**Possible association between helicobacter pylori infection and non-alcoholic fatty liver disease**

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Applied Physiology, Nutrition, and Metabolism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>apnm-2016-0499.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>23-Oct-2016</td>
</tr>
</tbody>
</table>
| Complete List of Authors: | Chen, Chang-Xi; The First Affiliated Hospital of Soochow University, Gastroenterology
Mao, Yu-Shan; The Affiliated Hospital of Medical College of Ningbo University, Endocrinology
Foster, Parker; National Institute of Health, NIH, Radiology and Imaging Sciences
Zhu, Zhong-Wei; Zhenhai Lianhua Hospital, Gastroenterology
Du, Juan; Zhenhai Lianhua Hospital, Gastroenterology
Guo, Chuan-Yong; Shanghai Tenth People's Hospital, Gastroenterology |
| Keyword: | Non-alcoholic fatty liver disease, Helicobacter pylori infection, ResNet Database, Pathway enrichment analysis, Network connectivity analysis |
Possible association between helicobacter pylori infection and non-alcoholic fatty liver disease

Chang-Xi Chen¹,²,#, Yu-Shan Mao³,## Parker Foster⁴, Zhong-Wei Zhu², Juan Du², Chuan-Yong Guo⁵*

¹Department of Gastroenterology, the First Affiliated Hospital of Soochow University, Suzhou, 215006, China
²Department of Gastroenterology, Zhenhai Lianhua Hospital, Ningbo, 315207, China
³Department of Endocrinology, the Affiliated Hospital of Medical College of Ningbo University, Ningbo, 315020, China
⁴Department of Radiology and Imaging Sciences, NIH, Bethesda, 20852, USA
⁵Department of Gastroenterology, Shanghai Tenth People’s Hospital, Tongji University School of Medicine, Shanghai, 200072, China

#These authors contributed equally to this study
*Correspondence to: Prof. Chuan-yong Guo, MD, PhD, Department of Gastroenterology, Shanghai Tenth People’s Hospital, Tongji University School of Medicine, Shanghai 200072, China.
Tel:+86-15232345674; fax:+86-21-66301173; E-mail: c.guo@gousinfo.com
Abstract

Possible association between Helicobacter pylori infection (HPI) and nonalcoholic fatty liver disease (NAFLD) has been proposed by several studies with inconsistent conclusions. Here, we studied the association between HPI and NAFLD at three levels: 1.) Genetic level; 2.) Small molecular level; and 3.) Clinical level. Relation data between diseases, genes, and small molecules were acquired from Pathway Studio ResNet Mammalian database. Clinical data were acquired from 2,263 elderly South Chinese subjects, including 603 NAFLD patients and 1660 non-NAFLD subjects. Results showed that HPI and NAFLD present significantly shared genetic bases (95 genes, p-value=2.5E-72), demonstrating multiple common genetic pathways (enrichment p-value<=$4.38E-20 for the top 10 pathways). Genetic network analysis suggested that mutual regulation may exist between HPI and NAFLD through 21 out of 95 genes. Furthermore, 85 out of the 95 genes manifested strong interaction with 12 small molecules/drugs that demonstrate effectiveness in treating both diseases. Clinical results showed that HPI rate in NAFLD group was significantly higher than that in the non-NAFLD group (51.9% vs. 43.6%; p-value= 4.9E-4). Multivariate logistic regression results supported the observations and suggested that HPI served as a risk factor for NAFLD in the experiment data studied (odds ratio: 1.387, p-value = 0.018). Results from this study support the hypothesis that complex biological association may exist between HPI and NAFLD, which partially explain the significant clinical co-occurrence in the elderly population of south China.
**Key words**: Non-alcoholic fatty liver disease, Helicobacter pylori infection, ResNet Database, Pathway enrichment analysis, Network connectivity analysis

