Appendix 1

Further information of DNA Extraction and Quantification

All samples were processed by the same experimenter under the same conditions, and sample status was blinded. Oral mucosal samples were collected and extracted in an aseptic laminar flow hood, and all steps were taken to ensure aseptic operation. Negative control samples (without DNA template) were used to detect possible reagent and environmental contamination in all sequencing batches. Furthermore, all samples were sequenced in the same batch.

Detailed Library Construction and Sequencing

The sequencing libraries were prepared following the PacBio amplicon Template Preparation and Sequencing protocols to amplify the 27F and 1492R regions. The input gDNA of 2 ng was PCR-amplified with 10× LA PCR Buffer II (Mg^{2+}-free), 2.5 mM of dNTP mix, 2.5 mM $MgCl_2$, 500 nM each of the F/R PCR primers, and 5 U of TaKaRa LA Taq (Takara, Kusatsu, Japan). The cycle conditions for PCR were 5 min at 94°C for heat activation, and 29 cycles of 30 s at 94 °C, 30 s at 53 °C, and 90 s at 72 °C, followed by a 5-min final extension at 72 °C. The primer pair with asymmetric barcoded adapters for the amplifications were as follows: 27F-F: 5'-AGRGTTYGATYMTGGCTCAG-3', 1492-R: 5'-RGYTACCTTGTTACGACTT-3'. The PCR product was purified with SMRTbell cleanup beads. The purified product was then quantified using Quant-IT PicoGreen (Invitrogen) and evaluated using the TapeStation D5000 Screen Tape (Agilent Technologies, Waldbronn, Germany). For PacBio Sequel IIe sequencing, 500 ng of pooled amplicon DNA was used for library preparation. A 10 µL library was prepared using PacBio SMRTbell prep kit 3.0. SMRTbell templates were annealed using Sequel II Bind Kit 3.1 and Int Ctrl 3.1. The Sequel II Sequencing Kit 2.0 and SMRT cells 8M Tray were used for sequencing. SMRT cells (Pacific Biosciences) using 10 h movies were captured for each SMRT cell using the PacBio Sequel IIe (Pacific Biosciences) sequencing platform by Macrogen (Seoul, Korea). The subsequent steps followed the PacBio Sample Net-Shared Protocol (https://www.pacb.com/).

