

Identification of a core set of model agnostic mRNA associated with nonalcoholic steatohepatitis (NASH)

Tyler Bissoondial¹, Prakash Narayan¹

¹ John F. Kennedy High School, Bellmore, New York

SUMMARY

Numerous studies have shown that nonalcoholic fatty liver disease (NAFLD) progression is greatly affected by dysregulation of the hepatic transcriptome. Studies using high throughput technologies such as RNA Sequencing and microarray have identified multiple dysregulated genes in NAFLD. However, these studies utilized vastly different models resulting in few consensus biomarker genes. Thus, in this investigation, we evaluated various datasets to find genes that are similarly expressed across heterogeneous murine models. We hypothesized that there exists a core set of genes dysregulated in nonalcoholic steatohepatitis (NASH), a subtype of NAFLD characterized by steatosis and inflammation. These genes are involved in glucose and lipid metabolism as well as in the inflammatory responses and are the drivers of the observed tissue pathology. To test this hypothesis, we obtained publicly available Gene Expression Omnibus (GEO) datasets from the National Center for Biotechnology Information (NCBI) database and analyzed them to find differentially expressed genes. We also cross-referenced the expression of these genes with published studies for further validation. Our dataset analyses identified 18 genes up-regulated and four down-regulated in at least six of seven datasets. Of these genes, glucokinase (*Gck*) was up-regulated, and the complement component C8 beta chain (*C8b*) was down-regulated in every murine dataset analyzed. Both exhibited a similar expression pattern in a human NAFLD dataset. Using approaches such as these, we believe that identifying consistently dysregulated mRNA can lead to the discovery of reliable biomarkers and potentially effective therapeutics in humans despite the heterogeneity of the experimental models used.

INTRODUCTION

In developed countries, obesity, insulin resistance, and multiple other environmental and genetic factors have led to an increase in the incidence and prevalence of nonalcoholic fatty liver disease (NAFLD), which starts with a condition wherein the liver accumulates 5-6% fat, called steatosis. This accumulation of fat can lead to inflammation, resulting in nonalcoholic steatohepatitis (NASH). Clinically, NASH and liver fibrosis are rarely symptomatic until the development of advanced disease (1). When left untreated, NASH can lead to liver fibrosis, which is the formation of an abnormally large amount of scar tissue in the liver (1). Cirrhosis or severe

liver scarring can occur when fibrosis becomes widespread and permanent, changing the liver's internal structure and impairing its function (2). The most malignant manifestation of NASH is the development of hepatocellular carcinoma (HCC), or primary liver cancer (3). HCC can develop with or without cirrhosis, although it is more common in patients with cirrhosis (4).

Currently, there is no method to predict if a patient will progress from NAFLD or NASH to fibrosis or HCC. Since it is challenging to study NAFLD in humans due to genetic heterogeneity and ethical constraints, animal models, such as murine models, are used primarily to understand the disease pathology, investigate development mechanisms, and test therapeutic strategies (5). There are different approaches to induce NAFLD in mice, such as dietary modification (high-fat diet or high-fat and sugar diet), genetic modification (leptin receptor-deficient *db/db* mice), or treatment with the steroid hormone dexamethasone (5). Moreover, since the cause of NAFLD is heterogeneous, it is difficult to determine the underlying molecular mechanism, especially when considering the multiple models used. Methods such as microarray and ribonucleic acid sequencing (RNA-Seq) allow for comprehensive profiling of genes associated with diseases (6). Both techniques analyze the transcriptome by measuring the relative mRNA level of hundreds of genes. Using a library of RNA probes, microarray measures hybridization with targeted transcripts. It can only assess the expression of transcripts that are on the curated list of RNA probes. For a more unbiased approach, RNA-Seq employs high-throughput shotgun sequencing of cDNA produced from all of the RNA transcripts in a sample. Unlike microarray, RNA-Seq can identify novel RNA or RNA variants (not limited to the RNA probes) (6).

Over the years, many studies have identified hepatic differentially expressed genes (DEGs) with high-throughput microarrays and RNA-Seq. While DEGs can be easily identified in a specific model used to study NAFLD, identifying genes that are consistently regulated between models have been difficult primarily due to the variability between studies (7). Some studies have tried to analyze multiple datasets from mice with similar histopathological characterization to identify core DEGs. Hou *et al.* analyzed two microarray GEO datasets (GSE52748 and GSE57425) and identified 293 DEGs that were up-regulated and 46 that were down-regulated in both (5). Similarly, Xiang *et al.* identified only 12 DEGs that were common between three different models (high-fat diet and sugar diet, leptin-receptor deficient (*db/db* mice), and dexamethasone-treated) by RNA-Seq (7).