**Introduction**

Non-alcoholic fatty liver disease (NAFLD) is a clinical syndrome related to insulin resistance, lipid changes, central fat mass and obesity. Recent studies considered NAFLD as a hepatic border of metabolic syndrome (Abenavoli et al. 2016). It can progress to liver cirrhosis and end-stage liver disease, and is one of the important causes of hidden-source cirrhosis (Chalasani et al. 2012). The prevalence rate of NAFLD in the general population is as high as 22.1–28.7%, leading to a serious threat to public health (Younossi et al. 2016). In recent years, the incidence of NAFLD has shown a clear increasing trend, becoming the primary cause of abnormal liver function and the first chronic liver disease in developed countries. The prevalence rate of NAFLD among Chinese adults is 15% (Fan 2013). A South Korean study of people aged 70–82 years found that the prevalence rate of NAFLD was 35.1%, higher than that in general population (20–30%) (Koehler et al. 2012). Several studies showed that NAFLD is related to liver lipid metabolic disturbance, insulin resistance, and diabetes (Fenandes et al. 2010).

Helicobacter pylori is a gram-negative bacillus that is usually found in human gastric
mucosal epithelium. The average rate of helicobacter pylori infection (HPI) among the global population is more than 50%, with higher infection rate in developing countries (Mandeville et al. 2009). HPI is thought to be one of the main pathogenic factors of gastric cancer, stomach inflammation, gastric lymphoma, and peptic ulcer (Fuccio et al. 2009). In recent years, however, a growing number of studies have found that HPI is also closely related to liver cancer and extra-gastrointestinal diseases involving lipid metabolism disorder (Aoshiba et al. 2013). Additionally, one study found that the intestinal microecological changes may be also involved in the pathogenesis of NAFLD (Carter and Karpen 2007). Moreover, several clinical studies have been conducted to explore the possible correlation between HPI and NAFLD (Takuma 2011; Polyzos et al. 2013; Okushin et al. 2015). However, the results were not consistent.

In this study, we integrated large scale ResNet relation data and clinical data for a comprehensive multiple-level analysis to test the hypothesis that NAFLD and HPI are biologically associated.

**Methods and Materials**

The study is laid out as follows: 1.) Study the association between HPI and NAFLD at genetic level 2.) Test the relation between HPI and NAFLD at small molecular level; 3.) Evaluate the clinical correlation between HPI and NAFLD.
HPI-Gene and NAFL-Gene data acquisition

HPI-Gene and NAFL-Gene relation data were acquired from Pathway Studio (PS) ResNet Mammalian database updated July 2016. All the supporting data were presented in Supplementary Data. The shared genes by both diseases were studied and analyzed with Pathway Enrichment Analysis (PEA) and Gene-Gene Interaction (GGI) analysis using Pathway Studio (www.pathwaystudio.com). In recent years, Pathway Studio has been widely used to study modeled relationships between proteins, genes, complexes, cells, tissues and diseases (Nikitin et al. 2003; Abdelmoez et al. 2016; Li et al. 2016; http://pathwaystudio.gousinfo.com/Mendeley.html).

The ResNet® Mammalian database is one part of PS ResNet Databases, which are a group of real-time updated network databases, includes curated signaling, cellular process and metabolic pathways, ontologies and annotations, as well as molecular interactions and functional relationships, which were extracted from the 35M+ references covering entire PubMed abstract and Elsevier full text journals using an automated natural language processing-based information extraction system, MedScan, with precision over 91% (Daraselia et al. 2004). Each ResNet relation within the database is supported with one or more references. By far, the PS ResNet Databases is the largest database among known competitors in the field (Lorenzi et al. 2014). More information about the PS ResNet databases please refer to: http://pathwaystudio.gousinfo.com/ResNetDatabase.html.

Test HPI-NAFLD association at small molecule level
Besides genetic association analysis, we also explored the clinical and pre-clinical drugs/small molecules that are effective in treating both HPI and NAFLD. In addition, we studied the drug-gene interaction between the identified drugs and genes related to both HPI and NAFLD. The drug-disease and drug-gene relations were also acquired from ResNet® Mammalian database as described above, with all supporting data provided in Supplementary Data.

