalyan, K., Sultanov, R., Govorun, V. M., and Arapidi, G. (2021). Benchmarking germline CNV calling tools from exome sequencing data. *Sci. Rep.* 11 (1), 14416. doi:10.1038/s41598-021-93878-2

Hannan, M. T., Menz, H. B., Jordan, J. M., Cupples, L. A., Cheng, C. H., and Hsu, Y. H. (2013). High heritability of hallux valgus and lesser toe deformities in adult men and women. *Arthritis care & Res.* 65, 1515–1521. doi:10.1002/acr.22040

Hecht, P. J., and Lin, T. J. (2014). Hallux valgus. *Hallux valgus* 98, 227–232. doi:10.1016/j.mcna.2013.10.007

Hsu, Y. H., Liu, Y., Hannan, M. T., Maixner, W., Smith, S. B., Diatchenko, L., et al. (2015). Genome-wide association meta-analyses to identify common genetic variants associated with hallux valgus in Caucasian and African Americans. *J. Med. Genet.* 52 (11), 762–769. doi:10.1136/jmedgenet-2015-103142

Hyslop, E., Mcinnes, I. B., Woodburn, J., and Turner, D. E. (2010). Foot problems in psoriatic arthritis: High burden and low care provision. *Ann. Rheum. Dis.* 69, 928. doi:10.1136/ard.2009.111971

Ionita-Laza, I., Rogers, A. J., Lange, C., Raby, B. A., and Lee, C. (2009). Genetic association analysis of copy-number variation (CNV) in human disease pathogenesis. *Genomics* 93, 22–26. doi:10.1016/j.ygeno.2008.08.012

Jayaraman, S., Reid, J. M., Hawse, J. R., and Goetz, M. P. (2021). Endoxifen, an estrogen receptor targeted therapy: From bench to bedside. *Endocrinology* 162 (12), bqab191. doi:10.1210/endocr/bqab191

Jia, J., Li, J., Qu, H., Li, M., Zhang, S., Hao, J., et al. (2021). New insights into hallux valgus by whole exome sequencing study. *Exp. Biol. Med. (Maywood)* 246 (14), 1607–1616. doi:10.1177/15353702211008641

Jordier, F., Gras, D., De Grandis, M., D'Journo, X. B., Thomas, P. A., Chanez, P., et al. (2019). HLA-H: Transcriptional activity and HLA-E mobilization. *Front. Immunol.* 10, 2986. doi:10.3389/fimmu.2019.02986

Jun, G., Flickinger, M., Hetrick, K. N., Romm, J. M., Doheny, K. F., Abecasis, G. R., et al. (2012). Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. *Am. J. Hum. Genet.* 91, 839–848. doi:10.1016/j.ajhg.2012.09.004

Khor, B., Gagnon, J. D., Goel, G., Roche, M. I., Conway, K. L., Tran, K., et al. (2015). The kinase DYRK1A reciprocally regulates the differentiation of Th17 and regulatory T cells. *Elife* 4, e05920. doi:10.7554/eLife.05920

Lee, C. H., Lee, S., Kang, H., Jung, D. E., Song, Y. M., Lee, K., et al. (2014). Genetic influences on hallux valgus in Koreans: The healthy twin study. *Twin Res. Hum. Genet.* 17, 121–126. doi:10.1017/thg.2014.10

Lessard, E., Fortin, A., Belanger, P. M., Beaune, P., Hamelin, B. A., and Turgeon, J. (1997). Role of CYP2D6 in the N-hydroxylation of procainamide. *Pharmacogenetics* 7, 381–390. doi:10.1097/00008571-199710000-00007

Li, H. (2013). "Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM (preprint),". arXiv:1303.3997.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Sequence Alignment/Map format SAMtools* 25, 2078–2079. doi:10.1093/bioinformatics/btp352

Liang, D., Peng, Y., Lv, W., Deng, L., Zhang, Y., Li, H., et al. (2014). Copy number variation sequencing for comprehensive diagnosis of chromosome disease syndromes. *J. Mol. Diagn* 16 (5), 519–526. doi:10.1016/j.jmoldx.2014.05.002

Lieberson, S., and Mendes, D. G. (1991). *Congenital hallux valgus*. NJ: SLACK Incorporated Thorofare.

Lio, D., Balistreri, C. R., Colonna-Romano, G., Motta, M., Franceschi, C., Malaguarnera, M., et al. (2002). Association between the MHC class I gene HFE polymorphisms and longevity: A study in Sicilian population. *Genes Immun.* 3 (1), 20–24. doi:10.1038/sj.gene.6363823

Macdonald, J. R., Ziman, R., Yuen, R. K., Feuk, L., and Scherer, S. W. (2014). The database of genomic variants: A curated collection of structural variation in the human genome. *Nucleic acids Res.* 42, D986–D992. doi:10.1093/nar/gkt958

Mann, R. A., and Coughlin, M. J. (1981). Hallux valgus-etiology, anatomy, treatment and surgical considerations. *Clin. Orthop. Relat. Res.* 157, 31–41. doi:10.1097/00003086-198106000-00008

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* 17, 10. doi:10.14806/ej.17.1.200

Mc, A. A., Flatley, B., Pennington, S. R., and Fitzgerald, O. (2015). Early biomarkers of joint damage in rheumatoid and psoriatic arthritis. *Arthritis Res. Ther.* 17, 141. doi:10.1186/s13075-015-0652-z

Mckenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi:10.1101/gr.107524.110

Namazi, M. R. (2009). Cytochrome-P450 enzymes and autoimmunity: Expansion of the relationship and introduction of free radicals as the link. *J. Autoimmune Dis.* 6, 4. doi:10.1186/1740-2557-6-4

Nitschke, K., Luxenburger, H., Kiraithe, M. M., Thimme, R., and Neumann-Haefelin, C. (2016). CD8+ T-cell responses in hepatitis B and C: The (HLA-) A, B, and C of hepatitis B and C. *Dig. Dis.* 34, 396–409. doi:10.1159/000444555

Nix, S., Smith, M., and Vicenzino, B. (2010). Prevalence of hallux valgus in the general population: A systematic review and meta-analysis. *J. Foot Ankle Res.* 321, 21. doi:10.1186/1757-1146-3-21

Ozden, F., Alkan, C., and Cicek, A. E. (2022). Polishing copy number variant calls on exome sequencing data via deep learning. *Genome Res.* 32 (6), 1170–1182. doi:10.1101/gr.274845.120

Perera, A. M., Mason, L., and Stephens, M. M. (2011). The pathogenesis of hallux valgus. *J. Bone Jt. Surg. Am.* 93, 1650–1661. doi:10.2106/JBJS.H.01630

Pique-Vidal, C., Sole, M. T., and Antich, J. (2007). Hallux valgus inheritance: Pedigree research in 350 patients with bunion deformity. *J. Foot Ankle Surg.* 46, 149–154. doi:10.1053/j.jfas.2006.10.011

Rojas-Villarraga, A., Bayona, J., Zuluaga, N., Mejia, S., Hincapie, M. E., and Anaya, J. M. (2009). The impact of rheumatoid foot on disability in Colombian patients with rheumatoid arthritis. *BMC Musculoskelet. Disord.* 10, 67. doi:10.1186/1471-2474-10-67

Rychlik-Sych, M., Baranska, M., Waszczykowska, E., Torzecka, J. D., Zebrowska, A., and Skretkowicz, J. (2013). Genetic polymorphisms of CYP2D6 oxidation in patients with autoimmune bullous diseases. *Postepy Dermatol Alergol.* 30 (4), 211–217. doi:10.5114/pdia.2013.37030

Sachidanandam, R., Weissman, D., Schmidt, S. C., Kakol, J. M., Stein, L. D., Marth, G., et al. (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409, 928–933. doi:10.1038/35057149

Sammarco, G. J., and Idusuyi, O. B. (2001). Complications after surgery of the hallux. *Clin. Orthop. Relat. Research®* 391, 59–71. doi:10.1097/00003086-200110000-00008

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., et al. (2021). The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49, D605–D612. doi:10.1093/nar/gkaa1074

Tao, T., Jiang, Y., Li, W., Li, Y., Du, J., and Gui, J. (2018). Association of vitamin D receptor gene TaqI, BsmI, FokI, and ApaI polymorphisms and susceptibility to hallux valgus in the Chinese population. *J. Foot Ankle Surg.* 57 (4), 753–758. doi:10.1053/j.jfas.2018.01.007

Wang, K., Li, M., and Bucan, M. (2007). Pathway-based approaches for analysis of genomewide association studies. *Am. J. Hum. Genet.* 81, 1278–1283. doi:10.1086/522374

Wit, J. M., van Duyvenvoorde, H. A., van Klinken, J. B., Caliebe, J., Bosch, C. A., Lui, J. C., et al. (2014). Copy number variants in short children born small for gestational age. *Horm. Res. Paediatr.* 82 (5), 310–318. doi:10.1159/000367712

Yang, J., Wang, J., Liang, X., Zhao, H., Lu, J., Ma, Q., et al. (2019). Relationship between genetic polymorphisms of the TNF gene and hallux valgus susceptibility. *Genet. Test. Mol. Biomarkers* 23 (6), 380–386. doi:10.1089/gtmb.2018.0269

Yang, T. L., Chen, X. D., Guo, Y., Lei, S. F., Wang, J. T., Zhou, Q., et al. (2008). Genome-wide copy-number-variation study identified a susceptibility gene, UGT2B17, for osteoporosis. *Am. J. Hum. Genet.* 83 (6), 663–674. doi:10.1016/j.ajhg.2008.10.006

Zarrei, M., Macdonald, J. R., Merico, D., and Scherer, S. W. (2015). A copy number variation map of the human genome. *Nat. Rev. Genet.* 16 (3), 172–183. doi:10.1038/nrg3871

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 23 May 2023
DOI 10.3389/fgene.2023.1169417

OPEN ACCESS

EDITED BY
Fang Fang Yu,
Zhengzhou University, China

REVIEWED BY
Peipei Pan,
University of California, San Francisco, United States
Zhiguang Ping,
Zhengzhou University, China

*CORRESPONDENCE
Xiong Guo,
guox@xjtu.edu.cn
Mikko J. Lammi,
mikko.lammi@umu.se

†These authors have contributed equally to this work

RECEIVED 19 February 2023
ACCEPTED 18 April 2023
PUBLISHED 23 May 2023

CITATION
Wang S, Wang Y, Li X, Yuan L, Guo X and Lammi MJ (2023), ATAC-seq reveals the roles of chromatin accessibility in the chondrocytes of Kashin–Beck disease compared with primary osteoarthritis. *Front. Genet.* 14:1169417. doi: 10.3389/fgene.2023.1169417

ATAC-seq reveals the roles of chromatin accessibility in the chondrocytes of Kashin–Beck disease compared with primary osteoarthritis

Sen Wang[1†], Yuanji Wang[2†], Xingyu Li[3], Linlin Yuan[1], Xiong Guo[1*] and Mikko J. Lammi[4*]

[1]School of Public Health, Health Science Center, Xi'an Jiaotong University, Xi'an, Shaanxi, China, [2]Department of Pharmacy, The First Affiliated Hospital of Northwest University, Xi'an, Shaanxi, China, [3]Department of Ophthalmology, Shaanxi Eye Hospital, Xi'an People's Hospital (Xi'an Fourth Hospital), Affiliated Guangren Hospital, School of Medicine, Xi'an Jiaotong University, Xi'an, Shaanxi, China, [4]Department of Integrative Medical Biology, University of Umeå, Umeå, Sweden

Objective: This study aimed to investigate the roles of accessible chromatin in understanding the different pathogeneses between Kashin–Beck disease (KBD) and primary osteoarthritis (OA).

Methods: Articular cartilages of KBD and OA patients were collected, and after tissue digestion, primary chondrocytes were cultured *in vitro*. Assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) was performed to compare the accessible chromatin differences of chondrocytes between KBD and OA groups. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were executed for the promoter genes. Then, the IntAct online database was used to generate networks of significant genes. Finally, we overlapped the analysis of differentially accessible region (DAR)-associated genes and differentially expressed genes (DEGs) obtained from whole-genomic microarray.

Results: We obtained 2,751 total DARs, which contained 1,985 loss and 856 gain DARs and belonged to 11 location distributions. We obtained 218 motifs associated with loss DARs, 71 motifs associated with gain DARs, 30 motif enrichments of loss DARs, and 30 motif enrichments of gain DARs. In total, 1,749 genes are associated with loss DARs, and 826 genes are associated with gain DARs. Among them, 210 promoter genes are associated with loss DARs, and 112 promoter genes are associated with gain DARs. We obtained 15 terms of GO enrichment and 5 terms of KEGG pathway enrichment from loss DAR promoter genes, and 15 terms of GO enrichment and 3 terms of KEGG pathway enrichment from gain DAR promoter genes. We obtained CAPN6 and other 2 overlap genes from loss DARs-vs-down DEGs, AMOTL1 from gain DARs-vs-down DEGs, EBF3 and other 12 overlap genes from loss DARs-vs-up DEGs, and ADARB1 and other 10 overlap genes from 101 gain DARs-vs-up DEGs. These overlap genes were built into 4 gene interaction networks.

Conclusion: *FGF7, GPD1L, NFIB, RUNX2,* and *VCAM1* were the overlapped genes from the DAR-associated genes and DEGs. These genes were associated with the abnormal chondrocyte function, which may play crucial roles in different processes between KBD and OA in the way of accessible chromatin.

