**Supplementary Information**

# An ubiquitous antibacterial toxin from human gut bacteria engenders neonatal colonization advantage

Ezequiel Valguarnera[1], Jessica L. Tung[1], Jesse J. Pak[1], Seongmi K. Russell[1], David W. Hecht[2], Drew J. Schwartz[1], and Juliane Bubeck Wardenburg[1]*.

[1]Washington University, Department of Pediatrics, St. Louis, MO 63110, USA.

[2]Loyola University Chicago, Stritch School of Medicine, Department of Medicine, Microbiology, and Immunology, Maywood, IL 60141, USA.

*Lead contact/Correspondence: Juliane Bubeck Wardenburg.

**Email:** jbubeck@wustl.edu

**This file includes:**

Supplementary Methods

Supplementary Figures

Supplementary References

**Supplementary Methods.**

Generation of an NCBI-retrieved BLAST database

For NCBI genomic data analysis, we downloaded 816 *B. fragilis* annotated genomes (combination of whole and partial sequences) available as of April 18th, 2023, generated a local BLAST database on Geneious v2020 and searched the database using the whole set of previously confirmed and newly identified effectors (Supplementary Dataset 1) using our clinical isolate collection.

Generation of a clinical isolate sub-collection and partial genomic assemblies

Patient samples were obtained from the clinical microbiologic laboratory, and only classified based on the culture date, body site, and hospital. A sub-collection of isolates was selected based on the most common *B. fragilis* body isolation sites. 2 ml 96-deep-well plates with 0.6 ml of BHI were inoculated with isolate vials and grown for 18-24 h. 30 µl from each well were used to inoculate a new plate with 1.2 ml of BHI per well and grown for 18-24 h. Each well was split into 0.25 ml of culture mixed with 0.25 ml of BHI 40% to generate frozen stocks, and 0.75 ml of culture that was pelleted for genomic DNA (gDNA) purification. gDNA was purified using a 96 well plate Blood and Tissue DNA Kit (Qiagen) and quantified with a Quant-it dsDNA Assay Kit (ThermoFisher). PCR-based sequencing libraries were generated (iGenomX, Twist Biosciences) and sequenced in an Illumina Hi-Seq 3000 run (Illumina) by the Genome Technology Access Center (GTAC), WUSTL. Short reads for each isolate were assembled *de novo* into contigs using Unicycler v0.4.7 [1], and contigs were annotated using Prokka v1.13.3 [2]. BLAST local databases were generated from annotated contigs in SequenceServer and Geneious (Dotmatics) [3]. Contigs were screened for virulence factors and antibiotic resistances, allowing to discriminate isolates that were contaminated/co-isolated with non-*Bacteroidales*. Non-*B. fragilis* were discriminated via whole proteome comparison entering Fasta aminoacid files for each isolate into Prot-Spam v1.0 and analyzing generated similarity matrixes [4].

Genetic characterization of *B. fragilis* isolates

For determination of T6SS genomic architectures (GAs), GA-specific structural genes (*tssB*, *tssC*, *tssN*) were used as BLAST queries. GA3-positive isolates were screened for known *effector* and *immunity* genes based on previous work [5], and isolates with no identical hits were manually screened for *eff-imm* neighboring genes *hcp2/3/4/5* (*locus I*) and *vgrG/tssN* (*locus II*). Sankey plots of identified GA3 *locus I/II* combinations were generated using SankeyMATIC (https://sankeymatic.com/). All isolates were screened for Acquired Interbacterial Defense (AID) genes, specifically arrangements AID-1 and AID-2 from *B. fragilis* 638R (Dataset S1). Other genes of interest, including *cfiA* (clade I/II

discrimination) [6], *bft1/2/3* and *bsaps* were also searched against the isolate collection to further define groups of clinical isolates to be tested for competition *in vitro* and *in vivo*.

DNA reagents and constructs

pEV1 was generated by linearizing the pWW3452 plasmid [7] with BamHI and inserting the promoter of the *B. thetaiotaomicron rpoD* gene (*bt1311*) via Gibson Assembly using the HiFi DNA Assembly Kit (New England Biolabs). pEV2 was generated in the same manner as pEV1, by BamHI-linearizing the pWW3515 plasmid [7] and inserting the same PCR product as for pEV1. PCR-amplified cloned DNA fragments were generated using Phusion or Q5 DNA polymerases (New England Biolabs). Cloning of pEV1/*locus I ei* constructs was carried out by linearizing pEV1 with XbaI and inserting different *locus I ei* PCR products containing a terminal 3' added BamHI site via Gibson Assembly. Subsequent rounds of cloning of *locus II ei* and accessory protein-coding inserts consisted of BamHI digestion of the preceding construct and ligation with a 3' BamHI-containing PCR product. For constructs that contained Flag-tagged genes, *gfp* was removed from pEV1 derivatives with a BamHI/XhoI digestion, since it was originally cloned in pWW3452 with both 6xHis and 3xFlag tags. Constructs containing *gfp-his-flag* showed a Western Blot band in whole cell lysates that corresponds to monomeric GFP when using anti-Flag and anti-Hcp antibodies (Hcp antigen for polyclonal sera is tagged with a 6xHis, leading to cross-reactivity to the GFP-6xHis band). pKNOCK/Δ*tdk* construct was generated by digesting pKNOCK with EcoRV and inserting two 1000 bp fragments flanking the thymidine kinase (*tdk*) coding sequence via Gibson Assembly. pExchange-*tdk*/Δ*locus I*, Δ*locus II*, Δl*ocus II-paar* constructs for deletions were generated by digesting pExchange-*tdk* [8] with EcoRV and inserting two 1000 bp fragments flanking the desired locus for deletion via Gibson Assembly. For constructs pExchange-*tdk*/Δ*vgrG* and pExchange-*tdk*/Δ*hcp23*, pExchange-*tdk* was digested with BamHI. pSIE1/Δ*tssC* construct for *tssC* deletions was generated by inverse PCR amplification of the pSIE1 backbone[9] using primers pSIE1_F and pSIE1_R, and insertion of two 1000 bp fragments flanking the *tssC* locus from strain NCTC 9343 via Gibson Assembly. For the generation of **pK**NOCK-***b****la*-***e****rmGb*-***s****acB* (pK-*bes*), pKNOCK-*bla-ermGb* was digested with KpnI and ligated via Gibson Assembly with PCR-amplified *rpoD* promoter and *sacB* gene fragments. *rpoD* promoter + RBS sequence was amplified from pEV1/*bte1* as template, *sacB* was amplified from pCVD442, a suicide vector for allelic exchange in *E. coli* hosts ($Amp^R$) [10]. For pK-*bes* deletion constructs, vector was digested with BamHI and ligated to two 1000 bp fragments flanking *bte1/bti1*, *bte3/bti3*, or *bte4/bti4*, amplified from isolate 13051 gDNA. pET28b cloning of *hcp1* for expression and purification of Hcp-6xHis was generated by ligating NcoI/XhoI-digested pET28b and NcoI/XhoI-digested *hcp1* PCR product without a stop codon to obtain a C-terminal 6xHis-tagged Hcp construct. For inducible expression of *bte3*, *bte3*-HA was PCR amplified from pEV1/*bte1_3*-HA and PCR fragments of *bte3* were ligated to pET22b digested with NdeI/XhoI or NcoI/XhoI for cytosolic or periplasmic

expression, respectively. For inducible expression of *Ecbte3*, *Ecbte3* was PCR amplified from *E. coli* Nissle 1917 gDNA and PCR fragments of *Ecbte3* were ligated to pET22b digested with NdeI/XhoI or NcoI/XhoI for cytosolic or periplasmic expression, respectively. For expression of *bti3* in *S*. Typhimurium and *V. cholerae*, pEXT20 was digested with BamHI and ligated to PCR-amplified *bti3* via Gibson Assembly. For expression of *bti3* in *S*. Typhimurium and *V. cholerae*, *bti3* was PCR amplified from ATCC 43858 gDNA and ligated via Gibson Assembly with pEXT20 digested with BamHI [11].

Generation of clean deletion mutants in *B. fragilis* NCTC 9343 (N1)

*B. fragilis* N1 strain was chosen as the parental strain for our study because it is genetically tractable, it carries an active GA3 alone and no AIDs that protect from GA3 effectors. N1 is not intrinsically resistant to clindamycin or rifampicin. Firstly, we generated a clean deletion of the thymidine kinase gene (Δ*tdk*) as a parental strain for all subsequent deletions. This mutation allows for growth in the presence of fluorodeoxyuridine (FUdR), providing a negative selection system [8]. We conjugated an *E. coli* S17-1 strain carrying pKNOCK/Δ*tdk* to the *wild type* N1 strain overnight in aerobiosis and plated for N1 exconjugants in BHI Gm Cln plates. Cln-resistant clones were back-diluted in BHI with no antibiotics for at least five daily passages, and plated in BHI 200 ug/ml FuDR to screen for *tdk* deletion mutants. Clones were tested to be Cln-sensitive. Deletion was confirmed by PCR. For subsequent N1 deletion mutants, we conjugated the *E. coli* S17-1 pExchange-*tdk* strain of choice with the parental N1 Δ*tdk* or Δ*tdk*-derivative strain. Cln-resistant clones were back-diluted 1:2000 in BHI with no antibiotics for at least five daily passages and plated in BHI 200 ug/ml FUdR to screen for required mutants. Clones were tested for Cln-sensitivity. Deletion was confirmed by PCR.

Generation of clean deletion mutants in *B. fragilis* clinical isolates

For deletion of *tssC* in clinical isolates 10765 and 13051, pSIE1/Δ*tssC* construct was conjugated via *E. coli* S17-1 strain. Cln-resistant clones were back-diluted in BHI with no antibiotics and plated in 100 ng/ml anhydrotetracycline (aTc) to screen for required mutants [9]. Clones were tested for Cln-sensitivity. Deletion was confirmed by PCR. For deletion of *bte1*, *bte3*, and *bte4* in 13051 isolate, we developed a negative selection allelic exchange vector that relies on growth inhibition by sucrose upon expression of the *B. subtillis sacB* gene that encodes the levansucrase enzyme. Levansucrase-based negative selection relies on the enzymatic reaction that converts intracellular sucrose into the toxic polysaccharide levan, thus has been widely used for allelic exchange in Gram-positive and Gram-negative bacteria[12,13]. pK-*bes*/Δ*bte1*, pK-*bes*/Δ*bte3* and pK-*bes*/Δ*bte4* constructs were conjugated and positively selected for genomic integration ($Cln^R$) as described above. Merodiploid clones were grown in BHI with no antibiotics for 16-24 hours, diluted in a range of 10-fold dilutions ($1.10^{-1}$ – $1.10^{-5}$) and plated in BHI 1.6% Sucrose. Plates were incubated anaerobically at 37°C for 24-48 hours, and sucrose-

resistant ($Suc^R$) clones were tested for Cln sensitivity ($Cln^S$). $Cln^S$ $Suc^R$ clones were tested for deletion by PCR. Note: Because *sacB* has evolved in free-living ubiquitous bacteria, selection by *sacB* at 37°C does not have full penetrance. Thus colony overgrowth must be avoided to minimize background growth of meridiploid clones. Alternatively, for BHI 1.6% Sucrose plates, anaerobic growth for 24 hours at 37°C followed by 24 hours at room temperature will minimize $Cln^R$ background growth.

