PNAS
www.pnas.org

Supplementary Information for

Bacteriophages evolve enhanced persistence to a mucosal surface

Wai Hoe Chin[1]**, Ciaren Kett[1], Oren Cooper[2], Deike Müseler[3], Yaqi Zhang[3], Rebecca S. Bamert[4], Ruzeen Patwa[1], Laura C. Woods[1], Citsabehsan Devendran[3], Denis Korneev[5], Joe Tiralongo[2], Trevor Lithgow[4], Michael J. McDonald[1], Adrian Neild[3] and Jeremy J. Barr[1]*

Jeremy J. Barr (Primary corresponding author)
Email: jeremy.barr@monash.edu

Wai Hoe Chin (Co-corresponding author)
Email: wai.chin@monash.edu

This PDF file includes:

Methods and materials
Figures S1 to S8
SI References

Other supplementary materials for this manuscript include the following:

Datasets S1 to S6

Methods and materials

Culture protocol for bacteria, phage and tissue culture cell lines. *Escherichia coli* strain B was used for all experiments and was grown in LB medium (10 g Tryptone, 10 g NaCl, 5 g yeast extract in 1 L of sterile dH_2O) at 37°C with agitation. T4 phage, which uses *E. coli* strain B as a replicative host, was used for all experiments except T4 replication-negative *43*$^-$ (DNA polymerase) and *44*$^-$ (polymerase clamp holder subunit) i.e. T4 *am43*$^-$*/44*$^-$ phage, that only uses amber-permissive host *E. coli* SupE to replicate. The cell line used was a human colon-derived tumorigenic goblet cell, HT29-MTX-E12, obtained from the European Collection of Authenticated Cell Cultures and cultured at 37°C with 5% CO_2 in complete media: DMEM with 10% FBS, 1× MEM non-essential amino acids and 1× penicillin-streptomycin antibiotics (ThermoFisher Scientific). Terminal cellular differentiation was induced with 10 µM N-[N-(3,5-Difluorophenacetyl)-L-alanyl]-S-phenylglycine t-butyl ester (DAPT; Sigma-Aldrich) while mucus-secretion was enhanced with 10 nM phorbol 12-myristate 13-acetate (PMA; Sigma-Aldrich).

Fabricating the gut-on-a-chip mould and device. A chip mould with 500 µm wide and 350 µm high channel was designed using SolidWorks® 2017 (Dassault Systèmes). The moulds were then 3D-printed and surface-salinized at Melbourne Centre for Nanofabrication (MCN), Victoria. The chips were manufactured by casting a 10:1 mixture of Sylgard™ PDMS and its curing agent respectively (Dowsil, USA), onto the moulds and were cured at 90°C until completely solidified. The chips were then removed, trimmed and their inlet and outlet ports were punched. Subsequently, the chips were washed in pentane and acetone to remove residual uncured PDMS followed by plasma bonding the chip onto a glass slide to enclose the chip channel. The chip channel was ethanol (80%v/v)-sterilised, UV-sterilised and pre-treated with 1:50 MaxGel™ ECM (Sigma-Aldrich). The channel was then seeded with 10 µl of HT29-MTX-E12 cells at 3.0×10^5 cells. The seeded chip was incubated statically for 16 hours to allow cell attachment. This was followed by perfusing the attached cells with complete media for 24

hours at 40 µl/hr flow rate to establish a confluent cell layer. The cell layer was then perfused with antibiotic-free media supplemented with cell-inducers DAPT and PMA, for another 24 hours at 120 µl/hr to purge residual antibiotic-containing media from the channel environment and to promote terminal cellular differentiation and mucus secretion by the cell layer. Perfusion was mediated by a 10-channel syringe pump (KD Scientific, USA) with the gut-on-a-chip maintained within an incubator at 37°C with 5% CO_2 to ensure optimal conditions for cell culture growth and maintenance in the channel.

Scanning electron microscopy (SEM) of the gut-on-a-chip mucosal epithelium. The HT29-MTX-E12 mucosal epithelium (cultivated as described previously) in the gut-on-a-chip was fixed with fresh methanol Carnoy solution [60% (v/v) absolute methanol, 30% chloroform and 10% glacial acetic acid] at 4°C for 3 hours. The PDMS encasing of the gut-on-a-chip was carefully trimmed and removed using a sharp scalpel to expose the fixed mucosal epithelium. The exposed epithelium was then immersed in 70% (v/v) ethanol for 30 minutes followed by further dehydration with graded ethanol at 85%, 95%, 100% and 100% for 20 minutes each. Residual ethanol was then evaporated with a critical point dryer and the sample was mounted on a standard SEM metal stub, followed by sample coating with a ~10 nm gold layer. The mucosal epithelium was imaged using a FEG-SEM ThermoFisher Elstar G4 (ThermoFisher) operating at an accelerating voltage of 2 kV using secondary electrons at a working distance of 4 mm in immersion mode via the through lens detector (TLD).

High temporal resolution gut-on-a-chip phage-bacteria sampling. An in-house automated dispensing platform was constructed to aid sample collection from the gut-on-a-chip over 24 hours at 30-minute intervals. The platform consisted of conveyer belts connected to 5V motors powered by an Arduino circuit board (Arduino, Italy). Two conveyer belt systems were aligned perpendicular to each other allowing motion along the X-Y plane. A custom-made tube holder was connected to the conveyer belt system that holds the gut-on-a-chip tube over the 96-well

plate to facilitate sample dispensing into wells. Time-steps for dispensing at 30-minute intervals were coded into Arduino in C++ using Arduino Integrated Development Environment (IDE). For a user-friendly interface, the code was translated onto a virtual switch board executable program using LabVIEW v.2020 (National Instruments, USA). The temporal experiment is initiated by perfusing the gut-on-a-chip with 10^4 colony forming units (CFU) of *E. coli* B followed by 10^4 PFU of T4 phages and the device was allowed to run for 24 hours under a 120 µl/hr flow rate whilst connected to the automated dispensing platform to collect egressing fluid samples. In between the gut-on-a-chip and the dispensing platform, the egressing fluid was channelled through an 80°C-heated blank chip to arrest phage and bacterial replication during their egress from the gut-on-a-chip before dispensing. Phages and bacteria from the heat-inactivated samples were quantified using qPCR using SYBR Green I Master with the Lightcycler® 480 (Roche). qPCR primers and cycling protocols for *E. coli* B were as described (1) using 1 µl of template. T4 protocols was adapted from (2) using forward primer: 5'- AGGAGTTATATCAACTGTAA - 3', and reverse primer: 5'- ATCTAGGATTCTGTACTGTT - 3', with the following cycling protocol: initial denaturation at 95°C for 5 minutes; 40 cycles at 95°C for 30 seconds, 56°C for 30 seconds, 72°C for 30 seconds; using 1 µl of template. Standard curves (Dataset S1) for phages were generated via 10-fold dilution in PCR-grade H_2O of a T4 phage stock at concentration 7×10^9 PFU/ml; hence, standardising the curve based on PFU/ml. Standard curve for bacteria was similarly diluted in 10-fold using genomic DNA extracted via GenElute Bacterial Genomic DNA kit following manufacturer's protocol (Sigma-Aldrich), from a known concentration of bacterial culture at 7×10^7 CFU/ml. Each chip replicate generates 48 unique timepoint heat-inactivated samples from sampling at 30-minute intervals over 24 hours. qPCR for each timepoint was carried in three technical replicate (n = 3) lending 144 qPCR sample reactions (48×3) per experimental replicate. There are three experimental replicates in total (N = 3). Due to the small sample volume, 1 µl of sample was directly used without prior DNA extraction to avoid sample loss and inaccuracy.

Phage experimental evolution in gut-on-a-chip. 10^4 PFU of T4 phages were perfused through the gut-on-a-chip followed by 10^4 CFU of *E. coli* B to supply the phages with hosts to replicate within the chip. The co-culture in each chip was maintained under a 120 µl/hr flow rate with antibiotic-free media for 24 hours under 37°C with 5% CO_2 conditions. Subsequently, the mucus and the cell layer were collected via washes with 1 × DPBS and 0.25% Trypsin (ThermoFisher Scientific). The chip sample was centrifuged to obtain the bacterial cell pellet, which was resuspended in 100 µl 1 × DPBS. The supernatant containing the phages was treated with 10% chloroform to obtain a purified phage lysate. Phages and bacteria were enumerated using soft-agar overlay assay and colony spot assay, respectively. For our phage passage protocol, 10^4 phage PFU were taken from the purified phage (supernatant) lysate to inoculate a new gut-on-a-chip with 10^4 ancestral *E. coli* B CFU. We adopted this passage protocol for a total of 5 passages. In our control experimental evolution, a shaking test-tube was used in place of the gut-on-a-chip within the flow set-up. The passage protocol in the control experiment was the same as the passages of the gut-on-a-chip phage experimental evolution.