https://dx.doi.org/10.21037/jgo-2025-432

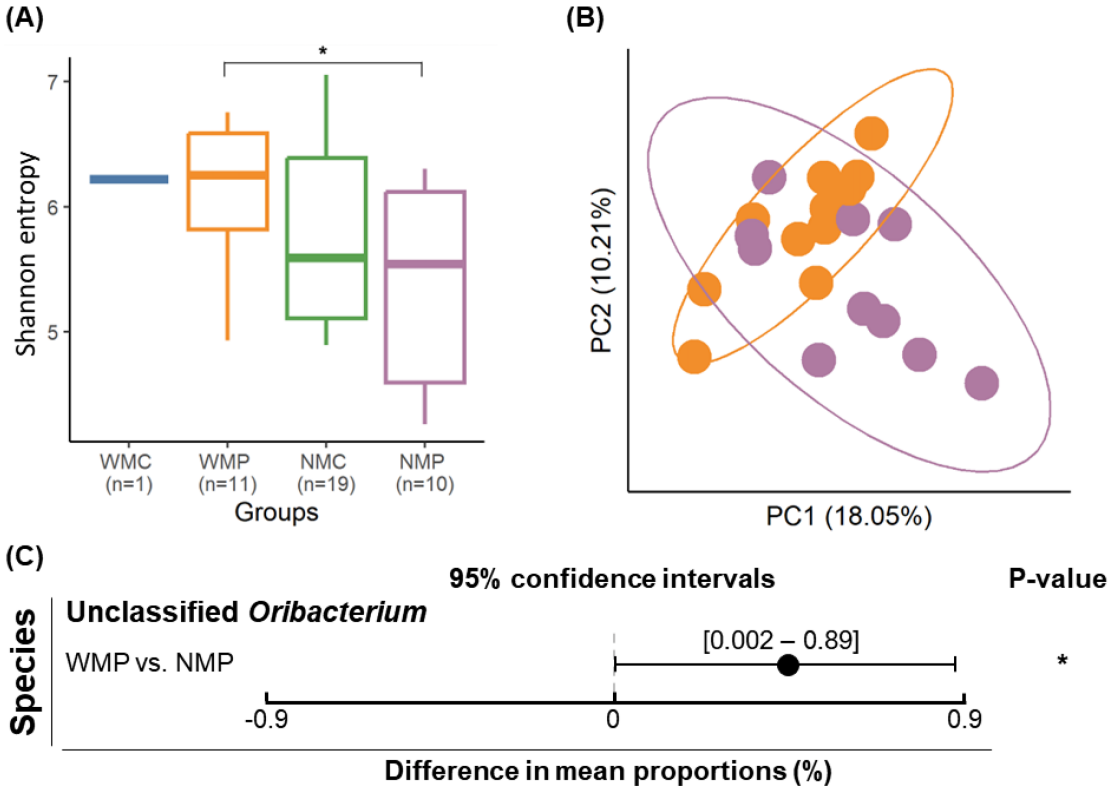

Figure S1 Oral microbiome differences in patients with esophageal squamous cell carcinoma based on malignant disease history. (A) α-Diversity analysis indicated significant differences between patients with (WMP) and without (NMP) previous malignant disease. (B) β-Diversity using two-dimensional principal coordinate analysis plots shows a clear separation between WMP and NMP. (C) Species-level analysis revealed significant variation between WMP and NMP for unclassified *Oribacterium*.

Table S1 Different taxonomies within malignant disease subgroups

Species	Group comparison	Group 1 mean (%)	Group 2 mean (%)	Difference in mean (%)	95% CI	P
Veillonella dispar	WMP *vs.* NMP	0.51	0.07	0.45	[0.002, 0.89]	0.049*

Group 1 refers to the first group in the "Group comparison" column, and Group 2 refers to the second group. *, P<0.05.