Complementation of *B. fragilis* strains

To generate complemented *B. fragilis* strains we conjugated the *E. coli* S17-1 strain carrying pEV1/pEV1-derivatives (cloned *effector-immunity* genes and accessory protein-coding genes) or pEV2/pEV2-derivatives (cloned *immunity* genes) of choice with the required *B. fragilis* strain. exconjugants were selected in BHI Gm Cln plates. Recipient strains for pEV2 were selected for Rif resistance by plating stationary phase liquid cultures in BHI Rif for 48 hours and selecting for colonies. For *B. thetaiotaomicron*, *B. ovatus*, and *P. vulgatus*, strains were conjugated with an *E. coli* S17-1 strain containing pFD340-Tet [14] and exconjugants were selected in BHI Gm Cln plates.

Hcp antisera production

pET28b/Hcp was transformed into *E. coli* BL21 cells, transformed cells were used to inoculate an overnight starter culture in LB. 1 liter cultures were inoculated 1:50 with the overnight culture and when OD600 = 0.4-0.7, *hcp-6xHis* expression was induced with 0.3 mM IPTG for 5 hours at 37°C. Cell pellets were resuspended in 50 mM HEPES 150 mM NaCl with protease inhibitors (cOmplete EDTA-free, Roche) and lysed using a continuous flow cell disruptor at 30 kpsi (Constant Systems), and centrifuged at 15000 x g to remove cell debris. Cleared lysates were incubated with Agarose-$Ni^{2+}$ resin and purified in a gravity column using Lysis buffer with 30 mM imidazole for washes and 250 mM for elution. Purified fractions were buffer exchanged to lysis buffer and used for intradermal injection into rabbits for a 70-day immunization scheme (Pocono Rabbit Farm and Laboratories, PA, USA). Total serum was tested for Hcp-specific signal by dot-blot and western blot.

SDS-PAGE and Western Blots

All protein samples were analyzed by in-house SDS-PAGE and transferred to 0.45 µm pore size nitrocellulose membranes using a semi-dry transfer apparatus (Bio-Rad Laboratories). Gel percentages were 15% for Hcp/RpoA analysis and 10% for T6 secretion complex proteins analyses (Bte3/Bfe3-HA, VgrG/Ate3/Pte3-Flag). In Figure S7, we used 15% bottom half/10% top half SDS-PAGE gels. Membranes were blocked in 3% skim milk in Tris-HCl 50 mM NaCl 150 mM pH 7.6 (TBS). Antibodies were incubated in blocking buffer with 0.1% Tween-20. We used TBS 0.1% Tween-20 (TBS-T) for

membrane washes. Antisera dilutions: anti-Hcp 1:2000 (rabbit polyclonal, in-house); anti-*E. coli* RpoA 1:2000 (mouse, clone 4RA2, Biolegend); anti-HA 1:2000 (rabbit, clone C29F4, Cell Signaling); anti-Flag 1:2000 (mouse, clone M2, Sigma); anti-Fpn 1:2000 (rabbit polyclonal, in-house); anti-BFT 1:2000 (rabbit polyclonal, in-house). For improved visualization of RpoA signal from *B. fragilis* using anti-*E. coli* RpoA, primary antibody was incubated overnight at 4°C.

Immunoprecipitations

100 ml of *B. fragilis* log phase cultures were pelleted and resuspended in 20 ml of TBS with protease inhibitors. Cells were disrupted at 30kpsi, centrifuged at 12000 x g for 20 min at 4°C, and supernatants were filtered using a 0.2 um filter and split into half. 10 ml of lysates were mixed with 40 µl of anti-Flag resin (Sigma), incubated for 3 h, and proteins were eluted with 40 µl of 2x non-reducing Laemmli buffer. 10 ml of lysates were incubated with anti-HA 1:1000 for 3 h at 4°C, 40 µl of protein A Sepharose resin was added and incubated for 3 h at 4°C. Proteins were eluted with 40 µl of 2x non-reducing Laemmli buffer. For mass spectrometry analysis of anti-Flag immunoprecipitations, Flag-tagged proteins were eluted by competition incubating the resin with 20 µg of 3xFlag peptide (Sigma) for 30 minutes at 4°C to minimize release of antibodies from the anti-Flag resin into the eluate fraction. Eluates were analyzed by SDS-PAGE and western blot and precipitated by TCA.

Mass spectrometry analysis

50 µl of 50 mM ammonium bicarbonate with 10% acetonitrile were added to the dry tubes containing the TCA-precipitated protein and gently vortexed. 10 µl (20 ng/µl) of modified sequencing-grade trypsin (Promega) was spiked into the solutions and the samples were placed overnight at 37ºC. Samples were acidified by spiking in 5 µl 20% formic acid solution and then desalted by STAGE tip. On the day of analysis, samples were reconstituted in 10 µl of HPLC solvent A. A nano-scale reverse-phase HPLC capillary column was created by packing 2.6 µm C18 spherical silica beads into a fused silica capillary (100 µm inner diameter x ~30 cm length) with a flame-drawn tip. After equilibrating the column each sample was loaded via a Famos auto sampler (LC Packings, San Francisco CA) onto the column. A gradient was formed and peptides were eluted with increasing concentrations of solvent B (97.5% acetonitrile, 0.1% formic acid). As peptides eluted, they were subjected to electrospray ionization and then entered into an LTQ Orbitrap Velos Elite ion-trap mass spectrometer (Thermo Fisher Scientific). Peptides were detected, isolated, and fragmented to produce a tandem mass spectrum of specific fragment ions for each peptide. Peptide sequences (and hence protein identity) were determined by matching protein databases with the acquired fragmentation pattern by the software program, Sequest[15] (Thermo Fisher Scientific). All databases include a reversed version of all the sequences and the data was filtered to between a one and two percent peptide false discovery rate.

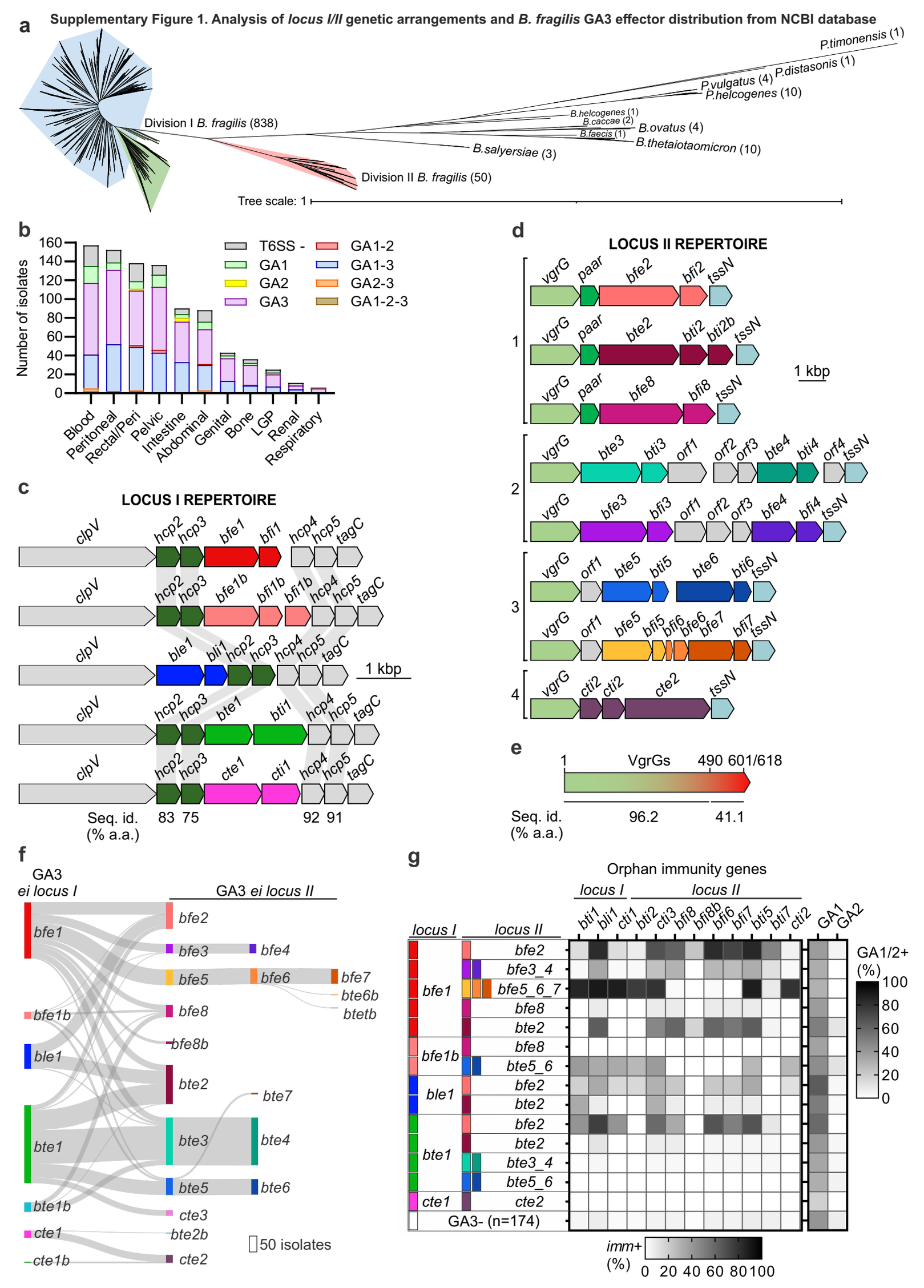
Selection of clinical isolates for T6-positive versus T6-positive competitions

We generated a list with all the potential NTBF and ETBF strains that shared a common GA3 *locus I/II* combination. All candidate strains carried a GA3 alone, no AIDs, and either no T6-independent toxins (BSAPs), or strain pairs were also matched following shared T6-independent toxin criteria. All candidates were tested for clindamycin sensitivity and ETBF candidates were tested for rifampicin sensitivity. Antibiotic-sensitive strains were tested for genetic tractability by conjugation with an *E. coli* S17-1 strain carrying pEV1 and screened for clindamycin resistant clones. Rif-resistant clones were derived from genetically-tractable ETBF strains and conjugated with an *E. coli* S17-1 strain carrying pEV2 or pEV2 with immunity protein-coding genes. Genetically-tractable NTBF strains were conjugated with an *E. coli* S17-1 strain carrying pEV1 or pEV1 with *locus I/II* combinations.