Phage DNA isolation, purification, sequencing and analyses. To obtain sufficient DNA yield for sequencing, phages from all transfers including the ancestral phage population were amplified to high titres ($\geq 10^9$/ml). The phages were amplified by inoculating 30 µl of phage lysate sample into 3 ml of *E. coli* B bacteria in exponential phase (OD_{600} = 0.3). The inoculum was incubated for a maximum of 4 hours at 37°C with agitation to maximise the probability of all phage genotypes in expanding without interference from host-induced bottlenecks at late stage incubations. This was followed by 10% chloroform treatment to purify the amplified phage lysate. Phages were concentrated and ultrapurified following the phage-on-tap protocol (3). 1 ml of each ultrapurified phage passage lysate was treated with 10 µl Ambion™ DNase I (ThermoFisher Scientific) and 20 µl RNase (Sigma-Aldrich) to eliminate bacterial genome contamination. Subsequently, the lysates underwent phage DNA extraction using Phage DNA

Isolation Kit (Norgen Biotek®, Canada) as per manufacturer protocol with the following modification to maximise DNA yield: 10 µl of 20 mg/ml Proteinase K (Sigma-Aldrich) per 1 ml of amplified phage lysate and incubated at 55°C for 1.5 hours. Phage DNA quality and concentrations were assessed via Nanodrop $A_{260/280}$ (ThermoFisher Scientific) readout and QuBit® Fluorometric Quantification High Sensitivity assay (ThermoFisher Scientific), respectively. Phage DNA samples were sequenced using Illumina HiSeq® 150bp paired-end chemistry (GeneWiz®, Hong Kong) and read alignments to T4 reference genome (4) (NCBI GenBank ID: MT984581.1) were performed via the Breseq (5) Polymorphism Mixed Population pipeline with filter settings turned off to maximise variant calling. *De novo* mutation hits were derived by comparing evolved phage population hits with ancestral background mutations using Breseq's -gdtools SUBTRACT and COMPARE commands. *De novo* mutations were assessed and manually counted to compare *de novo* mutation frequencies between gut-on-a-chip and test-tube experimentally evolved phage populations as described in Dataset S3.

Lytic phage recombination assay. T4 Δ21bp *goF* mutant was isolated from transfer 4 chip-evolved replicate 1 population by isolating phage plaques from soft-agar overlay. The phage isolates were PCR-screened and Sanger-sequenced with the flanking *goF* primers i.e. forward: 5' – GCATTAATCAGCATCAGTAC -3' and reverse: 5' – AAGACGGCACAACTTACTGG – 3', with the following PCR protocol: initial denaturation at 95°C for 10 minutes; 34 cycles at 95°C for 10 seconds, 57°C for 15 seconds, 72°C for 60 seconds; and final elongation at 72°C for 5 minutes. T4 *hoc* knockout (Δ*hoc*) phage was also PCR-amplified and sequence-confirmed using the flanking *hoc* primers i.e. forward: 5' – GCTGAAACTCCTGATTGGAAATCTCACCC – 3' and reverse: 5' – GCCCATAATACAGCCACTTCTTTTGCC – 3', with the following PCR protocol: initial denaturation at 95°C for 10 minutes; 34 cycles at 95°C for 30 seconds, 60°C for 60 seconds, 72°C for 90 seconds; and final elongation at 72°C for 10 minutes. The verified phages were amplified and chloroform-purified to high titre ($\geq 10^9$ PFU/ml), respectively. The

phages were diluted in SM buffer (5.8 g NaCl, 2.0 g $MgSO_4.7H_2O$, 50 ml 1 M Tris-HCl pH 7.4 in 1 L ddH_2O) to obtain a 1:1 phage mix containing ∆21bp *goF* and ∆*hoc* at 1×10^9 PFU/ml. 1 ml of the mixture was reserved as an initial condition control to test for 1:1 mix accuracy. The remaining mixture was used to prepare four experimental set-ups: two replicates of MOI = 10 and two replicates at MOI = 0.1. In MOI 10, 1 ml of the 1×10^9 PFU/ml mixture was added to 1 ml of 1×10^8 CFU/ml *E. coli* B; while in MOI 0.1, the phage mixture was diluted to 1×10^7 PFU/ml before adding to 1×10^8 CFU/ml *E. coli* B. The co-cultures were then incubated at 37°C with 150 rpm agitation for 30 minutes to allow a one-step T4 phage growth curve. The co-cultures were subsequently quenched with 10% chloroform. The phages in co-culture and the reserved initial condition phage mix were plated via soft-agar overlay. Single plaque cores were obtained from well-separated plaques, resuspended in 100 µl SM buffer, and PCR screened for recombinants (double mutant: ∆21bp *goF* + ∆*hoc* or WT recombinant T4 genotypes) using flanking *goF* primers and internal *hoc* primers. Internal *hoc* PCR primers were, forward: 5' - ACATTATCTACGCTCCAAGC – 3' and reverse: 5' - ATCTAGGATTCTGTACTGTT - 3', with the following protocol: 95°C for 10 minutes; 34 cycles at 95°C for 10 seconds, 56°C for 15 seconds, 72°C for 60 seconds; and final elongation at 72°C for 5 minutes. All PCR products were loaded on 2% agarose gel, stained with SYBR™ Gold Nucleic Acid Gel Stain (ThermoFisher Scientific), for 30 minutes at 60V and subsequently, 30 minutes at 50V to allow better separation between the WT and ∆21bp *goF* product. Both *goF* and *hoc* PCR products were matched to their sample of origin in the agarose gel run. The frequency of recombinants was quantified based on the *goF* PCR product size and the presence and absence of *hoc* PCR product.

Sequencing-based phage competition assay. Wildtype T4 phage and experimentally evolved D246N T4 mutant phage were isolated via plaque coring as previously described. The cores were resuspended in 100 µl of SM buffer and samples were PCR-amplified with flanking

hoc primers i.e. forward: 5’ – GCCCATAATACAGCCACTTCTTTTGCC – 3’ and reverse: 5’ – GCTGAAACTCCTGATTGGAAATCTCACCC – 3’, with the following protocol: initial denaturation at 95°C for 10 minutes; 30 cycles at 95°C for 30 seconds, 60°C for 60 seconds, 72°C for 90 seconds; and final elongation at 72°C for 10 minutes. The verified phages were amplified and chloroform-purified to high titre ($\geq 10^9$ PFU/ml), respectively. The amplified phages were diluted in antibiotic-free tissue culture media to obtain a 1:1 phage mix containing WT and D246N phages at 2×10^6 PFU (1×10^6 PFU each). 1×10^6 PFU of the phage mix was reserved as an initial condition i.e. T = 0 control. Three gut-on-a-chip replicates were each infused with 10^6 CFU *E. coli* B bacteria followed by 1×10^6 PFU phage mix at 120 µl/hr flow rate. The inoculated devices were maintained at 120 µl/hr for 24 hours and egressing fluid samples were collected for 1 hour at the 24-hour timepoint. Fluid samples were collected in 1 ml SM buffer to rapidly dilute the collected phages and bacteria to limit further phage adsorption during sample collection. Collected samples were then amplified, DNA-extracted, sequenced and analysed as previously outlined to track the frequency of D246N mutant phage as it competes with WT T4 phage over 24 hours. Selection coefficients were calculated as described in Dataset S4 based on absolute reads, obtained by multiplying read depth and coverage, of the mutation.

