1 Introduction

Breast milk jaundice (BMJ) commonly occurs in healthy breastfed infants. Prolonged unconjugated hyperbilirubinemia begins from postnatal 5-7 days, peaks on around the tenth day, and persists for more than 12 weeks. BMJ was first reported in 1963 by Newman et al. [1]. It is defined as a benign condition of prolonged unconjugated hyperbilirubinemia in healthy breastfed babies who are of normal body weight and who have normal urine production, and with no signs of disease. A diagnosis of BMJ can be made if the serum bilirubin levels of breastfed infants return to normal within the first three months, and no other causes of jaundice have been identified [2]. However, bilirubin encephalopathy caused by hyperbilirubinemia is far more likely to occur in breastfed infants [3-5].

The etiological mechanism of BMJ is not completely clear, although there are many hypotheses. It is noteworthy that gut microbiome composition and function are closely associated with the increased reabsorption of bilirubin in BMJ, and thus they can determine to a large extent the prognosis of this benign condition. Intestinal bilirubin absorption may be influenced by newly established intestinal bacterial populations in neonates. This population can reduce the production of bilirubin available for intestinal reabsorption by converting bilirubin glucuronides to urobilinoids [6, 7]. Recently, many studies have indicated that breast milk is a significant source of intestinal microbiota in the neonatal gut. Staphylococcus, Streptococcus, Lactobacillus, and Bifidobacterium are the most frequently detected bacterial...
genera in human milk [8-10]. Species of these genera as well as species of *Enterococcus* can be transferred from mother to infant through breastfeeding [11, 12], and *Bifidobacterium* and *Lactobacillus* can reduce serum bilirubin levels. Nevertheless, the precise involvement of the gut microbiome and the specific bacterial genera associated with the development of BMJ remains unclear due to the limitations of the sequencing methods currently used and the selection of the 16S rRNA region sequenced.

In the present study, we hypothesized that intestinal flora may be crucial to the development of BMJ. If this hypothesis was confirmed, then it would provide a theoretical basis for the prevention or treatment of BMJ, through interventions targeted to the specific gut microbiome involved in its development. Therefore, to understand better the involvement of the gut microbiome in BMJ, alterations in the fecal microbial communities in infants with BMJ were evaluated by 16S ribosomal RNA (rRNA) high-throughput sequencing.

# 2 Materials and methods

## 2.1 Study population

 Initially, 34 infants at the maternity units were screened and enrolled in the study based on the following inclusion criteria: fed by breast exclusively, jaundice remission from the time that breast feeding ceased; normal and full-term; without perinatal problems. The mothers of these infants had not received probiotic or antibiotic treatments during pregnancy or after birth, and the infants had not received probiotics, antibiotics, or probiotic-supplemented formulas after birth. Within postnatal days 14-35, the participants provided samples of infant feces.

Infants were excluded for the following reasons: fecal samples were not provided during postnatal days 14-35; diseases that are known risk factors of BMJ, including blood type incompatibilities, glucose-6-phosphate dehydrogenase deficiencies, Coombs test positivity, hemolytic disease; or diseases that may cause hyperbilirubinemia, such as diabetes mellitus or infection.

Finally, 12 infants with BMJ and prolonged jaundice were enrolled in the study. Twenty-two healthy infants with no jaundice from their birth in the hospital to the time of study initiation were included in the control group.

**Informed consent:** Informed consent has been obtained from all individuals included in this study.

**Ethical approval:** The research related to human use has been complied with all the relevant national regulations, institutional policies and in accordance the tenets of the Helsinki Declaration, and has been approved by the Institutional Review Board of the Children's Hospital of Chongqing Medical University (Chongqing, China).

## 2.2 Feces specimen collection

Each fecal sample obtained from the infants was transferred into 10-mL sterile tubes, which were then placed on ice and immediately sent to the laboratory for subsequent sequencing [13].

