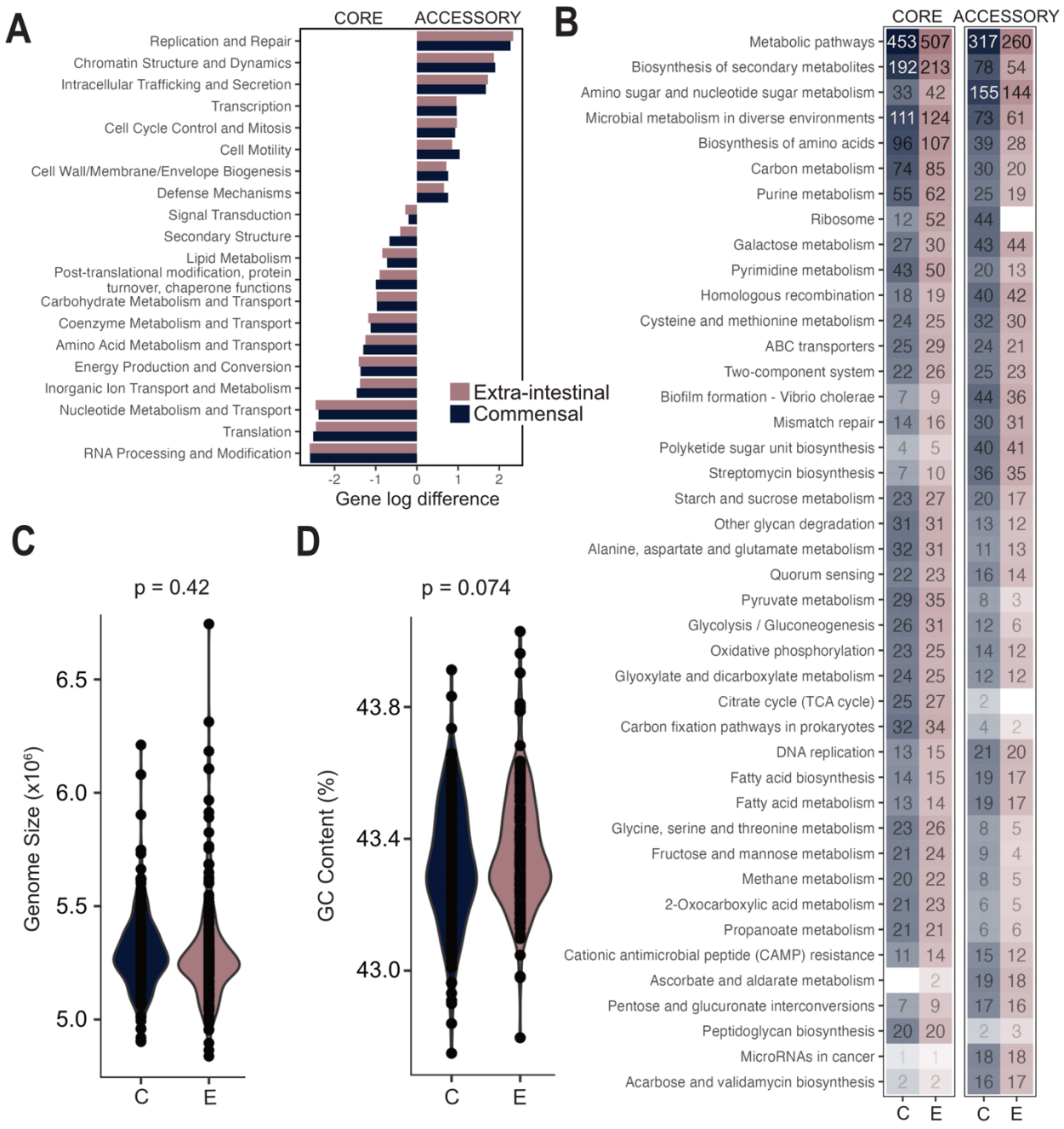


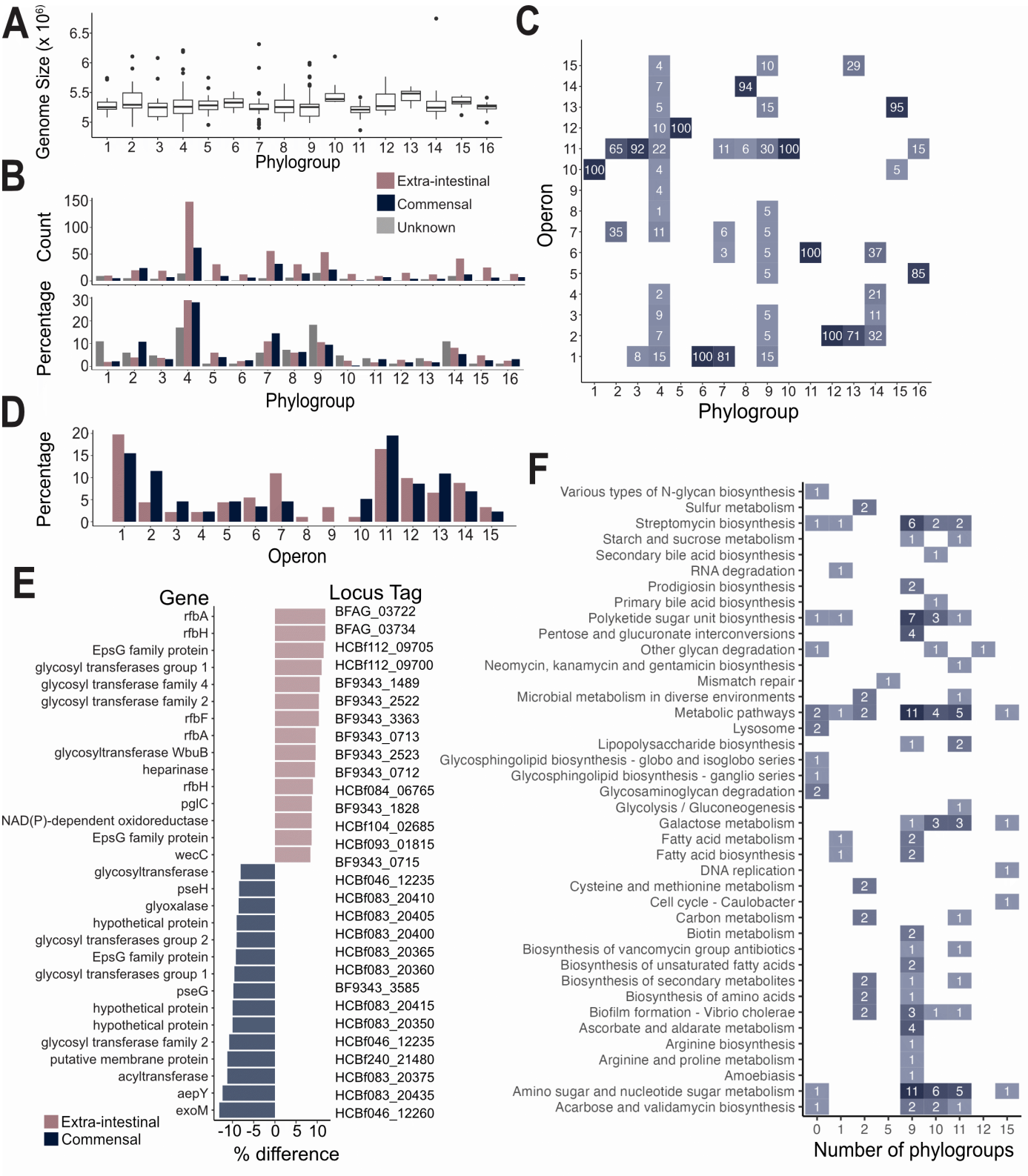
Supplemental Figure 1. Genetic diversity among commensal and extra-intestinal *B. fragilis* isolates

A) Bar graph of $\log_2$ odds ratio between accessory and core genes per COG category separated by commensal (navy) versus extra-intestinal (pink) strains.

