Supplementary Materials

Microbiomic Subprofiles Associate with *MDR1* Promoter Methylation in Head and Neck Squamous Cell Carcinoma

Gurkan Bebek[1, 4], Kristi L. Bennett[1], Pauline Funchain[1], Rebecca Campbell[1], Rahul Seth[2], Joseph Scharpf[2], Brian Burkey[2], & Charis Eng[1, 3, 5, 6, *]

[1]Genomic Medicine Institute, Lerner Research Institute, [2]Head and Neck Institute, [3]Taussig Cancer Institute, Cleveland Clinic, Cleveland, OH 44195, USA. [4]Center for Proteomics and Bioinformatics, [5]Department of Genetics and [6]Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, OH 44116, USA.

* Corresponding author. Charis Eng, Genomic Medicine Institute, Lerner Research Institute, 9500 Euclid Avenue, Cleveland, OH 44195, USA. Tel: 216/444-3440; Fax: 216/636-0655; Email: engc@ccf.org.

This file contains Supplementary Methods, Supplementary Figures and Supplementary Tables.

<u>Table of Contents:</u>

Supplementary Methods:

DNA extraction:

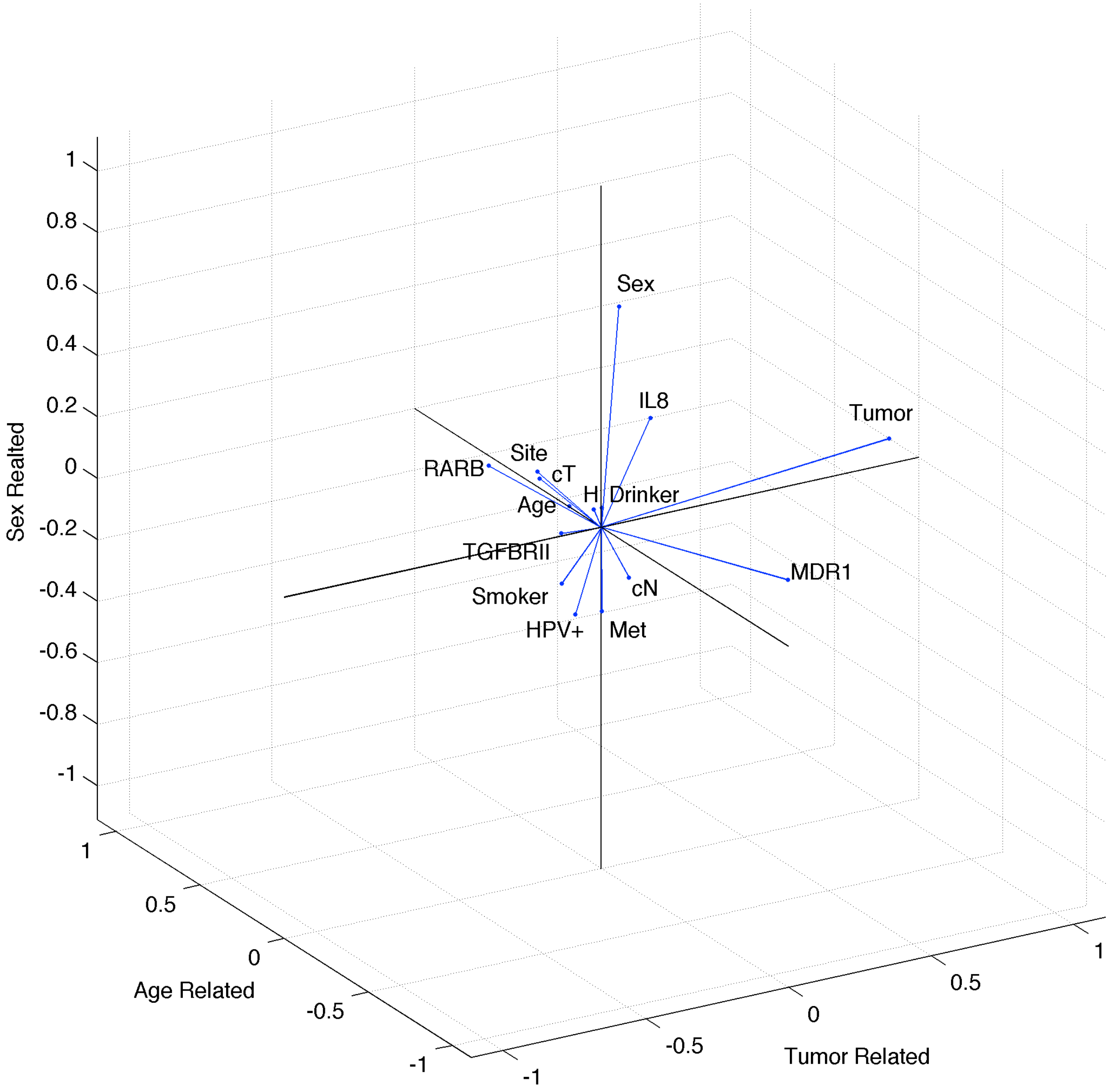
Fresh frozen tissue - Approximately 0.05g of each sample tissue was homogenized in 50mL of 2% trisodium citrate then centrifuged at 10,000g for 10 min. After the supernatants were removed, the pellets were re-suspended in 500μl of trisodium citrate. Centrifugation was repeated, the supernatant was removed, and the pellets were re-suspended in 400ml of enzymatic lysis buffer (20mM Tris HCl, 2 mM EDTA, 1.2% Triton X100, 20 mg/ml lysozyme, 1000U mutanolysin). These solutions were then incubated at 37°C for 1 hr. After 40μl of proteinase K (20mg/ml) and 200μl of Buffer AL (Qiagen) were added, the solutions were incubated at 56°C for 30 min. The QIAmp DNA mini kit (Qiagen, Valencia, CA, USA) was used to complete the DNA purifications.

