

Supplementary Table S1 Anosim analysis of difference among CF, ECC and SECC

Group	bray curtis		weight unfrac		unweight unfrac	
	R-value	P-value	R-value	P-value	R-value	P-value
ECC-SECC	0.01089	0.291	0.008051	0.316	0.06519	0.012
CF-SECC	0.03395	0.061	0.01136	0.266	0.1239	0.001
CF-ECC	-0.00151	0.455	-0.02371	0.946	0.03502	0.04

R-value shows the difference among CF, ECC and SECC. R-value less than zero represents that the difference of samples within group is larger than between groups, whereas value greater than zero indicates that the difference of samples between groups is larger than within group. *

Significant level P-value = 0.05.

Supplementary Table S2 Significant pathway between CF, ECC and SECC groups

up								
#Pathway	CF-vs-ECC_Up(37)	Reference(4362)	RichFactor	Pvalue	Qvalue	Pathway ID	Level 1	Level 2
Basal transcription factors	1	4	0.25	0.0335	0.3924	ko03022	Genetic Information Processing	Transcription
#Pathway	CF-vs-SECC_Up(145)	Reference(4402)	RichFactor	Pvalue	Qvalue	Pathway ID	Level 1	Level 2
Propanoate metabolism	9	65	0.1385	0.0002	0.0113	ko00640	Metabolism	Carbohydrate metabolism
Selenocompound metabolism	5	20	0.25	0.0004	0.0113	ko00450	Metabolism	Metabolism of other amino acids
Glyoxylate and dicarboxylate metabolism	7	60	0.1167	0.0033	0.0651	ko00630	Metabolism	Carbohydrate metabolism
Phosphotransferase system (PTS)	6	59	0.1017	0.0123	0.185	ko02060	Environmental Information Processing	Membrane transport
Biosynthesis of ansamycins	1	1	1	0.0329	0.2856	ko01051	Metabolism	Metabolism of terpenoids and polyketides

Nonribosomal peptide structures	1	1	1	0.0329	0.2856	ko01054	Metabolism	Metabolism of terpenoids and polyketides
Tropane, piperidine and pyridine alkaloid biosynthesis	2	9	0.2222	0.0333	0.2856	ko00960	Metabolism	Biosynthesis of other secondary metabolites

#Pathway	ECC-vs-SECC_Up(154)	Reference(4447)	RichFactor	Pvalue	Qvalue	Pathway ID	Level 1	Level 2
Propanoate metabolism	11	65	0.1692	0	0.0006	ko00640	Metabolism	Carbohydrate metabolism
Phosphotransferase system (PTS)	10	59	0.1695	0	0.0008	ko02060	Environmental Information Processing	Membrane transport
Fructose and mannose metabolism	9	59	0.1525	0.0002	0.0032	ko00051	Metabolism	Carbohydrate metabolism
Galactose metabolism	6	42	0.1429	0.0029	0.0438	ko00052	Metabolism	Carbohydrate metabolism
Glyoxylate and dicarboxylate metabolism	7	61	0.1148	0.0047	0.0567	ko00630	Metabolism	Carbohydrate metabolism
Nonribosomal peptide structures	1	1	1	0.0346	0.3463	ko01054	Metabolism	Metabolism of terpenoids and polyketides
Glycerolipid metabolism	3	24	0.125	0.0484	0.4148	ko00561	Metabolism	Lipid metabolism

down

#Pathway	CF-vs-ECC_Down(92)	Reference(4362)	RichFactor	Pvalue	Qvalue	Pathway ID	Level 1	Level 2
Cell cycle - Caulobacter	6	24	0.25	0	0.0003	ko04112	Cellular Processes	Cell growth and death
Bacterial secretion system	8	53	0.1509	0	0.0003	ko03070	Environmental Information Processing	Membrane transport
Geraniol degradation	2	9	0.2222	0.0144	0.2062	ko00281	Metabolism	Metabolism of terpenoids and polyketides
RIG-I-like receptor signaling pathway	1	1	1	0.0211	0.2267	ko04622	Organismal Systems	Immune system
Renin-angiotensin system	1	2	0.5	0.0417	0.2992	ko04614	Organismal Systems	Endocrine system

African trypanosomiasis	1	2	0.5	0.0417	0.2992	ko05143	Human Diseases	Infectious diseases: Parasitic
-------------------------	---	---	-----	--------	--------	---------	----------------	-----------------------------------

