Selection for nonspecific adhesion is a driver of FimH evolution increasing Escherichia coli biofilm capacity

Mari Yoshida, Stanislas Thiriet-Rupert, Leonie Mayer, Christophe Beloin, Jean-Marc Ghigo

microLife, 2022, **3**, 1–14
DOI: 10.1093/femsml/uqac001
Advance access publication date: 5 April 2022
Research Article

Selection for nonspecific adhesion is a driver of FimH evolution increasing *Escherichia coli* biofilm capacity

Mari Yoshida, Stanislas Thiriet-Rupert, Leonie Mayer, Christophe Beloin* and Jean-Marc Ghigo*

Institut Pasteur, Université de Paris, CNRS UMR 6047, Genetics of Biofilms laboratory, Paris F-75015, France
***Corresponding author:** Institut Pasteur, Université de Paris, CNRS UMR 6047, Genetics of Biofilms Laboratory, 75015 Paris, France. Tel: 00 33 1 40 61 34 18; 00 33 1 44 38 95 97; E-mail: cbeloin@pasteur.fr; jmghigo@pasteur.fr
One sentence summary: Experimental evolution showed that mutations in type 1 fimbriae FimH lectin are the main drivers of the acquisition of nonspecific surface adhesion and biofilm capacity in natural *Escherichia coli* isolates.
Editor: Bernt Eric Uhlin

Abstract

Bacterial interactions with surfaces rely on the coordinated expression of a vast repertoire of surface-exposed adhesins. However, how bacteria dynamically modulate their adhesion potential to achieve successful surface colonization is not yet well understood. Here, we investigated changes in adhesion capacity of an initially poorly adherent *Escherichia coli* strain using experimental evolution and positive selection for mutations improving adhesion and biofilm formation on abiotic surfaces. We showed that all identified evolved populations and clones acquired mutations located almost exclusively in the lectin domain of *fimH*, the gene coding for the α-d-mannose-specific tip adhesin of type 1 fimbriae, a key *E. coli* virulence factor. While most of these *fimH* mutants showed reduced mannose-binding ability, they all displayed enhanced binding to abiotic surfaces, indicating a trade-off between FimH-mediated specific and nonspecific adhesion properties. Several of the identified mutations were already reported in the FimH lectin domain of pathogenic and environmental *E. coli*, suggesting that, beyond pathoadaptation, FimH microevolution favoring nonspecific surface adhesion could constitute a selective advantage for natural *E. coli* isolates. Consistently, although *E. coli* deleted for the *fim* operon still evolves an increased adhesion capacity, mutants selected in the Δ*fim* background are outcompeted by *fimH* mutants revealing clonal interference for adhesion. Our study therefore provides insights into the plasticity of *E. coli* adhesion potential and shows that evolution of type 1 fimbriae is a major driver of the adaptation of natural *E. coli* to colonization.

Keywords: biofilm, *Escherichia coli*, type 1 fimbriae, experimental evolution, positive selection

Introduction

Adhesion and subsequent formation of biofilms on environmental or host surfaces enable bacteria to withstand natural mechanical fluxes and it is an essential step of most colonization or infection processes (Beloin *et al.* 2005, Van Houdt and Michiels 2005, Fronzes *et al.* 2008, Kline *et al.* 2009, Patel *et al.* 2017). In commensal and pathogenic *Escherichia coli*, adhesion and biofilm formation is achieved using a large arsenal of adhesion factors (Beloin *et al.* 2005, Kline *et al.* 2009), including proteinaceous adhesins protruding from the bacterial cell envelope and promoting nonspecific surface adhesion (e.g. amyloid curli), afimbrial autotransporter adhesins mediating aggregation via homotypic self-interactions (e.g. Ag43 adhesin) or polymeric fimbrial adhesins that specifically recognize oligosaccharidic ligands (e.g. type 1 fimbriae) (Korea *et al.* 2010, Larsonneur *et al.* 2016, Vo *et al.* 2017, Wernerburg and Thanassi 2018, Bhoite *et al.* 2019). *Escherichia coli* also secretes several polysaccharidic exopolymers, such as cellulose, ß-1,6-*N*-acetyl-D-glucosamine polymer and capsular polysaccharides, promoting or hindering adhesion on living or inert surfaces (Wang *et al.* 2004, Valle *et al.* 2006, Beloin *et al.* 2008, Matthysse *et al.* 2008, Romling and Galperin 2015). In addition to these well-characterized surface structures, the *E. coli* genome contains several partially characterized genes or operons encoding potential adhesion factors (Roux *et al.* 2005, Korea *et al.* 2010, Wurpel *et al.* 2013, Larsonneur *et al.* 2016). The expression of *E. coli* adhesins is controlled by diverse regulatory networks enabling rapid transcriptional change in response to stresses and environmental cues such as oxygen levels, pH and chemical gradients (Korea *et al.* 2010, Chagnot *et al.* 2013, Chahales and Thanassi 2015, Rossi *et al.* 2018). However, how *E. coli* controls its adhesive properties and coordinates the interplay between its adhesion factors while avoiding antagonistic physical interferences between adhesins is still poorly understood.

Here, we used a dynamic *in vitro* biofilm model amenable to experimental evolution to explore how *E. coli* adapts when subjected to positive selection for increased capacity to bind to an abiotic surface and form biofilms. We showed that, despite the diversity of *E. coli* adhesins, all end-point populations displaying increased biofilm biomass after positive selection almost exclusively acquired mutations in *fimH*, the gene coding for the α-D-mannose-specific tip adhesin of type 1 fimbriae enabling adhesion to mannosylated epithelial cells. The identified *fimH* mutations, including mutations found in environmental and clinical *E. coli* isolates, displayed enhanced capacity for initial adhesion to abiotic surfaces but reduced mannose-binding properties and outcompeted the strongest biofilm-forming mutant emerging from a selection performed in a strain lacking the whole *fim* operon. This indicates that selection of *fimH* mutants with increased nonspe-

Received: September 20, 2021. **Revised:** December 16, 2021. **Accepted:** March 31, 2022

cific adhesion capacity could provide commensal or pathogenic *E. coli* a selective advantage for surface colonization and persistence in the environment. Our study therefore provides direct insights into the plasticity of the *E. coli* adhesion repertoire and shows that biofilm formation is a powerful driver of the evolution of *E. coli* adhesion potential.

Materials and methods

Bacterial strains and growth conditions

Bacterial strains used in this study are listed in Table S7 (Supporting Information). *Escherichia coli* was grown in M63B1 minimum medium supplemented with glucose (0.4%) and kanamycin (20 μg/mL), and incubated at 37°C. All media and chemicals were purchased from Sigma-Aldrich Saint-Quentin-Fallavier, France.

Strain construction

The WT strains used in the experiments originated from an *E. coli* MG1655 K12 wild-type strain and were either tagged, at the lambda att site, with red (*mars*) or green fluorescent protein encoding genes (*gfpmut3*) (RFP/GFP) by P1 vir transduction. Strains without the whole *fim* operon (Δ*fimABCDEFGH*) were constructed by P1 vir phage transduction method from MG1655_Δ*fimAICDFGH::cat* (Korea *et al.* 2010) into the WT strains. The constitutive promoter controlling the *fim* operon (PcL*fim*) (Larsonneur *et al.* 2016) was transferred by lambda-red recombination into WT strains and individual *fim* mutants using pKOBEGA plasmid and lambda red recombination. The *ecpD/ecpR* mutant was reconstructed using P1 vir phage transduction into the MG1655 K12 wild-type strain tagged with either red (*mars*) or green fluorescent protein (*gfpmut3*). Primers used for genetic construction are listed in Table S8 (Supporting Information).

Biofilm formation in microfermenters

Continuous-flow biofilm microfermenters containing a removable glass spatula were used as described in Ghigo (2001), with internal agitation provided by filter-sterilized air-bubbling. Biofilm microfermenters were inoculated by placing the spatula in a culture solution adjusted to $OD_{600} = 1.0$ (containing 5.0×10^8 bacteria/mL) for 10 min. The spatula was then reintroduced into the microfermenter and biofilm culture was performed at 37°C in M63B1 with 0.4% glucose. Flow rate was then adjusted (30 mL/h) so that total time for renewal of microfermenter medium was shorter than bacterial generation time, thus minimizing planktonic growth by constant dilution of non-biofilm bacteria.

Positive selection procedure

At the beginning of the positive selection procedure (Day 0) – 80°C glycerol stocks of the parental *E. coli* GFP- or RFP-tagged strains were inoculated in culture tubes containing LB medium supplemented with kanamycin (20 μg/mL) for over-day 8 h culture, then transferred into M63B1 minimum medium supplemented with 0.4% glucose and kanamycin for overnight 15 h culture. Six removable glass spatulas were submerged into the GFP overnight culture adjusted to $OD_{600} = 1.0$ (~5.0×10^8 bacteria/mL) for 10 min, and six others with the RFP overnight culture. Each spatula was then reintroduced into one microfermenter and the bacteria adhered to the spatula were incubated for 8 h, during which fresh medium (M63B1 minimum medium supplemented with 0.4% glucose) flowed (30 mL/h) in the microfermenter, constantly diluting non-adhering, planktonic bacteria. After the 8 h incubation in the microfermenters, bacteria that developed as a biofilm onto the spatula were resuspended in culture tubes containing fresh minimum growth medium supplemented with glucose and kanamycin, and cultured overnight. A fraction of each overnight culture was used to prepare a glycerol stock that could be analyzed afterward, and the rest was then used for the inoculation for the next cycle of the positive selection experiment. See Supporting methods for additional details

Biofilm microtiter plate assay CV assay

The glycerol stocks of the target strains or populations were inoculated in LB medium supplemented with kanamycin (20 μg/mL) for over-day 5 h culture at 37°C, then inoculated in either a polystyrene Greiner 96-well plate or a polyvinyl chloride Corning plate at $OD_{600} = 0.01$ (5×10^6 cells/mL) in M63B1 minimum medium supplemented with 0.4% glucose and kanamycin. Incubation was done at 37°C for overnight, 16 h. The supernatant was removed from each well and Bouin solution (Sigma) was applied for 15 min to fix the biofilm attached to the well. Next the fixation solution was washed three times with water and biofilm was stained with 1% crystal violet solution (QCA) for 15 min. After removal of the crystal violet solution, biofilms were washed three times with water. For quantification of biofilm formation, dried stained biofilms were resuspended in 30% acetic acid and absorbance was measured at 585 nm using an infinite Tecan infinite M200 PRO plate reader. Biofilm formation in different strains is represented in values normalized to average biofilm formation in the control strains.