Factor analysis:

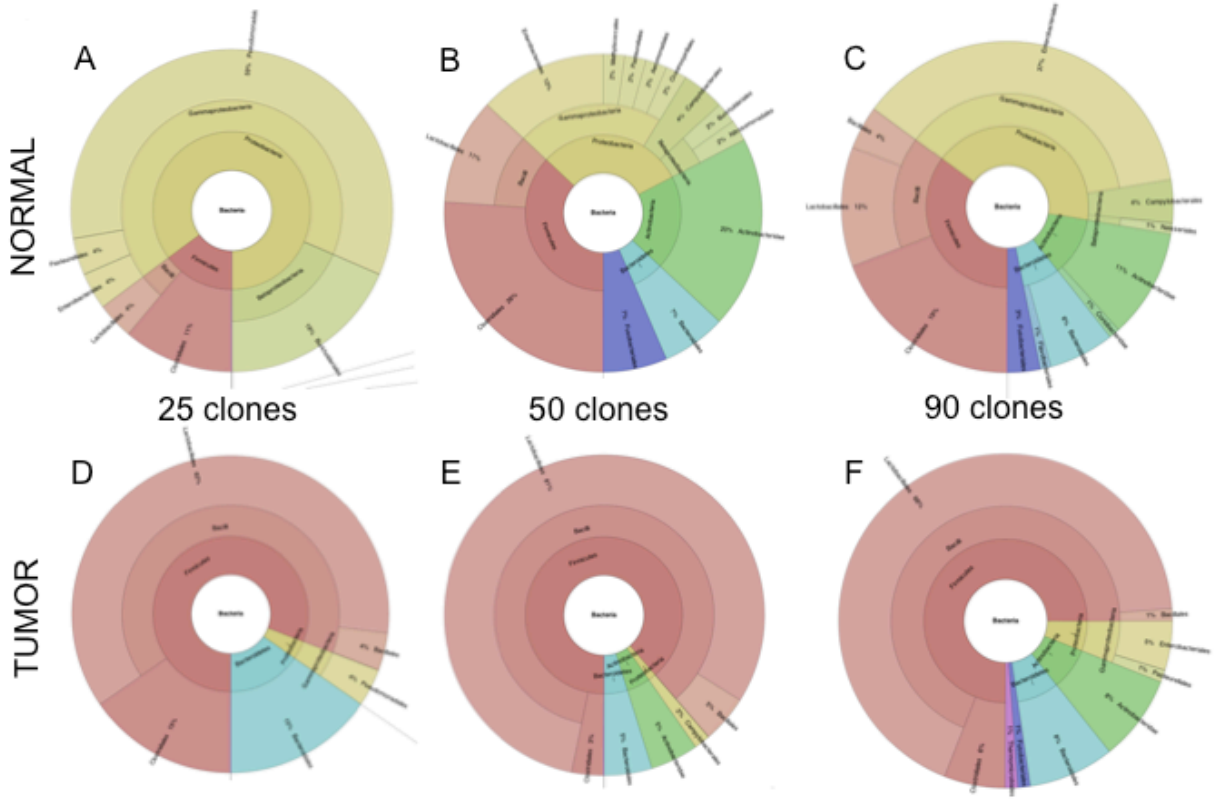
To reduce the number of variables and to detect the relationships between variables (i.e. classify variables), we carried out factor analysis using principle component extraction (Suppl. Fig. 1). Three significant factors (i) tumor, (ii) age, and (iii) sex related factors were apparent. Although arbitrary, the axis names explain the factors modestly. In this group of specimens analyzed, smoking, drinking, and age of patients were closely related. However, since these variables and variables such as tumor site had been placed in different factor groups, these were not closely related with HNSCC tumors when compared to normal mucosa. *RARB* methylation was on the same axis with age, and tumor site, indicating that not all tumors, but a subgroup has such alterations and that site and age have the largest effect on *RARB* methylation. cT and cN were closely related with outcome, i.e. HNSCC. *MDR1* and *IL8* methylation status of specimens were closer to the tumor axis validating the correlation of these genes and their association with HNSCC over normal mucosa, irrespective of microbial subflora. HPV status was on the same axis as sex. HPV status and *MDR1* methylation appearing in vastly different quadrants (Suppl. Fig.