KEYWORDS
Kashin–Beck disease, ATAC-seq, cartilage, chondrocyte, osteoarthritis

Introduction

Kashin–Beck disease (KBD) is one kind of endemic, chronic, and degenerative osteoarthropathy, which is characterized by apoptosis and necrosis of chondrocytes, degradation of cartilage, and extracellular matrix (Guo et al., 2014). Today, it is observed primarily in China, North Korea, and Siberia (Wang et al., 2020). In China, there were still a lot of people living in KBD endemic areas, 98.25 million people were at 379 counties, and 32.5 million people were at 2,047 villages, according to the 2022 Health Statistical Yearbook of China.

KBD is one special type of primary osteoarthritis (OA) with similar clinical manifestations at the late stage. The similar clinical manifestations include morning stiffness, joint pain, enlarged and shortened joints, and dysfunctional and deformed (Wang et al., 2013) joints. According to the diagnostic criteria of KBD of China (WS/T 207-2010 criteria), there are three grades (I–III) in KBD diagnosis. Grade I: thickening of finger joints; grade II: on the basis of grade I, deformity with short fingers (toes); and grade III: on the basis of grade II, deformity with short limb or short stature.

Assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) is a relatively new technique, which is used to define regions of accessible chromatin in cells (Galang et al., 2020). ATAC-seq owns several advantages compared with other sequencing technologies in testing the chromatin of DNase-seq (Merrill et al., 2022). It uses hyperactive Tn5 transposase to cut and ligate adapters simultaneously for high-throughput sequencing at regions of increased accessibility of chromatin. The mapping of inserted ends in genome-wide high-throughput sequencing allows multidimensional analysis of the regulatory landscape of chromatin in a relatively simple protocol for a standard sample of 50,000 cells in a few hours (Buenrostro et al., 2015).

To the best of our knowledge, we have found no other study related to the chromatin accessibility in the chondrocytes of KBD. This is the first time when such features of KBD have been analyzed using the ATAC-seq method. This study was performed to study chromatin accessibility signatures in the chondrocytes of KBD. We propose that changes of accessible chromatin in the chondrocytes may play crucial roles in understanding the different processes between KBD and OA.

Materials and methods

Cartilage sample collection

Articular cartilage samples were collected from 11 KBD patients (six females and five males, 55–69 years old, seven grade II and four grade III, respectively) and 11 primary OA patients (six females and five males, 54–67 years old, respectively) matched based on age and sex. The KBD patients from Xi'an city, Yongshou, and Linyou counties in Shaanxi Province of China were diagnosed as grades II and III based on the diagnostic criteria of KBD of China (WS/T 207-2010 criteria). Meanwhile, the primary OA patients came from the non-KBD-endemic areas at Xi'an in Shaanxi Province, according to the Western Ontario and McMaster Universities OA Index (WOMAC). All KBD and OA patients underwent knee arthroplasty or debridement in hospitals. Cartilage samples were stored in liquid nitrogen at −196°C after washing with physiological saline. This investigation was approved by the Human Ethics Committee of Xi'an Jiaotong University.

In vitro culture of chondrocytes

The cartilage samples of the KBD and OA patients were cut and digested with 0.25% trypsin and 0.2% collagenase type II (Sigma, Germany). The cartilage was incubated in trypsin solution overnight at 37°C. Chondrocytes were cultured in a DMEM/F12 culture medium (Gibco, Grand Island, NY), which contained 10% fetal bovine serum (Gibco, Grand Island, NY) in a CO_2 incubator. Primary cells were used in all experiments.

ATAC-seq experiment and bioinformatics analysis

The chondrocytes obtained from two KBD patients (one female and one male, 55 and 62 years old, respectively) and two OA patients (one female and one male, 57 and 63 years old, respectively) were used in the ATAC-seq experiment. The protocol followed the one described previously (Buenrostro et al., 2015). The methods included five steps: cell lysis, transposition, amplification, sequencing, and peak calling. Briefly, an ice–cold cell lysis buffer was added to the chondrocyte suspension to release nuclei. Tn5 transposed DNA was purified using AMPure DNA magnetic beads and then amplified by PCR. The qualified library was sequenced on an Illumina NovoSeq platform (San Diego, United States) with the PE150 mode. A trimmed read was aligned to the reference genome using Bowtie 2 (Langmead and Salzberg, 2012). All peak calling was performed using MACS-23 (Zhang et al., 2008).

We detected the differentially accessible regions (DARs) between the two sets of samples by comparing the signal values of ATAC-seq data. The DiffBind package in the R program was applied to test potential DARs in consensus peaksets. Significant DARs were considered peaks with a false discovery rate (FDR) < 0.05 and $\log_2$ fold enrichment >0.5. The ggplot2 package in the R program was utilized to build the volcano plot of DARs. Subsequently, we assessed the clustering of DAR signal values of individual samples between groups. pheatmap was applied to construct the clustering diagram of DARs. We completed the annotation and distribution statistics for the location of DARs in the genome. The ChIPseeker package in the R program was applied to plot the location distribution pie charts of DARs.

Motif scans and enrichment analyses of loss DARs (diminished accessible regions) and gain DARs (enhanced accessible regions)

were performed in the JASPAR database using the MEME Suite motifs and kept only when the *p*-value was <0.01 (Heinz et al., 2010).

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed for promoter genes. The clusterProfiler package in the R program was applied to build the GO enrichment diagram (*p*-value < 0.05).

Whole-genomic microarray analysis

Chondrocytes from four KBD (two females and two males, 58–68 years old, respectively) and four primary knee OA (two females and two males, 55–67 years old) patients were divided into four pairs for the microarray analysis. Total RNA was isolated using a minikit (Agilent, CA), following the protocol. Then, 1% agarose gel electrophoresis was performed to check the RNA integrity. Total RNA samples were converted into cDNA using the Amino Allyl MessageAmp aRNA Kit (Ambion, TX). Agilent 44 K human whole-genome oligonucleotide microarray and Feature Extraction 9.3 software were then utilized. A fold change of ≥2.5 or ≤0.4 was considered significant.

qRT-PCR analysis

In order to validate the microarray results of different experimental groups, we used significant DEGs named DOK5, TRPC6, and EPHA3 as target genes for qRT-PCR. The chondrocytes of five KBD patients (three females and two males, 55–69 years old) and five OA patients (three females and two males, 54–66 years old) were used in the experiments. Total RNA was extracted in the same way with the one in the microarray analysis.

We applied the ABI 7500 RT-PCR system was utilized for the qRT-PCR analysis. We used glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as an endogenous control. The primers and probes NM_018431 [for DOK5], NM_004621 [for TRPC6], and NM_005233 [for EPHA3] were chosen. Comparative analysis between the qRT-PCR data and the microarray data were implemented using paired *t*-tests in SPSS 19.0.

Transcriptome combined analysis between the ATAC-seq signal and microarray signal

We utilized a transcriptome-combined analysis to elucidate the relationship between the levels of the ATAC-seq signal and gene expression. After dividing the genes of each sample into five classes according to their expression levels, we observed the distribution trends of ATAC-seq signals in the gene body and their upstream and downstream 3-kb regions at different expression levels. Finally, we compared the DAR-associated genes and the DEGs to get the overlapped ones. The DEGs indicated the differential expression genes in chondrocytes between the KBD and OA groups.

The IntAct online database (https://www.ebi.ac.uk/intact/home) was used to build the gene networks of the overlap genes mentioned previously. We only selected the direct interaction from the *Homo sapiens* gene.

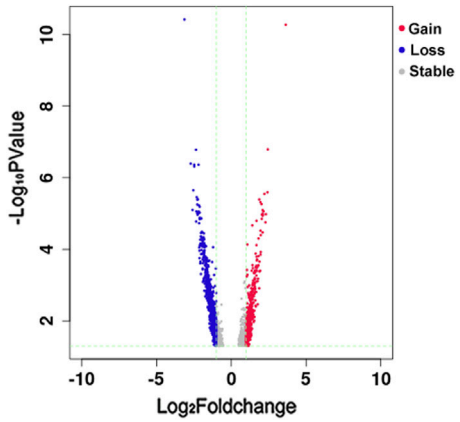


FIGURE 1
Volcano plot of Loss and Gain DARs. The horizontal axis is the $\log_2$ fold change of DARs, and the vertical axis is the $-\log_{10}$ *p*-value of DARs. Blue represents loss DAR, red represents gain DAR, and gray represents the area with no significant difference. The green trend lines are X = −1 (representing fold change = −2), X = 1 (representing fold change = 2), and Y = 1.3 (representing *p*-value = 0.05).

Results

Total peaks and DARs

We obtained 41,193 total peaks in the KBD group and 51,900 in the OA group, of which 37,359 were overlapping peaks between the two groups. By comparing the signal values of the ATAC-seq data between the two sets of samples, we obtained 2,751 total DARs (1,985 loss DARs and 856 gain DARs in the KBD group compared with the OA one). The volcano plot of loss and gain DARs, based on the $\log_2$ fold change and *p*-value of DARs, is shown in Figure 1.

Location distribution of DARs

The location distribution percentages of gain and loss DARs are shown in Figure 2. The locations included promoters, 5′UTR, 3′UTR, and six other items. No matter whether the loss DAR (Figure 2A) or gain DAR (Figure 2B) was analyzed, the trend of the location ranking was consistent.

The results of motif enrichment

We obtained 218 motifs associated with loss DARs and 71 motifs associated with gain DARs. Subsequently, we built 30 motif enrichments of loss DARs and 30 motif enrichments of gain DARs (shown in Figures 3A, B). The motifs are ranked by their adjusted *p*-value.

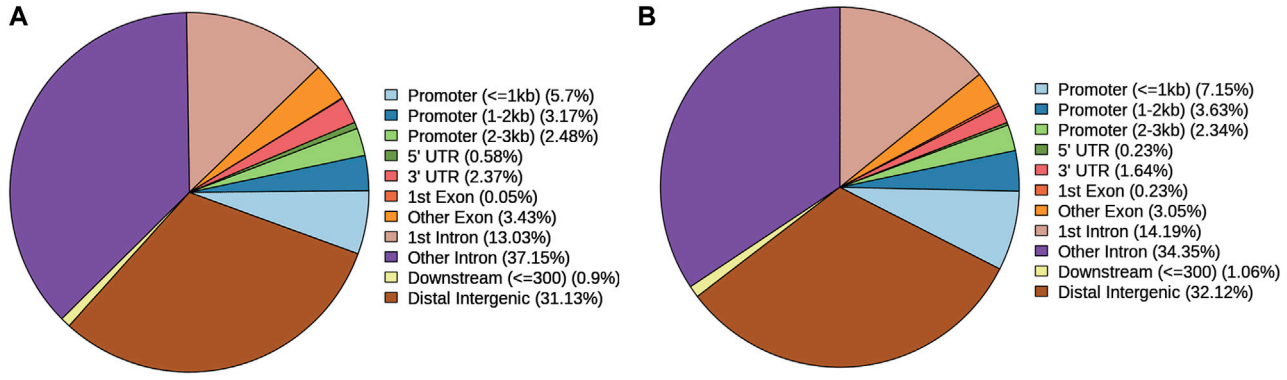


FIGURE 2
Location distribution percentages of **(A)** loss DARs and **(B)** gain DARs. Icons are the location types in the genome and their proportions of DARs.

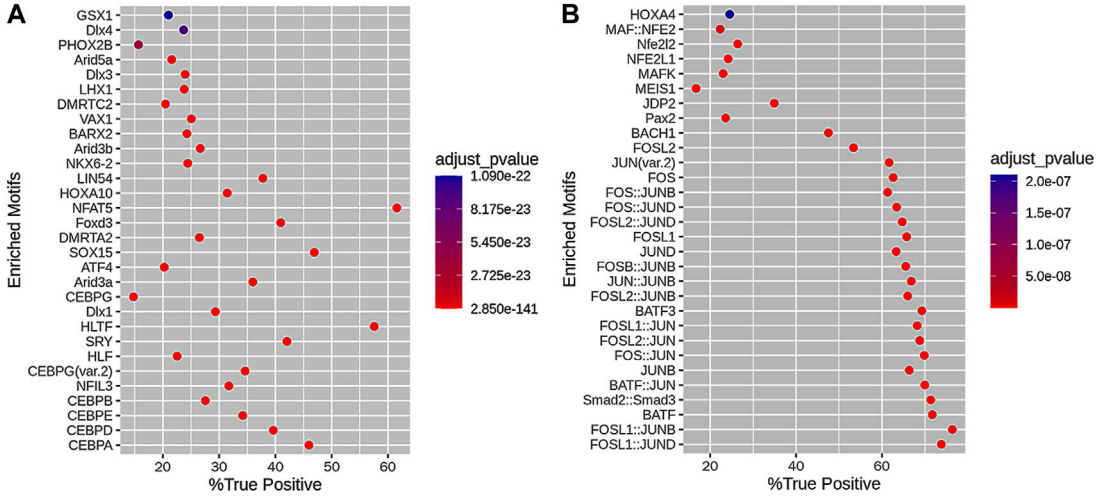


FIGURE 3
Motif enrichment results of **(A)** loss DARs and **(B)** gain DARs. The horizontal axis is the percentage of DARs with this motif in the DAR set, and the vertical axis is the motif. The color scale indicates the size of the adjusted *p*-value.

Annotation of DAR-associated genes

Between the KBD and OA groups, 1,749 genes were associated with loss DARs, and 826 genes, with gain DARs. Among them, 210 promoter genes associated with loss DARs and 112 promoter genes associated with gain DARs. We obtained 15 terms of GO enrichment for loss DAR promoter genes and 15 terms of GO enrichment for gain DAR promoter genes (shown in Figures 4A, B). The size and color of the circle represent the number of *p*-values and the gene number in GO.