#Pathway	CF-vs- SECC_Down(1 32)	Reference(4402)	RichFactor	Pvalue	Qvalue	Pathway ID	Level 1	Level 2
Cell cycle - Caulobacter	8	23	0.3478	0	0	ko04112	Cellular Processes	Cell growth and death
Propanoate metabolism	8	65	0.1231	0.0006	0.0253	ko00640	Metabolism	Carbohydrate metabolism Xenobiotics
Dioxin degradation	2	6	0.3333	0.0124	0.1599	ko00621	Metabolism	biodegradation and metabolism Metabolism of
Carotenoid biosynthesis	2	6	0.3333	0.0124	0.1599	ko00906	Metabolism	terpenoids and polyketides
Degradation of aromatic compounds	4	36	0.1111	0.0217	0.1599	ko01220	Metabolism	Global and overview maps
Valine, leucine and isoleucine degradation	4	37	0.1081	0.0237	0.1599	ko00280	Metabolism	Amino acid metabolism
Small cell lung cancer	1	1	1	0.03	0.1599	ko05222	Human Diseases	Cancers: Specific types
Influenza A	1	1	1	0.03	0.1599	ko05164	Human Diseases	Infectious diseases: Viral
p53 signaling pathway	1	1	1	0.03	0.1599	ko04115	Cellular Processes	Cell growth and death
Hepatitis B	1	1	1	0.03	0.1599	ko05161	Human Diseases	Infectious diseases: Viral
Toxoplasmosis	1	1	1	0.03	0.1599	ko05145	Human Diseases	Infectious diseases: Parasitic
Monoterpenoid biosynthesis	1	1	1	0.03	0.1599	ko00902	Metabolism	Metabolism of terpenoids and polyketides
Colorectal cancer	1	1	1	0.03	0.1599	ko05210	Human Diseases	Cancers: Specific types
Steroid biosynthesis	1	1	1	0.03	0.1599	ko00100	Metabolism	Lipid metabolism
Viral myocarditis	1	1	1	0.03	0.1599	ko05416	Human Diseases	Cardiovascular diseases

#Pathway	ECC-vs- SECC_Down(1 79)	Reference(4447)	RichFactor	Pvalue	Qvalue	Pathway ID	Level 1	Level 2
----------	-------------------------------	---------------------	------------	--------	--------	------------	---------	---------

Oxidative phosphorylation	11	74	0.1486	0.0002	0.0075	ko00190	Metabolism	Energy metabolism
Chloroalkane and chloroalkene degradation	3	14	0.2143	0.0168	0.4205	ko00625	Metabolism	Xenobiotics biodegradation and metabolism

Supplementary Table S3 Basic statistics of samples for metagenome sequencing

Sample Name	InsertSize(bp)	RawData(M)	CleanData(M)	Clean_Q20	NonHost(%)	Effective(M)
CF16	350	13,224.25	13,220.19	96.88	99.97	12,928.32
CF22	350	10,458.45	10,455.63	97.22	99.97	10,403.50
CF7	350	10,298.91	10,296.71	96	99.98	10,158.90
EC1	350	10,319.79	10,317.43	95.64	99.98	9,647.22
EC26	350	10,587.26	10,584.08	96.66	99.97	10,561.02
EC4	350	11,115.43	11,112.11	97	99.97	11,074.94
SECC17	350	11,022.50	11,019.01	97.06	99.97	10,980.76
SECC20	350	10,933.94	10,930.58	96.83	99.97	10,883.36
SECC28	350	10,233.19	10,229.93	96.79	99.97	9,457.76
SECC2	350	10,707.29	10,704.76	96.2	99.98	9,639.62