1) indicate that HPV+ patients have gone through similar alterations as HPV- patients, and therefore are not the differentiator of *MDR1* methylation associating with tumor.

Supplementary Figures

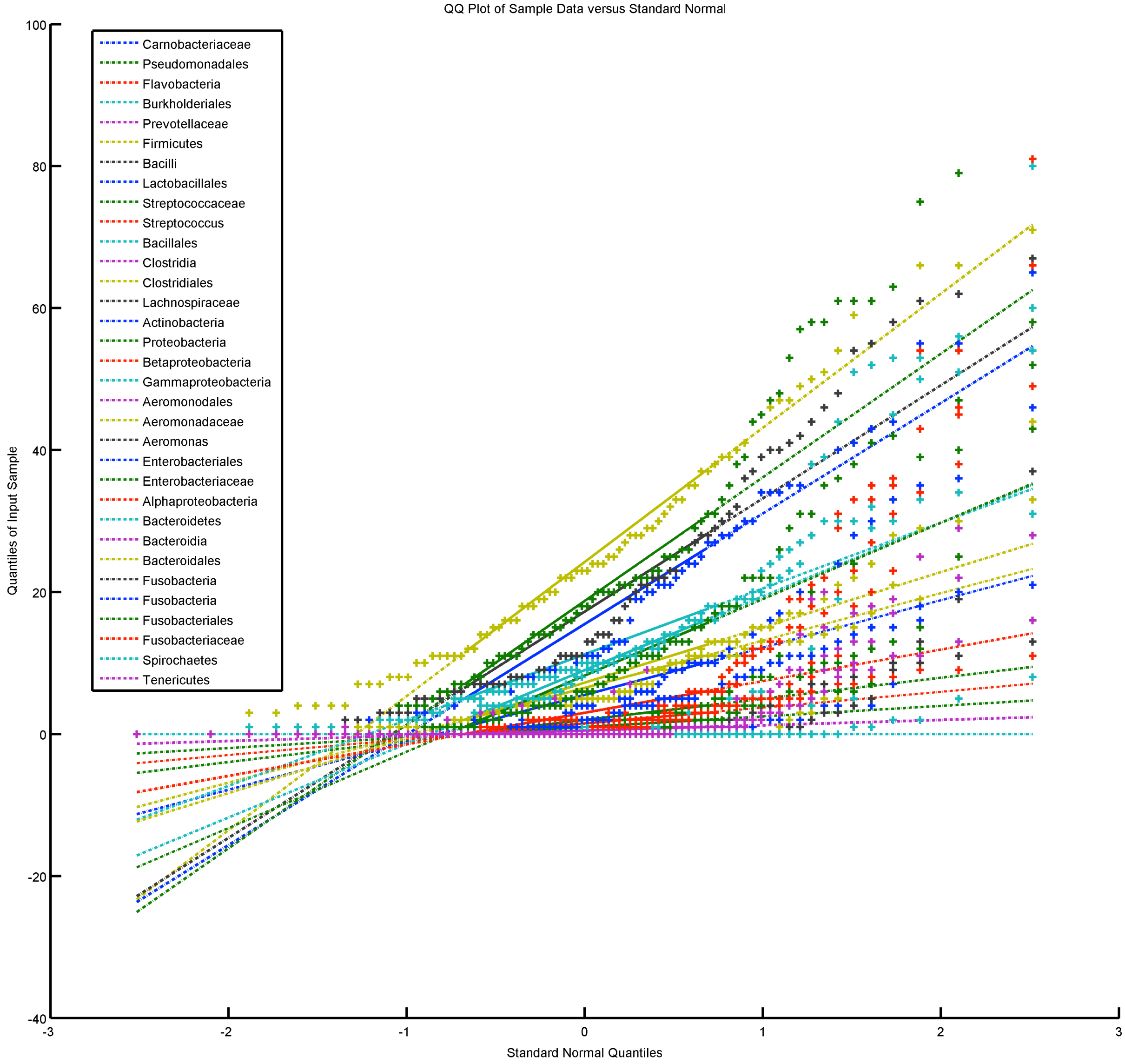


Supplementary Figure 1: Factor analysis identifies three components that underlie the clinical parameters, which are named tumor, age, and sex related based on their components.



Supplementary Figure 2: Microbiomic content, pictured as concentric pie graph with relative abundance of each phylum represented proportionally by color (1), showing adjacent normal (A,B,C) and HNSCC tumor tissue (D,E,F) from a single individual using approximately 25 (A,D), 50 (B,E), and 90 (C,F) clones. Taxonomy subdivides outward with each concentric layer.

- 🟥 Firmicutes
- 🟪 Fusobacteria
- 🟦 Bacteroidetes
- 🟩 Actinobacteria
- 🟨 Proteobacteria
- 🟪 Thermomicrobia



Supplementary Figure 3: QQ plot of the microbial populations are shown. Most bacterial populations follow a normal distribution with some outliers. This plot shows fit of the microbial data for multivariate analysis.

Supplementary Tables

Supplementary Table 1. Clinical information collected from specimen donors is shown. Age: Years old at diagnosis; Sex: M-Male, F-Female; Site: 1-Tongue, 2-Larynx, 3-Tonsil, 4-Floor of Mouth, 5-Oral cavity, 6-Hypopharynx, 7-Other; Histological Differentiation: 1-Well, 2-Well-to-Moderate, 3-Moderate, 4-Moderate-to-Poor, 5-Poor, 6-NOS(*Not Otherwise Specified*)/Other; Surgery: 0-No prior surgery, 1-Prior surgery; XRT: Radiation therapy; Met: Metastatic status; {XRT, Recurrence, Smoker, Drinker, Met}: 0-No, 1-Yes; cT: Clinical tumor stage; cN: Nodal involvement; HPV: HPV status (0-negative, 1-positive).

