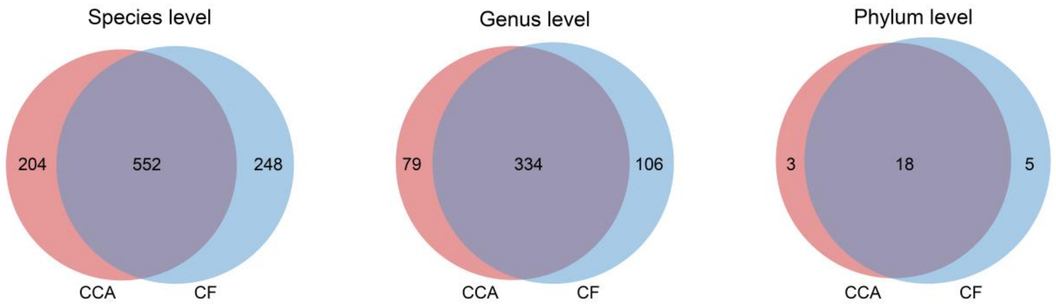

Figure S1 The respective numbers of phylotypes in CCA and CF. Venn diagrams showed the numbers of microbiota shared in CCA cohort (n=33) and CF (n=47), and the numbers of unique microbiota in either CCA (red) or CF (blue).

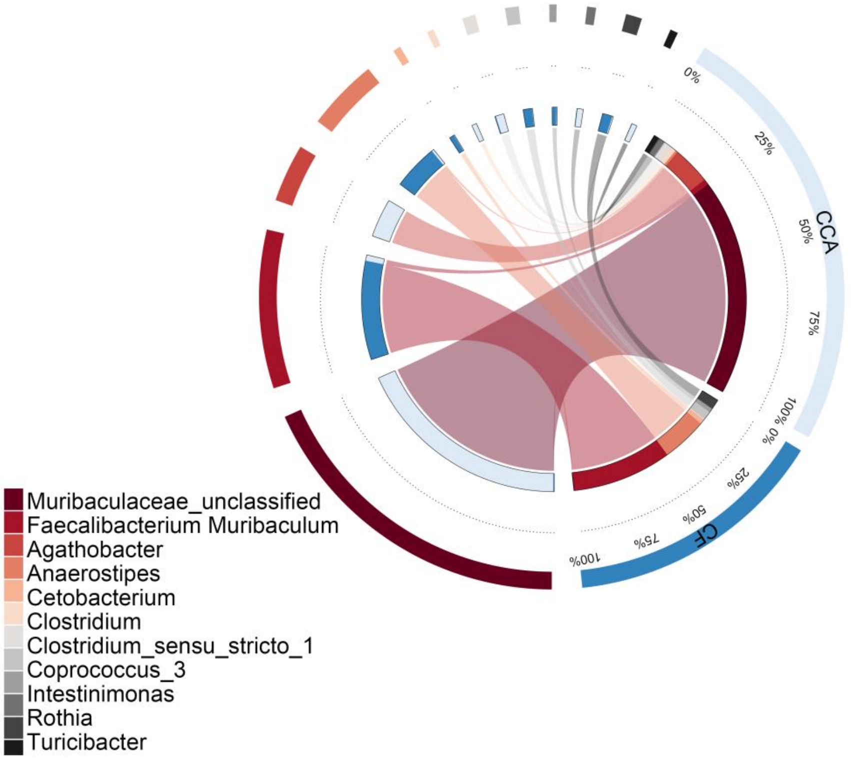

Figure S2 The distribution of 12 genera in CCA and CF. Circos diagram visualized the distribution of genera in CCA cohort (n=33) and CF (n=47).

