Analysis of the relationship between bile duct and duodenal microbiota reveals that dysbacteriosis is the main cause of primary common bile duct stones

Zhitang Lyu

Hebei University College of Life Sciences

Tingting Yu

Second Hospital of Hebei Medical University

Lichao Zhang

Second Hospital of Hebei Medical University

Lina Wang

Hebei University College of Life Sciences

Jiahui Shi

Hebei University College of Life Sciences

Jihong Li

Second Hospital of Hebei Medical University

Zhirong Li

Second Hospital of Hebei Medical University

Wei Zhang

Second Hospital of Hebei Medical University

Senlin Hou (✉ tinge2019@126.com)

Second Hospital of Hebei Medical University

Suning Hou

Second Hospital of Hebei Medical University

Research article

Keywords: primary common bile duct stones, ERCP, dysbacteriosis, bile duct microbiota, duodenal microbiota, 16S rRNA gene high-throughput sequencing, short chain fatty acids.

Posted Date: November 5th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-100839/v1>

License: (cc) (i) This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background Bacteria play an important role in the formation of primary CBD stones. While the composition and function of the microbiota of bile duct in patients with primary CBD stones could not be explained in the whole. Intestinal dysbacteriosis is related to the formation of many diseases, but there are relatively few reports on the effects of intestinal microbial community on the formation of primary CBD stones, but none of these studies have a normal population of biliary tract and gut microbiota as control.

Result We use the 16S rRNA gene high-throughput sequencing technology to analyze the microbial diversity and community composition of biliary and duodenal microbiota in 15 patients with primary CBD stones (EG) and 4 patients without biliary tract disease (CK), and we divided above samples into four subgroups: B (EG bile), D (EG duodenal juice), BCK (CK bile) and DCK (CK duodenal juice). Alpha diversity analysis showed that only the abundance and evenness in B were lower than that in DCK ($P < 0.05$); Beta diversity analysis showed that there were some differences in the composition between biliary microbiota and the duodenal microbiota, but the abundance of the main groups showed similarities. Proteobacteria and Firmicutes were the dominant bacteria at phylum level accounted for at least 75% of the total reads in each subgroup. It is noteworthy that *Clostridium sensu stricto* (VIP > 1), Lachnospiraceae_UCG-008 (VIP > 1), *Butyrivibrio* (VIP > 1) and *Roseburia* (VIP > 1) which could produce short chain fatty acids (SCFA), were significantly decreased in B ($p < 0.05$) than that in BCK.

Conclusion According to the results, dysbacteriosis of biliary and duodenal microbiota plays an important role in the formation of primary CBD stones. The decrease in abundance of certain major acid-producing bacteria affects the health of the biliary tract and thus leads to the formation of stones.

Background

As a benign digestive disease, primary CBD stones are endangering human health with high morbidity and serious complications in China. The morbidity of primary CBD stones accounts for 5–30% of patients with cholelithiasis[1], and its main component is brown pigment stones in China. The pathogenesis of primary CBD stones is not clear yet. Therefore, the study of the causes and mechanisms of primary CBD stones has always been the focus and hotspot in order to prevent and treat it better. Among the many risk factors, bacteria are closely related to the formation of primary CBD stones, and it had long been confirmed that bacteria are present in bile with CBD stones[2]. Bacterial products such as β -glucuronidase, phospholipase, bacterial mucus and biofilm play an indispensable role in the formation of primary CBD stones, and many related bacteria have been identified in patients' bile, such as *Escherichia coli*, *Enterococcus*, *Klebsiella* and so on[3]. In addition, the intestinal microflora has been considered as the main source of biliary bacteria, and mainly from the upper digestive tract[4]. Intestinal dysbacteriosis is associated with the development of various diseases such as metabolic syndrome, Crohn's disease, colon cancer, and Cholecystolithiasis[5–8]. We assumed that, the formation of CBD stones may be closely related to the intestinal dysbacteriosis.

The previous studies could not give an overall explanation of the composition and function of the biliary microbiota, especially the normal biliary microbiota, due to technical reasons. What's the relationship between biliary and gut microbiota and how do they affect the formation of primary CBD stones have rarely been reported. With the development of science and technology, high-throughput sequencing based on 16S rRNA gene PCR amplification has been widely used in microbial research because of its low cost, high throughput, high sensitivity, and rapid sequencing[9].

Based on the 16S rRNA gene high throughput sequencing technique, the bile and duodenal juice of the patients with primary CBD stones and CK in our hospital are sequenced using the Illumina Miseq sequencing platform, the composition and diversity of the microbiota were analyzed, and the differences between the bile duct bacteria and intestinal bacteria were compared, to explore its significance for the formation of primary CBD stones.

Results

Through the metagenomic DNA extraction, 16S rRNA gene amplification and amplicon sequencing of the 38 biliary bile and duodenal juice samples from 19 persons including 15 patients and 4 controls, a total of 1,811,427 good reads were obtained and 2,801 OTUs were annotated. Sequencing data statistics and diversity analysis each sample yielded 47669.13 sequences (range = 33338 ~ 76885, SD = 10186.88) on average.

The good reads coverage of each group was above 99%, and there was no significant difference in the amount of sequence and OTUs at phylum, genus and species among the four subgroups, namely EG group bile (B subgroup) and duodenal juice (D subgroup) and CK group bile (BCK subgroup) and duodenal juice (DCK subgroup).

The Species accumulation curves[10] for the total number of OTUs corresponding to each sample were drawn. It showed when the sample size increases, the slope of the curve gradually decreases and tends to flatten. At which point the total number of OTUs in the community will no longer increase significantly with the addition of new samples. It indicates that the current sample size was enough to reflect the richness of the community.