Patient ID	Age	Sex	Site	Hist Diff	Surgery	XRT	Chemo	Recurrence	Smoker	Drinker	cT	cN	HPV	Met
1	57	M	1	4	1	1	0	1	0	0	3	1	0	0
2	71	M	2	5	1	0	0	0	1	0	4	0	0	0
4	59	M	7	3	1	1	1	1	1	1	0	0	0	1
10	71	F	4	4	0	1	1	1	1	1	4	2	0	0
12	79	M	7	3	1	1	0	0	1	0	2	1	0	1
13	54	F	5	6	1	0	0	0	1	1	4	0	0	0
16	63	F	1	4	1	0	0	0	1	1	2	0	0	0
18	62	M	1	3	1	1	0	0	0	0	4	0	0	0
19	74	M	3	3	1	1	0	1	1	1	2	0	1	0
27	61	M	4	4	1	1	1	1	1	1	3	1	0	0
30	44	M	3	3	0	0	0	0	1	1	1	1	0	0
33	75	M	5	1	1	0	0	1	1	1	4	1	0	0
34	65	M	7	3	1	0	0	0	1	1	4	0	0	0
35	32	F	4	3	1	0	0	1	1	1	3	2	0	0
38	71	M	2	3	1	1	0	1	1	0	2	0	0	0
39	46	M	2	3	1	0	0	0	1	0	4	2	0	0
40	50	M	3	6	1	0	0	0	1	1	1	0	0	0
42	85	F	1	3	1	0	0	1	1	1	0	2	0	0
43	65	F	2	3	1	1	0	1	0	1	3	0	0	0
45	53	M	3	3	1	0	0	0	1	0	1	2	1	0
46	78	M	3	3	0	0	0	0	1	1	3	0	1	0
49	65	F	2	3	1	0	0	0	1	1	3	2	0	0
54	43	M	3	1	0	0	0	0	1	0	2	2	0	0
55	59	M	2	3	1	0	0	0	1	1	4	0	0	0
56	57	F	1	2	1	0	0	0	0	0	3	0	0	0
57	73	M	5	4	1	0	0	0	1	1	4	0	0	0
58	57	M	7	1	1	0	0	0	1	0	2	1	0	0
62	65	F	2	4	0	0	0	0	1	0	4	1	0	0
63	62	M	3	5	0	0	0	0	0	0	2	2	1	0
73	72	F	1	4	1	1	1	1	1	0	2	0	0	0
75	61	F	6	5	0	1	1	1	0	0	3	1	0	0
76	71	F	6	3	1	1	1	1	1	0	4	0	1	0
78	55	M	5	3	1	0	0	0	1	1	4	2	0	0
79	70	M	1	1	1	0	0	0	1	1	2	0	0	0
83	72	M	2	2	1	0	0	0	1	1	4	2	0	1
84	52	M	2	5	1	1	1	1	1	0	3	3	0	0
85	77	F	1	3	1	1	1	0	1	1	2	0	0	0
88	51	F	3	4	1	0	0	0	0	1	0	2	0	0
89	78	F	2	4	1	0	0	0	1	0	3	0	0	0
91	61	M	1	5	1	1	1	1	1	0	4	3	1	0
93	61	M	2	4	1	0	0	0	1	1	4	0	1	0
94	50	F	2	4	1	0	0	0	1	1	4	0	0	0

Supplementary Table 2: Microbial populations identified from normal mucosa samples are shown.

