Exploration of the core gene signatures and mechanisms between NAFLD and sarcopenia through transcriptomic level

Ziying Xu[1,2], Zihui Yu[2], Shang Li[3,4], Ziyan Tian[2],
Jing Yuan[2]* and Fuping You[1]*

[1]School of Basic Medical Sciences, Peking University, Beijing, China, [2]Department of Bacteriology,
Capital Institute of Pediatrics, Beijing, China, [3]Department of Orthopedics, General Hospital of
Chinese People's Liberation Army of China (PLA), Beijing, China, [4]National Clinical Research Center
for Orthopedics, Sports Medicine & Rehabilitation, Beijing, China

Introduction: The increased prevalence of non-alcoholic fatty liver disease (NAFLD) and sarcopenia among the elderly are facing a significant challenge to the world's health systems. Our study aims to identify the coexpressed genes in NAFLD and sarcopenia patients.

Methods: We downloaded the transcriptome data of NAFLD tissue from patients, as well as muscle tissues from sarcopenia patients, from the GEO database in order to investigate the shared transcriptional regulation mechanisms between these two diseases. Then, focusing on the genes that were frequently expressed in these diseases, together with GSVA and WGCNA, we utilized a range of analysis methods to identify the main co-expressed genes in both diseases by taking intersections. We investigated these changes after learning that they mostly affected lipid metabolism and oxidative stress injury pathways.

Results: By analyzing these genes and their interactions with transcription factors and proteins, we were able to identify 8 genes that share common patterns. From these 8 genes, we were possible to forecast potential future medicines. Our research raises the possibility of NAFLD and sarcopenia transcriptome regulatory pathways in aging populations.

Discussion: In conclusion, a complete transcription pattern mapping was carried out in order to identify the core genes, underlying biological mechanisms, and possible therapeutic targets that regulate aging in NAFLD and sarcopenia patients. It provides novel insights and proof in favor of decreasing the increased prevalence of sarcopenia in the elderly caused by NAFLD.

KEYWORDS

NAFLD (non-alcoholic fatty liver disease), sarcopenia, high throughput sequencing, bioinformatics analyses, co-expressed genes

1 Introduction

Non-alcoholic fatty liver disease (NAFLD) is characterized by the accumulation of lipids in the liver, which may further lead to the deterioration of liver fibrosis (1), cirrhosis and even liver cancer (2). According to statistics, about 25% of the world population is suffered from NAFLD, and patients attacked by NAFLD are becoming younger (3). Therefore, it is imperative to underlying the mechanisms of NAFLD and develop an effective treatment for it. Hepatic steatosis is age-related and associated with metabolic syndromes such as high-fat diet (4), flora disorders(5) and hyperlipidemia, as well as various toxins, drugs, and diseases. The pathogenesis of NAFLD is well established as a two-strike and multiple-strike theory but the molecular mechanism of the occurrence and development of NAFLD remains uncovered. (6). However, other aging diseases interrelated with lipid metabolism and fibrosis can aggravate the exacerbation of NAFLD.

Skeletal muscle accounts for 40% of the body weight, undertake 30% of the basic energy metabolism, and maintains behavioral functions. In adults, muscle loss begins at age 30 and accelerates after age 50. By age 60, muscle loss can reach 30 percent. The degeneration of muscle with the increase of age is defined as sarcopenia, which is accompanied by a series of pathological changes such as decreased muscle mass, fibrosis, and fat infiltration, seriously affecting the functional activities of the elderly and reducing life expectancy(7). The prevalence of sarcopenia in people over 80 years of age is as high as 50% and becoming a novel condition with direct life-threatening within the developing world (8) (9). Multiple factors are responsible for muscle glycolipid-metabolism disorder with aging. Reduced oxidative capacity or physical activity with aging both increases the proportion of lipids in body composition, and causes activation of inflammation and insulin resistance(10, 11). Numerous pro-inflammatory cascades are conjunct within muscle and visceral fat, which approach less muscle mass. Additionally, impaired insulin sensitivity can be further increased by muscle catabolism, resulting in abundant ectopic fat deposition within the muscle. Interstitial fibrosis is the other major histopathological change during the progress of sarcopenia, which contributes to the recession of force generation and enhances muscle stiffness.

Glycolipid-metabolism and fibrogenesis appear to be the intersection joint of NAFLD and sarcopenia. With a high degree of functional sharing, crosstalk and mutual regulation, one's metabolic disorders can lead to compensatory or even systemic metabolic disorders (12). Studies have shown that sarcopenia is an important indicator of the severity of NAFLD (13). Therefore, investigating the association and verifying the shared pathways between them provides a prospective way of creating novel age-related disease treatment strategies. (14).

Transcriptome analysis can determine and quantify changes in transcription levels in various states(15). A large number of applications in the life sciences have made transcriptomics widely used (16; 17). As the needs have changed, new techniques for transcriptome studies targeting low cell numbers and even more

accurately targeted sequencing have emerged (18). In disease research, transcriptome technology can help researchers more accurately understand the pathogenesis of diseases and the relationship between specific RNA and diseases. Based on clarifying the precise regulation of various genes in diseases, transcriptome technology can help in the development of new drugs and has important applications in the prevention and treatment of tumors. The integration and analysis of biological data by various bioinformatics tools are important means of life science research. For example, a network algorithm or Random Forest was used to predict patient-related biomarkers (19). Transcriptome data combined with dual disease analysis can be used to better understand the pathological molecular mechanisms between diseases and make more accurate drug predictions (20). The present study aimed to identify hub genes and a hot research topic to the link between NAFLD and Sarcopenia. Therefore, by obtaining transcriptome sequencing data from clinical patients of the two diseases from the GEO database, further joint analysis of their gene expression data was conducted. The differences and commonalities were preliminarily analyzed to clarify the disease characteristics of NAFLD and Sarcopenia. After that, the co-expressed genes of the two diseases were screened. Diversity statistical analysis methods were used to obtain the co-expressed genes and the pathways significantly associated with NAFLD and Sarcopenia. Finally, we integrate the results from the single analysis and intend to provide a basis for subsequent clinical-related research.

2 Materials and methods

2.1 Data processing

Two genome-wide transcriptome profiling using RNA-Seq (GSE167523, GSE167186) of NAFLD and sarcopenia samples were obtained from the GEO database by using Illumina high throughput sequencing platform. 98 NAFLD patients' gene expression profiles formed the GSE167523 data set. The 72 samples in the GSE167186 data set were patients with sarcopenia. The analytic workflow is shown in Figure 1.

2.2 Top 1000 expressed genes selection and gene set variation analysis

Counts in NAFLD and Sarcopenia data sets (GSE167523, GSE167186) were normalized treatment. First, the two data sets were integrated according to gene name. Using the cpm function of package R edgeR (V.3.38.4), Counts per million (CPM) were calculated, and log2 was performed. After the log2 (cpm+1) value is arranged from the largest to the smallest, the first 1000 genes are selected as the top 1000 genes. These genes were analyzed by GSVA (Gene set variation analysis) using R package GSVA (V.1.44.5), The reference gene sets were selected from Homo sapiens C5 (ontology gene sets) in the MSigDB database. Use
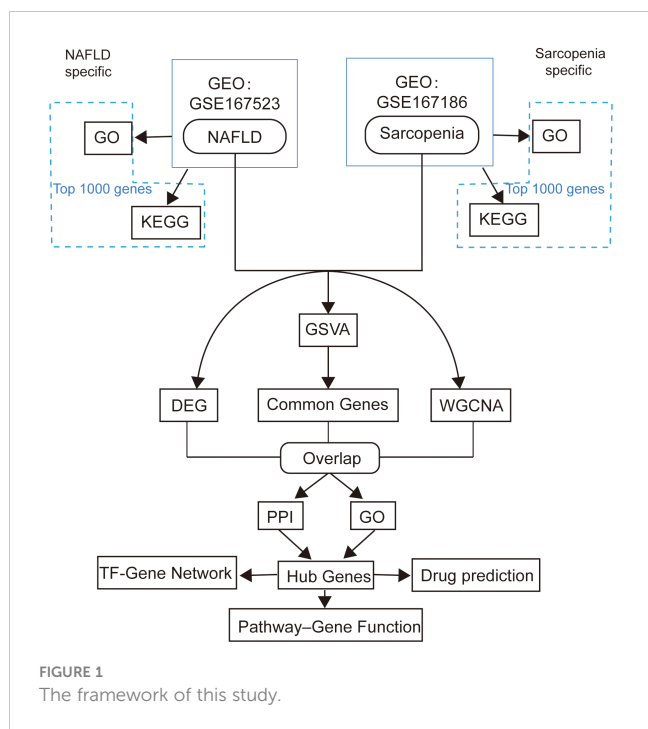
FIGURE 1
The framework of this study.

the GSVA function in the GSVA package. The analysis parameters are method="gsva", kcdf="Gaussian".