We applied the diversity index (Chao1, ACE, Simpson and Shannon) to reflect its alpha diversity. The Shannon curve in each sample reached plateau phase, which confirmed the sequence numbers is large enough to cover most of the microbial information in the sample, and abundance and evenness of the microbial species in the current study meet the requirements for sequencing and analysis. Although the Simpson, Chao 1 and ACE index demonstrated there was no striking difference in the microbial community abundance between groups, the Shannon index showed significant difference between B and DCK group ($p < 0.05$). Shannon index is more sensitive to microbiota richness and rare OTUs. Thus, it indicated that the microbiota richness in B group is lower than that in DCK, and the difference is mainly in the content of the rare species (Fig. 1).

Beta diversity analyses were used by employing Principal component analysis (PCA) and Multidimensional scaling (MDS) of the similarity of microbial community structure in different groups. The PCA [11] result was shown in Fig. 2, partial samples of each group gathered on the left side of the longitudinal axis (dotted line). The BCK group and DCK group were distributed on the horizontal axis (dotted line) and gathered together. While the samples of the B subgroup and D subgroup showed more discrete and distributed in different regions. Mapping data were obtained from UniFrac PCoA master coordinate analysis via QIIME [12, 13]. Based on the Unweighted UniFrac distance, almost all bile samples were on the left side of the vertical axis and the duodenal fluid samples were on the right side (Fig. 3.a), and Non-metric Multidimensional scaling (NMDS) analysis based on Unweighted UniFrac distance obtained similar mapping data (Fig. 3.c). On the contrary, based on the Weighted UniFrac distance, the distribution of all samples is more concentrated and some of them are clustered and overlapped (Fig. 3.b), and NMDS analysis based on Weighted UniFrac distance also showed similar results (Fig. 3.d).

We presented the proportion of shared and unique OTUs in each subgroup by using the Venn diagram visually (Fig. 4). There were 1205 shared OTUs among the four subgroups, occupying 43% of all OTUs (total 2801 OTUs). More than 70% OTUs were shared in B subgroup and D subgroup, but only about 60% shared in BCK subgroup and DCK subgroup. The most dominant shared phylum and genus were Proteobacteria (95%) and *Pseudomonas* (83%), respectively. In the Venn diagram, we could also observe the unique OTUs in each subgroup. The unique OTUs of genus in the B subgroup were mostly *Pseudomonas*, which corresponding species were unidentified. Only 6 unique OTUs of genus were observed in BCK subgroup: *Dechlorobacter*, *Halomonas*, *Pseudomonas* and other unidentified genera.

Analyses the microbial community structure in different groups

According to the OTU division and classification status, the specific microbial composition of samples at each classification level can be obtained. A total of 23 phyla were identified in this work. At each sample or group, Proteobacteria and Firmicutes accounted for at least 75% of all phyla (Fig. 5). The number of bacteria in B subgroup was less than that of the D subgroup and DCK subgroup at the phylum, genus and species level (Table 1).

Table 1
Comparison of bacterial number of different groups at the phylum, genus and species level.

	B	D	BCK	DCK
Phylum [#]	8.20 ± 1.94	11.73 ± 3.63	9.75 ± 2.36	13.25 ± 0.50
Genus [#]	52.40 ± 12.24	97.20 ± 24.43	66.50 ± 8.10	106.50 ± 20.04
Species [#]	26.07 ± 5.60	45.00 ± 11.46	29.50 ± 5.32	43.75 ± 9.67
Analysis by Kruskal-Wallis test.# means P < 0.05.				

1 Comparison of the microbial community structure in B subgroup and BCK subgroup

At the Phylum level, a total of 18 phyla were identified in the two groups. Proteobacteria was the most abundant phylum, accounting for more than 90% in each group.

Metastats(<http://metastats.cbcb.umd.edu/>) was utilized to analysis the microbial community difference between groups[14].Four phyla (Actinobacteria, Saccharibacteria, Deinococcus-Thermus and SR1) were discovered to have significant difference between B subgroup and BCK subgroup (Fig. 6.a).

A total of 232 genera were identified in the two groups, of which the most dominant genus *Pseudomonas* and sub-dominant *Escherichia-Shigella* accounted for more than 80% of the two groups, both of which belonged to the Gammaproteobacteria of Proteobacteria. Among 232 genera, *Actinomyces*, *Proteus*, *Clostridium sensu stricto*, *Klebsiella*, *Actinobacillus*, *Lachnospiraceae_UCG-008*, *Butyrivibrio*, *Roseburia*, *Porphyromonas*, *Streptococcus*, *Helicobacter*, *Prevotella*₁, *7*, *Enterobacter* et al. showed significant difference with $p < 0.05$. It is notable that except for *Brevundimonas* and *Prevotella*₁, the abundance of all the different genera in B subgroup was lower than that in BCK subgroup (Fig. 6.b).

2 Comparison of the microbial community structure in D subgroup and DCK subgroup

A total of 22 phyla were identified in the two groups, of which Proteobacteria was the most dominant phylum, accounting for more than 70% of OTUs in both groups, and Firmicutes was the sub-dominant phylum. No phylum showed significant difference.

A total of 291 genera were identified in the two groups, and the dominant genera were *Pseudomonas*, *Escherichia-Shigella* and *Streptococcus* in sequence. 14 genera were significant different with $p < 0.05$ between two groups. The abundance of *Escherichia-Shigella*, *Geobacillus*, *Stomatobaculum*, *Campylobacter* and *Solibacillus* in D subgroup was significantly increased, the remaining 9 genera such as *Ochrobactrum*, *Moraxella*, etc. increased significantly in DCK subgroup (Fig. 6.c).

3 Comparison of the microbial community structure in B subgroup and D subgroup

Similarly, Proteobacteria and Firmicutes were still the most abundant phyla. The bacterial phyla in bile were contained in the duodenal fluid, however the duodenal fluid additionally contained Nitrospirae and

Gracilibacteria. Eight phyla were significant different with $p < 0.05$ between B subgroup and D subgroup. In addition to the abundance of Proteobacteria and Cyanobacteria increased, the others decreased significantly in bile (Fig. 6.d).

