Journal of Pathology
J Pathol March 2024; **262:** 377–389
Published online 5 January 2024 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/path.6242

ORIGINAL ARTICLE

Integrated transcriptomic landscape of the effect of anti-steatotic treatments in high-fat diet mouse models of non-alcoholic fatty liver disease

Isabel Fuster-Martínez[1,2†], José F Català-Senent[3†], Marta R Hidalgo[3], Francisco J Roig[3,4], Juan V Esplugues[1,2,5], Nadezda Apostolova[1,2,5], Francisco García-García[3*] and Ana Blas-García[2,5,6*]

[1] *Departamento de Farmacología, Universitat de València, Valencia, Spain*
[2] *FISABIO (Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana), Valencia, Spain*
[3] *Computational Biomedicine Laboratory, Principe Felipe Research Center, Valencia, Spain*
[4] *Facultad de Ciencias de la Salud, Universidad San Jorge, Campus Universitario Villanueva de Gállego, Zaragoza, Spain*
[5] *CIBEREHD (Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas), Madrid, Spain*
[6] *Departamento de Fisiología, Universitat de València, Valencia, Spain*

**Co-senior and corresponding authors: A Blas-García, Departamento de Fisiología, Facultad de Medicina, Universidad de Valencia, Avda. Blasco Ibáñez n.15-17, 46010 Valencia, Spain. E-mail: ana.blas@uv.es; F García-García, Computational Biomedicine Laboratory, Principe Felipe Research Center, Calle de Eduardo Primo Yúfera, 3, 46012 Valencia, Spain. E-mail: fgarcia@cipf.es*

†These authors contributed equally to this work.

Abstract

High-fat diet (HFD) mouse models are widely used in research to develop medications to treat non-alcoholic fatty liver disease (NAFLD), as they mimic the steatosis, inflammation, and hepatic fibrosis typically found in this complex human disease. The aims of this study were to identify a complete transcriptomic signature of these mouse models and to characterize the transcriptional impact exerted by different experimental anti-steatotic treatments. For this reason, we conducted a systematic review and meta-analysis of liver transcriptomic studies performed in HFD-fed C57BL/6J mice, comparing them with control mice and HFD-fed mice receiving potential anti-steatotic treatments. Analyzing 21 studies broaching 24 different treatments, we obtained a robust HFD transcriptomic signature that included 2,670 differentially expressed genes and 2,567 modified gene ontology biological processes. Treated HFD mice generally showed a reversion of this HFD signature, although the extent varied depending on the treatment. The biological processes most frequently reversed were those related to lipid metabolism, response to stress, and immune system, whereas processes related to nitrogen compound metabolism were generally not reversed. When comparing this HFD signature with a signature of human NAFLD progression, we identified 62 genes that were common to both; 10 belonged to the group that were reversed by treatments. Altered expression of most of these 10 genes was confirmed *in vitro* in hepatocytes and hepatic stellate cells exposed to a lipotoxic or a profibrogenic stimulus, respectively. In conclusion, this study provides a vast amount of information about transcriptomic changes induced during the progression and regression of NAFLD and identifies some relevant targets. Our results may help in the assessment of treatment efficacy, the discovery of unmet therapeutic targets, and the search for novel biomarkers.

Keywords: transcriptomic signature; MAFLD; therapeutic targets; meta-analysis; treatment efficacy

Received 18 June 2023; Revised 20 October 2023; Accepted 28 November 2023

Conflict of interest statement: JVE has collaborated with Gilead, Abbvie, MSD, and Pfizer. No other conflicts of interest were declared.

Introduction

Non-alcoholic fatty liver disease (NAFLD), also known as metabolic-associated fatty liver disease, represents a major health problem worldwide due to its increasing prevalence and potential severity [1]. Given the lack of available specific pharmacologic treatments for NAFLD, the identification of new therapeutic strategies is an area of intense research [2,3]. Animal models of NAFLD are diverse, and among them high-fat diet (HFD) mouse models represent a fundamental preclinical tool to test novel treatments and define targets [4]. Prolonged HFD models, sometimes modified with fructose or cholesterol overconsumption, have been reported to most closely recapitulate the phenotype of human NAFLD [5].

Recently, detailed analyses of gene transcription changes in human liver tissue representing the full spectrum of NAFLD have provided important insight into the

pathophysiology of progressive fibrosing steatohepatitis and have revealed clinically relevant markers of disease regression that show that transcriptomic changes represent potentially tractable markers of disease progression [6,7]. Unfortunately, various studies have shown substantial differences between human and mouse hepatic transcriptomes, and so transcriptomic analysis of HFD animal models cannot be directly translated into clinical advances. Nevertheless, on a pathway level, gene expression patterns in the livers of HFD-fed mice have been shown to be more closely associated with human fatty liver disease than other models [8].

The aim of the present study was to better characterize the transcriptomic changes of HFD models and how they are modulated by pharmacologic treatments with anti-steatotic effects. To do this, we conducted a systematic review and meta-analysis of liver transcriptomic studies performed in HFD mouse models in order to identify novel NAFLD therapeutic targets and biomarkers.

Materials and methods

Selection of gene expression datasets

This systematic review of the Gene Expression Omnibus (GEO) and ArrayExpress databases was performed with the keywords 'high fat diet', 'nafld', and 'nash' [9,10]. The period reviewed ranged from January 2002 (generation of the public repositories) to May 2020. Results were filtered by: (1) organism: '*Mus musculus*', (2) study type: 'expression profiling by array' or 'expression profiling by high throughput sequencing', (3) tissue: 'liver'. To ensure the quality and transparency of the systematic review, PRISMA statement guidelines were followed [11]. In addition, the following exclusion criteria were applied: (1) information about inclusion criteria not available, (2) a diet with less than 35% kcal of fat, (3) any genetic alteration (*ob/ob*, $Ldlr^{-/-}$, etc.), (4) diets deficient in choline and/or methionine, (5) oophorectomy, (6) presence of other treatments that enhance liver damage, such as carbon tetrachloride or streptozotocin, (7) strain other than C57BL/6J, (8) fewer than three replicates in any diet, (9) diet duration of less than 8 weeks, (10) transcriptomic studies in which only non-coding RNA was analyzed.

Data acquisition and preprocessing

The studies included in the systematic review were downloaded and standardized. Gene names were converted to Gene Symbol nomenclature and the median expression was calculated when there were multiple values for the same gene. Individuals in each study were tagged as HFD, normal diet (ND), or HFD plus treatment (HFDt). After data normalization, an exploratory analysis was performed to detect possible batch effects and anomalous behavior among the data, using clustering and principal component analysis.

Differential gene expression (DGE)

After normalizing the data, different DGE analyses were performed for each individual study to detect differences in gene expression levels between groups. A comparison of HFD versus ND was made to obtain a transcriptomic signature of the pathology, while a comparison of HFDt versus HFD was performed to detect the anti-steatotic effect of the treatments. HFDt and ND were compared to assess the degree of reversion of the detected alterations. When different treatments were present in the same study, we performed a specific DGE for each of them. The limma R package [12] was used to calculate the differential expression statistics. p values were adjusted by the Benjamini & Hochberg method [13].

Meta-analysis of DGE

For the HFD versus ND comparison, meta-analysis techniques were used to obtain a robust transcriptomic HFD signature from the differential expression results of each individual study. Similarly, the DGE results obtained in the HFDt versus HFD comparisons of the evaluated studies were integrated through meta-analysis, to detect the common signal of the effect of the anti-steatotic treatments. The meta-analyses were performed following the methodology described by García-García [14]. In brief, the metafor R package [15] was used to combine the expression results of the corresponding individual studies, using a random-effects model [16]. Thus, the meta-analysis calculated a pooled measure of the expression effect for each gene across all the studies; to do this we used the logarithm of the fold-change (logFC), for which the less variable results were given greater weight, and tested its significance. We then adjusted the p value using the Benjamini & Hochberg method, and established 0.05 as the significance cut-off. To explore the impact of the anti-steatotic treatments on the transcriptomic HFD signature, we performed a correlation analysis of the common significant genes, using Pearson's correlation coefficient and a 95% confidence interval. Two treatments (resveratrol and metformin) were discarded from this analysis due to the low number of significant results.