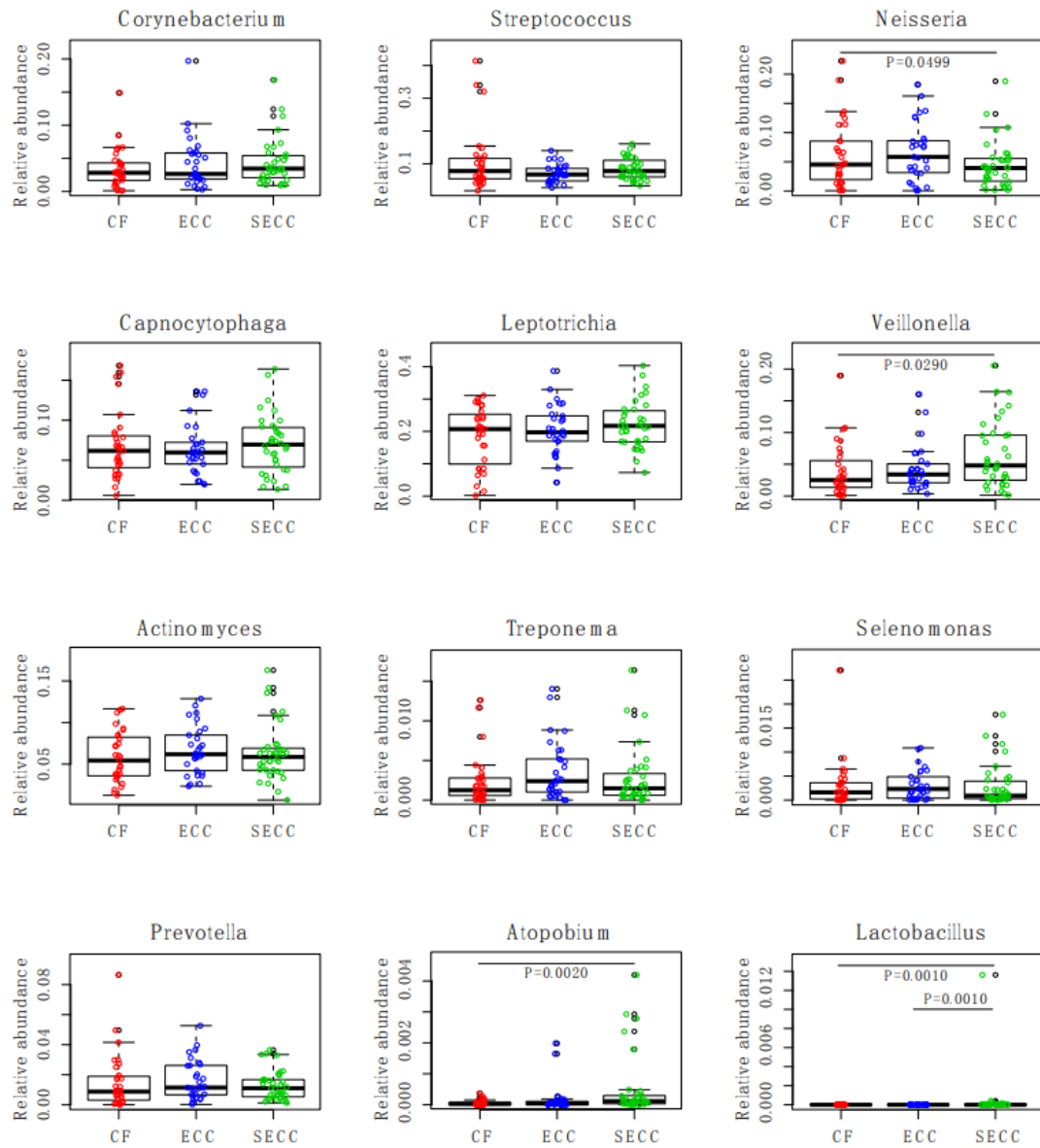


Figure S1. Genus biomarkers for caries.

The box depicts the interquartile range (IQR) between the first and third quartiles (25th and 75th percentiles, respectively) and the line inside denotes the median, whereas the points represent the abundance of sample.

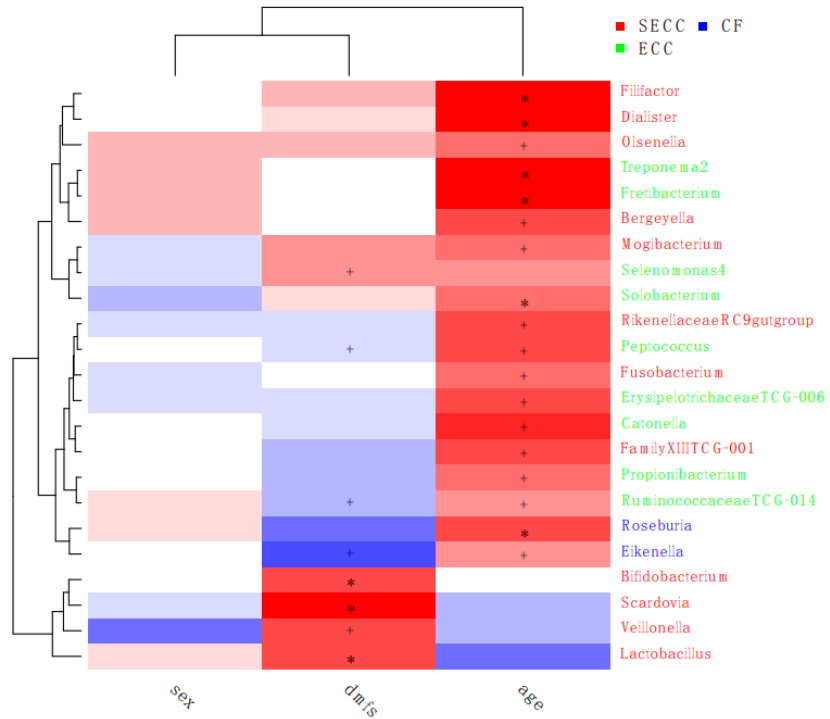
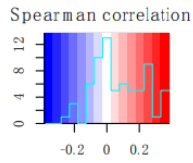