2.3 Analysis of inter-sample correlation and differentially expressed genes

When analyzing the correlation between samples, the vst function in package R DEseq2 (V.1.36.0) was first used to standardize the expression matrix. Then dist function was used to calculate the Pearson distance between samples, and the prcomp function was used for Principal Component Analysis (PCA). The DESeq function in using DEseq2 gene counts matrix analysis of differentially expressed genes, DEGs judgment standard for pvalue < 0.05 & (log2FoldChange > = 2 | log2FoldChange < = 2), The common genes are pvalue < 0.05 & (log2FoldChange >= -2 & log2FoldChange <= 2). For volcano mapping, the R package EnhancedVolcano (V.1.14.0) is used.

2.4 Weighted correlation network analysis

The log2 (cpm+1) matrix with an input file as the gene was constructed using an R package called "WGCNA". The power value was determined by the pickSoftThreshold function. Weight coexpression network uses the blockwiseModules function. The plotDendroAndColors function draws the clustering between samples. The labeledHeatmap function shows the correlation between the disease and gene Modules. The plotEigengeneNetworks function shows the correlation between each gene Module.

2.5 Gene Ontology and pathway enrichment analyses

GO is a database established by the Gene Ontology Consortium that provides simple annotations of gene products in terms of function, the biological pathways involved, and their location in the cell. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway is a database dedicated to storing information about genetic pathways in different species. KEGG's Orthopedic Annotated System (KOBAS) (http://kobas.cbi.pku.edu.cn) is a gene/protein functional annotation and functional enrichment Web server developed by Peking University, which collected functional annotation information of 4325 species. "GO terms" and "KEGG pathways" analyses use the "enrichGO" function and "enrichKEGG" function in the R package clusterProfiler (V.4.4.4), respectively, with the p-value cutoff set to 0.05. GO terms-genes network mapping uses the cnetplot function. The R packages GOplot (V.1.0.2) and ggplot2(V.3.3.6) are also used for visualization.

2.6 Determination and functional analysis of hub gene

Search Tool for the Retrieval of Interacting Genes (STRING; http://string-db.org)(11.0 version) Relationships between proteins of interest can be searched, such as direct binding relationships or co-existence of upstream and downstream regulatory pathways, to construct PPI networks with complex regulatory relationships. The TF-genes network is predicted by NetworkAnalyst software. DSigDB database was used to predict possible small-molecule drugs.

3 Results

3.1 GO and KEGG pathway analyses of NAFLD and sarcopenia

We conducted a preliminary analytic and statistical study on the disease data from NAFLD and sarcopenia. First, we performed enrichment analysis on the top 1000 genes from 98 NAFLD patients. The observations demonstrate a strong relationship between fat metabolism and energy metabolism in both molecular function, cellular component, and biological process (Figure 2A). And when we look at the KEGG data, we can see that these genes are related to some relevant metabolic pathways, such as liver alcoholic glycolysis degradation, fatty beta−alanine acid, cytochrome adducts P450, and other related metabolic pathways shown in Figure 2B. Genetic groups related to NAFLD were described.

Then we obtained transcriptome data from 81 patients with sarcopenia and selected the top 1000 genes for the enrichment analysis of GO and KEGG. Sarcopenia genes
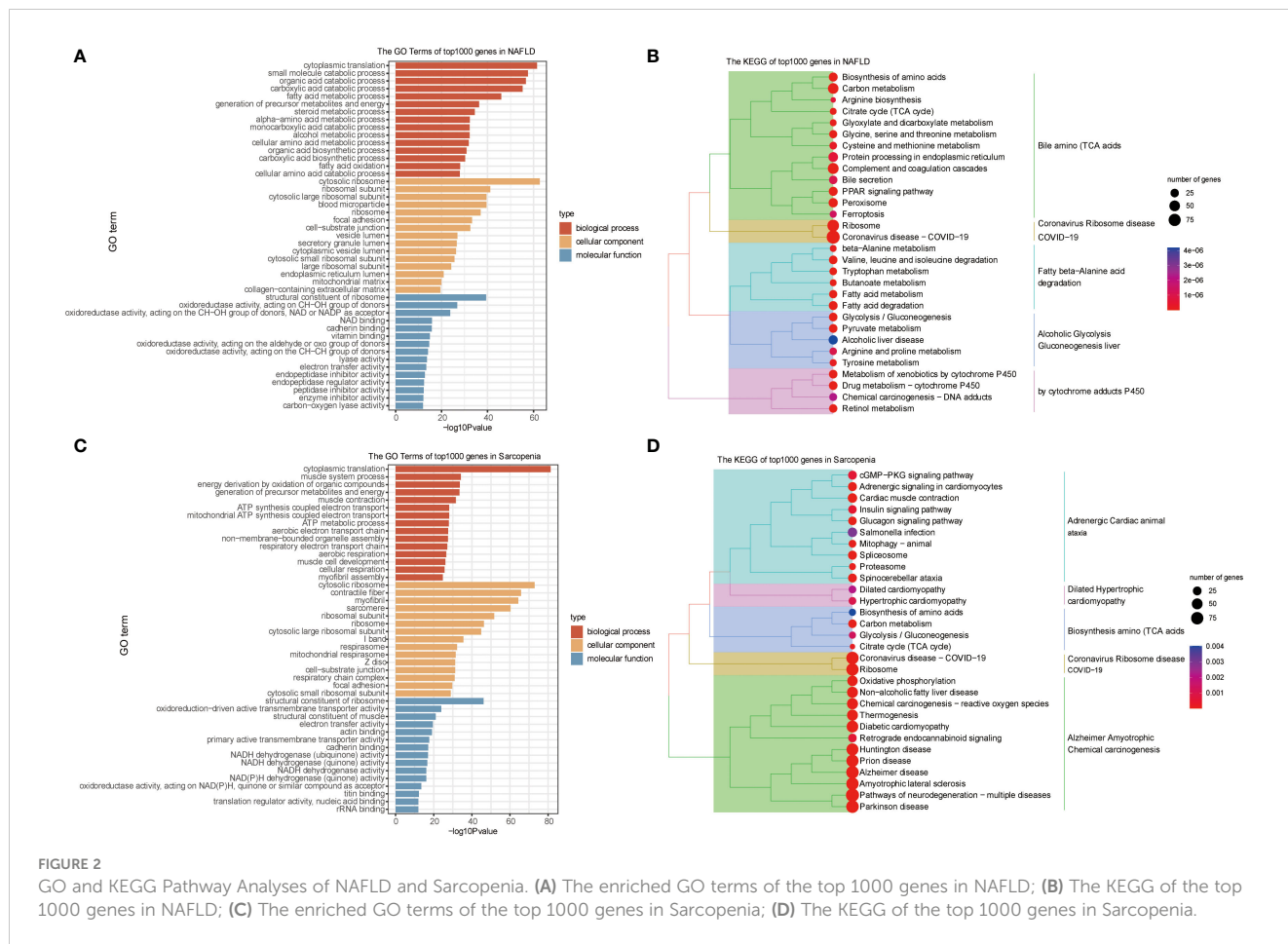
FIGURE 2

GO and KEGG Pathway Analyses of NAFLD and Sarcopenia. **(A)** The enriched GO terms of the top 1000 genes in NAFLD; **(B)** The KEGG of the top 1000 genes in NAFLD; **(C)** The enriched GO terms of the top 1000 genes in Sarcopenia; **(D)** The KEGG of the top 1000 genes in Sarcopenia.

were significantly correlated with energy metabolism and REDOX pathways in molecular function, cellular components, or biological processes (Figure 2C). Dilated Hypertrophic Cardiomyopathy, Biosynthesis of Amino Acids (TCA Acids), and Alzheimer's Amyotrophic Chemical Carcinogenesis were all strongly associated with KEGG enrichment (Figure 2D). This indicates that energy metabolism and redox pathways play a significant role in the disease features of NAFLD and sarcopenia, respectively.