Functional enrichment

To elucidate the molecular basis underlying the genetic differences found in the DGE, a functional enrichment of each individual study was performed using gene set analysis (GSA). The ontology used was gene ontology (GO) biological processes [17]. GSA was carried out using the logistic regression model implemented in the mdgsa R package [18]. The logarithm of the odds ratio (LOR) is the measure of effect in functional enrichment analysis and is obtained directly from logistic regression. Due to their hierarchical structure, the gene annotations with GO terms applied in the mdgsa package were propagated to inherit the annotations of the ancestor terms. Excessively specific or generic annotations (blocks smaller than 10 or larger than 500 genes, respectively) were

subsequently filtered out. Finally, functions with a Benjamini & Yekutieli-adjusted p value [19] under 0.05 were considered significant.

Meta-analysis of functional enrichment

The GSA results of the individual studies were integrated following the same procedure described above in 'Meta-analysis of DGE'. We performed one meta-analysis for each of the three abovementioned comparisons, in this case obtaining the LOR as a measure of the magnitude of change.

Summary and representation of altered GO terms in meta-analyses

To summarize the significantly altered GO terms in the three meta-analyses, we took advantage of the directed acyclic graph structure of the GO database, in which specific lower-level terms are linked to general higher-level terms. We selected a group of representative high-level GO terms with low overlapping rates, which we called categories, and classified each significant GO term in a category if it was an ancestor of the term. The selected categories were: cell death, cell population proliferation, cellular component organization or biogenesis, cytokine production, homeostatic process, immune system process, ion transport, lipid metabolic process, localization, metabolic process, multicellular organismal process, nitrogen compound metabolic process, regulation of metabolic process, and response to stress. GO terms without any of these ancestors were classified as other.

Cell culture

Hepatocyte-like HepaRG cells were used as a hepatocyte model. Undifferentiated HepaRG cells (HPRGC10) were purchased from Gibco (Thermo Fisher Scientific, Waltham, MA, USA) and cultured and differentiated as described previously [20]. To generate fatty acid (FA)-overloaded hepatocytes, cells were incubated in medium supplemented with 0.5 mM sodium oleate (O7501) and 0.25 mM sodium palmitate (P9767) (both from Sigma-Aldrich, Stenheim, Germany), conjugated with BSA (10775835001, Roche Life Science, Penzberg, Germany) (ratio 6:1), and 0.05 mM L-carnitine (C0158, Sigma-Aldrich) for 48 h.

LX2 cells, a human immortalized line of hepatic stellate cells (HSC), were purchased from Sigma-Aldrich (SCC064) and cultured in high glucose DMEM (41966-029, Gibco, ThermoFisher Scientific), supplemented with 10% heat-inactivated FBS (F7524, Sigma-Aldrich) and 50 µg/ml streptomycin and 50 U/ml penicillin (15070063, Gibco, Thermo Fisher Scientific). These cells were treated for 48 h with the profibrogenic stimulus TGF-β1 (130-095-066, Miltenyi Biotec, 2.5 ng/ml) or its vehicle, DMSO (D2650, Sigma Aldrich).

RNA extraction and RT-qPCR

Total mRNA isolation from cell cultures and real-time RT-PCR were performed as described previously [21]. Primer sequences of human genes were synthesized by Metabion (Planegg, Germany) and are shown in supplementary material, Table S1. *ACTB* (β-actin) was analyzed as a reference transcript. Data were analyzed using a paired t-test comparing the vehicle versus the treatment (FA-overload or TGF-β) of each biological replicate.

Results

Systematic review and selection of studies

Our dataset search identified and screened 415 results in GEO and 170 in ArrayExpress. Of these, 62 studies were assessed for eligibility as they contained at least one treatment with significant anti-steatotic effects in a HFD mouse model. Finally, 21 studies met our inclusion criteria and, therefore, were selected for the meta-analysis (Figure 1) [21–41]. They included liver samples from HFD-fed mice (HFD) and from mice fed with HFD and administered different anti-steatotic treatments (HFDt). Only 16 of these studies also contained liver samples from ND-fed animals. The treatments used in each study and the number of samples per group are detailed in Table 1. HFD murine models used in the selected studies varied in duration, diet composition, and sex and age of the animals, but they all shared the specific characteristics established by our inclusion criteria (supplementary material, Table S2). Further information about the clinical phenotypes achieved by each HFD model (weight gain, insulin resistance, hepatic inflammation, and fibrosis), and the effects on them of the different anti-steatotic treatments were compiled and included in the webtool and in the Zenodo repository (see Data availability statement). Exploratory analyses revealed no batch effect in the datasets.

HFD transcriptome signature

Analysis of differentially expressed genes (DEG) in the selected studies revealed diverse results. As shown in supplementary material, Table S3, most of the studies analyzed the expression of more than 20,000 genes and the number of DEG detected in the HFD versus ND comparison varied from two to 4,586. The integration of these data into the HFD versus ND DGE meta-analysis allowed us to define a transcriptomic signature of HFD mouse models consisting of 2,670 DEG, 2,427 of which were upregulated and 243 of which were downregulated. This HFD transcriptomic signature was very robust, as 2,324 of these DEG were observed in at least eight studies. The LogFC and p values of the HFD versus ND DEG are depicted in the volcano plot in Figure 2A. As expected, the top upregulated DEG included genes associated with steatotic liver (*Ly6d*, *Ubd, Cidea*, *Cfd*) or non-alcoholic steatohepatitis (NASH)

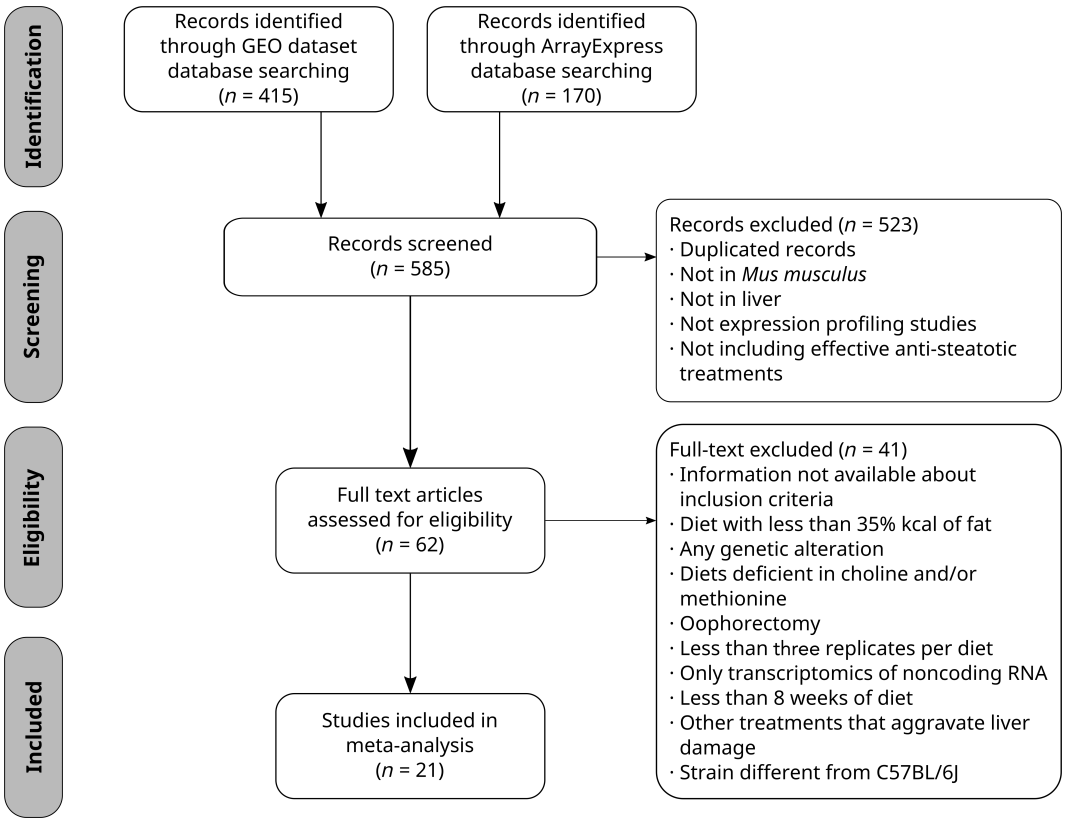


Figure 1. Flow diagram of the systematic review of the literature and selection of the studies for the meta-analysis.

Table 1. Summary of the studies selected for the meta-analysis, indicating the treatment used and the number of samples per group.