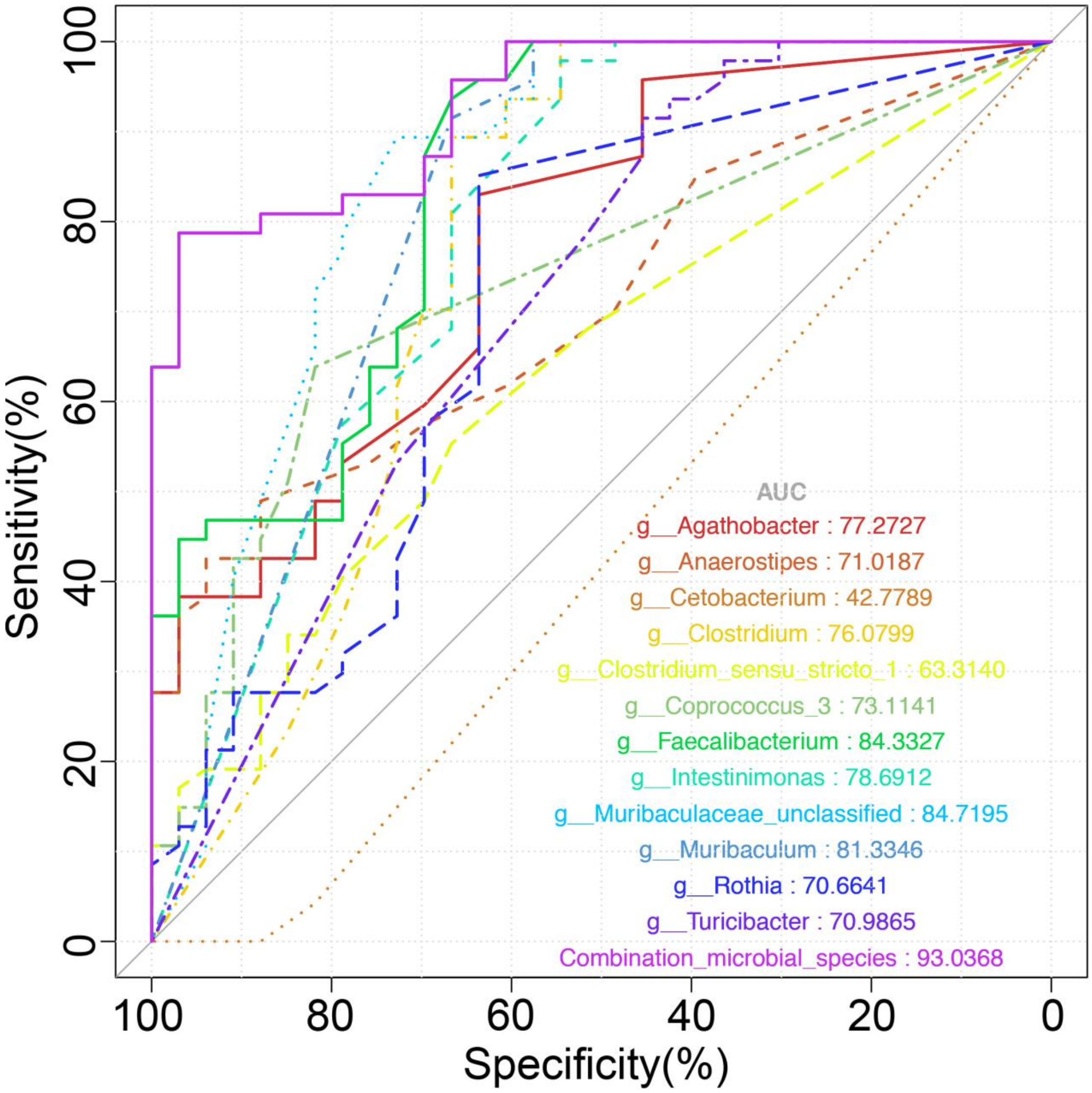

Figure S3 The respective AUC of 12 genera. With the combination of 12 genera and the respective genus, the effect of separating Cholangiocarcinoma from cancer free were evaluated by area under the ROC.