Tissue collection, preparation, and analysis

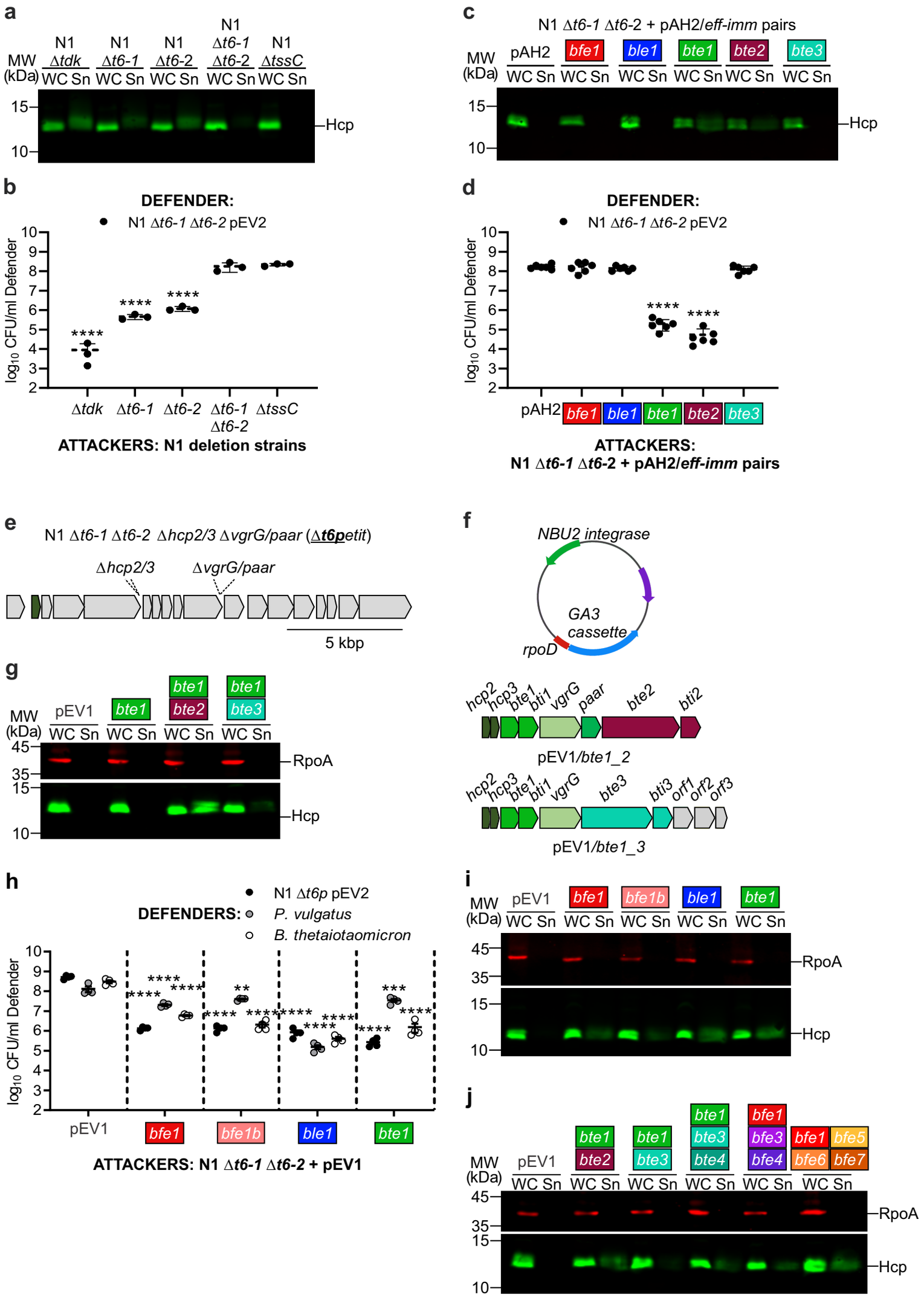
Colons were harvested from euthanized animals, photographed for gross pathology and length analyses, and then ceca were dissected for tissue mass and CFU analyses. Cecum sections were scraped with blunt scissors to remove excess material and mucus and washed through three Petri dishes with sterile PBS before being gently blotted of excess PBS with sterile gauze and placed into pre-weighed sterile 2 mL screw-cap tubes containing 1 mL of sterile PBS and a 1:1 mix of 2 mm and 1 mm beads. Homogenization was performed three times in an MP Biomedicals FastPrep 24 at 6 m/s for 30 s at 4°C, with a five-minute cool-down between homogenizations.

**Supplementary Figure 1. Analysis of *locus I/II* genetic arrangements and *B. fragilis* GA3 effector distribution from NCBI database**



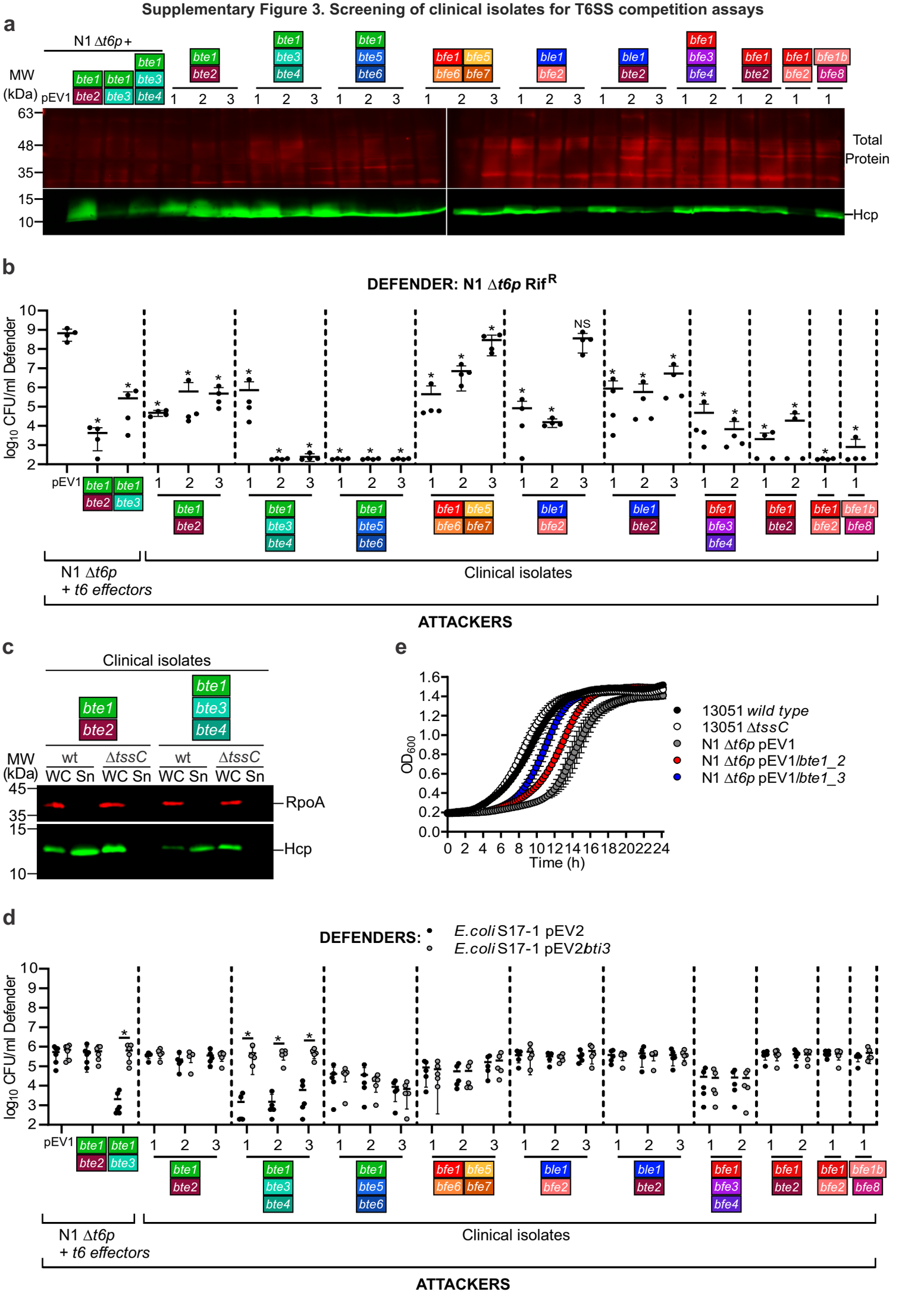
**Supplementary Figure 1. Analysis of *locus I/II* genetic arrangements and *B. fragilis* GA3 effector distribution from NCBI database. Related to Figure 1. a.** Phylogeny of 919 isolates and 8 control strains (*P. vulgatus* ATCC 8482, *B. thetaiotaomicron* VPI 5482, *B. fragilis* strains NCTC 9343, 638R, YCH46, ATCC 43858, ATCC 43859, HMW 616). Non-*B. fragilis* isolates were annotated via GTDB-Tk. Tree was visualized as unrooted using iTOL. **b.** Distribution of 882 *B. fragilis* isolates based on body isolation site and T6SS genomic architecture (GA) presence. LGP: Liver/Gallbladder/Pancreas. **c.** Schematic of most common *locus I* genetic arrangements. Gray-colored lines represent synteny of *hcp2*, *hcp3*, *hcp4* and *hcp5* genes. Bottom; aminoacidic sequence alignment of Hcp2, Hcp3, Hcp4 and Hcp5 corresponding to *locus II* arrangements **d.** Schematic of most common *locus II* genetic arrangements. **e.** Amino acid sequence alignment of VgrGs corresponding to *locus II* arrangements depicted in (**d**). Sequence length of different VgrG variants range between 601 and 618 amino acids (arrangement 1: 616 aa; arrangement 2: 601 aa; arrangement 3: 618 aa; arrangement 4: 610 aa). **f.** Sankey plot indicating proposed nomenclature and distribution of GA3 *ei* gene combinations across 557 GA3+ *B.fragilis* genomic assemblies retrieved from NCBI (637 T6SS+/816 genomic assemblies). **g.** Distribution of orphan immunity genes and genomic architectures 1 and 2 (GA1 and GA2) within *B. fragilis* clinical isolates. For orphan immunity genes, heatmap represents prevalence of orthologs of *locus I* and *locus II* cognate *imm* genes throughout isolates with different GA3 *ei* [39]. For GA1/GA2, heatmap represents prevalence of GA1 and GA2 throughout isolates with different GA3 *ei*.