	Microbial Populations																																
Patient ID	Carnobacteriaceae	Pseudomonadales	Flavobacteria	Burkholderiales	Prevotellaceae	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	Bacillales	Clostridia	Clostridiales	Lachnospiraceae	Actinobacteria	Proteobacteria	Betaproteobacteria	Gammaproteobacteria	Aeromonodales	Aeromonadaceae	Aeromonas	Enterobacteriales	Enterobacteriaceae	Alphaproteobacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Fusobacteria	Fusobacteria	Fusobacteriales	Fusobacteriaceae	Spirochaetes	Tenericutes
1	0	0	58	1	1	22	2	2	2	2	0	18	13	8	0	1	1	0	0	0	0	0	0	0	60	2	2	12	12	12	12	0	0
2	0	1	6	34	0	16	7	6	0	0	0	9	9	0	0	44	35	4	0	0	0	0	0	4	6	0	0	10	10	10	10	0	0
4	0	1	1	0	0	3	0	0	0	0	0	3	0	0	0	75	66	5	0	0	0	4	4	3	3	1	1	1	1	1	1	0	0
10	0	0	23	11	0	25	14	13	0	0	1	11	11	0	6	14	12	1	0	0	0	1	0	1	30	7	7	13	13	13	0	0	0
12	0	4	2	33	0	20	5	3	0	0	2	15	15	0	7	38	33	5	0	0	0	0	0	0	19	17	17	4	4	4	0	2	0
13	0	4	2	30	0	18	5	3	0	0	2	13	13	0	6	35	30	5	0	0	0	0	0	0	18	16	16	4	4	4	0	2	0
16	0	0	25	6	0	23	8	8	0	0	0	14	14	0	5	7	6	0	0	0	0	0	0	0	30	4	4	26	26	26	0	0	0
18	21	3	2	0	5	49	46	43	8	0	2	3	3	2	0	20	2	18	0	0	0	15	15	0	7	5	5	1	1	1	1	0	0
19	0	0	0	1	0	32	19	16	0	0	3	12	12	0	30	2	2	2	0	0	0	0	0	0	10	10	10	18	18	18	0	0	0
27	0	25	2	29	0	19	10	9	0	0	1	9	9	0	2	61	29	27	0	0	0	2	0	5	2	0	0	9	9	9	0	1	0
30	0	4	4	4	0	15	7	6	0	0	0	1	1	0	5	31	4	16	0	0	0	4	0	5	9	2	0	6	6	6	6	0	3
33	4	1	0	0	4	28	28	24	17	17	4	0	0	0	15	14	3	9	3	3	3	6	6	1	6	6	6	15	15	15	15	0	0
34	0	0	0	0	2	66	58	38	40	34	18	7	7	0	16	10	1	5	0	0	0	5	5	0	1	1	0	6	6	6	6	0	0
35	0	3	2	0	0	23	8	4	0	0	4	15	15	0	10	16	0	15	0	0	0	1	0	0	10	7	7	9	9	9	0	5	0
38	0	0	0	0	2	29	11	11	5	0	0	17	17	1	35	0	0	0	0	0	0	0	0	0	12	12	12	2	2	2	2	0	0
39	0	0	1	1	22	39	24	24	2	0	0	13	13	0	14	11	10	1	0	0	0	1	1	0	23	22	22	0	0	0	0	0	0
40	0	4	8	5	0	15	13	12	0	0	0	1	1	0	6	26	5	12	0	0	0	5	0	7	13	3	3	17	17	17	0	0	9
42	0	0	0	0	1	61	55	55	2	0	0	6	6	0	22	0	0	0	0	0	0	0	0	0	2	1	1	1	1	1	0	0	0