Table S1 Demographic characteristics of the cohorts

	CCA(n=33)	Caner-free(n=47)	P value
Demographics			
Age,years,mean±SD	67.3±8.8	52.8±13.6	<0.001^{owa}
Male gender, n (%)	17(51.5)	22(46.8)	0.678^{CS}
Smoking history, n (%)	5(15.1)	7(14.8)	0.975 ^{CS}
Drinking history, n (%)	8(24.2)	6(12.7)	0.184 ^{CS}
BMI,kg/m2, mean±SD	21.3±2.8	23.5±2.6	0.001^{owa}
hepatic disease history,n,(%)			
Cirrhosis	3(9.0)	0(0.0)	0.066 ^{F}
Fatty liver	4(12.1)	21(44.6)	0.004^{CS}
HBV-infected	5(15.1)	7(14.8)	0.975^{CS}
Tumor marker,meidan(P25,P75)			
AFP,µg/L	2.8(2,4.3)	—	—
CEA,ng/ml	4.1(2.5,20.1)	—	—
CA-199,µg/L	237.7(25,3842.3)	—	—
Hepatic function,meidan(P25,P75)			
ALT,U/L	60(18,150)	26(16,53)	0.042 ^{MW}
AST,U/L	57(23,135)	28(18,44)	0.003 ^{MW}
AKP,U/L	269(120.5,575.5)	88(70,103)	<0.001^{MW}
GGT,U/L	222(74,409)	34(21,85)	<0.001^{MW}
TB, µmol/L	27(12,145)	11(8,17)	<0.001^{MW}
ALB, g/L,(mean±SD)	34.3±4.5	40.3±3.9	<0.001^{OWA}
Child-Pugh,grade	6(5.5,7)	5(5,5)	<0.001^{MW}
Neoplasm staging			
TMN,stageI-II,n (%)	15(45.4)	—	—

Clinical characteristics of CCA and CF cohort.

Variables are significantly different between groups if $p < 0.05$ using OWA=one-way ANOVA test, CS = Chi-square test, F = Fisher’s exact test, MW = Mann Whitney test or KW= Kruskal-Wallis test.

Abbreviations: BMI, body mass index; AFP, alpha fetoprotein; CEA, carcinoembryonic antigen; CA19-9, Carbohydrate antigen199; ALT, alanine aminotransferase; AST, aspartate aminotransferase; AKP, alkaline phosphatase; GGT, gamma-glutamyltransferase; TB, total bilirubin; ALB, albumin; TMN, Tumor Node Metastasis; BCLC, Barcelona Clinic Liver Cancer.

Table S2 Differentially expressed bacteria in two cohort according to nonparametric tests

	Classification	log2FC	wilcox.test.p_value	regulation	mean	mean_CCA	mean_CF
1	g__Muribaculaceae_unclassified	7.93	0	up	5.88	14.16	0.06
2	g__Lachnospiraceae_NK4A136_group	7.97	0	up	2.86	6.88	0.03
3	g__Muribaculum	10.27	0	up	1.33	3.21	0.00
4	g__Ruminococcaceae_UCG-014	2.67	0	up	1.22	2.42	0.38
5	g__Alistipes	4.47	0	up	1.18	2.68	0.12
6	g__Prevotellaceae_UCG-001	9.48	0	up	0.98	2.36	0.00
7	g__Alloprevotella	7.20	0	up	0.45	1.08	0.01
8	g__Clostridium	3.43	0	up	0.31	0.67	0.06
9	g__Prevotellaceae_NK3B31_group	8.94	0	up	0.35	0.83	0.00
10	g__Ruminiclostridium	9.20	0	up	0.22	0.53	0.00
11	g__Ruminiclostridium_9	4.01	0	up	0.23	0.51	0.03
12	g__Intestinimonas	4.76	0	up	0.21	0.48	0.02
13	g__Candidatus_Saccharimonas	4.49	0	up	0.05	0.11	0.01
14	g__Eubacterium]_xylanophilum_group	6.73	0	up	0.16	0.38	0.00
15	g__Bilophila	3.31	0	up	0.18	0.38	0.04
16	g__Rikenellaceae_RC9_gut_group	5.07	0	up	0.08	0.18	0.01
17	g__Gastranaerophilales_unclassified	11.31	0	up	0.11	0.25	0.00
18	g__Anaerotignum	7.87	0	up	0.10	0.23	0.00
19	g__Clostridiales_Family_XIV._Incertae_Sedis_unclassified	8.41	0	up	0.17	0.41	0.00
20	g__Tyzzerella	6.41	0	up	0.11	0.26	0.00
21	g__GCA-900066575	4.20	0	up	0.04	0.09	0.00
22	g__Anaerotruncus	4.92	0	up	0.11	0.25	0.01
23	g__Acetatifactor	9.60	0	up	0.10	0.23	0.00
24	g__unclassified	5.34	0	up	0.03	0.08	0.00
25	g__Clostridiales_vadinBB60_group_unclassified	5.28	0	up	0.09	0.21	0.01
26	g__Lachnospiraceae_UCG-001	3.03	0	up	0.10	0.21	0.03
27	g__Ruminiclostridium_6	3.59	0	up	0.07	0.15	0.01
28	g__Kineothrix	10.49	0	up	0.06	0.14	0.00
29	g__Enterorhabdus	5.05	0	up	0.04	0.10	0.00
30	g__Mucispirillum	7.90	0	up	0.05	0.12	0.00

