Intrahepatic bacterial metataxonomic signature in non-alcoholic fatty liver disease

Silvia Sookoian, Adrian Salatino, Gustavo Osvaldo Castaño, Maria Silvia Landi, Cinthia Fijalkovsky, Martin Garaycoechea, Carlos Jose Pirola

ABSTRACT

Objective We aimed to characterise the liver tissue bacterial metataxonomic signature in two independent cohorts of patients with biopsy-proven non-alcoholic fatty liver disease (NAFLD) diagnosis, as differences in the host phenotypic features—from moderate to severe obesity—may be associated with significant changes in the microbial DNA profile.

Design and methods Liver tissue samples from 116 individuals, comprising of 47 NAFLD overweight or moderately obese patients, 50 NAFLD morbidly obese patients elected for bariatric surgery and 19 controls, were analysed using high-throughput 165 rRNA gene sequencing.

Results Liver bacterial DNA profile significantly differs between morbidly obese and non-morbidly obese patients with NAFLD. Bacteroidetes (p=1.8e-18) and Firmicutes (p=0.0044) were over-represented in morbidly obese patients and Proteobacteria (p=5.2e-10)—specifically Gammaproteobacteria and Alphaproteobacteria—and Deinococcus-Thermus (p=0.00012)—were over-represented in the non-morbidly obese cohort. Cohort-specific analysis of liver microbial DNA signatures shows patterns linked to obesity. The imbalance in Proteobacteria (Alpha or Gamma) among non-morbidly obese patients, and Peptostreptococcaceae, Verrucomicrobia, Actinobacteria and Gamma Proteobacteria DNA among morbidly obese patients was associated with histological severity. Decreased amounts of bacterial DNA from the Lachnospiraceae family were associated with more severe histological features. Proteobacteria DNA was consistently associated with lobular and portal inflammation scores. Microbial DNA composition corresponded to predicted functional differences.

Conclusion This is the first comprehensive study showing that the liver tissue of NAFLD patients contains a diverse repertoire of bacterial DNA (up to 2.5×10^4 read counts). The liver metataxonomic signature may explain differences in the NAFLD pathogenic mechanisms as well as physiological functions of the host.

INTRODUCTION

Non-alcoholic fatty liver disease (NAFLD) represents a major burden on health systems around the world, as it has become the most prevalent chronic liver disease. The NAFLD natural history is modulated by genetic and environmental factors, including the gut microbiota that can explain a fraction of the disease heritability. Ample body of evidence indicates that the magnitude and severity of obesity and type 2 diabetes play a significant role in the disease progression. In fact, NAFLD exhibits high degree of comorbidity with disorders of the metabolic syndrome and shares with them pathogenic mechanisms of disease, including a state of chronic systemic inflammation.

Much progress has been made in understanding NAFLD biology and pathogenesis. However, answers to some fundamental questions regarding mechanistic basis of disease remain elusive. For example, while epidemiological studies have
How might it impact on clinical practice in the foreseeable future?

► NASH subtypes might be distinguished by the microbial composition of the liver 16S rRNA gene profile.
► The correlation of disease phenotypes and host phenotypic features with tissue bacterial metataxonomic profiling will enable better understanding of NASH pathogenesis, as well as provide a platform on which to evaluate interventions.
► Liver bacterial metataxonomic profiling suggests that therapeutic options, including probiotic selection, should be precisely defined according to specific clinical scenarios, including the host phenome features.

robustly demonstrated that obesity is a major risk factor for NAFLD development, how and to what extent NAFLD interacts with the host phenotype is only partially understood. Results yielded by a meta-analysis on hepatic histology in obese patients undergoing bariatric surgery (BS) suggest that pathogenic mechanisms of NAFLD in severely obese individuals are linked to obesity status. Likewise, findings yielded by large studies exploring NAFLD histological features in morbidly obese patients demonstrate that severe obesity exerts strong influence on the liver disease phenotype.

Advances in molecular technologies have revolutionised our understanding of mechanisms of complex diseases. Likewise, high-throughput deep sequencing has allowed researchers to elucidate the impact of the gut microbiome on human health, including NAFLD and non-alcoholic steatohepatitis (NASH)—the severe histological form of the disease. Pioneering studies have consistently shown that obesity alters the gut microbiota. It has also been shown that severe—morbid—obesity per se imposes major changes in the gut microbiota, along with considerable systemic metabolic impairment.