Our study examined five RNA-Seq datasets from mice administered different high-fat diets for different durations to identify common DEGs. In addition, we compared the conserved DEGs from the five RNA-Seq datasets to

microarray DEGs identified by Hou *et al.* to add further selectivity in deriving potential biomarkers for NASH in murine models (5). We hypothesized that a core set of model-agnostic mRNA associated with NASH exists. We further hypothesized that some of these genes would likely be involved in gluconeogenesis, steroid biosynthesis, adipogenesis, and inflammatory responses since these processes are affected in NASH. In this study, we identified a set of genes that were differentially expressed with the same gene profile in multiple RNA-Seq datasets. Several of these genes have been implicated in development and progression of NAFLD.

RESULTS

Histological characterizations of tissues from mice used for RNA sequencing indicated that despite variability in treatments, mice fed high-fat diets developed at least steatosis with respect to the control or normal diets. Mice

used in the various studies however exhibited different stages of NAFLD. The mice used showed steatosis (GSE89296), early steatohepatitis (GSE135050), pronounced hepatocellular ballooning, lipoapoptosis, and progressive fibrosis (GSE164084) (9, 24, 25).

Principal component analysis (PCA) of the publicly available RNA Seq datasets revealed distinct clustering and separation of the control group (normal diet) from the high-fat diets (**Figure 1**). Despite some outliers seen among mice fed a normal diet or among the high-fat diet, there was sufficient separation between groups (normal vs. high fat) to use these datasets for further analyses.

To identify DEGs, we calculated the log₂ fold change (log₂FC) for each gene and performed student t-tests on the log₂ transformed data between dietary treatments. Each dataset contained a large number of genes that were differentially expressed between the two experimental conditions (p -value < 0.05, **Figure 2**). The log₂FC showed the

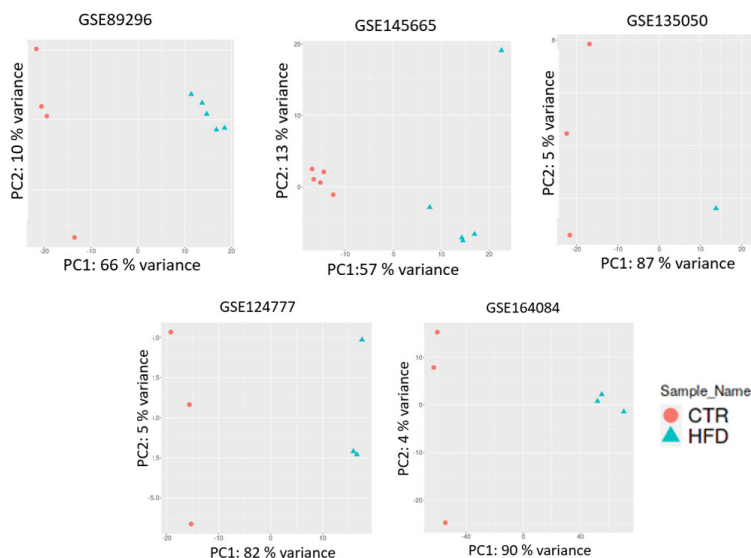


Figure 1: Principal Component Analysis (PCA) of the RNA Seq datasets. PCA plots were generated using iDEP. Normal diet (CTR, red) and High-Fat Diet (HFD, blue) are shown.