GEO ID	Treatment (dose)	Type of product	ND	HFD	HFDt
GSE28384 [22]	Amorfrutin 1 (100 mg/kg/day over the last 3 weeks of HFD)	Natural	-	4	4
GSE38141 [23]	Quercetin (0.05% of diet)	Natural	6	6	6
GSE38854 [24]	Amorfrutin 1 (37 mg/kg/day)	Natural	4	3	4
GSE50763 [25]	Dabigatran (0.01% of diet)	Synthetic	-	3	3
GSE51343 [26]	Quercetin (0.33% of diet)	Natural	-	11	10
GSE51432 [27]	β-Cryptoxanthin (0.003% of diet)	Natural	5	5	5
GSE57054 [28]	Amlexanox (25 mg/kg/day)	Synthetic	-	3	3
GSE57290 [29]	Branched-chain amino acids (3% of diet)	Natural	3	3	3
GSE66711 [30]	Lingoberries (20% of diet)	Natural	6	6	6
	Blackcurrant (20% of diet)	Natural			6
	Bilberries (20% of diet)	Natural			6
GSE84994 [31]	Apical sodium-dependent bile-acid transporter-inhibitor SC-435 (0.006% of diet)	Synthetic	4	4	4
GSE86080 [32]	Green tea extract (0.25% of diet)	Natural	3	3	3
GSE87729 [33]	Methotrexate (5 intraperitoneal injections of 50 mg/kg for 1 week after HFD)	Synthetic	4	4	5
GSE93819 [34]	Vitamin E (0.02% of diet)	Natural	5	5	5
GSE94593 [35]	Liposomal calcitriol (intravenously injected at weeks 12 and 13)	Natural	3	3	3
	Liposomal curcumin (intravenously injected at weeks 12 and 13)	Natural			3
GSE94790 [36]	Metformin (3 mg/kg/day for the last 5 weeks)	Synthetic	3	3	3
GSE108839 [37]	Digoxin (two intraperitoneal injections per week of 1,000 mg/kg)	Natural	3	3	3
GSE120484 [21]	Rilpivirine (5 mg/kg/day)	Synthetic	-	4	4
GSE126204 [38]	Nitro-fatty acids (via osmotic minipump for delivery of polyethylenglycol-solvated OA-NO2 at an infusion rate of 5 mg/kg/day)	Natural	3	3	3
GSE128850 [39]	Pure total flavonoids from citrus (50 mg/kg/day intragastric administration for the last 18 weeks of HFD)	Natural	3	3	3
GSE137365 [40]	D-allulose (5% of diet)	Natural	9	9	9
GSE137840 [41]	Resveratrol (100 mg/kg for 6 weeks)	Natural	4	4	4

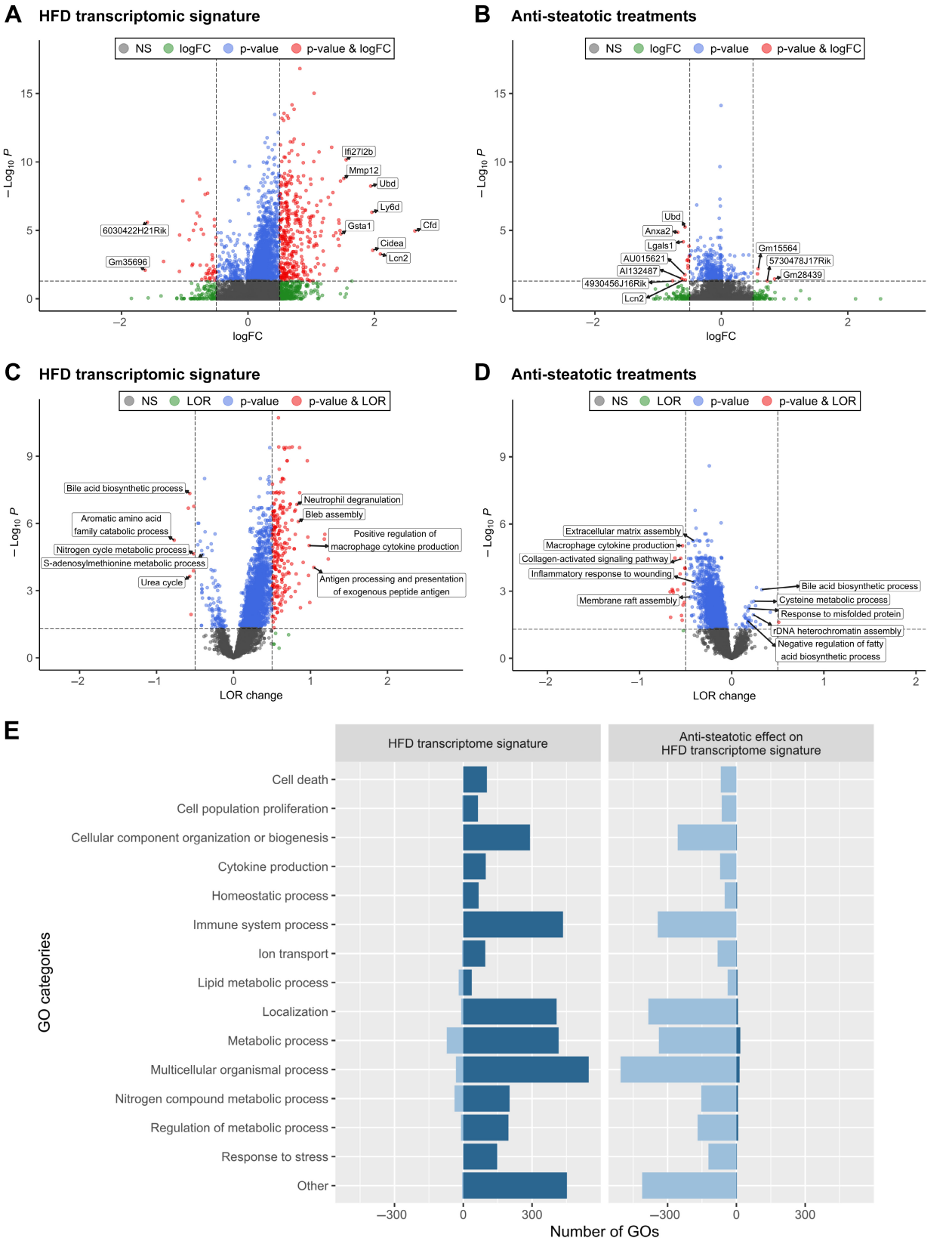


Figure 2. DEG and biological GO terms altered in the different meta-analyses performed. (A) Volcano plot showing the results of the HFD versus ND comparison at the gene expression level. (B) Volcano plot showing the results of the HFDt versus HFD comparison at the gene expression level. (C) Volcano plot showing the GO terms resulting from the HFD versus ND comparison. (D) Volcano plot showing the GO terms resulting from the HFDt versus HFD comparison. Blue dots represent genes or GO terms significantly altered [false discovery rate (FDR) < 0.05], red dots are those significantly changed (FDR < 0.05) and with a logFC/LOR > 0.5 or logFC/LOR < 0.5 and, finally, green dots represent those with a logFC/LOR > 0.5 or logFC/LOR < 0.5 that are not statistically significant. Dots with an absolute value of logFC/LOR > 3 are not displayed. (E) Number of significantly altered GO terms in the HFD transcriptome signature (HFD versus ND) and the effect of anti-steatotic therapies (HFDt versus ND) grouped by representative ancestor GO terms with low overlapping rates. Upregulated and downregulated GO terms are shown in dark and light.

progression (*Lcn2*, *Mmp12*). The HFDt versus HFD DGE meta-analysis revealed the hepatic transcriptomic profile of anti-steatotic treatments in the murine models compared with the one in HFD mice, detecting 54 upregulated DEG and 407 downregulated DEG (corresponding LogFC and *p* values are shown in Figure 2B). Interestingly, some of the top downregulated DEG were involved in the regulation of chronic inflammation (*Lgals1*) and fibrosis (*Lcn2*, *Anxa2*). Similarly, DEG from the HFDt versus ND meta-analysis are shown in supplementary material, Figure S1 and were analyzed later to determine the extent of reversion exerted by treatments.

Biological processes altered by HFD

To better understand the changes induced by HFD, we also carried out a functional meta-analysis of GO biological processes in HFD versus ND. As a result, we found 2,567 GO terms that were modified, 2,448 of which showed increased expression, whereas 119 were downregulated. LOR and *p* values of the HFD versus ND GO terms are depicted in the volcano plot in Figure 2C, which highlights some of the top downregulated processes, including the bile acid biosynthetic process, aromatic amino acid catabolic process, nitrogen cycle metabolic process, urea cycle, and S-adenosylmethionine metabolic process. Of note, a significant number of these processes were associated with the nitrogen cycle and amino acid metabolism.