**Supplementary Figure 2. Generation of a T6SS-defficient strain and genomic insertions for *effector/immunity* engineering**



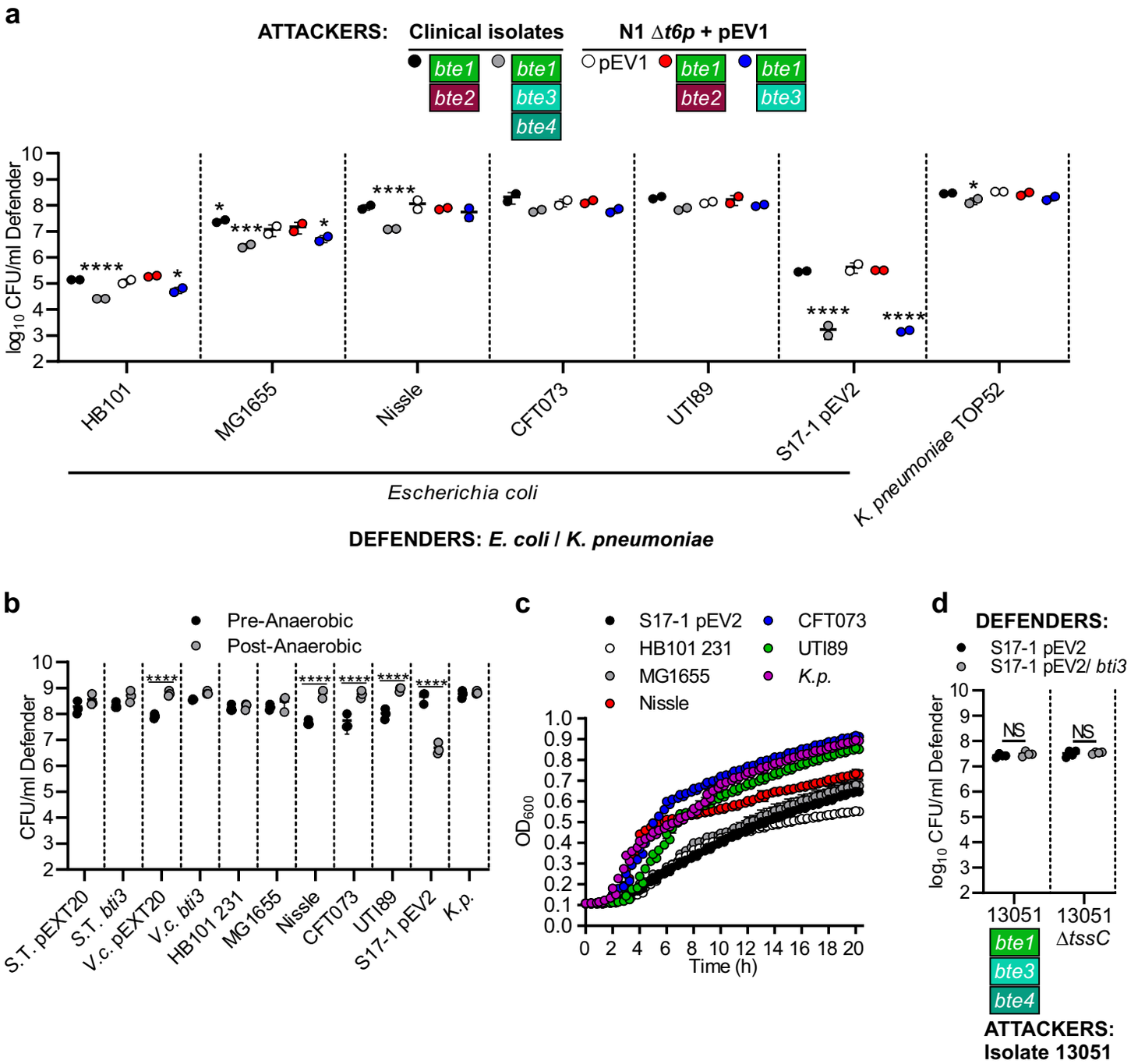
**Supplementary Figure 2. Generation of a T6SS-defficient strain and genomic insertions for *effector/immunity* engineering. Related to Figure 2. a, b.** Parental strain N1 Δ*tdk* and mutant strains Δ*t6-1*, Δ*t6-1*, Δ*t6-1* Δ*t6-2* and Δ*tssC* were analyzed by Western blot for secretion of Hcp (**a**) and used for *in vitro* competition assays against the Δ*t6-1* Δ*t6-2* strain (**b**). ****p value <0.0001; one-way ANOVA with Sidak's test vs Δ*tssC* strain. n=3. **c, d.** N1 Δ*t6-1* Δ*t6-2* complemented with pAH2 derivatives were analyzed by Western blot for secretion of Hcp (**c**) and used for *in vitro* competition assays against the Δ*t6-1* Δ*t6-2* strain (**d**). ****p value <0.0001; one-way ANOVA with Sidak's test vs pAH2-complemented strain. n=6. **e.** N1 Δ*t6petit* (Δ*t6p*) mutant strain was generated by additional deletions of *hcp2*, *hcp3*, *vgrG*, and *paar* genes on the Δ*t6-1* Δ*t6-2* background. **f.** Schematic of representative pNBU2-based insertions (pEV1) containing *btei1_2* or *btei1_3*. Constructs were derived from pWW3452, an insertion vector that contains *gfp* driven by a strong phage promoter[43]. Insertion of a *B. thetaiotaomicron rpoD* (*bt1311*) promoter upstream of *gfp* promoter generated pEV1. **g.** Western blot of whole cells and supernatants of N1 Δ*t6p* complemented with pEV1-derived constructs. **h, i.** *B. fragilis* N1 Δ*t6-1* Δ*t6-2* was complemented with pEV1 containing different *locus I* effectors. Strains were used for *in vitro* competitions against N1 Δ*t6p, B. thetaiotaomicron* and *P. vulgatus* (**h**) and analyzed by Western blot for secretion of Hcp (**i**). **p value =0.003, ***p value =0.0006, ****p value <0.0001; two-way ANOVA with Sidak's test of each attacker vs N1 Δ*t6-1* Δ*t6-2* pEV1, per *Bacteroidales* target strain. n=4. **j.** Western blot of N1 Δ*t6p* complemented with pEV1 containing *locus I/II* effector combinations. RpoA: RNA Polymerase A, cytosolic control. For all competition assays, data are presented as mean ± SD for n = independent competition spots.

**Supplementary Figure 3. Screening of clinical isolates for T6SS competition assays**



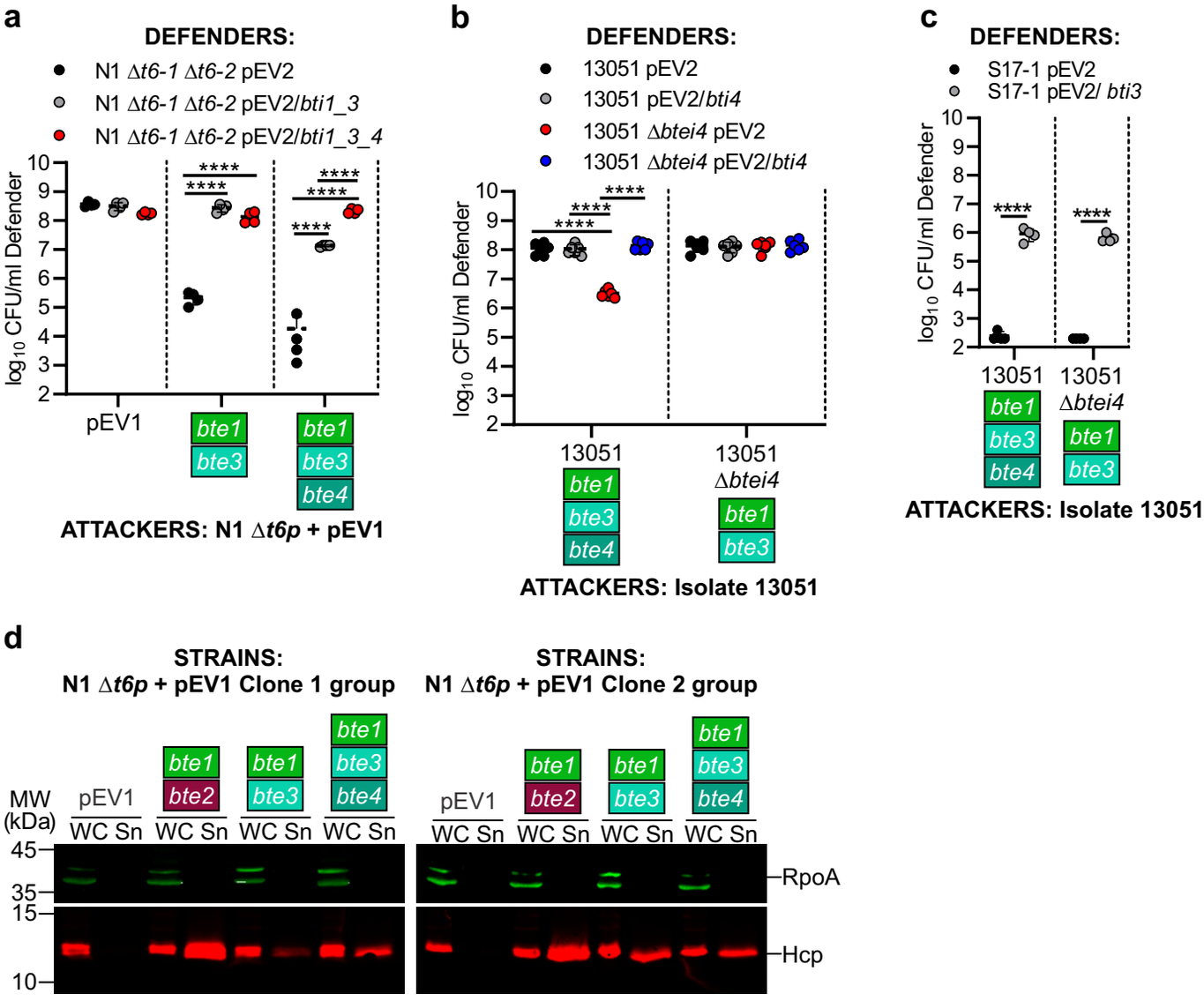
**Supplementary Figure 3. Screening of clinical isolates for T6SS competition assays. Related to Figure 2. a.** Anti-Hcp western blots of cell supernatants from N1 Δ*t6p* strains and clinical isolates representative of the 10 most abundant *locus I/II* combinations. REVERT: Total protein stain. **b, d**. *In vitro* competitions of N1 Δ*t6p* strains and clinical isolates against (**b**) *B. fragilis* NCTC 9343 (N1) Δ*t6p* strain, or (**d**) *E. coli* S17-1 with pEV2 (closed circles) or pEV2/*bti3* (gray circles). In (**b**) *p value <0.0001; one-way ANOVA with Dunnett's test vs. N1 Δ*t6p* pEV1. n=4. In (**d**), *p value <0.03, **p value = 0.006; two-way ANOVA with Sidak's test of pEV2 vs pEV2/*bti3*. n=5, except N1 Δ*t6p* pEV1, n=6. **c.** Clinical isolates 10765 (*bte1_2*) and 13051 (*bte1_3_4*) and their respective *tssC* deletion strains were grown to log phase and used for SDS-PAGE and anti-Hcp Western blots of whole cell lysates and cell supernatants. **e**. Clinical isolate 13051 (wt and Δ*tssC*) and N1 Δ*t6p* pEV1 derivatives were grown at 37°C in anaerobiosis. Growth curves were determined by $OD_{600}$ measurements every 15 minutes for 24 hours. Competition data are presented as mean ± SD.

