



Fig. S2. Differentially abundant species in biofilm-positive vs biofilm-negative samples. Frozen tissue from USA, MAL1, and MAL2 cohorts were screened by 16S rRNA gene amplicon sequencing and analyzed using Resphera Insight. **a** Biofilm positive CRC and normal flanking tissues demonstrated several significant shifts in the bacterial composition, including enrichment of *Clostridium ramosum*, *Gemmiger formicilis*, and *Dialister invisus*, and depletion of *Clostridium symbiosum*. **b** *B. fragilis*, *F. nucleatum*, *P. micra*, and human oral microbes (HOMD species) were not significantly different between biofilm positive and biofilm negative samples. Random-effects models with 95% CI above or below 0 (red diamonds) were considered statistically significant. Hedge's *g* difference statistic is shown on the X axes. The fixed effects model assumes there exists a single effect size shared by all included studies, while the random effects model allows for variation in the effect size from study to study. Heterogeneity analysis includes estimates of I^2 (percentage of variation reflecting true heterogeneity), τ^2 (random-effects between study variance), and p-value from Cochran's *Q* test for heterogeneity.