To summarize this functional profile, we grouped the significantly altered GO terms in categories defined by hierarchically superior GO terms (see Materials and methods). The number of significant terms in each category is shown in Figure 2E. As expected, terms upregulated by HFD included many of those related to metabolic processes, response to stress, cell death, and immune system response, supporting the idea that hepatocyte death and inflammation are induced by HFD. Furthermore, a great number of GO terms were related to increased biogenesis and cell proliferation, which might have been a reparative response to liver damage. On the other hand, the most abundant downregulated GO terms were those related to metabolism. Finally, significantly altered GO terms in the HFDt versus ND meta-analysis are shown in supplementary material, Figure S2.

Impact of treatments with anti-steatotic effect on HFD transcriptome signature

The HFDt versus HFD functional meta-analysis revealed 2,180 modified GO terms, 42 of which were upregulated and 2,138 of which were downregulated. LOR and *p* values of all GO terms are shown in Figure 2D, highlighting representative altered terms that revealed that anti-steatotic treatments not only downregulated processes directly related to lipid accumulation and biosynthesis, but also GO terms associated with advanced stages of the disease, including macrophage cytokine production and extracellular matrix assembly (Figure 2D). When these GO terms were grouped by the selected categories, we observed that these anti-steatotic treatments clearly induced the opposite effects to HFD (Figure 2E).

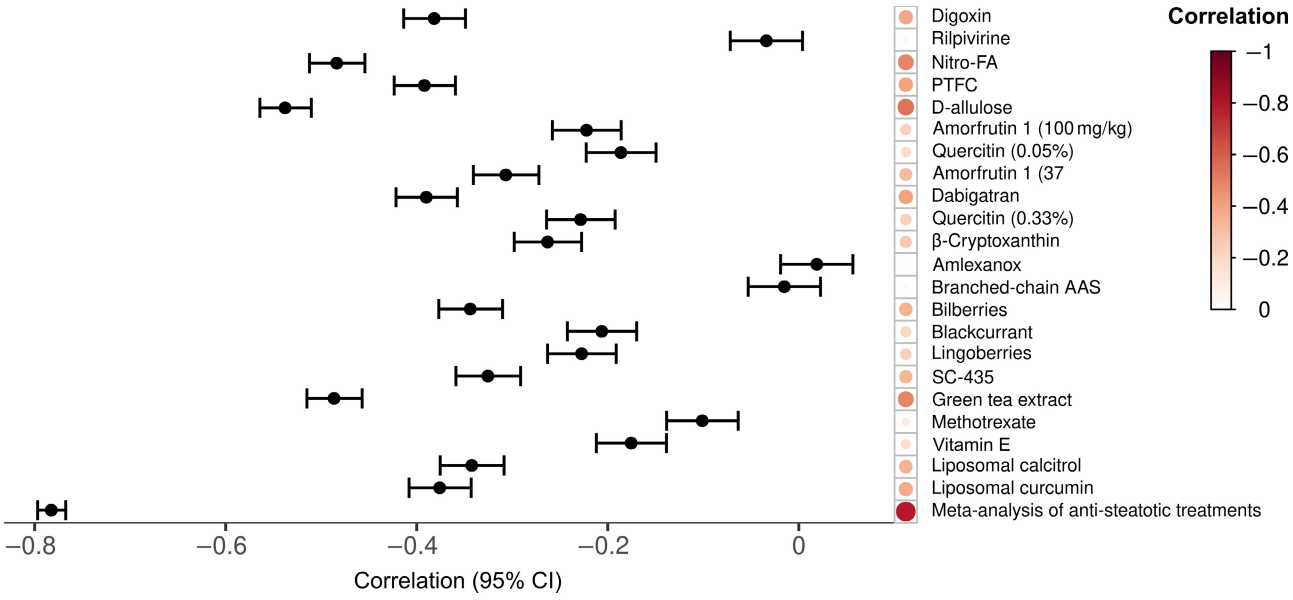


Figure 3. Correlation of the HFD transcriptomic signature with the effect of individual anti-steatotic treatments and their meta-analysis. Correlation coefficients of the HFD transcriptomic signature (HFD versus ND comparison from the meta-analysis) and their confidence intervals are shown for each of the differential expression analyses of the selected anti-steatotic treatments, as well as for the anti-steatotic effect signature (HFDt versus HFD comparison from the meta-analysis). Dot size and color intensity are proportional to the correlation's coefficients. Resveratrol and metformin treatments were not included in the correlation analysis due to the small number of common significant genes (<100).

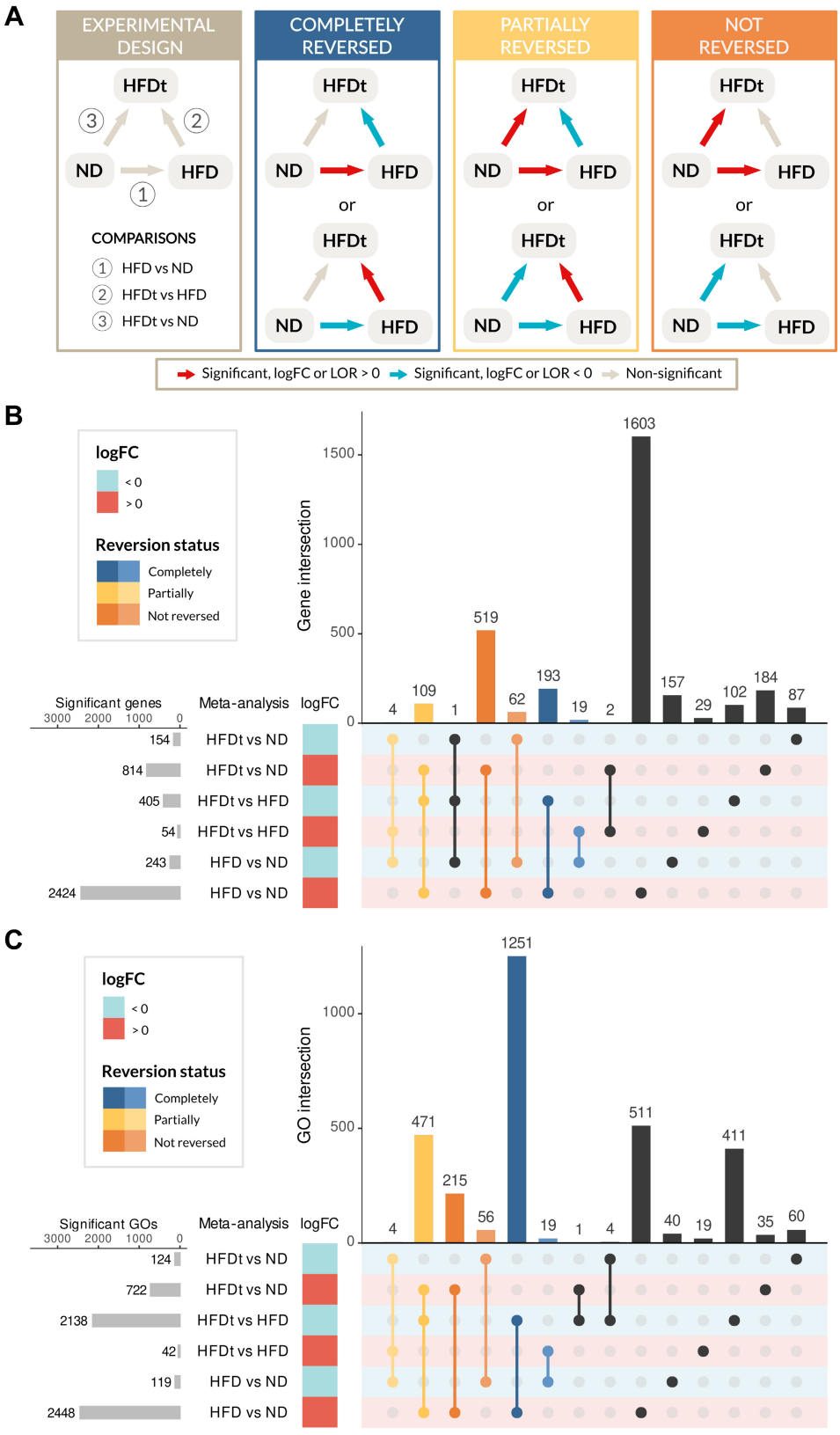


Figure 4. Completely reversed, partially reversed, and not reversed genes and biological processes observed in the different comparisons during the meta-analysis. (A) Graphical representation of the inclusion criteria for reversal definitions. (B and C) UpSet plots comparing (B) genes and (C) biological processes significantly modified in the meta-analysis and grouped according to the sign of logFC or LOR, respectively. Genes or biological processes shown in blue columns are those considered reversed (altered in the HFD versus ND comparison and in the opposite direction in the HFDt versus HFD comparison). Genes or biological processes considered not reversed are shown in orange columns (altered in the HFD versus ND comparison and still modified in the same direction in the HFDt versus ND comparison). Yellow columns indicate genes or biological processes partially reversed (altered in the HFD versus ND comparison and modified in the opposite direction in the HFDt versus HFD comparison, which remained altered in the HFDt versus ND comparison in the same direction as in HFD versus ND). Unclassified terms are depicted in black.