Competition assay for biofilm formation

The glycerol stocks of the target strains were inoculated in LB medium supplemented with kanamycin (20 μg/mL) for over-day 6 h culture at 37°C, then overnight in M63B1 minimum medium supplemented with 0.4% glucose and kanamycin. After overnight incubation, the cultures were adjusted to $OD_{600} = 1.0$ (~5.0×10^8 bacteria/mL). The target strains with different fluorescent tags were mixed in 1:1 ratio and the cell concentrations were verified by FACS and OD analyses. Each competition mix was then inoculated in either a polystyrene Greiner 96-well plate or a polyvinyl chloride Corning plate at $OD_{600} = 0.01$ (5×10^6 cells/mL) in M63B1 minimum medium supplemented with 0.4% glucose and kanamycin. Incubation was done at 37°C overnight, for 16 h. The supernatant was removed from each well and the biofilm was resuspended in 100 μL of PBS. The proportion of each strain was then assessed from this resuspension by FACS and the relative fitness of target strain was calculated as follows:

$$relative\ fitness = \frac{\left(\frac{Target_{post}}{Target_{pre}}\right)}{\left(\frac{ref_{post}}{ref_{pre}}\right)},$$

where $Target_{pre}$ and $Target_{post}$ are the cell concentrations of the target strain in the mixed culture before and after the overnight incubation, and ref_{pre} and ref_{post} are the cell concentrations of the reference strain in the mixed culture before and after the overnight incubation, respectively (Lenski *et al.* 1991).

Competition assay for growth in liquid cultures

The glycerol stocks of the target strains were inoculated in LB medium supplemented with kanamycin (20 μg/mL) for over-day 6 h culture, then for overnight in M63B1 minimum medium supplemented with 0.4% glucose and kanamycin. After overnight incubation, the cultures were adjusted to $OD_{600} = 1.0$ (~5.0×10^8 bacteria/mL). The target strains with the different fluorescent tags were mixed in 1:1 ratio and the cell concentrations were verified by FACS and OD analyses. Then they were transferred into

M63B1 minimum medium supplemented with 0.4% glucose and kanamycin at 1000 times dilution and incubated overnight. The concentrations of target and reference strains in the overnight culture was measured using FACS and spectrophotometer. The relative planktonic fitness during overnight culture were calculated as for competitions for biofilm formation assay (see above).

Competition assay for initial adhesion

Prior to the initial adhesion assay, the glycerol stocks of the target strains were inoculated in culture tubes containing LB medium supplemented with kanamycin (20 μg/mL) for over-day 6 h culture, then transferred into M63B1 minimum medium supplemented with 0.4% glucose and kanamycin for overnight culture at 37°C. The overnight culture was adjusted to $OD_{600} = 1.0$ (~5.0×10^8 bacteria/mL) and the target and reference strains with the different fluorescent tag were mixed in 1:1 ratio. The ratio was verified by FACS analysis using MACS Quant Myltenyi Bioanalyzer. Glass spatulas were inoculated with 15 mL of the mixed culture for 10 min and washed three times in 20 mL PBS. The bacteria attached to the glass spatula were resuspended in 10 mL minimum medium without glucose. The ratio of each strain in the resuspended solution was measured using FACS. The results were either expressed as (i) relative fitness for Fig. S2 (Supporting Information) and were calculated as for competitions for biofilm formation assay (see above), or (ii) for Fig. 4 as % gain within the population as follows:

$$(percent\ gain\ within\ the\ population)\ (\%)\ no\ number = (target_{adh} - target_{ino}) \times \frac{50}{target_{ino}},$$

where the measured percentage of the target strain in the inoculation mixture and the resuspension is shown as $target_{ino}$ and $target_{adh}$, respectively.

RNA extraction and RT-PCR

Three colonies of each strain were grown in LB medium supplemented with kanamycin (20 μg/mL) for over-day 6 h culture at 37°C, then overnight in 25 mL of M63B1 minimum medium supplemented with 0.4% glucose and kanamycin. RNA extractions were then performed using TRIzol Reagent (Ambion) following manufacturer instructions. Briefly, the equivalent of 5.0×10^7 cells were centrifuged for 5 min at 5000 rpm, the pellets were resuspended in 1 mL of TRIzol Reagent and incubated 5 min at room temperature. Two hundred microliters of chloroform was added and each tube was vortexed for 15 s and incubated for 5 min at room temperature before centrifugation for 15 min at 10 000 rpm and at 4°C. The upper phase containing the total RNA was then transferred in 500 μL of isopropanol. The tubes were mixed by inversion and incubated for 5 min at room temperature before centrifugation for 10 min at 10 000 rpm and at 4°C. The supernatant was discarded, and the pellets were washed by adding 1 mL of 70% ethanol. The tubes were mixed by inversion and incubated for 5 min at room temperature before centrifugation for 10 min at 10 000 rpm and at 4°C. After removing the supernatant, the pellets were air dried and resuspended in 50 μL of RNase free water. A DNase treatment was then applied to all samples using the TURBO DNase kit (Ambion). The resulting RNA was used as a matrix for reverse transcription using the AMV cDNA first strand synthesis kit (Roche). The produced cDNA was used to assess *ecpD* expression as well as 16S rRNA gene as a control.

Whole genome sequencing and analysis

Prior to genome extraction, the glycerol stocks of the target populations and clones were inoculated in LB medium for over-day culture till the OD_{600} reached around 1.0 (~5.0×10^8 bacteria/mL). The bacterial cells were collected from 2 mL of the culture and the genomic DNA was extracted using the Qiagen DNeasy Blood and Tissue kit. RNase digestion step was added during the genome extraction process. Sequencing libraries were prepared using Nextera XT or Nextera Flex DNA Library Preparation Kit. All samples were sequenced using Illumina HiSeq sequencer. Sequencing reads were preprocessed to remove low-quality or artifactual bases. We used fqCleaner v.0.5.0, a mini workflow implemented in Galaxy (Afgan *et al.* 2018) to process fastq files (quality trimming, duplicate and artifact filters). Mutations with a frequency superior to 5% were detected using *breseq* version 0.30.0 (Deatherage and Barrick 2014) with the consensus mode for clones sequences analyses and the polymorphism mode for population sequencing. In both cases default parameters were used.

Determination of the frequency of *fimH* mutation

The evolution of frequency of the identified mutations in *fimH* was performed using Sanger sequencing analysis of PCR products centered on the *fimH* region using the following oligonucleotides—oligo up: AGGATGACAGTGGCAACACA and oligo down: GTTTTGGCTTTTCGCACAAT. Small aliquots of the glycerol stocks corresponding to each positive selection cycle were diluted in water and used for PCR reaction (Thermo Phusion flash high-fidelity master mix). The PCR products were sent to Eurofins for purification and Sanger sequencing. The frequency of the mutations was calculated using QSVanalyzer (Carr *et al.* 2009), with a cutoff that does not allow detection of mutations lower than 5% frequency.

Determination of *fim* 'ON/OFF' status of evolved clones

Orientation of the 314 bp DNA segment harboring the *fimA* promoter (*fimS* region) was determined using a PCR-based assay using restriction fragment length dimorphism arising from the orientation-dependent location of a unique BstUI restriction site within the amplified DNA. Briefly, the switch (*fimS*) region was amplified from a sample of overnight cultures with oligonucleotides OL4 (5′ CCGTAACGCAGACTCATCCTC 3′) and OL20 (5′ GAGTTTTAATTTTCATGCTGCTTTCC 3′) to generate a 726 bp PCR product. DNA was amplified with Taq polymerase (Invitrogen) using the following PCR conditions: denaturing at 94°C for 5 min, followed by 30 cycles (94°C for 1 min, 58°C for 1 min and 72°C for 1 min) and a final extension of 10 min at 72°C. Samples were cooled at 8°C, 10 units of BstUI (Nex England Biolabs) were added to each reaction and incubation was conducted at 37°C for 3 h. Digested PCR products were resolved on 2% agarose gels. Using this assay, phase ON populations of bacteria yielded two DNA fragments 433 and 293 bp in length, whereas phase OFF populations yielded two fragments of 539 and 187 bp. Mixed populations contained a mixture of all four fragments.

Protein modelling and structure prediction

The 3D structure of the lectin domain of MG1655 FimH presented corresponds to the pdb model 1KLF (Hung *et al.* 2002). The 3D structures of the mutant FimH proteins were predicted using the PHYRE2 Protein Fold Recognition Server and visualized with the MacPymol Software (Schroedinger 2010, Kelley *et al.* 2015).

Type 1 fimbriae extraction

Surface exposed type 1 fimbriae was isolated by heat shock extraction: a 5 mL culture of each strain was grown in LB at 37°C for 16 h and OD_{600} = 10.0 equivalent of cell culture was harvested by centrifugation. The harvested cells were washed with 0.9% NaCl, collected by centrifugation and resuspended with 75 mM NaCl, 0.5 mM Tris–HCl, pH 7.4. The samples were incubated at 60°C for 20 min, cooled on ice for 3 min, centrifuged and then the detached adhesins present in the supernatant were precipitated with 10% TCA overnight. The proteins were collected by centrifugation (20 000 × *g*, 1 h, 4°C), then the acquired pellet was washed with 75% acetone and the proteins were dissolved in HEPES 10 mM.

Detection of FimA by western blot with anti-FimA antibodies

The heat extracted proteins from OD_{600} = 2 culture were suspended in 1× Laemmli buffer with 250 U of Benzonase Nuclease (Sigma E0114) and incubated for 5 min at 95°C. The protein extracts were run on Mini-PROTEAN TGX Stain-Free precast Gels (BioRad) in 1× TGX buffer and then transferred to nitrocellulose membrane using a Trans-Blot® Turbo Transfer System (Bio-Rad). Blocking was performed in a 5% solution of dry milk and 0.05% Tween 1× PBS (1× PBST) overnight at 4°C with agitation. The membranes were then incubated in 1× PBST with a polyclonal rabbit antiserum raised against FimA subunit of type 1 fimbriae (kindly given by Prof. Scott Hultgren) at 1:10 000 for 1 h at room temperature with agitation. Membranes were washed in 1× PBST and then incubated with the secondary antibody (anti-rabbit IgG conjugated with horse radish peroxidase at 1:10 000, Promega). After washing the excess secondary antibody, specific bands were visualized using the ECL prime detection method (GE Healthcare).

Yeast agglutination assay

The capacity of WT and *E. coli* mutants expressing FimH variant to bind yeast mannosylated proteins and agglutinate yeast cells was assessed as previously described (Hasman *et al.* 2000). See the Supporting Methods (Supporting Information) for details.

FimH sequences analysis

The FimH sequence from the ancestral strain (*E. coli* K12, strain MG1655) used in the positive selection experiment was used as a query for BLASTp searches (blast+ version 2.2.31; Camacho *et al.* 2009) against the NCBI nrprot database (version 2019-04-09) as well as against a custom database composed of 2067 sequenced genomes of *E. coli* (Table S2, Supporting Information) and 277 genomes from environmental strains published in Touchon *et al.* (2020). Only hits corresponding to *E. coli* and with an e-value lower than $1e^{-10}$ and sequence identity higher than 70% were kept. In our analysis, the detection of one mutation in two identical sequences coming from different origin was important. Therefore, the redundancy of the FimH sequences dataset was not addressed in terms of sequence identity, but rather in term of isolates. To avoid such duplicated sequences, FimH proteins with identical identifiers or strain information were discarded. The resulting nonredundant dataset was composed of 3543 sequences (Table S3A and B, Supporting Information).