43	1	1	4	4	1	37	32	27	3	3	2	4	4	0	1	20	4	16	0	0	0	5	0	0	19	15	15	1	1	1	0	0	2
45	0	3	7	2	0	12	11	10	0	0	1	1	1	0	9	20	2	7	1	0	0	1	0	9	14	14	14	11	11	11	0	0	6
46	0	3	1	2	0	16	14	12	0	0	2	2	2	0	24	12	2	5	0	0	0	1	0	5	2	1	1	13	13	13	13	0	0
49	2	5	3	5	1	47	37	34	0	0	2	10	10	1	2	22	6	16	0	0	0	11	3	0	16	13	13	2	2	1	1	0	0
54	0	0	0	0	0	4	3	2	0	0	2	1	1	0	5	11	0	10	0	0	0	9	0	0	1	1	1	0	0	0	0	0	0
55	0	4	2	1	0	26	22	20	2	2	2	4	4	0	1	18	1	16	1	0	0	7	0	1	7	5	5	31	31	31	6	0	0
56	0	0	0	0	0	8	7	7	0	0	0	1	1	0	5	33	4	18	0	0	0	2	0	8	10	4	2	11	11	11	0	0	8
57	0	1	0	0	4	20	16	11	2	2	5	4	4	0	2	8	0	8	0	0	0	0	0	0	24	24	24	35	35	35	35	0	1
58	0	0	0	0	9	28	22	19	4	4	2	5	5	0	0	11	0	11	0	0	0	0	0	0	9	9	9	29	29	29	7	0	1
62	2	4	4	11	2	35	31	28	4	0	3	4	4	1	5	21	11	9	0	0	0	4	0	1	17	13	13	8	8	8	0	0	0
63	0	4	8	1	0	10	9	9	1	0	0	1	1	0	12	23	2	11	0	0	0	0	0	9	9	1	1	15	15	15	15	0	9
73	0	0	1	0	0	63	61	55	52	49	1	7	7	7	0	23	0	20	9	9	9	11	11	2	1	0	0	1	1	1	1	0	0
75	3	3	0	3	1	54	48	30	8	0	15	4	4	0	3	15	4	11	0	0	0	5	3	0	10	10	10	4	4	4	2	0	0
76	2	0	0	0	15	18	16	10	1	1	5	2	2	0	1	4	0	4	0	0	0	4	2	0	30	30	30	42	42	42	13	0	1
78	0	0	0	0	28	22	21	21	13	3	1	1	1	0	46	1	0	1	0	0	0	1	1	0	28	28	28	0	0	0	0	0	0
79	7	1	0	0	0	24	14	13	0	0	0	10	10	0	4	53	0	51	29	0	0	18	0	2	2	2	2	1	1	1	0	0	0
83	0	2	0	0	0	14	1	1	1	0	0	13	13	13	2	61	0	53	25	25	1	25	25	8	0	0	0	0	0	0	0	0	0
84	0	1	0	2	4	23	13	13	11	10	1	10	10	10	1	48	2	39	5	5	4	12	12	5	10	10	10	0	0	0	0	0	0
85	0	12	1	8	0	44	41	40	0	0	1	2	2	0	2	39	21	16	0	0	0	3	0	2	4	3	3	0	0	0	0	0	0
88	0	0	0	0	0	7	6	6	0	0	0	1	1		5	26	4	11	0	0	0	0	0	11	18	4	4	16	16	16	0	0	16
89	20	1	0	0	1	47	42	41	3	0	1	4	4	4	2	5	1	4	0	0	0	2	2	0	1	1	1	31	31	31	31	0	1
91	16	4	0	0	1	30	20	19	3	0	1	10	10	10	10	45	19	26	9	9	0	13	13	0	1	1	1	7	7	7	7	0	0
93	1	2	1	0	6	7	7	2	1	1	0	5	5	1	13	10	0	11	3	3	3	5	5	1	20	20	18	43	43	43	43	0	0
94	1	1	0	2	4	22	11	9	8	6	2	11	11	11	4	58	3	52	37	37	37	10	10	2	10	6	6	2	2	2	2	0	0