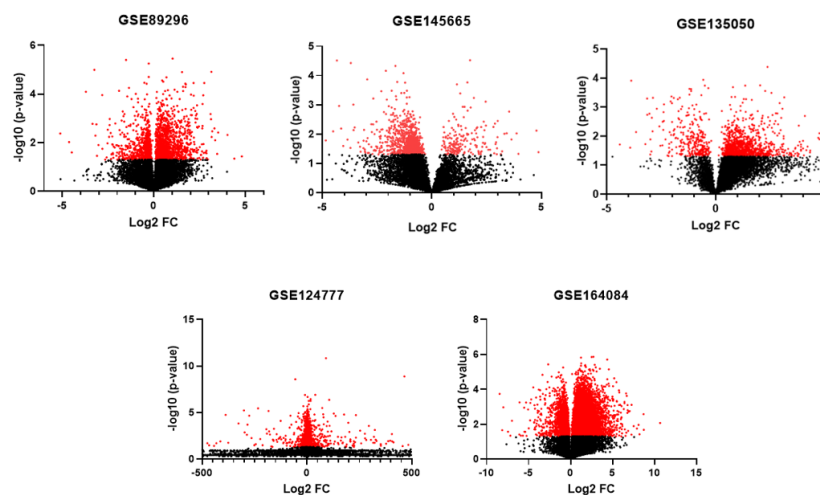


Figure 2: The volcano plots showing the differential expressed genes (DEGs) between the control diet and high-fat diet. Significantly expressed genes ($p < 0.05$, red) between the control and high-fat diet were calculated using Student T-tests. Plots were generated using GraphPad Prism.

same range (within ± 5 log₂FC) for four of the five datasets, except for GSE124777. This dataset *also* had the fewest number of significant DEGs, and the read counts for genes were scaled differently than the other four datasets. Despite the discrepancy, this dataset was included for comparison because of its previous characterization by Shon *et al.* (8).

Of the five datasets, GSE164084 had the most DEGs (>3000) and was previously characterized (9). Mice in this dataset were treated with a high-fat diet for the longest duration (30 weeks) (9). GSE89296 and GSE135050 showed a similar profile expression of differentially expressed genes (Figure 3). We further examined the datasets through hierarchical clustering, which revealed considerable heterogeneity, or variability, among the datasets (Figure 4). As seen in the dendrograms, genes with similar co-expression profiles were placed on the same branch or module of the clustering tree. Though datasets GSE89296 and GSE135050 displayed a similar number of expressed genes, there were more modules

in GSE89296 than in GSE135050. The number of modules does not reflect the number of DEGs but variability among the DEGs. As seen in Figure 4, there was substantial variability in gene expression in these two datasets even though the mice were treated with high-fat diets for the same amount of time (12 weeks) (24, 25). In addition, although GSE164084 had the most differentially expressed genes, it contained the fewest modules and the least amount of clustering, which seems to be affected by other factors besides the amount of time on a high fat diet.

A comparison of genes significantly up-regulated or down-regulated in the five datasets identified 20 that were down-regulated (log₂FC < 0) and 128 that were up-regulated (log₂FC > 0) in four of the five datasets. Among the down-regulated genes, we identified that 5 genes were in the cholesterol biosynthesis and metabolism pathways (*Mvk*, *Pmvk*, *Idl1*, *Slc10a1*, and *Hsd17b7*), 3 were associated with the complement and coagulation cascades (*Serpina1e*, *Serpina1d* and *C8b*), 11 were liver associated expression, and 10 were associated with cell signaling. Of the genes up-regulated, several were involved in the oxidation-reduction process, lipid metabolism, and AMPK signaling. Using the Database for Annotation Visualization and Integrated Discovery (DAVID), we also found that several of these genes were associated with lipid and carbohydrates metabolism (Figure 5). Of the 148 DEGs, only 9 genes showed the same expression profile in all five datasets (Table 1). No gene ontology (GO) enrichment could be found for these 9 DEGs. Our comparison of DEGs to those identified by Xiang *et al.* revealed that *Cyp2a22* was also up-regulated in two of the datasets that were analyzed (high fat and sugar diet and leptin deficient (*db/db*) mice) (7). *C8b*, *Mup7*, and *Serpina1e* were also down-regulated in these two models. None of the genes showed a similar profile in expression in the dexamethasone-treated model.

To further validate the genes identified from these datasets, the 148 DEGs (in four of the five datasets) were compared to genes previously identified in other studies. Hou *et al.* identified 293 genes up-regulated and 46 genes down-regulated in mice fed a high fat diet for 12 weeks, which they