All sequences were aligned using *mafft* version 7.407 (Yamada *et al.* 2016) with G-INS-i option and the resulting alignment was screened to identify mutations relative to the reference sequence from *E. coli* K12, strain MG1655, using custom perl script. The same alignment was used to compute the single amino-acids modifications frequencies for each sequence in the signal peptide, mannose lectin domain, linker part and the pilin domain. For each sequence, the mutation frequency was computed for the whole sequence, the signal peptide, the lectin domain and the pilin domain. These datasets were then compared using a one-way ANOVA. To identify positions with a significantly higher mutation frequency, the same method was performed using a sliding window approach.

Selection pressure analysis

See the Supporting Methods (Supporting Information).

Statistical analysis

Unpaired, nonparametric Mann–Whitney tests were performed using Prism 6.0 for Mac OS X (GraphPad Software, Inc.) for CV staining biofilm assay in which each experiment was performed at least 8–12 times. Unpaired *t*-test with Welch's corrections was performed in the case of continuous flow biofilm experiments in which each experiment was performed three times.

Results

Positive selection of *E. coli* mutants with increased biofilm capacity

We subjected the poor biofilm-forming strain *E. coli* MG1655 to positive selection for spontaneous mutants with increased adhesion capacity using continuous-flow biofilm microfermenters (Fig. 1A). We conducted 12 parallel evolution experiments with 15 cycles of 8 h biofilm formation followed by overnight planktonic growth (15 days). The capacity of biofilm formation in microfermenters was enhanced for 6 of the 12 evolved population, with a 10- to 100-fold increase for 4 of them as compared with the wild-type (WT) ancestor (G6, R4, R5 and R6) (Fig. 1B and C; Fig. S1A, Supporting Information). These six populations also displayed an enhanced capacity to form biofilms in microtiter plate (Fig. S1B, Supporting Information) showing that the evolved trait was robust over two different biofilm models. To study the dynamic of the evolution of the adhesion capacities of the different population we performed microtiter plate biofilm assays for each cycle of the different population (Fig. 1D). Biofilm formation started to increase at cycle 5 for the G6 population and around cycle 10 for the population G1, G3, R4, R5 and R6 (Fig. 1D). Consistently, the enhanced biofilm formation evolved trait was also observed at clone level since heterogeneous but significant increase in biofilm capacity compared with the ancestral WT strain was shown in 15 individual clones isolated from each of the six more adherent populations (Fig. 1E; Fig. S1B, Supporting Information).

End-point biofilm-positive mutants carry mutations in the gene encoding the type 1 fimbriae tip adhesin FimH

To identify the nature of the mutations leading to increased adhesion, we sequenced and compared the ancestral population to the genome of eight populations at the end of 15 selection cycles, including the six biofilm populations with enhanced biofilm forming capacity (G1, G3, G6, R4, R5 and R6) and six populations that showed moderate or no biofilm increase (G2, G4, G5, R1, R2 and R3) (Fig. S1B, Supporting Information). We also sequenced four individual clones isolated from each of the six biofilm positive populations and two of the six biofilm negative populations. The populations and clones with enhanced biofilm capacities all carried

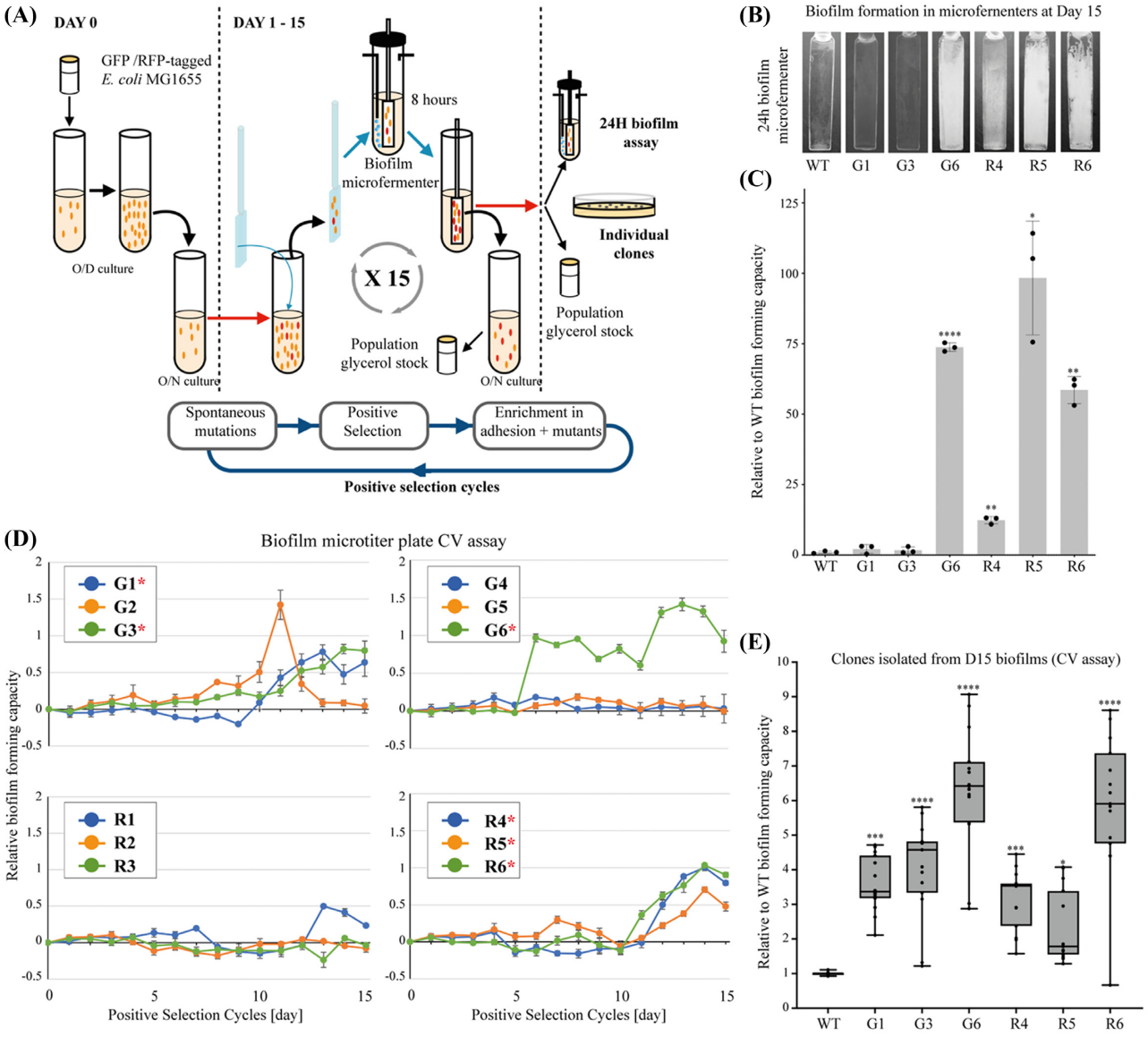

Figure 1. Selection of *E. coli* mutants with increased biofilm capacity. **(A)** Schematics of positive selection to identify GFP- or RFP-tagged *E. coli* MG1655 mutants with increased biofilm formation capacity. **(B)** Comparison of biofilm formation on biofilm microfermenters spatula of end-point populations. **(C)** Comparison of biofilm biomass of end-point populations and controls in biofilm microfermenters. The biofilm biomasses were collected from the spatula incubated for 24 h and the biofilm-forming capacities were calculated relative to the one obtained with the control ancestral WT (set to 1). **(D)** Evolution of biofilm forming capacity of bacterial population subjected to positive selection monitored by performing crystal violet biofilm assay using an aliquot of the population produced at each selection cycle. Comparison with the parental wild-type strain and normalization were performed by calculating relative value using WT *E. coli* MG1655 (poor biofilm former) and *E. coli* TG1 strains as follows: $(OD_{test} - OD_{WT})/(OD_{TG1} - OD_{WT})$. When the value is equal to 1 and 0, the capacity is as same as the TG1 strain and the WT MG1655 parental strain, respectively. The experiments were performed in triplicate. G1–6 populations derive from parental GFP-tagged MG1655. R1–6 populations derive from parental RFP-tagged MG1655. The selected populations with increased biofilm capacity are indicated with an asterisk. **(E)** Comparison of biofilm forming capacities of 15 clones isolated from the selected population at cycle 15 by crystal violet biofilm assay. Each point represents the data of an individual clone. Each solid line in the boxplot shows the median of the biofilm forming capacities, and the boxes illustrate the first and the third quartiles. The relative biofilm capacities were calculated using the capacities of the parental wild-type strain (median set to 1). All crystal violet biofilm assay experiments were conducted with at least eight replicates. Continuous flow biofilm experiments in microfermenters were performed three times (panel C) and statistics correspond to unpaired *t*-test with Welch's corrections. In panel E (12 replicates) statistics correspond to unpaired, nonparametric Mann–Whitney test comparing all conditions to ancestral WT. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$.

mutations at high frequencies located in the gene coding for FimH and no *fimH* mutation was detected in the six populations showing no increase in biofilm capacity after 15 cycles of positive selection (Fig. 2A; Table S1, Supporting Information). Mutations localized upstream of *flhD*, encoding a regulator of flagella production, were identified at low frequency in some biofilm positive populations. Although these mutations enhanced flagellin production, they had no impact on motility (Fig. S2, Supporting Information). In some populations that did not evolve increased biofilm capacity, we also identified several mutations in *waaB*, which codes for an enzyme adding a galactose on the outer core of LPS and reported to impact mucoidy or adhesion (Qian *et al.* 2014, Wang *et al.* 2015, Ren *et al.* 2016, Smith *et al.* 2017). However, we did not observe any difference between WT and *waaB* mutants in colony mucoidy