B) The top 30 core and top 30 accessory KEGG categories, ordered by most frequently observed KEGG KO category. C, commensal; E, extra-intestinal.

C) Genome size of commensal (C, n=281) and extra-intestinal (E, n=280) isolates, excluding assemblies derived from metagenomic assemblies, Welch's t-test, p=0.42.

D) GC content of commensal (C, n=281) and extra-intestinal (E, n=280) isolates, excluding assemblies derived from metagenomic assemblies, Welch's t-test, p=0.074.



Supplemental Figure 2. *B. fragilis* commensal and extra-intestinal strain distribution among phylogroups and niche specificity traits

A) Genome size of phylogroups 1-16. No statistically significant phylogroups by Welch's t-test at a corrected p-value of 0.05.

B) Distribution of *B. fragilis* isolates by isolation source across phylogroups. No statistically significant difference between isolation sources per phylogroup after chi-squared test with correction for multiple hypothesis testing.

C) Distribution of phylogroups per PSA operon structure expressed in percentage.

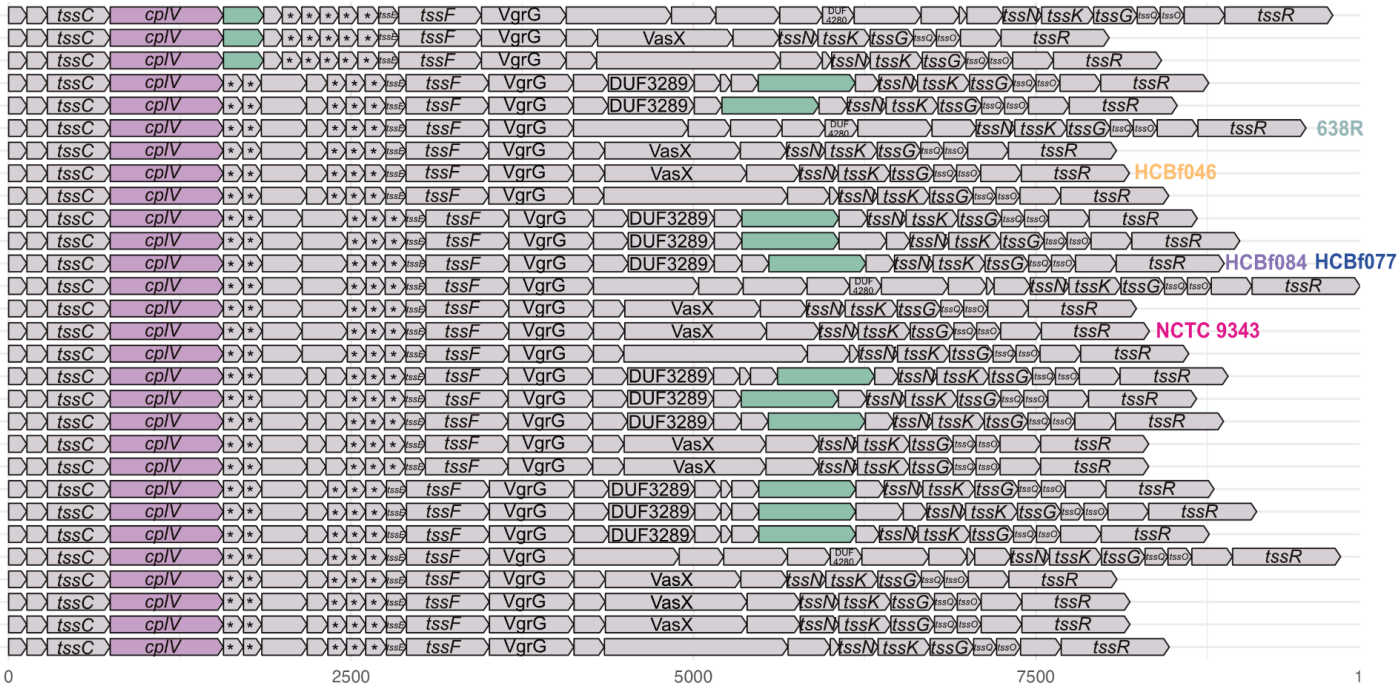
D) Distribution of commensal versus extra-intestinal *B. fragilis* isolates across PSA operons.