Earlier explorations of the biogeography of bacterial communities on the human body have demonstrated that microbiota varies systematically across body habitats. Consequently, studying the tissue microbiota of affected organs may open new research avenues, including the understanding of pathogenic mechanisms of human diseases. However, the impact of the liver microbiome on NAFLD biology and its potential interactions with the host phenotype are largely unknown, which greatly limits our mechanistic understanding of microbial associations with the disease. In conducting the present study, we hypothesised that the liver microbial DNA signature in NAFLD may explain a portion of the variance in the disease manifestation, including differences in disease mechanisms and histological phenotypes. Hence, we aimed to demonstrate that a diverse and dynamic repertoire of bacterial DNA inhabits the liver of patients with NAFLD. We reasoned that examining the basic biology of the hepatic metataxonomic profile in patients with NAFLD and diverse host phenotypic features—from moderate to severe obesity—may provide valuable insight into the disease mechanisms as well as physiological functions of the host. The goal of this investigation was thus to characterise the liver tissue 16S rRNA gene signature in two unique and independent cohorts—overweight or moderately obese, and morbidly obese patients—with biopsy-proven NAFLD diagnosis.

METHODS
Study design, patient selection and phenotypic characterisation

Using high-throughput 16S rRNA gene sequencing, we obtained pertinent liver metataxonomic information from the tissue samples of 116 individuals, namely 47 overweight or moderately obese patients with NAFLD, 50 NAFLD severely—morbidly—obese patients (body mass index (BMI) >40 kg/m²) elected for BS and 19 control subjects (non-NAFLD). Adopting a case–control design, we ensured that the control liver samples matched patients’ features in each cohort (morbidly obese and non-morbidly obese). Complete details of the study design, patient recruitment strategy, physical, anthropometric and biochemical evaluation, as well as histological assessment are shown in the online supplementary material.

All investigations performed as a part of the present study were conducted in accordance with the guidelines of the 1975 Declaration of Helsinki, as revised in 1993.

16S rRNA gene metagenomic sequencing, sequence data analysis and operational taxonomy unit (OTU) clustering

DNA was isolated from liver specimens by a manual protocol as explained in the online supplementary material. Extracted bulk DNA samples were amplified with barcoded primers; DNA libraries were constructed, and high-throughput sequencing was performed on an IlluminaMiSeq platform by Macrogen Inc (NGS Division) Seoul, South Korea. A negative (blank) control was used to test the potential presence of contaminant DNA and/or cross-contamination; the sample showed no product and did not pass the quality control analysis (online supplementary figure 1). We used qPCR based on a SYBR Green assay and universal primers to measure the 16S rRNA gene absolute abundance within each sample. Details of metagenomic sequencing, including Linear Discriminant Analysis (LDA) Effect Size (LEFSe) and functional prediction can be found in the online supplementary material.

Liver immunohistochemistry (IHC) for LPS staining

Analysis of LPS abundance and localisation was performed by IHC using an antibody that recognises the lipid A or endotoxin region of LPS from Escherichia coli. This antibody cross-reacts with Pseudomonas aeruginosa, Klebsiella pneumoniae, Salmonella enteritidis, Enterobacter aerogenes, Escherichia hermannii, Yersinia enterocolitica and Shigella sonnei. Further details can be found in the online supplementary material.

RESULTS

Clinical, biochemical, and histological features of NAFLD patients and controls are presented in table 1.

Liver microbial DNA diversity is related to the host phenome

The rarefaction curves obtained for the dataset (OTU based) flatten to the right after ~4000 reads sampled, indicating that the underlying microbial profiling’ richness is well covered by the sequence data, which was higher than ~7500 reads in all samples (online supplementary figure 2). As indicated by the plots of microbial DNA composition in figure 1A, significant differences exist in the microbial diversity between the two NAFLD patient cohorts. The liver 16S rRNA gene signature of morbidly obese patients with NAFLD exhibited a significantly (p=0.026, non-parametric analysis of variance one-way test) decreased microbial diversity at the order taxonomy level (Shannon’s diversity, H index: 2.34±0.34) compared with that...
of overweight or moderately obese patients with NAFLD (Shannon’s diversity, H index: 2.86±0.5) (figure 1B).

Principal coordinate analysis (PCoA) of Bray-Curtis beta diversities revealed significant differences between the two groups across all taxonomic levels. Permutational analysis of variance of Bray-Curtis beta diversities indicated that intersubject differences between the two cohorts were explained by microbial taxonomic variation (PERMDis2 p=1.19e–9) (figure 1C) at the order level. Phylogenetic beta diversity calculated using weighted UniFrac by PCoA also showed differences between groups (PERMDis2 p=2.86e–6) at the order level.

