Active colonisers of the gastrointestinal tract of Atlantic salmon farmed in a warm water region

Kate Hall[1], Chantelle E Reid[1,2], Qi Zhi Tan[1], Victoria Percival[1], Kamarul Zarkasi[3], Richard S Taylor[2], Andrew Bissett[2], Nicole Ruff[4], Barbara F Nowak[5], John P Bowman[1]*

[1]Centre for Food Safety and Innovation, Tasmanian Institute of Agriculture, University of Tasmania, Hobart, Tasmania 7005, Australia

[2]CSIRO, Hobart, Tasmania 7005, Australia

[3]School of Biological Sciences, Science University Malaysia, Penang, Malaysia

[4]Skretting Australia Pty Ltd, Cambridge, Tasmania 7170, Australia

[5]Institute of Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania 7001, Australia

*Corresponding author

John P. Bowman. Centre for Food Safety and Innovation, Tasmanian Institute of Agriculture, University of Tasmania, Hobart, Tasmania 7005, Australia. Email: john.bowman@utas.edu.au, Telephone: +61 03 6226 6380.

Supplementary figures S1-S5:

Figure S1. Location of farms from which salmon were sampled.
Figure S2. In silico DNA:DNA hybridization between Aliivibrio salmon isolates and other Aliivibrio strains via the GGDC model 2 method
Figure S3. Aligned 16S rRNA variants within the V1-V3 region for available Aliivibrio species including those from this study.
Figure S4. Distribution of Vibrionaceae taxa in Atlantic gut samples and from feed pellets, highlighting spatio-temporal variability and sampling time differences.
Figure S5. Comparison of concatenated cytolethal toxin subunit CdtA, CdtB and CdtC protein sequences from Atlantic salmon Aliivibrio isolates with Gram-negative pathogens possessing Cdt.