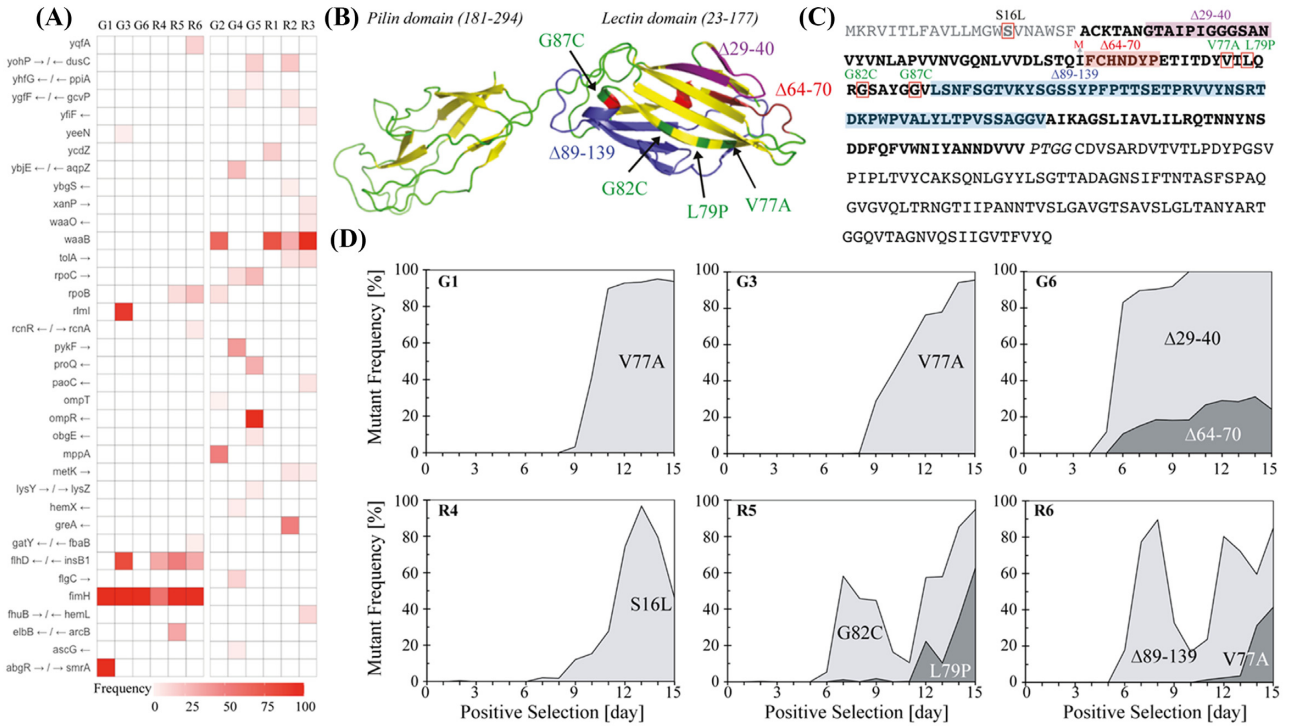

Figure 2. Nature and frequency of FimH mutations identified in evolved populations with increased biofilm capacity. **(A)** Frequency of mutation in targeted genes. Frequency of mutations inferred from population sequencing in the six populations with increased biofilm formation capacity (left) and the six populations without increased biofilm capacity (right). The frequency is color coded in red according to the legend at the bottom of the heat map and corresponds to the total frequency of all mutations in each gene within a population at cycle 15 of the experimental evolution. **(B)** Identified mutations in FimH protein structure. The locations of amino acid substitutions are indicated with black arrows and the in-frame deletions are indicated in purple (Δ29–40), red (Δ64–70) and blue (Δ89–139). **(C)** Identified FimH mutations are indicated above the FimH amino acid sequence. Single amino acid substitutions are highlighted with red boxes and three deletions are highlighted in pale purple, red and blue. The G87C mutation was identified in a distinct positive selection experiment using less selection cycles (4 instead of 15) and longer residence in the biofilm microfermenters at each cycle (18 h instead of 8 h). **(D)** Frequency of indicated FimH mutations acquired during the positive selection experiment in populations with increased adhesion capacity (G1, G3, G6, R4, R5, R6). The grey areas show the frequency of the mutants and the white area shows the frequency of the wild-type strains. The frequencies were calculated using Sanger sequencing of *fimH* region of sampled populations.

nor modification of biofilm formation (Fig. S3, Supporting Information).

FimH is the α-D-mannose-specific adhesin of type 1 fimbriae, a major *E. coli* adhesin enabling epithelial cell colonization and shown to be critical for biofilm formation on abiotic surfaces (Pratt and Kolter 1998). FimH is a 300 residues and 31.473 kDa protein consisting of a pilin domain (amino acid 181–294) and a mannose-binding lectin domain (amino acid 23–177) connected through a 4-amino acid linker peptide chain (Fig. 2B).

We identified three in-frame deletions (Δ29–40 and Δ64–70, and Δ89–139) and three non-synonymous substitutions (V77A, L79P, G82C) in the lectin domain of the FimH protein, as well as one non-synonymous substitution (S16L) in the signal peptide of the FimH protein (Fig. 2BC). Some *fimH* mutations coexisted within populations: L79P and G82C were found in R5 population, V77A and Δ89–139 in R6 population, and Δ29–40 and Δ64–70 in G6 population. By contrast, R4 population only had the S16 L substitution, and G1 and G3 populations only had V77A mutations (Table S1, Supporting Information).

In order to check if this outcome was specific to the parameters used in our selection protocol, we ran another experiment using different selection pressures with less selection cycles (4 instead of 15) and longer residence time in biofilm microfermenters at each cycle (18 h instead of 8 h). This protocol also resulted in the selection of *fimH* mutation (G87C), emphasizing the predominant role of this adhesin in enhanced binding capacities and biofilm formation (Fig. 2BC). In order to determine the dynamics of emergence and population spread of the identified *fimH* mutations over time, we performed a Sanger sequencing analysis of *fimH* region PCR products of populations collected at each of the 15 positive selection cycles that led to increased biofilm capacities (Fig. 2D). All *fimH* in-frame deletions and the G82C mutation emerged as early as the fourth/fifth cycle, while the other non-synonymous FimH substitutions emerged at cycle 6 or later.

The phenotypic differences between *fimH* mutants are not due to phase variable expression of the *fim* operon

We isolated individual clones carrying an independent mutation in FimH and measured their biofilm capacity using continuous biofilm microfermenters. While most tested evolved clones showed close to wild-type growth rate, three mutations (V77A, L79P and Δ29–40) slightly impacted growth (Fig. S4A and B, Supporting Information) and 1:1 competition between these 3 clones and WT confirmed this fitness default (Fig. S4C, Supporting Information). All clones but the one carrying L79P, demonstrated increased biofilm-forming capacity, i.e. 1.8–113 times more than WT ancestral strain after 24 h of incubation in microfermenters (Fig. 3A). The mutants encoding FimH deletions Δ64–70 and Δ89–139 displayed the highest biofilm formation, relative to the ancestral WT (Fig. 3B). This enhanced capacity to form biofilms could

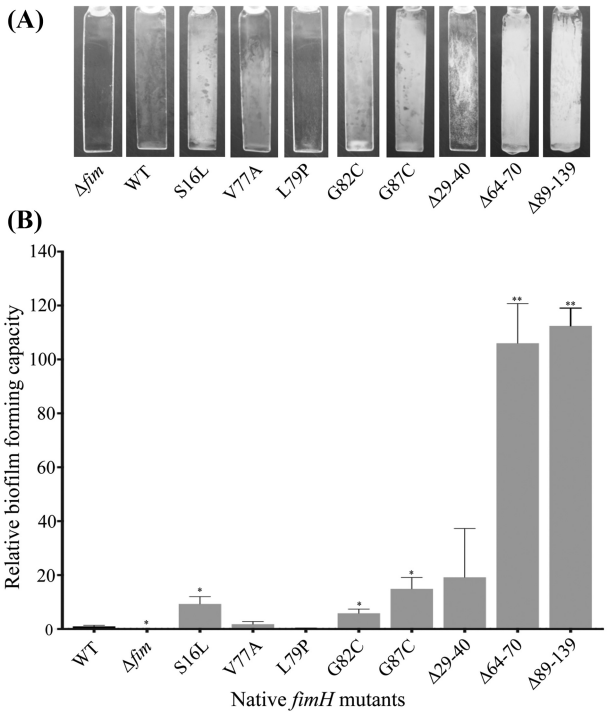

Figure 3. Biofilm formation capacity of individual *fimH* mutants. Biofilm formation of *fimH* mutants was compared after 24 h incubation in biofilm microfermenters. **(A)** Pictures of representative biofilm biomass formed on biofilm microfermenters spatulas. **(B)** The relative biofilm-forming capacities of individual *fimH* mutants was determined by comparison with the capacities of the Δ*fim* (0) and wild-type ancestor strains (1). Microfermenter biofilm experiments were performed in triplicate and statistics correspond to unpaired *t*-test with Welch's corrections. $^{*}P < 0.05$; $^{**}P < 0.01$.

be linked to phase variable expression of the *fim* operon that enable rapid on/off change of type 1 fimbriae production by inversion of the *fimA* promoter mediated by FimB and FimE recombinases (see *fim* region in Fig. S5A, Supporting Information) (Klemm 1986, Gally *et al.* 1996, van der Woude 2011). However, a test of the orientation of the *fim* promoter by PCR did not reveal any significant differences in ON/OFF status between wild-type parental and the evolved clones. The mutant clones were mainly OFF in planktonic culture and mainly ON in biofilm on microfermenter spatula, while none of the clones or the evolved populations displayed any mutations known to impact the *fim* switch, thus reflecting an enrichment for *fim* ON status in biofilm conditions (Fig. S5B and C, Supporting Information).

Evolved *fimH* mutations increase nonspecific initial adhesion

To characterize the phenotypic consequences of the identified *fimH* mutants in planktonic conditions and, considering the ON *fim* status of clones selected in biofilm conditions, we placed the WT and mutant *fim* operon (and therefore corresponding native or mutated *fimH* genes) under the control of the lambda PcL constitutive promoter (Korea *et al.* 2010). This led to normalized type 1 fimbriae expression, as shown using immunodetection with anti-FimA antibodies (Fig. S6, Supporting Information). Biofilm biomass formed in microfermenters of PcL-*fimH* mutants, compared with PcL-*fimH* WT, showed differences similar to those observed between (non-PcL) mutants and WT strains (compare Fig. 4A with Fig. 3), suggesting that, independently of *fim* phase variation, the selected FimH mutants displayed specific properties enhancing *E. coli* biofilm formation. Interestingly, selected mutants out-competed the wild-type strain for increased initial adhesion to glass in 1:1 competition between WT PcL-*fimH* and mutant PcL-*fimH* (Fig. 4B). This fitness advantage suggests that increased initial nonspecific adhesion is one of the main drivers of the enhanced capacity to form biofilms of the *fimH* mutant compared with the wild-type ancestral strain. However, we detected a negative correlation between *fimH* mutant initial adhesion capacity and their biofilm capacity (Spearman's rho = –0.76, *P*-value = 0.037; Fig. S7, Supporting Information), suggesting a trade-off between initial adhesion and biofilm maturation.

We also evaluated the ability of PcL-*fimH* mutants to bind mannosylated proteins using a yeast agglutination assay performed under static or agitated conditions, where the latter condition promotes catch-bond adhesion (Thomas *et al.* 2002). In both situations, all FimH mutations, except G87C, reduced yeast agglutination capacity, with some mutations displaying almost no agglutination, very similar to a Δ*fimH* strain agglutination phenotype (Fig. S8, Supporting Information). This indicates a clear trade-off between FimH-enhanced adhesion to abiotic surfaces and FimH adhesion to mannose.