Liver bacterial DNA profile significantly differs between morbidly obese and non-morbidly obese patients with NAFLD

While bacterial phyla Bacteroidetes, Firmicutes and Proteobacteria accounted for up to 90% of sequences on average, distribution of phyla and other taxonomic levels between the groups was uneven (figure 1D). Significant differences in the intercohort analysis results across most of the taxonomic levels seem to be driven by the differential abundance of the dominant phyla. Specifically, Bacteroidetes (false discovery rate (FDR): 1.8e–18) and Firmicutes (FDR: 0.0044) were significantly over-represented in the NAFLD-morbidly obese cohort, whereas Proteobacteria (FDR: 5.2e–10) and Deinococcus-Thermus (FDR: 0.00012) were over-represented among overweight or moderately obese patients with NAFLD (figure 1E). At the order level, Bacteroidia and Clostridia were over-represented in the NAFLD morbidity obese cohort, while Gamma, Alphaproteobacteria, Bacilli and Deinococci were over-represented in the NAFLD non-morbidly obese cohort (figure 2A). The dendrogram displayed in online supplementary figure 3 shows the phylogenetic relationships among these taxa.

To identify the biologically informative features that help in differentiating the two NAFLD cohorts, we applied the LEfSe method. LEfSe findings confirmed that Bacteroidetes, Clostridia, Alpha and Gammaproteobacteria were the features most likely to explain the differences between the two groups (figure 2B). Statistical comparison of taxon abundances across subsamples at the order, family and genus taxonomic levels, respectively, are shown in online supplementary tables 1-3. online supplementary figure 4 shows 25 of the top 100 features, some of which may assist in classifying patients into the NAFLD non-morbidly obese or NAFLD morbidly obese cohort based on area under the receiver operating characteristic (AUROC) analysis.

To gain further insight into the contribution of specific bacterial DNA profiles that explain the differences between the two cohorts, we used sparse partial least squares linear discriminant analysis (sPLS-DA), which takes into account the sparsity and compositionality of microbiome data. The multi-variable method identified similar bacterial DNA profiles as responsible for major differences, including component one that characterised the NAFLD morbidly obese cohort with Bacteroidales and Clostridiales, and to a lesser extent Verrucomicrobia, all of which accounted for −8% of the variance, and component two that characterised the NAFLD non-morbidly obese cohort with Gram-negative bacteria, including species of the Phylum Proteobacteria and Deinococcus-Thermus, accounting for −16% of the variance (figure 2C).

Results of the association between the liver tissue bacterial metatranscriptome profiling and BMI as continuous variable are shown in the online supplementary material, online supplementary table 4 and online supplementary figure 5.

Quantitative detection of total bacterial DNA

Total 16S rRNA gene load did not differ between NAFL and NASH patients in any of the cohorts; however, the liver of patients with NASH presented a modest (although non-significant)

**Table 1** Clinical and biochemical features of patients with NAFLD and the control group

<table>
<thead>
<tr>
<th>Variables</th>
<th>Control subjects</th>
<th>NAFL</th>
<th>NASH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of subjects</td>
<td>9</td>
<td>21</td>
<td>26</td>
</tr>
<tr>
<td>Female/male (n)</td>
<td>5/4</td>
<td>10/11</td>
<td>16/10</td>
</tr>
<tr>
<td>Age, years</td>
<td>43±8.8</td>
<td>49±4.1</td>
<td>46.7±13</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>24±8</td>
<td>30±5</td>
<td>34±6†</td>
</tr>
<tr>
<td>Type 2 diabetes (n)</td>
<td>0</td>
<td>6</td>
<td>15†</td>
</tr>
<tr>
<td>Fasting plasma glucose, mg/dL</td>
<td>87±10</td>
<td>105±26</td>
<td>124±39†</td>
</tr>
<tr>
<td>Fasting plasma insulin, μU/mL</td>
<td>5±2.5</td>
<td>14±7</td>
<td>19.6±12.4†</td>
</tr>
<tr>
<td>HOMA-IR index</td>
<td>1.02±0.5</td>
<td>3.5±1.7</td>
<td>6.1±6†</td>
</tr>
<tr>
<td>Total cholesterol, mg/dL</td>
<td>188±38</td>
<td>203±39</td>
<td>196±42</td>
</tr>
<tr>
<td>HDL-cholesterol, mg/dL</td>
<td>60±15</td>
<td>58±16</td>
<td>51.2±15</td>
</tr>
<tr>
<td>LDL-cholesterol, mg/dL</td>
<td>110±42</td>
<td>124±35</td>
<td>120±36</td>
</tr>
<tr>
<td>Triglycerides, mg/dL</td>
<td>101±24</td>
<td>143±94</td>
<td>148±67</td>
</tr>
<tr>
<td>AST, U/L</td>
<td>42±26</td>
<td>56±42</td>
<td>52±31†</td>
</tr>
<tr>
<td>ALT, U/L</td>
<td>32±11</td>
<td>40±23</td>
<td>87±59</td>
</tr>
</tbody>
</table>