https://dx.doi.org/10.21037/jgo-2025-432

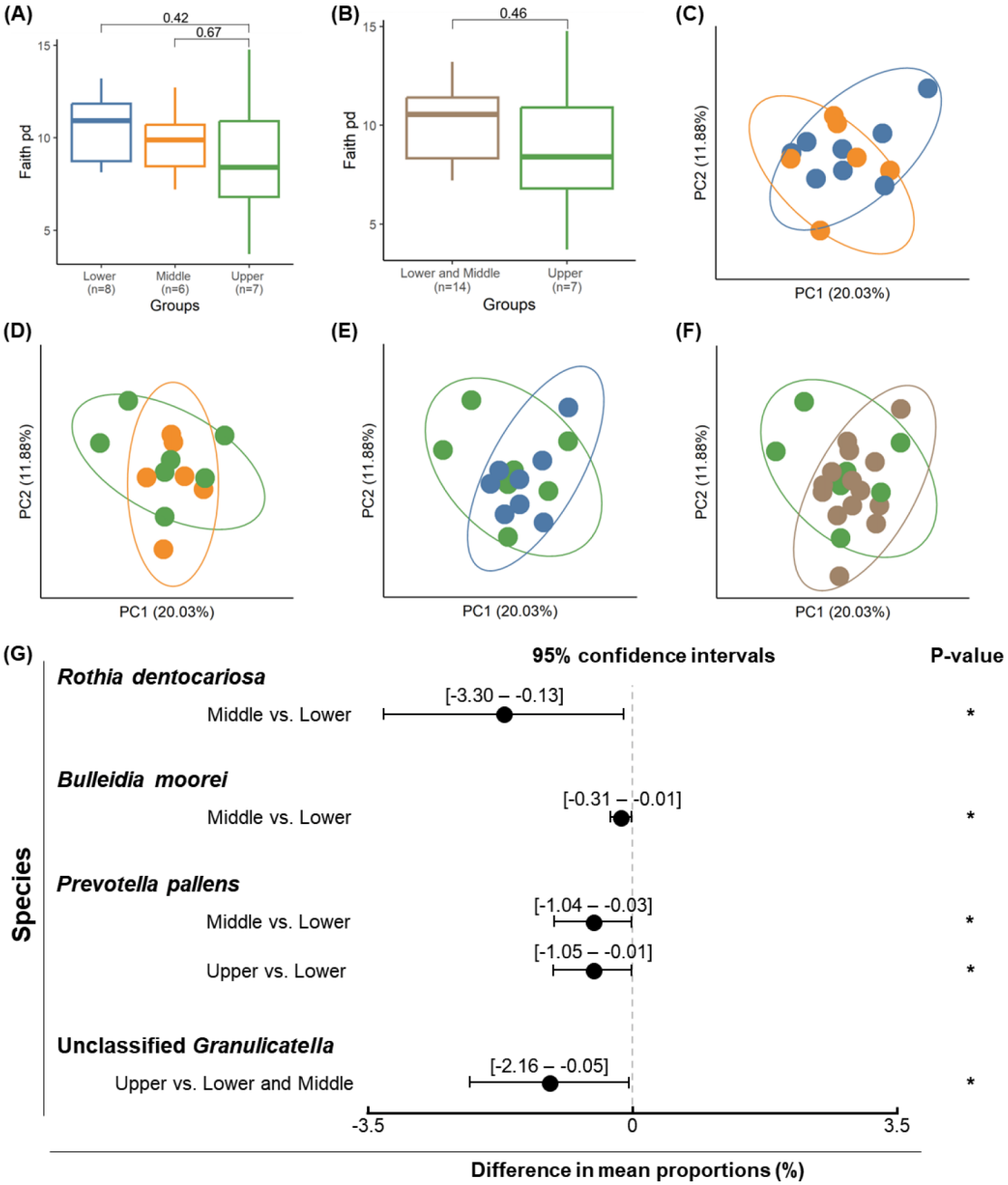

Figure S2 Association between tumor location and oral microbiome in patients with esophageal squamous cell carcinoma. (A) α-Diversity analysis showed no significant differences in microbial composition across tumor location subgroups. (B) α-Diversity analysis, combining the lower and middle subgroups against the upper subgroup, did not show significant differences. β-Diversity analysis using two-dimensional (2D) principal coordinate analysis (PCoA) plots revealed community-level differences between (C) lower versus middle, (D) middle versus upper, and (E) lower versus upper tumor locations. (F) 2D PCoA plots comparing the combined lower and middle subgroup to the upper subgroup showed similar results to Supplementary Figure S2E. (G) Taxonomic analysis showed that the abundance of *Rothia dentocariosa*, *Bulleidia moorei*, and *Prevotella pallens* was significantly different between the middle and lower locations, and *P. pallens* also showed significant differences between the upper and lower locations. A significant difference in the abundance of unclassified *Granulicatella* was identified between the upper and "lower and middle" locations.

https://dx.doi.org/10.21037/jgo-2025-432

Table S2 Different taxonomies within tumor location subgroups

Species	Group comparison	Group 1 mean (%)	Group 2 mean (%)	Difference in mean (%)	95% CI	P
Rothia dentocariosa	Middle *vs.* Lower	0.77	2.49	-1.72	[-3.30, -0.13]	0.036*
Bulleidia moorei	Middle *vs.* Lower	0	0.16	-0.16	[-0.31, -0.01]	0.040*
Prevotella pallens	Middle *vs.* Lower	0.12	0.65	-0.53	[-1.04, -0.03]	0.041*
	Upper *vs.* Lower	0.12	0.65	-0.53	[-1.05, -0.01]	0.046*
Unclassified Granulicatella	Upper *vs.* Lower and Middle	0.63	1.74	-1.1	[-2.16, -0.05]	0.041*

Group 1 refers to the first group in the "Group comparison" column, and Group 2 refers to the second group.

https://dx.doi.org/10.21037/jgo-2025-432