There were 295 genera were identified in the two groups, of which *Pseudomonas*, *Escherichia-Shigella* and *Clostridium sensu stricto_1* were the dominant genera in sequence. There are 15 unique genera in the bile, but there are up to 40 unique genera in the duodenal fluid. Even more striking is that 69 genera were significant different with $p < 0.05$ between two groups (Fig. 6.e). Among 69 genera, the abundance of *Pseudomonas* was significantly increased in the bile, others such as *Streptococcus*, *Veillonella*, *Lactobacillus*, *Prevotella_7* \ *Prevotella_6*, etc. increased significantly in duodenal fluid.

4 Comparison of the microbial community structure in BCK subgroup and DCK subgroup

A total of 18 phyla were identified in the two groups, of which Proteobacteria, Firmicutes and Actinobacteria were the dominant phyla in sequence. Only phylum Bacteroidetes was found obvious difference between bile and duodenal fluid in CK, of which the abundance was significantly increased in the duodenal fluid ($p < 0.05$) (Fig. 6.f).

At the genus level, we identified 234 genera from the two groups. Similarly, *Pseudomonas* and *Escherichia-Shigella* in the phylum Proteobacteria are in turn the two most dominant. 17 genera were found obvious significant difference between two groups (Fig. 6.g), of which only the abundance of *Aquabacterium* and *Enhydrobacter* was significantly increased in the duodenal fluid.

Key species screening and association analysis

1 Partial Least Squares Discriminant Analysis (PLS-DA)

The PLS-DA discriminant model was constructed based on the species abundance matrix and sample grouping data. Just like Fig. 7 showed, both bile samples and duodenal fluid samples received relatively better grouping models. At the same time, at the Genus level, we calculated the VIP (Variable importance in projection) coefficient for each genus ($VIP > 1$, the greater the value, the greater the contribution of the species to differences between groups). In bile samples, there were 75 genera that contributed to the bile grouping. Correspondingly, 87 contributing genera were discriminated for duodenal fluid grouping in the duodenal fluid.

2 Spearman Association Network Analysis of Dominant genera

We calculated the Spearman rank correlation coefficient between the dominant Genera, whose abundance was located in the top 50, and constructed the associated network for the related advantage where $\rho > 0.6$ and $P \text{ value} < 0.01$ [15]. The results showed (Fig. 8) that there were abundant correlated networks between EG and CK. Most of the bacterial genera existed in the same environment with collaborative manners. However the associations among bacterial genera in EG were more complicated. In CK, the relationship between the genera *Ralstonia* and *Ochrobactrum*, *Streptococcus* and

Granulicatella, *Acinetobacter* and *Alloprevotella* were antagonistic to each other (Fig. 8.a), but they showed cooperative relationships in EG (Fig. 8.b).

Microbial function prediction

In order to further understand the differences in the functions of biliary and intestinal microbiota, we performed functional prediction of the 16S rRNA gene sequence by PICRUSt in the KEGG database [16]. At KEGG level 3, we obtained a total of 304 pathways. 14 pathways were discovered to have significant difference in B subgroup and BCK subgroup ($p < 0.05$), and they were down-regulated in B subgroup than those in BCK subgroup. 8 pathways were significant different with $p < 0.05$ in duodenal fluid samples between D subgroup and DCK subgroup, including 6 up-regulated pathways and 2 down-regulated pathways.

The 6 major metabolic pathways, Metabolism, Genetic Information Processing, Environmental Information Processing, Cellular Processes, Organismal Systems and Human Diseases, all showed differences in some pathways at KEGG level 3 in bile samples between EG and CK. Microbiota in B subgroup had lower abundance of functions involved in Metabolism pathways such as Cysteine and methionine metabolism, Glycerophospholipid metabolism and Tetracycline biosynthesis than those in BCK subgroup. In addition, the abundance of cellular processes (Apoptosis and Endocytosis), environmental information processing (Two-component system and CAM ligands) and the immune system (RIG-I-like receptor signaling pathway) decreased significantly in B subgroup (Fig. 9.a).

In duodenal fluid samples (Fig. 9.b), microbiota in D subgroup had higher abundance of functions involved in Metabolism pathways (Purine metabolism, Metabolism of xenobiotics by cytochrome P450 and Naphthalene degradation) and human diseases (Bacterial invasion of epithelial cells, Measles). But the abundance of Digestive system (Pancreatic secretion) and nervous system (Glutamatergic synapse) decreased significantly in D subgroup.

Discussion

Due to differences in lifestyle, dietary structure, and living environment, primary CBD stones maintain a high incidence in China, which has brought about property damage and health hazards for our country and people. As an important condition for the occurrence of the disease, bacteria have been studied by scholars. Most previous experiments were based on Cholecystolithiasis or *in vitro* model to infer the relationship between the two. In addition, the intestinal microbiota plays a role in human diseases with its rich functions. There are a number of diseases such as Crohn's disease, colon cancer, and metabolic diseases that have been shown to be related to the intestinal flora [5–7]. The duodenal microbiota, as the nearest neighbor to the biliary tract, was considered to be the main source of possible biliary bacterial infections [17]. In this study, bile and duodenal fluid were directly obtained through ERCP techniques and diversity analysis was performed. Through the rigorous screening of the subjects, we obtained a total of 19 samples of patients. It is worth noting that 4 of them were pancreatic-related diseases patients

without hepatobiliary system disease. For the first time, the project used normal normal bile duct bile was used as CK to investigate the relationship between CBD and microbiota. The study was ethical and informed consents were obtained from the patients. To ensure that samples were not contaminated with each other was a difficult point in the sampling process. Thus, the duodenal fluid was first taken and then bile was extracted under the premise of strict aseptic processing to ensure the fidelity of this study.