Histological features

<table>
<thead>
<tr>
<th>Variables</th>
<th>Control</th>
<th>NAFL</th>
<th>NASH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree of steatosis (0–3)</td>
<td>0</td>
<td>1.5±0.7</td>
<td>2.2±0.4†</td>
</tr>
<tr>
<td>Lobular inflammation (0–3)</td>
<td>0</td>
<td>0.7±0.3</td>
<td>1.2±0.8‡</td>
</tr>
<tr>
<td>Hepatocellular ballooning (0–2)</td>
<td>0</td>
<td>0±0</td>
<td>0.8±0.6‡</td>
</tr>
<tr>
<td>Fibrosis stage</td>
<td>0</td>
<td>0±0</td>
<td>1.6±1.0†</td>
</tr>
<tr>
<td>NAFLD activity score (NAS)</td>
<td>0</td>
<td>2.5±1.19</td>
<td>4.3±1†</td>
</tr>
</tbody>
</table>

NAFLD morbidity obese patients

<table>
<thead>
<tr>
<th>Variables</th>
<th>Control</th>
<th>NAFL</th>
<th>NASH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of subjects</td>
<td>10</td>
<td>23</td>
<td>27</td>
</tr>
<tr>
<td>Female, %</td>
<td>6/4</td>
<td>12/11</td>
<td>15/12</td>
</tr>
<tr>
<td>Age, years</td>
<td>44±8.8</td>
<td>43±9</td>
<td>48±10</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>55±14</td>
<td>53±13</td>
<td>49±10</td>
</tr>
<tr>
<td>Type 2 diabetes (n)</td>
<td>4</td>
<td>8</td>
<td>18‡</td>
</tr>
<tr>
<td>Fasting plasma glucose, mg/dL</td>
<td>105±23</td>
<td>101±22</td>
<td>138±63‡</td>
</tr>
<tr>
<td>Fasting plasma insulin, μU/mL</td>
<td>11±8.4</td>
<td>13±7</td>
<td>35±45†</td>
</tr>
<tr>
<td>HOMA-IR index</td>
<td>2±1.2</td>
<td>3±1.7</td>
<td>16±40†</td>
</tr>
<tr>
<td>Total cholesterol, mg/dL</td>
<td>190±36</td>
<td>180±37</td>
<td>179±49</td>
</tr>
<tr>
<td>HDL-cholesterol, mg/dL</td>
<td>40±10</td>
<td>45±10</td>
<td>37±6</td>
</tr>
<tr>
<td>LDL-cholesterol, mg/dL</td>
<td>121±36</td>
<td>122±28</td>
<td>126±43</td>
</tr>
<tr>
<td>Triglycerides, mg/dL</td>
<td>128±65</td>
<td>154±57</td>
<td>191±102</td>
</tr>
<tr>
<td>ALT, U/L</td>
<td>19.7±8</td>
<td>28±27</td>
<td>43±20†</td>
</tr>
<tr>
<td>AST, U/L</td>
<td>19.6±8</td>
<td>29±20</td>
<td>31±14†</td>
</tr>
</tbody>
</table>

Histological features

<table>
<thead>
<tr>
<th>Variables</th>
<th>Control</th>
<th>NAFL</th>
<th>NASH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree of steatosis, %</td>
<td>0</td>
<td>1.7±0.8</td>
<td>2.19±0.8†</td>
</tr>
<tr>
<td>Lobular inflammation (0–3)</td>
<td>0</td>
<td>0.35±0.6</td>
<td>1.42±0.8†</td>
</tr>
<tr>
<td>Hepatocellular ballooning (0–2)</td>
<td>0</td>
<td>0.18±0.4</td>
<td>1.07±0.62†</td>
</tr>
<tr>
<td>Fibrosis stage</td>
<td>0</td>
<td>0.04±0.2</td>
<td>1.7±0.6†</td>
</tr>
<tr>
<td>NAFLD activity score (NAS)</td>
<td>0</td>
<td>2.2±1.44</td>
<td>4.7±2†</td>
</tr>
</tbody>
</table>

P value stands for statistical significance using Mann-Whitney U test, except for female/male proportion that p value stands for statistical significance using χ² test. Results are expressed as means±SD.

*P<0.001 indicates NAFL versus controls.

‡P<0.001 denotes comparisons between NAFL and control subjects.

‡P<0.001 indicates comparisons between NAFL and NASH.