## 2.3 DNA extraction and PCR amplification

An E.Z.N.A. DNA isolation Kit (Omega Bio-tek, Norcross, GA, USA) was used for extraction of bacterial DNA from the fecal samples, in accordance with the instructions of the manufacturer. The V3-V4 region of the bacterial 16S ribosomal DNA gene was selected for PCR amplification. The PCR amplification conditions were as follows: initial denaturation at 95°C for 2 min; then 25 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 30 s; and a final extension at 72°C for 5 min. The forward primer sequence used for PCR amplification was 5´-barcode-GTGCCAGCMGCCGCGG-3´, and the reverse primer sequence was 5´-CCGTCAATTCMTTTRAGTTT-3´, where barcode is an individual 8-nucleic acid base sequence for each sample. PCR reactions were performed in 20 μL of a mixture containing 0.8 μL of 5 μM of each primer, 0.4 μL of Pfu DNA polymerase, 4 μL of 5× Pfu buffer, 2 μL of 2.5 mM dNTP, and 10 ng of DNA templates.

## 2.4 Illumina MiSeq sequencing

An AxyPrep DNA gel extraction kit (Axygen Biosciences, Union City, CA, USA) was used to extract and purify the PCR product in accordance with the manufacturer’s instructions. The extracted DNA was quantified using a QuantFluor system (Promega, USA), and then was pooled in equimolar quantities and paired-end sequenced (2 × 250). An Illumina MiSeq platform was used in accordance with the manufacturer’s protocols. The raw reads were deposited in the NCBI Sequence Read Archive database (Accession Number: SRP064975).
2.5 Sequencing data processing

The raw Illumina sequencing data were processed with QIIME (Quantitative Insights into Microbial Ecology; version 1.9.1). The processing workflow started with FASTQ files, which were quality-filtered and demultiplexed using the following criteria: a sequence containing one or more N bases was ruled out; depending on the settings of a default parameter, the read was truncated at the base preceding the first set of stretches with a low quality, and the truncated sequence had to be ≥ 75 bases to be retained; a sequence whose barcode did not match exactly a barcode in the mapping file was removed; and a sequence that could not be accumulated was removed.

UPARSE software (version 7.1 http://drive5.com/uparse/) was used to cluster the operational taxonomic units (OTUs) based on 97% similarity. The UCHIME algorithm was utilized to remove chimeric sequences. RDP Classifier (a Bayesian classifier that provides taxonomic assignments from domain to genus) was used to analyze the taxonomy of each 16S rRNA gene sequence (http://rdp.cme.msu.edu/) against the SILVA (SSU123) 16S ribosomal RNA database, with a confidence threshold of 70% [14]. Mothur was employed to analyze the rarefaction curve and Simpson index curve [15]. Simpson’s Diversity Index was calculated to quantify the biodiversity (http://www.mothur.org/wiki/Simpson). The LEfSe (linear discriminant analysis effect size) method was used to analyze the statistical differences between the BMJ and control groups.

2.6 KEGG function prediction

PICRUSt (phylogenetic investigation of communities by reconstruction of unobserved states) was used to predict the functional composition of a microbial community metagenome from its 16S profile. For the 16S result, PICRUSt-compatible OTU tables (i.e., counts of OTUs on a per-sample basis) were constructed in accordance with the closed-reference OTU-picking protocol. In addition, KEGG orthologues were mapped according to the abundances of the KEGG Module; each module was correlated with a specific functional category.

2.7 Statistical analysis

All statistical analyses were performed using SPSS 18.0 software. Student’s t-test was applied to examine the differences between the BMJ case and control groups when the bacterial counts in feces had a normal distribution, and the results are expressed as mean ± standard error. The Wilcoxon rank-sum test was applied when the bacterial counts in the feces were not normally distributed, and the results are presented as mean ± standard deviation and median (upper and lower quartiles). P-values < 0.05 were considered statistically significant.

3 Results

3.1 Study population

The study population comprised 12 infants with prolonged BMJ and 22 healthy infants (BMJ infants show a 5/7 gender split compared to healthy control’s 10/12). There was no statistical difference between the two groups regarding day age, birth position, or gender ratio (Table 1). To estimate the species-level diversity and perform comparisons between the case and control groups, all sequences were clustered to form OTUs using Mothur with a similarity threshold of 97%. A total of 1,176,681 reads and 1215 OTUs were obtained from the 34 samples. The rarefaction curves of the Sobs index were then calculated for each sample (Figure 1). As can be seen from the beta diversities, there were significant differences in the bacterial composition profiles of the BMJ case and control groups (Adonis analysis; R² = 0.125; P < 0.001; Figure 2).