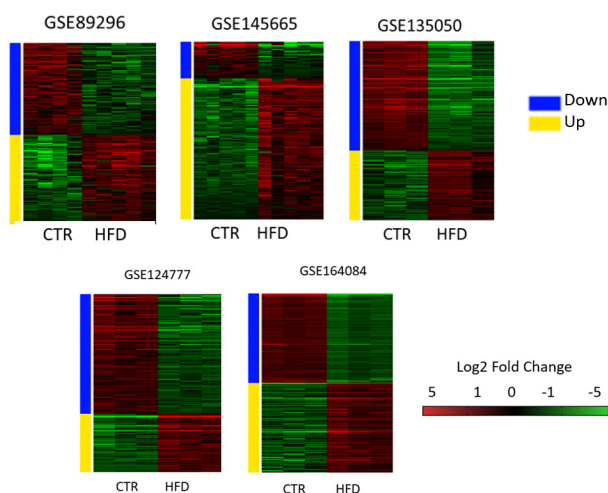


Figure 3: Heatmaps of DEGs between the control diet and high-fat diet. Heatmaps were generated using iDEP v0.94 to show DEGs between control diet (CTR) and High-Fat Diet (HFD).

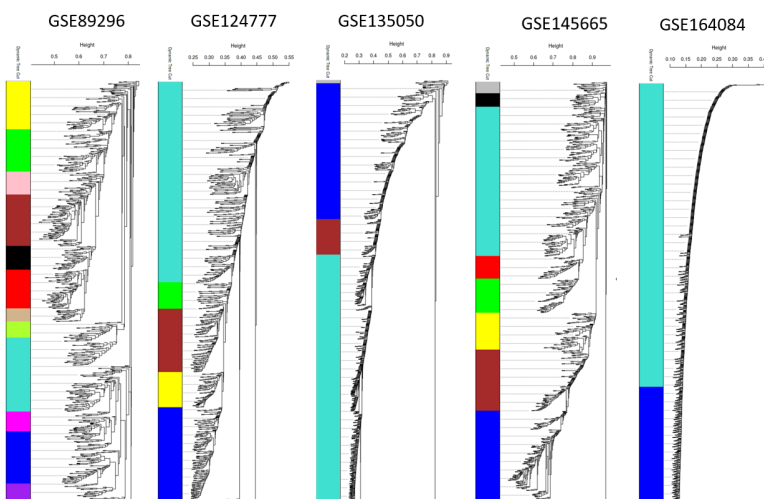


Figure 4: Dendrograms of the RNA Seq Datasets. Dendrograms were generated using iDEP v0.94 to show heterogeneity between datasets and hierarchical clustering among the transcripts expressed. Modules or clades are denoted arbitrarily by color and represents grouping of genes with similar co-expression profiles.

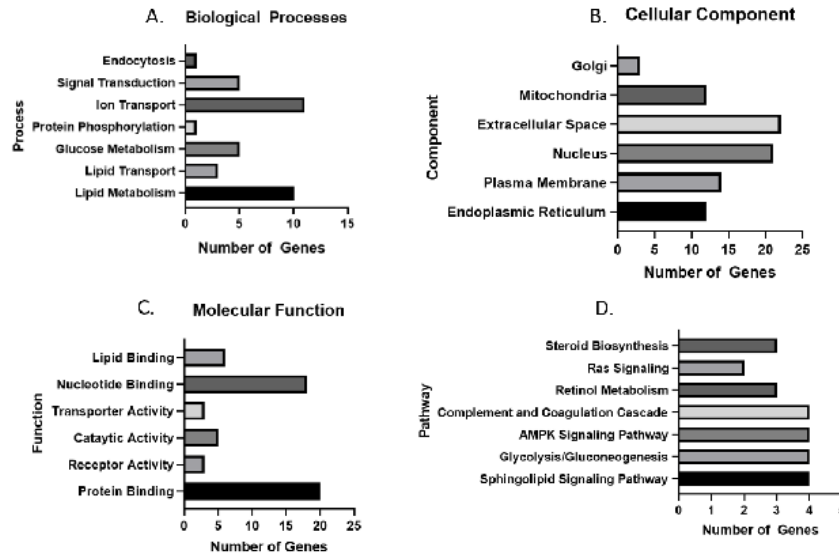


Figure 5: GO Pathways associated with the DEGs. GO pathways were determined using DAVID and organized into graphs. The graphs show the number of significantly associated ($p < 0.05$) genes in (A) biological processes, (B) cellular components, (C) molecular function, and (D) various metabolic pathways that are involved.