Mutations identified in FimH are also found in natural and clinical isolates

In order to assess whether the FimH mutations identified in our *in vitro* evolution experiments reflect evolutionary paths occurring in natural or clinical *E. coli* isolates, we retrieved all *E. coli* FimH sequences available in NCBI protein databank as well as 2067 sequenced genomes of *E. coli* (Table S2, Supporting Information). The resulting NCBI dataset composed of 3266 sequences (Table S3A, Supporting Information) shows a strong bias toward strains of human and animal origin (53.25% versus 1.81% for environmental strains, while no information is available for the remaining 45%), likely due to the predominant representation of health-related studies. To partly compensate for this skew, we added the *fimH* sequence of 277 *E. coli* environmental strains from a recent study (Touchon *et al.* 2020) (see the 'Materials and Methods' section; Table S3B, Supporting Information). When the 3543 total FimH protein sequences extracted from this dataset were compared with MG1655 FimH, 10.4% (369 FimH sequences) displayed amino-acid changes at the same positions of at least one of the eight mutations identified in our study (Table S3, Supporting Information). An enrichment for strains belonging to the B1 phylogroup was identified in these 369 sequences, relative to the whole dataset (Table S4A, Supporting Information). Interestingly, 64 FimH proteins showed mutations that were also identified in our study (one S16L, 18 V77A, one L79P, 42 G87C and two Δ29–40). The G87C mutation (= G66C when amino acid numbering starts after the signal peptide), a position previously reported to be under positive selection in uropathogenic *E. coli* strains, being the most widespread (Weissman *et al.* 2007, Chen *et al.* 2009). These results demonstrate that our study recapitulated some of the selection events driving natural evolution of FimH (Fig. 5A). The enrichment in B1 strains previously identified was not maintained in these 64 sequences, this phylogroup being rather depleted relatively to the whole dataset (Table S4A, Supporting Information). However, there was an enrichment for B2 and F strains (to which most ExPEC strains belong), suggesting that our experimental evolution reflects part of the selective pressures applied to ExPEC strains.

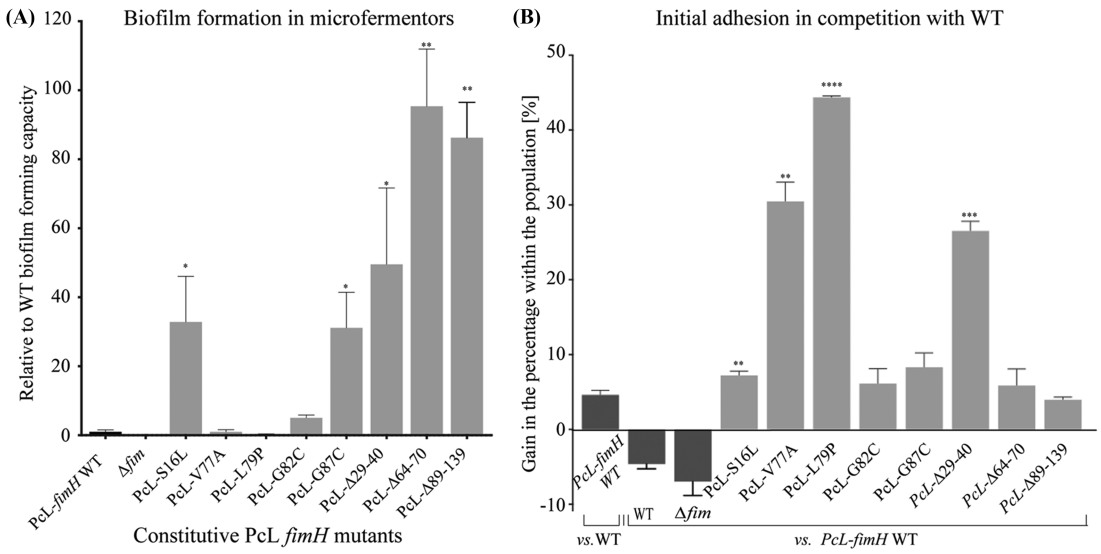

Figure 4. Biofilm and initial adhesion capacity of constitutive PcL *fimH* mutants. **(A)** Biofilm formation of Δ*fim* and PcL*fimH* WT strains (controls) and individual PcL*fimH* mutants after 24 h incubation in biofilm microfermenters. The relative biofilm-forming capacities of individual PcL *fimH* mutants was determined by comparison with the capacities of the Δ*fim* (0) and PcL*fimH* WT strains (1). All biofilm experiments were performed in triplicate. **(B)** Initial adhesion of PcL*fimH* mutants in competition with PcL*fimH* WT. The PcL*fimH* mutants were inoculated to the microfermenter spatula in 1:1 ratio with PcL*fimH* WT strain carrying either GFP or RFP fluorescent tags for 10 min, followed by washing, resuspension of the attached bacteria and CFU counting. The changes in the percentage of each strain within the population are shown. PcL*fimH* WT is also compared with the other control strains (WT and Δ*fim* strains, and PcL*fimH* WT strain with the different fluorescent tag). Statistics correspond to unpaired *t*-test with Welch's corrections comparing all conditions to ancestral WT. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$.

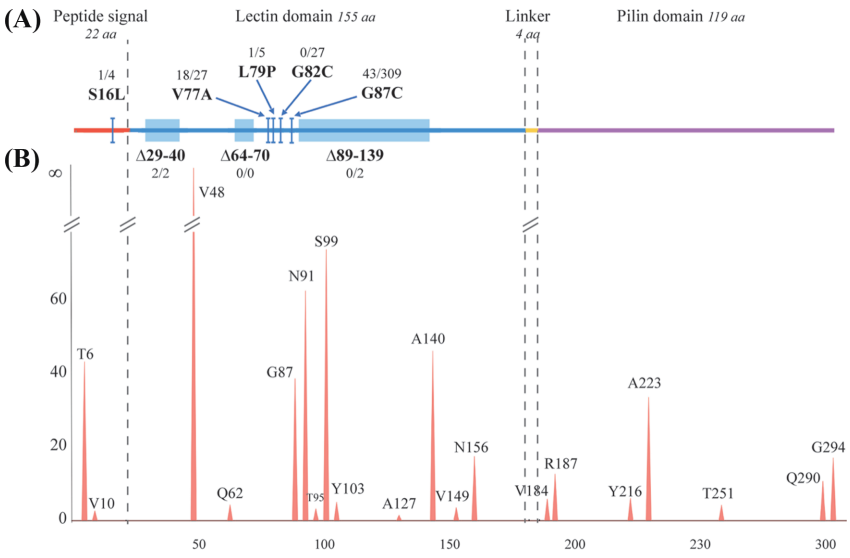

Figure 5. Analysis of the mutational spectrum of FimH sequences. **(A)** Linear structure and amino-acid size of FimH protein domains with the eight mutations selected during the experimental evolution. The two numbers provided for each mutation correspond to respectively, the number of mutations among the 3266 analyzed FimH sequences and the number of sequences having a mutation at the same position in this study (regardless of mutation type). **(B)** Negative log-transformed significance ($-\log_{10}$(*P*-value)) of the mutation frequency at each position of the sequence relatively to the frequency calculated over all sequences from Database sequences. A one-way ANOVA was performed, the resulting *P*-values were adjusted and a false discovery rate of 0.05 was considered as significant. Only positions with a significantly higher mutation frequency are plotted.

FimH mutational landscape diversity highlights its potential functional plasticity

All selected mutations in our study were found in the FimH lectin domain and none in its pilin domain. To test whether this was a general trend in FimH sequences or rather specific to our experimental settings, we screened all available FimH sequences for single amino-acid changes and calculated their frequency for each structural part of the FimH sequence (the signal peptide, the lectin domain, the four amino-acids linker and the pilin domain). The majority of these mutations were found in the lectin domain and were significantly higher (1.47-fold, one-way ANOVA: df = 1, *P*-value < $2.2e^{-16}$) than expected, while the pilin domain accumulated significantly less mutations (–2.31-fold, one-way ANOVA: df = 1, *P*-value < $2.2e^{-16}$) (Fig. 5B). To analyze the *E. coli* FimH mutational landscape in more details, the same analysis was carried out for each position of the sequence and 20 of them were found

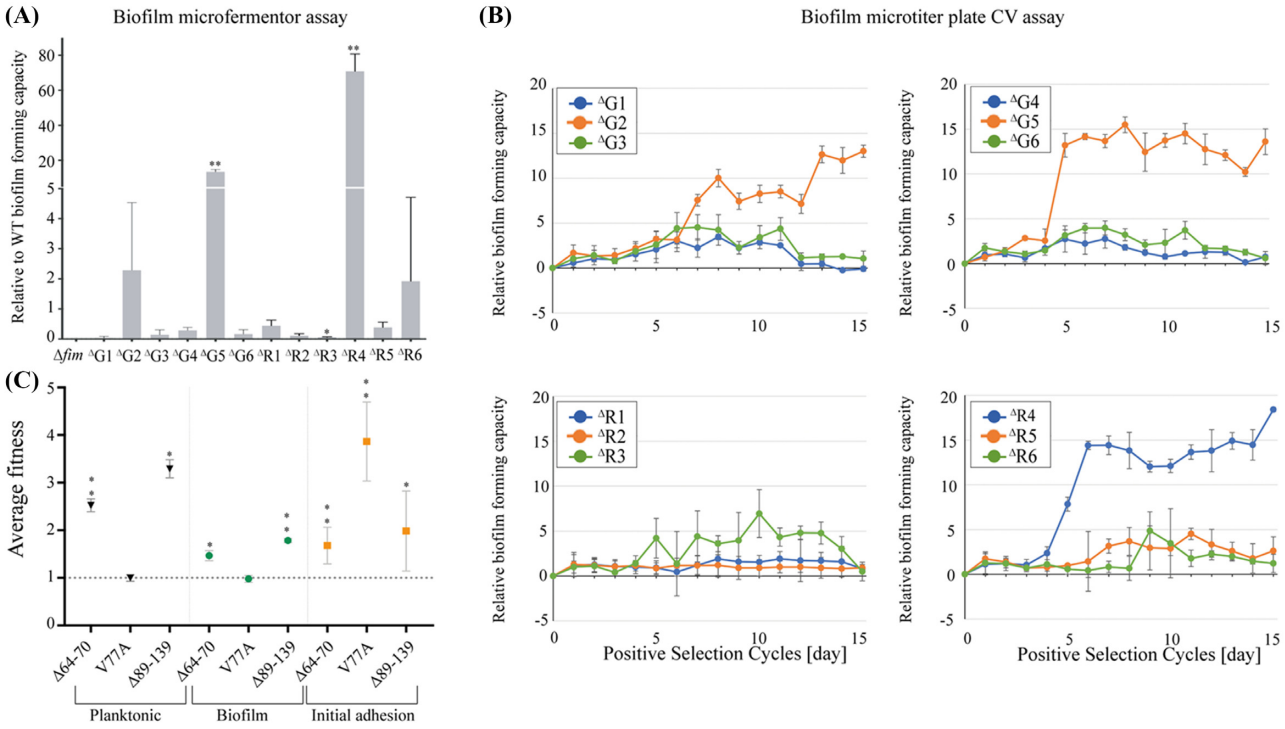

Figure 6. Evolution of biofilm forming capacity of Δ*fim* bacterial populations subjected to positive selection for adhesion. **(A)** The relative biofilm forming capacities of the populations at the end of the $^{\Delta fim}$positive selection experiment, after 24 h in microfermenters. The relative capacities are calculated using the capacities of the wild-type (1) and Δ*fim* ancestral strains (0). All microfermenter biofilm experiments were performed in triplicate. Statistics correspond to unpaired *t*-test with Welch's corrections. **(B)** The biofilm forming capacities were monitored at each cycle during the $^{\Delta fim}$positive selection experiment with the crystal violet assay of an aliquot of the enriched population as well as the parental wild-type strains. The relative capacities were calculated using *E. coli* MG1655Δ*fim* (very poor biofilm former) and *E. coli* MG1655 strain as follows: $(OD_{test} - OD_{\Delta fim})/(OD_{MG1655} - OD_{\Delta fim})$. When the value is equal to 1 and 0, the capacity is as same as the WT and the Δ*fim* strains, respectively. Each panel shows three populations: $^{\Delta fim}$G1-3, $^{\Delta fim}$G4-6, $^{\Delta fim}$R1-3 and $^{\Delta fim}$R4-6. All crystal violet biofilm assay experiments were conducted with at least eight replicates. **(C)** Comparison of fitness for three *fimH* mutants under native promoter competed against the best biofilm former mutant from the Δ*fim* evolution experiment (ecpR IS2(–) +5 bp/ecpD G89D). The strains were competed for planktonic growth, biofilm formation and initial adhesion on the microfermentor glass spatula as described in the 'Materials and Methods' section. Dotted line represents *ecpD/ecpR* mutant fitness set to 1. Statistics correspond to unpaired, nonparametric Mann–Whitney test comparing all conditions to WT Δ*fim*. $^{*}P < 0.05$; $^{**}P < 0.01$.

to be statistically more prone to mutation (Fig. 5B; Table S4B, Supporting Information). Interestingly, some of the mutations identified at these hotspots were found to co-occur in the same FimH sequences (Fig. S9, Supporting Information). Different phylogroup enrichments were also identified at these polymorphic sites (Table S4C, Supporting Information).