Supplementary Table 3: Microbial populations identified from tumor samples are shown.

	Microbial Populations																																
Patient ID	Carnobacteriaceae	Pseudomonadales	Flavobacteria	Burkholderiales	Prevotellaceae	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	Streptococcus	Bacillales	Clostridia	Clostridiales	Lachnospiraceae	Actinobacteria	Proteobacteria	Betaproteobacteria	Gammaproteobacteria	Aeromonodales	Aeromonadaceae	Aeromonas	Enterobacteriales	Enterobacteriaceae	Alphaproteobacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Fusobacteria	Fusobacteria	Fusobacteriales	Fusobacteriaceae	Spirochaetes	Tenericutes
1	0	0	45	0	4	12	1	0	0	0	1	11	11	8	0	0	0	0	0	0	0	0	0	0	50	5	5	22	22	22	22	0	0
2	0	0	2	22	1	4	1	1	0	0	0	3	3	0	0	79	22	53	0	0	0	0	0	4	3	1	1	2	2	2	2	0	5
4	0	1	5	0	0	4	1	1	0	0	0	4	4	0	1	61	54	1	0	0	0	0	0	2	8	3	3	4	4	4	4	0	0
10	0	6	18	15	0	19	8	8	0	0	0	11	11	0	4	25	16	7	0	0	0	1	0	1	21	3	3	22	22	22	0	0	0
12	0	5	6	54	0	11	6	5	0	0	1	5	5	0	2	63	54	6	0	0	0	0	0	3	8	2	2	3	3	3	0	0	0
13	0	4	3	19	0	39	6	5	0	0	1	33	33	0	3	30	19	10	0	0	0	5	0	1	16	13	13	2	2	2	0	1	0
16	0	0	9	12	0	25	18	8	0	0	9	7	7	0	10	14	12	1	0	0	0	1	0	0	14	5	5	22	22	22	0	0	0
18	0	0	2	0	0	41	23	20	0	0	2	14	14	0	6	1	0	1	0	0	0	1	0	0	32	29	29	7	7	7	0	0	0
19	0	0	3	0	0	38	22	21	0	0	1	12	12	0	14	4	2	0	0	0	0	0	0	0	18	15	15	10	10	10	0	0	0
27	0	0	0	0	0	46	21	21	0	0	0	25	25	0	36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	1	0	1	2	5	4	3	2	0	0	0	1	1	0	0	7	2	2	0	0	0	0	0	3	6	5	5	4	4	4	4	0	2
33	1	0	2	0	4	31	27	25	24	23	1	4	4	1	19	11	8	2	0	0	0	0	0	0	6	5	5	3	3	3	3	0	0
34	1	0	1	0	2	27	2	0	0	0	1	25	25	0	0	14	0	1	0	0	0	1	1	0	7	5	5	47	47	47	46	0	0
35	0	10	3	6	0	29	23	23	0	0	0	6	6	0	1	31	6	24	0	0	0	14	0	1	14	11	11	5	5	5	0	0	3
38	1	0	1	0	2	18	7	6	4	0	1	11	11	1	28	4	0	4	0	0	0	2	2	0	22	21	21	18	18	18	12	0	0
39	0	0	1	0	12	24	11	9	3	3	1	13	13	0	9	8	6	2	0	0	0	1	0	0	45	44	44	0	0	0	0	0	1
40	0	1	11	3	0	10	8	8	0	0	0	1	1	0	10	17	4	5	0	0	0	1	0	5	12	1	1	22	22	22	0	0	4
42	0	0	20	0	17	24	2	1	1	0	1	21	21	9	0	0	0	0	0	0	0	0	0	0	51	17	17	4	4	4	4	8	0