In the field of microbial research, the widespread application of high-throughput sequencing technology has broken the limitations of traditional culture techniques. In our study, the 16S rRNA gene high-throughput sequencing technology was used for biliary and intestinal (duodenal) microbiota in patients with CBD stones and in patients without hepatobiliary system diseases. Through Illumina platform sequencing, a total of 1,811,427 effective sequencing sequences were obtained, and a total of 2801 bacterial OTUs were identified, The Goods coverage (%) of each group was above 99%, that is, the amount of sequencing of each sample is nearly saturated. In addition, the Species accumulation curves indicate that the current sample size was enough to reflect the richness of the community. The above two fully illustrate that this study has a very high reliability and reference value. 748,340 high-quality gene sequences were obtained in B subgroup, and 2062 bacterial OTUs were identified, which is similar to previous studies by Shen et al[18]. Interestingly, at the same time, we obtained a large number of effective high-quality sequences of 172,025 and 1597 OTUs in BCK subgroup (without hepatobiliary system disease). In the past, based on bacterial culture, it was found that there was no bacteria in normal bile which without biliary tract infections, because of its potent antimicrobial properties[19]. This study has broken through the limitations of this conclusion. Some scholars have also demonstrated the existence of bacteria in bile without common bile duct stones. For example, Pereira P et al. found the presence of bile bacterial communities in PSC patients[20]. Liu et al. found bacterial structures in acute liver failure and normal mice[21]. However, none of the patients enrolled in this study had the above-mentioned liver and biliary tract diseases, which provided valuable resources for the study of normal biliary microbiota.

The sequence quantity and the number of OTUs identified indicated that the bile duct bile and duodenal juice are rich in bacteria, regardless of whether the patients have hepatobiliary disease. Our result showed that the Alpha diversity of bile microbiota in the experiment group was lower than that in CK. But there was also study showing that the Alpha diversity of biliary microbiota was higher than that of normal intestinal flora and had a higher richness[8]. The factors for this difference may be related to the sites of the bile samples (gallbladder or common bile duct) and the sites of the intestinal microbiota samples (fecal or duodenal fluid). At the same time, we observed that in the same group, the Alpha diversity of bile and duodenal microbiota was similar, at least for their respective dominant bacterium.

We found that the most dominant phylum in bile and duodenal fluid was Proteobacteria in obtained 23 phyla. Other phyla with higher abundances included Firmicutes, Actinobacteria, Bacteroidetes, etc., this is similar to previous studies[4, 22]. The difference is that Synergistetes (0.63%) with a high abundance in B subgroup were not found in BCK subgroup. Synergistetes, the main pathogens of oral diseases such as periodontitis[23], may also participate in the formation of CBD stones. The analysis of the composition of the biliary and duodenal microbiota showed that phyla contained in the bile could be found in the

duodenal fluid almost. Similar result was also found in previous work of Ye et al.[4], their work showed the biliary microbiota had a comparatively higher similarity with the duodenal microbiota. However, the abundance of Proteobacteria and Cyanobacteria in the bile of the experiment group was significantly higher than that in the duodenal fluid. It shows that although the composition of biliary and duodenal bacterial community is similar, but the composition ratio is different.

At the genus level, the most dominant genus was *Pseudomonas* of the phylum Proteobacteria, followed by *Escherichia-Shigella*. This is different from previous research results, Ye et al. found three Enterobacteriaceae genera (*Escherichia*, *Klebsiella*, and an unclassified genus) and *Pyramidobacter* were abundant in bile of six gallstone patients, while Li et al. found *Acinetobacter* and *Prevotella* were dominant genera in the duodenum of nine healthy volunteers[4, 24]. According to Metastats and PLS-DA, *Pseudomonas* (VIP < 1) and *Escherichia-Shigella* (VIP < 1) were discovered to have significant difference in bile between EG and CK, and were not critical to bile grouping. In our view, the two most abundant genera may only exist as normal flora in bile, not as a risk factor for CBD stones. At the same time, *Escherichia-Shigella* (VIP > 1) was found to be more abundant in the duodenal fluid of EG than that of CK. It shows that *Escherichia-Shigella* plays an important role in the duodenum of patients with CBD stones. The bile or duodenal fluid in EG contained more genera than CK. The results showed that although the floras of the samples were relatively conservative at phylum level, there was dysbacteriosis and the imbalance of the composition ratio in the bile or duodenal fluid on genus level in EG. Meanwhile, the associations among bacterial genera in EG were more complicated by Spearman Association Network Analysis. It further proves that dysbacteriosis occurs in both the biliary tract and the duodenum in the relationship among bacterial genera.

Studies have shown that changes in the bacterial community of the duodenum were associated with various diseases such as small intestinal bacterial overgrowth (SIBO), irritable bowel syndrome (IBS) and celiac disease (CD) [25]. It was found that *Clostridium sensu stricto* (VIP > 1), Lachnospiraceae_UCG-008 (VIP > 1), *Butyrivibrio* (VIP > 1), *Roseburia* (VIP > 1), which can produce short chain fatty acids (SCFA), were significantly reduced in bile in EG compared with CK. SCFA has the functions of sterilization, bacteriostasis, and provide energy for host antibodies to improve immunity[26, 27]. SCFA not only affects the intestinal epithelium, but has a certain effect on airway epithelium, enteric nervous system (ENS), central nervous system (CNS), and peripheral neurons[28]. The occurrence of multiple diseases is associated with a decrease in the proportion of "good bacteria" that can produce SCFA in the intestine, such as inflammatory bowel disease[29], type 2 diabetes[30], Asthma[31] and so on. So we can conclude that the low expression of "good bacteria" in bile has a negative impact on biliary health in patients with common bile duct stones, which ultimately leads to the development of stones. In addition, *Klebsiella* (VIP > 1) and *Helicobacter* (VIP > 1) have been reported in previous studies to be associated with CBD stone formation. For example, *Klebsiella pneumoniae* and *Helicobacter pylori* have been cultured by bacteria or identified by PCR technique in bile from patients with CBD stones[18, 32]. Similar results were obtained in this study. But the two genera were more abundant in BCK subgroup. This phenomenon suggests that the disorder of the biliary bacteria community may be more important than the number of related pathogenic bacteria in the formation of CBD stones.