ALT and AST, serum alanine and aspartate aminotransferase; BMI, body mass index; HDL, high-density lipoprotein; HOMA, homeostatic model assessment; IR, insulin resistance; LDL, low-density lipoprotein; NAFL, non-alcoholic fatty liver; NAFLD, non-alcoholic fatty liver disease; NASH, non-alcoholic steatohepatitis.
Cohort-specific analysis of liver microbial DNA signatures and liver histology

Analysis of the 16S rRNA gene signature in each individual cohort revealed that the liver bacterial metataxonomic signature was significantly influenced by the host phenome (morbid obesity vs non-morbid obesity).

When controls (non-NAFLD) were compared with patients with NAFLD, we found that the hepatic microbial DNA profile varied greatly in taxa composition, indicating that some taxa may be associated with a protective effect against fatty liver. Nevertheless, a distinctive pattern was observed in severely obese individuals (figure 3A,B). For example, cohort-specific analysis showed that some bacterial DNA was seen most frequently in the liver specimens of non-NAFLD morbidly obese subjects (figure 3A), including *Brevibacillus, Dialister, Prevotella two* and *Aliivibrio*. Conversely, in the liver of non-NAFLD non-morbidly obese subjects, the most abundant taxa included *Marivita* (order Rhodobacterales) and *Lachnoclostridium* (order Clostridiales).

The within-cohort analysis of bacterial DNA composition aimed at elucidating the NAFLD severity (NAFL vs NASH) showed associations with histological traits across the entire spectrum, including features of more aggressive stages such as ballooning degeneration and fibrosis. Detailed metataxonomic analysis at the genus and family levels and their relationship with main histological traits are shown in online supplementary tables 5 and 6.

Specifically, we observed that the presence of bacterial DNA of *Roseibacillus* (Verrucomicrobia/Verrucomicrobiae), *Peptostreptococcus* (Firmicutes/Clostridia), *Bifidobacterium* (Actinobacteria/Actinobacteria) and *Streptomyces* (Actinobacteria/Actinobacteria) was particularly dominant in the NAFLD morbidly obese cohort, and the more dramatic increases were
Liver bacterial DNA profiling is influenced by the host phenome. (A) Box plot showing the level (median and SD) of differential bacterial DNA abundance (taxonomy: class) between groups. P value denotes non-parametric one-way analysis of variance (Wilcoxon rank test) statistics corrected by false discovery rate. (B) Cladogram representing the LDA-LEfSe results, comparing taxonomic composition of liver bacterial DNA in the samples obtained from morbidly obese and non-morbidly obese cohorts, including controls, NAFL and NASH patients. (C) sPLS-DA classification performance of liver microbial DNA composition in morbidly obese versus non-morbidly obese patients with NAFLD based on the relative abundance of bacterial taxa at the class level. Graph shows the first two components with 95% confidence level ellipse plots. LDA, Linear Discriminant Analysis; LEfSe, Linear Discriminant Analysis (LDA) Effect Size; NAFL, non-alcoholic fatty liver; NASH, non-alcoholic steatohepatitis.

Joint analysis of metataxonomic profiles suggests that major differences in bacterial DNA composition of the liver tissue in patients with NASH are driven by enrichment with Proteobacteria taxa in the Gamma class

Next, we focused on the differences in liver bacterial DNA composition across the histological stages of disease severity, from the control (non-NAFLD) liver to NASH in the whole population, which was normalised per sample and cohort values (morbid obesity and non-morbid obesity) according to the LEfSe method. All features were analysed by Kruskal-Wallis test, including whether the values in different classes were differentially distributed.
The LEfSe approach was used to identify the taxa that most likely explain differences in the disease severity. The pairwise Wilcoxon test was therefore conducted to assess whether all pairwise comparisons between groups significantly agree with the class-level trend. Finally, an LDA model was developed, yielding relative differences among classes that were used to rank the features. It revealed that the most differentially abundant bacterial taxa in control liver specimens belong to phylum Firmicutes. In contrast, over-represented taxa in NASH specimens included Gammaproteobacteria and Pseudomonadales (figure 3C, online supplementary figure 7). Taxonomic representation of statistically and biologically consistent differences among control liver, NAFL and NASH is shown in figure 3D. Overabundance of Gammaproteobacteria DNA in the liver of patients with NASH was further validated by an independent quantitative molecular approach (RT-qPCR) using primer sequences with high coverage and specificity for Pseudomonadaceae, including Xanthomonadaceae/STenotrophomonas (figure 3E). Depletion of Lachnospiraceae DNA in the liver of NASH patients was further validated by qPCR (figure 3F).