Molecular cloning of recombinant Hoc protein expression strains. Wildtype Hoc T4 phage and D246N Hoc T4 phage genomic DNA were extracted as described above. The respective *hoc* genes were PCR-amplified using primers designed with NcoI/SpeI restriction sites i.e. forward: 5’ – CCTCCATGGCGATGACTTTTACAGTTGATATAAC – 3’ and reverse: 5’ – TTGACTAGTTATGGATAGGTATAGATGATAC – 3’, with the following protocol: initial denaturation at 98°C for 5 minutes; 36 cycles at 98°C for 30 seconds, 58°C for 30 seconds, 72°C for 120 seconds; and final elongation at 72°C for 5 minutes. The amplified *hoc* products were gel-extracted following manufacturer’s protocol (GenElute™ Gel Extraction Kit, Sigma

Aldrich). Wildtype and D246N *hoc* genes were individually cloned in-frame to expression vector pPROEX-HTb, containing an N-terminal hexa-His sequence. Briefly, the amplified *hoc* product and pPROEX-HTb were digested with NcoI and SpeI (New England Biolabs) at 37°C overnight, followed by ligation at room temperature for 2 hours. The ligated expression vector was transformed into NEB 5α Competent *E. coli* as per manufacturer's protocol (New England Biolabs) and plated on LB medium supplemented with 100 µg/ml ampicillin, where colonies were PCR-screened as above mentioned. PCR-positive colonies were grown and the vector was extracted using GenElute Plasmid Miniprep Kit following manufacturer's protocol (Sigma-Aldrich). The vector was then transformed into expression strain *E. coli* BL21(DE) Star as follows. *E. coli* BL21(DE) Star was grown in LB medium to OD_{600} 0.4 at 37°C. 5 ml of culture was centrifuged at 4°C and the pellet was washed thrice with 1 ml ice-cold 10% glycerol between centrifugations. The pellet was resuspended in 50 µl of ice-cold 10% glycerol and added with 3 µl of the expression vector. The mixture was transferred into a 0.1 cm electroporation cuvette (BioRad) and pulsed at 1.8 kV. Electroporated cells were recovered in 1 ml pre-warmed LB medium for 1 hour at 37°C and subsequently plated on LB medium supplemented with 100 µg/ml ampicillin to recover Hoc expression strains.

Recombinant Hoc protein expression, purification and modelling. Hoc expression strains were grown in Terrific Broth (with shaking) to OD_{600} 0.8 at 37°C. Expression was induced with 0.2 mM IPTG, incubation temperature dropped to 18°C and cells collected by centrifugation the following morning. Cell pellets were resuspended in 20 mM Tris pH8, 300 mM NaCl, 20 mM imidazole, 0.5 mM $MgCl_2$, 1× complete EDTA-free protease inhibitor (Roche) and lysed through an Avestin Emulsiflex C3 cell press. Following centrifugation at 18000 ×g the soluble fraction was applied to a 5 ml HisTrap HP column (GE Healthcare). The column was washed and protein eluted along a gradient using 20 mM Tris pH8, 400 mM NaCl, 1 M imidazole. The peak fraction (eluting at ~150 mM imidazole) was pooled and further purified over size

exclusion chromatography on a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) equilibrated in SEC buffer (20 mM Tris pH 8.0 and 150 mM NaCl). The Hoc proteins each eluted as a single monomeric peak and were run on reducing SDS-PAGE and verified by anti-His western (R&D Systems) (Fig.S6). Proteins were concentrated to 1 mg/ml, EDTA added to 0.5 mM final and aliquots snap-frozen in liquid nitrogen. The structural model of the T4 Hoc protein was generated using Phyre2 server (6) and modelled upon the crystal structure of the three N-terminal IgG domains of phage RB49 Hoc protein (PDB ID: 3SHS). The capsid binding domain could not be accurately modelled due to a lack of solved structural homologues.

Glycan array printing. Glycan arrays consisting of 150 diverse glycans (DextraLabs) in the absence of spacers were taken from existing glycan libraries (7–9). Glycans were amine functionalized as previously described (10) and subsequently printed as described (11). Briefly, glycosylamines were suspended in 1:1 dimethylformamide (DMF) : dimethyl sulfoxide (DMSO) at a concentration of 500 µM and printed onto SuperEpoxy3 glass slides (ArrayIt) using a SpotBot Extreme array spotter (ArrayIT) in a six-pin subarray print per glass slide format. All glycans were printed in replicates of four, including four AlexaFlour 555/647 and FITC control spots, per subarray using 946MP4 pins and a contact time of 1 second at 50% relative humidity, with pins being reloaded after every 8 spots. DMF : DMSO was also printed as blanks controls. The printed arrays were subsequently acetylated in 25% (v/v) acetic anhydride in methanol at 4˚C for 15 min, and then neutralized in 1:1 ethanolamine : DMF. Finally, glycan arrays were washed with 100% ethanol and dried in an empty 50 mL tube by centrifugation for 5 min at 200 ×g. Glycan arrays were vacuum sealed and stored at 4°C.

T4 phage labelling and glycan array hybridization. To label T4 phages (wildtype [WT], D246N Hoc or Δ*hoc),* stocks were diluted to 10^8 phages/mL in SM buffer and allowed to incubate with SYBR green dye (1:1000) (Molecular Probes) in the dark at 4°C for 1 hour. Excess dye was removed by three consecutive washes with 1 mL of SM buffer using an Amicon ultrafiltration tube (100 kDa). A buffer-exchange through three consecutive washes with 1 mL of array phosphate buffered saline (PBS) (50 mM PBS, 1.8 mM $MgCl_2$ and 1.8 mM $CaCl_2$, pH 7.4) was similarly performed using Amicon ultrafiltration tubes (100 kDa) (Merck). SYBR-labelled phages were prepared fresh daily and immediately applied to glycan arrays after buffer-exchange. Before hybridizations, glycan array slides were blocked in 0.5% BSA in array PBS for 5 min at room-temperature (RT). After washing with array PBS, slides were dried through centrifugation and a Gene Frame (1.7 × 2.8 cm, 125 µL, Abgene) was used to isolate the arrays prior to the addition of the labelled phage. 10^8 of either SYBR labelled WT, D246N Hoc or Δ*hoc* T4 phages were applied to individual glycan arrays as a 1 mL bubble and allowed to hybridize at RT for 1 hour in the dark. In the final 5 minutes of incubation, phages were fixed through the addition of formaldehyde into the same bubble (final concentration 4%). Following hybridization, glycans arrays were gently washed three times for 5 min in array PBS and finally dried through centrifugation.

WT and D246N Hoc protein labelling and glycan array hybridization. Labelling of recombinant WT and D246N Hoc proteins was performed using their respective hexa-His-tags. Here, 1 µg of each protein was incubated at a molar ratio of 1:2:4 with anti-His-tag mouse monoclonal antibody (Cell Signalling Technology), anti-mouse-IgG-Alexa647 conjugated rabbit polyclonal antibody (Life Technologies) and goat conjugated anti-rabbit-IgG-Alexa647 polyclonal antibody (Life Technologies) in 1 mL Array PBS. This complex was allowed to hybridize in the dark at 4˚C for 15 min. As described previously, gene frames were used to isolate glycan arrays, and Alexa647 labelled recombinant Hoc proteins were applied as a

bubble for 1 hour at RT and allowed to hybridize. Glycan arrays were subsequently washed for three times for 5 min in array PBS, and dried through centrifugation.

Fluorescent image acquisition and data processing. Fluorescence intensities of the array spots were measured with the Innoscan 1100AL (Innopsys) scanner using either the 488 nm (SYBR) or 635 nm (A647) laser excitation wavelength depending on the sample. The Image analysis was carried out using the inbuilt imaging software, MAPIX (Innopsys). Raw glycan signals were exported into Microsoft Excel 2016. The mean background was calculated from the average of DMF/DMSO blanks on the array plus three standard deviations. This was subtracted from each glycan to generate an adjusted signal. A one tailed t-test was performed with significance set at $p = 0.05$. Binding events confirmed across 3 arrays were compiled as heatmaps representing t-test and fold increases above background.