To bring further insight into the evolution of FimH sequences in *E. coli*, we built a phylogenetic tree (Fig. S10, Supporting Information) and analyzed it with the *codeml* program in PAML v4.9 (Yang 2007). For this purpose, we divided the tree into five subgroups (Sg1 to Sg5) based on phylogroups distribution along the tree (see methods). Overall, the evolutionary analysis showed a prevalence for purifying selection (Table S5A, Supporting Information). However, sequences belonging to Sg3 showed signs of episodic positive selection. Interestingly, most of the sites identified to be positively selected in this group were located in the lectin domain including positions 82 and 87 corresponding to evolved sites in our experimental evolution (Table S5A and B, Supporting Information). Moreover, Sg3 was enriched in strains belonging to phylogroup B2 (Table S4D, Supporting Information) as well as comprised all sequences having amino acids Ser-70 and Asn-78 (Ser-91 and Asn-99 in this study) described to be related to UPEC strains (Chen *et al.* 2009). Finally, a site partition analysis revealed that, although being both under purifying selection, the lectin and the pilin domains underwent a different evolutionary path in Sg3 and Sg5 (Table S5C and D, Supporting Information). Overall, these results demonstrate that the lectin domain of FimH is the most subjected to mutations both in laboratory evolved strains and in natural and clinical isolates, which could reflect complementary selection pressure in environments where bacteria need to withstand natural fluxes and perturbations.

In absence of *fimH*, evolution toward biofilm formation involves a broader range of mutations

Despite the diversity of *E. coli* K-12 surface structures known to contribute to biofilm formation (Korea *et al.* 2011), mutations leading to increased biofilm formation revealed by our study were surprisingly targeted to *fimH*. To test the evolution toward increased adhesion capacity in absence of *fimH*, we deleted the whole *fim* operon (*fimABCDEFGH*; see Fig. S5A, Supporting Information) and we ran 12 parallel positive selection experiments of the Δ*fim* strain for 15 cycles. We observed an evolution toward significant and relevant increased biofilm formation in both biofilm models for 3 out of 12 Δ*fim* populations, $^{\Delta}$G2, $^{\Delta}$G5 and $^{\Delta}$R4 (Fig. 6), with $^{\Delta}$G5 and $^{\Delta}$R4 achieving 14- and 71-times increased biofilm capacity in biofilm microfermenters (Fig. 6A). The analysis of the mutations found in these three biofilm-positive evolved Δ*fim* populations and in corresponding clones revealed mutations affecting biofilm associated functions such as chaperone-usher fimbrial surface structure (*yqiG*, *ecpR* and *ecpD*), autotransporter adhesin

(*ycgV*, *flu/agn43*) and flagellum (*flgH*, *ecpR*) (Table S6, Supporting Information). Whereas the *ecp* operon is not expressed in *E. coli* K12 MG1655 (Lehti *et al.* 2013), biofilm-promoting mutations in *ecpD* were only found in clones also displaying an IS insertion right after *ecpR*, the regulator of the operon. This IS insertion resulted in *ecp* operon expression (Fig. S11, Supporting Information), potentiating further mutations in *ecpD*, leading to increased biofilm formation in Δ*fim* population $^{\Delta}$R4. These results therefore showed that, in absence of type 1 fimbriae, different mutations in surface structure and adhesins could lead to increased biofilm capacity.

fimH mutants outcompete mutants selected in the absence of *fimH*

To investigate the potential origin of the lack of mutation diversity in evolution experiments performed in the WT background compared with those performed in a Δ*fim* background, we reconstructed the *ecpD/ecpR* mutant from the $^{\Delta}$R4 population (the highest biofilm-forming population from the Δ*fim* evolution) in a WT background. This mutant was then used in competition assays against the natural *fimH* mutants with highest biofilm capacity (Δ64–70 and Δ89–139) and increased, but lower biofilm capacity (V77A). These competition experiments were performed for three critical steps of our selection protocol: initial adhesion on the glass spatula inserted in biofilm microfermenters, biofilm formation and growth in liquid culture. We showed that, during biofilm formation in microtiter plates as well as growth in liquid culture, both Δ64–70 and Δ89–139 *fimH* mutants outcompeted the *ecpD/ecpR* mutant (Fig. 6C). Whereas there was no advantage of the V77A mutation in these conditions, competition for initial adhesion showed a very clear advantage for the V77A *fimH* mutant and, in a lesser extent, for both Δ64–70 and Δ89–139 mutants (Fig. 6C). These observed fitness advantages of the *fimH* mutants could account for the out-competition of other adhesin mutations in our evolution experiments.

Discussion

In this study, we showed that despite the vast arsenal of proteinaceous and macromolecular surface structures known to contribute to biofilm formation in *E. coli*, *in vitro* experimental evolution selecting for increased biofilm-forming capacities systematically led to the acquisition, at high frequency, of mutations in the type 1 fimbriae tip adhesin gene *fimH*, one of the first *E. coli* appendages implicated in biofilm formation on abiotic surfaces (Pratt and Kolter 1998). The restricted mutational landscape revealed by our experiments showed that type 1 fimbriae are the main *E. coli* adhesins contributing to biofilm formation. However, mutations located upstream of *flhD* were also identified at lower frequency in some biofilm positive evolved populations. Although these mutations do not impact bacterial motility, they increase the production of the flagellin FliC, and may therefore contribute to the observed increased biofilm capacity. In addition, we cannot exclude that mutations in other genes could have emerged and have been counter selected before the last selection cycles or been present below the 5% threshold detection of our *breseq* analyses. For instance, the G2 population showed an episodic increase in adhesion capacity at day 11 that was not detected afterward (Fig. 1B). This correlated with the presence of an in-frame deletion in *flu*, the gene coding for the self-recognition *E. coli* adhesin Antigen 43 (Diderichsen 1980, Henderson *et al.* 1997). We hypothesize that this mutation and other potential *fimH*-independent mutants could be outcompeted by emerging *fimH* mutations leading to a strong bias toward type 1 fimbriae mutations. In support of this hypothesis, we showed that, although positive selection for increased adhesion in a strain deleted for the *fimA-H* operon also identified mutations in genes encoding adhesins, the strongest biofilm-former amongst these mutants was outcompeted by *fimH* mutants for growth, initial adhesion and biofilm formation. In addition, we observed that mutation of the tip-pilus adhesin *ecpD* was always observed together with the insertion of an IS element that increases expression of the otherwise cryptic *ecp* pilus operon. This suggests that adhesin expression may first need to be unlocked before productive mutations could enhanced their adhesive properties and our selection could favor bacteria in which *fim* expression was locked ON, so more likely to evolve toward increased adhesion. We also identified in some populations that did not evolve enhanced biofilm capacity some mutations in *waaB*, a gene coding for an enzyme adding galactose to the LPS outer core (Qian *et al.* 2014). Whereas mutations in *waaB* mutations were shown to have no or little impact on *E. coli* biofilm formation (Wang *et al.* 2015, Smith *et al.* 2017), *waaB* mutants were also reported to display increased colanic acid production (Ren *et al.* 2016). Mucoidy could prevent the emergence of biofilm-positive clones by masking type 1 fimbriae in our positive selection experiments. However, we did not observe any difference between WT and *waaB* mutants in mucoidy nor biofilm formation and the cause and consequence of the selection of *waaB* mutations will therefore need to be explored further.

FimH is an allosterically regulated mannose-binding protein and FimH-dependent mannose-binding is considered as important for adhesion to mannosylated cell surface receptors *in vivo* as well as *in vitro* biofilm maturation possibly via the recognition of mannose-rich biofilm matrix component (Rodrigues and Elimelech 2009). We did not identify any mutations in the other components of type 1 fimbriae such as the major pilin FimA composing the shaft of the fimbriae or the minor pilins FimF and FimF. This indicates that these proteins do not contribute to adhesion *per se* in our experiments. In presence of shear forces, FimH mannose-binding capacity is enhanced by catch-bond mechanisms (Thomas *et al.* 2002, Le Trong *et al.* 2010), which could provide a selective advantage in the turbulent conditions used to select for increased-adhesion mutants. However, while most identified *fimH* mutants displayed an enhanced capacity for initial adhesion to abiotic surface they exhibited reduced mannose-binding capacity and *fimH* mutants with deletion encompassing part of FimH mannose-binding pocket were amongst the strongest biofilm-formers. This shows that FimH-dependent mannose-binding and catch bond do not significantly contribute to the increased biofilm capacities of the mutants identified in our study. FimH was early demonstrated to contribute to both specific and nonspecific adhesion (Pratt and Kolter 1998) and our results show that the lectin domain of FimH contributes to both types of adhesion, with residues of this lectin domain directly engaged in interaction with abiotic surfaces. This revealed a trade-off between attachment to mannose and abiotic surfaces that could be due to the fact that residues that are mutated or deleted in our selected *fimH* mutants are either at the vicinity of the mannose-binding pocket and could directly impact the specificity of the interaction through structural modification (Fig. S12, Supporting Information). Hydrophobic interactions also play a key role in bacterial adhesion (Rosenberg and Kjelleberg 1986) and mutations reducing hydrophobicity could enhance FimH interactions to abiotic surfaces. Such mutations were selected during our evolution experiments: substitution of hydrophobic by hydrophilic amino acid residue (V77A, L79P), producing the mutants with the high-

est initial adhesion on the spatula (Fig. 5), substitution of nonpolar by polar amino-acid (G82C and G87C) or deletion of stretches of hydrophobic residues (Δ89–139 correspond to the deletion of 38 hydrophobic residues out of 51). Hence, increasing FimH hydrophilicity could contribute to improve adhesion of FimH mutants on microfermenter glass spatula. In the case of the S16 L mutant in FimH signal peptide, alteration of signal peptide could increase FimH transport efficiency and type 1 fimbriae exposition (Ronald *et al.* 2008). However, we could not detect any differences in the quantity of surface exposed FimA in this FimH S16 L mutant. Alternatively, the selected mutations in *fimH* could improve surface contact and adhesion due to changes in the tertiary structure of FimH (Fig. S12, Supporting Information).