**Supplementary Figure 4. Targeting of non-*Bacteroidota* species by *B.fragilis* N1-derived strains and clinical isolates**



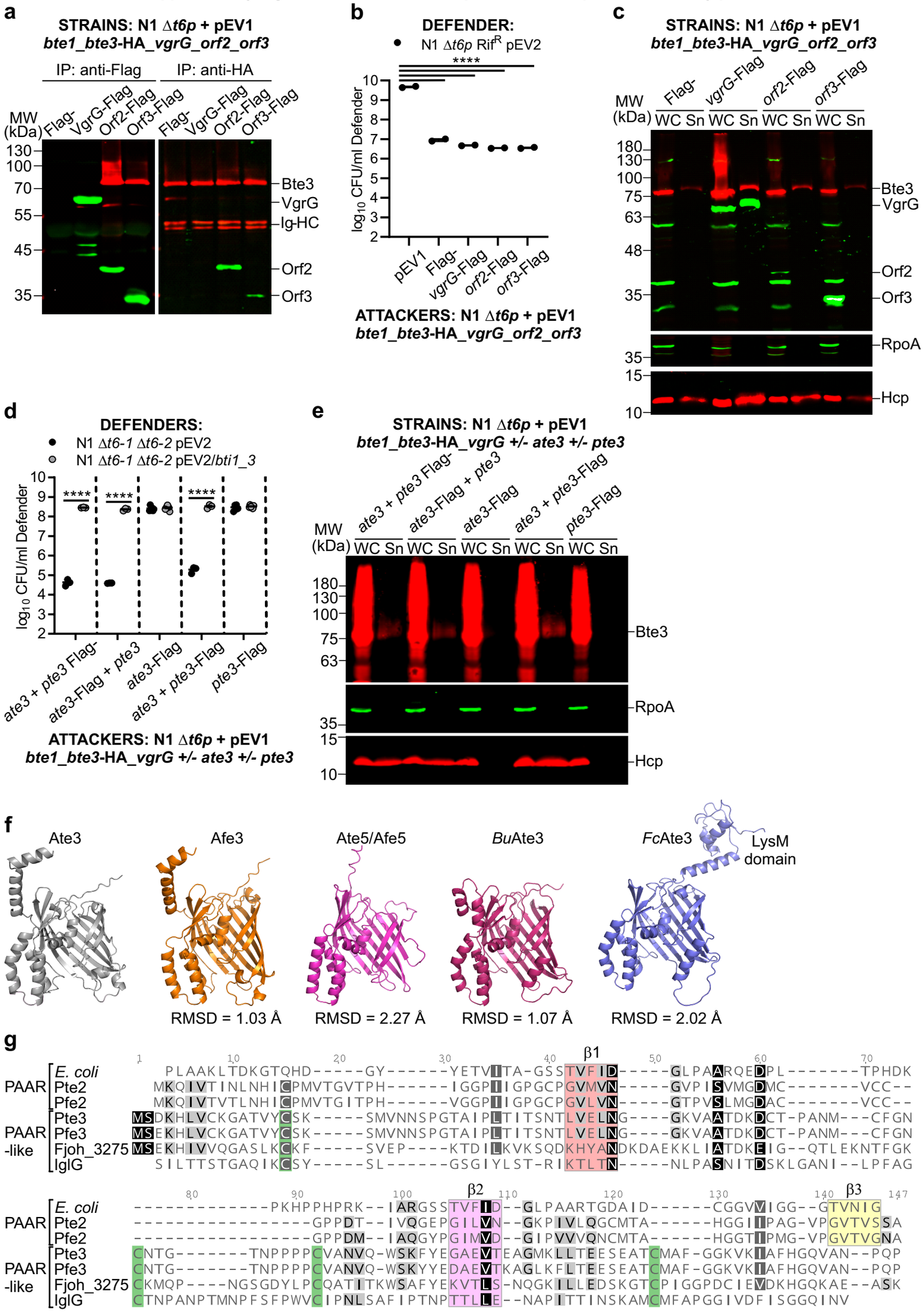
**Supplementary Figure 4. Targeting of non-*Bacteroidota* species by *B.fragilis* N1-derived strains and clinical isolates. Related to Figure 2. a, e.** *In vitro* competitions of clinical isolates (*bte1_2* or *bte1_3_4*) combinations and N1 Δ*t6p* strains against *E. coli* strains and *K. pneumoniae* TOP52 strain (**a**), or Gram-positive species and *C. albicans* strain SN425 (**e**). In (**a**), *p value <0.04, ***p value = 0.0001, ****p value <0.0001; two-way ANOVA with Dunnett's test of t N1 ∆*t6p* pEV1 vs each attacker strain. n=2. In (**e**), *p value <0.0001; two-way ANOVA with Dunnett's test of N1 ∆*t6p* pEV1 vs each attacker strain. n=2. **b.** *Gammaproteobacteria* species and strains were grown to stationary phase and plated in LB media for CFU enumeration (Pre-Anaerobic, filled circles). Cultures were standardized to 1 $OD_{600}$, plated in BHI and incubated for 16-20 hours at 37°C in anaerobiosis. Bacterial growth spots were resuspended and plated in LB for CFU enumeration (Post-anaerobic). *p value <0.0001; two-way ANOVA with Sidak's test of Pre-anaeobic vs Post-anerobic for each strain. n=3. **c.** Bacterial species and strains from (**a**) were grown at 37°C in aerobiosis. Growth curves were determined by $OD_{600}$ measurements every 15 minutes for 24 hours. **d**. Clinical isolate 13051 (*bte1_bte3_bte4*) wt and ∆*tssC* strains were used for *in vitro* competitions in liquid media against *E. coli* S17-1 strain with pEV2 or pEV2 containing *bte3*. p>0.94; two-way ANOVA with Sidak's test of pEV2 vs *bti3* for each attacker strain. All competition data are presented as mean ± SD.