3.2 Gene set variation analysis in NAFLD and sarcopenia

To further investigate the similarity between the two diseases, we integrated the genes of NAFLD and Sarcopenia. The enriched GSVA of the two diseases showed the main pathways as follows, according to the clustering analysis of the genes expressed in the two diseases: Purine nucleotide salvage, Fat-soluble vitamin catabolic process, Lipoxygenase pathway, Long-chain fatty acyl CoA biosynthetic process, Long-chain fatty acyl CoA metabolic process (Figure 3). A high similarity between the two diseases was identified by GSVA.

3.3 Differential gene expression in NAFLD and sarcopenia

Principal component analysis (PCA), which was used to further evaluate the transcriptome of these two diseases, revealed that the differences between the two diseases were more meaningful than the differences between the two diseases themselves (sFigure1A). The Pearson distance analysis, as shown in Figure 4A, further supported this result. Figure 4B indicates the variations in overall gene expression between the two diseases. The statistics show that there are numerous overlap genes between the two diseases, which will need to be further investigated. Searching at the differences between the two diseases and the differentially expressed genes in sarcopenia versus NAFLD, it is fairly obvious from GO terms that the majority of the genes associated with sarcopenia's high expression are those that are participated in muscle system processes, muscle contraction, muscle organ development, and other components of muscle development (Figure 4C). However, the metabolism of small molecules, sterols, alcohol, and other fat and disease-related metabolic pathways were all strongly expressed by NAFLD (Figure 4D). Similar results to those in the GO term were shown in the KEGG enrichment (sFigure 1C). Sarcopenia is mostly overexpressed in pathways
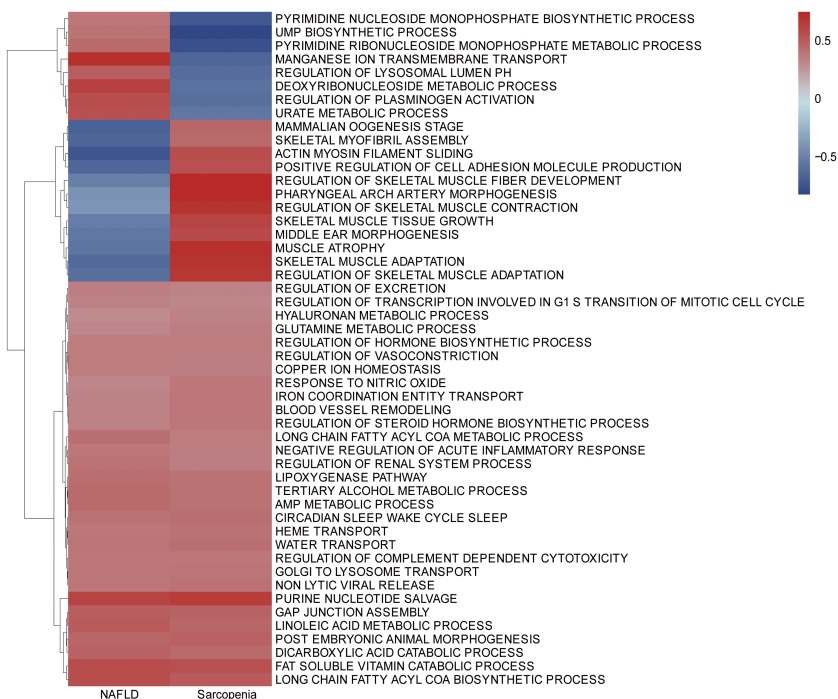
FIGURE 3

Gene Set Variation Analysis (GSVA) in NAFLD and Sarcopenia. GSVA of top 1000 genes in NAFLD and Sarcopenia.

FIGURE 4

Differential gene expression in NAFLD and Sarcopenia. **(A)** The correlation heatmap of NAFLD and Sarcopenia; **(B)** Volcanic map showing the differentially expressed genes of NAFLD and Sarcopenia; **(C)** The GO term of the up-regulated pathways in sarcopenia versus NAFLD; **(D)** The GO term of the up-regulated pathways in NAFLD versus sarcopenia.

linked to muscle growth relative to NAFLD; while compared to sarcopenia, NAFLD is overexpressed in lipid metabolism-related pathways.

3.4 Common genes analyses in NAFLD and sarcopenia

We further investigate the relationship between the two diseases in consideration of the common genes depicted in Figure 4B. The resulting Go term revealed the common genes of NAFLD and Sarcopenia, regardless of their molecular function, cellular component, or biological process, by clustering the shared genes. These mainly enriched processes involve ribonucleoprotein complex biogenesis, ribosome biogenesis, ncRNA processing, histone modification, rRNA metabolic process, transcription coregulator activity, On DNA binding transcription factor binding, and other pathways, which indicates that these two diseases are strongly connected to epigenetic changes (Figure 5A). After evaluating the KEGG enrichment of common genes, we observed that various relevant pathways were enrichment, as well as metabolic pathways of numerous significant diseases (Figure 5B). Two diseases associated with nucleic acid metabolism and epigenetic modifications by GO enrichment.