Unexpectedly, our study also revealed a negative correlation between the strength of initial adhesion displayed by the evolved *fimH* mutants and their capacity to form mature biofilms. This trade-off is particularly clear in the case of mutations at position V77A and L79P, leading to strong increase in initial adhesion with almost no positive impact on biofilm formation capacity. By contrast, the Δ64–70 and Δ89–139 mutations showed low gain in initial adhesion but led to strong increase in biofilm formation. These results illustrate the complexity of the mechanisms at play during biofilm formation. Strong initial adhesion could deeply impact bacterial metabolism and significantly delay biofilm maturation (Geng *et al.* 2014). Alternatively, but these hypotheses are not mutually exclusive, a too strong attachment to the surface might not be optimal to favor later cell-to-cell interaction and matrix production, but also the dynamism, including movement of cells, that might be necessary during biofilm maturation.

Type 1 fimbriae were shown to contribute to both pathogenic and commensal *E. coli* colonization of biotic surfaces. FimH-mediated adhesion enables commensal *E. coli* to adhere to buccal and intestinal epithelia as part of the normal bacterial flora but also enables pathogenic strains to colonize various mannosylated tissues (Ofek *et al.* 1977, Gbarah *et al.* 1991, Kukkonen *et al.* 1993). Although mutations in the FimH pilin domain are likely negatively selected because of its critical functional role, the mannose lectin domain displays a higher genetic plasticity, which could lead to functional diversity, conferring selective advantage by modifying and diversifying substrate-binding capacity (Sokurenko *et al.* 1994, 1997). Consistently, several phenotypic variants of the FimH lectin domain have been identified in clinical urinary or intestinal *E. coli* isolates (Bouckaert *et al.* 2006, Weissman *et al.* 2007, Stahlhut *et al.* 2009, Iebba *et al.* 2012, Dreux *et al.* 2013, Szunerits *et al.* 2016). These variants are often single point, non-synonymous amino acid substitutions found in the lectin domain but not affecting the mannose-binding pocket directly. While adaptive mutations identified in our study were also found in the lectin domain of the protein, only position 87 (G87C) has been previously described in pathoadaptive variants (Bouckaert *et al.* 2006, Weissman *et al.* 2007, Dreux *et al.* 2013, Szunerits *et al.* 2016) and described as a mutational hotspot with amino acid changes to arginine, alanine, serine and cysteine (Weissman *et al.* 2007). The G87S and G87C variants are the only mutants that display moderate increases in mannose binding, but still display catch-bond properties under flow conditions (Weissman *et al.* 2007).

While FimH variants previously identified in clinical isolates were mostly associated with tissue tropism (Weissman *et al.* 2007, Iebba *et al.* 2012, Dreux *et al.* 2013, Szunerits *et al.* 2016), our study shows that some mutations isolated in natural and clinical strains can also be selected for increased initial adhesion to abiotic surface and, as a consequence, biofilm formation capacity. This suggests that evolution of FimH in natural and/or clinical environments can also be driven by abiotic environmental selection pressures. The observation that FimH sequence distribution does not strictly reflect phylogroup clustering (Fig. S10, Supporting Information) further suggests that this gene is subjected to selection pressures specific to each strain environment and, consequently, the need for specific adaptations. Supporting this idea, the mutational landscape analysis of FimH sequences in *E. coli* revealed 20 hotspots constituting a potential source of functional diversity. These different hotspots have various phylogroup enrichment suggesting that these mutations are related to diverse ecological contexts and environments. For instance, the co-occurrence of S91N and N99S related to UPEC strains (Chen *et al.* 2009) was also identified in environmental strains, which is consistent with the growing evidence supporting the cycling of *E. coli* between the host gut and the environment (Jang *et al.* 2017). Although the frequency of this lifestyle switch and how long each *E. coli* clone resides in the gut and in the environment is not well known, *E. coli* is able to grow, evolve and adapt in both of these environments. Interestingly, conditions in the mammalian gut are relatively stable as compared with the one found in the outside host environment, and experimental evolution of *E. coli* in the gut did not select for mutation in *fimH* (Fabich *et al.* 2011, Barroso-Batista *et al.* 2014, Lourenço *et al.* 2016, Lescat *et al.* 2017, Ghalayini *et al.* 2018), in contrast with studies on plant-associated *E. coli* demonstrating an evolution toward increased capacity to form biofilm in this environment (Meric *et al.* 2013). Hence, rapidly changing conditions encountered by *E. coli* outside of the host could impose strong selection pressures on environmental *E. coli*, potentially leading to their rapid evolution (Bergholz *et al.* 2011, Luo *et al.* 2011). The fact that FimH mutations selected in our experiment correspond to naturally occurring mutations enriched in B2 and F phylogroups (mainly ExPEC strains), suggests that mutations favoring nonspecific *versus* specific adhesion could be selected even in strains known for their tissue tropism. One can therefore speculate about the nature of the main *in vivo* and/or environmental selection pressure driving the evolution sequenced *E. coli* clones. Unfortunately, the clear lack of environmental *E. coli* strains in the databases hinders genomic analysis. Indeed, among the 3543 FimH sequences included in our analysis 1256 were of unknown origin, 1192 were isolated from human, 690 from animals, 64 from food-related samples and only 341 from the environment. This therefore emphasizes the need for an increased number of sequenced and annotated genomes of environmental *E. coli* strains.

From an evolutionary perspective, the benefits of FimH mutations affecting mannose-binding capacity greatly depend on the environmental surfaces colonized by *E. coli*. FimH variants with increased mannose affinity under shear stress could be advantageous for bladder colonization, while decreased concentrations of free soluble mannose in the intestine could impose a very different selection pressure. Our findings show that *fimH* can undergo rapid microevolution, leading to increased nonspecific adhesion and biofilm formation independently of mannose-binding capacity. The plasticity of FimH adhesin should therefore be considered when designing anti-biofilm strategies on medical equipment as the use of pilicides inhibiting type 1 pili production may be preferred to mannose derivatives inhibiting FimH specific binding to mannose (Cegelski *et al.* 2008). This extended FimH mutational landscape in biofilms could reflect a natural strategy to diversify *E. coli* surface binding capacity, resist physical and chemical disruptions and potentially providing selective advantage for persistence during *E. coli* cycling between hosts and the environment.

Acknowledgments

We thank Rebecca Stevick, Médéric Diard and Olaya Rendueles for critical reading of the manuscript. We are grateful to Prof. Scott Hultgren for kindly providing the anti-FimA antiserum and to Dr Olaya Rendueles for the initial help with the analysis of the mutations.

Supplementary data

Supplementary data are available at *FEMSML* online.

Data availability

The data underlying this article are available in github at https://github.com/Sthiriet-rupert/FimH_Evolution and in GenBank Database at https://www.ncbi.nlm.nih.gov/genbank/ and can be accessed with the BioProject accession number PRJNA714528.

Author contributions

M.Y., C.B. and J.-M.G. designed the experiments. M.Y., L.M. and S.T.-R. performed the experiments. M.Y., C.B., S.T.-R., L.M. and J.-M.G. analyzed data. J.-M.G, C.B., M.Y. and S.T.-R. wrote the paper.

Funding

This work was supported by an Institut Pasteur grant, by the French government's Investissement d'Avenir Program, Laboratoire d'Excellence 'Integrative Biology of Emerging Infectious Diseases' (grant no. ANR-10-LABX-62-IBEID) and by the Fondation pour la Recherche Médicale (FRM grant no. DEQ20180339185). M.Y. was supported by Institut Pasteur Roux Cantarini fellowship. S.T.-R was supported by the French National Research Agency (ANR), project EvolTolAB (ANR-18-CE13-0010).

Conflict of interest statement. None declared.

References

Afgan E, Baker D, Batut B *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* 2018;**46**:W537–44.

Barroso-Batista J, Sousa A, Lourenço M *et al.* The first steps of adaptation of *Escherichia coli* to the gut are dominated by soft sweeps. *PLos Genet* 2014;**10**:e1004182.

Beloin C, Da Re S, Ghigo J-M. Colonization of abiotic surfaces. *EcoSal Plus* 2005;**1**.

Beloin C, Roux A, Ghigo JM. *Escherichia coli* biofilms. *Curr Top Microbiol Immunol* 2008;**322**:249–89.

Bergholz PW, Noar JD, Buckley DH. Environmental patterns are imposed on the population structure of *Escherichia coli* after fecal deposition. *Appl Environ Microbiol* 2011;**77**:211–9.

Bhoite S, Van Gerven N, Chapman MR *et al.* Curli biogenesis: bacterial amyloid assembly by the type VIII secretion pathway. *EcoSal Plus* 2019;**8**.

Bouckaert J, Mackenzie J, De Paz JL *et al.* The affinity of the FimH fimbrial adhesin is receptor-driven and quasi-independent of *Escherichia coli* pathotypes. *Mol Microbiol* 2006;**61**:1556–68.

Camacho C, Coulouris G, Avagyan V *et al.* BLAST+: architecture and applications. *BMC Bioinf* 2009;**10**:421.

Carr IM, Robinson JI, Dimitriou R *et al.* Inferring relative proportions of DNA variants from sequencing electropherograms. *Bioinformatics* 2009;**25**:3244–50.

Cegelski L, Marshall GR, Eldridge GR *et al.* The biology and future prospects of antivirulence therapies. *Nat Rev Microbiol* 2008;**6**:17–27.

Chagnot C, Zorgani MA, Astruc T *et al.* Proteinaceous determinants of surface colonization in bacteria: bacterial adhesion and biofilm formation from a protein secretion perspective. *Front Microbiol* 2013;**4**:303.

Chahales P, Thanassi DG. Structure, function, and assembly of adhesive organelles by uropathogenic bacteria. *Microbiol Spectr* 2015;**3**.

Chen SL, Hung CS, Pinkner JS *et al.* Positive selection identifies an *in vivo* role for FimH during urinary tract infection in addition to mannose binding. *Proc Natl Acad Sci USA* 2009;**106**:22439–44.

Deatherage DE, Barrick JE Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol Biol* 2014;**1151**:165–88.

Diderichsen B. flu, a metastable gene controlling surface properties of *Escherichia coli*. *J Bacteriol* 1980;**141**:858–67.

Dreux N, Denizot J, Martinez-Medina M *et al.* Point mutations in FimH adhesin of Crohn's disease-associated adherent-invasive *Escherichia coli* enhance intestinal inflammatory response. *PLoS Pathog* 2013;**9**:e1003141.

Fabich AJ, Leatham MP, Grissom JE *et al.* Genotype and phenotypes of an intestine-adapted *Escherichia coli* K-12 mutant selected by animal passage for superior colonization. *Infect Immun* 2011;**79**:2430–9.

Fronzes R, Remaut H, Waksman G. Architectures and biogenesis of non-flagellar protein appendages in Gram-negative bacteria. *EMBO J* 2008;**27**:2271–80.