Figure S2. Association between genus biomarkers and decayed, missing, and filled surface (dmfs) index.

Heatmap of the spearman correlation between genus and dmfs, age, gender. BH adjusted p-value is denoted: +, q-value<0.1; *, q-value<0.05; **, q-value<0.01.

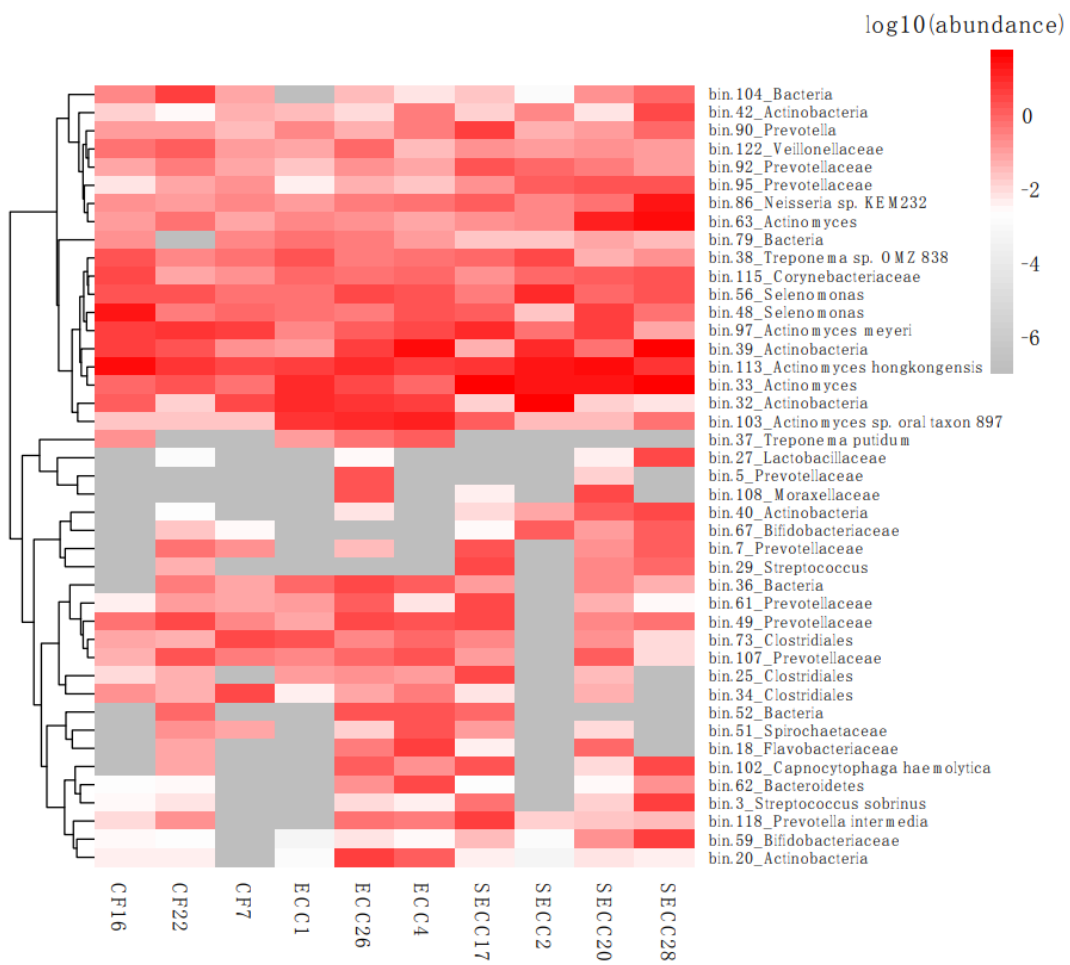


Figure S3. MAGs population distribution.