Surface plasmon resonance detection. Surface plasmon resonance (SPR) experiments to confirm glycan hits and elucidate differences in binding affinity between the WT and D246N Hoc proteins were performed using a Pioneer FE SPR system (Pioneer). WT and D246N Hoc proteins were loaded onto channels 1 and 2 of a HisCap biosensor (Satorious) and channel 3 was blank immobilized to enable reference subtraction in PBS. A minimum of 5000 relative units (RU) of either Hoc protein was immobilized using the nitrilotriacetic acid (NTA)-Nickel capture system modified from reference (12). Here, the hexa-His-tag allows capture of the Hoc proteins in the correct orientation and subsequent covalent crosslinking prevents protein from dissociating over the course of the SPR run. In brief, nickel was bound to the HisCap biosensor using $NiSO_4$ in running buffer. The carboxymethylated dextran (CMD) surface was then activated using N-hydroxysuccinimide (NHS)/1-ethyl-3(3-dimethylaminopropyl)-carbodiimide hydrochloride (EDC). Each protein was subsequently immobilized at flow rate of 10 µL/min for 10 min. Uncoupled amine reactive sites of the CMD were blocked through an injection of

ethanolamine and finally 0.35 M EDTA was injected to remove any poorly associated protein. A maximum concentration of 100 µM of selected glycans was tested using a OneStep analysis programmed using the Pioneer instrument. OneStep was performed with 75% loop volume and a 3% sucrose control. Glycans were flowed at 40 µL/min with a dissociation time of 180s (Fig.S7). Subsequent regeneration of the surface was performed with TE buffer for 60s at 50 µL/min and 60s dissociation. Blanks were run periodically every 2 cycles. Analysis of SPR sensorgrams to determine glycan dissociation constants (K_D) was performed separately with the Qdat analysis software package (Biologic Software, Campbell, Australia). All analyses were performed on two independently prepared HisCap chips with each protein loaded twice and glycans tested in duplicate per run. SPR responses less than 5 RU were deemed insignificant and attributed to non-specific interaction of the glycan with the positively charged HisCap chip surface.

Phage retention and washout assay. A 1:1 phage mix consisting of D246N Hoc T4 phage and a WT T4 Hoc phage (am*43*$^-$*/44*$^-$) was prepared in antibiotic-free tissue culture media at 1 mM final glycan concentration of glycans α-1,6-mannobiose (DextraLabs) or Lacto-N-fucopentaose I (DextraLabs). The 1:1 phage ratio was verified by plating on *E. coli* SupE and *E. coli* B lawns in triplicates where, the amber phage only plaques on an amber mutant permissive host, *E. coli* SupE while D246N::Δ*goF* phage plaques on both *E. coli* SupE and *E. coli* B. Hence, we were able to quantify the D246N::Δ*goF* phage (on *E. coli* B) and the amber mutant phage via subtraction (total plaques from *E. coli* SupE – total plaques from *E. coli* B). Three replicate gut-on-a-chips were infused with 1×10^7 PFU/ml of 1:1 phage-glycan mix for 1 hour at 120 µl/hr (Figs.4C and S8A). After which, the devices were perfused with sterile antibiotic-free tissue culture media supplemented with 1 mM of the appropriate glycan for 4 hours. Device effluents were collected in equal volumes of SM buffer every 15 minutes for the first hour and every 30 minutes for the subsequent hours. The phage timepoints were

quantified by spot-plating the device effluents on both *E. coli* B and *E. coli* SupE lawns to assess for phage washout.

Statistical analysis

All analyses for statistical significance were performed with Prism software (GraphPad Prism 9). Unless stated otherwise, all experiments were performed with three experimental replicates (N = 3) with three repeated measures i.e. technical replicates (n = 3). Data are plotted as mean ± SEM of experimental replicates (N) or technical replicates (n) as defined within the figure legends. Statistical details for t-tests were reported in text with the appropriate figure; this includes t values, degrees of freedom (d.f.) and P values. Significant values are designated as $P < 0.05$ and non-significant values as $P \geq 0.05$, unless stated otherwise.