The correlation between our HFD transcriptomic signature (HFD versus ND) and the transcriptomic profile of each treatment with anti-steatotic effect compared to its HFD control (HFDt versus HFD) was generally negative. In fact, a statistically significant negative correlation was detected for 19 of the 22 treatments evaluated. Furthermore, the core set of DEG identified by the HFDt versus HFD meta-analysis had the strongest negative correlation with the HFD transcriptomic signature (Figure 3). This proves that this HFD transcriptomic signature was at least partially reversed by the treatments with an anti-steatotic effect. In addition, different treatments displayed distinct degrees of a negative correlation with the HFD transcriptomic signature, suggesting their variable effectiveness in reversing the transcriptomic alterations induced by HFD.

To better understand the impact of the different treatments on the transcriptome, we used the core sets of DEG and statistically altered biological processes from the different meta-analyses carried out (HFD versus ND, HFDt versus HFD, and HFDt versus ND) to classify them according to the following definitions (Figure 4A):

1. Completely reversed: genes or biological processes significantly altered in the HFD versus ND and HFDt versus HFD comparisons, with opposite directions of change.
2. Not reversed: genes or biological processes significantly modified in the HFD versus ND and HFDt versus ND comparisons, with equal directions of change.
3. Partially reversed: genes or biological processes significantly changed in the three comparisons, in the same direction in HFD versus ND and HFDt versus ND, and in the opposite direction in HFDt versus HFD.
4. Unclassified: genes or biological processes that did not fulfill the criteria to be included in any of the other three categories.

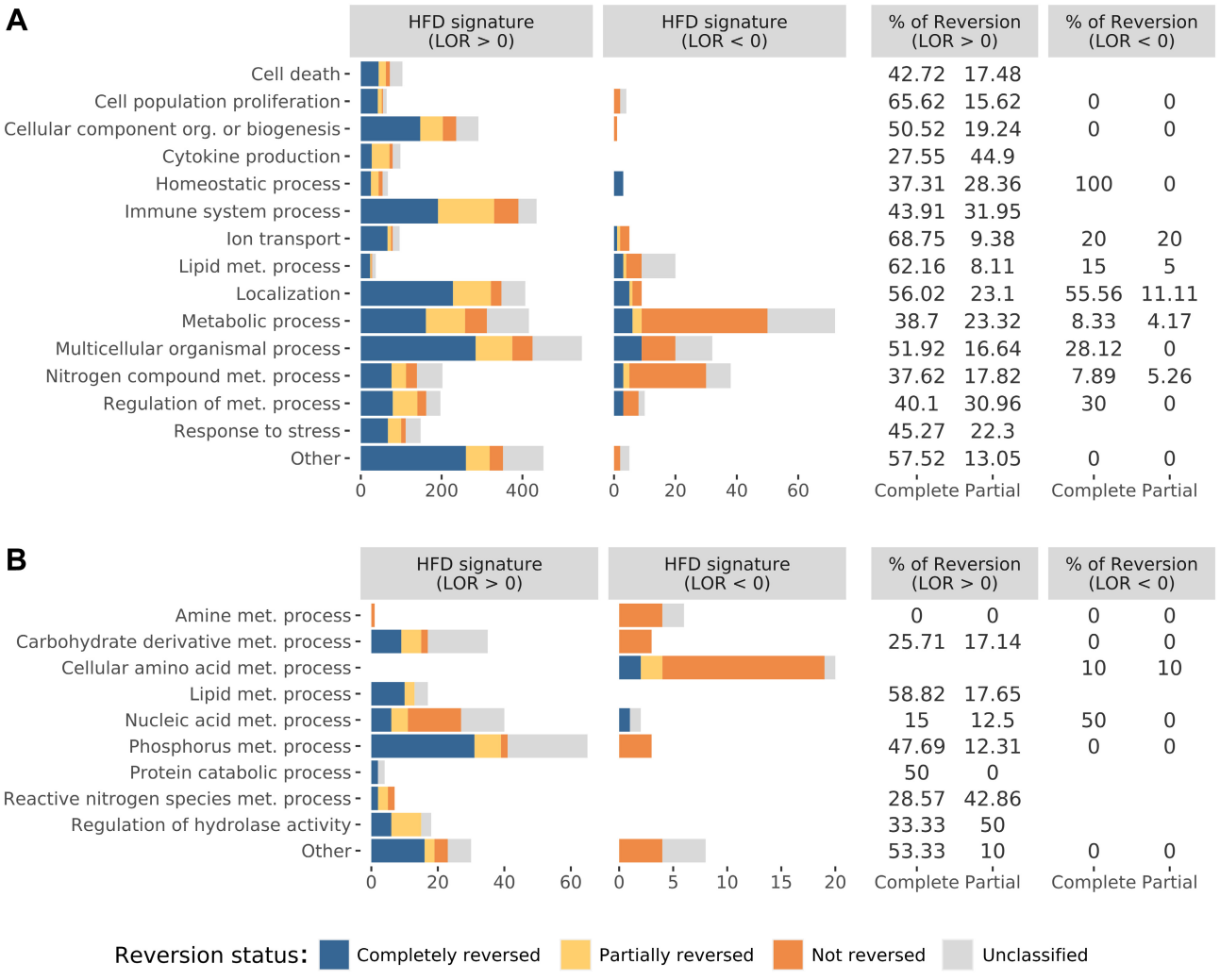
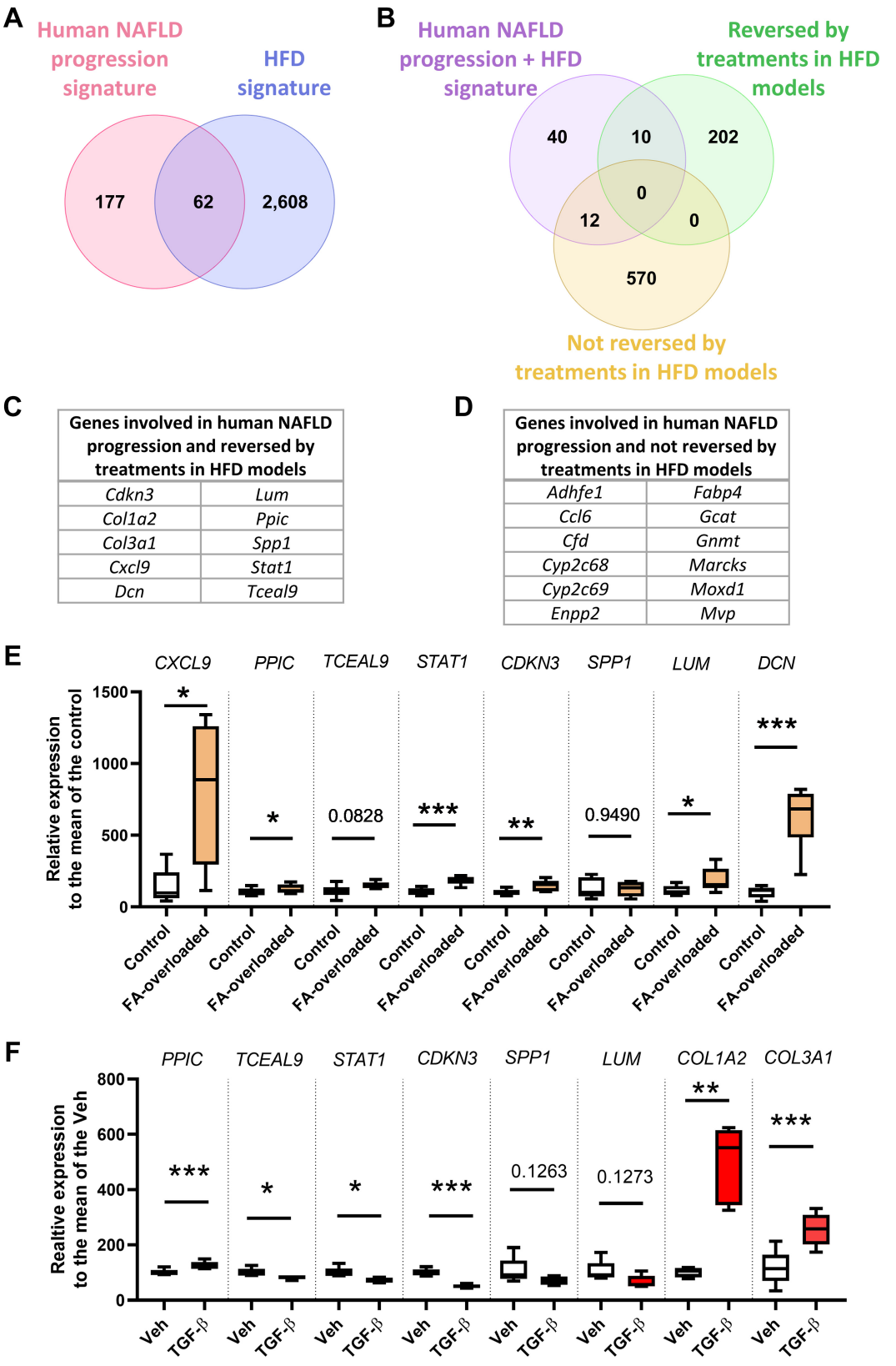