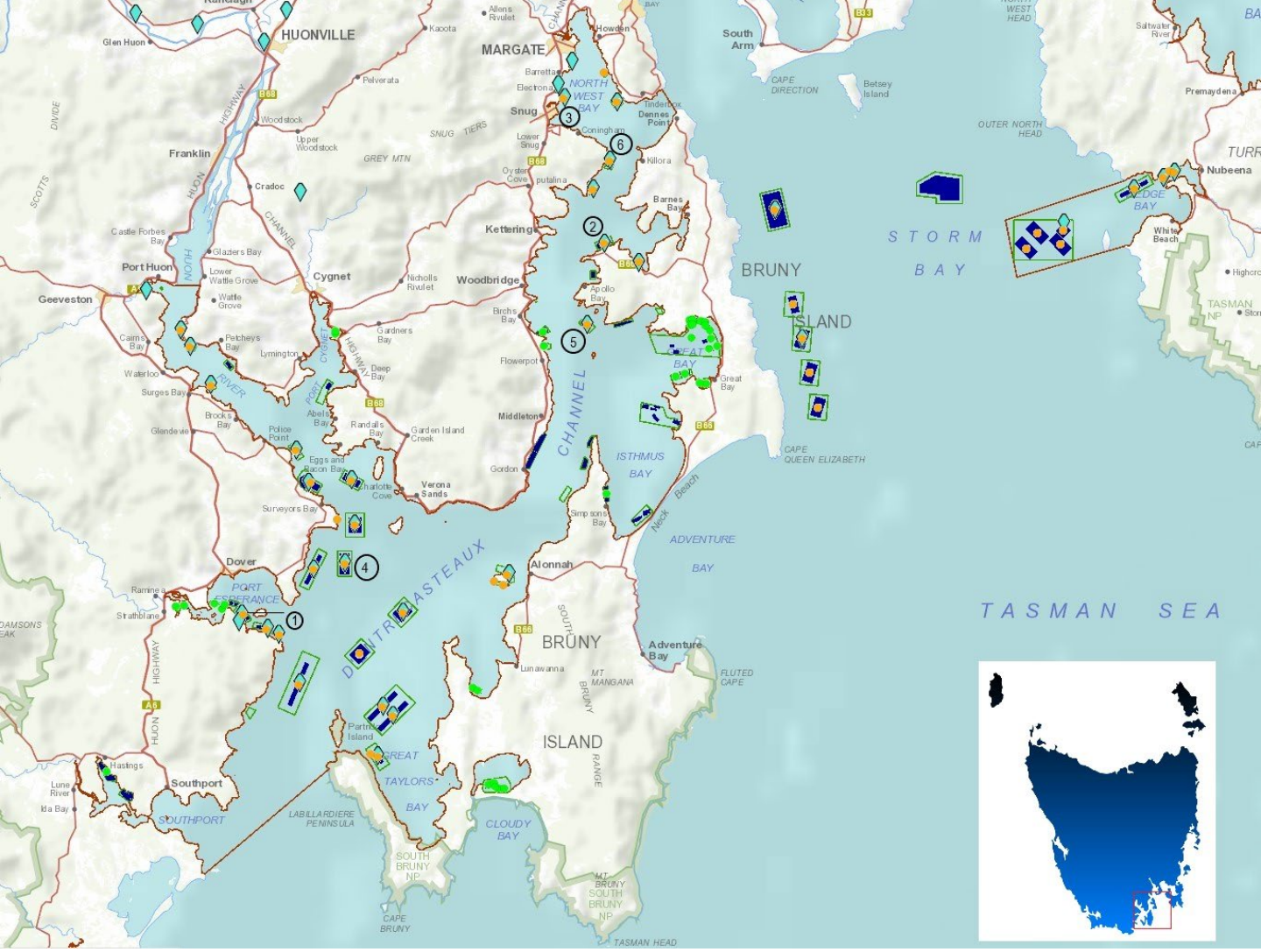

Figure S1. Location of farms from which samples were collected for this study and antecedent published studies as shown in Table 1. Farms sampled were at (1) Meads Creek, Port Esperance; (2) Robert's Point lease, D'Entrecasteaux Channel; (3) North West Bay; (4) Red Cliffs, D'Entrecasteaux Channel; (5) Soldier's Point, D'Entrecasteaux Channel; Sheppard's lease, D'Entrecasteaux Channel (6).