The comparison of biliary microbiota and duodenum microbiota indicated that gut microbiota showed a higher diversity of genera, both in EG and in CK. At the same time, the specificity and complexity of the two bacterial communities could be seen from as many as 60 differential genera between biliary microbiota and duodenum microbiota in EG. According to the Spearman Association Network Analysis of Dominant genera in EG and the normal group, we found that the relationship between *Ralstonia* and *Ochrobactrum*, *Streptococcus* and *Granulicatella*, *Acinetobacter* and *Alloprevotella* were antagonistic to each other in CK, but they showed cooperative relationships in EG. This shift in bacterial relationship may be the biological basis for the occurrence and development of biliary tract diseases (e.g. primary CBD stones). The reasons for this shift are not referenced in the relevant literature, may be related to the elimination or reduction of specific genus, or may be affected by other genera. These speculations need further research to explore.

The relationship between biliary and intestinal microbial communities has been a subject of continuous research and discussion. Bile bacterial infection was an important factor in the formation of CBD stones. Based on human anatomy and related studies, biliary microbiota in patients with CBD stones originated from the intestinal tract[33, 34]. In recent years, research has also tried to confirm this view. Ye et al. provided evidence that the biliary flora in patients with CBD stones originated in the intestine and mainly from the upper gastrointestinal tract by the 16S rRNA gene amplicon sequence[4]. In this study, the beta diversity analysis of the sample community structure showed that the biliary and the duodenum microbiota showed some similarities in both EG and CK, especially in terms of bacterial population abundance (slightly different in bacterial composition), moreover this similarity was more evident in the EG. At the same time, a Venn diagram showed "core microbiome" (1205 shared OTUs), which was first proposed by Turnbaugh et al[35]. Thus we can conclude that the most of biliary microbiota and the duodenal microbiota have a common source. More importantly, this study confirmed the existence of self-microbiota in the normal biliary tract, which constituted a microenvironment of bile together with bile salts, bile acids, and so on. In addition, though the abundance of unique OTUs was low in four subgroups, we still have the possibility to find biomarkers related CBD stones among them.

In order to gain more understanding of the biliary and duodenal microbiota, we performed microbial function prediction by PICTUS. Both the duodenal and the biliary microbiota have gained rich function pathways. There are differences in some pathways at KEGG level 3 in bile samples between EG and CK, such as cellular processes (Apoptosis, Endocytosis), environmental information processing (Two-component system, CAM ligands) and the immune system (RIG-I-like receptor signaling pathway), of which the abundance in EG decreased. We believe that some genera appeared to have a decline of self-repair ability and environmental adaptability in bile in EG and were in a "pathological state." In duodenal fluid samples, microbiota in EG had higher abundance of functions involved in Naphthalene degradation and Bacterial invasion of epithelial cells. Studies have shown that *Pseudomonas* sp. JLR11 and *E. coli* were related to pathway of Naphthalene degradation[36, 37]. Some bacteria such as *Shigella*, *Streptococcus* and *Listeria* could enter the epithelial cells through pathway of Bacterial invasion of epithelial cells[38–40]. It is indicated that the increase of bacteria with invasive ability of duodenal microbiota may indirectly lead to the formation of the stone, which may also explain the increase in the

abundance of *Escherichia-Shigella* in the duodenal fluid of EG. However, in the Infectious Diseases pathway, Shigellosis pathways and pathogenic *Escherichia coli* infection pathways had no difference in duodenal microbiota between two groups. This is also a good illustration of why the high abundance of *Escherichia-Shigella* in duodenal flora does not cause related diseases.

This study further explored the composition of the biliary and intestinal microbiota in patients with primary CBD stones and the differences in key genera. And we had a completely new understanding of the biological communities in the normal biliary tract. However, this experiment still has its limitations. The small sample size may lead to certain deviations. As far as research methods are concerned, in the future, metagenomic sequencing, animal experiments and other methods may be needed to study and validate the biliary flora, so as to further clarify the role of bacteria on the formation of primary CBD stones.

Conclusion

In conclusion, our results suggest the existence of the intrinsic flora in normal bile and give a more comprehensive disclosure of its composition. The disorder of the biliary flora is of great importance in the formation of CBD stones, which is mainly reflected in the reduction of some dominant bacterium and the change of the association network of dominant genera. These are new understandings of the biliary microbiota and pathogenic mechanisms. Equally important is the discovery of reduction of bacteria which could produce short chain fatty acids (SCFA), the decrease may affect the health of the biliary tract and lead to the formation of stones. Maybe this discovery will guide us in the prevention and treatment of primary CBD stones in the future.

Methods

Study Design and Sample Collection

We selected the inpatients at the Second Hospital of Hebei Medical University, through clinical symptoms, signs, B-ultrasonography, abdominal CT and other imaging examinations, and finally confirmed by ERCP as primary CBD stones patients (EG) without biliary tract infections. In EG, there were 15 cases, including 10 males and 5 females, aged 68.13 ± 13.53 years, and 4 patients without hepatobiliary diseases (CK), all were male and aged 39.25 ± 12.09 years old. Inclusion criteria were as follows: (1) Patients with related clinical manifestations of choledocholithiasis and diagnosed according to the relevant diagnostic methods, or without any hepatobiliary diseases. (2) No major diseases such as severe heart, brain, kidney, etc. (3) Good condition, can withstand ERCP. (4) Express understanding and support for this study and sign informed consent. Exclusion criteria: (1) antibiotics or probiotics were applied within 3 months before operation. (2) Those who combine cholecystolithiasis. (3) Patients with common bile duct obstruction caused by malignant tumors.

The bile (B group and BCK group) and duodenal juice (D group and DCK group) were respectively extracted from two groups of patients under ERCP.

The entire ERCP process was strictly aseptic and all experimental materials were sterilized according to the procedure. During the entry of the duodenoscope (TJF240/JF260V; Olympus), any inhalation was strictly prohibited to avoid contamination. Before duodenal fluid was drawn, the duodenoscope channel was washed with 10 ml acidified water ($\text{PH} = 2.5 \pm 0.2$) and 20 ml physiological saline at the level of the duodenum, and used balloon extraction catheter to draw at least 2 ml of intestinal fluid within 3 cm of the duodenal papilla. Then, the channels and papilla were flushed again with 20 ml of normal saline, and 5–10 ml of bile was extracted from the common bile duct using a Tripe Lumen Sphincterotome. The samples were placed in sterile sputum cup and immediately stored in the -80°C refrigerator for later bacterial DNA extraction and high throughput sequencing analysis.