Subsequently, we obtained five terms of the KEGG pathway enrichment of loss DAR promoter genes and three terms of the KEGG pathway enrichment of gain DAR promoter genes (shown in Figures 5A, B). The size and color of the circle represent the number of *p*-values and the gene number in the pathway.

Whole-genomic microarray analysis

The microarray analysis revealed 2,236 upregulated genes and 137 downregulated genes in the chondrocytes of KBD patients compared to the OA patients (shown in Supplementary Table S1).

qRT-PCR analyses

The genes *DOK5*, *TRPC6*, and *EPHA6* were selected for validation of the gene expression results in the microarray. It is confirmed that all three genes were upregulated in the KBD chondrocytes, consistent with the microarray data (shown in Supplementary Figure S1).

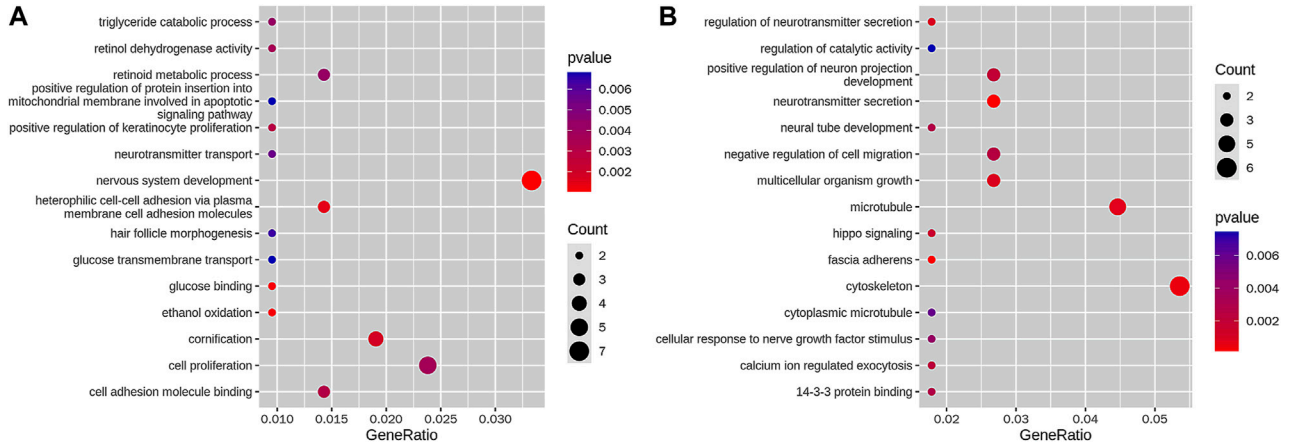


FIGURE 4
GO enrichment of **(A)** loss DAR promoter genes and **(B)** gain DAR promoter genes. The horizontal axis is a gene ratio, which is the ratio of the gene number in the GO term to the GO term in the gene of the species, and the vertical axis is the GO term; the color in the figure is the size of the *p*-value, and the circle is the size of the number of genes.

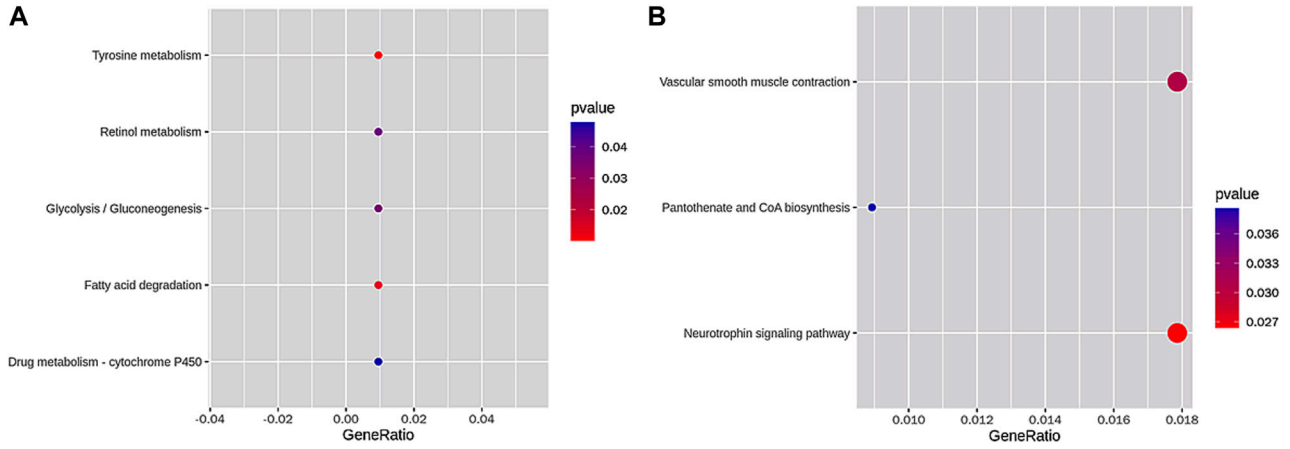


FIGURE 5
KEGG pathway enrichment of **(A)** loss DAR promoter genes and **(B)** gain DAR promoter genes. The horizontal axis is the gene ratio, which is the ratio of the gene number in the pathway to the pathways in the gene of the species. The vertical axis is the pathway. The color in the figure represents the size of the *p*-value, and the circle indicates the number of genes in the pathway.

Relationship between ATAC-Seq signal intensity and gene expression levels

From the overlap analysis of DAR-associated genes and DEGs, we obtained 3 overlapping genes from 207 loss DARs versus 732 downregulated DEGs (*CAPN6, EVL,* and *PTPN6*); 1 overlapping gene from 111 gain DARs versus 734 downregulated DEGs (*AMOTL1*); 13 overlapping genes from 197 loss DARs versus 5,646 upregulated DEGs (*EBF3, FGF7, FOLH1, GLRA3, JDP2, LGI2, MBNL1, NFIB, NLGN1, RUNX1T1, RUNX2, SPINK6,* and *VCAM1*), and 11 overlapping genes from 101 gain DARs versus 5,648 upregulated DEGs (*ADARB1, BMP2K, CALD1, DYNLT1, GPD1L, IGF2BP2, NRAP, PEG10, RIMS1, SLC22A3,* and *VNN1*).

Gene interaction networks of the overlap genes

We obtained four gene interaction networks from each overlapping gene group. The network containing *EVL* and *PTPN6* is shown in Figure 6A, while the network containing *AMOTL1* is shown in Figure 6B. The network containing *MBNL1, NLGN1, RUNX1T1,* and *RUNX2* is shown in Figure 6C,

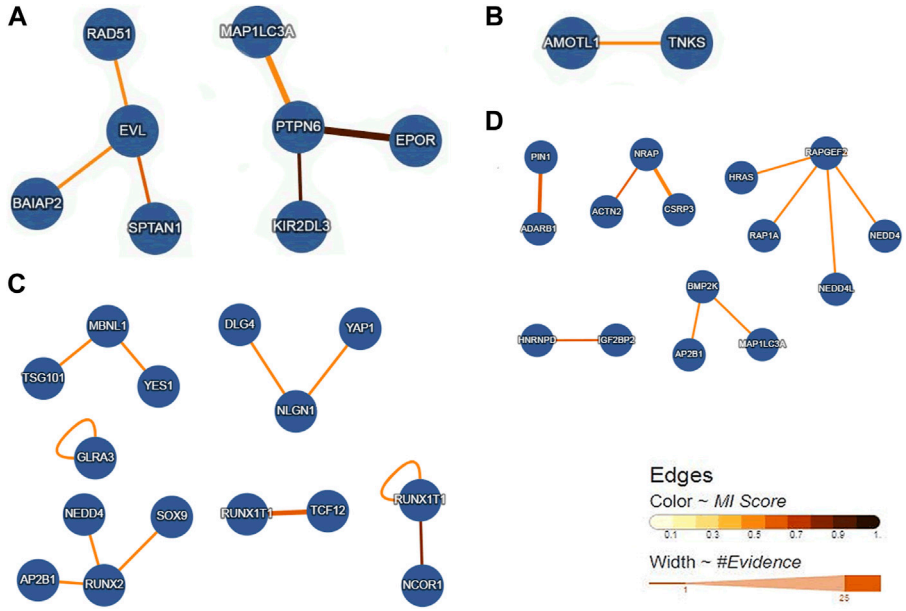


FIGURE 6
Gene interaction networks of the overlap genes. **(A)** Networks containing the overlap genes from Loss DAR-vs-down DEGs. **(B)** Networks including the overlap genes from gain DAR-vs-down DEGs. **(C)** Networks containing the overlap genes from loss DAR-vs-up DEGs. **(D)** Networks containing the overlap genes from gain DAR-vs-up DEGs.

and *ADARB1, BMP2K, IGF2BP2,* and *NRAP* can be seen in Figure 6D.

Discussion

This study revealed that *FGF7, GPD1L, NFIB, RUNX2,* and *VCAM1* were overlapping genes between the DAR-associated genes and DEGs. *FGF7, NFIB, RUNX2,* and *VCAM1* belonged to loss DAR and upregulated DEGs. *GPD1L* belonged to gain DARs and upregulated DEGs. These significant genes are associated with the abnormal chondrocyte function, which may play crucial roles in KBD in the way of accessible chromatin.

It has been shown that growth plate chondrocytes express the fibroblast growth factor 7 (FGF7) (Lazarus et al., 2007). FGF7 is also expressed significantly higher in OA meniscal cells than the normal meniscal cells (Sun et al., 2010). In this study, the relation of the FGF7 in comparison with loss DARs and upregulated DEGs was observed, indicating that its abnormal expression in the KBD is more serious than that in OA cells. The diminished accessible region may cause upregulation.

Articular cartilage is a tissue under a low oxygen tension, and the chondrocytes are adapted to the hypoxic condition. Specific hypoxia-inducible factors (HIFs) are involved in chondrogenesis, regulated by transcription factors such as SOX9 (Zhang et al., 2015). In this role, HIF-1α is an important regulator which contributes to the synthesis of the chondrocytes' extracellular matrix (Zhai et al., 2017). The gene *GPD1L* has been implicated in circ-0001946/miR-21/GPD1L/HIF-1α axis. *GPD1L* is a main regulator for the hydroxylation of HIF-1α, which can lead to the decrease of HIF-1α expression (Kelly et al., 2011). HIF-1α can protect the articular cartilage in the way of promoting the chondrocyte phenotype, supporting the adaptation to a hypoxic environment and maintaining the chondrocyte viability (Zhang et al., 2015). Although primary chondrocytes can resist heat shock protein (Hsp)-associated stress (Kaarniranta et al., 2001), HIF-1α-induced HSP70 overexpression can increase the expression of ECM genes and the cell viability and protect chondrocytes from apoptosis. It may indicate that *GPD1L* participates in the modulation of chondrocyte apoptosis (Tsuchida et al., 2014). Meanwhile, *GPD1L*-knockdown chondrocytes presented a higher apoptotic rate (Zhai et al., 2017). *GPD1L* was involved in gain DARs and upregulated DEGs. It appears to indicate that the damage to the KBD chondrocytes is more serious than the OA chondrocytes.

This study revealed that *NFIB* was involved in loss DARs and upregulated DEGs. Consistent with the transient activation of the *NFIB* gene in chondrogenesis, dominant-negative mutations in *NFIB* can interfere with chondrogenesis (Uchihashi et al., 2007). The overexpression of *NFIB* increased SOX9 and COL2A1 expression (Nagy et al., 2011). Previous findings indicated that the over-expression of *NFIB* can increase the expression of genes related to the synthesis of ECM and proliferation of chondrocytes. However, the over-expression of *NFIB* can decrease the expression of genes related to the degradation of ECM and differentiation of chondrocytes (Pan et al., 2021). *NFIB* was associated with loss DARs and upregulated DEGs. It may indicate the abnormal chondrocyte function of KBD is more serious than that of OA.

It is well known that *RUNX2* can regulate the proliferation and differentiation of chondrocytes that promote endochondral ossification of chondrocytes to a hypertrophic-like state (Zhang et al., 2022). It is an important transcription factor for the maturation of chondrocytes, which induces the expression of collagen X in the maturation process (Zhang et al., 2019). *RUNX2* was involved in loss DARs and upregulated DEGs. This suggests a stronger involvement of *RUNX* in regulating the chondrocyte function in KBD and OA.

VCAM1 is a known pro-inflammatory gene (Harasymowicz et al., 2021). Its influence for chondrocytes, which was tested by flow cytometry, includes inflammation and metabolic alterations (Ragni et al., 2020). In addition, *VCAM1* revealed the relation of loss DARs and upregulation of DEGs. It is interesting that the repression of *VCAM1* limited the inflammation and apoptosis in the murine OA model related to the PI3K/Akt pathway (Jiang et al., 2023).

In summary, a number of associations between DARs of chromatin and DEGs were revealed in this study. They may provide a new insight into understanding the different processes between KBD and OA. Further study is warranted to identify the roles of the finding in this study.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA953668.

Ethics statement

The studies involving human participants were reviewed and approved by the Human Ethics Committee, Xi'an Jiaotong University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

XG conceived and designed the research; SW performed most of the experiments and wrote the manuscript; YW analyzed the data on the experiments and edited the manuscript; XL prepared the figures; LY collected the samples of a cartilage and cultured the chondrocyte; ML interpreted results of the experiments and revised the language of the manuscript. All of the authors approved the final version of the manuscript.

Funding

This study was supported by the Natural Science Basic Research Program of Shaanxi (Province 2021JM-037).