**Evaluate clinical correlation between HPI and NAFLD**

**Participants**

Originally, there were 2,740 (aged ≥ 60 years) subjects recruited from south China area (Ningbo, Zhejiang province). All of these participants were through physical examination in Medical Examination Center of Lianhua Hospital, Ningbo, during Mar. to Nov. 2013. After data quality control, 447 subject were excluded from the study, resulting a total of 2,263 subjects left for the study (male/female: 1424/839; sex ratio 1.7: 1; age ≥ 60; average age: 69 ± 7 years). The exclusion criteria were described as follows: those with specific liver disease that may result in NAFLD or abnormal liver function, such as viral hepatitis, hepatolenticular degeneration, autoimmune liver disease, or drug-induced liver diseases; those who had a long history of drinking with alcohol consumption in men >140 g/wk and in women >70 g/wk; gastric cancer; peptic ulcer; recent (≤ 4weeks) use of medications including proton pump inhibitors, bismuth containing compounds, antibiotics, probiotics, non-steroidal anti-inflammatory drugs, or glucocorticoids; H. pylori eradication therapy; previous gastrointestinal surgery; and those with incomplete
data for various reasons. Before participation, the whole procedure was explained to all participants, with verbal informed consent obtained from each subject. The study was reviewed and approved by the Ethics Committee of the Zhenhai Lianhua Hospital, Ningbo, China (approval no. 20120213).

**Diagnosis of NAFLD**

The diagnosis of NAFLD was based on the criteria suggested by the Chinese Liver Disease Association, and the clinical diagnostic standards (Fan et al. 2011) for the determination of NAFLD as described as follows: (1) diffuse echo enhancement near the liver, and the echo was stronger than that of the kidney; (2) the intrahepatic duct structure was not clear; and (3) the liver far-field echo gradually reduced.

**Detection of HPI**

*H. pylori* were detected using the C\textsuperscript{13} urea breath test (HCBT - 01 tester, Shenzhen Zhonghe Headway Biological Technology Co. Ltd, China). All the participants were examined after fasting serum samples were withdrawn in the morning. The specific steps were as follows: (1) collect baseline breath samples; (2) take 75-mg C\textsuperscript{13} urea capsules with 50 mL drinking water; (3) sit still for 30 min to collect measured breath samples; (4) measure the breath samples before and after oral administration of C\textsuperscript{13} urea capsules by C\textsuperscript{13} breath mass spectrometer; (5) test results expressed as Delta Over Baseline (DOB) value; and (6) when DOB was ≥ 4.0, it could be judged that HPI was positive.
Acquisition of other clinical features

Basic clinical features of each participant were collected, including: age, gender, history of liver disease, alcohol consumption, smoking history, current use of hepatoprotective drugs, and a personal medical history including hypertension, hyperlipidemia, diabetes, anemia, gastrointestinal disease, liver disease, and recent history of medication. Examination was carried out in the morning, in a fasting state. Standard methods were used to measure height, weight, waistline, and blood pressure. Ten milliliters of fasting venous blood from each subject was used to detect blood biochemical indexes, including serum biochemical measures and hemoglobin A1c, which were measured by an Au 640 automatic biochemical detector (Olympus, Tokyo, Japan). HbA1c was measured by an HLC-723 G7 automatic analyzer (Tosoh, Tokyo, Japan). The reagent and quality control products were all provided by Tosoh.

Statistical analysis

SPSS version 18.0 statistical software was used for data analysis. Data with a normal or near normal distribution were expressed as mean ± SD. t test was used for comparison between groups. Distribution of data was expressed by M (P25 and P75), and Mann-Whitney U test was adopted for comparison between groups. Chi-square test was used for classification of variables; Multivariate regression analysis was used for risk factor analysis; and stepwise logistic regression analysis (backward: wald; entry: 0.05; removal: 0.10) was used to evaluate the risk factors for NAFLD, with the following factors included prior to running the analysis: age, gender, fasting plasma glucose, uric
acid, AST, ALT, alkaline phosphatase, γ-glutamyltransferase, TC, triglyceridemia, HDL-C, LDL-C, body height, body weight, body mass index, waist circumference, systolic blood pressure, diastolic blood pressure, serum HbA1c, and HPI. The results were expressed by odds ratio and 95% confidence interval. $P < 0.05$ indicated statistical significance.