**Supplementary Figure 5. The *btei4* locus encodes an effector immunity pair**



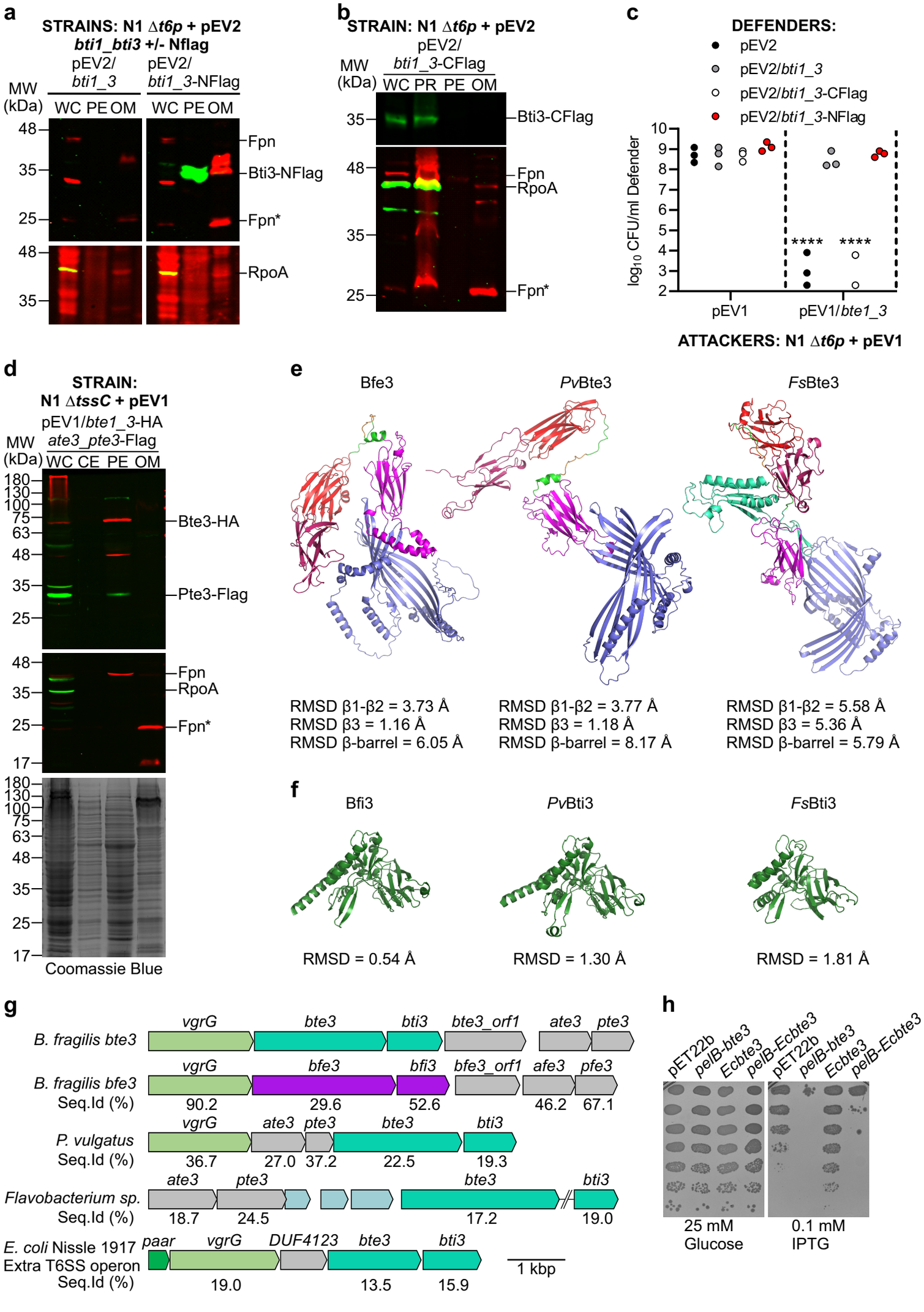
**Supplementary Figure 5. The *btei4* locus encodes an effector immunity pair. Related to Figure 3. a.** *In vitro* competitions of N1 Δ*t6p* with pEV1 containing *bte1_3* or *bte1_3_4* against N1 Δ*t6-1* Δ*t6-2* with pEV2 containing *bti1_3* or *bti1_3_4* immunity gene combinations. ****p value <0.0001; two-way ANOVA with Tukey's test of all vs. all defenders. n=4. **b**. 13051 isolate and Δ*btei4* mutant strain were used for *in vitro* competitions against 13051 wt and Δ*btei4* complemented with pEV2 or pEV2 containing *bti4*. ****p value <0.0001; two-way ANOVA with Dunnett's test of 13051 Δ*btei4* pEV2 vs each defender. n=6. **c**. 13051 isolate and Δ*btei4* mutant strain were used for *in vitro* competitions against *E. coli* S17-1 strain with pEV2 or pEV2 containing *bte3*. ****p value <0.0001; two-way ANOVA with Dunnett's test of pEV2 vs *bti3*. n=4. **d**. Anti-Hcp Western blots of cell supernatants from independent clones of N1 Δ*t6p* strains complemented with pEV1 containing *bte1_2*, *bte1_3* or *bte1_3_4*. All competition data are presented as mean ± SD for n = independent biological replicates.

**Supplementary Figure 6. Bte3 secretion requires effector-specific accessory proteins**



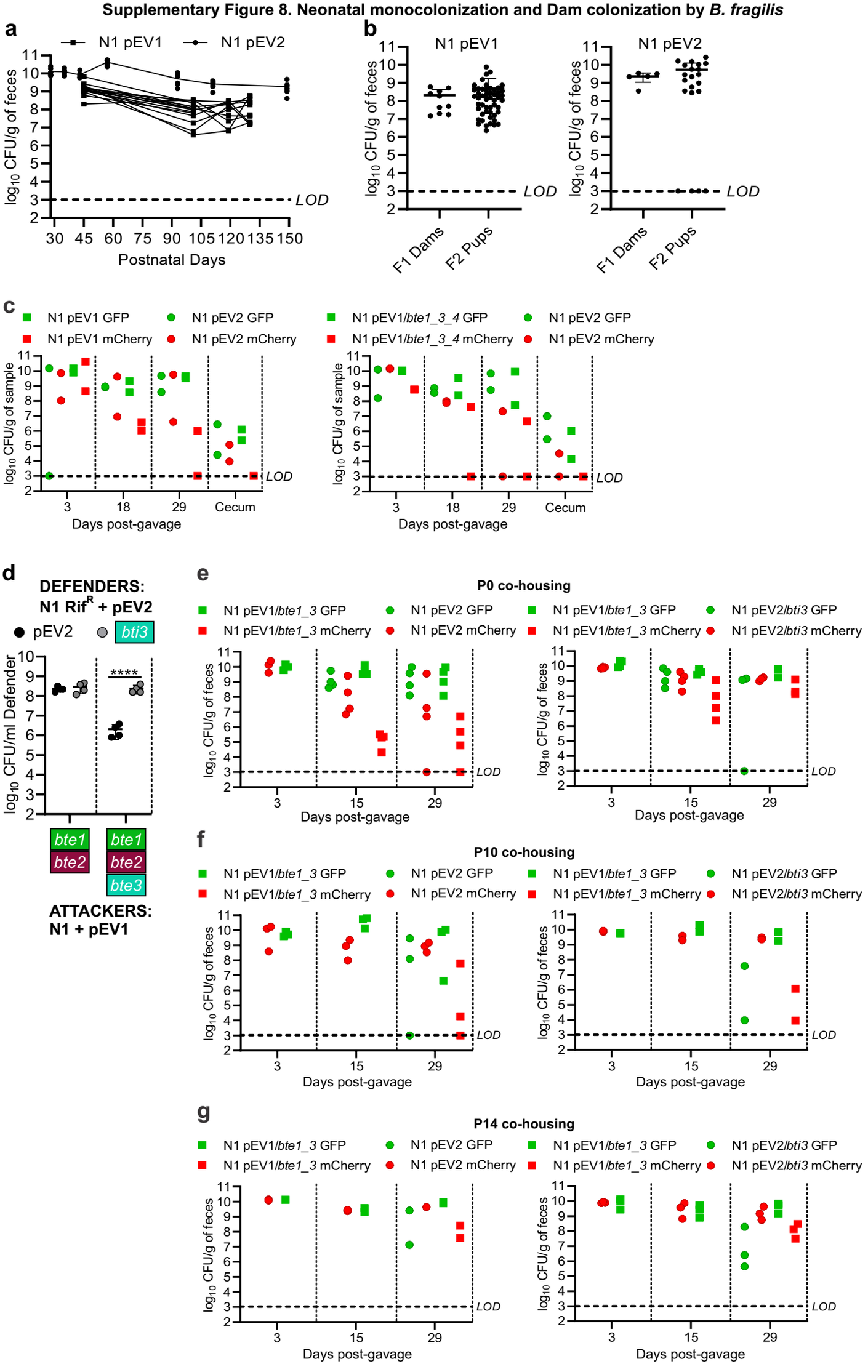
**Supplementary Figure 6. Bte3 secretion requires effector-specific accessory proteins. Related to Figure 3. a-c.** N1 Δ*t6p* was complemented with *bte3*-HA and Flag-tagged accessory genes *vgrG*, *orf2*, or *orf3* and used for (**a**) immunoprecipitations with anti-Flag or anti-HA antibodies and Western blot analysis (Ig-HC: Immunoglobulin heavy chain), (**b**) *in vitro* competitions against the N1 Δ*t6p* pEV2 strain, and (**c**) anti-Flag/HA/Hcp/RpoA Western blots of whole cells and supernatants. In (**b**) ****p value <0.0001; one-way ANOVA with Dunnett's test vs. N1 Δ*t6p* pEV1. n=2. **d**, **e**. N1 Δ*t6p* was complemented with *bte3*-HA and *vgrG* with Flag-tagged *ate3* or *pte3*, in the presence or absence of *pte3* or *ate3*, respectively. Strains were used for (**d**) *in vitro* competitions against N1 Δ*t6p* strain with pEV2 or pEV2/*bti1_3*, and (**e**) anti-Flag/HA/Hcp/RpoA Western blots of whole cells and supernatants. In (**d**) ****p value <0.0001; two-way ANOVA with Dunnett's test of N1 Δ*t6p* pEV2 vs pEV2/*bti1_3*. n=3 except for *ate3-Flag*/*pte3-Flag*, n=6. **f.** Ate3 structural orthologs by Foldseek from *B. fragilis* (Afe3/Ate5/Afe5), other Bacteroidota members, including gut commensal *B. uniformis* (*Bu*Ate3) and free-living *Flavobacterium columnare* (*Fc*Ate3). Root Mean Square Deviation (RMSD) values between Ate3 and structural orthologs were calculated using PyMOL. **g.** PROMALS3D structure-based sequence alignment of PAAR orthologs *Ec*PAAR (PDB ID: 4gIW), Pte2 and Pfe2, and PAAR-like orthologs Bte3, Bfe3, Fjoh_3275 from *F. johnsoniae* and IglG from *F. novicida.* Conserved beta sheets are indicated with boxes: β1 (red), β2 (magenta) and β3 (yellow). Conserved cysteine residues in PAAR-like proteins are highlighted in green. Competition data are mean ± SD for n = independent biological replicates.