E) Top 15 capsular polysaccharide genes which are differentially present in commensal compared with extra-intestinal strains.

F) KEGG categories of genes core to a number of phylogroups (x-axis).

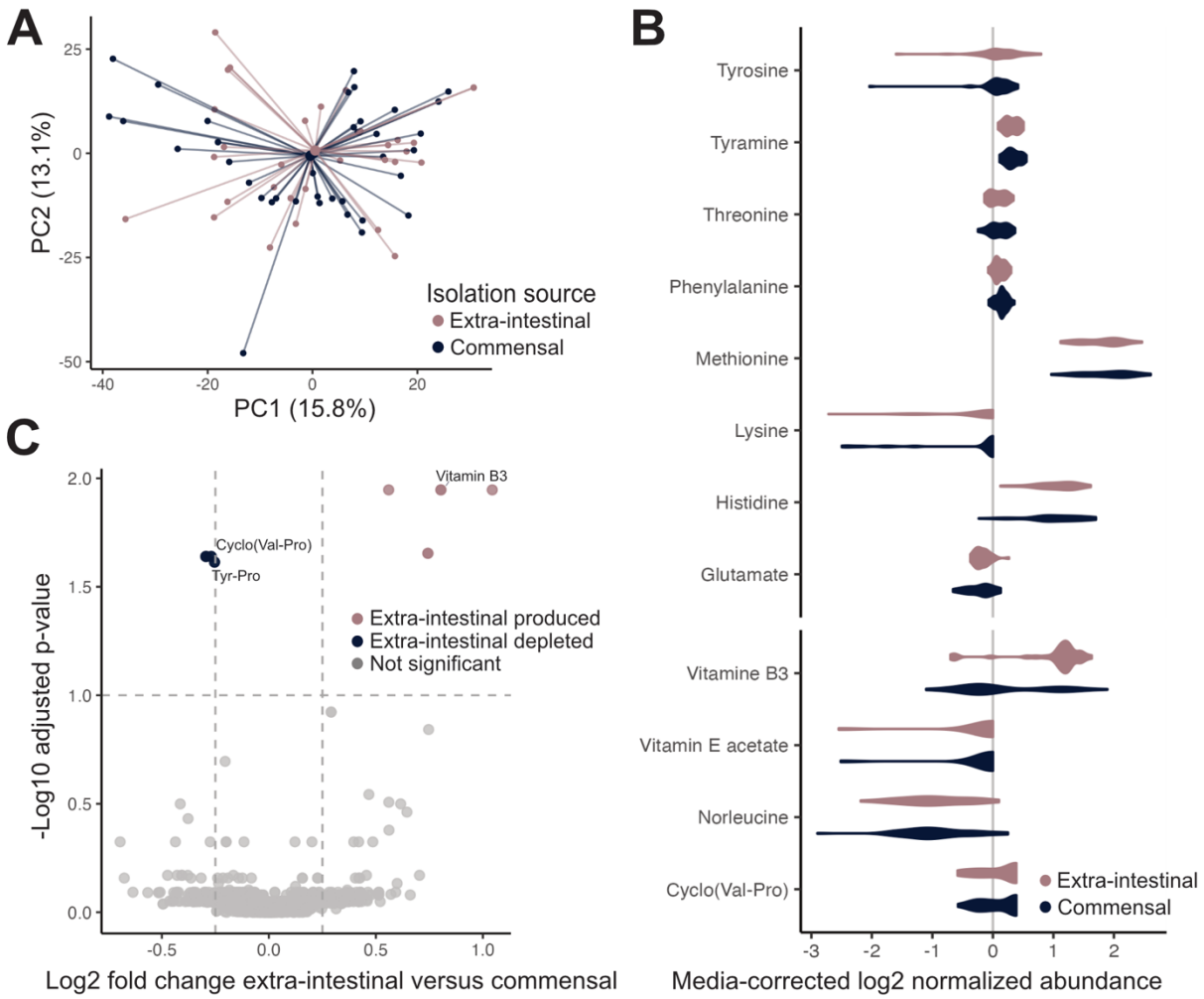
Supplemental Figure 3. Polysaccharide operon structures among *B. fragilis* strains