Results

Shared genetic bases between HPI and NAFLD

We conducted a systematic analysis on the HPI-Gene and NAFLD-Gene ResNet relation data acquired to identify genes associated with HPI and NAFLD. Results showed that 267 genes were associated with HPI, supported by 913 scientific references from 1992 to June 2016. We present the full list of the 267 genes and the 913 supporting references in Supplementary Table S1a and Table S1b, respectively, including titles of the references and the related sentences where these relations were identified. For NAFLD, we identified 687 genes supported by 2,832 references from 1999 to June 2016. The gene list and the supporting references are presented in Supplementary Table S2a and Table S2b, respectively. To note, we observed a significant overlap of 95 genes between the HPI-genes and NAFLD-genes (Fisher Exact test p-value=2.5E-72), as shown in Fig. 1. We provide the full list of these 95 genes and the supporting references in Supplementary Table S3a and Table S3b, respectively. The network (Fig. 1) and the supporting references (Supplementary Table S3b) can be used to identify the detailed
descriptions regarding how these genes may contribute to the overall disease state in NAFLD. For an instance, activators of NRF2 has been found to attenuate progression of nonalcoholic steatohepatitis-related fibrosis and loss of NRF2 could lead to rapid onset and exacerbation of steatohepatitis (Shimozono et al. 2013). Both TLR2 and palmitic acid activate the inflammasome in Kupffer cells and macrophages in the development of NAFLD (Miura et al. 2013).

**Put Fig 1 above here**

To test the functional profile of the 95 genes associated with both HPI and NAFLD, we conducted a Pathway Enrichment Analysis (PEA) using Pathway Studio with all these 95 genes as input. Due to lack of space, we only present the top 10 significantly enriched pathways (p-value ≤ 4.38E-20) in Table 1. The full list of 108 pathways/gene sets enriched with p-value < 1e-10 (with 90/95 unique genes) are listed in **Supplementary Table S4**, including the symbols of the enriched genes in each pathway.

**Put Table 1 above here**

From PEA results we observed that 5 pathways/GO terms (49 unique genes) were related to drug effects, 4 (45 unique genes) to cell growth and proliferation, 3 (50 unique genes) to cell apoptosis, 3 (43 unique genes) to immune system, 3 (37 unique genes) to transcription factors, 2 (20 unique genes) to protein phosphorylation, 1 (34 unique genes)
to neuro system, and 1 (13 unique genes) to protein kinase. To note, many of these pathways have been implicated in both HPI and NAFLD, such as the response to lipopolysaccharide (GO ID: 0032496) (Matsushita et al. 2016; Lepper et al. 2005), inflammatory response (GO ID: 0006954) (Tilg and Moschen 2010), ageing (GO ID: 0016280) (Bertolotti et al. 2014; Safe et al. 1993), response to hypoxia (GO ID: 0001666) (Suzuki et al. 2014; Kang et al. 2014), and cytokine activity (GO ID: 0005125) (Tanahashi et al. 2000; Tilg 2010). For more detailed information of these significantly enriched pathways, please refer to Supplementary Table S4.

Our observations suggested that HPI and NAFLD share multiple genetic pathways, through which a large group of genes play roles affecting the pathogenic development of both diseases.

Possible co-regulations between HPI and NAFLD

Further network analysis using Pathway Studio showed that, 21 out of the 95 genes and the two diseases, HPI and NAFLD, may be mutually influencing each other, as shown in Fig. 2, demonstrating 128 relations supported by 1,080 references (see Supplementary Table S5). The ‘network building’ module in Pathways Studio can be used to identify both upstream (regulating: ‘→’) and downstream (being regulated: ‘←’) relationships between one entity and another. In the network presented in Fig. 2, these 21 genes demonstrated both upstream regulation of HPI and NAFLD and downstream connection with the two diseases (influenced by the disease status). Therefore, the pathogenesis of
HPI and NAFLD could lead to changes of these 21 genes, which in turn affect the development of the other disease. Detailed description regarding how HPI could affect NAFLD through influencing these gene were provided in the ‘Discussion’ Section. For more detailed description of each of the 128 disease-gene-disease relations presented in Fig. 2, please refer to the references presented in Supplementary Table S5. Our observation suggest that HPI and NAFLD may possess functional associated with each other through a genetic network of multiple genes (see Fig. 2).