The rarefaction curves approached a plateau with the current sequencing, and the sequencing coverage reached 99.88%, indicating that the depth of the sequencing was sufficient to cover all types of bacteria. The alpha diversity was analyzed based on the richness index (Ace and Chao),

| Table 1. Clinical features of the patients in the BMJ and control groups *. |
|-----------------|-----------------|-------------|-----|
|                 | BMJ             | Control     | Statistic | P   |
| Collection time after birth, d | 24.7 ± 2.2      | 26.1 ± 3.4   | 0.35   | 0.73 |
| Delivery, cesarean section/eutocia | 7/5             | 14/8        | —       | 1    |
| Infant gender, male/female | 5/7             | 10/12       | —       | 1    |

*Reported as n/n, unless noted otherwise.
diversity index (Shannon and Simpson), and evenness index (Shannoneven) (Table 2). The Simpson index of the BMJ case group (0.946 ± 0.538) was significantly lower than that of the control group (1.453 ± 0.739; \( P < 0.05 \); Table 3). This result shows the presence of a significantly higher microbial diversity in the BMJ infants than in the controls.

### 3.2 Overall phylogenetic profiles of gut microbiota

The gene copy number of glutamate-5-semialdehyde dehydrogenase in the BMJ group (10,696.58 ± 6108.19) was significantly higher than that of the control group (6982.82 ± 2179.33; \( P < 0.05 \)). In addition, the *Escherichia/Shigella* concentrations in the control group were higher than those in the jaundiced group (\( P < 0.05 \)). No statistical difference between these two groups was observed in the concentrations of *Bifidobacterium*, *Enterococcus*, *Streptococcus*, and *Staphylococcus* (Figure 3).

### 3.3 Abundance of bacteria in the BMJ and control groups

The LEfSe method was used to analyze the statistical differences between the BMJ and control groups regarding microbial communities. This method was especially useful for the analysis of data where the number of samples was much lower than the number of species. Cladograms of the taxa with least discriminant analysis (LDA) values > 2.0 are depicted in Figure 4. There were 10 groups of bacteria that were abundant in the fecal samples of the healthy control infants (\( P < 0.05 \)), namely, *Escherichia/Shigella*, *Bacteroidaceae*, *Bacteroides*, *Clostridiaceae*, *Clostridium*, *Lactobacillaceae*, *Lactobacillus*, *Ferruginibacter*, *Marmoricola*, and *Geobacillus*. The lineages of the first 5 of these groups had an LDA value ≥ 4. There were also 10 groups of bacteria that were plentiful in the samples of feces from the BMJ infants, namely, *Rothia*, *Haemophilus*, *Actinomycetales*, *Actinomyces*, *Corynebacteriaceae*, *Propionibacteriaceae*, *Anoxybacillus*, and *Synergistaceae*.

### Table 2. Summary of diversity indices and coverage estimators of BMJ and control groups.

<table>
<thead>
<tr>
<th></th>
<th>BMJ</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage, %</td>
<td>99.89 ± 0.23</td>
<td>99.87 ± 0.17</td>
</tr>
<tr>
<td>Shannon</td>
<td>1.91 ± 0.23</td>
<td>1.14 ± 0.75</td>
</tr>
<tr>
<td>Simpson</td>
<td>0.95 ± 0.54</td>
<td>1.45 ± 0.74</td>
</tr>
<tr>
<td>Ace</td>
<td>170.82 ± 120.33</td>
<td>232.64 ± 175.97</td>
</tr>
<tr>
<td>Chao</td>
<td>163.67 ± 93.57</td>
<td>212.50 ± 81.38</td>
</tr>
<tr>
<td>Shannoneven</td>
<td>0.27 ± 0.03</td>
<td>0.28 ± 0.02</td>
</tr>
</tbody>
</table>

Figure 1. Rarefaction curves for OTU calculated using Mothur software.