3.5 Weighted correlation network analysisof NAFLD and sarcopenia

The scale-free network, adjacency matrix, and topological overlap matrix (sFigure 2A) were all constructed after the two groups of data were clustered using the Pearson correlation coefficient. Removal of the outliers, a sample clustering tree (sFigure 2B) was established. Finally, Figure 6A displays 12 modules based on average hierarchical clustering and dynamic
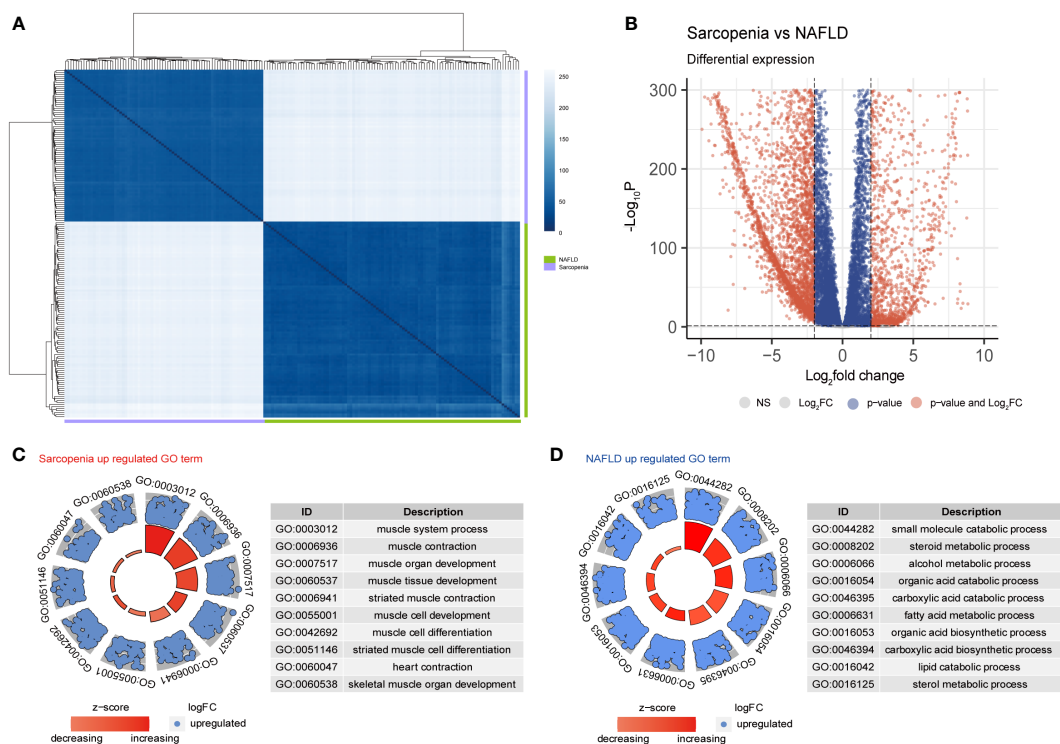
tree pruning (the grey module is often regarded as an undefined module). We found that the blue and turquoise modules, which were selected as clinically significant modules for further analysis, were significantly correlated with NAFLD and sarcopenia. We investigated the connection between characterizing genes. Information regarding the pairing relationships between gene co-expression modules can be obtained from characteristic genes. The characteristic genes were grouped. The results demonstrated that the 11 modules can be grouped into two clusters in Figure 6B and that each of the module combinations (blue and pink, and turquoise and yellow) exhibit a high level of interactive connectedness. We enriched the modules for GO terms by combining them with clinical characteristics (Figure 6C). Blue modules were found to be significantly correlated with histone modification and RNA splicing, whereas brown modules were associated with gastrointestinal diseases, green modules with olfactory dysfunction, gray modules with miRNA regulation, red modules with cofactor 2, and turquoise modules with lipid metabolism. Which, the turquoise module also indicated that the organic acid catabolic process, carboxylic acid catabolic process, small molecule catabolic process, cellular lipid catabolic process, and alcohol metabolic process were significantly correlated (Figure 6D). Blue module revealed that protein methylation, protein alkylation, RNA splicing, and RNA splicing *via* transesterification processes were all strongly related to both diseases (Figure 6E). WGCNA shows that metabolism-related processes and behaviors such as RNA shearing are closely associated with both diseases.

3.6 Protein-protein interaction network

Therefore, intersection analysis was conducted on the genes in the obtained GSVA, DEG of common genes, and the modules obtained by WGCNA. 126 genes were screened out from these

FIGURE 5

Common Genes Analyses in NAFLD and Sarcopenia. **(A)** The GO term of all common genes in NAFLD and Sarcopenia; **(B)** The KEGG of all common genes in NAFLD and Sarcopenia.
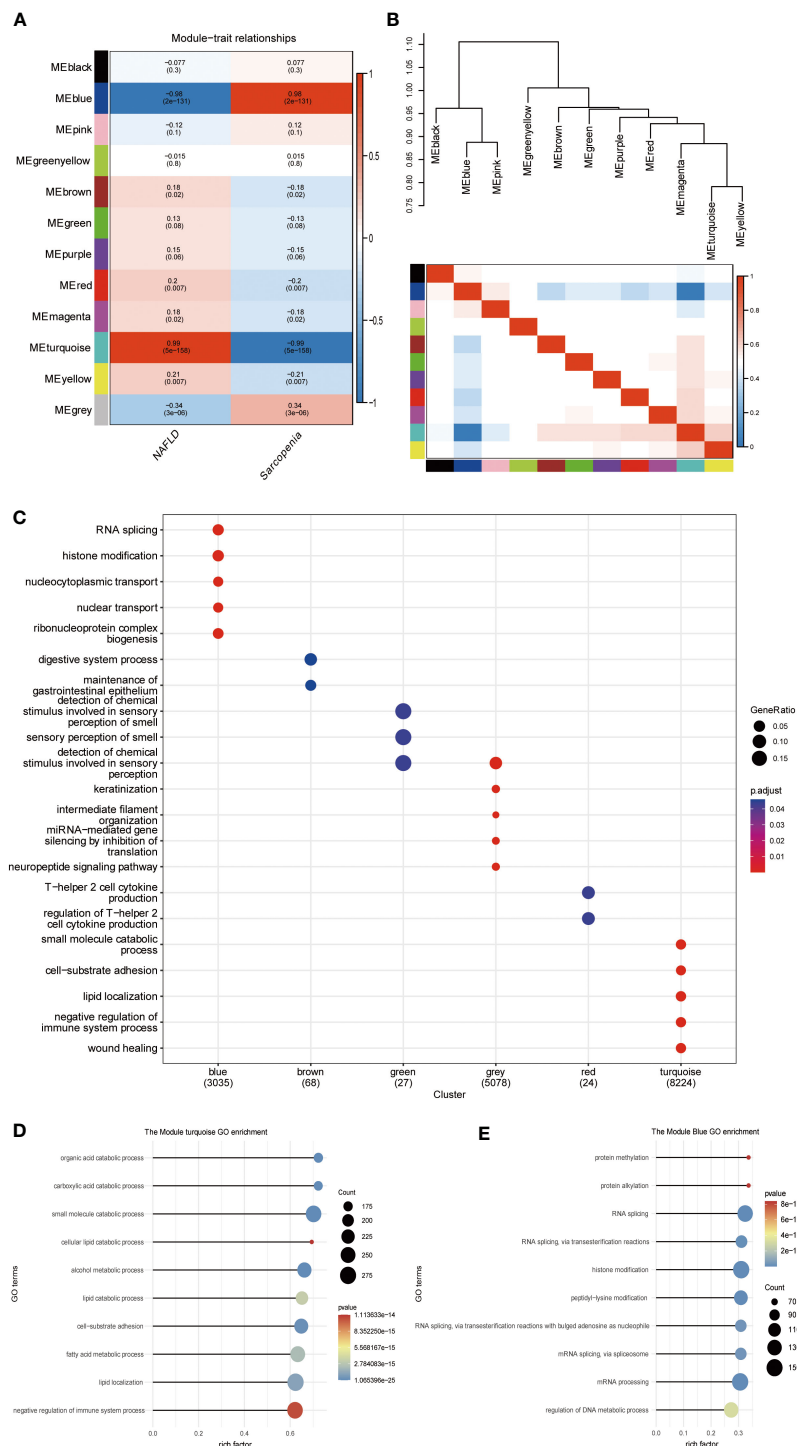
FIGURE 6

WGCNA of NAFLD and Sarcopenia. **(A)** Module−trait associations. Each row corresponds to a module, and each column corresponds to a trait. Each cell contains the corresponding correlation and P value. The table is color-coded by correlation according to the color legend; **(B)** Eigengene dendrogram and eigengene adjacency plot; **(C)** Gene Ontology analysis; **(D)** Gene Ontology analysis of the genes involved in the turquoise module; **(E)** Gene Ontology analysis of the genes involved in the blue module.