31	g__Prevotella	3.61	0	up	0.02	0.05	0.00
32	g__Coprococcus	5.27	0	up	0.02	0.04	0.00
33	g__Ruminococcaceae_UCG-009	6.37	0	up	0.03	0.08	0.00
34	g__Desulfovibrionaceae_unclassified	7.40	0	up	0.02	0.05	0.00
35	g__A2	9.10	0	up	0.02	0.05	0.00
36	g__Harryflintia	5.14	0	up	0.02	0.04	0.00
37	g__GCA-900066225	2.27	0	up	0.02	0.03	0.01
38	g__Peptococcus	2.63	0	up	0.01	0.02	0.00
39	g__Defluviitaleaceae_UCG-011	3.01	0	up	0.01	0.03	0.00
40	g__Fournierella	6.32	0	up	0.02	0.04	0.00
41	g__Thalassospira	7.53	0	up	0.01	0.02	0.00
42	g__Family_XIII_UCG-001	2.34	0	up	0.01	0.02	0.00
43	g__Erysipelotrichaceae_unclassified	2.14	0	up	0.01	0.01	0.00
44	g__Dehalobacterium	4.79	0	up	0.00	0.01	0.00
45	g__Oxalobacter	4.70	0	up	0.00	0.01	0.00
46	g__Clostridiales_unclassified	5.68	0.00	up	1.13	2.66	0.05
47	g__Odoribacter	4.41	0.00	up	0.14	0.31	0.01
48	g__Duncaniella	6.16	0.00	up	0.09	0.22	0.00
49	g__Peptococcaceae_unclassified	5.47	0.00	up	0.06	0.13	0.00
50	g__Candidatus_Soleaferrea	2.52	0.00	up	0.04	0.07	0.01
51	g__Marvinbryantia	3.96	0.00	up	0.02	0.05	0.00
52	g__Robinsoniella	5.96	0.00	up	0.01	0.02	0.00
53	g__Ruminococcaceae_UCG-010	2.01	0.00	up	0.02	0.03	0.01
54	g__Papillibacter	2.32	0.00	up	0.00	0.01	0.00
55	g__Turicibacter	4.89	0.00	up	0.18	0.42	0.01
56	g__Murimonas	6.57	0.00	up	0.02	0.04	0.00
57	g__Bacteroidetes_unclassified	2.45	0.00	up	0.01	0.02	0.00
58	g__Eubacterium]_oxidoreducens_group	3.26	0.00	up	0.00	0.00	0.00
59	g__Butyricimonas	2.09	0.00	up	0.04	0.07	0.02
60	g__Oscillibacter	3.38	0.00	up	0.21	0.46	0.04
61	g__Lactococcus	2.14	0.00	up	0.05	0.08	0.02
62	g__Pseudoflavonifractor	2.38	0.01	up	0.00	0.01	0.00
63	g__Ruminococcaceae_UCG-007	4.46	0.01	up	0.00	0.00	0.00
64	g__Atopobiaceae_unclassified	2.16	0.01	up	0.01	0.02	0.00
65	g__Eubacterium]_ventriosum_group	2.57	0.01	up	0.27	0.54	0.09
66	g__Catabacter	2.47	0.01	up	0.00	0.01	0.00
67	g__Cetobacterium	10.33	0.02	up	0.16	0.39	0.00