DEGs Down-Regulated in All 5 RNA-Seq Datasets	<i>C8b, Mup7, Pmvk, Serpina1e</i>
DEGs Up-Regulated in All 5 RNA-Seq Datasets	<i>Cfd, Cyp2a22, Gbe1, Gck, Cpr98</i>
DEGs Down-Regulated RNA-Seq and Microarray Datasets	<i>Apom, Avpr1a, C8b, Nnmt</i>
DEGs Up-Regulated RNA-Seq and Microarray Datasets	<i>Adora, Cd5l, Cybb, Fam126a, Fitm1, Gck, Gpc1, Ly86</i>

Table 1: Genes that were differentially expressed in the RNA-Seq and microarray datasets. *C8b* and *Gck* were differentially expressed in all seven datasets.

identified by comparing microarray studies (GSE52748 and GSE57425) (5). When we compared the 148 genes identified from the five RNA-Seq datasets to 339 genes identified from two microarray datasets by Hou *et al.*, we found 22 genes consistently differentially expressed among the seven datasets (**Table 1**). Of these genes, *Gck* was up-regulated in all seven datasets and *C8b* was down-regulated in all of the datasets. No significant enrichment was detected in the GO pathways among the 22 genes. We further analyzed the genes using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING), which revealed possible co-expression interactions between a few of the proteins (*Plin4*, *Cidea* and *Fitm1*).

GENEVESTIGATOR® is a database that allows users to

assess the expression of genes transcriptionally from publicly available datasets (microarray or RNA-Seq) (10). To test the validity of the 22 genes identified above, we evaluated them using GENEVESTIGATOR®. The 18 up-regulated genes were also significantly up-regulated with different degrees of expression in the mouse GEO dataset GSE110404 (p -value < 0.05). The 4 down-regulated genes were also significantly down-regulated in GSE110404, indicating that the 22 genes are reliable indicators for NAFLD in mice regardless of the dataset used (**Figure 6**). When these 22 DEGs were evaluated in the human dataset (GSE126848), there was less consistency in expression compared to the murine models. Most notably, conserved expression of *Gck* and *C8b* was found in all murine datasets analyzed. Therefore, the up-

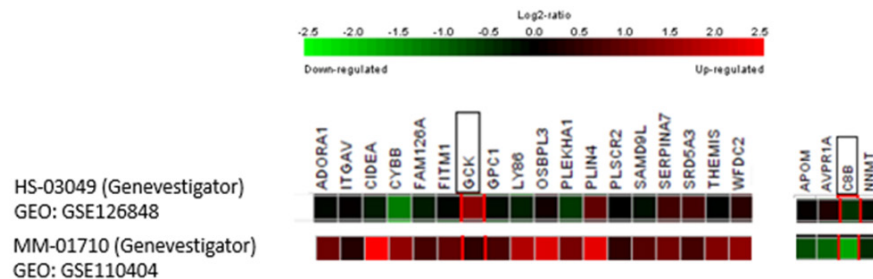


Figure 6: GENEVESTIGATOR® analysis of the 22 DEGs identified. Genes differentially expressed ($p < 0.05$) in RNA-Seq and microarray datasets are shown. The color in the boxes show the log2FC, with brighter colors indicating greater change. *Gck* and *C8b* were differentially expressed with same expression pattern in the murine datasets evaluated.

regulation of *Gck* and the down-regulation of *C8b* could likely be biomarkers for NAFLD progression in both humans and mice (Figure 6).

DISCUSSION

Identifying consistently dysregulated mRNA expression can lead to the discovery of reliable biomarkers and potentially effective therapeutics in humans despite the heterogeneity of the experimental models used. The DEGs identified in this study are related to NASH since the mice used for RNA-Seq displayed various histological manifestation of NAFLD. Moreover, correlating the genes identified from the five RNA-Seq datasets with the two microarray datasets (GSE52748 and GSE57425) analyzed by Hou *et al.* provide further support that these 22 DEGs can be used as potential biomarkers for NAFLD when compared to the controls used in this study (5).