Put Fig. 2 above here.

Small molecular level association

Besides genetic association analysis, we also studied the small molecular/drugs that are effective in treating both HPI and NAFLD, using the Drug-Disease relation data acquired from the PS ResNet Mammalian database. Results showed that 12 drugs have been shown effective in treating both HPI and NAFLD (Fig. 3 (a)), with 6 of them studied in clinical trials, including Polypeptide antibiotic agent, rifampicin, N-acetylcysteine, Vitamin D, Vitamin E and Fe2+ (highlighted in pink in Fig. 3(a); source: ClinicalTrials.gov). We present the 136 supporting references, including ClinicalTrials.gov identifiers, for the drug-disease relations in Fig. 3 (a) in Supplementary Table S6.

Put Fig. 3 above here.

Additionally, we studied the relation of these 12 drugs and the 95 genes associating with
both HPI and NAFLD, as shown in Fig. 3 (b). Results showed that 85 out of 95 genes presented strong relation with all the 12 drugs, supported by over 9,000 references (see Supplementary Table S7). These drug-gene association helps in understanding the mechanism of the linkage between these 85 genes and the disease (HPI and NAFLD). For detailed description of the disease-drug-gene relations, please refer to the references underlying each relation, which were presented in (Supplementary Table S6 and Table S7).

Clinical results

Co-incidence of HPI and NAFLD

The HPI and NAFLD disease status are presented in Table 2. Among the 2,263 participants, there were 603 NAFLD cases (26.6%; 392 males and 211 females) and 1660 non-NAFLD controls. For HPI, we observed 1036 HPI positives (case) and 1227 HPI negatives (controls). Additionally, within the NAFLD case group, the HPI rate was 51.9% (313/603), while the number was 43.6% (723/1660) in the NAFLD control group (see Table 2). A significant HPI difference between the NAFLD case group and the NAFLD control group was observed (Fisher-Exact test p-value = 4.9E-4), suggesting a possible clinical association between these two diseases.

Put Table 2 above here

To eliminate any possible co-regulation effects of other risk factors when considering the
influence of HPI on NAFLD, we performed a multivariate logistic regression (MLR) analysis to test the relation between NAFLD and all clinical parameters. Only these showing a p-value <0.05 from the MLR were taken as risk factors (see Table 3). Results showed that besides HPI, multiple other parameters may also serve as potential risk factors for NAFLD, including gender, age, uric acid, AST, ALT, γ-Glutamyltransferase, triglyceride, body mass index, waist circumference, and serum HbA1C. To note, for HPI, the odds ratio = 1.387 with P-value = 0.018, suggesting that NAFLD cases were linked to HPI for the participants in this study.

Discussion

Recently studies posed possible association between HPI and liver pathological damage (Polyzos et al. 2013; Takuma 2011). HPI has been reported to promote insulin resistance by inducing chronic inflammation and affecting insulin-regulating gastrointestinal hormones (Aydemir et al. 2005), whilst insulin resistance is a key element in NAFLD occurrence (Yoo et al. 2013; Gunji et al. 2009; Park et al. 2010). Moreover, HPI could induce local infiltration of lymphocytes and granulocytes by releasing inflammatory factors such as IL-6, IL-1, and TNF-α, which induces inflammation and aggravates liver damage (Polyzos et al. 2013; Park et al. 2011). In this study, associations between HPI and NAFLD were studied at three levels using comprehensive data integration analysis. Our results support the early observations by Takuma and by Polyzos et al. and suggest
that HPI and NAFLD may possess biological association.

ResNet relation data between genes, diseases (HPI/NAFLD) and small molecules/drugs were acquired from Pathway Studio ResNet relation database. In recent years, Pathway Studio has been widely used to study modeled relationships between proteins, genes, complexes, cells, tissues and diseases (Nikitin et al. 2003; Abdelmoez et al. 2016; Li et al. 2016; http://pathwaystudio.gousinfo.com/Mendeley.html). Besides the full lists of genes and drugs, we also presented the supporting references for each disease-gene-drug relation in Supplementary Tables, which could be used to located detailed description of how a candidate gene/drug is related to HPI and/or NAFLD.