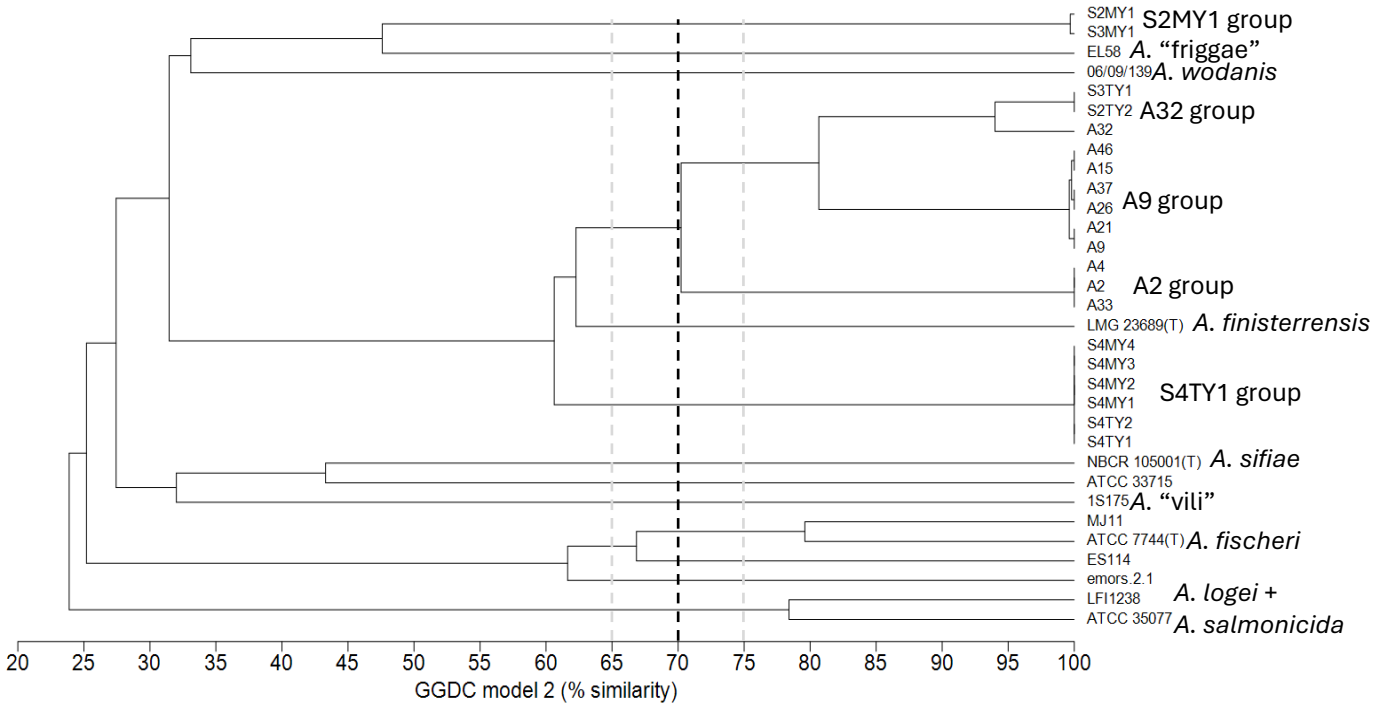

Figure S2. Group average-based tree based on similarities between *Aliivibrio* strain genomes. The similarities were generated with Meier-Kolthoff et al. GGDC model 2, which simulates *in silico* DNA:DNA hybridization. The classic demarcation for species boundaries of 70% is shown as a dashed line. The main features of the plot is that the A2, A9, and A32 groups cluster at the 70% boundary while the type strain of *A. finisterrensis* (LMG 23689^T) clusters with these groups at 58 to 65% similarity. The S4TY1 group clusters with LMG 23689^T at only 54% and 60-62% with the A2/A9/A32 groups. Strains of *A. logei* and *A. salmonicida* (neither are the actual type strains) cluster at 78%, while *A. fischeri* strains cluster at 60 to 80%. Group S2MY1 clusters closest to A. "friggae" EL581 at 43% similarity. The 60% similarity level approximately represents 94-95% average nucleotide identity. As a result, conservatively, the A2/A9/A32, S4TY1 and more obviously S2MY1 strains groups all represent separate species, however *A. fischeri* may not be a single species if this stance is taken. This situation is suggested in the GTDB database (gtdb.ecogenomics.org) where 6 strains (emors.2.1 – the placeholder representative, emors.5.2, emors.5.1o, emors.2.2, emors.5.1t, emors.1.2) are designated as *A. fischeri*_A while the 71 other strains form *A. fischeri sensu stricto*. Strain emors.2.1 (as shown above) has 60% similarity with the *A. fischeri* type strain ATCC 7744^T. GTDB uses an ANI circumscription radius of 95% to define a species though to retain as many species names as possible this boundary ANI values up to 97% are used. Increasingly ANI is used to demarcate bacterial species but this needs to be related to other data due to the fuzzy nature of bacterial species boundaries. A. "vili", A. "friggae" and other species in quotations from the study of Klemetsen and colleagues are placeholder designations but based on the data constitute distinct species.