68	g__Stenotrophomonas	3.91	0.03	up	0.00	0.00	0.00
69	g__Acinetobacter	5.14	0.04	up	0.01	0.02	0.00
70	g__Curvibacter	3.54	0.05	up	0.00	0.00	0.00
71	g__Faecalibacterium	-4.14	0	down	5.08	0.47	8.32
72	g__Erysipelotrichaceae_UCG-003	-2.43	0	down	1.40	0.39	2.11
73	g__Agathobacter	-4.41	0	down	2.31	0.18	3.81
74	g__Coprococcus_3	-2.63	0.00	down	0.22	0.06	0.34
75	g__Intestinibacter	-3.70	0.00	down	0.03	0.00	0.05
76	g__Faecalibaculum	-4.61	0.00	down	0.00	0.00	0.00
77	g__Rothia	-3.41	0.00	down	0.51	0.08	0.82
78	g__Anaerostipes	-3.41	0.00	down	0.25	0.04	0.40
79	g__Dialister	-3.28	0.00	down	0.27	0.04	0.43
80	g__Gemella	-2.70	0.00	down	0.03	0.01	0.05
81	g__F0332	-3.81	0.01	down	0.00	0.00	0.00
82	g__Desulfovibrio	-2.68	0.01	down	0.05	0.01	0.08
83	g__Ruminococcaceae_UCG-013	-3.06	0.01	down	0.49	0.09	0.78
84	g__Clostridium_sensu_stricto_8	-2.17	0.01	down	0.00	0.00	0.00
85	g__Candidatus_Microthrix	-2.58	0.01	down	0.00	0.00	0.00
86	g__Peptostreptococcus	-2.50	0.01	down	0.02	0.00	0.02
87	g__Paraprevotella	-5.30	0.02	down	0.02	0.00	0.04
88	g__Mitochondria_unclassified	-2.17	0.02	down	0.00	0.00	0.00
89	g__Holdemanella	-2.56	0.02	down	0.49	0.13	0.74
90	g__Clostridium_sensu_stricto_1	-4.54	0.02	down	0.44	0.03	0.73
91	g__Burkholderia-Caballeronia-Paraburkholderia	-2.70	0.02	down	0.00	0.00	0.00
92	g__Saccharimonadales_unclassified	-2.70	0.02	down	0.00	0.00	0.00
93	g__Rhizobacter	-3.64	0.03	down	0.01	0.00	0.02
94	g__Lachnospiraceae_UCG-004	-7.17	0.03	down	0.01	0.00	0.01
95	g__Corynebacterium_1	-2.32	0.03	down	0.00	0.00	0.00
96	g__Devosia	-2.46	0.03	down	0.00	0.00	0.00
97	g__Allisonella	-3.31	0.04	down	0.01	0.00	0.02
98	g__Romboutsia	-3.31	0.04	down	0.39	0.06	0.62

	Phylum	log2FC	wilcox.test.p_value	regulation	mean	mean_CCA	mean_CF
99	p__Bacteroidetes	3.10	0	up	13.72	28.50	3.33
100	p__unclassified	5.34	0	up	0.03	0.08	0.00
101	p__Deferribacteres	7.90	0	up	0.05	0.12	0.00
102	p__Cyanobacteria	4.94	0.00	up	0.11	0.26	0.01
103	p__Patescibacteria	2.08	0.00	up	0.07	0.12	0.03

No 1-98: Differentially expressed bacteria
No 99-103: Differentially expressed bacterial phyla

Table S3 Co-differentially expressed bacteria in two cohort according to nonparametric tests and LEFSe