intersection genes for subsequent analysis (Figure 7A). Therefore, The PPI network of the intersection DEGs was constructed using String (Figure 7B). We analyzed the enrichment top GO pathway by looking at the GO of intersection genes and found that these genes and blood vessel remodeling, regulation of transcription

involved in G1/S transition of the mitotic cell cycle, regulation of hormone biosynthetic process, and other vascular regulation and hormone anabolic pathways (Figure 7C). Two pairs of genes with high and low expression were filtered out by combining the results of Figures 7B, C. The results showed that the PPI
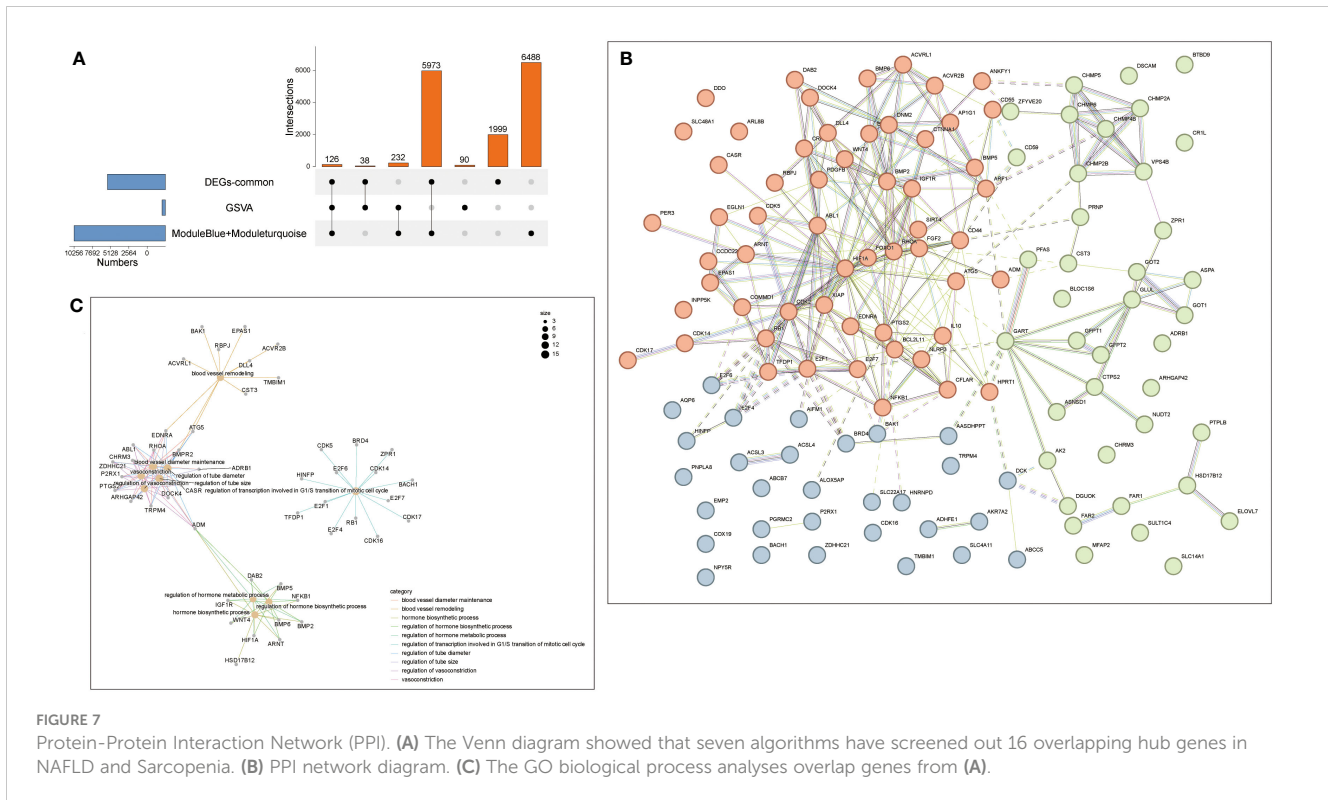
FIGURE 7

Protein-Protein Interaction Network (PPI). **(A)** The Venn diagram showed that seven algorithms have screened out 16 overlapping hub genes in NAFLD and Sarcopenia. **(B)** PPI network diagram. **(C)** The GO biological process analyses overlap genes from **(A)**.

network of these 126 genes correlated with energy, and hormone anabolism.

3.7 Pathway–gene functional network to screen hub gene

The expression patterns of NAFLD and Sarcopenia were found to be positively correlated after reviewing the expression values of all genes. This finding implies that the two diseases may be attached by a co-regulatory network and that further research into the co-regulatory mechanisms of the two diseases is essential (Figure 8A). Figure 8B shows the high expression of *HIF1A* and *ATG5*, as well as *ADM* and *CST3*, in the two diseases. *HIF1A* and *ATG5* were also strongly connected with the reoxidation-reduction of the disease. *ADM* and *CST3* were linked to hormonal disorders. The mutually compatible receptors *BMP2* and *BMPR2*, which are connected to protease hydrolysis, etc., were two pairs of genes with low expression in common. *TFDP1* and *E2F6* were two genes that also have significant regulatory functions in transcription and translation (Figure 8C). Additionally, we constructed a TF-target regulatory network diagram based on the eight-node genes, and Figure 8D clearly illustrates the correlation between the eight genes. With a high degree of linkage, *CST3*, *TFDP1*, *ADM*, and *BMPR2* could be especially noteworthy in the TF-target network. To give thorough treatment for patients who also have NAFLD and sarcopenia, several additional TFs were included, and medicines prediction was also done based on these genes. The top 10 potential medicines were displayed in Table 1. Future treatments for both diseases may be based on these 8 node genes.

4 Discussion

Investigating embryonic development, we learned that mesodermal differentiation is the principal source of muscle formation (21), while the liver consists of endoderm-derived hepatobiliary cell lineage and various mesodermal-derived cells (22), and liver development, from liver specification to liver maturation, requires close interaction with cells of mesodermal origin. This also implies that the progenitor cells of both organs have a strong commonality, and the origin from the same germ layer indicates that they are also functionally very closely related.

To better identify genes that are co-regulated in both diseases, we used the GSVA, common gene in DEGs analysis, and WGCNA analyses to jointly identify genes that are expressed in high abundance in both diseases and involved in disease development. Our research focused on the analysis of the co-expressed genes in both tissues to identify potential therapeutic strategies. We conducted an enrichment analysis for the top 1000 genes associated with each disease after homogenizing the obtained data. While Sarcopenia was more closely associated with energy metabolism and reoxidation reduction, we could see that the principal genes expressed in NAFLD were still associated with lipid metabolism, glycometabolism, and energy metabolism (Figure 2). As the body's primary metabolic organs, the liver and muscle can also be considered as potentially sharing some of the same functions (23). After realizing this commonality, we investigated the variations and consistency of the data in more detail. We proceeded by analyzing the differences between the two diseases and the co-expressed genes. We could see that sarcopenia and NAFLD were distinguishable in that sarcopenia had a higher
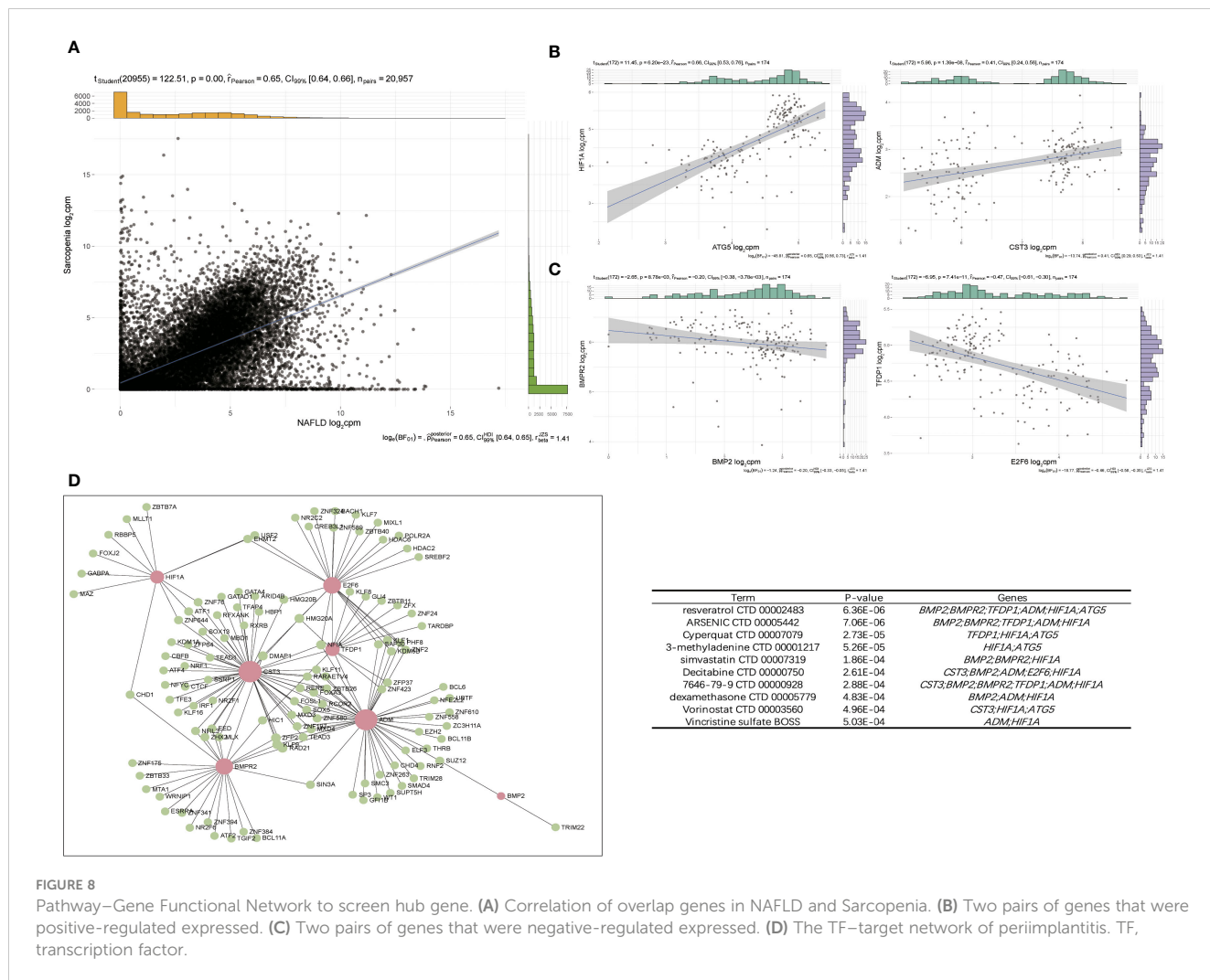
FIGURE 8
Pathway−Gene Functional Network to screen hub gene. **(A)** Correlation of overlap genes in NAFLD and Sarcopenia. **(B)** Two pairs of genes that were positive-regulated expressed. **(C)** Two pairs of genes that were negative-regulated expressed. **(D)** The TF−target network of periimplantitis. TF, transcription factor.