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1169417/full#supplementary-material

References

Buenrostro, J. D., Wu, B., Chang, H. Y., and Greenleaf, W. J. (2015). ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* 109, 21.29.1–21.29.9. doi:10.1002/0471142727.mb2129s109

Galang, G., Mandla, R., Ruan, H., Jung, C., Sinha, T., Stone, N. R., et al. (2020). ATAC-seq reveals an Isl1 enhancer that regulates sinoatrial node development and function. *Circ. Res.* 127, 1502–1518. doi:10.1161/CIRCRESAHA.120.317145

Guo, X., Ma, W. J., Zhang, F., Ren, F. L., Qu, C. J., and Lammi, M. J. (2014). Recent advances in the research of an endemic osteochondropathy in China: Kashin-Beck disease. *Osteoarthr. Cartil.* 22, 1774–1783. doi:10.1016/j.joca.2014.07.023

Harasymowicz, N. S., Azfer, A., Burnett, R., Simpson, H., and Salter, D. M. (2021). Chondrocytes from osteoarthritic cartilage of obese patients show altered adiponectin receptors expression and response to adiponectin. *J. Orthop. Res.* 39, 2333–2339. doi:10.1002/jor.24993

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., et al. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* 38, 576–589. doi:10.1016/j.molcel.2010.05.004

Jiang, Y., Zhang, L., and Tian, H. (2023). MicroRNA-149 improves osteoarthritis via repression of VCAM-1 and inactivation of PI3K/AKT pathway. *Exp. Gerontol.* 174, 112103. doi:10.1016/j.exger.2023.112103

Kaarniranta, K., Holmberg, C. I., Lammi, M. J., Eriksson, J. E., Sistonen, L., and Helminen, H. J. (2001). Primary chondrocytes resist hydrostatic pressure-induced stress while primary synovial cells and fibroblasts show modified Hsp70 response. *Osteoarthr. Cartil.* 9, 7–13. doi:10.1053/joca.2000.0354

Kelly, T. J., Souza, A. L., Clish, C. B., and Puigserver, P. (2011). A hypoxia-induced positive feedback loop promotes hypoxia-inducible factor 1alpha stability through miR-210 suppression of glycerol-3-phosphate dehydrogenase 1-like. *Mol. Cell Biol.* 31, 2696–2706. doi:10.1128/MCB.01242-10

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi:10.1038/nmeth.1923

Lazarus, J. E., Hegde, A., Andrade, A. C., Nilsson, O., and Baron, J. (2007). Fibroblast growth factor expression in the postnatal growth plate. *Bone* 40, 577–586. doi:10.1016/j.bone.2006.10.013

Merrill, C. B., Pabon, M. A., Montgomery, A. B., Rodan, A. R., and Rothenfluh, A. (2022). Optimized assay for transposase-accessible chromatin by sequencing (ATAC-seq) library preparation from adult *Drosophila melanogaster* neurons. *Sci. Rep.* 12, 6043. doi:10.1038/s41598-022-09869-4

Nagy, A., Kenesi, E., Rentsendorj, O., Molnar, A., Szenasi, T., Sinko, I., et al. (2011). Evolutionarily conserved, growth plate zone-specific regulation of the matrilin-1 promoter: L-Sox5/Sox6 and nfi factors bound near TATA finely tune activation by Sox9. *Mol. Cell Biol.* 31, 686–699. doi:10.1128/MCB.00019-10

Pan, D., Qian, B., Zhao, D., and Yao, B. (2021). Nfib promotes chondrocyte proliferation and inhibits differentiation by mildly regulating Sox9 and its downstream genes. *Mol. Biol. Rep.* 48, 7487–7497. doi:10.1007/s11033-021-06767-2

Ragni, E., Perucca, O. C., De Luca, P., Mondadori, C., Vigano, M., Colombini, A., et al. (2020). Inflammatory priming enhances mesenchymal stromal cell secretome potential as a clinical product for regenerative medicine approaches through secreted factors and EV-miRNAs: The example of joint disease. *Stem Cell Res. Ther.* 11, 165. doi:10.1186/s13287-020-01677-9

Sun, Y., Mauerhan, D. R., Honeycutt, P. R., Kneisl, J. S., Norton, J. H., Hanley, E. J., et al. (2010). Analysis of meniscal degeneration and meniscal gene expression. *BMC Musculoskelet. Disord.* 11, 19. doi:10.1186/1471-2474-11-19

Tsuchida, S., Arai, Y., Takahashi, K. A., Kishida, T., Kubo, T., Honjo, K., et al. (2014). HIF-1α-induced HSP70 regulates anabolic responses in articular chondrocytes under hypoxic conditions. *J. Orthop. Res.* 32, 975–980. doi:10.1002/jor.22623

Uchihashi, T., Kimata, M., Tachikawa, K., Koshimizu, T., Okada, T., Ihara-Watanabe, M., et al. (2007). Involvement of nuclear factor I transcription/replication factor in the early stage of chondrocytic differentiation. *Bone* 41, 1025–1035. doi:10.1016/j.bone.2007.08.028

Wang, S., Duan, C., Zhang, F., Ma, W., and Guo, X. (2013). Regulatory gene networks and signaling pathways from primary osteoarthritis and Kashin-Beck disease, an endemic osteoarthritis, identified by three analysis software. *Gene* 512, 89–96. doi:10.1016/j.gene.2012.10.006

Wang, S., Gao, Z., Liu, H., Meng, P., Wu, C., Lammi, M., et al. (2020). Roles of glycoprotein glycosylation in the pathogenesis of an endemic osteoarthritis, Kashin–Beck disease, and effectiveness evaluation of sodium hyaluronate treatment. *Turk J. Med. Sci.* 50, 1028–1037. doi:10.3906/sag-1903-137

Zhai, X., Meng, R., Li, H., Li, J., Jing, L., Qin, L., et al. (2017). miR-181a modulates chondrocyte apoptosis by targeting glycerol-3-phosphate dehydrogenase 1-like protein (GPD1L) in osteoarthritis. *Med. Sci. Monit.* 23, 1224–1231. doi:10.12659/msm.899228

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., et al. (2008). Model-based analysis of ChIP-seq (MACS). *Genome Biol.* 9, R137. doi:10.1186/gb-2008-9-9-r137

Zhang, F. J., Luo, W., and Lei, G. H. (2015). Role of HIF-1 alpha and HIF-2 alpha in osteoarthritis. *Jt. bone, spine* 82, 144–147. doi:10.1016/j.jbspin.2014.10.003

Zhang, F., Lammi, M. J., Shao, W., Zhang, P., Zhang, Y., Wei, H., et al. (2019). Cytotoxic properties of HT-2 toxin in human chondrocytes: Could T(3) inhibit toxicity of HT-2? *Toxins (Basel)* 11, 667. doi:10.3390/toxins11110667

Zhang, D., Deng, X., Liu, Y., Zhang, Y., Wang, H., Zhang, M., et al. (2022). MMP-10 deficiency effects differentiation and death of chondrocytes associated with endochondral osteogenesis in an endemic osteoarthritis. *Cartilage* 13, 19476035221109226. doi:10.1177/19476035221109226

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 28 June 2023
DOI 10.3389/fgene.2023.1161817

OPEN ACCESS

EDITED BY
Qiling Yuan,
Xi'an Jiaotong University, China

REVIEWED BY
Michael D. Gallagher,
Massachusetts Institute of Technology, United States
Zhenwei Zhou,
Boston University, United States
Yuancheng Pan,
Xiamen University, China

*CORRESPONDENCE
Wayne Yuk-Wai Lee,
waynelee@cuhk.edu.hk

†These authors have contributed equally to this work

RECEIVED 16 February 2023
ACCEPTED 12 June 2023
PUBLISHED 28 June 2023

CITATION
Wu Z, Yang KG, Lam T-P, Cheng JCY, Zhu Z and Lee WY-W (2023), Genetic insight into the putative causal proteins and druggable targets of osteoporosis: a large-scale proteome-wide mendelian randomization study.
Front. Genet. 14:1161817.
doi: 10.3389/fgene.2023.1161817

Genetic insight into the putative causal proteins and druggable targets of osteoporosis: a large-scale proteome-wide mendelian randomization study

Zhichong Wu[1,2,3†], Kenneth Guangpu Yang[2,3,4,5,6†], Tsz-Ping Lam[3], Jack Chun Yiu Cheng[3], Zezhang Zhu[1,2,3] and Wayne Yuk-Wai Lee[2,3,4,5,6*]

[1]Division of Spine Surgery, Department of Orthopedic Surgery, Nanjing Drum Tower Hospital, The Affiliated Hospital of Nanjing University Medical School, Nanjing, China, [2]Musculoskeletal Research Laboratory, SH Ho Scoliosis Research Laboratory, Department of Orthopaedics and Traumatology, Faculty of Medicine, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China, [3]Joint Scoliosis Research Centre of the Chinese University of Hong Kong and Nanjing University, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China, [4]Prince of Wales Hospital, Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China, [5]Center for Neuromusculoskeletal Restorative Medicine, CUHK InnoHK Centres, Shatin, Hong Kong SAR, China, [6]Key Laboratory for Regenerative Medicine, School of Biomedical Sciences, Faculty of Medicine, Ministry of Education, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China

Background: Osteoporosis is a major causative factor of the global burden of disease and disability, characterized by low bone mineral density (BMD) and high risks of fracture. We aimed to identify putative causal proteins and druggable targets of osteoporosis.

Methods: This study utilized the largest GWAS summary statistics on plasma proteins and estimated heel BMD (eBMD) to identify causal proteins of osteoporosis by mendelian randomization (MR) analysis. Different GWAS datasets were used to validate the results. Multiple sensitivity analyses were conducted to evaluate the robustness of primary MR findings. We have also performed an enrichment analysis for the identified causal proteins and evaluated their druggability.

Results: After Bonferroni correction, 67 proteins were identified to be causally associated with estimated BMD (eBMD) ($p < 4 \times 10^{-5}$). We further replicated 38 of the 67 proteins to be associated with total body BMD, lumbar spine BMD, femoral neck BMD as well as fractures, such as RSPO3, IDUA, SMOC2, and LRP4. The findings were supported by sensitivity analyses. Enrichment analysis identified multiple Gene Ontology items, including collagen-containing extracellular matrix (GO:0062023, $p = 1.6 \times 10^{-10}$), collagen binding (GO:0005518, $p = 8.6 \times 10^{-5}$), and extracellular matrix structural constituent (GO:0005201, $p = 2.7 \times 10^{-5}$).

Conclusion: The study identified novel putative causal proteins for osteoporosis which may serve as potential early screening biomarkers and druggable targets. Furthermore, the role of plasma proteins involved in collagen binding and extracellular matrix in the development of osteoporosis was highlighted. Further studies are warranted to validate our findings and investigate the underlying mechanism.

KEYWORDS

mendelian randomization, osteoporosis, BMD, genetic, drug

Introduction

Characterized by low bone mineral density (BMD) and impaired bone microarchitecture, osteoporosis is the most common skeletal disorder (Compston et al., 2019). The prevalence of osteoporosis worldwide was reported to be 18.3%, and the number keeps increasing rapidly due to the aging of the population (Salari et al., 2021). Low BMD and osteoporosis may lead to the fragility of bone and increased risks of fracture, resulting in substantially higher healthcare burden, risks of disability, and mortality (Cummings and Melton, 2002; Shen et al., 2022). The typical treatment regimen for osteoporosis encompasses pharmaceutical interventions, including bisphosphonates and denosumab. However, these agents have the potential to elicit adverse effects, including gastrointestinal disturbance, musculoskeletal pain, and the rare complication of osteonecrosis of the jaw. Emerging evidence on molecular drug targets provide new viewpoint to treatment of OP.

Proteins are the major source of druggable targets and serve as biomarkers of complex traits (Santos et al., 2017; Storm et al., 2021). Previous observational studies have established correlations between protein levels and BMD (Martínez-Aguilar et al., 2019; Huang et al., 2020; Al-Ansari et al., 2022), but the findings may be biased by confounders and reverse effects, leaving the causal relationship remains unclear. Mendelian randomization (MR), similar to randomized control trials (RCT), is a novel statistical approach that uses genetic variants (i.e., protein quantitative trait loci, pQTLs) as proxies to infer the causality between exposures (i.e., proteins) and outcomes (i.e., BMD) (Sanderson et al., 2022). Dissecting causal relationships not only lead to the identification of the driving proteins of human disease but also provide strong evidence of druggable targets. Due to the limited sample size and lack of statistical power of GWAS on proteins, Previous MR studies focused more on the effect of gene methylation and gene expression on complex traits (Liu et al., 2021a; Liu et al., 2021b; Chen et al., 2022), leaving the direct effect of proteins being largely uncharacterized. With significant technological advances in proteomic profiling and a substantial reduction in the cost of genomic sequencing, genetic determinants of over 4,000 proteins have recently been reported by large-scale GWAS studies of plasma protein which offers us a valuable opportunity to comprehensively address the causal relationship between proteins and multiple diseases (Suhre et al., 2017; Sun et al., 2018; Ferkingstad et al., 2021). Through the integration of genomic and proteomic data by MR, a few studies have successfully identified druggable targets and biological pathways of human disease, such as childhood neurodevelopmental disorders (Yang et al., 2022), COVID-19 (Palmos et al., 2022; Zheng et al., 2022) and depression (Deng et al., 2022). However, genomic insights into the causal proteins, as well as druggable target of osteoporosis have not been comprehensively explored.

In this study, we aimed to prioritize causal proteins and druggable targets of osteoporosis using large-scale proteome-wide mendelian randomization by combining the largest publicly available GWAS datasets on plasma proteins and osteoporosis-related traits.