Validation of main findings by an independent molecular approach

We validated NGS bacterial DNA abundances of eight representative taxa (Alpha and Gamma Proteobacteria, Actinobacteria, Peptostreptococcus, Verrucomicrobia, Veillonella, Lachnospiraceae and Bacteroides/Prevotella) that showed the most significant differences in the analysis of association with the disease severity and related histological traits. Validation was performed by RT-qPCR using taxon-specific primers listed in the online supplementary material.

Results of copy number abundances of taxa associated with severe histological traits were consistent with the findings of 16S rRNA gene metagenomic sequencing, as all effects remained in the same direction as those yielded by the NGS analysis. There was no detectable amplification in non-template controls in any of the assays. A summary of cohort-specific validation results can be found in online supplementary figures 8,9.

LPS derived from Gram-negative bacteria localise in the portal tract

We hypothesised that Gammaproteobacteria DNA found in tissue samples—particularly Enterobacteriales—originates in

Figure 3  Cohort-specific and joint analysis of liver microbial DNA signatures. (A) Analysis of liver microbial DNA signatures in morbidly obese patients (bariatric surgery cohort). (B) Analysis of liver microbial DNA signatures in non-morbidly obese patients. Measurements across sample groups are compared by the non-parametric Kruskal-Wallis statistical test. Significantly different taxa are shown as bar chart. SE is depicted by error bars. Pairwise comparisons are performed by Mann-Whitney U test and are annotated as *p<0.05 and **p<0.01. Filter: include top 300 most abundant taxa. Control: non-NAFLD subjects. (C) LDA score in the joint analysis of liver microbial DNA signatures: visualisation of differential features with statistical and biological significance, ranked according to the effect size. (D) Cladogram representing the LDA LEfSe results, comparing taxonomic composition of control, NAFL and NASH liver tissue. Main features in the cladogram are represented on phylogenetic trees. The cladogram reports the taxa (highlighted by small circles and by shading) showing different abundance values (according to the LEfSe results) in the three groups as described in the results section; for each taxon, the colour denotes the class with higher median for both the small circles and the shading. These graphical outputs were generated by the LEfSe visualisation modules publicly available at the Galaxy website. (E and F) Validation of results by qPCR in the joint analysis. A non-parametric Kruskal-Wallis test was used for pairwise comparison of the subject groups. Bars represent mean±SEM. Data were analysed using GraphPad Prism version 4.00 for Windows (GraphPad software, San Diego, California, USA). The results were considered statistically significant at p<0.05. *In silico PCR on the Silva database (https://www.arb-silva.de/search/testprime) shows that sequence primers used to validate Gammaproteobacteria DNA abundance have 96% coverage and 82% specificity for Pseudomonadaceae and 99% coverage and 80% specificity for Xanthomonadaceae/STenotrophomonas. (G–I) Lipopolysaccharides (LPS) from Gram-negative Proteobacteria (Enterobacteriales) localise in the portal tracts. Immunohistochemistry analysis was performed to assess the abundance and localisation of LPS in the liver. (A) Representative sample of non-NAFLD liver; (B and C) representative samples of NAFLD (NAFL and NASH fibrosis, respectively). LPS expression was assessed in five non-NAFLD samples and 16 samples of patients with NAFLD (NAFL n=8; NASH n=8). Original magnification 200×. LDA, Linear Discriminant Analysis; LEfSe, Linear Discriminant Analysis Effect Size; NAFL, non-alcoholic fatty liver; NASH, non-alcoholic steatohepatitis; Pt, portal tract.
the gut and translocates to the liver through the portal system. Therefore, the liver tissue was examined for the presence of another bacterial constituent, bacterial LPS. As positive control tissue, we included a tissue specimen of necrotic colon cancer in which we observed marked immunoreactivity of LPS (online supplementary figure 10). Immunohistochemical analysis revealed that LPS immunostaining localised in the portal tracts with absent or minimal immunoreactivity in liver zones 2 or 3 (figure 3H/I). Clusters of LPS particles were more frequently seen in areas of portal inflammation. In addition, compared with non-NAFLD liver specimens (0.2±0.27 mean, SD) and liver samples of patients with NAFL (0.75±0.27), the liver of patients with NASH (2.05±0.8) presented significantly higher number of portal tracts stained positive for antibodies against LPS (p=0.0004) (figure 3 H/I).

Liver bacterial DNA profiling shows distinctive predicted functional signatures

Specialised pathways were observed in each patient subsample, which were modulated by the host phenome-related factors (severe obesity vs non-severe obesity). Details can be found in the online supplementary material and online supplementary figure 11−15.

DISCUSSION

NAFLD specialists have long relied on liver biopsy to infer the clinical impact of histological findings, but revolutionary advances in molecular and sequencing techniques have provided new avenues for enhancing our understanding of the disease pathogenesis. Here, we extended the analysis of the disease beyond risk factors and histological features to include largely unknown features of the liver bacterial metataxonomic profile. We focused on severe (morbid) obesity and non-morbid obesity and confirmed that the liver microbial DNA signature of NAFLD is dramatically modulated by the host phenotype. Although the liver microbiota composition in terms of main phyla was comparable between the morbidly obese and non-morbidly obese NAFLD cohorts, we observed significant differences in the abundance profiling and diversity measurements and in the predicted functionality of bacterial DNA.

Specifically, we found that, compared with non-morbidly obese patients, the liver of morbidly obese patients with NAFLD showed decreased Alpha diversity. In addition, we noted significant differences in the Beta diversity between the two cohorts, as measured by the non-phylogeny-based Bray-Curtis method and the weighted UniFrac, which uses the OTU’s phylogeny. Analysis of other relevant clinical covariates indicated that there were no associations with sex, type 2 diabetes or arterial hypertension.

Furthermore, we found that bacterial DNA profile of the liver of morbidly obese patients with NAFLD is characterised by over-abundance of species of the phyla Bacteroidetes and Firmicutes, more specifically the classes Bacteroidia and Clostridia, respectively, that represented ~8% of the diversity variance in the PCoA 1. In contrast, in the liver of non-morbidly obese NAFLD patients, we noted overabundance of species of the phyla Proteobacteria and Deinococci—specifically, Alpha and Gammaproteobacteria, Bacillli and Deinococci, which represented ~16% of the diversity variance in the PCoA 2.

Assessment of differences in the liver bacterial DNA composition across the full disease severity spectrum, including the control liver, and the liver of patients with NAFL and NASH in the whole population normalised per sample and cohort values (BMI) according to the LEfSe method, showed that the disease severity is associated with Proteobacteria in both Alpha and Gamma classes. Specifically, the most pronounced fold changes in microbial DNA associated with NASH were seen for Gammaproteobacteria, Pseudomonadaceae and Pseudomonadales. These findings are in agreement with those yielded by previous explorations of human NAFLD microbiota signatures in non-liver tissues, particularly the gut lumen microbiome,19 and the circulating microbiota of patients with NASH20 or alcoholic liver disease.21

Assessment of liver bacterial DNA abundance and histological traits highlighted the specific impact of the architectural metataxonomic signature in each patient cohort, indicating that the uneven distribution of bacterial DNA species differentially influenced the histological phenotypes. Ruminoclostridium and Alloprevotella depletion was associated with NASH in non-morbidly obese patients, whereas increased abundance of Xanthomonadaceae family, as well as some particular members of the Clostridia, Actinobacteria and Verrucomicrobiae classes—including Peptostreptococcus, Bifidobacterium and Roseibacillus—was associated with NASH among morbidly obese patients. Significantly greater proportion of the anaerobes Peptostreptococcus was noted in the samples of morbidly obese patients with NASH relative to NAFL. This genus is recognised as GI inhabitants,31 and recent studies have shown increased levels of Peptostreptococcus sp. in faecal samples of patients with ulcerative colitis22 and colorectal cancer.32 In addition, marked expansion of Verrucomicrobia DNA was seen in the liver specimens of morbidly obese patients with NASH—a finding that was validated by qPCR using a primer set with high specificity (99.9%) for Roseibacillus (Verrucomicrobiae) (online supplementary figure 8). The implication of Roseibacillus in the context of human diseases remains largely unknown; these Gram-negative taxa have been isolated in aquatic and terrestrial environments.33

The overall analysis showed consistent association of Proteobacteria (Alpha or Gamma) DNA with features of the disease severity, including lobular and portal inflammation, liver fibrosis and the NAFLD Activity Score (NAS) score. This is highly relevant in light of available evidence suggesting that Gram-negative bacteria, particularly Proteobacteria, are involved in the genesis of endotoxaemia—also described as metabolic endotoxaemia—35, 36 and inflammation,36, 37 high levels of ethanol production,14, 20 evasion from immune system surveillance and defence mechanisms or molecular mimicry18 and predisposition to cardiovascular (CV) disease.39

We investigated the putative presence of bacterial LPS in the liver and found that Gram-negative-derived LPS staining was more frequently observed in the liver of NASH patients with severe fibrosis than in NAFL or non-NAFLD samples and we demonstrated that LPS predominantly localised in the portal tracts. These results suggest that Gram-negative-derived DNA (and other constituents) found in the liver may also enter the liver through the portal system. Hence, it is highly plausible to assume that gut represents the main source of Gram-negative Gammaproteobacteria, particularly Enterobacteriales.