Since NAFLD is a continuous spectrum of pathologies (steatosis, fibrosis, cirrhosis and HCC) and the DEGs identified were from mice with different histological manifestations, these 22 DEGs may be used to assess multiple stages of NAFLD. In fact, several of them have been shown to be also associated with different stages of NAFLD in other studies. Genes that were identified in this study that are associated with hepatic steatosis in mice include *adenosine receptor A1 (ADORA1)* and *nicotinamide N-methyltransferase (NNMT)*. *ADORA1* was shown to be positively correlated with liver steatosis in NAFLD patients (11). *NNMT* is expressed predominantly in the liver and gene variants have been associated with the development of NASH. It has been shown to promote liver steatosis and fibrosis (12). In both humans and mice, an increase in *cell death-inducing DFFA-like effector a (Cidea)* is also associated with an increase in hepatic steatosis. Overexpression of *Cidea* in mice resulted in increased lipid accumulation and knockout of *Cidea* reduced lipid accumulation (15). *Perilin 4 (Plin4)* is a major lipid droplet (LP) protein and is selectively expressed in adipocytes (16). Knockout of *Plin4* in mice resulted in decrease expression of genes involved in lipogenesis in the liver (17). Besides steatosis, DEGs identified in this study has been shown to play a role in development of liver fibrosis and hepatocellular carcinoma. *CD5-like protein (CD5L)* has been shown to be involved in liver fibrosis (19). *Sterile Alpha Motif Domain-containing 9-like (SAMD9L)*, *Steroid 5 alpha-reductase 3 (SRD5A3)*, and *Apolipoprotein M (ApoM)* have been shown to play roles in the development of hepatocellular carcinoma (20-22). Although *Gck* and *C8b* were consistently deregulated

in all of the datasets, there are no studies implicating these genes in specific stages of NAFLD. *Gck* however has been linked to insulin sensitivity in humans and mice (13).

Our investigation had some limitations that need to be addressed in future studies. Not all of the datasets included physiological data describing the stages of NAFLD. Additionally, many of these genes were not validated by other techniques, like qRT-PCR. Future investigations would examine the expression of these genes in various murine models and clinical specimen by qRT-PCR and to determine if these are expressed continuously or in specific stages of NAFLD.

Over the last decade, obesity and insulin resistance have been associated with an increase in the occurrence of NAFLD (14). The exact mechanisms of NAFLD remain unresolved and often conflicting mechanisms are identified as a result of different models. DEGs found in one study may not be shown to be involved in another or have an opposite direction of expression. Moreover, the methods used (RNA-Seq vs. microarrays) can also influence the identification of DEGs. This study identified DEGs regardless of the methodology and the treatment used in the murine models. The 22 DEGs identified in our study shared the same gene profile in many datasets, and several of these have been implicated in numerous studies to be involved in NAFLD. *Gck* and *C8b* seem to be particularly useful biomarkers for assessing NAFLD in both mice and humans, and the *Cidea* and *Plin4* seem to belong to an important hub for lipid localization and metabolism. Therefore, the DEGs identified in this study could be used as potential diagnostic and therapeutic targets for NAFLD.

MATERIALS AND METHODS

Obtaining RNA-Seq Datasets

In order to obtain RNA-Seq datasets, the GEO datasets were queried on the NCBI database. Specific keywords were used to find these datasets, such as “NASH,” “NAFLD,” “high fat diet,” “fast food diet,” and “hepatocellular carcinoma.” Five datasets were identified with accession numbers: GSE145665 (23), GSE135050 (24), GSE89296 (25), GSE124777 (8) and GSE164084 (9). All of the datasets obtained are open-source to the public. A summary of the treatments of mice with high-fat diets is indicated in Table 2. All datasets were downloaded and converted into a .CSV file. Some of the reads in the dataset were corrupted when downloading, indicated by a zero read for each sample of the data. These

GEO Datasets	GSE145665	GSE89296	GSE135050	GSE124777	GSE164084
Species	<i>Mus musculus</i>	<i>Mus musculus</i>	<i>Mus musculus</i>	<i>Mus musculus</i>	<i>Mus musculus</i>
Strain	C57BL/6	C57BL/6N	C57BL/6	C57BL/6J	C57BL/6J
Tissue	Liver	Liver	Liver	Liver	Liver
Age	6 weeks	7 weeks	6 months	5 weeks	8 weeks
Treatment Protocol	High-fat diet (60.9% fat)	High-fat Western diet (40% fat)	High-fat diet (40% fat)	High-fat diet (45% fat)	Fat, fructose, and cholesterol diet
Duration of Treatment	24 weeks	12 weeks	12 weeks	16 weeks	30 weeks
Sample Size	5 control, 5 high-fat diet	4 control, 5 high-fat diet	3 control, 3 high-fat diet	3 control, 3 high-fat diet	3 control, 3 high-fat diet
Library Strategy	RNA-Seq	RNA-Seq	RNA-Seq	RNA-Seq	RNA-Seq

Table 2: Summary of the GEO datasets obtained. Each of the datasets were generated from mice to get relatively uniform data. The treatments of the mice differed but were all relatively similar in that they were fed a high fat diet.

genes were removed from the data entirely, and any excess samples from other treatments that were not being used were removed from the datasets, so the only data present were the control samples and the high-fat diet samples.