Methods

Source of exposure and outcome datasets

Exposures for MR were plasma-circulating proteins, and outcomes were bone mineral density and fractures (Table 1). The largest and most comprehensive GWAS summary datasets on plasma protein were obtained from the consortium of deCODE genetics (https://www.decode.com/summarydata/) (Ferkingstad et al., 2021). The detailed description of the datasets can be found in the original report. In brief, the pQTL datasets consisted of the associations between genome-wide genetic variants and 4,719 plasma proteins tagged by 4,907 aptamers in 35,559 Icelanders adjusted for age, sex, and sample age. The plasma protein levels were measured with the SomaScan version 4 assay (SomaLogic).

In the discovery phase for the causal proteins of osteoporosis, we utilized the largest GWAS datasets on BMD estimated from quantitative heel ultrasounds (eBMD), which includes 426,824 individuals (Morris et al., 2019). Secondary MR analysis was performed to validate the identified proteins on other osteoporosis-related traits, including total bone BMD (TBBMD), femoral neck BMD(FNBMD), lumbar spine BMD (LSBMD), and fractures. The datasets of eBMD and fracture were obtained from UK Biobank (https://www.ukbiobank.ac.uk/) which is a large-scale biomedical database and research resource, containing genetic and phenotypic information from nearly 500,000 participants. A total of 426,795 white British individuals, comprising 53,184 cases and 37,361 controls, were included in the fracture GWAS (Morris et al., 2019). The identification of fracture cases was based on hospital-based fracture diagnosis and self-reported fracture within the past 5 years. The summary-level data on dual-energy X-ray absorptiometry (DXA) derived TBBMD (n = 56,284) (Medina-Gomez et al., 2018), FNBMD (n = 49,988) and LSBMD (n = 44,731) (Zheng et al., 2015a) were accessed from GEnetic Factors for OSteoporosis Consortium (GEFOS) (http://www.gefos.org/).

To the best of our knowledge, most participants were of European ancestry and there is no sample overlap between exposure and outcome datasets which minimizes the bias from population stratification and overlapped participants.

Selection of instrumental variables

The selection of instrumental variables (IVs) were based on three assumptions (Sanderson et al., 2022). First, the IVs should be associated with exposures (proteins). Second, the IVs are supposed to be independent of confounders. Third, the effects of IVs on outcomes were only mediated through exposures.

To satisfy the first assumption, we selected genetic variants that were associated with corresponding proteins using a stringent p-value threshold ($p < 5 \times 10^{-8}$). To avoid weak instrument bias, we excluded those SNPs with weak strength (F statistic < 10). To meet the second and third assumptions, we removed pleiotropic SNPs [associated with more than five proteins as suggested by previous studies (Zheng et al., 2020; Yang et al., 2023)] and only included SNPs that explain substantially larger variance of exposure than outcomes as indicated by the Steiger filtering test ($p < 0.05$). Given the

TABLE 1 Source and descriptive information of GWAS data used in mendelian randomization.

Phenotypes	Ancestry	Sample size	Method of measurement	Adjustments	PMID
Plasma proteins	European	35,559	SomaScan	Age, sex and sample age	34857953
eBMD (Primary analysis)	European	426,824	Ultrasound	Age, sex, genotyping array, assessment center, and ancestry informative principal components 1 to 20	30598549
TBBMD (Secondary analysis)	European	56,284	DXA	Age, sex, weight, height, genomic principal components, study-specific covariates	29304378
LSBMD (Secondary analysis)	Predominantly European (~70%)	44,731	DXA	Sex, age, age^2, weight	26367794
FNBMD (Secondary analysis)	Predominantly European (~70%)	49,988	DXA	Sex, age, age^2, weight	26367794
Fracture (Secondary analysis)	European	426,795 (53,184 cases)	Hospital-based and self-reported fracture history	Age, sex, genotyping array, assessment center, and ancestry informative principal components 1 to 20	30598549

complex linkage disequilibrium (LD) structure in the region of human major histocompatibility complex (MHC) (chr6: 28477897-33448354), the SNPs within the region were removed. To ensure the imputation quality, only genetic variants with minor allele frequency >1% were included (Zheng et al., 2015b). After harmonizing exposure and outcome datasets, we performed linkage disequilibrium (LD) clumping using 1,000 genomes European reference panel to select conditionally independent genetic instruments (r2 < 0.1) (Auton et al., 2015).

There are two types of pQTL according to their genomic location, namely, cis-pQTL (SNPs within 1MB from gene start position) and trans-pQTL (SNPs more than 1MB away from gene start position) (Zheng et al., 2020; Wingo et al., 2021; Ghanbari et al., 2022; Yazdanpanah et al., 2022). The cis-pQTLs were defined as variants located near or within the corresponding protein-coding gene, while trans-pQTLs were distant from the corresponding gene. Comparing with trans-pQTL, cis-pQTLs are more likely to directly affect the genes and are less likely to be prone to horizontal pleiotropy. Therefore, only cis-pQTL within 1MB from gene start position were included in this study.

Two sample mendelian randomization

To ensure the reliability of MR results, minimize chance finding induced by false positive SNP-protein associations and enable further sensitivity analyses which require at least three instrumental variables (IVs), only proteins with ≥3 instrumental variables (IVs) were included in the following MR analysis.

For the main analysis, MR estimates of each SNP was evaluated by Wald ratios, which were subsequently meta-analyzed using Inverse Variance Weighted (IVW) method. Random-effect IVW model will be used to account for the potential heterogeneity. Furthermore, MR-Egger regression and weighted median model were adopted as sensitivity analyses to evaluate the robustness of MR results. MR-Egger regression combines Wald ratio together into a meta-regression while adjusted for any directional pleiotropy. Generally, MR-Egger regression is more robust to potential pleiotropy, but the statistical power is much lower when compared with IVW method (Burgess and Thompson, 2017). Weighted median model can give a consistent estimate of exposures on outcomes even if up to 50% if IVs were invalid (Bowden et al., 2016). The MR-Egger intercept was conducted to assess the directional horizontal pleiotropy. Cochrane's Q value was used to assess the heterogeneity. Steiger directionality test was used to evaluate the causal direction between circulating proteins and BMD or fracture.

MR analysis was conducted by TwoSampleMR R package (https://mrcieu.github.io/TwoSampleMR/). The results reaching the threshold of p-value < 4×10^{-5} (Bonferroni-corrected for 1,245 proteins) were defined as significant in the discovery analysis, by which we identified a list of putative causal proteins.

Replications of the identified proteins

The causal effects of all the plasma-circulating proteins which reached significance in the primary analysis using MR-IVW method were further replicated on other

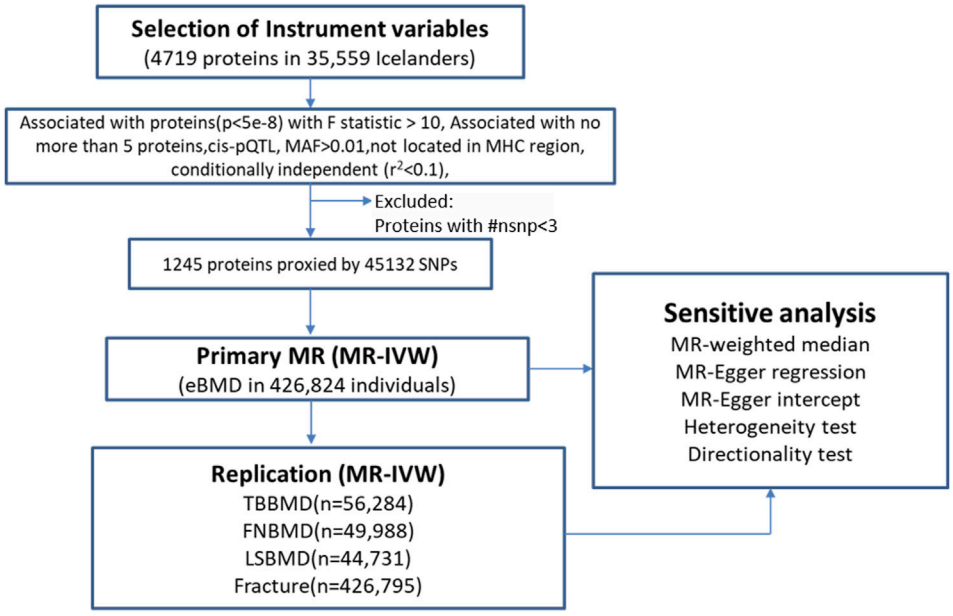


FIGURE 1
The overall study design and analysis process. MR, mendelian randomization; ##nsnp number of SNPs (IVs); MAF minor allele frequency; MHC, major histocompatibility complex; MR-IVW, mendelian randomization-Inverse Variance Weighted.

osteoporosis-related traits, including TBBMD, LSBMD, FNBMD and fractures. TBBMD, LSBMD, FNBMD were measured by DXA which is different from the method of measurement for eBMD, the corresponding population is also independent of the eBMD cohort, and the sample size (n = 44,751–56,284) is much smaller than eBMD cohort. The MR methods used in replication phase were identical to those used in the discovery cohort. For the replication and MR-sensitivity analyses, the nominal p-value < 0.05 was considered significant.

GO enrichment analysis and network analysis

To further investigate the biological role of the identified proteins as well as their interactions, we performed Gene Ontology (GO) enrichment analysis and Protein-protein interaction (PPI) network analysis based on the corresponding coding genes. GO enrichment analysis was conducted using enrichGO() function in clusterProfiler package with default settings and the results were adjusted by Benjamini-Hochberg (BH) method. PPI network analysis was conducted using search tool for the retrieval of interacting genes/proteins (STRING) database with default settings (https://string-db.org/). By default, only interactions with medium confidence (>0.4) are presented.

Druggable genome and known drugs

The druggable genome be defined as a set of protein-coding genes that can or theoretically can be modulated by therapeutic compounds. The products of druggable genes either were already targeted by existing proteins and drugs or have structural and functional properties suggestive of druggability. We obtained the list of druggable genes from the study by Finan et al. (2017) which classified the genes into three tiers according to their druggability. Tier 1 includes the targets of drugs in clinical use or clinical development. Genes in tier 2 have not already be targeted by existing drugs but encode peptides with high sequence homology to tier 1 genes. Tier 3 genes encodes extracellular proteins and members of key drug target families. We further searched for updated information on the drugs targeting the identified putative causal proteins in Open Target platform (https://www.opentargets.org/). The Open Targets Platform is a comprehensive tool that promote identification or drug targets by integrating multiple database.

Results

Primary MR analysis identified 67 putative causal proteins on BMD

The overview of the study design was presented in Figure 1. Based on the up-to-date largest and the most comprehensive GWAS studies, we conducted a two-sample MR analysis to discover putative causal proteins regulating bone mineral density. Following the criteria for the selection of instrumental variables (IVs), there are enough IVs ($n \geq 3$) for 1,245 aptamers representing 1,215 unique proteins to perform MR analysis, with the number of IVs ranging from 3 to 242 (Supplementary Tables S1, S2). The F-statistic for all the IVs were larger than 10 (Supplementary Table S1), supporting those cis-pQTLs were strong instruments. After corrected for multiple tests by Bonferroni method ($p < 0.05/1245$ or 4×10^{-5}), our primary MR analysis by IVW method revealed 67 unique proteins, represented by 72 aptamers, to be associated

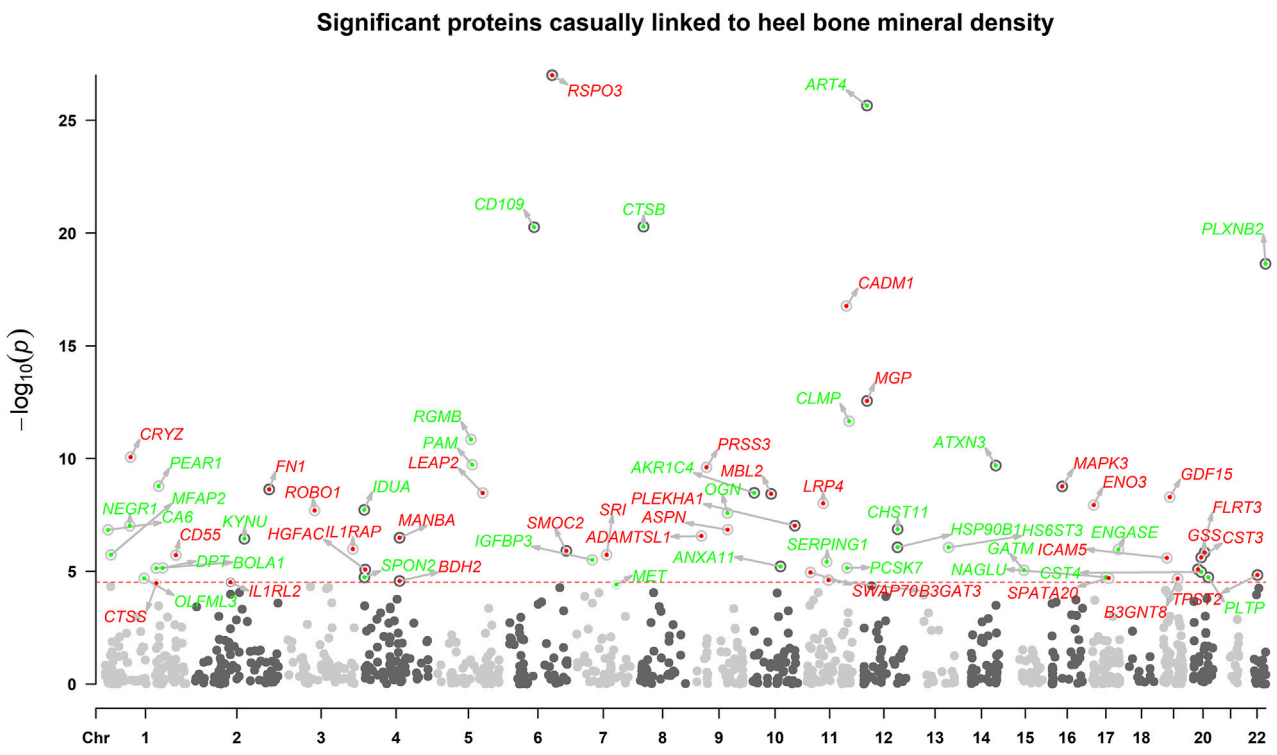


FIGURE 2
The manhattan plot illustrates the associations between circulating proteins. *X*-axis indicates the position in the chromosome, *y*-axis indicates negative log10-transformed *p*-values. Horizontal dotted line indicates the threshold of Bonferroni-adjusted significance. Red points indicate that the proteins were positively associated with eBMD, while green points indicate negative associations.

with estimated heel bone mineral density (eBMD) (Figure 2; Supplementary Table S2). Among the 67 proteins, 34 proteins have positive effects on eBMD while the remain 33 proteins were negatively associated with eBMD.