At a genetic level, we identified significantly shared gene-bases between HPI and NAFLD (95 overlapped genes, Fisher-exact test p-value=2.5E-72). Pathway analysis showed that these 95 genes significantly enriched 108 pathways (p-value<1e-10), many of which have been implicated to be linked to both HPI and NAFLD, such as the response to lipopolysaccharide (GO ID: 0032496), inflammatory response (GO ID: 0006954), ageing (GO ID: 0016280), response to hypoxia (GO ID: 0001666), and cytokine activity (GO ID: 0005125) (Matsushita et al. 2016; Lepper et al. 2005; Tilg and Moschen 2010; Bertolotti et al. 2014; Safe et al. 1993; Suzuki et al. 2014; Kang et al. 2014; Tanahashi et al. 2000; Tilg 2010). These results suggest that HPI and NAFLD share
multiple genetic pathways, through which a large group of genes working together influence the pathogenic development of both diseases.

More importantly, we observed a 21-gene network, through which HPI and NAFLD could affect the disease status of one another (Fig. 2). Genetic linkage between HPI and NAFLD can be identified from the network and the supporting references (Supplementary Table S5). For example, studies showed that HPI activates expression of MyD88 (Ryu et al. 2013) which has been shown as a crucial signaling molecule that promotes nonalcoholic steatohepatitis and fibrosis (Miura et al. 2010). On the other hand, decreased expression level of TLR4 has been shown to attenuate nonalcoholic steatohepatitis and fibrosis (Chen et al., 2014), which were activated in case of HPI (Ryu et al. 2013). These observation provides further support for the hypothesis that these two diseases may be biologically associated.

At small molecule level, we observed that there were 12 drugs/small molecules effective in treating both HPI and NAFLD (Fig. 3 (a)). Noteworthy, 6 out of the 12 drugs have been through clinical trials (Supplementary data Table S8). Additionally, we noted that 85 out of 95 genes shared by both diseases presented strong association with all the 12 drugs, supported by over 9,000 references (Supplementary data Table S9). To note, there were multiple drugs used to treat NAFLD and/or HPI, and the 12 common drugs are
not necessarily the first line treatment choice in clinical. The presentation of these drugs was with the purposes of identifying possible common regulators at small molecular level, which could add new insights into the understanding of the possible shared biological mechanisms of the two diseases.

Consistent with the genetic and molecule level studies, our clinical study showed that HPI and NAFLD demonstrated significant co-incidence (p-value = 4.9E-4). Specifically, the prevalence of NAFLD among the 2, 263 subjects was 26.6% (603), with 53% (313) also diagnosed with HPI. Whilst in the 1,660 non-NAFLD subjects, only 43.6% were diagnosed with HPI. Additionally, MLR results supported these observations and suggested that HPI served as one of the risk factors for NAFLD in the experiment data studied (odds ratio: 1.387, p-value = 0.018). To note, there are multiple factors linked to the risk of NAFLD, and HPI may only partially contribute to the development and progression of the disease (see Table 3), which helps to explain the fact that countries with lower prevalence of HPI, such as the United States, still have high incidence of NAFLD (Ashtari et al. 2015; Thirumurthi et al. 2012). Through MLR, influences of other NAFLD risk factors could be removed, which helps to gain clearer observation between HPI and NAFLD.

Nevertheless, this study has several limitations that need future work. First, the relation data between diseases, drugs, and genes were acquired from PS ResNet database.
Although supported by thousands of references, it is possible that some shared genetic evidence remain un-covered. Second, the 2,263 subjects were from elder south China population. The purpose of selecting subjects of senior age was with the consideration that both HPI and NAFLD are age sensitive (Safe et al. 1993; Bertolotti et al. 2014). Thus using elder subject group alone can reduce the influence of age on the results, which helps to gain a clearer observation of possible relations between HPI and NAFLD. However, cross-age subjects from other geographic regions should be tested to gain a more comprehensive conclusion.