Heatmap of relative abundance of MAGs.

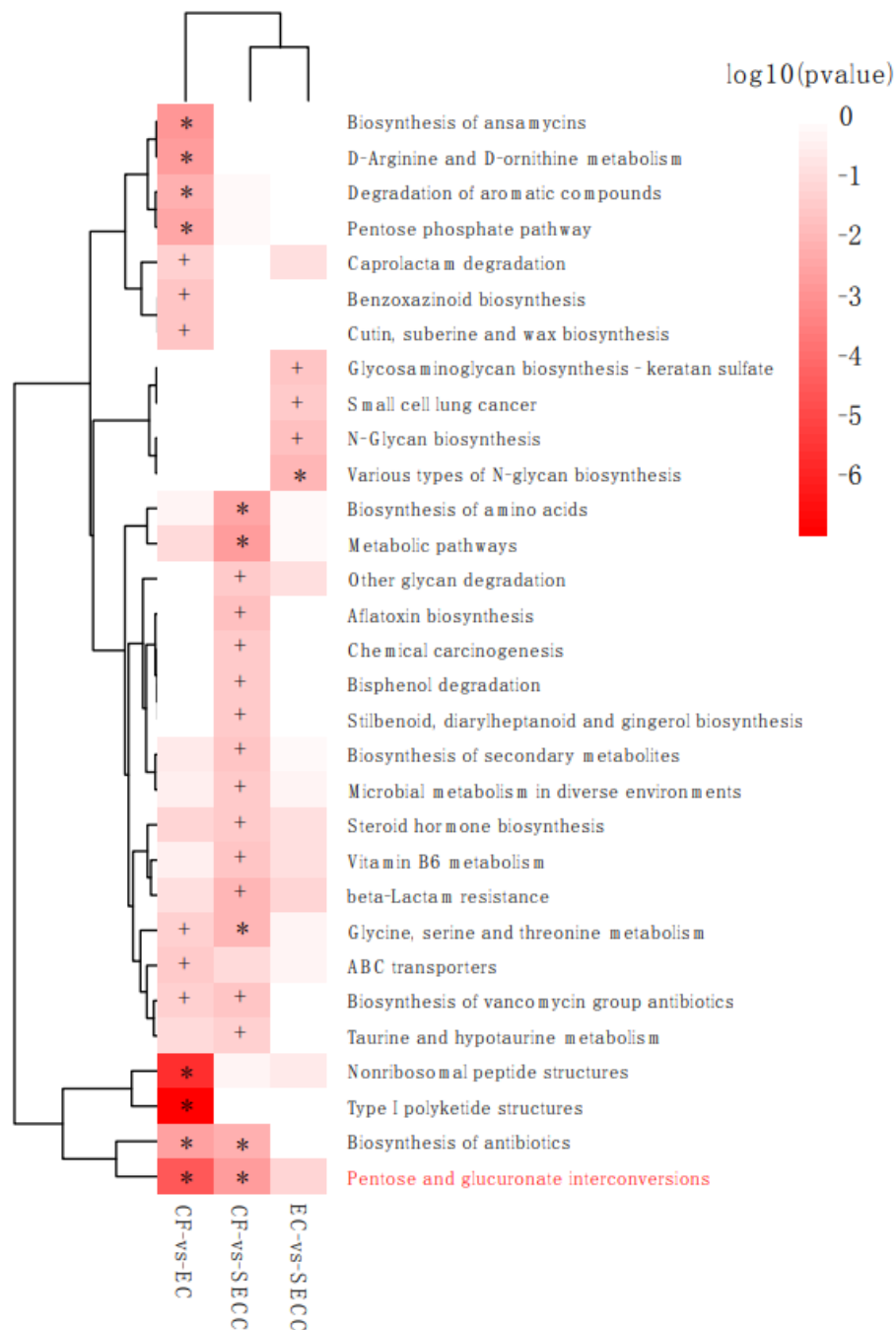


Figure S4. Association between pathway and caries.

Heatmap of the spearman correlation between pathway and caries. Wilcox test was used to compare groups. BH adjusted p-value is denoted: +, q-value<0.1; *, q-value<0.05; **, q-value<0.01.