All patients in this study expressed their understanding and signed informed consent and obtained approval from the Ethics Committee of the Second Hospital of Hebei Medical University.

16S rRNA gene amplification and sequencing of amplicons

The genomic DNA of bile and duodenal fluid samples were extracted by means of Soil DNA Kit (Omega). The obtained DNA was subjected to 0.8% agarose gel electrophoresis for molecular size determination, and the DNA was quantified by ultraviolet spectrophotometer. The DNA is stored in the -20°C refrigerator for subsequent analysis after passing the test. In order to ensure the quality of sequencing, the insertion fragment range of the best sequencing was 200–450 bp by Miseq sequence reading length. In this experiment, the V3 ~ V4 region of 16SrRNA gene was selected for sequencing. The PCR primers used were 515F (5'-barcode-GTGCCAGCMGCCGCGG-3') and 907R (5'-CCGTCAATTCMTTTRAGTTT-3'). Conventional PCR thermal cycling program: Initial denaturation at 98°C for 30 s, Denaturation at 98°C for 15 s, Annealing at 50°C for 30 s, Extension at 72°C for 30 s, The above three processes were carried out for 25 cycles and lastly extended for 5 min at 72°C , finally stored at $4-10^{\circ}\text{C}$ conventionally. The amplification results were performed on 2% agarose gel electrophoresis. After passing the test, the target fragment was excised and the target fragment was recovered using the Axygen Gel Recovery Kit. The PCR products were quantified on a Microplate reader (BioTek, FLx800) using the Quant-iTPicoGreen dsDNA Assay Kit, and then mixed according to the amount of data required for each sample. The obtained PCR products above were used to construct a sequencing library by using Illumina trusEq. The Qualified sequencing library (index sequence cannot be repeated) was diluted according to the gradient, then mixed according to the required amount of sequencing, and denatured by NaOH into single strands for on-machine sequencing. Barcoded V3-V4 amplicons were sequenced using the pair-end method by Illumina Miseq[41]. The optimal sequencing length was chosen to be 200–450 bp. Using FLASH software[42] to pair pair-end sequences based on the initial screening according to the overlapping bases: the overlapped bases of both Read 1 and Read 2 sequences are required to be ≥ 10 bp in length, and base mismatches are not allowed. Afterwards, the valid sequences were obtained according to the Index information corresponding to each sample. Finally, use QIIME software (Quantitative Insights Into

Microbial Ecology, v1.8.0)[43] to identify the error sequence and remove it. Amplification and sequencing of 16S rRNA gene was completed by Personal Biotechnology Co., Ltd. (Shanghai, China).

OTU clustering and statistical analysis

Operational Taxonomic Unit(OTU)[44]usually refers to a set of human-defined sequence similarity thresholds. The sequences from one or more samples are merged, and the sequences whose similarity is higher than the threshold will be merged into one OTU. Using the UCLUST sequence comparison tool in QIIME software[45], the sequences obtained above were merged and OTU-divided with a sequence similarity of 97%, and selected the highest abundance sequence in each OTU as the representative sequence of the OTU. The OTU with an abundance value less than 0.001% of the all sample total sequencing was removed[46], then used Silva database as the template sequence for the identification of the OTU taxonomic status.

Statistical analyses were performed using the SPSS 23.0 software. The Fisher's exact test was used to enumeration data, and the Kruskal-wall is test to measurement data. The OTU was classified by QIIME software using the Silva database and the bacterial annotation was carried out. Differences within or between groups were compared by Alpha diversity, Beta diversity, Metastats analysis, PLS-DA analysis, and Spearman correlation network analysis. PICRUST was used to predict bacterial function through the KEGG database.

Abbreviations

CBD	Common Bile Duct
ERCP	Endoscopic retrograde cholangiopancreatography
EG group	Patients with primary CBD stones
CK group	Patients without hepatobiliary diseases
B group	Bile of patients with primary CBD stones
BCK group	Bile of without hepatobiliary diseases
D group	Duodenal juice of patients with primary CBD stones
DCK group	Duodenal juice of without hepatobiliary diseases
rRNA	Ribosomal Ribonucleic Acid
PCR	Polymerase chain reaction
SCFA	Short chain fatty acids
OTU	Operational Taxonomic Unit
PCA	Principal component analysis
PCoA	principal co-ordinates analysis
MDS	Multidimensional scaling
NMDS	Non-metric Multidimensional scaling
PLS-DA	Partial Least Squares Discriminant Analysis
VIP	Variable importance in projection
KEGG	Kyoto Encyclopedia of Genes and Genomes
PICRUST	Phylogenetic Investigation of Communities by Reconstruction of Unobserved States

Declarations

Compliance with ethical standards

Conflict of interest and Competing Interests The authors have no conflict of interest and competing Interests to declare.

Ethical approval This study was approved by the ethics committee of the second hospital of Hebei Medical University, Shijiazhuang, P. R. China.

Informed consent Informed consent was obtained from all individual participants for the information included in this article.

Author Contributions

The project idea was conceived and the project was designed by S.L.H.; Zh.T.L. managed the project. T.T.Y., L.C.Zh., J.H.L. and W.Zh. performed sample collection and clinical study. L.N.W., Zh.R.L., S.N.H. and T.T.Y. performed DNA experiments. Zh.T.L. and S.L.H designed the analysis. T.T.Y., L.N.W. and J.H.Sh. performed the data analysis. Zh.T.L and T.T.Y. wrote the paper. All authors have read and approved the manuscript.

Acknowledgements

This work was supported by Hebei Provincial Key Discipline of Bioengineering, Biology Discipline of Hebei Provincial Level with Distinctive Characteristic, as well as First-class Discipline Construction Project of Hebei Province. The funder, Zhitang Lyu, who managed the project was the first author in our study.