PSB-H operon structure per all high-quality isolates in the *B. fragilis* pangenome (n=262). Genes are annotated with gene names, if annotated by bakta. Colored by COG category, annotated with information on select strains profiled throughout the study.



Supplemental Figure 4. T6SSiii GA3 operon structures present in *B. fragilis* isolates

T6SSiii GA3 operon structure per all isolates in the *B. fragilis* pangenome (n=813), genes annotated with gene names if annotated by bakta, colored by COG category (grey, unknown; green, cell wall/membrane/envelope biogenesis; pink, replication and repair), annotated with information on phylogroup composition per operon.

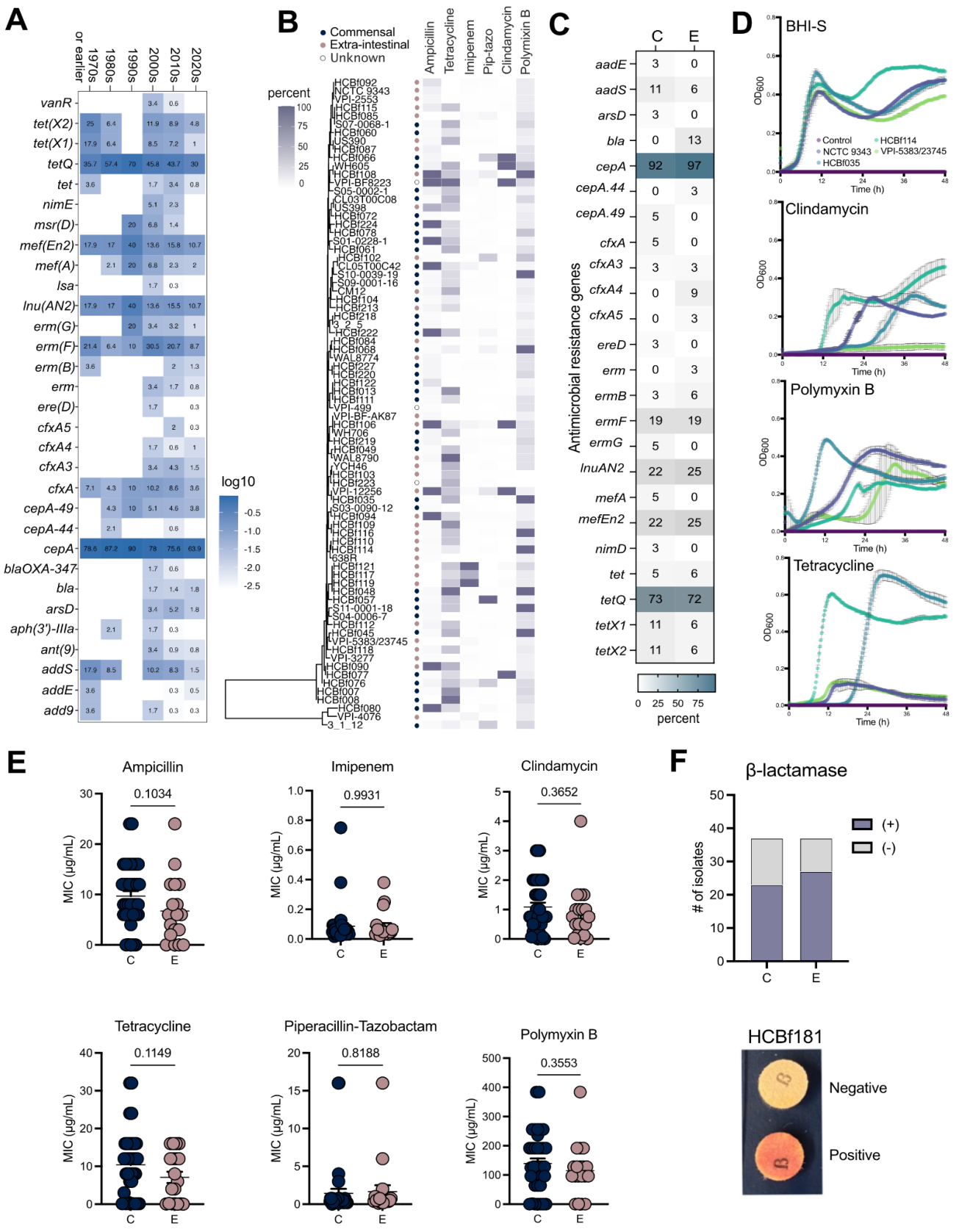


Supplemental Figure 5. Metabolic features among commensal and extra-intestinal *B. fragilis*

A) PCA plot of media-corrected normalized abundances of strains (total, n=69; extra-intestinal=38; commensal=31) connected to the centroid of each group, colored by isolation source, (PERMANOVA, p=0.992).

B) Metabolite production of annotated metabolites across *B. fragilis* strains normalized against the media control (n=73), comparing extra-intestinal (n=33) and commensal (n=40) strains.

C) Volcano plot of differentially abundant metabolites of extra-intestinal (n=33) compared with commensal (n=40) strains. P-values on the y-axis are from a Wilcoxon rank-sum test comparing the average abundance of metabolites across strains versus a media control (dashed lines at Benjamini-Hochberg adjusted p-value ≤ 0.1 and log2-fold change ≥ 0.25).