As mendelian randomization may be subjected to pleiotropy of the IVs, several sensitivity analyses were adopted. The effects on eBMD were in the same direction and the estimates were similar for all the identified proteins between IVW MR, weighted median MR and MR-egger regression (Supplementary Table S2). To be noted, all the identified proteins remained to be significant in weighted median MR. Despite low power to detect causal effects, MR-egger method revealed 45 out of the 67 proteins remained significant. Cochrane's Q test showed low to moderate heterogeneity for the associations between proteins and eBMD. Little evidence of directional horizontal pleiotropy ($p > 0.05$) was found by MR-egger intercept analyses for all the proteins except ART4, CTSB, PRSS3, GDF15, HSP90B1, CST4, TPST2, and NAGLU. Steiger directionality test indicated that MR estimate of causal direction was accurate (All $p < 0.05$).

38 proteins replicated to be associated with other osteoporosis-related traits

Out of 67 identified proteins, 38 proteins were validated to be significantly associated with TBBMD, FNBMD, LSBMD and/or fracture ($p < 0.05$) and in the same direction for all the significant associations, including ADAMTSL1, AKR1C4, ART4, ASPN, B3GAT3, B3GNT8, CA6, CD109, CLMP, CST4, CTSB, CTSS, ENGASE, FLRT3, HS6ST3, HSP90B1, ICAM5, IDUA, IGFBP3, LRP4, MBL2, MET, MFAP2, MGP, NAGLU, OGN, OLFML3, PLEKHA1, PLXNB2, ROBO1, RSPO3, SERPING1, SMOC2, SPATA20, SPON2, SRI, and TPST2 (Figure 3; Supplementary Table S3). The results were largely consistent in multiple sensitivity analyses (Figure 3; Supplementary Table S3). Furthermore, the genetically predicted causal effects of the 67 identified proteins on BMD at different anatomical sites demonstrate positive correlations (r = 0.33–0.65), and exhibit negative associations with fracture risk (r = −0.72–−0.33) (Supplementary Table S4).

GO enrichment analysis and PPI networks

GO enrichment analysis revealed that the putative causal proteins were involved in 37 GO items, including 12 items in the group of cellular components items and 25 items in the group of molecular function (Figure 4; Supplementary Table S5). The top GO item was collagen-containing extracellular matrix (GO:0062023, $p = 1.6 \times 10^{-10}$, p.adjust $= 2.1 \times 10^{-8}$) with 14 genes involved, including CTSB, MGP, FN1, MBL2, GDF15, OGN, ASPN, HSP90B1, SMOC2, MFAP2, SERPING1, ANXA11, DPT, CTSS. Other items, such as

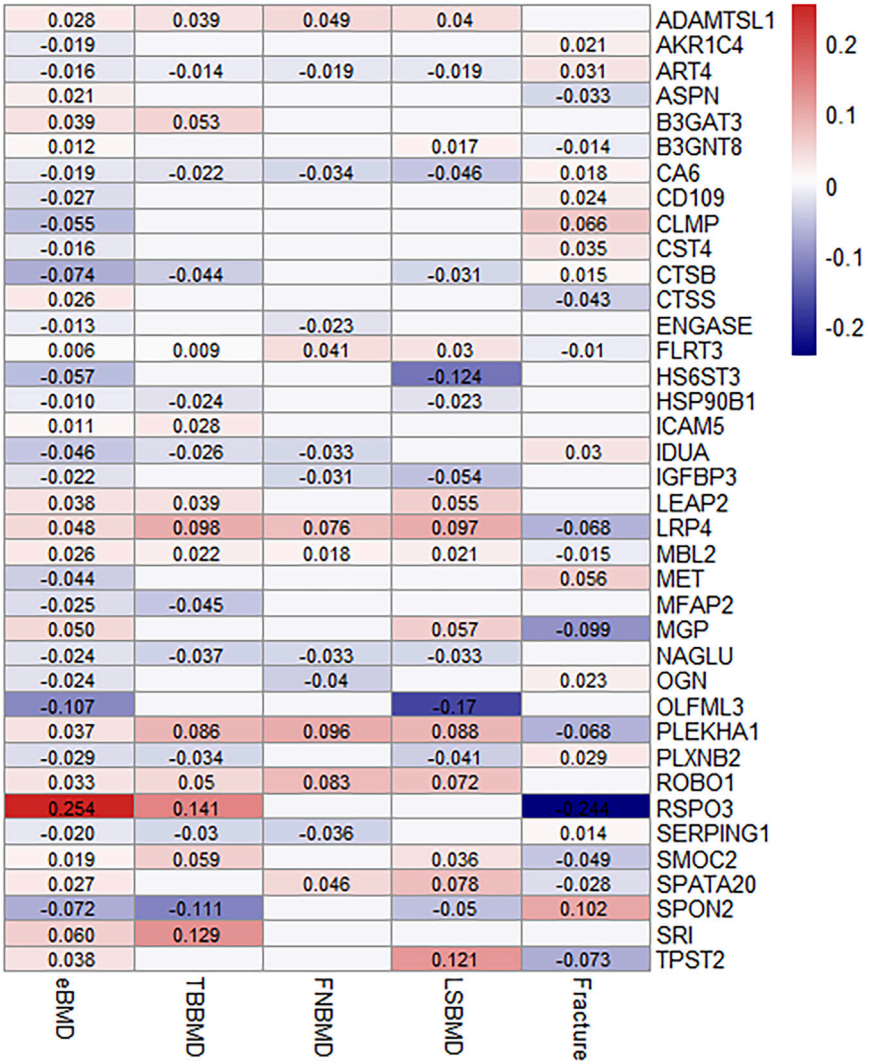


FIGURE 3
MR-IVW effect estimates (beta value) of proteins that were identified and replicated on osteoporosis-related traits. For aesthetic purposes, the beta values for the insignificant associations were set as zero and not displayed.

collagen binding (GO:0005518, $p = 8.6 \times 10^{-5}$, *p*.adjust = 0.006), extracellular matrix structural constituent (GO:0005201, $p = 2.7 \times 10^{-5}$, *p*.adjust = 0.003) has also been highlighted. We have also performed another GO enrichment analysis of non-significant proteins, while collagen binding extracellular matrix structural constituent were not among the top GO terms. We conducted an additional Gene Ontology (GO) enrichment analysis on proteins that did not show statistical significance. Our results (Supplementary Figure S1; Supplementary Table S6) did not identify collagen binding or extracellular matrix structural constituent as significant GO terms in the top-ranking results. Moreover, Many of the top enriched Gene Ontology (GO) terms for these non-causal proteins showed specificity, such as secretory granule lumen, blood microparticle, heparin binding, and receptor ligand activity. We have noticed much lower *p*-value in the enrichment analysis for non-causal proteins which may be attributed to the substantially larger sample size (1,133 vs. 67 proteins). Protein-protein interaction networks (Supplementary Figure S2) indicate FN1 and MAPK3 were highly connected with other proteins and may play a key role in the modulation of BMD.

Druggability of the identified proteins

As shown in Supplementary Table S7, a total of 42 genes encoding putative causal proteins of BMD were listed in the druggable genome, 9 genes in tier 1 group (CTSB, MAPK3, FN1, CHST11, CA6, GSS, CD55, CTSS, MET), 7 genes in tier 2 group and the remaining 26 genes in tier 3 group. By searching Open Target Platform, we identified multiple drugs with investigational or approved indications targeting MAPK3, FN1, CA6, CTSS, MET, RSPO3, IL1RAP and IL1RL2 (Supplementary Table S8). For example, hepatocyte growth

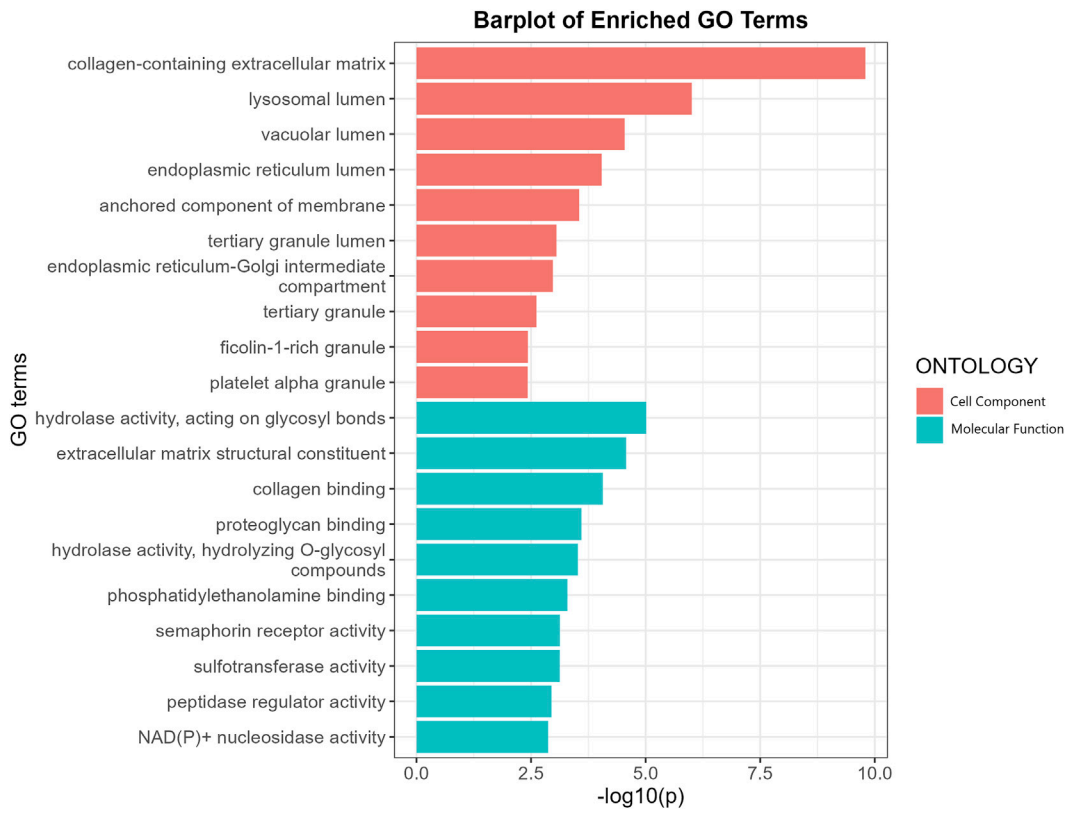


FIGURE 4
GO enrichment analysis of the putative causal proteins of BMD. *X*-axis indicates negative log10-transformed *p*-values, y a-axis indicates GO items. CC, Cellular component; MF, Molecular function.

factor receptor (HGFR), encoded by MET gene, was negatively associated with BMD (beta = −0.04, $p = 3.9 \times 10^{-5}$), Cabozantinib may be repurposed to treat osteoporosis as an inhibitor of HGFR which was supported by real-world observation (Pan et al., 2020; Ratta et al., 2021).

Discussion

In this study, we undertook two-sample MR analysis to investigate the causal effect of 1,245 circulating proteins from up to date largest proteomics GWAS (35,559 individuals) on multiple osteoporosis-related traits. We identified 67 proteins that were causally associated with eBMD and 38 out of the 67 proteins have been successfully replicated in other osteoporosis-related traits. The findings were supported by multiple sensitivity analysis. Further Enrichment analysis highlights the role of proteins involved in collagen binding and extracellular matrix in the development of osteoporosis. Evaluation of druggable genomics provided further information for the chance of drug repurposing to treat osteoporosis or monitoring and preventing side effects of drugs that may induce osteoporosis.