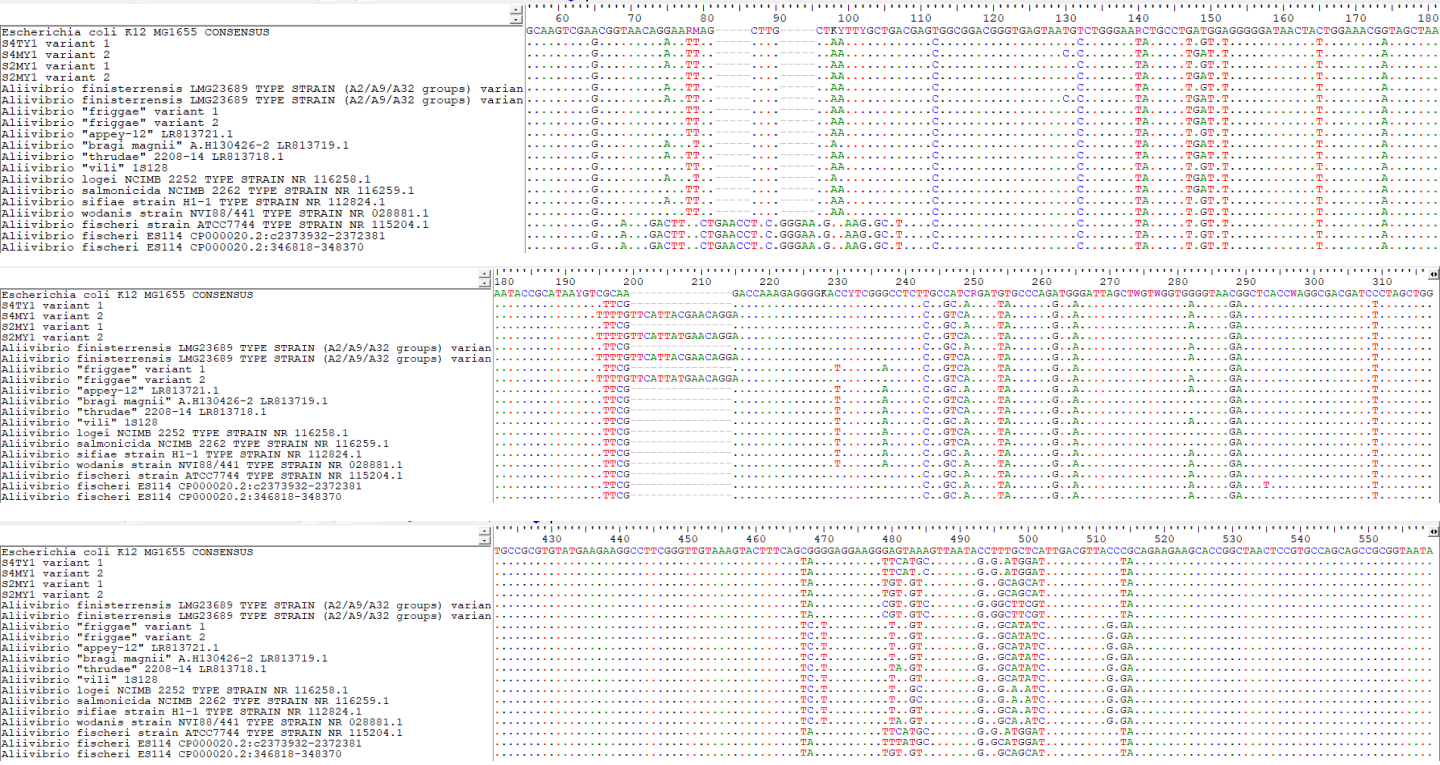

Figure S3. Alignment of sections of the V1-V3 region of 16S rRNA genes covering the equivalent to *E. coli* positions 10 to 534 for *Aliivibrio* species, including Atlantic salmon isolate groups. The two main variants are shown for the three Atlantic salmon gut isolate groups (A. finisterrensis, S2MY1, S4TY1) visualising the regions of hypervariability in the V1-V3 region for these strains and related species. This data resulted in two different variants (491 bp, 506 bp) representing OTUs for *A. finisterrensis*, S2MY1, S4TY1, *A. fischeri* and *A. "friggae"*. These were subsequently added to a localised database to aid in detecting them in the sequence data. Longer (506 bp) variants were not found in the sequence survey data for various species likely due the species being present at low abundance levels and potentially due to biases in joining forward and reverse reads for these longer sequences.