expression of genes mainly related to muscle development. The high gene expression of NAFLD relative to sarcopenia was enriched in the adipose metabolism pathways associated with NAFLD illness itself (Figure 4E), which also reflected the specificities of each disease. This was associated with the gene expression of the muscle itself (Figure 4D). We were particularly interested in the relationship between the two diseases in our research. The major pathways of co-expressed enrichment of these two diseases were identified by GSVA, and the results revealed that the enrichment pathway was not only significantly related to epigenetics but also

TABLE 1 Top 10 drug predictions of hub genes.

Term	Overlap	P-value
resveratrol CTD 00002483	6/1602	6.36E-06
ARSENIC CTD 00005442	5/854	7.06E-06
Cyperquat CTD 00007079	3/160	2.73E-05
3-methyladenine CTD 00001217	2/028	5.26E-05
simvastatin CTD 00007319	3/305	1.86E-04
Decitabine CTD 00000750	5/1801	2.61E-04
7646-79-9 CTD 00000928	6/3095	2.88E-04
dexamethasone CTD 00005779	3/422	4.83E-04
Vorinostat CTD 00003560	3/426	4.96E-04
Vincristine sulfate BOSS	2/86	5.03E-04

involved in other metabolic diseases and immune-related pathways (Figure 3). This suggests that the gene expression of most metabolic diseases is very similar, and the cause of metabolic diseases may be related to REDOX-related pathways (24), epigenetic modification, or the mutual regulation of various RNAs (25, 26), as well as activating the immune system (27, 28).

WGCNA was used for further analysis to check the relationship between these two diseases in more depth. In the highly correlated turquoise and blue modules, we found that the key genes for the two diseases were still enriched in the epigenetic modification and lipid metabolism pathways (Figure 6D, E). Additionally, it is consistent with the preliminary results. In addition to checking more relevant genes and more precisely identifying and verifying the core genes of the two diseases, we chose the intersection of genes obtained by various analysis methods. These 126 genes were shown to be significantly enriched for the cell cycle, angiogenesis, and hormone anabolism pathways. We further checked into the co-expression of these genes and observed 4 pairs of genes out of a large number that were concurrently positive- or negative-regulated. *HIF-1A* activates the transcription of numerous genes, including those involved in energy metabolism, angiogenesis, apoptosis, and other genes whose protein products increase oxygen delivery or facilitate metabolic adaptation to hypoxia, serving as a master regulator of cellular and systemic homeostatic response to hypoxia (29); *ATG5* encoded protein participates in several cellular functions, including the production of autophagic vesicles, mitochondrial quality control following oxidative damage, inhibition of the innate antiviral immune response, and proliferation and development of lymphocytes (30). We also observed that the preprohormone *ADM*, which is produced by this gene, can be broken down into two physiologically active peptides: adrenomedullin and pro-adrenomedullin N-terminal 20 peptide. Adrenomedullin is a 52 AA peptide having a variety of activities, such as vasodilation, hormone secretion regulation, angiogenesis stimulation, and antibacterial action. It also plays a significant role in oxidative stress (31); *CST3* inhibitors appear to have preventive properties in a variety of human fluids and secretions, but they also play a crucial regulatory role in the development of cancer and other diseases (32) (Figure 8B). These genes have a strong connection to the REDOX of the disease. Therefore, synergistic high expression in NAFLD and sarcopenia is significant.

However, there are fewer studies related to the direct occurrence of RNA splicing in NAFLD, but lipid accumulation, as well as obesity, are closely associated with the development of NAFLD. Which is the main cause of increased alternative RNA splicing in the liver. Gene expression data from the liver and muscle of Pihlajamaki et al. provided that obese patients found substantial downregulation of RNA splicing genes, suggesting that the expression of RNA splicing-related genes is negatively associated with liver lipids accumulation and hyperinsulinemia and that altered expression of RNA splicing factors may contribute to obesity-related phenotypes (33). Also, NAFLD, especially sarcopenia, as a disease of the elderly, is significantly associated with increased RNA splicing (34). For example, Li et al. published an article in Cell Metabolism demonstrating that death-associated protein kinase-related apoptosis-inducing kinase-2 (*DRAK2*) can inhibit the

phosphorylation of *SRSF6* by the *SRSF* kinase *SRPK1*, and regulates selective splicing of mitochondrial function-related genes (35).

The activation of BMP signaling in skeletal muscle is significant in maintaining muscle mass as well as muscle-nerve interaction during cachexia and the aging process (36, 37). Restoring *BMP* activity ameliorates cancer-mediated muscle wasting and sarcopenia (36, 37). The activity of *BMP* receptors in muscles induced hypertrophy was dependent on Smad1/5-mediated activation of mTOR signaling (38). *TFDP-1* is a heterodimerization partner for members of the E2F family of transcription factors and up-regulates *E2F*-mediated transcriptional activation (39). *E2F/TFDP-1* regulates the expression of various cellular promoters, particularly gene products that are involved in the cell cycle (40). The combination of *TFDP1* with *E2F*s can promote liver regeneration by regulating *MYCN* transcription (41). Elevated expression of *TFDP1* was associated significantly with larger tumor size and down-regulation of *TFDP1* inhibited the growth of Hep3B cells. In conclusion, overexpression of *TFDP1* may contribute to the progression of some HCCs by promoting the growth of the tumor cells (40). Murine and human HCC data indicate significant correlations of *STMN1* expression with *E2F1/TFPD1* and with *KPNA2* expression and their association with poor prognosis in HCC patients (42). These four genes are negatively regulated and there are opposite regulatory patterns, and we checked their roles and found that the mechanisms are also different in the two diseases.