**Supplementary Figure 7. Bte3 and Bti3 functional, localization, and structural orthology**



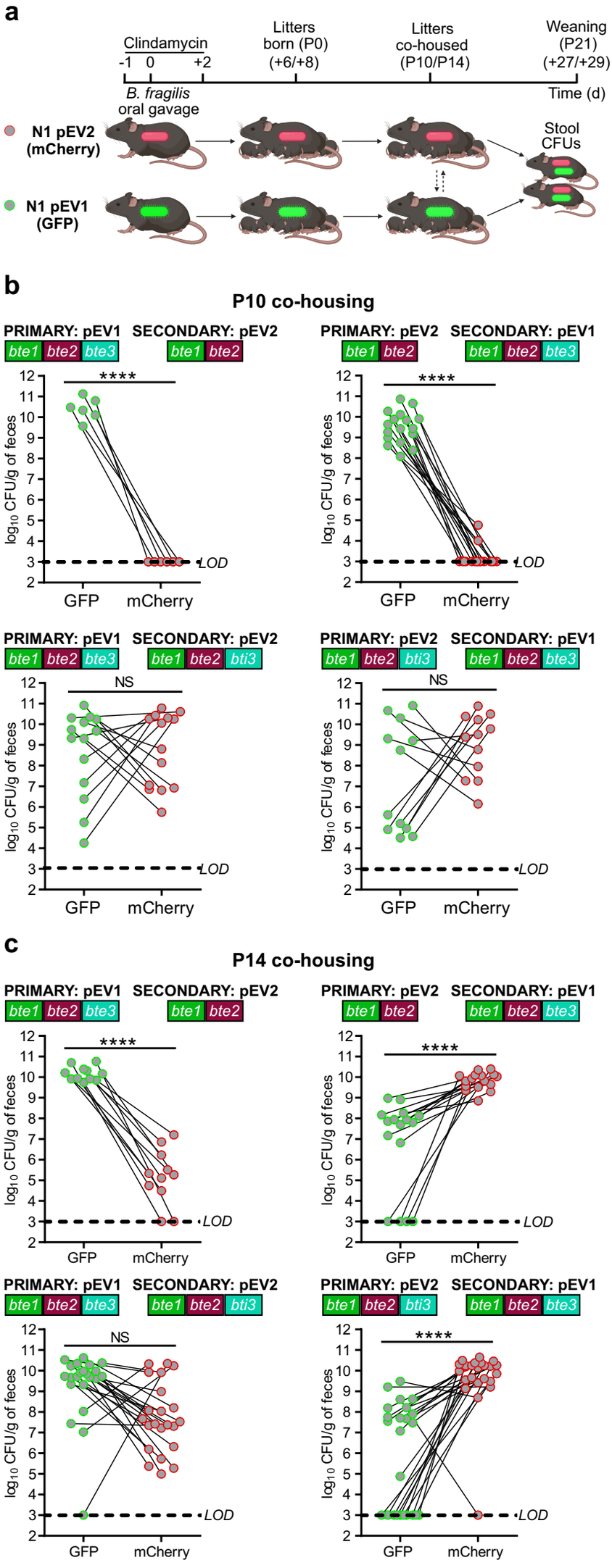
**Supplementary Figure 7. Bte3 and Bti3 functional, localization, and structural orthology. Related to Figure 3. a, b.** Western blots of subcellular fractions from N1 Δ*t6p* complemented with pEV2 containing (**a**) *bti1_3*-NFlag or (**b**) *bti1_3*-CFlag. WC: whole cells; Pe: periplasm; OM: outer membrane; Pr: protoplast. RpoA: RNA Polymerase subunit A. Fpn: Fragipain. **c.** N1 Δ*t6p* was complemented with pEV2 (closed circles) or pEV2 containing *wild type bti3* (gray circles), *bti1_3*-CFlag (open circles) or *bti1_3*-NFlag (red circles). Strains were used for *in vitro* competitions against N1 Δ*t6p* pEV1 or pEV1/*bte1_3*. ****p value <0.0001; two-way ANOVA with Dunnett's test vs. pEV2/*bti1_3*. n=3. **d.** Western blots of subcellular fractions from N1 Δ*tssC* complemented with pEV1 containing *bte3*-HA and Flag-tagged *pte3* genes. CE: Cell envelope, fraction containing periplasm and outer membrane. Bottom panel: Coomassie stained total protein pattern. **e, f.** AlphaFold2 models of Bte3 (**e**) and Bti3 (**f**) orthologs in *B.fragilis* (Bfe3/Bfi3), other gut commensal (*Pv*Bte3/*Pv*Bti3 from *Phocaiecola vulgatus*) and free living (*Fs*Bte3/*Fs*Bti3 from *Flavobacterium* sp.) *Bacteroidota*. Root Mean Square Deviation (RMSD) values were calculated using PyMOL. **g**. Genetic organization of *bte3/bti3* orthologs. Identity percentages are from aminoacidic sequence comparisons between ortholog genes within the *B. fragilis bte3* cluster. *Flavobacterium* sp. operon is missing five hypothetical function *orfs* for figure size constraints (tilted bars). T6SS-2: *E. coli*-specific T6SS architecture 2. *DUF4123*: T6SS chaperone domain in *Pseudomonadota* [60]. **h**. *E. coli* BL21 expressing empty vector (pET22b), cytosolic *Ec*Bte3, or periplasmic *pelB-Ecbte3* were plated under repressing (25 mM Glucose) or inducing (0.1 mM IPTG) conditions. Images are representative of n=3 biological replicates. Competition data are mean ± SD.

**Supplementary Figure 8. Neonatal monocolonization and Dam colonization by *B. fragilis***



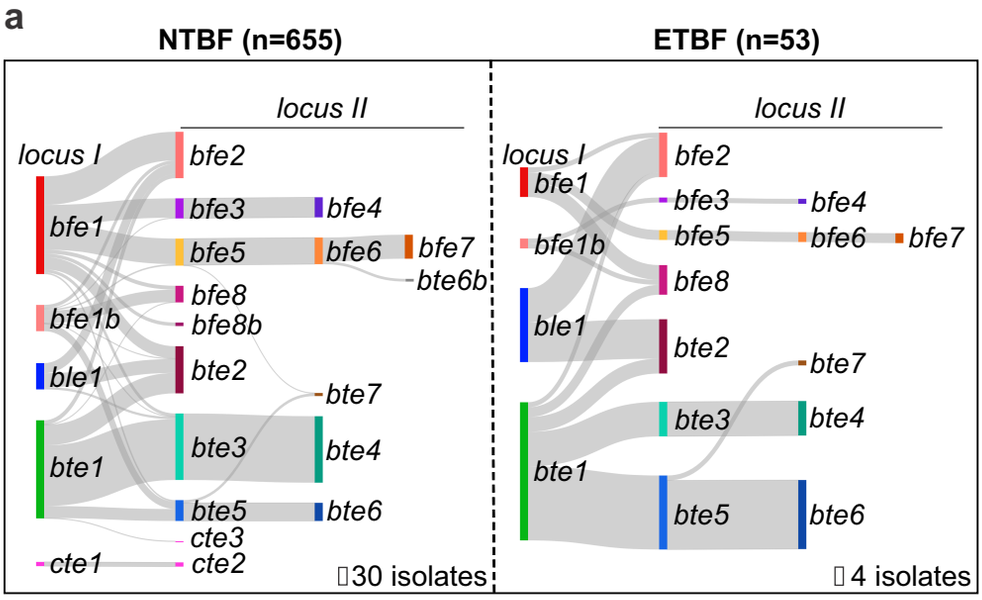
**Supplementary Figure 8. Neonatal monocolonization and Dam colonization by *B. fragilis*. Related to Figure 4. a.** Stool CFU of litters monocolonized by maternal N1 pEV1 or N1 pEV2 after weaning at postnatal day 21. **b**. Stool CFU of N1 pEV1 or N1 pEV2 strains from F1 Dams and F2 neonates at postnatal day 21. **c.** N1 vs N1 strain combinations in dams, represented by stool CFUs at different days post inoculation, together with cecum CFUs when pups were sacrificed for analysis. Related to Figure 4b. **d**. N1 with pEV1 or *bte1_3* were used for *in vitro* competitions against N1 with pEV2 constructs (pEV2 or pEV2/*bti3*). ****p value <0.0001; Sidak's test of N1 pEV2 vs N1 pEV2/*bti3* for each attacker strain. n=4. **e-g**. Dam stool CFU for N1 vs N1 strain combinations at different days post inoculation, related (**e**) Figure 4c, (**f**) Figure 4d, (**g**) Figure 4e. Data represents two or more combined independent experiments. All applicable data are presented as mean ± SD.

**Supplementary Figure 9. Neonatal cross-colonization by *B. fragilis* in delayed co-housing experiments**



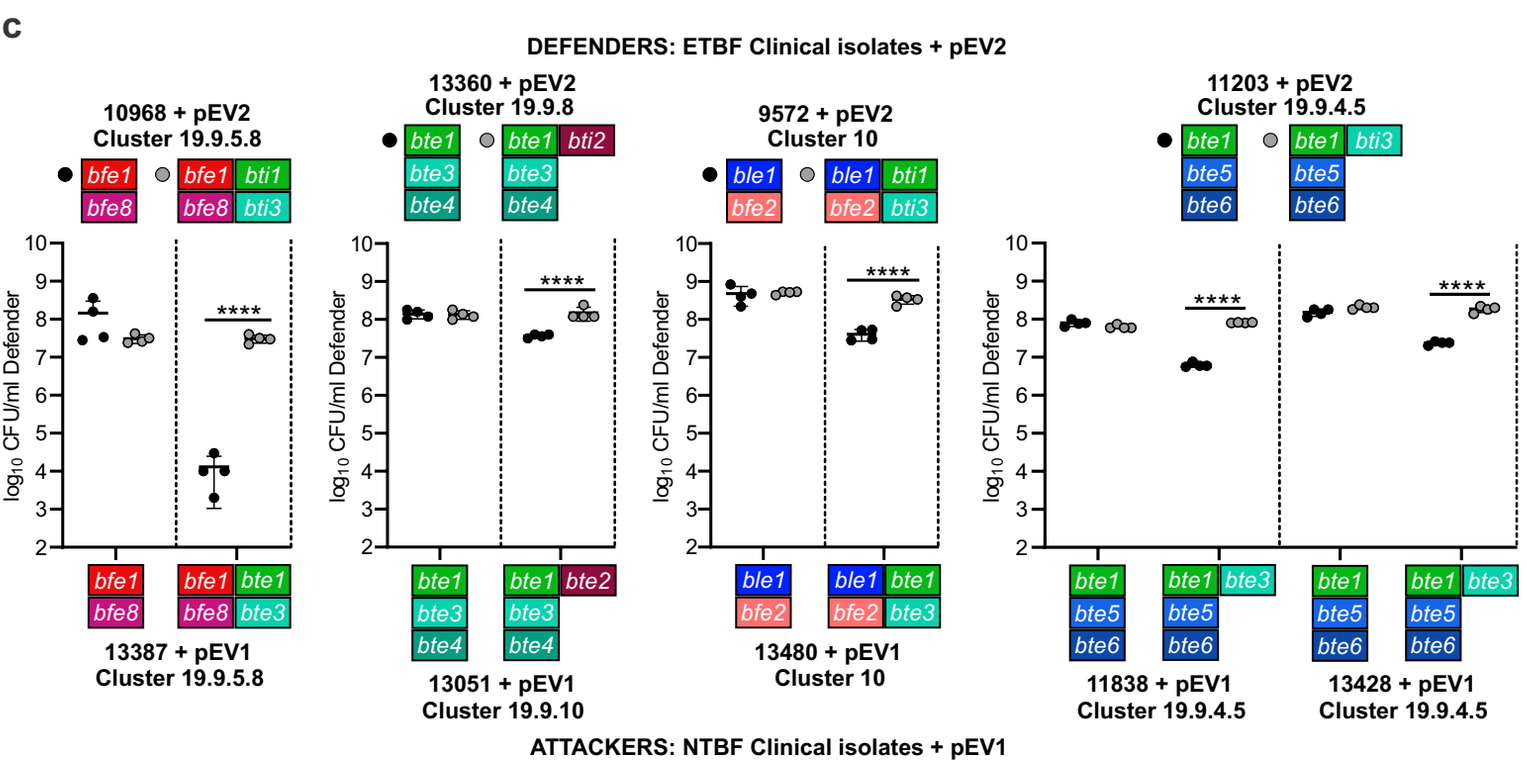
**Supplementary Figure 9. Neonatal cross-colonization by *B. fragilis* in delayed co-housing experiments. Related to Figure 4. a.** Schematic of the neonatal colonization/competition model using N1 complemented with pEV1 + *bte3* or pEV2- (mCherry) constructs. Monocolonized litters with pEV1 or pEV2 strains were co-housed at P10 or P14 until P21. Time (x-axis) = days post-inoculation. P0/P10/P14/P21: postnatal day 0/10/14/21. Graphical elements: Created in BioRender. Valguarnera, E. (2026) https://BioRender.com/87rjycc. **b, c**. Stool CFU determination for stool samples at P21 of neonates co-housed at (**b**) P10 or (**c**) P14. Top panels: co-housing of N1 pEV1 + *bte3* and pEV2. Bottom panels: co-housing of N1 pEV1 + *bte3* and pEV2 + *bti3*. Left panels: primarily exposure to pEV1 + *bte3*. Right panels: primarily exposure to pEV2. ****p value <0.0001; two-tailed paired t test for pEV1 vs pEV2. In (**b**), n=6 (1: pEV1/*bte1_3*; 2: pEV2), n=17 (1: pEV2; 2: pEV1/*bte1_3*), n=14 (1: pEV1/*bte1_3*; 2: pEV2/*bti3*), n=12 (1: pEV2/*bti3*; 2: pEV1/*bte1_3*). In (**c**), n=11 (1: pEV1/*bte1_3*; 2: pEV2), n=16 (1: pEV2; 2: pEV1/*bte1_3*), n=22 (1: pEV1/*bte1_3*; 2: pEV2/*bti3*), n=24 (1: pEV2/*bti3*; 2: pEV1/*bte1_3*). Data represents three combined independent experiments. Data are means ± SD for n = individual neonatal mice.