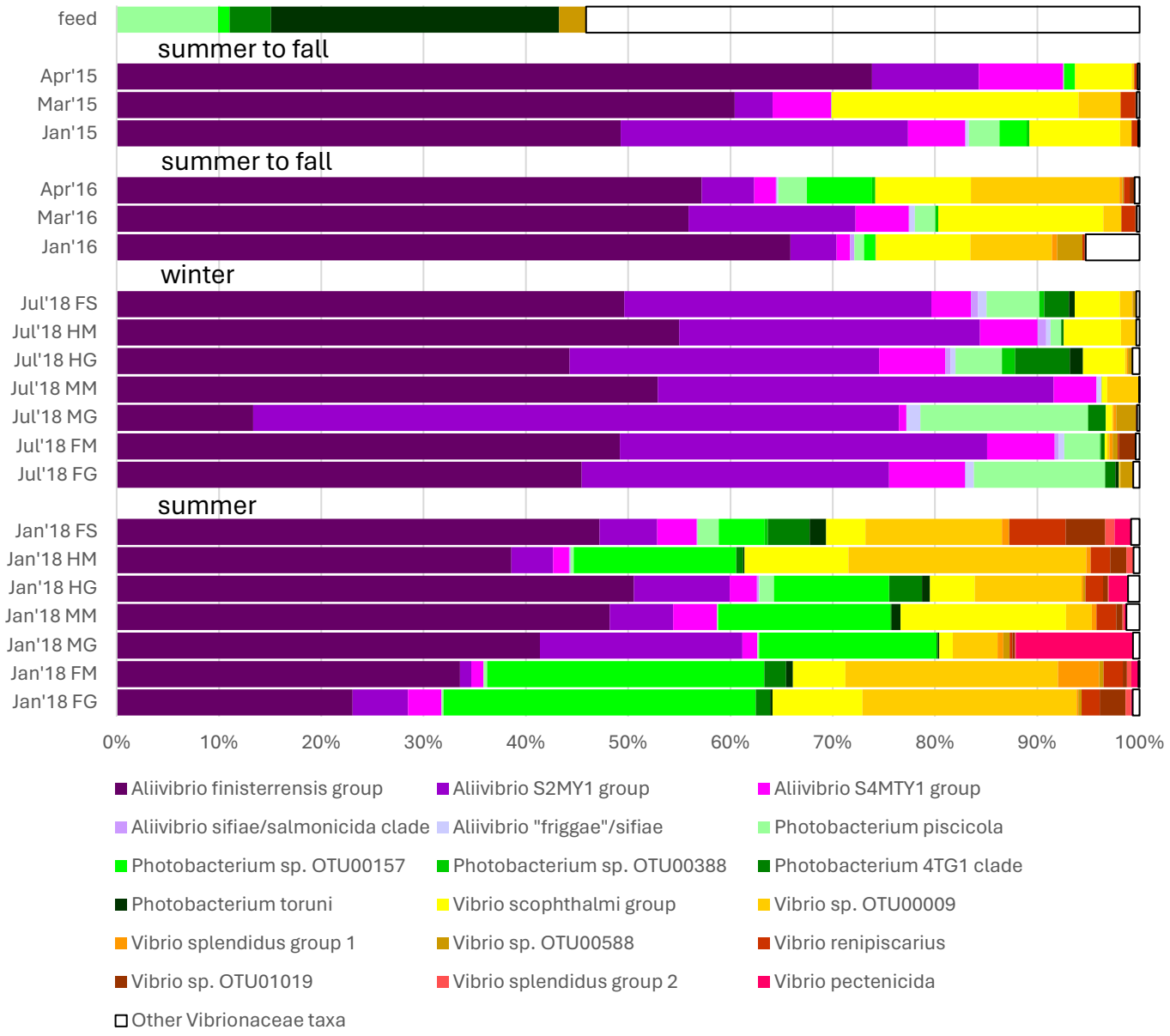

Figure S4. Distribution of Vibrionaceae taxa in Atlantic salmon gut samples collected during three different years from farms in the D'Entrecasteaux Channel, Tasmania. Feed samples came from the 2018 survey. Samples from 2015 and 2016 all were obtained by faecal stripping. 2018 samples were from dissected gut samples – abbreviations (FS, faecal stripped; FG, fore-gut digesta; FM, fore-gut mucosa; MG, mid-gut digesta; MM, mid-gut mucosa; HG, hind-gut digesta; HM, hind-gut mucosa. Only the predominate and prevalent Vibrionaceae taxa are shown. For these OTU designations are arbitrary for the study but when mentioned indicate taxa of unknown species or unresolvable 16S rRNA gene variants from known species. *Aliivibrio*"friggae"/*sifiae* and *A. sifiae/A. salmonicida* OTUs are shown as separate groups for comparison though could belong to only one, either or both species. For other Vibrionaceae taxa these mainly include reads mapping to a range of *Photobacterium* and *Vibrio* species. The noteworthy features for this data is the difference in profiles occurring in summer and winter. Winter (Jul'18) samples are much richer in the *Aliivibrio* S2MY1 group and *P. piscicola*. Summer samples appear to exhibit an increase in the abundance and diversity of *Vibrio* spp. Feed samples include Vibrionaceae DNA though it represents only 3.0% of reads. Most of these reads come from *Photobacteriu*m (2.3%). Of salmon gut microbiome OTUs from feed map to salmon OTUs *P. toruni*, *P. piscicola*, OTU00157, S4TG1 and *Vibrio* sp. OTU00588. The other Vibrionaceae detected in feed map to taxa found either at very low abundances in digesta and mucosa samples or were different to sequences in the localised database.