We also examined the TF regulatory network for these 8 genes, and we found that several of the transcription factors among these genes had strong connections to fibrosis, damage, and fat metabolism (Figure 8D). Resveratrol, which ranked top among such genes to predict small molecule medicines, was discovered to have a beneficial preventative impact on obesity-induced diet in NAFLD and NASH patients (43) (Table 1). It can also improve the validation status of skeletal muscles (44). Resveratrol is also a highly significant healthcare product in daily life, demonstrating the necessity of a daily supplement. In addition, the last few medicines are also widely used. This evidence can support the continued usage of previously prescribed medicines.

When compared to other research, our study still has several limitations. For example, disease development may be regulated at various histological levels, and we have only conducted a preliminary investigation of the co-regulatory mechanisms of NAFLD and sarcopenia at the transcriptome level. For instance, studies on DNA/RNA methylation have been applied to explain how so many diseases develop (39). Our analysis also revealed a strong correlation between both diseases and lipid metabolism as well as oxidative stress, demonstrating the importance of further metabolomic research (26, 45). What's more, the research should really be based on healthy samples to obtain differentially expressed genes and then compare them. However, since healthy human liver and muscle samples are not easy to obtain, we only collected partial liver control datasets, but considering that the liver's gene expression is affected by sex and age (46) (47), we were unable to find a dataset that could be matched exactly. The dataset was not available for muscle. Therefore, the study was mainly conducted on the expression profile of the disease. Therefore, it is necessary to update the data in healthy subjects if they are available subsequently.

In summary, our work demonstrates the potential transcriptome regulatory mechanisms of NAFLD and sarcopenia. Through a thorough mapping of the transcription pattern, the key genes, molecular processes, and potential therapeutic targets that cause NAFLD and sarcopenia were examined. It offers a new perspective and supporting evidence to decrease the high incidence of NAFLD in sarcopenia patients.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be listed below. Repository/Repositories Accession Number Gene Expression Omnibus GSE167523 Gene Expression Omnibus GSE167186.

Author contributions

JY and FY designed the study. ZX and ZY performed data analysis. ZX prepared the figures and tables. ZX, ZY, and SL wrote the manuscript and approved the final draft. ZX, ZY, and ZT participated in data interpretation and analysis. ZX, ZY, and SL were involved in proofreading and deep editing and approved the final manuscript. JY and FY devised the main conceptual idea, supervised the project, performed proofreading and deep editing of the manuscript, and approved the final draft. All authors contributed to the article and approved the submitted version.

Funding

Acknowledgments

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/fendo.2023.1140804/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Differential gene expression in NAFLD and Sarcopenia. (A) The PCA gene expression profile in NAFLD and Sarcopenia patients; (B) The KEGG enrichment of the up-regulated pathway in sarcopenia; (C) The KEGG enrichment of the up-regulated pathway in NAFLD.

SUPPLEMENTARY FIGURE 2
Determination of soft-threshold power in the WGCNA. (A) Analysis of the scale-free index for various soft-threshold powers (β). (B) Analysis of the mean connectivity for various soft-threshold powers.

References

1. Estes C, Razavi H, Loomba R, Younossi Z, Sanyal AJ. Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. *Hepatology* (2018) 67:123–33. doi: 10.1002/hep.29466

2. Collaborators, G. B. D. Cirrhosis. 'The global, regional, and national burden of cirrhosis by cause in 195 countries and territories 1990-2017: a systematic analysis for the global burden of disease study 2017' g. *Lancet Gastroenterol Hepatol* (2020) 5:245–66. doi: 10.1016/S2468-1253(19)30349-8

3. Lazarus JV, Mark HE, Villota-Rivas M, Palayew A, Carrieri P, Colombo M, et al. The global NAFLD policy review and preparedness index: Are countries ready to address this silent public health challenge? *J Hepatol* (2022) 76:771–80. doi: 10.1016/j.jhep.2021.10.025

4. Younossi Z, Tacke F, Arrese M, Chander Sharma B, Mostafa I, Bugianesi E, et al. Global perspectives on nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. *Hepatology* (2019) 69:2672–82. doi: 10.1002/hep.30251

5. Yuan J, Chen C, Cui J, Lu J, Yan C, Wei X, et al. Fatty liver disease caused by high-Alcohol-Producing klebsiella pneumoniae. *Cell Metab* (2019) 30:675–88 e7. doi: 10.1016/j.cmet.2019.08.018

6. Feng J, Qiu S, Zhou S, Tan Y, Bai Y, Cao H, et al. mTOR: A potential new target in nonalcoholic fatty liver disease. *Int J Mol Sci* (2022) 23. doi: 10.3390/ijms23169196

7. Cruz-Jentoft AJ, Sayer AA. Sarcopenia. *Lancet* (2019) 393:2636–46. doi: 10.1016/S0140-6736(19)31138-9

8. Shaw SC, Dennison EM, Cooper C. Epidemiology of sarcopenia: Determinants throughout the lifecourse. *Calcif Tissue Int* (2017) 101:229–47. doi: 10.1007/s00223-017-0277-0

9. Dennison EM, Sayer AA, Cooper C. Epidemiology of sarcopenia and insight into possible therapeutic targets. *Nat Rev Rheumatol* (2017) 13:340–47. doi: 10.1038/nrrheum.2017.60

10. Zhang N, Valentine JM, Zhou Y, Li ME, Zhang Y, Bhattacharya A, et al. Sustained NFkappaB inhibition improves insulin sensitivity but is detrimental to muscle health. *Aging Cell* (2017) 16:847–58. doi: 10.1111/acel.12613

11. Al Saedi A, Debruin DA, Hayes A, Hamrick M. Lipid metabolism in sarcopenia. *Bone* (2022) 164:116539. doi: 10.1016/j.bone.2022.116539

12. Li AA, Kim D, Ahmed A. Association of sarcopenia and NAFLD: An overview. *Clin Liver Dis (Hoboken)* (2020) 16:73–6. doi: 10.1002/cld.900

13. Chun HS, Kim MN, Lee JS, Lee HW, Kim BK, Park JY, et al. Risk stratification using sarcopenia status among subjects with metabolic dysfunction-associated fatty liver disease. *J Cachexia Sarcopenia Muscle* (2021) 12:1168–78. doi: 10.1002/jcsm.12754

14. Li S, Zhou C, Xu Y, Wang Y, Li L, Pelekos G, et al. Similarity and potential relation between periimplantitis and rheumatoid arthritis on transcriptomic level: Results of a bioinformatics study. *Front Immunol* (2021) 12:702661. doi: 10.3389/fimmu.2021.702661

15. Costa V, Angelini C, De Feis I, Ciccodicola A. Uncovering the complexity of transcriptomes with RNA-seq. *J BioMed Biotechnol* (2010) 2010:853916. doi: 10.1155/2010/853916

16. Li Z, Jiang N, Lim EH, Chin WHN, Yeoh AE. Role of transcriptome sequencing in clinical diagnosis of b-cell acute lymphoblastic leukemia. *Leukemia* (2021) 35:2135–37. doi: 10.1038/s41375-021-01185-6

17. Cao S, Lin Y, Yang Z. Vaccinia virus transcriptome analysis by RNA sequencing. *Methods Mol Biol* (2019) 2023:157–70. doi: 10.1007/978-1-4939-9593-6_10

18. Maynard KR, Collado-Torres L, Weber LM, Uytingco C, Barry BK, Williams SR, et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat Neurosci* (2021) 24:425–36. doi: 10.1038/s41593-020-00787-0

19. Li Y, Qi D, Zhu B, Ye X. Analysis of m6A RNA methylation-related genes in liver hepatocellular carcinoma and their correlation with survival. *Int J Mol Sci* (2021) 22. doi: 10.3390/ijms22031474