Multiple proteins that showed strong evidence to be associated with osteoporosis-related traits were supported by previous studies, such as RSPO3, IDUA, SMOC2. RSPO3 (R-spondin 3) is a secreted protein that modulates WNT-signaling pathway, and the pathway has been widely recognized to be the major determinant of bone formation and accumulation of bone mass (Marini et al., 2022). Previous GWAS study has identified locus in RSPO3 region was associated with BMD (Duncan et al., 2011). RSPO3 was highly expressed in osteoblast, and osteoblasts-specific inactivation of RSPO3 in mice induced decreased trabecular bone mass via downregulation of WNT-signaling targets (Nilsson et al., 2021). The role of RSPO3 as positive regulator of BMD was confirmed by our large-scale MR study. We have also observed a negative effect of IDUA (alpha-L-iduronidase) on BMD which is consistent with several animal studies which showed that IDUA knockout mice have thickened abnormally formed bones and increased BMD (Clarke et al., 1997; Kim et al., 2015). SMOC2 (SPARC related modular calcium binding 2 protein) is an extracellular matrix glycoprotein which promotes matrix assembly and angiogenic activity (Rocnik et al., 2006). In humans, pathogenic mutations in SMOC2 have been implicated in severe dental anomalies and skeletal dysplasia, characterized by microdontia, dentin dysplasia, reduced alveolar/jaw bone density, and flatten lumbar vertebrae (Bloch-Zupan et al., 2011; Alfawaz et al., 2013). Zebrafish models with smoc2 knockdown showed reduced expression of bone morphogenetic protein (BMP)

target genes and exhibited craniofacial hypoplasia (Melvin et al., 2013; Mommaerts et al., 2014). Morkmued et al. (2020) demonstrated that deletion of SMOC2 in mouse lead to impaired bone healing and age-dependent bone loss via osteoclast activation. Our MR study further strengthens the positive effect of SMOC2 on multiple osteoporosis-related traits in the general population.

Genetically predicted higher concentration of circulating LRP4 (Low-density lipoprotein receptor-related protein 4) was positively associated with higher eBMD, TBBMD, FNBMD, FNBMD and negatively associated with lower risks of fracture in our study. However, the result was contrary to some previous studies. LRP4, belonging to lipoprotein-related protein family, is a membrane receptor. Leupin et al. (2011) found that LRP4 can interact with Sclerostin (SOST) and enhance the sclerostin-mediated inhibition of Wnt/β-Catenin Signaling and vitro mineralization. They have also identified two missense mutations (R1170W and W1186S) in sclerosteosis patients (Leupin et al., 2011). Knock-in of the mutations in mice recapitulated the high BMD phenotype (Bullock et al., 2019). Chang et al. (2014) generated osteoblast/osteocyte-specific LRP4 conditional knockout mice by flanking Lrp4 exon 1, and the mice exhibited significantly increased total bone mineral density. Here are some possible reasons for the discrepancy between our findings and these studies. Intracellular and extracellular LRP4 may exert opposite effects on BMD, plasma LRP4 may bind to the circulating SOST and attenuate the inhibition of Wnt Signaling induced by the interaction between SOST and LRP4 in the cytomembrane. Mutations in different domainss of LRP4 may also have different effects on BMD, this is supported by another study (Choi et al., 2009) which revealed that Lrp4 deficient mutant mice, generated by introducing a stop codon into the exon 36 of LRP4, exhibited shortened total femur length and reduced BMD. Furthermore, the present MR study was based on the information on plasma LRP4 levels from the general population and the condition is distinct from loss-of-function mutations in LRP4 among specific patients. Therefore, the comprehensive and exact roles of LRP4 in bone metabolism remains to be explored.

For mendelian randomization, replication in other independent datasets is important. Storm et al. (2021) successfully replicated 15 of 31 putative drug targets for Parkinson's disease and Chen et al. (2022) replicated 4 of 11 causal genes for aortic aneurysms. In this study, 38 out of 67 proteins were successfully replicated in other osteoporosis-related traits, supporting their important role in the regulation of bone metabolism. Low statistical power in the validation cohort, difference between methods of measurement and site-specific effect on BMD may be the reasons why other proteins were not replicated. Thus, the potential causal effect of those proteins cannot be excluded. One example is CADM1 (Cell Adhesion Molecule 1) which showed positive effect on eBMD ($\beta = 0.17$, $p =$ 1.74E-17) in the primary analysis but failed to be replicated. CADM1 is a ubiquitously expressed gene involved in many biological processes, including cancer and spermatogenesis, and revealed as an osteoblast-specific marker for osteosarcoma (Inoue et al., 2013). Using a co-expression network for mineralizing osteoblasts, Sabik et al. (2020) identified CADM1 as the core gene modulating BMD. CADM1 has also been shown to inhibit osteoclastic bone resorption under the regulation of RANKL and NFATc1, and decreased bone mass was exhibited in Cadm1-deficient mice (Nakamura et al., 2017). These studies supported CADM1 as a protective factor for osteoporosis.

GO enrichment analysis of the putative causal proteins of osteoporosis highlighted the role of collagen and extracellular matrix in the regulation of BMD. Collagen is the primary structural components of bone which enables adhesiveness of cells and assembly of the extracellular matrix. Mutations in the multiple collagen encoding genes, like COL1A1, COL1A2, COL2A1, have been well documented in a range of mendelian bone fragility disorders characterized by low bone mass and microarchitectural deterioration of bone tissue, including osteogenesis imperfecta, Ehlers–Danlos syndrome and hypochondrogenesis. Randomized controlled trials have shown that specific collagen peptides have anabolic influence on bone formation and BMD in Postmenopausal Women. Extracellular matrix (ECM) is non-cellular structure secreted by cells into the extracellular space, and has been revealed to influence functional characteristics of the mature bone by regulating cell adhesion, proliferation, and responses to cytokines. As an essential glycoprotein of the extracellular matrix, FN1 (Fibronectin 1) was highlighted as a key causal proteins in the PPI network analysis. Therefore, our result reinforced and expanded the effect of collagen and ECM on BMD among general population.

We have observed a recently published MR study (Han et al., 2023) which also investigated the potential causal effect of plasma proteins on bone mineral density. However, it is essential to emphasize that our study significantly diverges from that particular study in terms of methodology, results, and conclusions. Specifically, our study differs notably in the selection of instrumental variables for exposures. While that similar study selected all SNPs with $p \leq 5 \times 10^{-8}$ across all the genomic region, we only selected cis-pQTLs, which are variants located near or within the corresponding protein-coding gene. Cis-pQTLs have higher biological plausibility, are more likely to directly affect the genes, and are less likely to be prone to horizontal pleiotropy compared to trans-pQTLs. In addition, we only included proteins with more than 3 instrumental variables, which allowed for sensitivity analyses. Our different selection criteria resulted in distinct results from the study by Han et al. (2023). For instance, in their study, they presented contradictory results with no proper explanation that elevated genetically predicted abundance of LRP4 was associated with higher eBMD and lower TB-BMD. In contrast, in our study, with the only inclusion of cis-pQTLs, elevated genetically predicted abundance of LRP4 was consistently associated with higher eBMD, TB-BMD, FN-BMD, LS-BMD, and lower risks of fracture. Moreover, we adopted the Bonferroni correction for multiple tests to minimize false positive rates, which is stricter than the false-discovery rate (FDR) correction method used by Han et al. Additionally, we used multiple datasets to

replicate our primary results, including GWAS for TB-BMD, FN-BMD, LS-BMD, and fractures. These measures lowered the number of identified proteins and made our results more conclusive and credible. To further support our findings and enable their application in clinical practice, we investigated the druggability of the identified proteins, which was not included in the study by Han et al. (2023). Interestingly, some of our findings have been supported by previous observations, which further strengthen the credibility and reliability of our results. For example, hepatocyte growth factor receptor (HGFR), encoded by the MET gene, was negatively associated with BMD (beta = −0.04, $p = 3.9 \times 10^{-5}$). Cabozantinib, an inhibitor of HGFR, may be repurposed to treat osteoporosis, as supported by real-world observations.

Overall, our study has some strengths. MR approach can largely reduce the measured or unmeasured confounding and reverse causation compared with traditional observational studies. Therefore, the identified proteins are more like to be the "true" causal factors rather than bystanders. The large sample size of the GWAS datasets used in the study enable us to select multiple IVs for each protein which increased statistical power and reduce chance findings. Only cis SNPs which has high biological plausibility were selected as IVs. Hence, the risks of pleiotropy were minimized which strength MR results.

There are also several limitations in the present study need to be acknowledged. Firstly, the potential influence of directional horizontal pleiotropy could not be completely excluded, although we select IVs using stringent criteria and performed multiple sensitivity analyses which accounted for the horizontal pleiotropy. Secondly, the GWAS data utilized in the study were mainly obtained from the cohorts of European ancestry. Considering the difference of genetic architecture among different ethnic groups, caution should be exercised when generalizing our results to other ancestries. Further population-specific MR studies were warranted to cross-validate our findings in non-Europeans. Thirdly, we used GWAS summary data on eBMD rather than BMD in the primary MR analysis because of the larger sample size, which may induce bias. However, there is a high degree of genetic concordance between eBMD and DXA derived-BMD and we replicated the primary findings in independent GWAS datasets on DXA-derived BMD at individual bone sites and overall fracture risks. Finally, to minimize the false positive rates, we adopted Bonferroni correction for multiple tests and some proteins that may play an important role in BMD may be missed.

In conclusion, our large-scale MR analysis provided evidence that genetically predicted levels of 67 circulating proteins were associated with eBMD, and 38 of the 67 proteins were validated to be associated with other osteoporosis-related traits. Proteins involved in collagen binding and extracellular matrix play important role in the pathogenesis of osteoporosis. The study broadens the causal proteins for osteoporosis which may serve as early screening biomarkers and/or druggable targets. Further studies are warranted to validate our findings.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

ZW, Investigation, data curation, formal analysis, writing—review and editing; KY, Methodology, validation, visualization; T-PL, Writing—review and editing, supervision; JC, Writing—review and editing, supervision; ZZ, Writing—review and editing, supervision; WL, Conceptualization, writing—review and editing, resources. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Start-up grant from Chinese University of Hong Kong (Ref. Nos. 4930991 and 4930992); 2020 Rising Star Award from American Society for Bone and Mineral Research; General Research Fund (Ref. No. 14104620) and Research Matching Grant Scheme, University Grants Committee; Area of Excellence (Ref. No. AoE/M-402/20), University Grants Committee; Center for Neuromusculoskeletal Restorative Medicine (Ref. No. CT1.1), Health@InnoHK program, Innovation Technology Commission, Hong Kong SAR, China.

Acknowledgments

We thank all the study participants and research staff for their contributions and commitment to the present study.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1161817/full#supplementary-material

References

Al-Ansari, M. M., Aleidi, S. M., Masood, A., Alnehmi, E. A., Abdel Jabar, M., Almogren, M., et al. (2022). Proteomics profiling of osteoporosis and osteopenia patients and associated network analysis. *Int. J. Mol. Sci.* 23. doi:10.3390/ijms231710200

Alfawaz, S., Fong, F., Plagnol, V., Wong, F. S., Fearne, J., and Kelsell, D. P. (2013). Recessive oligodontia linked to a homozygous loss-of-function mutation in the SMOC2 gene. *Arch. Oral Biol.* 58, 462–466. doi:10.1016/j.archoralbio.2012.12.008

Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., et al. (2015). A global reference for human genetic variation. *Nature* 526, 68–74. doi:10.1038/nature15393

Bloch-Zupan, A., Jamet, X., Etard, C., Laugel, V., Muller, J., Geoffroy, V., et al. (2011). Homozygosity mapping and candidate prioritization identify mutations, missed by whole-exome sequencing, in SMOC2, causing major dental developmental defects. *Am. J. Hum. Genet.* 89, 773–781. doi:10.1016/j.ajhg.2011.11.002

Bowden, J., Davey Smith, G., Haycock, P. C., and Burgess, S. (2016). Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.* 40, 304–314. doi:10.1002/gepi.21965

Bullock, W. A., Hoggatt, A. M., Horan, D. J., Elmendorf, A. J., Sato, A. Y., Bellido, T., et al. (2019). Lrp4 mediates bone homeostasis and mechanotransduction through interaction with sclerostin *in vivo*. *iScience* 20, 205–215. doi:10.1016/j.isci.2019.09.023

Burgess, S., and Thompson, S. G. (2017). Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur. J. Epidemiol.* 32, 377–389. doi:10.1007/s10654-017-0255-x

Chang, M. K., Kramer, I., Huber, T., Kinzel, B., Guth-Gundel, S., Leupin, O., et al. (2014). Disruption of Lrp4 function by genetic deletion or pharmacological blockade increases bone mass and serum sclerostin levels. *Proc. Natl. Acad. Sci. U. S. A.* 111, E5187–E5195. doi:10.1073/pnas.1413828111

Chen, Y., Xu, X., Wang, L., Li, K., Sun, Y., Xiao, L., et al. (2022). Genetic insights into therapeutic targets for aortic aneurysms: A mendelian randomization study. *EBioMedicine* 83, 104199. doi:10.1016/j.ebiom.2022.104199

Choi, H. Y., Dieckmann, M., Herz, J., and Niemeier, A. (2009). Lrp4, a novel receptor for Dickkopf 1 and sclerostin, is expressed by osteoblasts and regulates bone growth and turnover *in vivo*. *PLoS One* 4, e7930. doi:10.1371/journal.pone.0007930

Clarke, L. A., Russell, C. S., Pownall, S., Warrington, C. L., Borowski, A., Dimmick, J. E., et al. (1997). Murine mucopolysaccharidosis type I: Targeted disruption of the murine alpha-L-iduronidase gene. *Hum. Mol. Genet.* 6, 503–511. doi:10.1093/hmg/6.4.503

Compston, J. E., McClung, M. R., and Leslie, W. D. (2019). Osteoporosis. *Lancet* 393, 364–376. doi:10.1016/s0140-6736(18)32112-3

Cummings, S. R., and Melton, L. J. (2002). Epidemiology and outcomes of osteoporotic fractures. *Lancet* 359, 1761–1767. doi:10.1016/s0140-6736(02)08657-9

Deng, Y. T., Ou, Y. N., Wu, B. S., Yang, Y. X., Jiang, Y., Huang, Y. Y., et al. (2022). Identifying causal genes for depression via integration of the proteome and transcriptome from brain and blood. *Mol. Psychiatry* 27, 2849–2857. doi:10.1038/s41380-022-01507-9