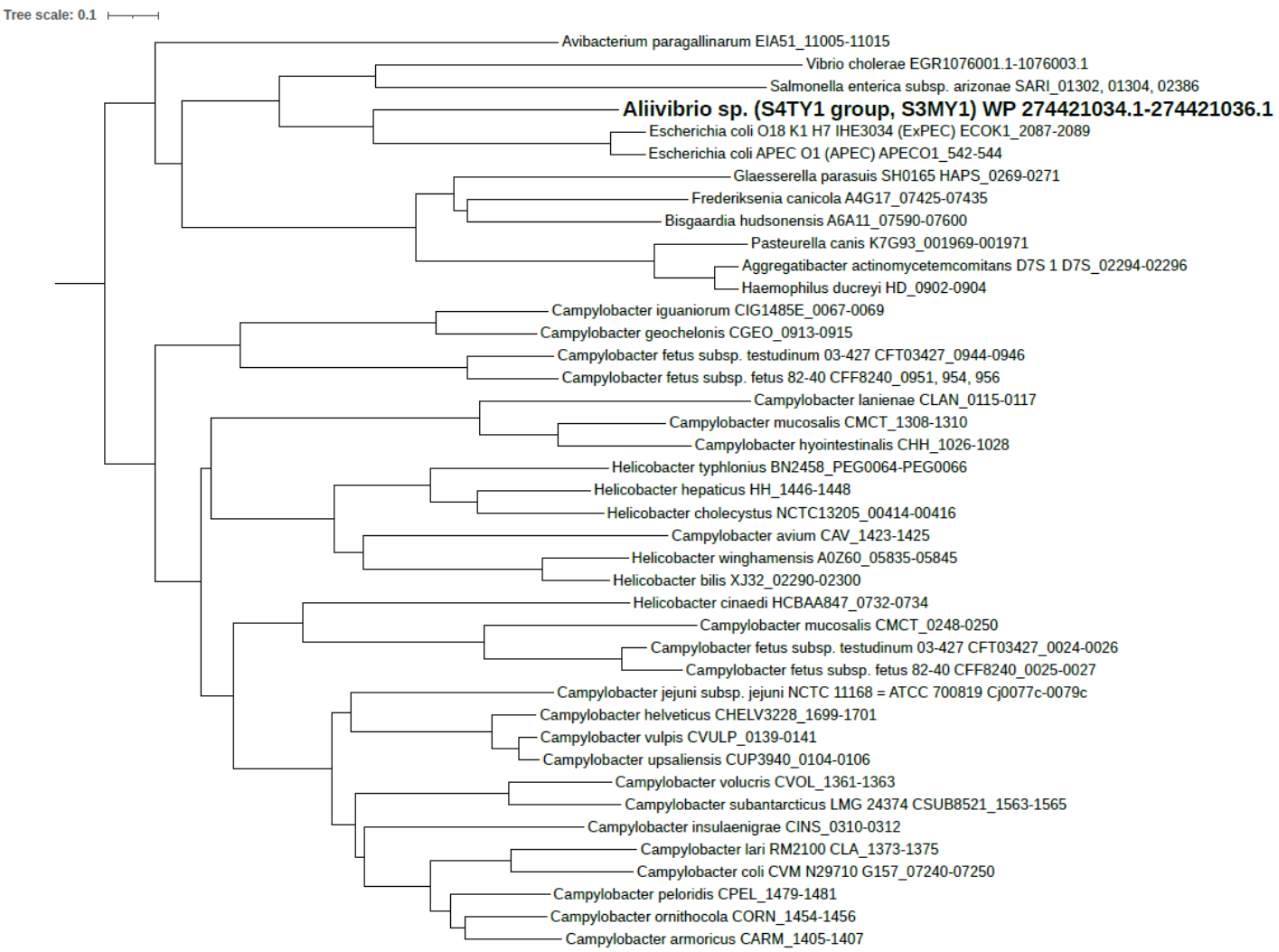

Figure S5. Unweighted group average-based tree based on concatenated cytolethal toxin subunit CdtA, CdtB and CdtC protein sequences. Only bacteria that included an intact operon are shown, mostly belonging to the order Campylobacterales. Genomes of most taxa shown are annotated in detail in the KEGG database. The CdtABC homologs present in *Aliivibrio* strain S3MY1 and strains of the S4TY1 group had identical amino acid sequences. Closest matches (55% identity overall) was to proteins of pathogenic *E. coli* strains (ExPEC – extraintestinal pathogenic E. coli, APEC – avian pathogenic E. coli). The Cdt sequences from a human faecal *Vibrio cholerae* isolate (no. 633012) deposited into NCBI by the England Public Health in 2020 (PM Ashton and colleagues) was more distantly related but in the same cluster with Cdt from *Salmonella enterica*. The expression, function and impact of Cdt in Atlantic salmon is currently unknown. Nor is it know the frequency of *cdt* genes in *Aliivibrio* or other fish-associated bacteria.