Supplemental Figure 6. Antimicrobial resistance profile among commensal and pathogenic *B. fragilis*

A) The percentage of antimicrobial gene presence-absence, grouping *B. fragilis* isolates by decade of isolation (1970s or earlier, n=28; 1980s, n=47; 1990s, n=10; 2000s, n=348; 2010s, n=393; 2020s, n=393).

B) Phylogenetic tree of *B. fragilis* strains used in antimicrobial resistance testing with proportion of resistance to ampicillin, tetracycline, imipenem, piperacillin-tazobactam (pip-tazo), clindamycin and polymyxin B.

C) Percent of antimicrobial resistance gene distribution among *B. fragilis* strains (commensal, n=37; extra-intestinal, n=32). Unknowns were excluded from the analysis.

D) Growth rate of representative strains in no antibiotic media control, clindamycin (0.03125 µg/mL), polymyxin B (400 µg/mL), and tetracycline (0.1875 µg/mL).

E) Comparison of antimicrobial minimal inhibitory concentration (MIC, µg/mL) among commensal (C) and extra-intestinal (E) *B. fragilis* strains. Unknowns were excluded from the analysis. Unpaired t-test with Welch's correction.

F) Beta-lactamase production among commensal (C) and extra-intestinal (E) *B. fragilis* strains and representative positive reaction for beta-lactamase production (n=37 per group). p=0.457 from a Fisher's exact test.