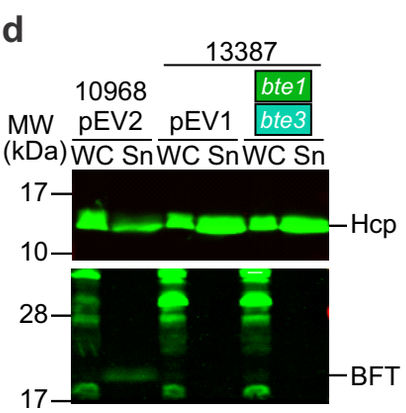
**Supplementary Figure 10. T6SS competition assays between NTBF and ETBF clinical strains**

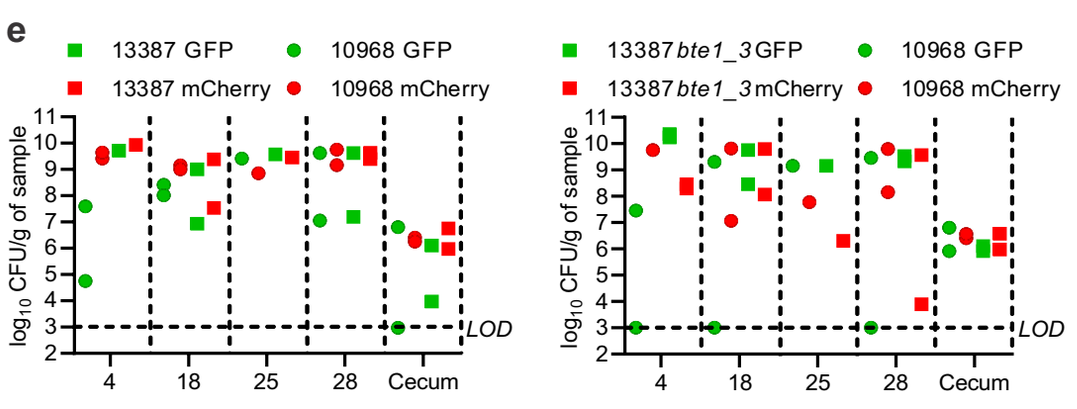



b

| T6SS GA3 | NTBF | ETBF | *bsap* |
|---|---|---|---|
| *bfe1_8* | 2 | 1 | |
| *ble1_bfe2* | 1 | 2 | |
| *ble1_bte2* | 2 | 3 | 1 |
| *bte1_5_6* | 5 | 2 | 1;4 |
| *bte1_5_6* | 3 | 1 | 1 |
| *bte1_2* | 1 | 1 | 1 |
| *bte1_3_4* | 2 | 1 | |
| *bte1_3_4* | 2 | 2 | 1 |
| *bte1_3_4* | 15 | 1 | 1;4 |
| *bte1_3_4* | 15 | 1 | 1;*ubb* |
| Total | 48 | 15 | |



**Supplementary Figure 10. T6SS competition assays between NTBF and ETBF clinical strains. Related to Figure 4.** **a.** Sankey plot of GA3 *ei* combinations separated by *bft* absence (NTBF) or presence (ETBF). Vertical bar height represents absolute counts. Generated using SankeyMATIC (https://sankeymatic.com/). **b.** Table of NTBF/ETBF isolates matched by GA3 *ei* combinations and other known non-T6SS antibacterial toxins (*Bacteroidales* secreted *a*ntimicrobial *p*roteins, *bsap*). **c.** *In vitro* competitions of matched NTBF/ETBF T6SS-isogenic pairs. NTBF strains were complemented with pEV1 or pEV1/*bte1_3* or pEV1/*bte1_2.* ETBF strains were complemented with pEV2 (black circles), pEV2/*bti1_3* or pEV2/*bti1_2* (gray circles). ****p value <0.0001; two-way ANOVA with Sidak's test of ETBF pEV2 vs ETBF pEV2/*bti.* n=4. **d.** Anti-Hcp (top panels) and anti-BFT (lower panels) Western blots of whole cells (WC) and supernatants (Sn) from 10968 and 13387 strains. BFT: C-terminal active secreted BFT. **e**. Dam stool and cecum CFU for 13387 vs 10968 strain combinations at different days post inoculation. Left panels: T6SS-isogenic strain combinations. Right panels: T6SS-anisogenic strain combinations. Shown data corresponds to two combined independent experiments. Data are presented as mean ± SD.

**Supplementary References**


1 Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* **13**, e1005595 (2017). https://doi.org/10.1371/journal.pcbi.1005595

2 Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068-2069 (2014). https://doi.org/10.1093/bioinformatics/btu153

3 Priyam, A. *et al.* Sequenceserver: A Modern Graphical User Interface for Custom BLAST Databases. *Mol Biol Evol* **36**, 2922-2924 (2019). https://doi.org/10.1093/molbev/msz185

4 Leimeister, C.-A. *et al.* Prot-SpaM: fast alignment-free phylogeny reconstruction based on whole-proteome sequences. *GigaScience* **8** (2018). https://doi.org/10.1093/gigascience/giy148

5 Verster, A. J. *et al.* The Landscape of Type VI Secretion across Human Gut Microbiomes Reveals Its Role in Community Composition. *Cell Host Microbe* **22**, 411-419.e414 (2017). https://doi.org/10.1016/j.chom.2017.08.010

6 Wallace, M. J., Jean, S., Wallace, M. A., Burnham, C.-A. D. & Dantas, G. Comparative Genomics of Bacteroides fragilis Group Isolates Reveals Species-Dependent Resistance Mechanisms and Validates Clinical Tools for Resistance Prediction. *mBio* **13**, e03603-03621 (2022). https://doi.org/doi:10.1128/mbio.03603-21

7 Whitaker, W. R., Shepherd, E. S. & Sonnenburg, J. L. Tunable Expression Tools Enable Single-Cell Strain Distinction in the Gut Microbiome. *Cell* **169**, 538-546.e512 (2017). https://doi.org/10.1016/j.cell.2017.03.041

8 Koropatkin, N. M., Martens, E. C., Gordon, J. I. & Smith, T. J. Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. *Structure* **16**, 1105-1115 (2008). https://doi.org/10.1016/j.str.2008.03.017

9 Bencivenga-Barry, N. A., Lim, B., Herrera, C. M., Trent, M. S. & Goodman, A. L. Genetic Manipulation of Wild Human Gut Bacteroides. *J Bacteriol* **202** (2020). https://doi.org/10.1128/jb.00544-19

10 Donnenberg, M. S. & Kaper, J. B. Construction of an eae deletion mutant of enteropathogenic Escherichia coli by using a positive-selection suicide vector. *Infect Immun* **59**, 4310-4317 (1991). https://doi.org/10.1128/iai.59.12.4310-4317.1991

11 Dykxhoorn, D. M., St Pierre, R. & Linn, T. A set of compatible tac promoter expression vectors. *Gene* **177**, 133-136 (1996). https://doi.org/10.1016/0378-1119(96)00289-2

12 Schäfer, A. *et al.* Small mobilizable multi-purpose cloning vectors derived from the Escherichia coli plasmids pK18 and pK19: selection of defined deletions in the chromosome of Corynebacterium glutamicum. *Gene* **145**, 69-73 (1994). https://doi.org/10.1016/0378-1119(94)90324-7

13 Zhou, P., G, C. B., Hu, B. & Wu, C. Development of SacB-based counterselection for efficient allelic exchange in Fusobacterium nucleatum. *Microbiol Spectr* **13**, e0206624 (2025). https://doi.org/10.1128/spectrum.02066-24

14 Hecht, A. L. *et al.* Strain competition restricts colonization of an enteric pathogen and prevents colitis. *EMBO Rep* **17**, 1281-1291 (2016). https://doi.org/10.15252/embr.201642282

15 Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* **5**, 976-989 (1994). https://doi.org/10.1016/1044-0305(94)80016-2