Analysis of the GEO datasets

The mRNA counts from the GEO datasets were normalized integrated Differential Expression and Pathway analysis or iDEP (26). iDEP 0.94 is a web-based tool used to analyze RNA-seq data that was used to create visualizations, such as heatmaps and PCA plots, in addition to the previously mentioned normalization plots (26). DEGs were determined using Microsoft Excel using the formula = TTEST ((sham Levels), (high-fat diet Levels), 2, 2) for a 2-sample equal variance, 2 tailed, t-test. The log2FC was calculated using the formula = LOG ((mean NASH level/mean high-fat diet level), 2). The $-\log_{10}$ *p*-value was calculated using the formula = -LOG (*p*-value, 10). Volcano plots were created using GraphPad Prism version 9.1.1. The DAVID v6.8 was used to identify the Gene Ontology (GO) of the differentially expressed genes in multiple datasets. STRING v11.5 was used to create networks of differentially expressed mRNA and determine the pathways these genes play a role in. GENEVESTIGATOR[®] was used to validate the expressions of selected DEGs (10).

Received: December 27, 2021

Accepted: March 3, 2022

Published: October 13, 2022

REFERENCES

- van Koppen, Arianne, *et al.* "Uncovering a predictive molecular signature for the onset of NASH-related fibrosis in a translational NASH mouse model." *Cellular and Molecular Gastroenterology and Hepatology*, vol. 5, no. 1, 6 Oct. 2017, pp. 83-98. doi:10.1016/j.jcmgh.2017.10.001
- Schuppan, Detlef, and Nezam H. Afdhal. "Liver cirrhosis." *The Lancet*, vol. 371, no. 9615, 8 Mar. 2009, pp. 8-14. doi:10.1016/S0140-6736(08)60383-9
- Gerhard, Glenn S., *et al.* "Differentially expressed mRNAs and lncRNAs shared between activated human hepatic stellate cells and nash fibrosis." *Biochemistry and Biophysics Reports*, vol. 22, 12 Mar. 2020. doi:10.1016/j.bbrep.2020.100753
- Friedman, Scott L. "Focus. An emerging conceptual scaffold for NASH." *Journal of Hepatology*, vol. 56, no. 3, 9 Nov. 2011, pp. 511-512. doi:10.1016/j.jhep.2011.11.001
- Hou, Chao, *et al.* "Bioinformatics analysis of key differentially expressed genes in nonalcoholic fatty liver disease mice models." *Gene Expression*, vol. 19, no. 1, 14 Dec. 2018, pp. 25-35. doi:10.3727/105221618X15341831737687
- Mantione, Kirk J., *et al.* "Comparing bioinformatic gene expression profiling methods: microarray and RNA-Seq." *Medical Science Monitor Basic Research*, vol. 20, 23 Aug. 2014, pp. 138-141. doi:10.12659/MSMBR.892101
- Xiang, Liping, *et al.* "Comparison of hepatic gene expression profiles between three mouse models of Nonalcoholic Fatty Liver Disease." *Genes & Diseases*, vol. 9, no. 1, 17 Feb. 2021, pp. 201-215. doi:10.1016/j.gendis.2021.02.008
- Shon, Jong Cheol, *et al.* "Plasma lipidomics reveals insights into anti-obesity effect of Chrysanthemum morifolium Ramat leaves and its constituent luteolin in high-fat diet-induced dyslipidemic mice." *Nutrients*, vol. 12, no. 10, 29 Sep. 2020. doi:10.3390/nu12102973
- Furuta, Kunimaro, *et al.* "Lipid-induced endothelial vascular cell adhesion molecule 1 promotes nonalcoholic steatohepatitis pathogenesis." *The Journal of Clinical Investigation*, vol. 131, no. 6, 21 Jan. 2021. doi:10.1172/JCI143690
- Hruz, T. *et al.* "Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes" *Advances in Bioinformatics*, 7 Apr. 2008, pp. 1-5. doi:10.1155/2008/420747
- Hoyles, Lesley, *et al.* "Molecular phenomics and metagenomics of hepatic steatosis in non-diabetic obese women." *Nature Medicine*, vol. 24, no. 7, 25 Jun. 2018, pp. 1070-1080. doi:10.1038/s41591-018-0061-3