Supplementary information figures

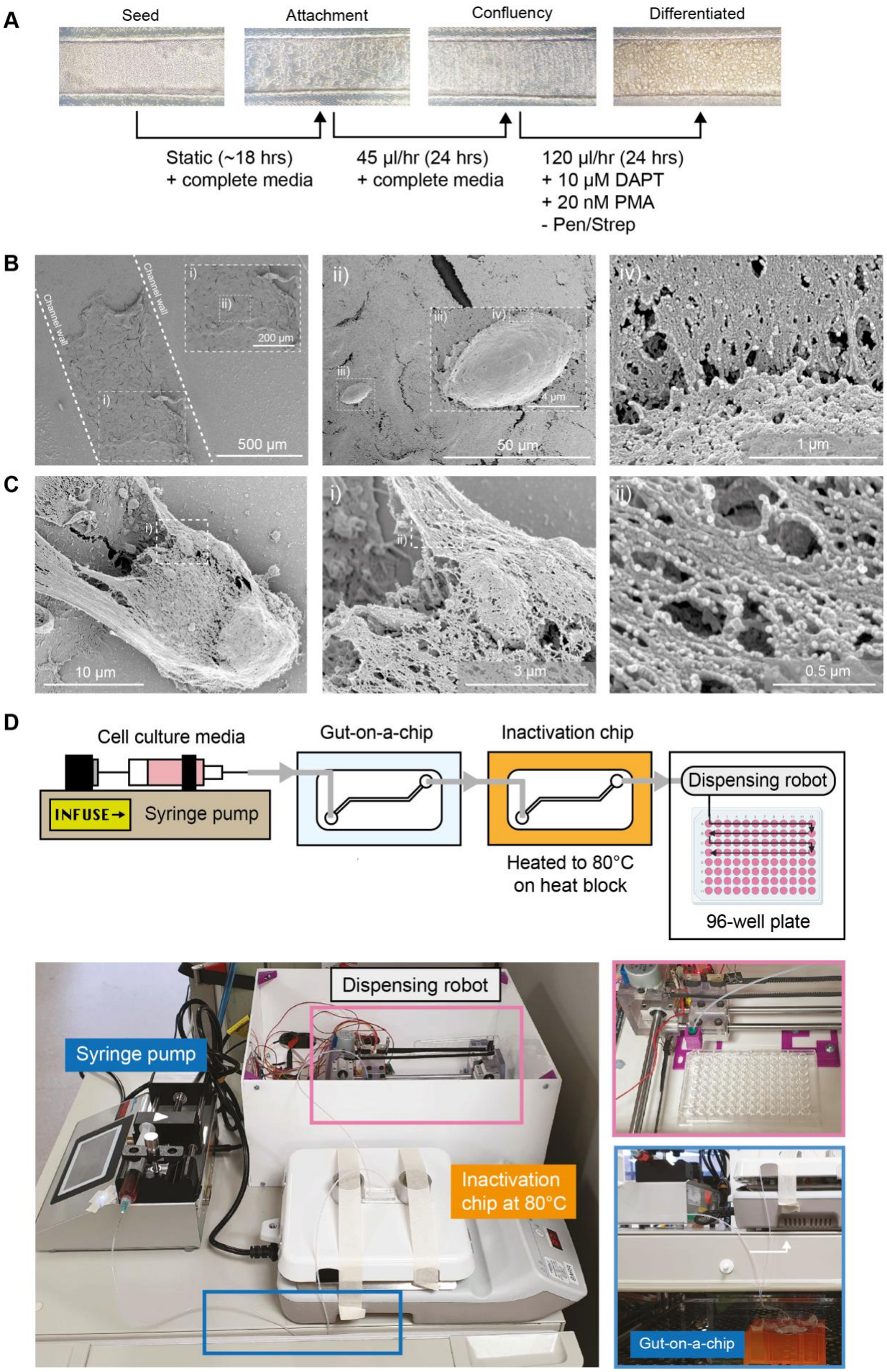

Fig.S1: A) HT29-MTX-E12 cell culture in the gut-on-a-chip. Cells were seeded to a density of 3.0×10^7 cells/ml in the device channel and allowed to incubate under static conditions for ~18 hours for attachment. Attached cells were subsequently fed with tissue culture media at low perfusion rate of 45 µl/hr until the cell layer has established confluency (~24 hours). Following that, antibiotic-free (-Pen/Strep) media, supplemented with inducers – 10 µM N-[N-(3,5-Difluorophenacetyl)-L-alanyl]-S-phenylglycine t-butyl ester (DAPT) and 20 nM 12-myristate 13-acetate (PMA) – to promote terminal cellular differentiation and mucus production, respectively were perfused at a 120 µl/hr to mimic physiological shear stresses of the *in vivo* gut. The resulting acclimatised cell layer possessed extensive goblet cell morphologies and exhibited mucus production visible from alterations in cell layer shade and opacity under light microscopy. B) Scanning electron microscopy (SEM) of the HT29-MTX-E12 mucosal epithelium in a gut-on-a-chip from the anterior perspective i.e. top-view. The parallel lines indicate the position of the channel walls before removal of the gut-on-a-chip PDMS encasing for ethanol graded dehydration and imaging. Numbered dotted boxes represents the focused and magnified fields of vision of the mucosal blanket of the epithelium (i and ii) leading up to a high magnification image (iii and iv) of the net-like structure of the mucus glycoprotein layer. C) Up-close SEM of a single HT29-MTX-E12 cell within the gut-on-a-chip. Numbered dotted boxes represents the focused and magnified fields of vision on the mucosal network enveloping a single cell (i) and the net-like structure of the mucosal network (ii). D) Automated sampling platform for high temporal resolution sampling from the gut-on-a-chip. The gut-on-a-chip was perfused by a syringe pump where egressing fluid from the device was channelled to an “inactivation chip” – an empty device placed on a heat-block at 80°C – to inactivate phages and bacteria within the egressing fluid sample. The inactivated sample was subsequently channelled to a dispensing robot which dispenses the inactivated sample across a sterile 96-well plate over 24 hours, which were then quantified for phage and bacteria abundances via qPCR. The robot was manufactured in-house, consisting of Arduino-driven stepper motors and conveyer belts that moves along the horizontal X-Y plane following written code with relevant time-steps.

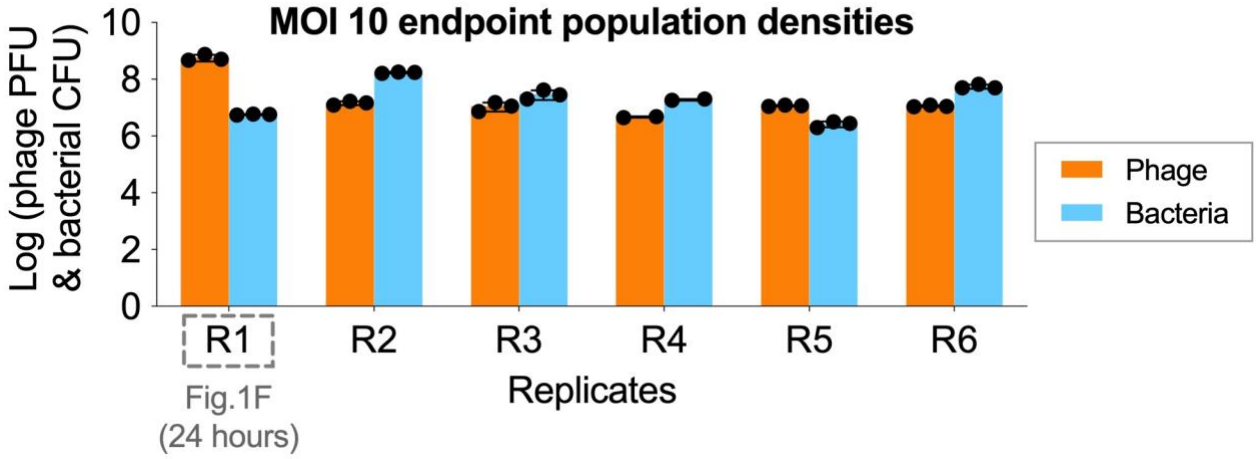

Fig.S2: Phage-bacteria densities recovered at 24-hour endpoint from six replicate devices (N = 6) initiated at MOI 10. Data points were plotted as a bar chart with error bars representing mean ± SEM of three technical replicates (n = 3; except replicate 4 i.e. R4, where n = 2). Grey dotted box on R1 corresponds to the endpoint population density at 24 hours within the MOI 10 device in Fig.1F.

A. *goF* nucleotide sequence:

```
Mutant   1   ATGGCTATTAAATTTGAAGTTAATAAATGGTATCAATTTAAAAATAAACAAGCTCAAGAA  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Ancest   1   ATGGCTATTAAATTTGAAGTTAATAAATGGTATCAATTTAAAAATAAACAAGCTCAAGAA  60

                   1)Sub:T
Mutant  61   AATTTTATTAAAGACCATACTGATAACGGAATCTATGCACGACGTTTAGGTATGGAGCCT  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Ancest  61   AATTTTATTAAAGACCATACTGATAACGGAATCTATGCACGACGTTTAGGTATGGAGCCT  120

                         2)5aa-del
Mutant 121   TTTAAAATTTTAGATGCTGATTATCTTGGGCGTCCTACTAAAATTATGACATCCATAGGT  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Ancest 121   TTTAAAATTTTAGATGCTGATTATCTTGGGCGTCCTACTAAAATTATGACATCCATAGGT  180

Mutant 181   GTACTCAAACGTTGTGCCGGCGGTGATATCCTTGACGAAAACTTTATCTGGCTCTCTACT  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Ancest 181   GTACTCAAACGTTGTGCCGGCGGTGATATCCTTGACGAAAACTTTATCTGGCTCTCTACT  240


                3)Sub:A                                              7aa-del
Mutant 241   AACGAAGCTGGGTTCTTTGATGAAGTGGAAAATCCATATCAGGCGGTTG----------  289
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Ancest 241   AACGAAGCTGGGTTCTTTGATGAAGTGGAAAATCCATATCAGGCGGTTGAAGAGCAAGAG  300

Mutant 290   ----------AAGAGCAAGAACAAATAGAAGATTTCACAGAATTCCCAGTCATGAAAGTT  339
                       ||||||||||||||||||||||||||||||||||||||||||||||||||||
Ancest 301   CAGGAAGAGAAAGAGCAAGAACAAATAGAAGATTTCACAGAATTCCCAGTCATGAAAGTT  360
      4)Sub:T (early STOP codon: UAG)

Mutant 340   ACTATTGAAAATAATGATCAGGCGTGGTCTTTATATCAGATGTTGAAAGCTTACTTTAAG  399
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Ancest 361   ACTATTGAAAATAATGATCAGGCGTGGTCTTTATATCAGATGTTGAAAGCTTACTTTAAG  420

Mutant 400   GAATAA  405
             ||||||
Ancest 421   GAATAA  426
```

B. GoF protein sequence:

```
Mutant   1   MAIKFEVNKWYQFKNKQAQENFIKDHTDNGIYARRLGMEPFKILDADYLGRPTKIMTSIG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Ancest   1   MAIKFEVNKWYQFKNKQAQENFIKDHTDNGIYARRLGMEPFKILDADYLGRPTKIMTSIG  60

Mutant  61   VLKRCAGGDILDENFIWLSTNEAGFFDEVENPYQAVE-------EQEQIEDFTEFPVMKV  113
             ||||||||||||||||||||||||||||||||||||        ||||||||||||||||
Ancest  61   VLKRCAGGDILDENFIWLSTNEAGFFDEVENPYQAVEEQEQEEEKEQEQIEDFTEFPVMKV  120

Mutant 114   TIENNDQAWSLYQMLKAYFKE  134
             ||||||||||||||||||||||
Ancest 121   TIENNDQAWSLYQMLKAYFKE  141
```

Fig.S3: A) Nucleotide sequence of complete (ancestral; Ancest) *goF* gene from the National Centre for Biotechnology Information (NCBI) aligned with Sanger-sequenced *goF* mutant (Mutant). The alignment shows the position of the in-frame 21bp-deletion (Δ21bp) in yellow. Other previously characterised mutations are shown in blue, numbered and labelled in bold (13). B) The corresponding translated protein sequence of the ancestral *goF* gene aligned with the mutant *goF* protein sequence,

demonstrating the seven amino acids eliminated by the in-frame Δ21bp in yellow. The green zones represent the acidic region with predicted homology to RNA-binding proteins and RNA helicases (14).