43	0	3	2	1	9	71	67	65	2	0	1	4	4	3	2	6	1	5	0	0	0	2	2	0	11	9	9	0	0	0	0	0	0
45	0	4	6	1	0	8	5	5	0	0	0	3	3	0	8	18	1	9	0	0	0	1	0	5	15	15	15	6	6	6	0	0	11
46	0	3	3	2	0	7	5	5	0	0	0	2	2	0	21	7	2	3	0	0	0	0	0	1	7	4	4	6	6	6	6	0	3
49	4	5	1	1	9	59	54	34	22	2	19	5	5	5	0	13	1	12	0	0	0	6	0	0	10	9	9	0	0	0	0	0	0
54	0	0	0	0	0	11	9	9	7	0	0	1	1	1	0	4	0	4	0	0	0	2	2	0	5	5	5	0	0	0	0	0	0
55	0	4	3	2	0	51	40	34	4	0	5	11	11	1	0	25	2	23	2	0	0	8	2	0	8	5	5	3	3	3	1	0	0
56	0	0	0	0	0	16	11	11	0	0	0	5	5	0	17	21	2	8	0	0	0	0	0	6	7	2	2	11	11	11	0	0	11
57	0	0	0	0	19	14	9	6	6	1	3	4	4	3	0	6	2	4	0	0	0	0	0	0	19	19	19	38	38	38	33	0	0
58	1	3	0	1	5	11	8	2	5	2	0	3	3	0	0	57	1	56	0	0	0	8	8	0	5	5	5	11	11	11	2	0	1
62	0	7	7	1	0	28	23	22	0	0	0	4	4	0	0	29	3	25	0	0	0	15	0	0	11	4	4	13	13	13	4	0	1
63	0	2	0	4	0	11	7	7	0	0	0	2	2	0	11	21	4	13	0	0	0	6	0	2	1		1	5	5	5	0	0	4
73	0	1	1	0	4	50	44	44	39	38	0	6	6	5	1	15	2	11	5	5	5	4	4	5	5	4	4	0	0	0	0	0	1
75	2	0	0	2	11	66	62	30	24	5	31	0	0	0	4	2	2	0	0	0	0	0	0	0	11	11	11	1	1	1	0	0	0
76	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	5	5	81	81	81	81	0	0
78	0	0	0	0	13	37	36	35	31	31	0	1	1	0	33	0	0	0	0	0	0	0	0	0	13	13	13	0	0	0	0	0	0
79	4	0	0	5	7	33	29	28	8	0	1	4	4	2	20	22	5	13	6	0	0	5	4	4	13	13	13	3	3	3	3	0	0
83	0	0	0	0	0	11	7	4	0	0	3	4	4	1	5	25	2	17	10	0	0	7	0	4	2	2	2	41	41	41	0	0	0
84	1	0	0	10	0	35	28	27	20	20	1	7	7	7	1	58	11	38	13	13	13	10	10	7	2	2	2	0	0	0	0	0	0
85	0	6	1	4	0	40	39	35	0	0	4	0	0	0	9	16	6	10	0	0	0	4	0	0	1	0	0	3	3	3	0	0	1
88	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19	5	6	0	0	0	3	0	6	18	7	7	13	13	13	0	0	13
89	4	5	2	0	2	45	40	39	1	0	1	5	5	2	1	20	0	19	0	0	0	11	1	1	8	6	6	8	8	8	2	0	2
91	2	43	0	0	1	3	3	3	0	0	0	0	0	0	1	47	3	44	1	1	1	0	0	0	1	1	1	36	36	36	36	0	0
93	0	0	0	0	3	33	30	29	27	27	1	3	3	1	10	22	9	13	2	2	2	9	9	2	9	9	4	18	18	18	18	0	0
94	0	1	2	0	3	8	3	3	3	1	0	5	5	5	1	80	0	80	19	19	19	58	58	1	5	3	3	0	0	0	0	0	0

Reference:

1 Ondov, B.D., Bergman, N.H. and Phillippy, A.M. (2011) Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics*, 12, 385.