In summary, results from this study support the hypothesis that HPI and NAFLD may possess biological association, leading to co-incidence in clinical trials. Our results may partially explain the significant co-incidence of HPI and NAFLD in the elderly population of south China. To our knowledge, this is the first study integrating large scale ResNet relation data and large sample sized clinical data to explore possible association between HPI and NAFLD, which may provide new insights into the current field of HPI-NAFLD correlation study.

**Conflict of interest statement**

The authors declare no conflicts of interest associated with this manuscript.
Acknowledgements

This study was partly funded by the Social Development Program of Ningbo, China (No. 2013C50044, No.2011C50021); The Social Development and Science & Technology Program of Zhenhai, Ningbo, China (No. 2013S1002); and Medical Science & Technology Foundation of Zhenhai, Ningbo, China (No. 2014020)
References


https://mc06.manuscriptcentral.com/apnm-pubs


Table 1. Molecular function pathways/ groups enriched by 95 genes reported

<table>
<thead>
<tr>
<th>Pathway/gene set name</th>
<th>Hit type</th>
<th>GO ID</th>
<th># of Entities</th>
<th>Overlap</th>
<th>p-value</th>
<th>Jaccard similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>response to lipopolysaccharide</td>
<td>biological_process</td>
<td>003249</td>
<td>6</td>
<td>252</td>
<td>32</td>
<td>1.4E-38</td>
</tr>
<tr>
<td>inflammatory response</td>
<td>biological_process</td>
<td>000695</td>
<td>4</td>
<td>404</td>
<td>35</td>
<td>2.25E-36</td>
</tr>
<tr>
<td>aging</td>
<td>biological_process</td>
<td>001628</td>
<td>0</td>
<td>254</td>
<td>27</td>
<td>2.93E-30</td>
</tr>
<tr>
<td>response to drug</td>
<td>biological_process</td>
<td>001703</td>
<td>5</td>
<td>509</td>
<td>33</td>
<td>4.89E-30</td>
</tr>
<tr>
<td>response to hypoxia</td>
<td>biological_process</td>
<td>000166</td>
<td>6</td>
<td>259</td>
<td>25</td>
<td>7.63E-27</td>
</tr>
<tr>
<td>cytokine activity</td>
<td>molecular_function</td>
<td>000512</td>
<td>5</td>
<td>237</td>
<td>24</td>
<td>2.09E-25</td>
</tr>
<tr>
<td>lipopolysaccharide-mediated signaling pathway</td>
<td>biological_process</td>
<td>003166</td>
<td>3</td>
<td>33</td>
<td>14</td>
<td>2.63E-25</td>
</tr>
<tr>
<td>MyD88-dependent toll-like receptor signaling pathway</td>
<td>biological_process</td>
<td>000275</td>
<td>5</td>
<td>83</td>
<td>15</td>
<td>1.23E-20</td>
</tr>
<tr>
<td>positive regulation of nitric oxide biosynthetic process</td>
<td>biological_process</td>
<td>004542</td>
<td>9</td>
<td>48</td>
<td>13</td>
<td>1.55E-20</td>
</tr>
<tr>
<td>response to activity</td>
<td>biological_process</td>
<td>001482</td>
<td>3</td>
<td>69</td>
<td>14</td>
<td>4.38E-20</td>
</tr>
</tbody>
</table>

Note: For each pathway/Go term, the p-value was calculated using Fisher-Exact test against the hypothesis that a randomly selected gene group of same size (95) can generate a same or higher overlap with the corresponding pathway/Go term.
Table 2: HPI and NAFLD disease status for the 2663 participants

<table>
<thead>
<tr>
<th></th>
<th>HPI</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>NAFLD</td>
<td>313</td>
<td>290</td>
<td>603</td>
</tr>
<tr>
<td>Controls</td>
<td>723</td>
<td>937</td>
<td>1660</td>
</tr>
<tr>
<td>Total</td>
<td>1036</td>
<td>1227</td>
<td>2263</td>
</tr>
</tbody>
</table>
Table 3 Multivariate logistic regression analysis results for NAFLD