Figure 2. Unweighted UniFrac principal coordinates analysis (PCoA) plot comparing sample distribution between the two cohorts. Green and red dots represent healthy controls and BMJ infants, respectively; c1: BMJ group; case1: control group.
Table 3. Relative abundance of the microbial community, Simpson index, and glutamate-5-semialdehyde dehydrogenase content in the BMJ case and healthy control groups. *

<table>
<thead>
<tr>
<th></th>
<th>BMJ</th>
<th>Control</th>
<th>Statistic</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Escherichia/Shigella</em></td>
<td>0.165 ± 0.212</td>
<td>0.429 ± 0.346</td>
<td>−2.755</td>
<td>0.01</td>
</tr>
<tr>
<td><em>Bifidobacterium</em></td>
<td>0.184, 0.425</td>
<td>0.039, 0.146</td>
<td>−1.622</td>
<td>0.105</td>
</tr>
<tr>
<td><em>Enterococcus</em></td>
<td>0.005, 0.215</td>
<td>0.002, 0.013</td>
<td>−1.55</td>
<td>0.121</td>
</tr>
<tr>
<td><em>Streptococcus</em></td>
<td>0.005, 0.046</td>
<td>0.012, 0.148</td>
<td>−1.297</td>
<td>0.195</td>
</tr>
<tr>
<td><em>Staphylococcus</em></td>
<td>0.005, 0.037</td>
<td>0.004, 0.016</td>
<td>−0.18</td>
<td>0.857</td>
</tr>
<tr>
<td>Simpson index</td>
<td>0.946 ± 0.538</td>
<td>1.453 ± 0.739</td>
<td>−2.089</td>
<td>0.045</td>
</tr>
<tr>
<td>GPR</td>
<td>10,696.58 ± 6108.19</td>
<td>6,982.82 ± 2179.33</td>
<td>2.592</td>
<td>0.014</td>
</tr>
</tbody>
</table>

* Expressed as mean ± standard deviation or median with interquartile range.

GPR, glutamate-5-semialdehyde dehydrogenase

4 Discussion

The present study provides a detailed comparison of breastfed infants, with or without jaundice, of their fecal bacterial communities. Twelve infants with BMJ were selected for a gut microbiome composition analysis according to strict screening standards, and contributed significant data together with 22 healthy infants. The high-throughput sequencing covered more than 99.84% of genes, a sequencing depth that is sufficient for the comparison of taxon richness. The gradually flat rarefaction curves of OTUs may have been caused by real diversity, and not experimental biases.

The ratio of *Escherichia* or *Shigella* in the gut microbiome was significantly lower, and the intestinal...
16S ribosomal RNA-based gut microbiome composition analysis in infants with breast milk jaundice

showing that obstructive jaundice correlated with higher levels of bacterial diversity [16]. In a report concerning 29 healthy Chinese 2-month-old infants [17], the diversity of the fecal microbiome was significantly higher than that of neonates, and the most abundant bacterial genera microbial flora diversity was remarkably higher, in the BMJ infants, compared with the healthy infants. There were no significant differences in flora diversity among the 22 normal infants. Our results are consistent with those of a report on the occlusion of biliary stents,
were Veillonella, Bacteroides, and Lactobacillus. It is worth mentioning that the population size of Escherichia and Enterococcus actually slightly lower. Therefore, the bacterial diversity in infants at different ages may vary in significance with regard to the incidence of BMJ.

In addition, in several studies on necrotizing enterocolitis, a low alpha-diversity of gut flora and no obvious difference in the intestinal bacterial composition was found between patients with necrotizing enterocolitis and controls [18, 19]. These findings may explain the significantly higher diversity of intestinal microbial flora in BMJ infants compared with the healthy control infants. However, an important limitation of the PICRUSt approach in interpreting data and making predictions is that PICRUSt’s ability to detect patterns depends on the input data used. The software cannot distinguish a variation at the strain level if the marker gene sequence used is identical among strains, and it cannot detect gene families if those genes are not included in the input genomic data used, or if the pathway annotations are currently poor.