Gally DL, Leathart J, Blomfield IC. Interaction of FimB and FimE with the fim switch that controls the phase variation of type 1 fimbriae in *Escherichia coli* K-12. *Mol Microbiol* 1996;**21**:725–38.

Gbarah A, Gahmberg CG, Ofek I *et al.* Identification of the leukocyte adhesion molecules CD11 and CD18 as receptors for type 1-fimbriated (mannose-specific) *Escherichia coli*. *Infect Immun* 1991;**59**:4524–30.

Geng J, Beloin C, Ghigo J-M *et al.* Bacteria hold their breath upon surface contact as shown in a strain of *Escherichia coli*, using dispersed surfaces and flow cytometry analysis. *PLoS One* 2014;**9**:e102049.

Ghalayini M, Launay A, Bridier-Nahmias A *et al.* Evolution of a dominant natural isolate of *Escherichia coli* in the human gut over the course of a year suggests a neutral evolution with reduced effective population size. *Appl Environ Microbiol* 2018;**84**:e02377–17.

Ghigo J-M. Natural conjugative plasmids induce bacterial biofilm development. *Nature* 2001;**412**:442–5.

Hasman H, Schembri MA, Klemm P. Antigen 43 and type 1 fimbriae determine colony morphology of *Escherichia coli* K-12. *J Bacteriol* 2000;**182**:1089–95.

Henderson IR, Meehan M, Owen P. Antigen 43, a phase-variable bipartite outer membrane protein, determines colony morphology and autoaggregation in *Escherichia coli* K-12. *FEMS Microbiol Lett* 1997;**149**:115–20.

Hung C-S, Bouckaert J, Hung D *et al.* Structural basis of tropism of *Escherichia coli* to the bladder during urinary tract infection. *Mol Microbiol* 2002;**44**:903–15.

Iebba V, Conte MP, Lepanto MS *et al.* Microevolution in fimH gene of mucosa-associated *Escherichia coli* strains isolated from pediatric patients with inflammatory bowel disease. *Infect Immun* 2012;**80**:1408–17.

Jang J, Hur H-G, Sadowsky MJ *et al.* Environmental *Escherichia coli*: ecology and public health implications—a review. *J Appl Microbiol* 2017;**123**:570–81.

Kelley LA, Mezulis S, Yates CM *et al.* The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 2015;**10**:845.

Klemm P. Two regulatory fim genes, fimB and fimE, control the phase variation of type 1 fimbriae in *Escherichia coli*. *EMBO J* 1986;**5**: 1389–93.

Kline KA, Fälker S, Dahlberg S *et al.* Bacterial adhesins in host–microbe interactions. *Cell Host Microbe* 2009;**5**:580–92.

Korea C-G, Badouraly R, Prevost M-C *et al.* *Escherichia coli* K-12 possesses multiple cryptic but functional chaperone-usher fimbriae with distinct surface specificities. *Environ Microbiol* 2010;**12**: 1957–77.

Korea C-G, Ghigo J-M, Beloin C. The sweet connection: solving the riddle of multiple sugar-binding fimbrial adhesins in *Escherichia coli*: multiple *E . coli* fimbriae form a versatile arsenal of sugar-binding lectins potentially involved in surface-colonisation and tissue tropism. *Bioessays* 2011;**33**:300–11.

Kukkonen M, Raunio T, Virkola R *et al.* Basement membrane carbohydrate as a target for bacterial adhesion: binding of type I fimbriae of *Salmonella enterica* and *Escherichia coli* to laminin. *Mol Microbiol* 1993;**7**:229–37.

Larsonneur F, Martin FA, Mallet A *et al.* Functional analysis of *Escherichia coli* Yad fimbriae reveals their potential role in environmental persistence. *Environ Microbiol* 2016;**18**:5228–48.

Le Trong I, Aprikian P, Kidd BA *et al.* Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like beta sheet twisting. *Cell* 2010;**141**:645–55.

Lehti TA, Bauchart P, Kukkonen M *et al.* Phylogenetic group-associated differences in regulation of the common colonization factor Mat fimbria in *Escherichia coli*. *Mol Microbiol* 2013;**87**:1200–22.

Lenski RE, Rose MR, Simpson SC *et al.* Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2000 generations. *Am Nat* 1991;**138**:1315–41.

Lescat M, Launay A, Ghalayini M *et al.* Using long-term experimental evolution to uncover the patterns and determinants of molecular evolution of an *Escherichia coli* natural isolate in the streptomycin-treated mouse gut. *Mol Ecol* 2017;**26**:1802–17.

Lourenço M, Ramiro RS, Güleresi D *et al.* A mutational hotspot and strong selection contribute to the order of mutations selected for during *Escherichia coli* adaptation to the gut. *PLos Genet* 2016;**12**:e1006420.

Luo C, Walk ST, Gordon DM *et al.* Genome sequencing of environmental *Escherichia coli* expands understanding of the ecology and speciation of the model bacterial species. *Proc Natl Acad Sci USA* 2011;**108**:7200–5.

Matthysse AG, Deora R, Mishra M *et al.* Polysaccharides cellulose, poly-beta-1,6-*n*-acetyl-D-glucosamine, and colanic acid are required for optimal binding of *Escherichia coli* O157:H7 strains to alfalfa sprouts and K-12 strains to plastic but not for binding to epithelial cells. *Appl Environ Microbiol* 2008;**74**:2384–90.

Méric G, Kemsley EK, Falush D *et al.* Phylogenetic distribution of traits associated with plant colonization in *Escherichia coli*. *Environ Microbiol* 2013;**15**:487–501.

Ofek I, Mirelman D, Sharon N. Adherence of *Escherichia coli* to human mucosal cells mediated by mannose receptors. *Nature* 1977;**265**:623–5.

Patel S, Mathivanan N, Goyal A. Bacterial adhesins, the pathogenic weapons to trick host defense arsenal. *Biomed Pharmacother* 2017;**93**:763–71.

Pratt LA, Kolter R. Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Mol Microbiol* 1998;**30**:285–93.

Qian J, Garrett TA, Raetz CRH. In vitro assembly of the outer core of the lipopolysaccharide from *Escherichia coli* K-12 and *Salmonella typhimurium*. *Biochemistry* 2014;**53**:1250–62.

Ren Ge, Wang Z, Li Ye *et al.* Effects of lipopolysaccharide core sugar deficiency on colanic acid biosynthesis in *Escherichia coli*. *J Bacteriol* 2016;**198**:1576–84.

Rodrigues DF, Elimelech M. Role of type 1 fimbriae and mannose in the development of *Escherichia coli* K12 biofilm: from initial cell adhesion to biofilm formation. *Biofouling* 2009;**25**:401–11.

Römling U, Galperin MY. Bacterial cellulose biosynthesis: diversity of operons, subunits, products, and functions. *Trends Microbiol* 2015;**23**:545–57.

Ronald LS, Yakovenko O, Yazvenko N *et al.* Adaptive mutations in the signal peptide of the type 1 fimbrial adhesin of uropathogenic *Escherichia coli*. *Proc Natl Acad Sci USA* 2008;**105**:10937–42.

Rosenberg M, Kjelleberg S. Hydrophobic interactions: role in bacterial adhesion. *Adv Microb Ecol* 1986;**9**:353–93.

Rossi E, Cimdins A, Lüthje P *et al.* "It's a gut feeling": *Escherichia coli* biofilm formation in the gastrointestinal tract environment. *Crit Rev Microbiol* 2018;**44**:1–30.

Roux A, Beloin C, Ghigo J-M. Combined inactivation and expression strategy to study gene function under physiological conditions: application to identification of new *Escherichia coli* adhesins. *J Bacteriol* 2005;**187**:1001–13.

Schroedinger LLC. PyMOL. (PyMOL molecular graphics system) 2010.

Smith D, Price J, Burby P *et al.* The production of curli amyloid fibers is deeply integrated into the biology of *Escherichia coli*. *Biomolecules* 2017;**7**:75.

Sokurenko EV, Chesnokova V, Doyle RJ *et al.* Diversity of the *Escherichia coli* type 1 fimbrial lectin. Differential binding to mannosides and uroepithelial cells. *J Biol Chem* 1997;**272**:17880–6.

Sokurenko EV, Courtney HS, Ohman DE *et al.* FimH family of type 1 fimbrial adhesins: functional heterogeneity due to minor sequence variations among fimH genes. *J Bacteriol* 1994;**176**: 748–55.

Stahlhut SG, Tchesnokova V, Struve C *et al.* Comparative structure-function analysis of mannose-specific FimH adhesins from *Klebsiella pneumoniae* and *Escherichia coli*. *J Bacteriol* 2009;**191**:6592–601.

Szunerits S, Zagorodko O, Cogez V *et al.* Differentiation of Crohn's disease-associated isolates from other pathogenic *Escherichia coli* by fimbrial adhesion under shear force. *Biology (Basel)* 2016;**5**:14.

Thomas WE, Trintchina E, Forero M *et al.* Bacterial adhesion to target cells enhanced by shear force. *Cell* 2002;**109**:913–23.

Touchon M, Perrin A, De Sousa JAM *et al.* Phylogenetic background and habitat drive the genetic diversification of *Escherichia coli*. *PLos Genet* 2020;**16**:e1008866.

Valle J, Da Re S, Henry N *et al.* Broad-spectrum biofilm inhibition by a secreted bacterial polysaccharide. *Proc Natl Acad Sci USA* 2006;**103**:12558–63.

Van Der Woude MW. Phase variation: how to create and coordinate population diversity. *Curr Opin Microbiol* 2011;**14**:205–11.

Van Houdt R, Michiels CW. Role of bacterial cell surface structures in *Escherichia coli* biofilm formation. *Res Microbiol* 2005;**156**:626–33.

Vo JL, Martínez Ortiz GC, Subedi P *et al.* Autotransporter adhesins in *Escherichia coli* pathogenesis. *Proteomics* 2017;**17**:1–10.

Wang X, Preston JF, Romeo T. The pgaABCD locus of *Escherichia coli* promotes the synthesis of a polysaccharide adhesin required for biofilm formation. *J Bacteriol* 2004;**186**:2724–34.

Wang Z, Wang J, Ren Ge *et al.* Influence of core oligosaccharide of lipopolysaccharide to outer membrane behavior of *Escherichia coli*. *Mar Drugs* 2015;**13**:3325–39.

Weissman SJ, Beskhlebnaya V, Chesnokova V *et al.* Differential stability and trade-off effects of pathoadaptive mutations in the *Escherichia coli* FimH adhesin. *Infect Immun* 2007;**75**: 3548–55.

Werneburg GT, Thanassi DG. Pili assembled by the chaperone/usher pathway in *Escherichia coli* and Salmonella. *EcoSal Plus* 2018;**8**.

Wurpel DJ, Beatson SA, Totsika M *et al.* Chaperone-usher fimbriae of *Escherichia coli*. *PLoS One* 2013;**8**:e52835.

Yamada KD, Tomii K, Katoh K. Application of the MAFFT sequence alignment program to large data-reexamination of the usefulness of chained guide trees. *Bioinformatics* 2016;**32**:3246–51.

Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 2007;**24**:1586–91.