<table>
<thead>
<tr>
<th>Variables</th>
<th>β</th>
<th>SE</th>
<th>Wald X²</th>
<th>p-value</th>
<th>OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male gender</td>
<td>−0.563</td>
<td>0.147</td>
<td>14.647</td>
<td>0.000</td>
<td>0.570</td>
<td>0.427-0.760</td>
</tr>
<tr>
<td>Age</td>
<td>−0.109</td>
<td>0.05</td>
<td>4.708</td>
<td>0.03</td>
<td>0.897</td>
<td>0.812-0.990</td>
</tr>
<tr>
<td>Uric acid</td>
<td>0.003</td>
<td>0.001</td>
<td>15.183</td>
<td>0.000</td>
<td>1.003</td>
<td>1.002-1.005</td>
</tr>
<tr>
<td>AST</td>
<td>0.022</td>
<td>0.009</td>
<td>5.803</td>
<td>0.016</td>
<td>0.978</td>
<td>0.961-0.996</td>
</tr>
<tr>
<td>ALT</td>
<td>0.045</td>
<td>0.007</td>
<td>36.911</td>
<td>0.000</td>
<td>1.046</td>
<td>1.031-1.062</td>
</tr>
<tr>
<td>γ-Glutamyltransferase</td>
<td>0.006</td>
<td>0.002</td>
<td>7.352</td>
<td>0.007</td>
<td>1.006</td>
<td>1.002-1.011</td>
</tr>
<tr>
<td>Triglyceride</td>
<td>0.388</td>
<td>0.060</td>
<td>42.236</td>
<td>0.000</td>
<td>1.474</td>
<td>1.311-1.656</td>
</tr>
<tr>
<td>Body mass index</td>
<td>0.262</td>
<td>0.028</td>
<td>90.476</td>
<td>0.000</td>
<td>1.300</td>
<td>1.231-1.373</td>
</tr>
<tr>
<td>Waist circumference</td>
<td>0.035</td>
<td>0.009</td>
<td>15.268</td>
<td>0.000</td>
<td>1.035</td>
<td>1.017-1.053</td>
</tr>
<tr>
<td>Serum HbA1c</td>
<td>0.205</td>
<td>0.071</td>
<td>8.371</td>
<td>0.004</td>
<td>1.228</td>
<td>1.068-1.411</td>
</tr>
<tr>
<td>HPI</td>
<td>0.267</td>
<td>0.113</td>
<td>5.552</td>
<td>0.018</td>
<td>1.387</td>
<td>1.046-1.729</td>
</tr>
<tr>
<td>Constant</td>
<td>12.402</td>
<td>0.965</td>
<td>165.052</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Note: β is partial regression coefficient; SE is standard error of partial regression coefficient; OR is odds ratio; CI is confidence interval. P-value specifies the possibility that a parameter is not a risk factor for NAFLD.
Figure Legends

Fig. 1 The 95 genes associated with both HPI and NAFLD. (a) Venn diagram of the HPI-genes and NAFLD-genes; (b) Network between 95 genes and HPI and NAFLD.

Fig. 2 The 21 genes mutually affecting both HPI and NAFLD.

Fig. 3 The 12 drugs/small molecules effective in treatment of with both HPI and NAFLD. (a) Network between the 12 drugs and HPI and NAFLD; All drugs have been through pre-clinical study; the 6 pink-color-highlighted drugs were also through clinical trials; (b) Network between 85/95 genes and the 12 drugs.
Fig. 1 The 95 genes associated with both HPI and NAFLD. (a) Venn diagram of the HPI-genes and NAFLD-genes; (b) Network between 95 genes and HPI and NAFLD.
Fig. 2 The 21 genes mutually affecting both HPI and NAFLD.
Fig. 3 The 12 drugs/small molecules effective in treatment of with both HPI and NAFLD. (a) Network between the 12 drugs and HPI and NAFLD; All drugs have been through pre-clinical study; the 6 pink-color-highlighted drugs were also through clinical trials; (b) Network between 85/95 genes and the 12 drugs.