Figure 5. Number of GO terms altered in the HFD signature, classified by common ancestors. (A) The biological processes upregulated or downregulated in our HFD transcriptomic signature were grouped in representative high-level GO terms with low overlapping rates. (B) Upregulated (top) and downregulated (bottom) GO terms included in the nitrogen compound metabolic process term, subsequently grouped by common ancestors. The number of GO terms completely reversed, partially reversed, or not reversed by the treatments within each category are represented in blue, yellow, and orange tones, respectively. Unclassified terms are depicted in gray. Percentages of complete or partial reversion are shown for each category.



Genes involved in human NAFLD progression and reversed by treatments in HFD models	
Cdkn3	*Lum*
Col1a2	*Ppic*
Col3a1	*Spp1*
Cxcl9	*Stat1*
Dcn	*Tceal9*

Genes involved in human NAFLD progression and not reversed by treatments in HFD models	
Adhfe1	*Fabp4*
Ccl6	*Gcat*
Cfd	*Gnmt*
Cyp2c68	*Marcks*
Cyp2c69	*Moxd1*
Enpp2	*Mvp*

Figure 6. Comparison between the HFD murine model and a human NAFLD progression signature. (A) Venn diagram comparing the HFD mouse model transcriptomic signature with a human NAFLD progression transcriptomic signature from the literature [7]. (B) Venn diagram comparing the 62 common genes between human NAFLD progression and HFD mouse model transcriptomic signatures with those whose expression was reversed or not by the anti-steatotic treatments in HFD mouse models according to our meta-analysis. (C and D) Common genes between human NAFLD progression and HFD mouse model transcriptomic signatures whose expression was (C) reversed or (D) not by the treatments. (E and F) Validation of gene expression in *in vitro* models of (E) FA-overloaded human hepatocytes ($n = 6$, mean ± SEM) and (F) human profibrogenic HSC (treated with TGF-β) ($n = 5$, mean ± SEM) for those candidate genes whose expression was reversed by treatments. Data were analyzed using a paired *t*-test (*p value < 0.05, **p value < 0.01, ***p value < 0.001).

The specific number of genes and biological processes in each definition are shown in Figure 4B,C, respectively. A significant proportion of the genes in this HFD transcriptomic signature could not be included in any of these groups, as they were significantly changed in the HFD versus ND comparison but not in the HFDt versus HFD or HFDt versus ND comparisons. Consequently, among the 2,427 DEG upregulated in HFD murine models in our meta-analysis, 193 were completely reversed, 109 were partially reversed, and 519 were not reversed; among the 243 DEG downregulated by HFD murine models, 19 were totally reversed, four were partially reversed, and 62 were not reversed.

Unlike the genes, most of the GO terms could be classified by our definitions. Among the 2,448 GO terms upregulated in HFD, 1,251 were completely reversed by the treatments, 471 were partially reversed, and 215 were not reversed. In addition, among the 119 downregulated by HFD, 19 were completely reversed, four were partially reversed, and 56 were not reversed (Figure 4C). When classified in the selected GO categories (Figure 5), we observed that most of the biological processes upregulated with HFD were reversed among most of the categories. Unsurprisingly, the categories widely reversed by the anti-steatotic treatments included lipid metabolism, cell death, cell proliferation, immune system, cytokine production, and response to stress. On the other hand, the processes downregulated by HFD showed a much lower percentage of reversal, especially in processes related to nitrogen compound metabolism, with the amino acid metabolic process representing the category that included most of the not reversed GO terms.

Comparison of the HFD transcriptomic signature obtained with human data

To evaluate if the HFD mouse models were similar to NAFLD in humans, we compared the HFD model transcriptome signature with a human NAFLD progression signature from another meta-analysis [7] that contained 218 human genes, which were equivalent to 239 murine orthologs. We found that 62 murine orthologs from these human data, which correspond to 60 human genes, were also present in our HFD model signature (Figure 6A). These genes have been associated in humans to specific aspects of NAFLD progression (including the degree of steatosis, inflammation, odds of NASH over simple steatosis, fibrosis, NAFLD activity score, and progression to hepatocellular carcinoma) [7]; furthermore, 36 of these human genes have also been associated with NAFLD progression in a transcriptomic analysis containing samples from 206 NAFLD patients [6] (supplementary material, Table S4). Among them, 10 were included in the group of genes whose expression was reversed by the treatments and 12 were included in the group whose expression was not reversed (Figure 6B); all of these genes are listed in Figure 6C,D. Of note, in the HFDt versus HFD comparison, forest plots revealed the expression of these 10 genes reversed by the treatments to be quite homogeneous in the different studies included in the HFD murine meta-analysis (supplementary material, Figure S3).

Some of the abovementioned reversed genes are already known to participate in the progression of NAFLD (*Stat1*, *Spp1*, *Cxcl9*, *Col1a2*, *Col3a1*), but our analysis revealed other genes whose role in the progression of liver disease is not yet described. To further investigate these genes, we analyzed their expression in *in vitro* models of FA-overloaded hepatocytes (in the human HepaRG cell line) and of HSC activated with TGF-β (in the human LX2 cell line). In line with the results obtained in whole liver samples (and submitted to our meta-analysis), we observed that all the genes found to be reversed by the treatments (Figure 6C) were also significantly upregulated in steatotic hepatocytes, except for *TCEAL9* and *SPP1*, although there was a tendency for *TCEAL9* expression to increase (Figure 6E). Conversely, only the expression of collagens and *PPIC* was increased in activated LX2 cells, whereas the expression of genes involved in cell cycle arrest (*STAT1* and *CDKN3*) was decreased, as was that of the transcription elongation factor *TCEAL9* (Figure 6F).

Discussion

The lack of effective strategies for the early diagnosis and treatment of NAFLD fuels intensive clinical and preclinical research in which validated mouse models are employed to identify novel biomarkers and therapeutic targets. In this study, we used an innovative framework to compare and extract relevant information concerning the impact of anti-steatotic treatments on murine NAFLD. This integrative analysis provides valuable data for generating new hypotheses regarding preclinical HFD models and new treatments for NAFLD. First, we identified a robust transcriptomic signature of HFD models of NAFLD using a meta-analysis of 16 different datasets that only included data from prolonged nutritional models without genetic alterations or concomitant liver insults to mimic the natural progression of human NAFLD. However, the studies we selected displayed some experimental variability, as some combined HFD with high levels of cholesterol, sugar, and/or bile acids. Hence, our transcriptomic signature contains mainly NAFLD-inducing genes that are common to these different diets rather than representing a precise transcriptomic signature for each specific diet. Although other murine HFD transcriptomic signatures have previously been proposed [42], the high number of studies included in this meta-analysis provides more robust data. Importantly, apart from a large number of genes whose participation in the progression of the disease are wet-lab validated in the literature, this signature includes many novel genes whose role in NAFLD has

not previously been described. Although we cannot rule out that a part of these new genes is associated with interspecies differences (murine versus human), our results open new avenues for further research about the pathophysiology of NAFLD.