Mother’s body contains some bacterial DNA that can be transmitted to infants and can affect the composition of infants’ gut microbiota. For example, DNA of the common gut bacteria Lactobacillus and Bifidobacterium was reportedly detected in human placentas [20, 21], and a strain of Enterococcus faecium was transferred from pregnant mice to the fetal umbilical cord blood [22]. In another examination, the prevalence of Bifidobacterium in the bacterial gut communities of breastfed infants was obviously affected by mother’s fucosyltransferase 2 status [23]; hence, the types of bacteria residing in an infant’s gut may depend on the bacterial DNA in the mother’s gut. Breastfeeding may influence the bacterial flora in the intestinal tract of infants and is thus important to the incidence of BMJ.

Altered gut microbiota composition is associated with various diseases in infants, such as eczema [24], autism spectrum disorders [25], inflammatory bowel disease [26], and diabetes [27]. Using high-throughput sequencing in the present study, we determined the presence of diverse bacterial genera, of which Escherichia/Shigella accounted for the highest proportion in healthy breastfed infants. The harmless strain Escherichia coli is part of the normal flora of the gut and is the most abundant bacteria in the infant intestine throughout the first year after birth [28]. Intestinal flora is a complex microbial community colonized in the digestive tracts of humans. In our study, Escherichia was found to be more abundant in healthy breastfed infants than in infants with BMJ, which suggests that Escherichia is also involved in preventing BMJ.

Intestinal microflora was reported to affect the serum bilirubin level in a Gunn rat model of hyperbilirubinemia [29]. After oral administration of an antibiotic, the fecal urobilinoids disappeared, and the serum bilirubin level went up [30]. In the human gut, bilirubin is degraded into urobilinoids by microflora. Bilirubin can be metabolized selectively by UDP-glucuronosyltransferase 1A1 and transformed into a water-soluble glucuronide [31]. Beta-glucuronidase in the gut brush border converts conjugated bilirubin to the unconjugated form, which is reabsorbed. This enzyme is also present in breast milk, which contributes to the development of neonatal jaundice [32-34].

In the present study, many metabolic pathways were predicted to be related to the presence of BMJ, including the bacterial invasion of epithelial cells, fatty acid elongation in mitochondria, lipopolysaccharide biosynthesis, and carbohydrate metabolism (Figure 5). In addition, the level of glutamate-5-semialdehyde dehydrogenase in the BMJ group was significantly higher than in the control group. The expression of this predicted gene was also reported to be 3.20- and 2.03-fold-up higher, respectively, under bile and acidified bile stress [35]. These data indicate that the enzyme is involved in the urea cycle, and metabolism of amino groups is exceedingly important to the occurrence of BMJ [36].

Neonatal breastfeeding jaundice correlates highly with intestinal bilirubin absorption, and higher levels can lead to an increase in enterohepatic circulation [37, 38]. Elevated enterohepatic circulation is currently considered the most probable mechanism for neonatal jaundice caused by breastfeeding [39]. Intestinal bacteria can convert bilirubin glucuronides to urobilinoids, and therefore a low amount of bilirubin is reabsorbed in the intestinal tract. Probiotics such as Saccharomyces boulardii lowered the serum bilirubin level of healthy neonates with jaundice [40]. Furthermore, Bifidobacterium bifidum was predominant in the intestine of healthy breastfed infants [41] and was detected in higher quantities in the fecal samples of the control group than in those of the BMJ group [42]. Nevertheless, in our study, no obvious difference of B. bifidum was found in infants’ feces between the control and the BMJ group. This was likely due to the appearance of Bifidobacterium among the gut microbial communities three months after birth [28], whereas most of our research subjects were less than one month of age.

In the current study, a detailed comparison between the gut microbiome composition in infants with or without BMJ was made by high-throughput sequencing. The percentage of Escherichia/Philligella in the gut microbiome of BMJ infants was significantly lower than that in the
control infants, implying that *Escherichia/Shigella* may be related to the development of BMJ.

**Acknowledgments:** This study was supported by grants from the National Nature Science Foundation of China (No. 81370744 and No. 81571483), Doctoral Degree Funding from the Chinese Ministry of Education (No. 20135031100009), funds from State Key Discipline Project (No. 2011-873), and grants from Clinical Research Foundation of Children’s Hospital of Chongqing Medical University (No. [2014]254cyj2014-11).

**Conflict of interest:** Authors state no conflict of interest

**References**