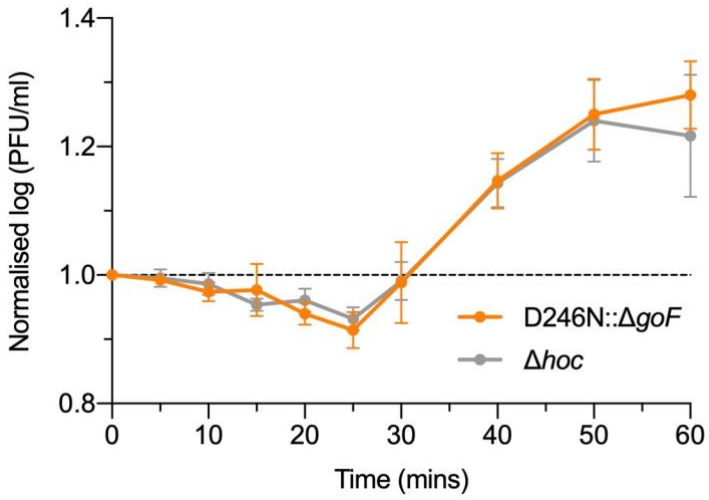

Fig.S4: Independent one-step growth curves of D246N::Δ*goF* and Δ*hoc* phages. PFU/ml values were logged and normalised to the initial timepoint (t = 0) i.e. plotted values represent the log-change relative to phage density at t = 0. The dashed line at 1.0 represents the expected value in the absence of change from the initial phage density. Datapoints and error bars were plotted as mean ± SEM of three independent biological replicates (N = 3) for each phage genotype tested.

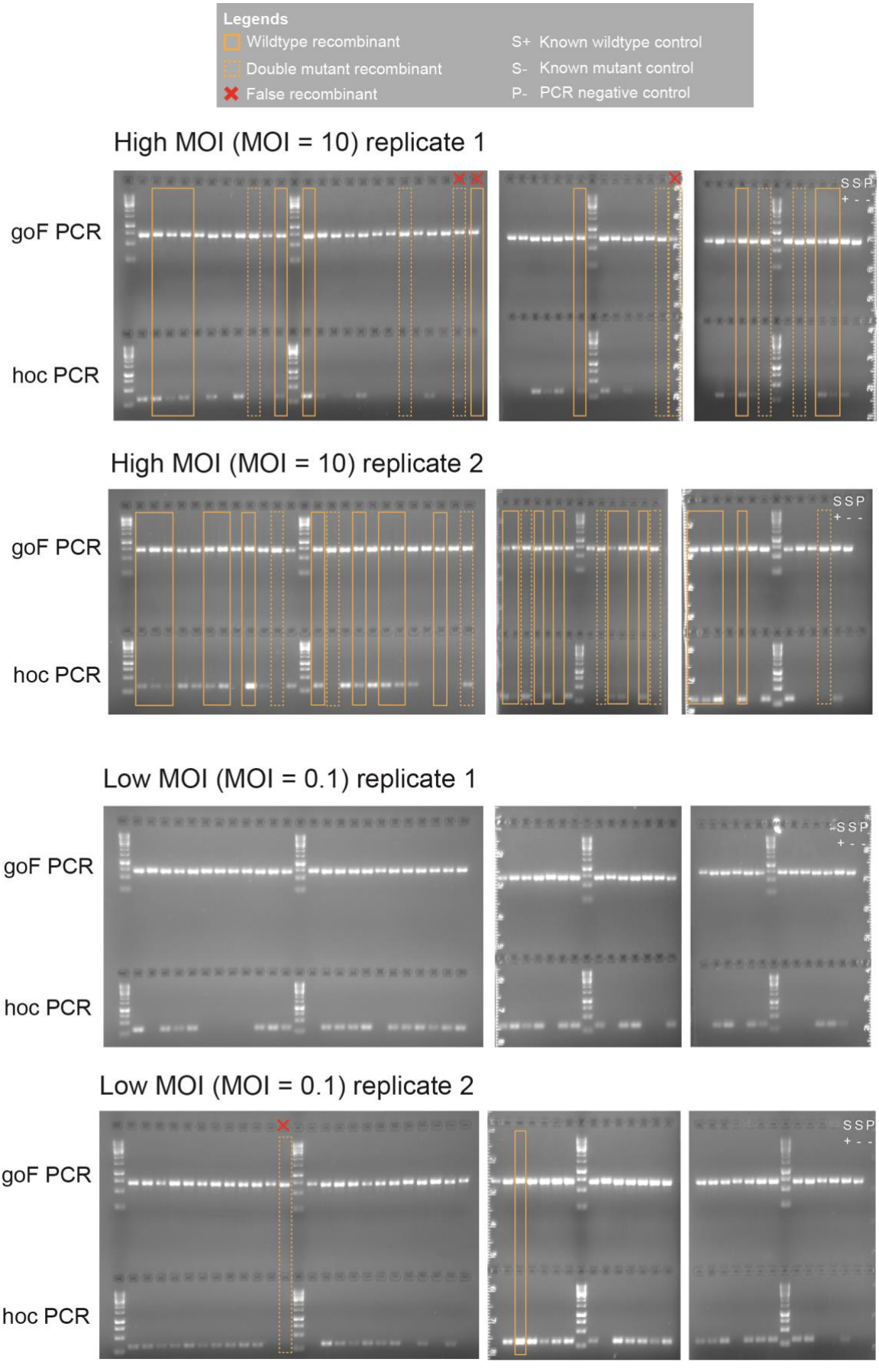
Legends
Wildtype recombinant
Double mutant recombinant
False recombinant
S+ Known wildtype control
S- Known mutant control
P- PCR negative control
High MOI (MOI = 10) replicate 1
goF PCR
hoc PCR
SSP
+ - -
High MOI (MOI = 10) replicate 2
goF PCR
hoc PCR
SSP
+ - -
Low MOI (MOI = 0.1) replicate 1
goF PCR
hoc PCR
SSP
+ - -
Low MOI (MOI = 0.1) replicate 2
goF PCR
hoc PCR
SSP
+ - -

Fig.S5: PCR gels of phage recombinant screenings from recombination experiment at high and low multiplicities-of-infections (MOIs 10 and 0.1), with two independent replicates per MOI condition. A total of 49 phage isolates were screened per condition per replicate alongside three controls: i) known wildtype sample control (S+), known mutant control (S-), and sample negative PCR control (P-) to assess for contamination. PCR for *goF* on the top row of the gels discriminates between wildtype (ancestral) and the mutant phage genotypes by observing for band shifts corresponding to the ∆21bp *goF* mutation. Meanwhile, PCR for *hoc* on the bottom row of the gels discriminates between wildtype and ∆*hoc* phage genotypes for the presence or absence of PCR product, respectively. Solid rectangles indicate wildtype recombinants, possessing both wildtype alleles of the *goF* and *hoc* gene, while dashed rectangles indicate double mutant recombinants, possessing both ∆21bp *goF* and ∆*hoc* mutations. All recombinants were verified via a separate PCR and false recombinants were labelled accordingly (i.e. red crosses) and were discounted from the total recombinant count. We note that the false recombinants were largely arising from ambiguous amplified products in the initial PCR screen.

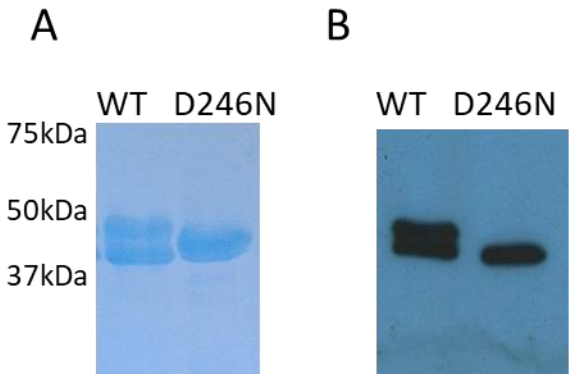

Fig.S6: A) Coomassie stained SDS-PAGE of purified His-WT Hoc and His-D246N Hoc. WT Hoc purified as a doublet. B) Subsequent probing of Western transfer with anti-His antibodies (R&D Systems) shows that both the full-length and partially degraded WT Hoc are His-tagged, implying C-terminal degradation of the protein. The D246N Hoc mutant has also undergone truncation from the C-terminus to produce a stable N-terminally His-tagged Hoc protein of the same size as the WT degradation product. These events did not impact upon the three N-terminal Ig-like domains, with the loss of approximately 5kDa from the C-terminus of the fourth domain (the capsid binding domain).