As expected, the different anti-steatotic treatments studied induced a clear reversion of metabolic processes, especially in upregulated lipid metabolic processes, response to stress, and proinflammatory pathways (immune system process and cytokine production). Conversely, processes related to amino acid metabolism and the urea cycle were not reversed by the treatments, although they were greatly decreased in the livers of HFD animals. It is important to note that previous studies in various NAFLD animal models and clinical settings have demonstrated that liver fibrosis and steatosis are associated with a decrease in the function of urea cycle enzymes and the detoxification of ammonia [43,44]. Furthermore, altered plasma amino acid concentrations have been described in NAFLD patients and are probably associated with insulin resistance and dysregulated protein catabolism [45–47]. Our results suggest that the anti-steatotic treatments analyzed herein do not usually affect these specific biological processes. As the beneficial effects of these treatments do not seem to reverse the alterations induced by diet in these metabolic pathways, they constitute potential targets for the development of new compounds to be used to treat NAFLD in combination with other drugs with anti-steatotic effects. Of note, the heterogeneous nature of the anti-steatotic treatments used hinders the discussion of the multiple potential mechanisms by which they may induce their beneficial effect, as no significant associations were observed when classifying them by any of their known similarities (i.e. natural versus synthetic products). Consequently, future work using these findings should focus on better understanding the plethora of cellular mechanisms altered during NAFLD regression and their potential crosstalks, in order to identify novel therapeutic targets and protective mechanisms.

HFD mouse models are considered the most similar to human NAFLD [5,8], as they recapitulate steatosis, inflammation, and, to some extent, fibrosis, as well as the weight gain and insulin resistance commonly associated with NAFLD in humans [4,5]. Despite this, we detected a substantial difference between our HFD transcriptomic signature and a human transcriptomic signature of NAFLD progression described on a previous meta-analysis [7], which is in line with other reports suggesting that animal models are merely an approximation of human disease and do not recapitulate the transcriptomic changes of this condition in their entirety in humans [8,42,48].

Nevertheless, we detected 62 mouse genes in our HFD transcriptomic signature whose human ortholog was present in the previously reported transcriptomic signature of NAFLD progression [7] and analyzed their behavior in two studies in which patients were classified according to disease progression [6,7]. Some of these genes might be of clinical interest because of their potential as biomarkers and/or new pharmacologic targets for NAFLD. In line with this, *PPIC* expression was upregulated both in human NAFLD progression and in HFD models, and this increase was validated *in vitro* in FA-overloaded hepatocytes and activated HSC. This is relevant, as *PPIC* encodes for an isomerase that seems to be implicated in protein folding. Interestingly, this gene is reported to play an important role in the differentiation of B lymphocytes and invariant NK T cells [49], as well as in hepatic fibrosis, as its knockdown in a CCl_4-induced mouse model and in TGF-β-activated HSC was shown to attenuate fibrosis and reduce extracellular matrix deposition [50]. Together, this evidence encourages a better understanding of its role in hepatocytes and its implication in NAFLD. Another gene found with this approach is *TCEAL9*, which encodes a WW domain binding protein. Although its specific role in liver diseases is unknown, *TCEAL9* has recently been included in a candidate gene signature that correlated with the histologic severity of fibrosis in patients with chronic liver disease [51]. Confirmation of such an association and identification of its specific function in the progression of the disease requires further investigation. Finally, the group of 12 genes that were not reversed in the HFD models may represent an important niche of unexplored therapeutic targets, and their regulation by specific drugs may have beneficial effects in NAFLD, both alone and in combination with other drugs. For instance, *Gnmt*, the main gene involved in liver S-adenosylmethionine catabolism, is downregulated in NAFLD progression and fibrosis [52,53]; curiously, S-adenosylmethionine metabolism was one of the top downregulated processes in our meta-analysis that was not reversed by treatments. Similarly, *Mvp* deficiency has been shown to exacerbate HFD-induced obesity and metabolic inflammation, which correlates with a worsening of glucose and insulin tolerance and fatty liver disease. Given that the mechanism involved is related to the enhancement of NF-κB-mediated signaling, further studies are needed to test whether targeting MVP can be exploited to treat metabolic disorders without altering the host defense system [54].

The fact that most of the studies employed in this meta-analysis included data from only male mice ruled out the possibility of analyzing sex-related differences. Clearly, this is a limitation of our work, which is, however, very common in biomedical research due to the lack of representation of animals of both sexes in numerous studies in public data repositories. Proper incorporation of a sex perspective in research should be encouraged to identify sex-linked disparities and provide more accurate and clinically applicable results.

In conclusion, this meta-analysis provides a vast amount of information about transcriptomic changes induced during the progression and regression of NAFLD. Our results endorse a generalized framework for the assessment of treatment efficacy, the discovery

of unmet therapeutic targets, and the search for novel biomarkers.

Acknowledgements

The authors thank Brian Normanly for his English-language editing and help with the preparation of the manuscript. This research was funded by the Instituto de Salud Carlos III (CIBER CB06/04/0071 and IMP/00019 – co-funded by the European Regional Development Fund of the European Union), Generalitat Valenciana (AICO2021/017 and CIPROM/2021/044), FISABIO (UGP-21-236), and Spanish Ministry of Science and Innovation, co-funded by the European Regional Development Fund of the European Union (PID2021-127945OB-I00 and PID2021-124430OA-I00). IF-M is a recipient of a Predoctoral Trainee Research Grant (FPU18/02435, Ministerio de Educación, Cultura y Deporte).

Author contributions statement

All authors contributed to the study conception and design. Material preparation, data collection and analysis were carried out by IF-M, JFC-S, MRH and FJR. The first draft of the manuscript was written by IF-M, JFC-S, MRH, FG-G and AB-G; all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data availability statement

All datasets analyzed in the current study are available from Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/). All data generated during this study are provided as supplementary material and/or are available in the webtool resource http://bioinfo.cipf.es/metafun-HFD/ and the Zenodo open data repository (http://doi.org/10.5281/zenodo.6669962).

References

1. Younossi Z, Tacke F, Arrese M, *et al*. Global perspectives on nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. *Hepatology* 2019; **69:** 2672–2682.
2. Lee YA, Friedman SL. Inflammatory and fibrotic mechanisms in NAFLD-implications for new treatment strategies. *J Intern Med* 2022; **291:** 11–31.
3. Lazarus JV, Mark HE, Anstee QM, *et al*. Advancing the global public health agenda for NAFLD: a consensus statement. *Nat Rev Gastroenterol Hepatol* 2022; **19:** 60–78.
4. Nevzorova YA, Boyer-Diaz Z, Cubero FJ, *et al*. Animal models for liver disease – a practical approach for translational research. *J Hepatol* 2020; **73:** 423–440.
5. Im YR, Hunter H, de Gracia HD, *et al*. A systematic review of animal models of NAFLD finds high-fat, high-fructose diets most closely resemble human NAFLD. *Hepatology* 2021; **74:** 1884–1901.
6. Govaere O, Cockell S, Tiniakos D, *et al*. Transcriptomic profiling across the nonalcoholic fatty liver disease spectrum reveals gene signatures for steatohepatitis and fibrosis. *Sci Transl Med* 2020; **12:** eaba4448.
7. Ryaboshapkina M, Hammar M. Human hepatic gene expression signature of non-alcoholic fatty liver disease progression, a meta-analysis. *Sci Rep* 2017; **7:** 12361.
8. Teufel A, Itzel T, Erhart W, *et al*. Comparison of gene expression patterns between mouse models of nonalcoholic fatty liver disease and liver tissues from patients. *Gastroenterology* 2016; **151:** 513–525.
9. Barrett T, Wilhite SE, Ledoux P, *et al*. NCBI GEO: archive for functional genomics data sets-update. *Nucleic Acids Res* 2012; **41:** D991–D995.
10. Athar A, Füllgrabe A, George N, *et al*. ArrayExpress update – from bulk to single-cell expression data. *Nucleic Acids Res* 2019; **47:** D711–D715.
11. Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; **151:** 264–269.
12. Ritchie ME, Phipson B, Wu D, *et al*. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; **43:** e47.
13. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Stat Methodol* 1995; **57:** 289–300.
14. García-García F. Métodos de análisis de enriquecimiento funcional en estudios genómicos, PhD Dissertation: Universitat de València 2016.
15. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010; **36:** 1–48.
16. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7:** 177–188.
17. Ashburner M, Ball CA, Blake JA, *et al*. Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat Genet* 2000; **25:** 25–29.
18. Montaner D, Dopazo J. Multidimensional gene set analysis of genomic data. *PLoS One* 2010; **5:** e10348.
19. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 2001; **29:** 1165–1188.
20. Blas-García A, Marti-Rodrigo A, Victor VM, *et al*. The purine analogues abacavir and didanosine increase acetaminophen-induced hepatotoxicity by enhancing mitochondrial dysfunction. *J Antimicrob Chemother* 2016; **71:** 916–926.
21. Martí-Rodrigo A, Alegre F, Moragrega ÁB, *et al*. Rilpivirine attenuates liver fibrosis through selective STAT1-mediated apoptosis in hepatic stellate cells. *Gut* 2020; **69:** 920–932.
22. Weidner C, de Groot JC, Prasad A, *et al*. Amorfrutins are potent antidiabetic dietary natural products. *Proc Natl Acad Sci U S A* 2012; **109:** 7257–7262.
23. Kobori M, Masumoto S, Akimoto Y, *et al*. Chronic dietary intake of quercetin alleviates hepatic fat accumulation associated with consumption of a western-style diet in C57/BL6J mice. *Mol Nutr Food Res* 2011; **55:** 530–540.
24. Meierhofer D, Weidner C, Hartmann L, *et al*. Protein sets define disease states and predict in vivo effects of drug treatment. *Mol Cell Proteomics* 2013; **12:** 1965–1979.
25. Kopec AK, Joshi N, Towery KL, *et al*. Thrombin inhibition with dabigatran protects against high-fat diet-induced fatty liver disease in mice. *J Pharmacol Exp Ther* 2014; **351:** 288–297.
26. den Hil H-v, Elise F, van Schothorst EM, *et al*. Quercetin decreases high-fat diet induced body weight gain and accumulation of hepatic and circulating lipids in mice. *Genes Nutr* 2014; **9:** 418.
27. Kobori M, Ni Y, Takahashi Y, *et al*. β-Cryptoxanthin alleviates diet-induced nonalcoholic steatohepatitis by suppressing inflammatory gene expression in mice. *PLoS One* 2014; **9:** e98294.