Duncan, E. L., Danoy, P., Kemp, J. P., Leo, P. J., McCloskey, E., Nicholson, G. C., et al. (2011). Genome-wide association study using extreme truncate selection identifies novel genes affecting bone mineral density and fracture risk. *PLoS Genet.* 7, e1001372. doi:10.1371/journal.pgen.1001372

Ferkingstad, E., Sulem, P., Atlason, B. A., Sveinbjornsson, G., Magnusson, M. I., Styrmisdottir, E. L., et al. (2021). Large-scale integration of the plasma proteome with genetics and disease. *Nat. Genet.* 53, 1712–1721. doi:10.1038/s41588-021-00978-w

Finan, C., Gaulton, A., Kruger, F. A., Lumbers, R. T., Shah, T., Engmann, J., et al. (2017). The druggable genome and support for target identification and validation in drug development. *Sci. Transl. Med.* 9. doi:10.1126/scitranslmed.aag1166

Ghanbari, F., Yazdanpanah, N., Yazdanpanah, M., Richards, J. B., and Manousaki, D. (2022). Connecting genomics and proteomics to identify protein biomarkers for adult and youth-onset type 2 diabetes: A two-sample mendelian randomization study. *Diabetes* 71, 1324–1337. doi:10.2337/db21-1046

Han, B. X., Yan, S. S., Yu, H., Xu, Q., Zhao, Q. G., Ma, X. L., et al. (2023). Causal effects of plasma proteome on osteoporosis and osteoarthritis. *Calcif. Tissue Int.* 112, 350–358. doi:10.1007/s00223-022-01049-w

Huang, D., Wang, Y., Lv, J., Yan, Y., Hu, Y., Liu, C., et al. (2020). Proteomic profiling analysis of postmenopausal osteoporosis and osteopenia identifies potential proteins associated with low bone mineral density. *PeerJ* 8, e9009. doi:10.7717/peerj.9009

Inoue, T., Hagiyama, M., Enoki, E., Sakurai, M. A., Tan, A., Wakayama, T., et al. (2013). Cell adhesion molecule 1 is a new osteoblastic cell adhesion molecule and a diagnostic marker for osteosarcoma. *Life Sci.* 92, 91–99. doi:10.1016/j.lfs.2012.10.021

Kim, C., Kwak, M. J., Cho, S. Y., Ko, A. R., Rheey, J., Kwon, J. Y., et al. (2015). Decreased performance in IDUA knockout mouse mimic limitations of joint function and locomotion in patients with Hurler syndrome. *Orphanet J. Rare Dis.* 10, 121. doi:10.1186/s13023-015-0337-3

Leupin, O., Piters, E., Halleux, C., Hu, S., Kramer, I., Morvan, F., et al. (2011). Bone overgrowth-associated mutations in the LRP4 gene impair sclerostin facilitator function. *J. Biol. Chem.* 286, 19489–19500. doi:10.1074/jbc.M110.190330

Liu, Y., Li, B., Ma, Y., Huang, Y., Ouyang, F., and Liu, Q. (2021a). Mendelian randomization integrating GWAS, eQTL, and mQTL data identified genes pleiotropically associated with atrial fibrillation. *Front. Cardiovasc Med.* 8, 745757. doi:10.3389/fcvm.2021.745757

Liu, Y., Jin, G., Wang, X., Dong, Y., and Ding, F. (2021b). Identification of new genes and loci associated with bone mineral density based on mendelian randomization. *Front. Genet.* 12, 728563. doi:10.3389/fgene.2021.728563

Marini, F., Giusti, F., Palmini, G., and Brandi, M. L. (2022). Role of Wnt signaling and sclerostin in bone and as therapeutic targets in skeletal disorders. *Osteoporos. Int.* 34, 213. doi:10.1007/s00198-022-06523-7

Martínez-Aguilar, M. M., Aparicio-Bautista, D. I., Ramírez-Salazar, E. G., Reyes-Grajeda, J. P., De la Cruz-Montoya, A. H., Antuna-Puente, B., et al. (2019). Serum proteomic analysis reveals vitamin D-binding protein (VDBP) as a potential biomarker for low bone mineral density in Mexican postmenopausal women. *Nutrients* 11. doi:10.3390/nu11122853

Medina-Gomez, C., Kemp, J. P., Trajanoska, K., Luan, J., Chesi, A., Ahluwalia, T. S., et al. (2018). Life-course genome-wide association study meta-analysis of total body BMD and assessment of age-specific effects. *Am. J. Hum. Genet.* 102, 88–102. doi:10.1016/j.ajhg.2017.12.005

Melvin, V. S., Feng, W., Hernandez-Lagunas, L., Artinger, K. B., and Williams, T. (2013). A morpholino-based screen to identify novel genes involved in craniofacial morphogenesis. *Dev. Dyn.* 242, 817–831. doi:10.1002/dvdy.23969

Mommaerts, H., Esguerra, C. V., Hartmann, U., Luyten, F. P., and Tylzanowski, P. (2014). Smoc2 modulates embryonic myelopoiesis during zebrafish development. *Dev. Dyn.* 243, 1375–1390. doi:10.1002/dvdy.24164

Morkmued, S., Clauss, F., Schuhbaur, B., Fraulob, V., Mathieu, E., Hemmerlé, J., et al. (2020). Deficiency of the SMOC2 matricellular protein impairs bone healing and produces age-dependent bone loss. *Sci. Rep.* 10, 14817. doi:10.1038/s41598-020-71749-6

Morris, J. A., Kemp, J. P., Youlten, S. E., Laurent, L., Logan, J. G., Chai, R. C., et al. (2019). An atlas of genetic influences on osteoporosis in humans and mice. *Nat. Genet.* 51, 258–266. doi:10.1038/s41588-018-0302-x

Nakamura, S., Koyama, T., Izawa, N., Nomura, S., Fujita, T., Omata, Y., et al. (2017). Negative feedback loop of bone resorption by NFATc1-dependent induction of Cadm1. *PLoS One* 12, e0175632. doi:10.1371/journal.pone.0175632

Nilsson, K. H., Henning, P., El Shahawy, M., Nethander, M., Andersen, T. L., Ejersted, C., et al. (2021). RSPO3 is important for trabecular bone and fracture risk in mice and humans. *Nat. Commun.* 12, 4923. doi:10.1038/s41467-021-25124-2

Palmos, A. B., Millischer, V., Menon, D. K., Nicholson, T. R., Taams, L. S., Michael, B., et al. (2022). Proteome-wide Mendelian randomization identifies causal links between blood proteins and severe COVID-19. *PLoS Genet.* 18, e1010042. doi:10.1371/journal.pgen.1010042

Pan, T., Martinez, M., Hubka, K. M., Song, J. H., Lin, S. C., Yu, G., et al. (2020). Cabozantinib reverses renal cell carcinoma-mediated osteoblast inhibition in three-dimensional coculture *in vitro* and reduces bone osteolysis *in vivo*. *Mol. Cancer Ther.* 19, 1266–1278. doi:10.1158/1535-7163.Mct-19-0174

Ratta, R., Verzoni, E., Mennitto, A., Pantano, F., Martinetti, A., Raimondi, A., et al. (2021). Effects of cabozantinib on bone turnover markers in real-world metastatic renal cell carcinoma. *Tumori* 107, 542–549. doi:10.1177/0300891620969817

Rocnik, E. F., Liu, P., Sato, K., Walsh, K., and Vaziri, C. (2006). The novel SPARC family member SMOC-2 potentiates angiogenic growth factor activity. *J. Biol. Chem.* 281, 22855–22864. doi:10.1074/jbc.M513463200

Sabik, O. L., Calabrese, G. M., Taleghani, E., Ackert-Bicknell, C. L., and Farber, C. R. (2020). Identification of a core module for bone mineral density through the integration of a Co-expression network and GWAS data. *Cell. Rep.* 32, 108145. doi:10.1016/j.celrep.2020.108145

Salari, N., Darvishi, N., Bartina, Y., Larti, M., Kiaei, A., Hemmati, M., et al. (2021). Global prevalence of osteoporosis among the world older adults: A comprehensive systematic review and meta-analysis. *J. Orthop. Surg. Res.* 16, 669. doi:10.1186/s13018-021-02821-8

Sanderson, E., Glymour, M. M., Holmes, M. V., Kang, H., Morrison, J., Munafò, M. R., et al. (2022). Mendelian randomization. *Nat. Rev. Methods Prim.* 2, 6. doi:10.1038/s43586-021-00092-5

Santos, R., Ursu, O., Gaulton, A., Bento, A. P., Donadi, R. S., Bologa, C. G., et al. (2017). A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* 16, 19–34. doi:10.1038/nrd.2016.230

Shen, Y., Huang, X., Wu, J., Lin, X., Zhou, X., Zhu, Z., et al. (2022). The global burden of osteoporosis, low bone mass, and its related fracture in 204 countries and territories, 1990-2019. *Front. Endocrinol. (Lausanne)* 13, 882241. doi:10.3389/fendo.2022.882241

Storm, C. S., Kia, D. A., Almramhi, M. M., Bandres-Ciga, S., Finan, C., Hingorani, A. D., et al. (2021). Finding genetically-supported drug targets for Parkinson's disease

using Mendelian randomization of the druggable genome. *Nat. Commun.* 12, 7342. doi:10.1038/s41467-021-26280-1

Suhre, K., Arnold, M., Bhagwat, A. M., Cotton, R. J., Engelke, R., Raffler, J., et al. (2017). Connecting genetic risk to disease end points through the human blood plasma proteome. *Nat. Commun.* 8, 14357. doi:10.1038/ncomms14357

Sun, B. B., Maranville, J. C., Peters, J. E., Stacey, D., Staley, J. R., Blackshaw, J., et al. (2018). Genomic atlas of the human plasma proteome. *Nature* 558, 73–79. doi:10.1038/s41586-018-0175-2

Wingo, T. S., Liu, Y., Gerasimov, E. S., Gockley, J., Logsdon, B. A., Duong, D. M., et al. (2021). Brain proteome-wide association study implicates novel proteins in depression pathogenesis. *Nat. Neurosci.* 24, 810–817. doi:10.1038/s41593-021-00832-6

Yang, J., He, X., Qian, L., Zhao, B., Fan, Y., Gao, F., et al. (2022). Association between plasma proteome and childhood neurodevelopmental disorders: A two-sample mendelian randomization analysis. *EBioMedicine* 78, 103948. doi:10.1016/j.ebiom.2022.103948

Yang, H., Chen, L., and Liu, Y. (2023). Novel causal plasma proteins for hypothyroidism: A large-scale plasma proteome mendelian randomization analysis. *J. Clin. Endocrinol. Metab.* 108, 433–442. doi:10.1210/clinem/dgac575

Yazdanpanah, N., Yazdanpanah, M., Wang, Y., Forgetta, V., Pollak, M., Polychronakos, C., et al. (2022). Clinically relevant circulating protein biomarkers for type 1 diabetes: Evidence from a two-sample mendelian randomization study. *Diabetes Care* 45, 169–177. doi:10.2337/dc21-1049

Zheng, H. F., Forgetta, V., Hsu, Y. H., Estrada, K., Rosello-Diez, A., Leo, P. J., et al. (2015a). Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. *Nature* 526, 112–117. doi:10.1038/nature14878

Zheng, H. F., Rong, J. J., Liu, M., Han, F., Zhang, X. W., Richards, J. B., et al. (2015b). Performance of genotype imputation for low frequency and rare variants from the 1000 genomes. *PLoS One* 10, e0116487. doi:10.1371/journal.pone.0116487

Zheng, J., Haberland, V., Baird, D., Walker, V., Haycock, P. C., Hurle, M. R., et al. (2020). Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nat. Genet.* 52, 1122–1131. doi:10.1038/s41588-020-0682-6

Zheng, J., Zhang, Y., Zhao, H., Liu, Y., Baird, D., Karim, M. A., et al. (2022). Multi-ancestry Mendelian randomization of omics traits revealing drug targets of COVID-19 severity. *EBioMedicine* 81, 104112. doi:10.1016/j.ebiom.2022.104112

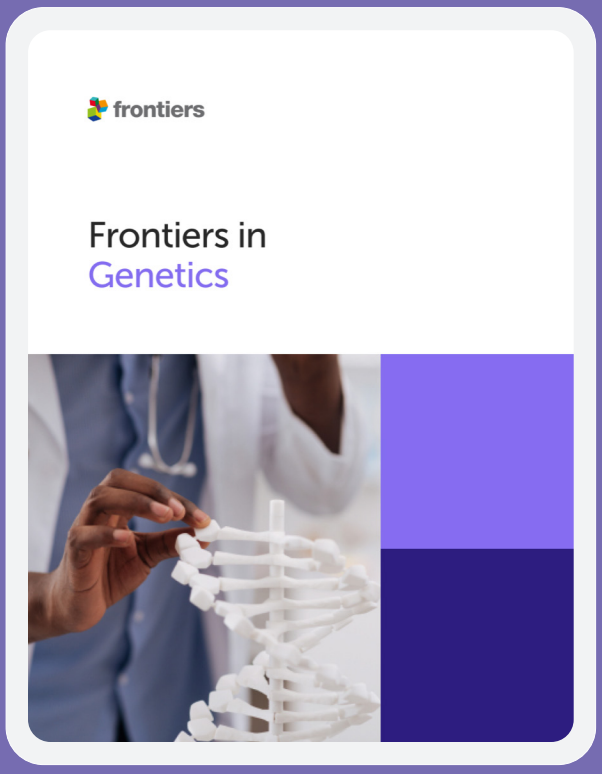
Frontiers in Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact