RESEARCH ARTICLE

Altered gut microbial energy and metabolism in children with non-alcoholic fatty liver disease

Sonia Michail1,2,*, Malinda Lin1,2, Mark R. Frey1,2, Rob Fanter1,3, Oleg Paliy5, Brian Hilbush4 and Nicholas V. Reo5

1Department of Pediatrics, Division of Gastroenterology, Hepatology, and Nutrition at Children’s Hospital Los Angeles, 4650 Sunset Blvd. MS#78, Los Angeles, CA 90027, USA, 2Keck School of Medicine at the University of Southern California, Los Angeles, CA 90089, USA, 3Childrens Hospital Los Angeles Proteomics Core, 4650 Sunset Blvd. MS#78, Los Angeles, CA 90027, USA, 4Real time Genomics, 999 Bayhill Dr #101, San Bruno, CA 94066, USA and 5Department of Biochemistry & Molecular Biology, Boonshoft School of Medicine, Wright State University, 3640 Colonel Glenn Hwy, Dayton, OH 45435, USA

*Corresponding author: University of Southern California, Los Angeles, CA 90089, USA; Children’s Hospital of Los Angeles, 4650 Sunset Blvd MS#78, Los Angeles, CA 90027, USA. Tel: +323-361-5924; Fax: +323-361-2425; E-mail: sonia.michail@hotmail.com

One sentence summary: This study describes the gut microbial profile of obese children with and without fatty liver and identifies several key metabolites that have potential to contribute to the pathophysiology of this common disease.

ABSTRACT

Obesity is becoming the new pediatric epidemic. Non-alcoholic fatty liver disease (NAFLD) is frequently associated with obesity and has become the most common cause of pediatric liver disease. The gut microbiome is the major metabolic organ and determines how calories are processed, serving as a caloric gate and contributing towards the pathogenesis of NAFLD. The goal of this study is to examine gut microbial profiles in children with NAFLD using phylogenetic, metabolomic, metagenomic and proteomic approaches. Fecal samples were obtained from obese children with or without NAFLD and healthy lean children. Stool specimens were subjected to 16S rRNA gene microarray, shotgun sequencing, mass spectroscopy for proteomics and NMR spectroscopy for metabolite analysis. Children with NAFLD had more abundant Gammaproteobacteria and Prevotella and significantly higher levels of ethanol, with differential effects on short chain fatty acids. This group also had increased genomic and protein abundance for energy production with a reduction in carbohydrate and amino acid metabolism and urea cycle and urea transport systems. The metaproteome and metagenome showed similar findings. The gut microbiome in pediatric NAFLD is distinct from lean healthy children with more alcohol production and pathways allocated to energy metabolism over carbohydrate and amino acid metabolism, which would contribute to development of disease.

Key words: NASH; obesity; gut microbiome; carbohydrate; energy metabolism

INTRODUCTION

The incidence of obesity among children in the United States has increased 3-fold in the last few decades (Kuczmarski et al. 2002; Ogden et al., 2006). The obesity epidemic is paralleled by an increase of non-alcoholic fatty liver disease (NAFLD) in the pediatric population, affecting at least 3% of children and adolescents in the United States and Asia (Strauss et al., 2000; Park et al., 2005). NAFLD is the most common cause of liver disease
in children and adolescents in the United States (Schwimmer et al., 2006). The term NAFLD includes a spectrum of histological features from simple steatosis to inflammation with cellular injury to frank cirrhosis. Because the definitive diagnosis requires liver biopsy, which is not feasible in pediatric studies, increased serum aminotransferase levels and increased echogenicity on radiographic studies have been used as diagnostic tools (Patton et al., 2006). Non-alcoholic steatohepatitis (NASH) can potentially progress to cirrhosis, liver failure and hepatocellular carcinoma (White et al., 2012; Stepanova et al., 2013).

The pathophysiology of NAFLD remains unclear, but the gut microbiome is considered an important contributor. The gut microbiome houses bacterial communities that direct host digestion and energy homeostasis (Backhed et al., 2005; Sonnenburg et al., 2005; Flint et al., 2008). The ‘obese microbiota’ is characterized by an increased capacity to harvest energy from the diet (Turnbaugh et al., 2006). Studies on obese mice showed a 50% reduction in Bacteroidetes and an increase in Firmicutes by a comparable amount, demonstrating that ‘obesogenic microbiota’ is distinct from ‘lean microbiota’ (Ley et al., 2005) and that the differences in microbiota composition in mice can determine response to a high-fat diet and whether those mice will develop hepatic steatosis (Le Roy et al., 2013), as well as how the disease will progress (Henao-Mejia et al., 2012). Comparable observations are noted in human studies. Microbiota of obese adults had fewer Bacteroidetes (Mouzaki et al., 2013) and more Firmicutes resulting in difference in Bacteroidetes-to-Firmicutes ratio between obese and non-obese subjects, and the pattern was reversed with weight loss (Ley et al., 2006). Some studies suggest that the ‘manipulation of gut microbial communities could be another approach in the treatment of obesity’ (Jumpertz et al., 2011). Recent studies support this notion and further demonstrate that microbiota can be transmissible and could prevent the development of obesity and obesity-associated metabolic phenotypes in experimental animals (Ridaura et al., 2013; Walker and Parkhill 2013).

The objective of this study was to compare the fecal gut microbiome community structure, composition and function in obese children with and without NAFLD to those of healthy children. Microarray analysis was used to characterize microbial communities based on the 16S rRNA gene. The functional capacity was determined using high-throughput metagenomic sequencing and functional differences were determined using MS proteomic and NMR metabolomic analyses of fecal samples.

**METHODS**

This study was approved by the Institutional Review Board at the Children’s Hospital Los Angeles and Dayton Children’s Hospital. Lean healthy children and obese children with and without a clinical diagnosis of NAFLD based on ultrasound findings and elevated transaminases were recruited into the study. Other liver diseases were excluded. We could not definitely exclude early liver disease in the obese group without clinical markers of NAFLD since confirmation of absence of NAFLD in the obese subjects would require a liver biopsy. For ethical issues, we could not obtain liver biopsies in this subset of patients for the purpose of this study.

Body mass index (BMI) for healthy children were less than the 85% for age, and the BMI for obese children with and without clinical NAFLD were greater than the 95% for age. The average aspartate and alanine transaminases were 96 and 157 IU/L, respectively, in the NAFLD group. BMI was calculated as weight in kilograms divided by height in meters squared. AST and ALT were performed through local commercial labs which utilize the modified IFCC method (Lustig et al., 1988). There were no dietary restrictions for any subjects during this study and there was no significant difference in caloric intake, carbohydrate, protein or fat intake between the three groups. None of the children received antibiotics for at least six months prior to the study.

**Gut microbial 16S rRNA gene microarray**

All fecal specimens were homogenized and immediately frozen at −80 °C. Triplicate samples of each stool were processed. Total genomic DNA (gDNA) was isolated using ZR Fecal DNA Isolation Kit (Zymo Research Corporation) according to manufacturer’s protocol. gDNA was amplified, pooled together, fragmented and then hybridized to Microbiota Array as we previously described (Palyi et al., 2009; Agans et al., 2011; Michail and Kenche 2011; Michial et al., 2012).

**High-throughput DNA sequencing**

Samples were prepared using sequencing based on the standard protocol for the Ion One Touch™ System and loaded onto Ion 318™ chips and run on the Ion Torrent Personal Genome Machine™ (Life Technologies, Carlsbad, CA) under standard conditions. The genomic center at Children’s Hospital of Los Angeles is CLIA certified, and all CLIA quality control measures were applied. Shotgun sequencing of whole gDNA were performed on DNA samples without prior 16S rRNA gene amplification.

**Proteomic processing**

Fecal specimens were thawed at 4 °C and microbial cells extracted by differential centrifugation (Apajalahti et al., 2002; Verberkmoes et al., 2009). Microbial cell pellets (−100 mg wet weight) were washed, and then lysed via probe sonication in wash buffer with 0.25% w/v Rapigest. Lysed samples were centrifuged at 15 K rcf for 5 min, and the protein concentration was determined via bicinchoninic acid colorimetric assay (Chutipongtanate et al., 2012). Thirty mcg were fractionated via SDS-PAGE and in-gel digested overnight with trypsin. Tryptic peptides were purified via Varian C18 OMIX tips and analysed using an Eksigent nanoLC-2D coupled to a Thermo Orbitrap XL tandem mass spectrometer. Human and microbial proteins were identified from raw MS data using Proteome Discoverer and Scaffold software packages in conjunction with the UniProt human proteome and the METADB database, respectively. The METADB database has been previously described by Verberkmoes et al. (2009).

**NMR metabolite analysis**

Stools were frozen at −80 °C until further analysis. Fecal water was extracted using a weighed sample of thawed stool material (200 mg) and adding five volumes of sterile phosphate buffer (1.9 mM Na2HPO4, 8.1 mM NaH2PO4, pH 7.4). The mixture was homogenized by vortex mixing for 5 min. The fecal slurry was centrifuged at 4500 rpm for 10 min and the supernatant was filtered through a GDX filter (Whatman). The filtrate was centrifuged at 16 000 g for 15 min and the supernatant retained for NMR analysis.

NMR studies were conducted using a Varian Inova 600 NMR. Briefly, 550 μl of fecal extract was mixed with 150 μl of a solution of 9 mM trimethylsilylproponic acid (TSP) in D2O, which was used as a chemical shift and quantification reference. A 1H NMR spectrum was acquired at 600 MHz and 25 °C using a NOESY pulse sequence (first increment) with water suppression during the relaxation delay and mixing time (the total
interpulse delay ~ 11 s). Spectral intensities were normalized to the total sum of all metabolite signals and integrated peak areas were measured for specific resonances of interest (known metabolite signals).

**Statistical analysis**

**Quantification of NMR metabolites**

Quantification of specific metabolite resonances was accomplished using an interactive spectral deconvolution algorithm in MATLAB adapted from our previously described methods (Anderson et al., 2009, 2011). The deconvolution tool fits a defined spectral region using a combination of tunable baseline shapes (spline, v-shaped, linear or constant) and a Gauss–Lorentz peak-fitting function. Integrated areas for peaks of interest are then output to a text file. All metabolite peak intensities were corrected for equivalent number of protons and normalized relative to the TSP signal intensity; but signals were not corrected for partial T1 saturation. Since all extracts were prepared from an identical stool weight, relative comparisons between samples are fully valid, but absolute quantitative amounts of each metabolite are not reported. Analyses of proton NMR data included multivariate data analyses of the full spectra and of 12 measured fecal metabolites. To obtain average levels of metabolites among all samples within each sample type, weighted mean around median was calculated as we described previously (Agans et al., 2011). P values were adjusted for multiple comparisons using Bonferroni corrections.

**Metabolic profiling, sequencing and statistical analysis**

We utilized the HUMAnN analysis pipeline together with RTG’s mapx translated nucleotide search algorithm to infer community-wide metabolic pathway and gene abundances in the study. Filtered sequence data were searched against the KEGG database with mapx, using the following parameters: w = 8, a = 3 (sensitivity settings equivalent to BLASTX—blastall with default settings) and identity threshold at 70%. Output from mapx was processed by the HUMAnN pipeline (Abubucker et al., 2012).

**Proteomic analysis**

Proteins were quantitated in Scaffold according to number of assigned spectra, using 95% minimum protein and peptide identification probabilities, and a minimum peptide count of two. Protein quantification was averaged across three samples obtained from three separate stool specimens for control and NAFLD samples, and the resulting values were used for analysis, including functional category mapping via the databases of clusters of orthologous groups (COG).

**Microarray statistical analysis**

Signals from 852 separate bacteria were obtained from subjects in all groups. The bacteria were classified into 120 different genera, 53 families, 25 orders, 14 classes and 8 phyla. The signal from each bacterium for each patient was then divided by its genus average copy number to obtain an adjusted signal for each bacterium. The 852 adjusted signals were summed and the percent of the total adjusted signal was determined for all bacteria. Adjusted signals were then summed by genus, family, order, class and phylum, and the percentages of the total signal were determined for each hierarchical classification level.

### Table 1. Characteristics of healthy children and obese children with NAFLD.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Healthy</th>
<th>NAFLD</th>
<th>Obese no evidence of NAFLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>14 (53.8)</td>
<td>6 (46.0)</td>
<td>8 (72.7)</td>
</tr>
<tr>
<td>Female</td>
<td>12 (46.2)</td>
<td>7 (54.0)</td>
<td>3 (27.3)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>26 (100)</td>
<td>12 (92.4)</td>
<td>11 (100)</td>
</tr>
<tr>
<td>Other</td>
<td>0 (0)</td>
<td>1 (7.6)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Age (y)</td>
<td>13.3 ± 2.7</td>
<td>13.6 ± 3.0</td>
<td>13.2 ± 3.8</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>155.5 ± 16.9</td>
<td>164.4 ± 18.0</td>
<td>159.2 ± 19.1</td>
</tr>
<tr>
<td>BMI (kg/m2)</td>
<td>19.4 ± 3.7(a)</td>
<td>40.8 ± 4.0(b)</td>
<td>31.2 ± 4.5(c)</td>
</tr>
</tbody>
</table>

Values listed are total for the variable (percent of total value n) or Means ± SD. Across rows different superscript letters indicate means that are significantly different (P < 0.01) using analysis of variance (ANOVA) with Bonferroni correction.

Lack of superscript letters indicate means that are not significantly different, P > 0.05

For statistical analyses, all adjusted signals were transformed using the natural logarithmic transformation. Comparisons between groups were made with repeated measures analysis of variance. P values were adjusted for multiple comparisons using Bonferroni corrections.

**RESULTS**

**Demographic characteristics**

All three groups of children had similar characteristics except for a higher BMI (Table 1) in the two NAFLD and the obese children without evidence of NAFLD. All children with NAFLD had evidence of fatty liver change on ultrasonography and elevated AST and ALT levels. Three children underwent liver biopsy in our cohort; two had evidence of NASH and the third had steatosis without NASH. Some of the obese children with or without NAFLD had coexisting functional abdominal pain. Family history of obesity was observed in 78, 69 and 31% of children with NAFLD, obese without NAFLD and healthy children, respectively, while 34, 15 and 4% had a family history of fatty liver.

**16S rRNA gene analysis of bacterial populations**

Differences exist at all levels when comparing the microbiota profiles of the NAFLD and healthy groups. Table 2 demonstrates changes noted at phylum and class levels. At the phylum level, children with NAFLD harbor more Gamma- and Epsilonproteobacteria than the healthy group. At the genus level, children with NAFLD harbor a more abundant member of the phylum Bacteroidetes—Prevotella, which is detected at lower levels in healthy children as well as obese children without evidence of NAFLD (Table 2). There was no difference in Firmicutes and Bacteroidetes or their ratio in any of the groups. However, obese children without NAFLD had lower Clostridia and higher Alphaproteobacteria.

**Nuclear magnetic resonance measurements of small metabolites**

Data were analysed by spectral integration of a few specific signals assigned to various chemicals of interest, shown in
Table 2. Microbiota profile in healthy and obese children with and without fatty liver. Percent (mean ± SD) by phylum and class/genus.

<table>
<thead>
<tr>
<th>Phylum</th>
<th>Class (subclass)/[genus]</th>
<th>No. of species</th>
<th>Healthy (n = 26)</th>
<th>NAFLD (n = 13)</th>
<th>Obese without evidence of NAFLD (n = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actinobacteria</td>
<td>Actinobacteria [Actinobacteridae]</td>
<td>20</td>
<td>1.02 ± 0.12</td>
<td>1.10 ± 0.34</td>
<td>1.14 ± 0.22</td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>Actinobacteria [Coriobacteridae]</td>
<td>9</td>
<td>1.05 ± 0.11</td>
<td>1.13 ± 0.22</td>
<td>1.04 ± 0.28</td>
</tr>
<tr>
<td>Bacteroidetes</td>
<td>Bacteroidia</td>
<td>142</td>
<td>14.97 ± 2.43</td>
<td>16.59 ± 5.9</td>
<td>16.28 ± 2.77</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>Bacillia [Prevotella]</td>
<td>25</td>
<td>0.19 ± 0.04a</td>
<td>7 ± 0.13a</td>
<td>0.21 ± 0.03a</td>
</tr>
<tr>
<td></td>
<td>Clostridia</td>
<td>564</td>
<td>76.66 ± 2.51a</td>
<td>74.81 ± 5.83ab</td>
<td>74.65 ± 2.72b</td>
</tr>
<tr>
<td></td>
<td>Erysipelotrichi</td>
<td>39</td>
<td>2.24 ± 0.39a</td>
<td>0.30 ± 0.03b</td>
<td>2.49 ± 0.27a</td>
</tr>
<tr>
<td>Fusobacteria</td>
<td>Fusobacteria</td>
<td>3</td>
<td>0.13 ± 0.01a</td>
<td>0.15 ± 0.09b</td>
<td>0.14 ± 0.07a</td>
</tr>
<tr>
<td>Lentisphaerae</td>
<td>Lentisphaeria</td>
<td>1</td>
<td>0.06 ± 0.01</td>
<td>0.06 ± 0.03</td>
<td>0.08 ± 0.01</td>
</tr>
<tr>
<td>Proteobacteria</td>
<td>Alphaproteobacteria</td>
<td>9</td>
<td>0.17 ± 0.03a</td>
<td>0.10 ± 0.34b</td>
<td>0.20 ± 0.03a</td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>18</td>
<td>0.61 ± 0.15</td>
<td>0.52 ± 0.26</td>
<td>0.70 ± 0.28</td>
<td></td>
</tr>
<tr>
<td>Deltaproteobacteria</td>
<td>5</td>
<td>0.30 ± 0.04</td>
<td>0.28 ± 0.06</td>
<td>0.30 ± 0.05</td>
<td></td>
</tr>
<tr>
<td>Epsilonproteobacteria</td>
<td>6</td>
<td>0.27 ± 0.04a</td>
<td>0.30 ± 0.03b</td>
<td>0.27 ± 0.03a</td>
<td></td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>11</td>
<td>0.66 ± 0.14a</td>
<td>0.73 ± 0.10b</td>
<td>0.77 ± 0.09f</td>
<td></td>
</tr>
<tr>
<td>Spirochaetes</td>
<td>Spirochaetia</td>
<td>4</td>
<td>0.13 ± 0.01</td>
<td>0.13 ± 0.04</td>
<td>0.13 ± 0.02</td>
</tr>
<tr>
<td>Verrucomicrobia</td>
<td>Verrucomicrobiae</td>
<td>1</td>
<td>0.07 ± 0.05</td>
<td>0.04 ± 0.03</td>
<td>0.05 ± 0.04</td>
</tr>
</tbody>
</table>

Across rows, different superscript letters indicate means that are significantly different (p < 0.05) using analysis of variance (ANOVA) with Bonferroni correction.

Table 3. Relative metabolite levels measured by 1H NMR of fecal water extracts from obese children with NAFLD, obese children without evidence of NAFLD and normal healthy children. Values are Mean ± SE of integrated signal intensities (arbitrary units) normalized to TSP standard.

<table>
<thead>
<tr>
<th>METABOLITE</th>
<th>Healthy (n = 22)</th>
<th>Obese with NAFLD (n = 11)</th>
<th>Obese without evidence of NAFLD (n = 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethanol</td>
<td>15.3 ± 4.1a</td>
<td>32.1 ± 9.8b</td>
<td>12.4 ± 4.2ab</td>
</tr>
<tr>
<td>Uric acid</td>
<td>7.2 ± 0.8</td>
<td>6.1 ± 1.6</td>
<td>8.9 ± 1.8</td>
</tr>
<tr>
<td>Acetate</td>
<td>550.0 ± 47.1a</td>
<td>417.9 ± 44.7b</td>
<td>497.5 ± 95.8ab</td>
</tr>
<tr>
<td>Formate</td>
<td>7.0 ± 0.5a</td>
<td>4.2 ± 1.3b</td>
<td>7.0 ± 1.0ab</td>
</tr>
<tr>
<td>Glucose</td>
<td>17.8 ± 5.1</td>
<td>23.4 ± 3.0</td>
<td>27.4 ± 7.6</td>
</tr>
<tr>
<td>Fumarate</td>
<td>1.4 ± 0.2</td>
<td>1.4 ± 0.5</td>
<td>1.6 ± 0.5</td>
</tr>
<tr>
<td>Lactate</td>
<td>17.9 ± 2.7</td>
<td>11.7 ± 2.3</td>
<td>15.1 ± 1.6</td>
</tr>
<tr>
<td>Alanine</td>
<td>34.9 ± 5.2ab</td>
<td>27.0 ± 3.9a</td>
<td>42.9 ± 8.1ab</td>
</tr>
<tr>
<td>Butyrate</td>
<td>221.0 ± 30.5</td>
<td>226.0 ± 36.8</td>
<td>260.1 ± 75.9</td>
</tr>
<tr>
<td>Propionate</td>
<td>250.2 ± 30.9</td>
<td>227.3 ± 44.0</td>
<td>249.3 ± 45.4</td>
</tr>
<tr>
<td>Valerate</td>
<td>19.9 ± 3.0ab</td>
<td>9.9 ± 2.3b</td>
<td>12.2 ± 2.2ab</td>
</tr>
<tr>
<td>Lysine</td>
<td>44.3 ± 8.1a</td>
<td>62.8 ± 16.1ab</td>
<td>81.7 ± 29.6b</td>
</tr>
</tbody>
</table>

Across rows, different superscript letters indicate means that are significantly different (p < 0.05) using analysis of variance (ANOVA) with Bonferroni correction.

Table 3. Some short chain fatty acids (SCFA), specifically formate, acetate, and valerate were less abundant in the NAFLD group relative to healthy controls. However, other SCFA (butyrate and propionate) were unaffected. The NAFLD group also had significantly higher levels of ethanol than healthy children. Obese children without evidence of NAFLD had higher lysine metabolite compared to healthy children. There was no statistical difference observed in lactate among the three groups.

Shotgun gut microbial metabolite sequencing

Subsequently, we identified metabolic pathways that are exclusively present in the healthy (n = 11) or NAFLD groups (n = 7, Tables 4A and B and Tables S4 and S5 in the Supporting Information). The average number of sequences obtained for each sample was 2 850 000 sequences with a GC content of 51 ± 4. We also noted pathways that were more abundant in either population. There are more pathways for energy metabolism and lipid synthesis and fewer pathways for amino acid metabolism, especially in relation to the urea cycle and elimination of ammonia in the NAFLD group. Methane oxidation was found to be many folds higher in children with NAFLD compared to healthy children. This finding suggests that energy metabolism in the gut microbiome of children with NAFLD is much more efficient than in healthy children. This is also supported by the proteomic findings of more pathways dedicated to energy production (Fig. 2).

Metaproteomics

Noting differences in metabolism at the genomic level through sequencing, we attempted to identify such associations at the protein level using proteomics. The bacterial proteins identified
**Table 4A.** Shotgun genomics: pathways identified exclusively in children with NAFLD.

<table>
<thead>
<tr>
<th>Metabolic pathway</th>
<th>KEGG description</th>
</tr>
</thead>
<tbody>
<tr>
<td>M00030: lysine biosynthesis, 2-oxoglutarate = 2-aminoadipate = lysine</td>
<td>Amino acid metabolism: amino acid synthesis</td>
</tr>
<tr>
<td>M00032: lysine degradation, lysine = saccharopine = acetoacetyl-CoA</td>
<td>Amino acid metabolism: amino acid degradation</td>
</tr>
<tr>
<td>M00142: complex I (NADH dehydrogenase), NADH dehydrogenase I</td>
<td>Energy metabolism; ATP synthesis</td>
</tr>
<tr>
<td>M00319: manganese/zinc/iron transport system</td>
<td>ABC transporters-iron and B12 TRANSPORT</td>
</tr>
</tbody>
</table>

**Table 4B.** Shotgun genomics: pathways more abundant in NAFLD.

<table>
<thead>
<tr>
<th>Metabolic pathway</th>
<th>Fold change</th>
<th>KEGG description</th>
</tr>
</thead>
<tbody>
<tr>
<td>M00174: methane oxidation, methylotroph, methane = CO2</td>
<td>4.87E+05</td>
<td>Energy metabolism</td>
</tr>
<tr>
<td>M00287: PTS system, galactosamine-specific II component</td>
<td>111</td>
<td>Galactose metabolism</td>
</tr>
<tr>
<td>M00082: fatty acid biosynthesis, initiation</td>
<td>6.67</td>
<td>Lipid synthesis</td>
</tr>
<tr>
<td>M00012: glyoxylate cycle</td>
<td>5.56</td>
<td>Carbohydrate and Lipid metabolism</td>
</tr>
<tr>
<td>M00096: CS isoprenoid biosynthesis, non-mevalonate pathway</td>
<td>3.23</td>
<td>Carbohydrate and Lipid metabolism</td>
</tr>
</tbody>
</table>

**Figure 1.** Distribution of COG functional categories in healthy children and obese children with NAFLD.

were classified into COG functional categories. The most common COG functional categories identified in all groups were related to carbohydrate and energy metabolism (Fig. 1). The abundance of proteins expressed in different groups was compared, while the total number of proteins for each analysis was the same. There are significantly higher numbers of COG pathways for energy production in the NAFLD than in the healthy group (Fig. 2). This is also consistent with the metagenomic findings described above. The healthy group has more pathways allocated for amino acid and carbohydrate metabolism, but these two categories are still highly represented within the NAFLD group. Because more pathways...
Figure 2. Total assigned MS/MS spectra for COG-catalogued proteins that were found to be differentially expressed by a factor of at least three. C: energy production and conversion; CQ: COG functional categories C and Q (secondary metabolites biosynthesis, transport and catabolism); E: amino acid transport and metabolism; ET: COG functional categories E and T (signal transduction mechanisms); G: carbohydrate transport and metabolism; H: coenzyme transport and metabolism; J: translation, ribosomal structure and biogenesis; O: post translational modifications, protein turnover, chaperones; U: intracellular trafficking, secretion, and vesicular support.

are devoted to energy production, the NAFLD group appears to have fewer metabolic pathways for carbohydrate and amino acid metabolism, which may account for the effects seen on SCFA. COG categories more highly represented in healthy children include intracellular trafficking, translation and ribosomal biogenesis, inorganic ion transport and metabolism (Fig. 2). Proteins expressed 3-fold higher in NAFLD children that are relevant to energy metabolism include NAD-dependent aldehyde dehydrogenases and glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase (Table S2 in the Supporting Information).

Through the human database, we identified human proteins more highly expressed in the NAFLD group (Table S3 in the Supporting Information). Alpha amylase 1 and pancreatic triacylglycerol lipase are more highly expressed in the NAFLD children and are relevant to carbohydrate and lipid metabolism.

DISCUSSION

The human gut microbiome is composed of a complex structure of trillions of organisms that contribute to the processing of different components of our diet such as carbohydrates, fats and proteins thus serving as the gate that permits energy and caloric retention and subsequent development of obesity. Both genetics and environment play a role in obesity and its sequela, such as NAFLD; however, genetic susceptibility cannot fully explain the epidemic in the United States. Because adherence to lifestyle changes is difficult, it is important to explore alternative targets for therapy. As we uncover more about the gut microbiome and its role in host metabolism and disease, its relevance in the obesity epidemic increases.

Microbiome and host metabolism

The gut microbiota influences host metabolism through a variety of mechanisms. Gut microbiota ferment indigestible carbohydrates, producing SCFA. SCFA are precursors in lipid production and gluconeogenesis, activate hormones that inhibit gut motility, increase nutrient absorption and stimulate appetite (Machado and Cortez-Pinto 2012). Frost and colleagues demonstrate that acetate activates acetyl-CoA carboxylase and changes in the expression profiles of regulatory neuropeptides that cause appetite suppression (Frost et al., 2014). The microbiome influences expression of Angiopoietin-like protein 4 and AMP-activated protein kinase, two important regulatory factors in fatty acid production and adipose tissue deposition (Backhed et al., 2004, 2007). In addition, there is evidence that the gut microbiota can direct increases in hepatic triglyceride production (Backhed et al., 2004). When the gut microbiota is exposed to a high-fat diet, a higher plasma lipopolysaccharide level is achieved. This state has been termed ‘metabolic endotoxemia’, which promotes liver steatosis, inflammation and fibrogenesis (Cani and Delzenne 2009; Laugerette et al., 2011; Zhu et al., 2013). With such a wide range of implications, altering the gut microbiota can be useful in preventing metabolic consequences (Delzenne and Cani 2008). This concept has already proven its efficacy in experimental animals (Ridaura et al., 2013). The emerging field of metagenomics and metaproteomics permits studying the functions of microbial communities by sampling directly from their habitat (Maron et al., 2007). Analysing fecal samples from lean children and obese children with NAFLD, we identified differences in bacterial populations, gene and protein expression, and metabolite levels. We observed an increase in bacterial subgroups responsible for SCFA and
ethanol production and a shift towards energy metabolism in obese children with NAFLD. Although some SCFA were observed to be less abundant, ethanol was significantly higher in the stool from obese children with NAFLD.

Gut microbial phylogenetic differences

There are significant differences among the gut bacterial populations in healthy and obese children with and without NAFLD. Namely, increases in the genus Prevotella and phylum Gammaproteobacteria have clinical importance and were found to be increased in children with NAFLD but not healthy or obese children without evidence of NAFLD. Prevotella is known to play a role in carbohydrate fermentation, producing exogenous SCFA (Salminen et al., 1998; Pouteau et al., 2003; Zhang et al., 2009). The present pediatric study, however, demonstrates the complexity of the relationship between gut microbes and metabolites, as differential effects were observed in SCFA. While we observed a decrease in some SCFA, none of the SCFA measured were found to be elevated in the NAFLD groups. Zhang et al. have highlighted the potential importance of Prevotellaceae in the obesity-associated microbiota, suggesting that H2-producing members of this family coexist with H2-oxidizing methanogenic Archaea in the gastrointestinal tracts of obese individuals (Zhang et al., 2009). This syntrophic relationship may increase host energy extraction from indigestible carbohydrates, as an increase in H2-oxidizing methanogenesis facilitates fermentation. Therefore, Prevotellaceae populations may be an important factor in the association between obesity, fatty liver, ethanol and SCFA production and the gut microbiota. This is consistent with adult studies that suggest that Prevotella populations are under-represented in non-obese subjects (Hoyles and McCartney 2009). However, in our pediatric population, obese children without NAFLD did not have a statistically significant increase in Prevotella suggesting that its role is more specific in NAFLD rather than obesity in general. We do not detect significant differences in the distribution of Firmicutes and Bacteroidetes between healthy and NAFLD groups as prior adult studies have demonstrated (Ley et al., 2006; Jumpertz et al., 2011). This may be related to the patient population being younger and early in the stages of obesity compared to adults.

Alcohol production

Furthermore, both Gammaproteobacteria (Ren et al., 2007) and Prevotella have been associated with endogenous alcohol production, which could account for the high-ethanol signal detected in the NAFLD group. Our findings represent one of the early pediatric studies describing this phenomenon and is consistent with animal studies that demonstrate production of ethanol in microbiota of obese mice with fatty liver, which was attenuated with Neomycin (Cope et al., 2000). Endogenous alcohol production through bacterial fermentation of dietary carbohydrates may contribute to the pathogenesis of fatty liver by delivering a constant source of oxidative stress. This is demonstrated in studies that describe high-blood ethanol levels in obese mice models and in patients with NAFLD (Cope et al., 2000; Volynets et al., 2012; Zhu et al., 2013). Zhu and Baker showed that children and adolescents with NASH had more alcohol-producing bacteria, higher serum ethanol levels and increased gene expression of pathways involved in alcohol catabolism (Zhu et al., 2013). Furthermore, obesity and fat deposition in the liver potentiate the damaging effects of alcohol on the liver (Xu et al., 2011). The abundance of Gammaproteobacteria classes in the NAFLD groups provides clues for future mechanistic studies that could elucidate the initial pathological changes that contribute to the development of NAFLD.

Energy metabolism

Our data show that the NAFLD group has a greater proportion of pathways dedicated to energy production and conversion and a smaller proportion for amino acid and carbohydrate transport and metabolism in comparison to the lean group. The metaproteome and metagenome data show similar findings. Looking at specific protein expression, NAD-dependent aldehyde dehydrogenase is 3-fold higher in the NAFLD group. This protein can impact glycolysis, gluconeogenesis, fatty acid and amino acid metabolism. But more importantly, it may play a direct role in the production of ethanol from acetaldehyde and is likely responsible for the elevated ethanol levels observed in the NAFLD group. The predominance of these pathways supports the idea that obese microbiota are active in energy extraction. Experimental models showed that less energy was lost in the stool in obese mice than their lean littermates, suggesting more efficient energy extraction by the obese microbiome. Conventionalized mice actually had a higher metabolic rate than their germ-free counterparts. Interestingly, despite a higher metabolic rate, there was no accompanying increase in tissue high-energy phosphate stores in conventionalized mice, implying the presence of inefficient metabolism (Backhed et al., 2004; Turnbaugh et al., 2006). We have evidence that the NAFLD microbiome is highly equipped for energy metabolism but cannot explain how the energy is utilized or allocated. Obese children with NAFLD had similar dietary intake compared to healthy children yet developed obesity. This may be explained on the basis of harboring a microbiome that is more efficient in making and preserving calories as noted in the findings of this study showing that gut microbial energy conservation and production was higher allowing the gut microbiome to conserve calories for the host with subsequent weight gain.

Carbohydrate and lipid metabolism

Although the NAFLD group has a smaller proportion of functional groups dedicated to carbohydrate transport and metabolism, these pathways still represent one of the principal functional groups in NAFLD children. This finding has been suggested by previous animal and human studies. Comparative metagenomic analysis of genetically lean and obese mice showed that the cecal microbiome from obese mice was enriched in enzymes involved in carbohydrate digestion (Ley et al., 2006). In a human study comparing metabolic pathways of the microbiome of lean and obese twins, the obese human gut was enriched in pathways involved in microbial processing of carbohydrates (Turnbaugh et al., 2009).

From our human database analysis, there is a 3-fold increase in pancreatic triacylglycerol lipase and alpha amylase 1 in the NAFLD group. Pancreatic triacylglycerol lipase is central in the digestion and absorption of dietary fat, and alpha amylase 1 is responsible for carbohydrate digestion. Alpha amylase 1 inhibitors are being studied to reduce the risk of heart attacks, treat diabetes type II and glucose intolerance (Chiasson et al., 2003). Amylase inhibition is known to induce carbohydrate tolerance, satiety and weight loss (Laugier et al., 2011). Pancreatic triacylglycerol lipase is the main enzyme responsible for the hydrolysis and absorption of triacylglycerol from our diet. Lipid hydrolases are also targets for pharmacotherapy; Orlistat is
a gastrointestinal lipase inhibitor which results in a reduction in the absorption of exogenous fat (Finer et al., 2000). These differences have clinical consequences and may contribute to obesity and NAFLD. The significance of pancreatic triacylglycerol lipase is further noted in mice deficient in pancreatic triacylglycerol lipase and consuming a Western diet and those who were found to have lower fasting cholesterol levels. Because of their reduced dietary lipid absorption, these mice had a significantly reduced susceptibility to high-fat/high-cholesterol diet-induced obesity (Gilham et al., 2007).

CONCLUSION
Our study highlights the complexity of the gut microbial ecology as it contributes to the development of obesity and fatty liver in the pediatric population. It demonstrates that no one technology alone can demonstrate the complexity of its structure or the mechanisms by which gut bacteria influence metabolism and energy balance. Looking at the picture from multiple perspectives—phylogenetic, metabolomic, metagenomic and proteomic—helps elucidate the role the gut microbiome plays in the development of NAFLD. We have yet to determine the mechanisms resulting in NAFLD, and our study adds to existing data that the process is multifactorial. This is one of few studies that focus specifically on the pediatric population. Our population is unique because we are evaluating children with early stages of NAFLD, which may be different from long-standing liver disease seen in adults. We also studied obese children without clinical features of NAFLD, revealing similarities between the two obese groups. These observations support the idea of a continuum in the diagnosis of fatty liver disease and progression to clinical NAFLD. For future studies, it is important to study obese patients without NAFLD as a separate group to identify potential factors that lead to NAFLD in some obese patients but not others and to pinpoint the early changes in the microbiome that result in NAFLD. Our study identifies differences in bacterial subtypes in NAFLD patients that have previously been associated with increased ethanol production and SCFA fermentation, which may have implications for liver damage and increased energy extraction. Our metagenomic and proteomic data show more bacterial pathways devoted to energy production and conversion that also supports the idea of providing more energy to the host. Results of this study may potentially provide a tool to distinguish children with NAFLD who have a gut microbial profile that is more efficient in processing energy and harbors bacteria that are capable of producing ethanol and SCFA. These findings can also help shape future interventional efforts to modify the microbiome in a way that would reverse the offending metabolite profile in NAFLD.

SUPPLEMENTARY DATA
Supplementary material is available at FEMSEC online.

ACKNOWLEDGEMENTS
The authors wish to thank Adrienne Stolfi for her assistance with some of the statistical analyses, Sean McCann, Harsha Kenche and Daniel Homer for technical assistance and Dr Tim Triche for his assistance with genomics. Part of the initial work was performed while first author was affiliated with Wright State University.

FUNDING
This work was supported by the National Institutes of Health grants NCCAM AT003423 and NICHD HD065575 and funding through the Clinical and Translational Science Institute (CTSI), Children’s Hospital Los Angeles and MemorialCare Foundation.

Conflict of interest statement. None declared.

REFERENCES