28. Reilly SM, Ahmadian M, Zamarron BF, *et al*. A subcutaneous adipose tissue-liver signalling axis controls hepatic gluconeogenesis. *Nat Commun* 2015; **6:** 6047.
29. Takegoshi K, Honda M, Okada H, *et al*. Branched-chain amino acids prevent hepatic fibrosis and development of hepatocellular carcinoma in a non-alcoholic steatohepatitis mouse model. *Oncotarget* 2017; **8:** 18191–18205.
30. Heyman-Lindén L, Seki Y, Storm P, *et al*. Berry intake changes hepatic gene expression and DNA methylation patterns associated with high-fat diet. *J Nutr Biochem* 2016; **27:** 79–95.
31. Rao A, Kosters A, Mells JE, *et al*. Inhibition of ileal bile acid uptake protects against nonalcoholic fatty liver disease in high-fat diet-fed mice. *Sci Transl Med* 2016; **8:** 357ra122.
32. Choi J, Kim YJ, Ryu R, *et al*. Effect of green tea extract on systemic metabolic homeostasis in diet-induced obese mice determined via RNA-seq transcriptome profiles. *Nutrients* 2016; **8:** 640.
33. Myers CE, Hoelzinger DB, Truong TN, *et al*. Chemotherapy can induce weight normalization of morbidly obese mice despite undiminished ingestion of high fat diet. *Oncotarget* 2017; **8:** 5426–5438.
34. Ni Y, Nagashimada M, Zhuge F, *et al*. Astaxanthin prevents and reverses diet-induced insulin resistance and steatohepatitis in mice: a comparison with vitamin E. *Sci Rep* 2015; **5:** 17192.
35. Maradana MR, Yekollu SK, Zeng B, *et al*. Immunomodulatory liposomes targeting liver macrophages arrest progression of nonalcoholic steatohepatitis. *Metabolism* 2018; **78:** 80–94.
36. Guo J, Zhou Y, Cheng Y, *et al*. Metformin-induced changes of the coding transcriptome and non-coding RNAs in the livers of non-alcoholic fatty liver disease mice. *Cell Physiol Biochem* 2018; **45:** 1487–1505.
37. Ouyang X, Han S, Zhang J, *et al*. Digoxin suppresses pyruvate kinase M2-promoted HIF-1α transactivation in steatohepatitis. *Cell Metab* 2018; **27:** 339–350.e3.
38. Rom O, Xu G, Guo Y, *et al*. Nitro-fatty acids protect against steatosis and fibrosis during development of nonalcoholic fatty liver disease in mice. *EBioMedicine* 2019; **41:** 62–72.
39. Hong W, Li S, Wu L, *et al*. Prediction of VEGF-C as a key target of pure total flavonoids from citrus against NAFLD in mice via network pharmacology. *Front Pharmacol* 2019; **10:** 582.
40. Han Y, Yoon J, Choi M. Tracing the anti-inflammatory mechanism/triggers of d-allulose: a profile study of microbiome composition and mRNA expression in diet-induced obese mice. *Mol Nutr Food Res* 2020; **64:** e1900982.
41. Shu L, Hou G, Zhao H, *et al*. Resveratrol improves high-fat diet-induced insulin resistance in mice by downregulating the lncRNA NONMMUT008655.2. *Am J Transl Res* 2020; **12:** 1–18.
42. Cazanave S, Podtelezhnikov A, Jensen K, *et al*. The transcriptomic signature of disease development and progression of nonalcoholic fatty liver disease. *Sci Rep* 2017; **7:** 17193.
43. De Chiara F, Heebøll S, Marrone G, *et al*. Urea cycle dysregulation in non-alcoholic fatty liver disease. *J Hepatol* 2018; **69:** 905–915.
44. Gallego-Durán R, Ampuero J, Pastor-Ramírez H, *et al*. Liver injury in non-alcoholic fatty liver disease is associated with urea cycle enzyme dysregulation. *Sci Rep* 2022; **12:** 3418.
45. Gaggini M, Carli F, Rosso C, *et al*. Altered amino acid concentrations in NAFLD: impact of obesity and insulin resistance. *Hepatology* 2018; **67:** 145–158.
46. Winther-Sørensen M, Galsgaard KD, Santos A, *et al*. Glucagon acutely regulates hepatic amino acid catabolism and the effect may be disturbed by steatosis. *Mol Metab* 2020; **42:** 101080.
47. Masoodi M, Gastaldelli A, Hyötyläinen T, *et al*. Metabolomics and lipidomics in NAFLD: biomarkers and non-invasive diagnostic tests. *Nat Rev Gastroenterol Hepatol* 2021; **18:** 835–856.
48. Jiang C, Li P, Ruan X, *et al*. Comparative transcriptomics analyses in livers of mice, humans, and humanized mice define human-specific gene networks. *Cell* 2020; **9:** 2566.
49. Paiva RS, Ramos CV, Azenha SR, *et al*. Peptidylprolyl isomerase C (Ppic) regulates invariant natural killer T cell (iNKT) differentiation in mice. *Eur J Immunol* 2021; **51:** 1968–1979.
50. Yang X, Shu B, Zhou Y, *et al*. Ppic modulates CCl_4-induced liver fibrosis and TGF-β-caused mouse hepatic stellate cell activation and regulated by miR-137-3p. *Toxicol Lett* 2021; **350:** 52–61.
51. Pantano L, Agyapong G, Shen Y, *et al*. Molecular characterization and cell type composition deconvolution of fibrosis in NAFLD. *Sci Rep* 2021; **11:** 18045.
52. Varela-Rey M, Martínez-López N, Fernández-Ramos D, *et al*. Fatty liver and fibrosis in glycine N-methyltransferase knockout mice is prevented by nicotinamide. *Hepatology* 2010; **52:** 105–114.
53. Fernández-Tussy P, Fernández-Ramos D, Lopitz-Otsoa F, *et al*. miR-873-5p targets mitochondrial GNMT-complex II interface contributing to non-alcoholic fatty liver disease. *Mol Metab* 2019; **29:** 40–54.
54. Ben J, Jiang B, Wang D, *et al*. Major vault protein suppresses obesity and atherosclerosis through inhibiting IKK-NF-κB signaling mediated inflammation. *Nat Commun* 2019; **10:** 1801.

SUPPLEMENTARY MATERIAL ONLINE

Figure S1. Volcano plot showing alterations in gene expression detected in the HFDt versus ND comparison

Figure S2. Bar chart representing the number of significantly altered GO terms in the HFDt versus ND meta-analysis, grouped by GO categories following the hierarchical structure of the GO database

Figure S3. Forest plots showing gene expression in the HFDt versus HFD comparison of those genes that were altered during human NAFLD progression [7] and reversed by the anti-steatotic treatments analyzed in our meta-analysis

Table S1. Primers for human genes used in RT-qPCR

Table S2. Main characteristics of the studies selected for the meta-analysis

Table S3. Number of genes analyzed and significant genes for transcriptomic signature comparison, for each study

Table S4. Relevant information about the human ortholog genes found to be common in our HFD transcriptomic hepatic signature and the signature of human NAFLD progression [7]