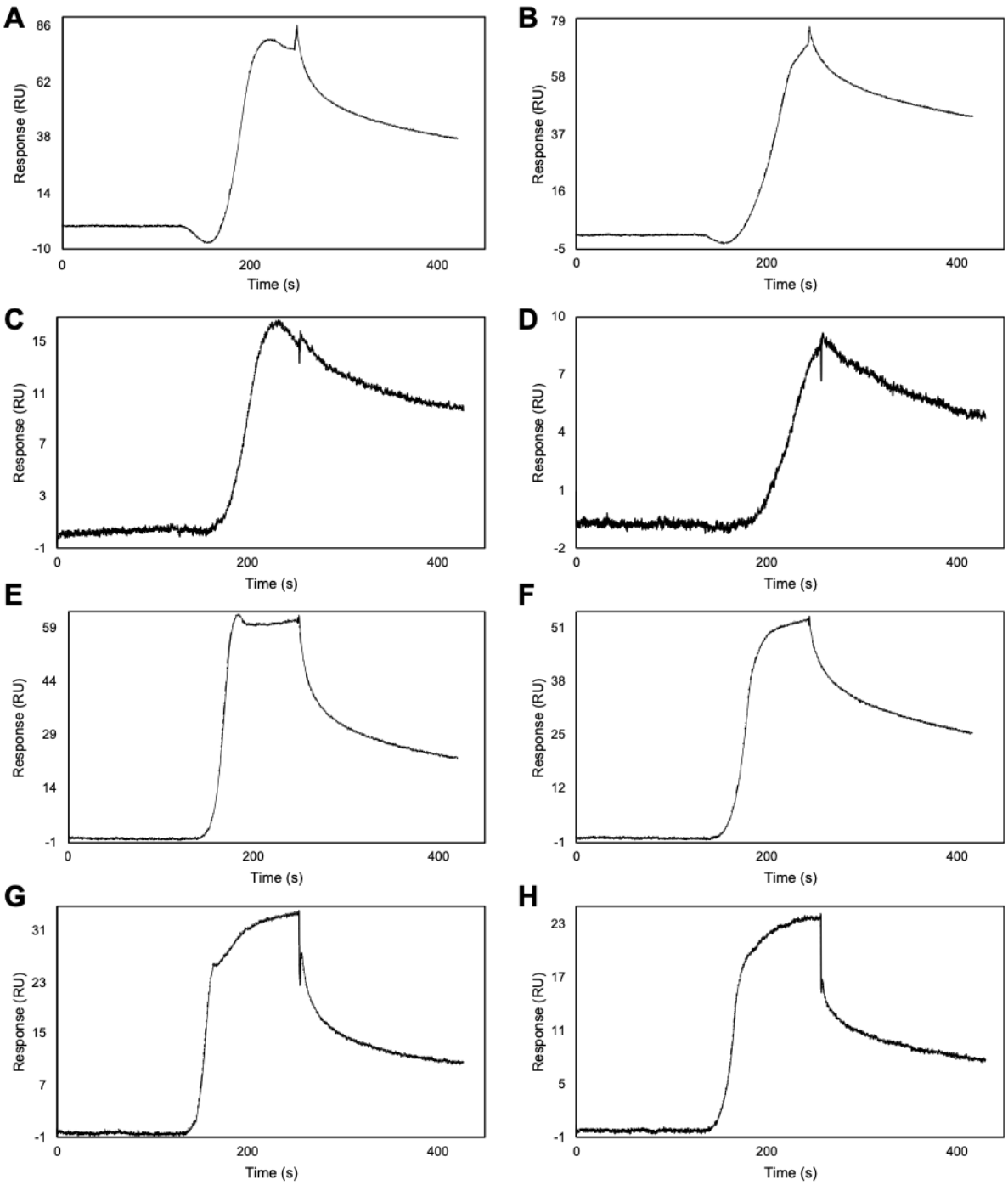

Fig.S7: SPR sensorgrams of the WT Hoc (left column) and D246N Hoc mutant (right column) with various glycans using OneStep analysis. (A-B) 2'-Fucosyllactose; (C-D) Lacto-N-difucohexaose II; (E-F) Lewis^{y}, and (G-H) Lacto-N-fucopentaose II.

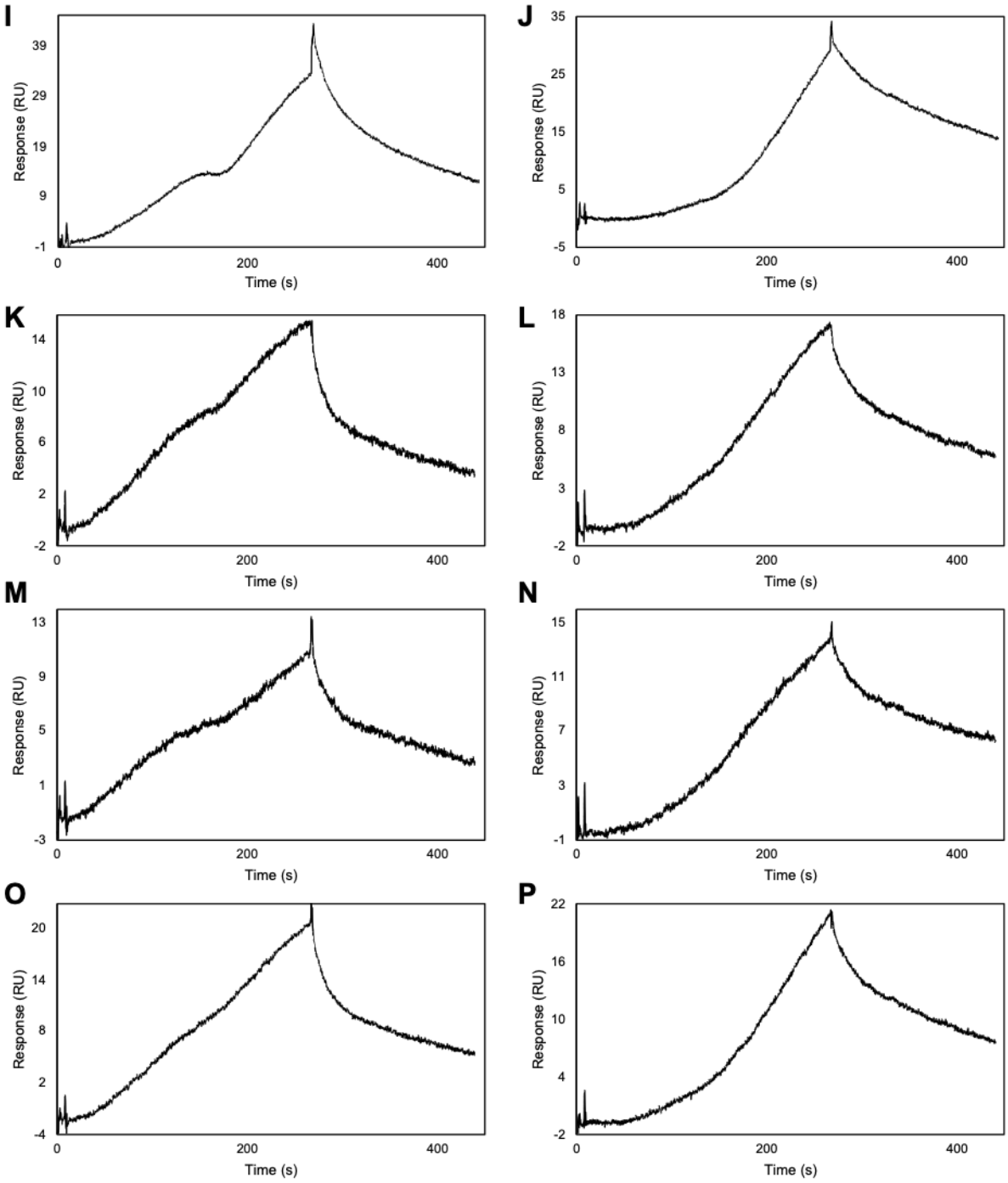

Fig.S7 (*continued*): SPR sensorgrams of the WT Hoc (left column) and D246N Hoc mutant (right column) with various glycans using OneStep analysis. (I-J) Lacto-N-fucopentaose I; (K-L) Lewis^{a}; (M-N) Blood Group A Trisaccharide, and (O-P) Lewis^{x}.