Availability of data and materials 16S rRNA sequence reads of 38 samples have been submitted to the NCBI SRA metadata under accession code SRP155620.

References

1. Lu J, Xiong XZ, Cheng Y, Lin YX, Zhou RX, You Z, Wu SJ, Cheng NS: **One-stage versus two-stage management for concomitant gallbladder stones and common bile duct stones in patients with obstructive jaundice.** *Am Surg* 2013, **79**(11):1142-1148.
2. Stewart L, Smith AL, Pellegrini CA, Motson RW, Way LW: **Pigment gallstones form as a composite of bacterial microcolonies and pigment solids.** *Ann Surg* 1987, **206**(3):242-250.
3. Fang-Peng YE, Shi ZC, Ying WX: **Distribution and drug susceptibility of pathogens cultured from common bile duct stones patients complicated with biliary tract infections.** *Chinese Journal of Nosocomiology* 2015.
4. Ye F, Shen H, Li Z, Meng F, Li L, Yang J, Chen Y, Bo X, Zhang X, Ni M: **Influence of the Biliary System on Biliary Bacteria Revealed by Bacterial Communities of the Human Biliary and Upper Digestive Tracts.** *PLoS One* 2016, **11**(3):e0150519.
5. Belcheva A, Irrazabal T, Robertson SJ, Streutker C, Girardin SE, Edelmann W, Guttman DS, Philpot DJ, Martin A: **Gut Microbiota and the Etiology of Colon Cancer.** *Environ Mol Mutagen* 2015, **56**(1-2):S46-S46.
6. Øyri SF, Múzes G, Sipos F: **Dysbiotic gut microbiome: A key element of Crohn's disease.** *Comparative Immunology Microbiology & Infectious Diseases* 2015, **43**:36-49.
7. Tilg H, Moschen AR: **Gut Microbiome, Obesity, and Metabolic Syndrome:** Springer International Publishing; 2016.
8. Tao W, Zhang Z, Liu B, Hou D, Yun L, Jie Z, Peng S: **Gut microbiota dysbiosis and bacterial community assembly associated with cholesterol gallstones in large-scale study.** *BMC Genomics* 2013, **14**(1):669.

9. Kenneth Nelson F, Snyder M, Gardner AF, Hendrickson CL, Shendure JA, Porreca GJ, Church GM, Ausubel FM, Ju J, Kieleczawa J: **Introduction and Historical Overview of DNA Sequencing**: John Wiley & Sons, Inc.; 2011.
10. Chao A, Shen TJ: **Nonparametric prediction in species sampling**. *Journal of Agricultural Biological & Environmental Statistics* 2004, **9**(3):253-269.
11. Ramette A: **Multivariate analyses in microbial ecology**. *FEMS Microbiol Ecol* 2007, **62**(2):142-160.
12. Lozupone C, Knight R: **UniFrac: a new phylogenetic method for comparing microbial communities**. *Appl Environ Microbiol* 2005, **71**(12):8228-8235.
13. Lozupone CA, Hamady M, Kelley ST, Knight R: **Quantitative and Qualitative β Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities**. *Appl Environ Microbiol* 2007, **73**(5):1576-1585.
14. White JR, Nagarajan N, Pop M: **Statistical methods for detecting differentially abundant features in clinical metagenomic samples**. *PLoS Comput Biol* 2009, **5**(4):e1000352.
15. Faust K, Raes J: **Microbial interactions: from networks to models**. *Nature Reviews Microbiology* 2012, **10**(8):538.
16. Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Dan K, Reyes JA, Clemente JC, Burkepile DE, Thurber RLV, Knight R: **Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences**. *Nat Biotechnol* 2013, **31**(9):814-821.
17. WANG OY: **Relationship between duodeno-biliary reflux and pigment gallstone formation**. *Chinese Journal of General Surgery* 2017, **26**(8):1057-1064.
18. Shen H, Ye F, Xie L, Yang J, Li Z, Xu P, Meng F, Li L, Chen Y, Bo X *et al*: **Metagenomic sequencing of bile from gallstone patients to identify different microbial community patterns and novel biliary bacteria**. *Sci Rep* 2015, **5**:17450.
19. Begley M, Gahan CG, Hill C: **The interaction between bacteria and bile**. *FEMS Microbiol Rev* 2010, **29**(4):625-651.
20. Pereira P, Aho V, Arola J, Boyd S, Jokelainen K, Paulin L, Auvinen P, Färkkilä M: **Bile microbiota in primary sclerosing cholangitis: Impact on disease progression and development of biliary dysplasia**. *PLoS One* 2017, **12**(8):e0182924.
21. Liu HJ, Yu L, Hu YX, Cheng ML: **[Microflora changes in upper digestive tract and bile in acute liver failure mice]**. *Zhonghua Yi Xue Za Zhi* 2017, **97**(21):1664-1667.
22. Stearns JC, Lynch MDJ, Senadheera DB, Tenenbaum HC, Goldberg MB, Cvitkovitch DG, Croitoru K, Morenoghagsieb G, Neufeld JD: **Bacterial biogeography of the human digestive tract**. *Sci Rep* 2011, **1**(2840):170.
23. Yu XL, Chan Y, Zhuang LF, Lai HC, Lang NP, Lacap-Bugler DC, Leung WK, Watt RM: **Distributions of Synergistetes in clinically-healthy and diseased periodontal and peri-implant niches**. *Microb Pathog* 2015, **94**:90-103.