Genus	Nonparametric tests		LEFSe	
	log2FC	wilcox.test.p_value	LDA_values	p_value
g__Muribaculaceae_unclassified	7.93	0	4.86	0.00
g__Lachnospiraceae_NK4A136_group	7.97	0	4.55	0.00
g__Faecalibacterium	-4.14	0	4.56	0.00
g__Muribaculum	10.27	0	4.22	0.00
g__Agathobacter	-4.41	0	4.31	0.00
g__Alistipes	4.47	0	4.10	0.00
g__Alloprevotella	7.2	0	3.75	0.00
g__Clostridium	3.43	0	3.61	0.02
g__Prevotellaceae_NK3B31_group	8.94	0	3.60	0.00
g__Ruminiclostridium	9.2	0	3.45	0.00
g__Ruminiclostridium_9	4.01	0	3.40	0.00
g__Intestinimonas	4.76	0	3.39	0.00
g__Bilophila	3.31	0	3.28	0.00
g__Gastranaerophilales_unclassified	11.31	0	3.12	0.00
g__Anaerotignum	7.87	0	3.09	0.00
g__Tyzzerella	6.41	0	3.16	0.00
g__Anaerotruncus	4.92	0	3.11	0.00
g__Acetatifactor	9.6	0	3.09	0.00
g__Clostridiales_vadinBB60_group_unclassified	5.28	0	3.04	0.00
g__Clostridiales_unclassified	5.68	0.0001	4.16	0.00
g__Coprococcus_3	-2.63	0.0001	3.14	0.00
g__Odoribacter	4.41	0.0001	3.16	0.00
g__Duncaniella	6.16	0.0001	3.05	0.00
g__Turicibacter	4.89	0.0003	3.35	0.00
g__Rothia	-3.41	0.0009	3.55	0.00
g__Anaerostipes	-3.41	0.0012	3.26	0.00
g__Oscillibacter	3.38	0.0017	3.32	0.00
g__Dialister	-3.28	0.0029	3.19	0.00
g__Cetobacterium	10.33	0.0158	3.20	0.02
g__Holdemanella	-2.56	0.0176	3.54	0.02
g__Clostridium_sensu_stricto_1	-4.54	0.0198	3.61	0.02

g__Romboutsia	-3.31	0.0376	3.52	0.04

Table S4　Differentially expressed bacteria in ICC and other CCA according to nonparametric tests

Genus	log2FC	wilcox.test. p_value	regulation	mean	mean_ICC	mean_CCA
g__Murimonas	-3.73	0.02	down	0.04040	0.00550	0.07320
g__Asaccharobacter	-Inf	0.04	down	0.00510	0.00000	0.01000
g__Ezakiella	-Inf	0.04	down	0.00510	0.00000	0.00990
g__Garciella	-Inf	0.04	down	0.00060	0.00000	0.00120

Table S5　Correlations between between significantly different genus based on Sparcc

Genus	Genus	value	abs
g__Lachnospiraceae_NK4A136_group	g__Muribaculaceae_unclassified	0.57	0.57
g__Muribaculum	g__Muribaculaceae_unclassified	0.47	0.47
g__Alistipes	g__Muribaculaceae_unclassified	0.43	0.43
g__Muribaculaceae_unclassified	g__Lachnospiraceae_NK4A136_group	0.57	0.57
g__Muribaculum	g__Lachnospiraceae_NK4A136_group	0.41	0.41
g__Muribaculaceae_unclassified	g__Muribaculum	0.47	0.47
g__Lachnospiraceae_NK4A136_group	g__Muribaculum	0.41	0.41
g__Muribaculaceae_unclassified	g__Alistipes	0.43	0.43

Table S6　The correlation coefficient between 12 genus and clinical variables based on Spearman's rank correlation

	Age	BMI	ALT	AST	AKP	GGT	TB	ALB
g__Muribaculaceae_unclassified	0.33	-0.14	-0.08	0.05	0.26	0.19	0.19	-0.33
g__Faecalibacterium	-0.29	0.18	-0.01	-0.11	-0.16	-0.21	-0.19	0.43
g__Muribaculum	0.25	-0.14	0.02	0.15	0.25	0.25	0.28	-0.37
g__Agathobacter	-0.29	0.09	-0.07	-0.20	-0.18	-0.15	-0.13	0.35
g__Clostridium	0.28	-0.23	-0.08	0.05	0.28	0.19	0.21	-0.25
g__Intestinimonas	0.22	-0.24	-0.03	0.12	0.30	0.21	0.27	-0.26
g__Coprococcus_3	-0.29	0.15	0.10	-0.04	-0.08	-0.01	-0.07	0.32
g__Turicibacter	0.22	-0.21	0.07	0.20	0.24	0.21	0.30	-0.09
g__Rothia	-0.18	0.00	-0.01	-0.10	-0.14	-0.13	-0.12	0.13
g__Anaerostipes	-0.35	0.04	0.06	0.03	-0.19	-0.11	-0.02	0.32
g__Cetobacterium	0.13	-0.06	0.17	0.14	0.30	0.22	0.21	-0.12
g__Clostridium_sensu_stricto_1	-0.08	0.10	-0.11	-0.10	-0.18	-0.16	-0.17	0.36