We identified some individual bacterial DNA that might be associated with protection against liver inflammation. This beneficial effect was observed for members of the Lachnospiraceae family, the Lachnospiraceae NK4A136 and Lachnospiraceae FCS020, as well as a member of the Bacteroidia class—the Alloprevotella genus that was significantly depleted in the liver of NAFLD-non-morbidly obese patients with more severe histological phenotypes. The validation analysis showed significantly reduced levels of Alloprevotella in patients with ballooning degeneration and fibrosis as well. In agreement with our observation, recent
evidence indicates that administration of Lachnospiraceae attenuates obesity, inflammation and dysbiosis in mice fed high fat diet. 40 Capsular faecal transplant of beneficial taxa such as Ruminococcaceae and Lachnospiraceae was correlated with cognitive improvement and decrease in inflammation in patients with cirrhosis. 41 Likewise, the taxa Alloprevotella has been associated with decreased lifetime CV risk. 42 We have found, however, that different genera within the family Prevotellaceae—for example Prevotella 9 and Alloprevotella—had opposite effects on liver histological features. This observation is in agreement with available CV research findings. 42

Nevertheless, it is worth noting that the presence of bacterial DNA in the liver tissue does not necessarily imply existence of live bacteria. It is highly plausible that bacterial DNA could be present as debris in immune cells, specifically in Kupffer cells. 43

Limitations and possible implications

The main limitation of our study is its cross-sectional design, which does not permit any causality inferences. Specifically, we cannot confirm that the disease severity and progression is caused by the observed microbial signature (bacterial DNA). Similarly, we cannot confirm whether the injured—diseased—and/or, for instance, treatment and prevention of infections of the disease and for expanding therapeutic options, including the gut microbiota.

In conclusion, this is the first comprehensive study showing that the liver tissue of NAFLD patients contains a diverse repertoire of bacterial DNA (up to 2.5 x 10^5 reads counts after exhaustive filtering). Measurement of 16S rRNA gene content by qPCR using universal primers showed that total bacterial DNA of the liver tissue is not negligible. More importantly, the most pronounced differences within cohorts between mild and severe histological features were independently validated by RT-qPCR.

The liver microbial DNA signature may explain differences in NAFLD pathogenic mechanisms, as well as physiological functions of the host. NAFLD subtypes might be distinguished by the microbial composition of the liver microbiome. Our results may have important implications for unravelling the molecular basis of the disease and for expanding therapeutic options, including probiotic selection or BS that should be more precisely defined according to specific clinical scenarios, such as the obesity status and/or, for instance, treatment and prevention of infections of the mouth.

Author affiliations

1 Institute of Medical Research A Lanari, University of Buenos Aires Faculty of Medicine, Buenos Aires, Argentina
2 Institute of Medical Research (IDIM), Department of Clinical and Molecular Hepatology, National Scientific and Technical Research Council, Buenos Aires, Argentina
3 Institute of Medical Research (IDIM), Department of Molecular Genetics and Biology of Complex Diseases, National Scientific and Technical Research Council, Buenos Aires, Argentina
4 Liver Unit, Medicine and Surgery Department, Hospital General de Agudos Dr Abel Zubizarreta, Buenos Aires, Argentina
5 Department of Surgery, Hospital El Cruce, Buenos Aires, Argentina

Twitter Silvia Sookooian @Ssookooian and Carlos Jose Pirola @cpirola, @panat

Contributors SS: study concept and design; data acquisition; performed liver biopsies and collected biological material; histological evaluation; immunohistochemistry (IHC); RT-qPCR, data analysis and interpretation; general study supervision; manuscript drafting; and securing funding. AS: bioinformatics analysis of sequencing data and data processing and presentation. MSL and CF: sample processing; MG and GOC: performed liver biopsies and collected biological samples. CJP: study concept and design, biological material collection and data acquisition; RT-qPCR; data analysis and interpretation; statistical analysis; IHQ; manuscript drafting; and general study supervision and securing funding.

Funding This study was partially supported by grants PID-C2012-0061, PICT 2014-0432, PICT 2014-1816, PICT 2015-0551 and PICT 2016-0135 (Agencia Nacional de Promoción Científica y Tecnológica, FONCyT), CONICET Proyectos Unidades Ejecutoras 2017, PUE 0055.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Biological specimens, including blood samples and liver biopsies from all subjects included in this study, were obtained with written, informed consent in accordance with the Institutional Review Board-approved protocols (protocol numbers: 104/HG2AZ/09, 89/100 and 120/2012).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iDs
Silvia Sookooian http://orcid.org/0000-0001-5929-5470
Carlos Jose Pirola https://orcid.org/0000-0001-8234-4058

REFERENCES