24. Li G, Yang M, Zhou K, Zhang L, Tian L, Lv S, Jin Y, Qian W, Xiong H, Lin R *et al*: **Diversity of Duodenal and Rectal Microbiota in Biopsy Tissues and Luminal Contents in Healthy Volunteers.** *J Microbiol Biotechnol* 2015, **25**(7):1136-1145.
25. Wang ZK, Yang YS: **Upper gastrointestinal microbiota and digestive diseases.** *World J Gastroenterol* 2013, **19**(10):1541-1550.
26. Wan X, Wang X, Li N: **[Research progression of short chain fatty acid].** *Zhonghua Wei Chang Wai Ke Za Zhi* 2015, **18**(9):958-960.
27. Kim M, Qie Y, Park J, Chang HK: **Gut Microbial Metabolites Fuel Host Antibody Responses.** *Cell Host & Microbe* 2016, **20**(2):202.
28. Koh A, Vadder FD, Kovatcheva-Datchary P, Ckhd FB: **From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites.** *Cell* 2016, **165**(6):1332-1345.
29. Frank DN, Amand ALS, Feldman RA, Boedeker EC, Harpaz N, Pace NR: **Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases.** *Proc Natl Acad Sci U S A* 2007, **104**(34):13780.
30. Zhao L, Zhang F, Ding X, Wu G, Yan YL, Wang X, Fu H, Xue X, Lu C, Ma J: **Gut bacteria selectively promoted by dietary fibers alleviate type 2 diabetes.** *Science* 2018, **359**(6380):1151-1156.
31. Zaiss MM, Alexis R, Luc L, Kumar DL, Ilaria M, Kerstin S, Alessandra P, Laure M, Walker AW, Jacques R: **The Intestinal Microbiota Contributes to the Ability of Helminths to Modulate Allergic Inflammation.** *Immunity* 2015, **43**(5):998-1010.
32. Song M, Wu J, Wang P, Huang XD, Zhang H, Sun SB, Fan Y, Liao YS, Li H: **Detection of Helicobacter pylori DNA in bile and gallstone of patients with primary choledocholithiasis.** *World Chinese Journal of Digestology* 2010, **18**(35):3809-3812.
33. Lee DK, Tarr PI, Haigh WG, Lee SP: **Bacterial DNA in mixed cholesterol gallstones.** *Am J Gastroenterol* 1999, **94**(12):3502-3506.
34. Su Y, Wu S, Fan Y, Jin J, Zhang Z: **The preliminary experimental and clinical study of the relationship between the pigment gallstone and intestinal mucosal barrier.** *J Gastroenterol Hepatol* 2009, **24**(8):1451-1456.
35. Turnbaugh PJ, Ley RE, Hamady M, Fraserliggett C, Knight R, Gordon JI: **The human microbiome project: exploring the microbial part of ourselves in a changing world.** *Nature* 2007, **449**(7164):804.
36. Bai J, Yang J, Liu P, Yang Q: **Transformation pathway of 2,4,6-trinitrotoluene by Escherichia coli nitroreductases and improvement of activity using structure-based mutagenesis.** *Process Biochemistry* 2015, **50**(5):705-711.
37. Pascual J, Udaondo Z, Molina L, Segura A, Estevenúñez A, Caballero A, Duque E, Ramos JL, Van DP: **Draft Genome Sequence of Pseudomonas putida JLR11, a Facultative Anaerobic 2,4,6-Trinitrotoluene Biotransforming Bacterium.** *Genome Announc* 2015, **3**(5).
38. Iwai H, Kim M, Yoshikawa Y, Ashida H, Ogawa M, Fujita Y, Muller D, Kirikae T, Jackson PK, Kotani S: **A bacterial effector targets Mad2L2, an APC inhibitor, to modulate host cell cycling.** *Cell* 2007, **130**(4):611-623.

39. Wang B, Li S, Dedhar S, Cleary PP: **Paxillin phosphorylation: bifurcation point downstream of integrin-linked kinase (ILK) in streptococcal invasion.** *Cell Microbiol* 2007, **9**(6):1519–1528.
40. Mostowy S, Cossart P: **Cytoskeleton rearrangements during Listeria infection: Clathrin and septins as new players in the game.** *Cell Motil Cytoskeleton* 2009, **66**(10):816-823.
41. Derakhshani H, Tun HM, Khafipour E: **An extended single-index multiplexed 16S rRNA sequencing for microbial community analysis on MiSeq illumina platforms.** *J Basic Microbiol* 2016, **56**(3):321-326.
42. Mago T, Salzberg SL: **FLASH: fast length adjustment of short reads to improve genome assemblies.** *Bioinformatics* 2011, **27**(21):2957.
43. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI: **QIIME allows analysis of high-throughput community sequencing data.** *Nature Methods* 2010, **7**(5):335.
44. Blaxter M, Mann J, Chapman T, Thomas F, Whitton C, Floyd R, Abebe E: **Defining operational taxonomic units using DNA barcode data.** *Philosophical Transactions Biological Sciences* 2005, **360**(1462):1935-1943.
45. Edgar RC: **Search and clustering orders of magnitude faster than BLAST.** *Bioinformatics* 2010, **26**(19):2460.
46. Bokulich NA, Sathish S, Faith JJ, Dirk G, Gordon JI, Rob K, Mills DA, Gregory CJ: **Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing.** *Nature Methods* 2013, **10**(1):57.

Figures

Figure 1

Microbial alpha diversity with a box plot exhibiting the community diversity.

Figure 2

PCA analysis of the principal component of the similarity of microbial community structure in different groups.

a.

b.

c.

d.

Figure 3

Multidimensional scaling (MDS) of the similarity of microbial community structure in different groups. (a. Unweighted UniFracPCoA; b. weighted UniFracPCoA; c. Unweighted UniFrac NMDS; d. weighted UniFrac NMDS.)

Figure 4

A Venn diagram showing shared and unique OTUs among the four subgroups.

Figure 5

The microbial community structure in Phylum level in different samples

Figure 6

Comparison of microbial community structure at Phylum and Genus level in different groups by Metastats.

a.

b.

Figure 7

The PLS-DA discriminant model was constructed in bile group (a) and duodenal fluid group (b).

a.

b.

Figure 8

Network plots highlight correlations among bacterial genera in EG (a) and in CK (b). (Red lines refer to cooperative correlation and green lines refer to antagonistic correlation)

b.

Figure 9

Difference of functional pathway at KEGG level 3 in bile samples (a), duodenal fluid samples (b) between EG and CK.