- Komatsu, Motoaki, *et al.* "NNMT activation can contribute to the development of fatty liver disease by modulating the NAD+ metabolism." *Scientific Reports*, vol. 8, no. 1, 5 Jun. 2018, pp. 1-15. doi:10.1038/s41598-018-26882-8
- Bechmann, Lars P., *et al.* "Glucokinase links Krüppel-like factor 6 to the regulation of hepatic insulin sensitivity in nonalcoholic fatty liver disease." *Hepatology*, vol. 55, no. 4, 16 Nov. 2011, pp. 1083-1093. doi:10.1002/hep.24793
- Wang, Jiao, *et al.* "Identification of hub genes and key pathways of dietary advanced glycation end products induced nonalcoholic fatty liver disease by bioinformatics analysis and animal experiments." *Molecular Medicine Reports*, vol. 21, no. 2, 9 Dec. 2019, pp. 685-694. doi:10.3892/mmr.2019.10872
- Zhou, Linkang, *et al.* "Cidea promotes hepatic steatosis by sensing dietary fatty acids." *Hepatology*, vol. 56, no. 1, 25 Jan. 2012, pp. 95-107. doi:10.1002/hep.25611
- Itabe, Hiroyuki, *et al.* "Perilipins: a diversity of intracellular lipid droplet proteins." *Lipids in Health and Disease*, vol. 16, no. 1, 28 Apr. 2017, pp. 1-11. doi:10.1186/s12944-017-0473-y
- Kvam, Ingvild Solberg. "Characterisation of Plin4 null mice-Absence of Plin4 results in lower hepatic expression of genes involved in lipogenesis." *University of Oslo*, May 2019. Master's thesis
- Su, Shaoyong, *et al.* "DNA methylation of the LY86 gene is associated with obesity, insulin resistance, and inflammation." *Twin Research and Human Genetics*, vol. 17, no. 3, 15 Apr. 2014, pp. 183-191. doi:10.1017/thg.2014.22
- Bárcena, Cristina, *et al.* "CD5L is a pleiotropic player in liver fibrosis controlling damage, fibrosis and immune cell content." *EBioMedicine*, vol. 43, 8 May 2019, pp. 513-524. doi:10.1016/j.ebiom.2019.04.052
- Wang, Qun, *et al.* "SAMMD9L inactivation promotes cell proliferation via facilitating G1-S transition in hepatitis B virus-associated hepatocellular carcinoma." *International Journal of Biological Sciences*, vol. 10, no. 8, 17 Jul. 2014, pp. 807-816. doi:10.7150/ijbs.9143
- Mai, Qicong, *et al.* "Steroid 5 alpha-reductase 3 (SRD5A3) promotes tumor growth and predicts poor survival of human hepatocellular carcinoma (HCC)." *Open-Access Impact Journal on Aging*, vol. 12, no. 24, 31 Dec. 2020. doi:10.18632/aging.104142
- Bai, Yaping, *et al.* "ApoM is an important potential protective factor in the pathogenesis of primary liver

- cancer." *Journal of Cancer*, vol. 12, no. 15, 4 Jun. 2021, pp. 4661-4671. doi:10.7150/jca.53115
23. Lu, Yifei, *et al.* "Integrative transcriptomics and metabolomics explore the mechanism of kaempferol on improving nonalcoholic steatohepatitis." *Food & Function*, vol. 11, no. 11, 14 Oct. 2020. doi:10.1039/D0FO02123G
24. Emole, Nancy Magee. "Uncovering Mechanisms Behind Simple Steatosis Progression to Non-Alcoholic Steatohepatitis." *University of Kansas*, 2020. PhD dissertation
25. Luo, Yuwen. "The Role of Sugar-sweetened Water in the Progression of Nonalcoholic Fatty Liver Disease." *Auburn University*, 8 Dec. 2016. PhD dissertation
26. Ge, Steven Xijin, Eun Wo Son, and Runan Yao. "iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data." *BMC Bioinformatics*, vol. 19, no. 1, 19 Dec. 2018, pp. 1-24. doi:10.1186/s12859-018-2486-6

Copyright: © 2022 Bissoondial and Narayan. All JEI articles are distributed under the attribution non-commercial, no derivative license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>). This means that anyone is free to share, copy and distribute an unaltered article for non-commercial purposes provided the original author and source is credited.