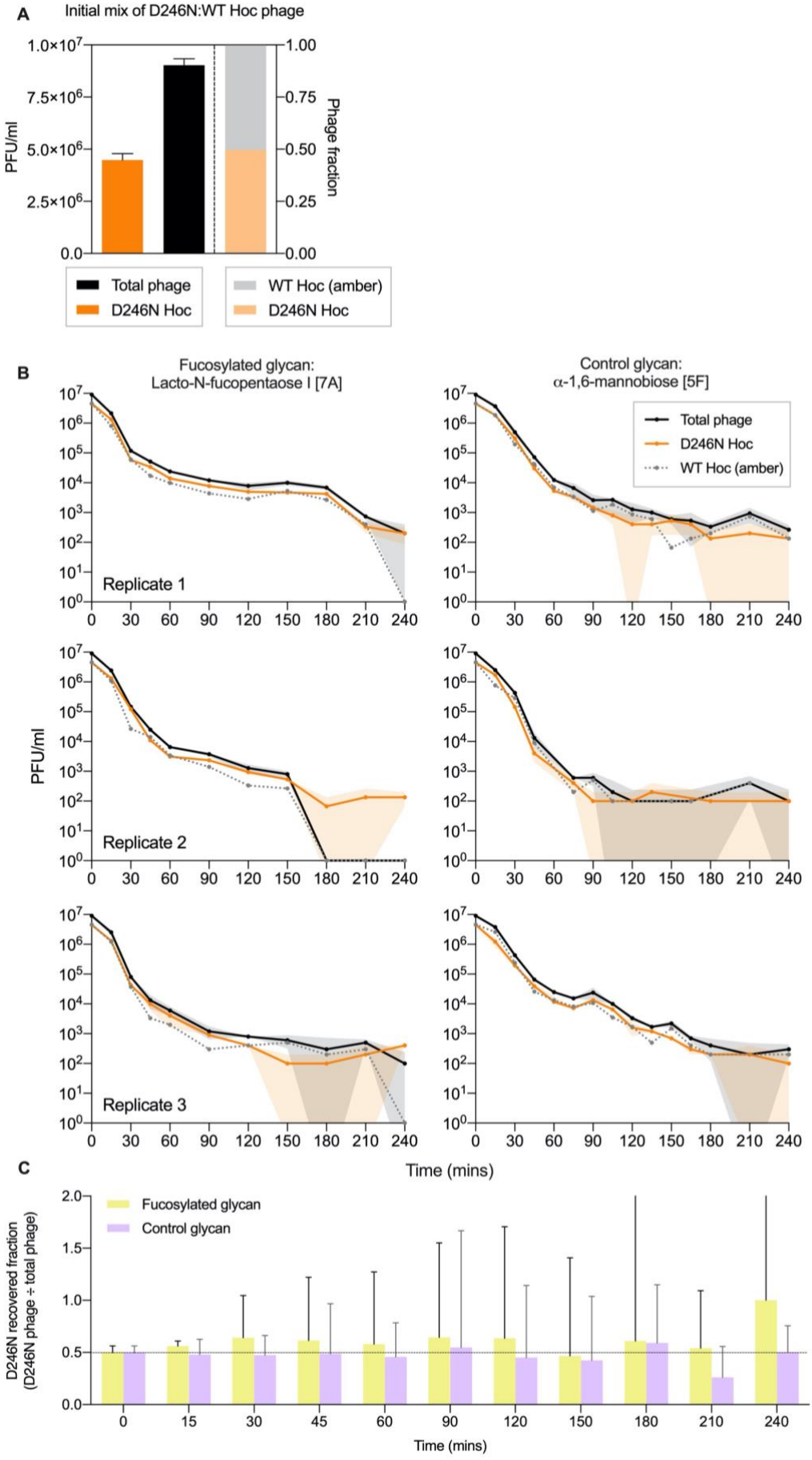

Fig.S8: A) Phage titre and1:1 fraction of D246N::Δ*goF* phage and wildtype (WT) Hoc with amber *43*$^{-}$ */44*$^{-}$ mutation phage inoculated in the gut-on-a-chip replicates for competitive phage-glycan washout.

Individual replicate data for competitive phage-glycan washout from the gut-on-a-chip between WT Hoc (amber *43⁻/44⁻* mutant) and experimentally evolved D246N Hoc mutant under 1 mM A) control glycan: α-1,6-mannobiose and B) fucosylated glycan: Lacto-N-fucopentaose I treatments. C) Comparison on the D246N::Δ*goF* phage fraction recovered from the gut-on-a-chip between the fucosylated and control glycan perfusion per sampled timepoint. Lines and shaded regions in panel B for total phage and D246N::Δ*goF* phage fraction are plotted as mean ± SEM of three technical replicates per timepoint (n = 3). The dotted line in panel B is a representative fraction of the WT Hoc (amber *43⁻/44⁻* mutant) phage by subtracting the D246N::Δ*goF* phage fraction from the total phage number. Bar chart in panel C is plotted as mean ± SEM of three biological replicates per timepoint (N = 3). The dotted line in panel C at 0.5 represents the expected fraction if D246N::Δ*goF* phage and the WT Hoc (amber *43⁻/44⁻* mutant) phages were recovered in equal proportions.

Dataset S1: qPCR standard curves

Refer to: SI_Dataset 1_qPCR_std_curves.xlsx

Dataset S2: List of fixed ancestral and standing background mutations

Refer to: SI_Dataset 2_background_muts.xlsx

Dataset S3: List of *de novo* mutations.

Refer to: SI_Dataset 3_denovo_muts.xlsx

Dataset S4: Selection coefficient calculations

Refer to: SI_Dataset 4_Selection_coefficient_analysis.xlsx

Dataset S5: Whole phage and Hoc protein glycan array heatmaps.

Refer to: SI_Dataset 5_glycan array.xlsx

Dataset S6: Mutational calling scripts (Breseq and mutational comparisons)

Refer to: SI_Dataset 6_mutational_calling_script.txt

SI References:

1. C. Lee, S. Lee, S. G. Shin, S. Hwang, Real-time PCR determination of rRNA gene copy number: Absolute and relative quantification assays with Escherichia coli. *Appl. Microbiol. Biotechnol.* **78**, 371–376 (2008).
2. M. Fittipaldi, *et al.*, Discrimination of infectious bacteriophage T4 virus by propidium monoazide real-time PCR. *J. Virol. Methods* **168**, 228–232 (2010).
3. N. Bonilla, *et al.*, Phage on tap-a quick and efficient protocol for the preparation of bacteriophage laboratory stocks. *PeerJ* **2016**, e2261 (2016).
4. D. Subedi, J. J. Barr, Temporal Stability and Genetic Diversity of 48-Year-Old T-Series Phages. *mSystems* **6** (2021).
5. D. E. Deatherage, C. C. Traverse, L. N. Wolf, J. E. Barrick, Detecting rare structural variation in evolving microbial populations from new sequence junctions using breseq. *Front. Genet.* **5**, 1–16 (2015).
6. L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2016).
7. O. Blixt, *et al.*, Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17033–17038 (2004).
8. N. X. Arndt, J. Tiralongo, P. D. Madge, M. Von Itzstein, C. J. Day, Differential carbohydrate binding and cell surface glycosylation of human cancer cell lines. *J. Cell. Biochem.* **112**, 2230–2240 (2011).
9. M. E. Huflejt, *et al.*, Anti-carbohydrate antibodies of normal sera: Findings, surprises and challenges. *Mol. Immunol.* **46**, 3037–3049 (2009).
10. C. J. Day, *et al.*, Differential Carbohydrate Recognition by Campylobacter jejuni Strain 11168: Influences of Temperature and Growth Conditions. *PLoS One* **4**, e4927 (2009).
11. O. Cooper, *et al.*, Functional Microarray Platform with Self-Assembled Monolayers on 3C-Silicon Carbide. *Langmuir* **36**, 13181–13192 (2020).
12. F. S. Willard, D. P. Siderovski, Covalent immobilization of histidine-tagged proteins for

surface plasmon resonance. *Anal. Biochem.* **353**, 147–149 (2006).

13. B. Sanson, M. Uzan, Sequence and characterization of the bacteriophage T4 comC alpha gene product, a possible transcription antitermination factor. *J. Bacteriol.* **174**, 6539–6547 (1992).

14. E. S. Miller, *et al.*, Bacteriophage T4 Genome. *Microbiol. Mol. Biol. Rev.* **67**, 86–156 (2003).