20. Lv Y, Zhang T, Cai J, Huang C, Zhan S, Liu J. Bioinformatics and systems biology approach to identify the pathogenetic link of long COVID and myalgic Encephalomyelitis/Chronic fatigue syndrome. *Front Immunol* (2022) 13:952987. doi: 10.3389/fimmu.2022.952987

21. Burton PM. Insights from diploblasts; the evolution of mesoderm and muscle. *J Exp Zool B Mol Dev Evol* (2008) 310:5–14. doi: 10.1002/jez.b.21150

22. Yang L, Li LC, Lamaoqiezhong, Wang X, Wang WH, Wang YC, et al. The contributions of mesoderm-derived cells in liver development. *Semin Cell Dev Biol* (2019) 92:63–76. doi: 10.1016/j.semcdb.2018.09.003

23. da Silva Rosa SC, Nayak N, Caymo AM, Gordon JW. Mechanisms of muscle insulin resistance and the cross-talk with liver and adipose tissue. *Physiol Rep* (2020) 8:e14607. doi: 10.14814/phy2.14607

24. Wong BW, Marsch E, Treps L, Baes M, Carmeliet P. Endothelial cell metabolism in health and disease: impact of hypoxia. *EMBO J* (2017) 36:2187–203. doi: 10.15252/embj.201696150

25. Sodum N, Kumar G, Bojja SL, Kumar N, Rao CM. Epigenetics in NAFLD/NASH: Targets and therapy. *Pharmacol Res* (2021) 167:105484. doi: 10.1016/j.phrs.2021.105484

26. Liu JC, Dong SS, Shen H, Yang DY, Chen BB, Ma XY, et al. Multi-omics research in sarcopenia: Current progress and future prospects. *Ageing Res Rev* (2022) 76:101576. doi: 10.1016/j.arr.2022.101576

27. Nelke C, Dziewas R, Minnerup J, Meuth SG, Ruck T. Skeletal muscle as potential central link between sarcopenia and immune senescence. *EBioMedicine* (2019) 49:381–88. doi: 10.1016/j.ebiom.2019.10.034

28. Romero-Gomez M. Non-alcoholic steatohepatitis. *Med Clin (Barc)* (2022) 159:388–95. doi: 10.1016/j.medcli.2022.06.017

29. Rashid M, Zadeh LR, Baradaran B, Molavi O, Ghesmati Z, Sabzichi M, et al. Up-down regulation of HIF-1alpha in cancer progression. *Gene* (2021) 798:145796. doi: 10.1016/j.gene.2021.145796

30. Vara-Perez M, Rossi M, Van den Haute C, Maes H, Sassano ML, Venkataramani V, et al. BNIP3 promotes HIF-1alpha-driven melanoma growth by curbing intracellular iron homeostasis. *EMBO J* (2021) 40:e106214. doi: 10.15252/embj.2020106214

31. Li MY, Zhu XL, Zhao BX, Shi L, Wang W, Hu W, et al. Adrenomedullin alleviates the pyroptosis of leydig cells by promoting autophagy *via* the ROS-AMPK-mTOR axis. *Cell Death Dis* (2019) 10:489. doi: 10.1038/s41419-019-1728-5

32. Kim YI, Shin HW, Chun YS, Cho CH, Koh J, Chung DH, et al. Epithelial cell-derived cytokines CST3 and GDF15 as potential therapeutics for pulmonary fibrosis. *Cell Death Dis* (2018) 9:506. doi: 10.1038/s41419-018-0530-0

33. Pihlajamaki J, Lerin C, Itkonen P, Boes T, Floss T, Schroeder J, et al. Expression of the splicing factor gene SFRS10 is reduced in human obesity and contributes to enhanced lipogenesis. *Cell Metab* (2011) 14:208–18. doi: 10.1016/j.cmet.2011.06.007

34. Kadota Y, Jam FA, Yukiue H, Terakado I, Morimune T, Tano A, et al. Srsf7 establishes the juvenile transcriptome through age-dependent alternative splicing in mice. *iScience* (2020) 23:101242. doi: 10.1016/j.isci.2020.101242

35. Li Y, Xu J, Lu Y, Bian H, Yang L, Wu H, et al. DRAK2 aggravates nonalcoholic fatty liver disease progression through SRSF6-associated RNA alternative splicing. *Cell Metab* (2021) 33:2004–20 e9. doi: 10.1016/j.cmet.2021.09.008

36. Sartori R, Hagg A, Zampieri S, Armani A, Winbanks CE, Viana LR, et al. Perturbed BMP signaling and denervation promote muscle wasting in cancer cachexia. *Sci Transl Med* (2021) 13. doi: 10.1126/scitranslmed.aay9592

37. Uezumi A, Ikemoto-Uezumi M, Zhou H, Kurosawa T, Yoshimoto Y, Nakatani M, et al. Mesenchymal Bmp3b expression maintains skeletal muscle integrity and decreases in age-related sarcopenia. *J Clin Invest* (2021) 131. doi: 10.1172/JCI139617

38. Winbanks CE, Chen JL, Qian H, Liu Y, Bernardo BC, Beyer C, et al. The bone morphogenetic protein axis is a positive regulator of skeletal muscle mass. *J Cell Biol* (2013) 203:345–57. doi: 10.1083/jcb.201211134

39. Huang J, Sun W, Wang Z, Lv C, Zhang T, Zhang D, et al. FTO suppresses glycolysis and growth of papillary thyroid cancer *via* decreasing stability of APOE mRNA in an N6-methyladenosine-dependent manner. *J Exp Clin Cancer Res* (2022) 41:42. doi: 10.1186/s13046-022-02254-z

40. Yasui K, Okamoto H, Arii S, Inazawa J. Association of over-expressed TFDP1 with progression of hepatocellular carcinomas. *J Hum Genet* (2003) 48:609–13. doi: 10.1007/s10038-003-0086-3

41. Fan Z, Kong M, Miao X, Guo Y, Ren H, Wang J, et al. An E2F5-TFDP1-BRG1 complex mediates transcriptional activation of MYCN in hepatocytes. *Front Cell Dev Biol* (2021) 9:742319. doi: 10.3389/fcell.2021.742319

42. Drucker E, Holzer K, Pusch S, Winkler J, Calvisi DF, Eiteneuer E, et al. Karyopherin alpha2-dependent import of E2F1 and TFDP1 maintains protumorigenic stathmin expression in liver cancer. *Cell Commun Signal* (2019) 17:159. doi: 10.1186/s12964-019-0456-x

43. Milton-Laskibar I, Cuevas-Sierra A, Portillo MP, Martinez JA. Effects of resveratrol administration in liver injury prevention as induced by an obesogenic diet: Role of ruminococcaceae. *Biomedicines* (2022) 10. doi: 10.3390/biomedicines10081797

44. Sirago G, Toniolo L, Crea E, Giacomello E. A short-term treatment with resveratrol improves the inflammatory conditions of middle-aged mice skeletal muscles. *Int J Food Sci Nutr* (2022) 73:630–37. doi: 10.1080/09637486.2022.2027889

45. Masoodi M, Gastaldelli A, Hyotylainen T, Arretxe E, Alonso C, Gaggini M, et al. Metabolomics and lipidomics in NAFLD: biomarkers and non-invasive diagnostic tests. *Nat Rev Gastroenterol Hepatol* (2021) 18:835–56. doi: 10.1038/s41575-021-00502-9

46. Marcos R, Correia-Gomes C, Miranda H, Carneiro F. Liver gender dimorphism–insights from quantitative morphology. *Histol Histopathol* (2015) 30:1431–7. doi: 10.14670/HH-11-648

47. Sato S, Solanas G, Peixoto FO, Bee L, Symeonidi A, Schmidt MS, et al. Circadian reprogramming in the liver identifies metabolic pathways of aging. *Cell* (2017) 170:664–77 e11. doi: 10.1016/j.cell.2017.07.042