Rescue of bacterial motility using two and three-species FliC chimeras

Jacob Scadden[1], Divyangi Pandit[1], Yoshiyuki Sowa[2,3], Pietro Ridone[1], Matthew AB Baker[1*]

[1]School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, Australia
[2]Department of Frontier Bioscience, Hosei University, Tokyo, Japan
[3]Research Center for Micro-Nano Technology, Hosei University, Tokyo, Japan

*Corresponding Author

Abstract

The bacterial flagellar filament acts as a propeller to drive most bacterial swimming. The filament is made of flagellin, known as FliC in *Escherichia coli, Salmonella* Typhimurium and *Pseudomonas aeruginosa*. FliC consists of four domains, the highly conserved core D0 and D1 domains and the hypervariable outer D2 and D3 domains. The size and structure of the outer domains varies, being completely absent in some bacterial species. Here we sought to identify outer domains from various species which were compatible such that they could form functional filaments to drive motility. We calculated a phylogeny of 211 representative flagellin amino acid sequences and generated two outer domain deleted variants and six chimeric *fliC* mutants using domains from *E. coli*, *Salmonella* Typhimurium, *P. aeruginosa*, *Collimonas fungivorans*, *Helicobacter mustelae* and *Mesorhizobium* sp. ORS3359. Four of the chimeric *fliC* mutants rescued motility in a *fliC* disrupted strain, all of which contained the *Salmonella* Typhimurium D2 domain. Overall, we demonstrate the interchangeability of the outer domains, in particular that domains from different species can be interchanged to form functional filaments that propel bacterial swimming.

Introduction

Motility provides organisms the ability to exploit new niches, to evade predation and increased opportunities for the exchange of genetic material (1). The flagellar filament is a major component of the bacterial flagellar motor (BFM) taking the form of a large helical protein complex which is formed from approximately 30,000 flagellin monomers, known as FliC in *Escherichia coli* (2,3). FliC interacts with both the flagellar hook junction proteins (FlgK and FlgL) and filament cap proteins (FliD) that form part of the extracellular components of the BFM (4–6). Flagellin monomers interact to form a 1-start helix consisting of 11 subunits per two turns which comprises the filament (7).

FliC is comprised of four domains, D0, D1, D2 and D3 (8,9). Domains D0 and D1 are found on the N- and C-termini of the protein and are highly conserved across multiple bacterial phyla, such as *Proteobacteria*, *Bacillota* and *Actinobacteria* (10). These "core" domains are formed of α-helices and interact with each other as part of the supercoiled filament (7–9). Within the D0 domain is a 22-residue segment which is the recognition signal for the flagellar export chaperone, FliS (11). Domains D2 and D3, also termed "outer domains", are variable in both sequence and length even at a species level (3). These outer domains consist primarily of β-sheets and were shown in *Salmonella* Typhimurium to have no interactions between each other (8,12). In some cases, flagellin monomers from other bacterial species such as *Bacillus cereus*, *Agrobacterium tumefaciens* and *Treponema pallidum* do not contain a D2 or D3 domain and still produce a functional flagellar filament capable of motility (10,13–17). It has been shown that at the species level of *Escherichia* and *Salmonella* there is a large level of diversity in the outer domains, suggesting that the outer domains are acquired via the process of

evolution (18,19). This implies that the outer domains of flagellin can be incorporated into host flagellins and could provide competitive advantages to increase motility efficiency in new environments or evade immune system recognition.

In *Pseudomonas aeruginosa* PAO1 the outer domains of flagellin increase the motility of bacteria in more viscous environments due to inter-domain interactivity (20). Likewise, in *E. coli* O157:H7, *E. coli* O127:H6, *Achromobacter* and *Sinorhizobium meliloti* the outer domains can dimerise or tetramerise to form a screw-like surface which can increase motility through stabilisation an intermediate waveform and prolonging tumbling (21). Furthermore, interactions in the outer domains of *E. coli* can increase tumbling time (21). Inter-species flagellin chimeras between Salmonella strains where able to form functional filaments (22). Of 13 FliC variants, six were motile and it was shown that alternating the outer domains could impact motility (22).

Here we tested the ability of chimeric FliC to support motility where we replace the outer domains of *E. coli* with outer domains from multiple, differing bacterial phyla. We show that the outer domains of *E. coli* FliC are not required to drive swimming motility, and that chimeric FliC can form functional filaments that spread on swim plates and swim in solution at equivalent speeds to wild type FliC.

Materials and Methods

Bacterial strains, plasmids and growth conditions

Table S1 describes the strains and plasmids created and used in this study. The *E. coli* strains SYC29 was constructed from the chemotactic wild-type strains RP437, using the λ red recombination system, as previously described (23,24). *E. coli* strains were grown in Lysogeny Broth (LB) (10 g/L NaCl, 10 g/L Bacto-tryptone, 5 g/L Yeast Extract). For 2% (w/v) or 0.3% (w/v) LB agar plates 20 g/L or 3 g/L agar was added, respectively. *E. coli* strains were grown in tryptone Broth (TB) to aid in induction of motility (10 g/L Bacto-tryptone, 5 g/L NaCl). Antibiotic selection used ampicillin (AMP) (50 µg/ml), chloramphenicol (CAM) (25 µg/ml) and/or tetracycline (TET) (10 µg/ml). Induction for gene expression from plasmids required L-arabinose (0.02% (w/v)) and/or Isopropyl ß-D-1-thiogalactopyranoside (IPTG) (0.1 mM).

Phylogenetic analysis of FliC

To investigate the structural diversity of the outer domains of FliC and their presence/absence across the tree of life, eight flagellin amino acid sequences (Table S2) were used as reference sequences for a BLASTP search in Uniprot (25) (E-Threshold - 10; Matrix – BLOSUM62; Filter – None; Gapped – Yes; Hits – 1000; HSPs per hit – All). All hits for each BLASTP search were selected and used in ID mapping from database UniProtKB AC/ID to UniRef90 (26). The clusters were sorted by member size and each list was manually curated to select 226 representative FliC sequences for further phylogenetic analysis, inclusive of 15 FliD amino acid sequences as an outgroup. The 226 sequences were downloaded in fasta format and uploaded to CIPRES Science Gateway (27). Alignment of the representative sequences was performed using MUSCLE (3.7) (28) and viewed using AliView (29). The alignment was used to generate a phylogenetic tree using IQTree (2.3.2) (30) using CIPRES (for parameters see Table S3). The phylogenetic tree was visualised and annotated in FigTree (v1.4.4).

Design and structural prediction of chimeric FliC

To study the effect of outer domain manipulation on bacterial motility, a rational design approach for each chimeric FliC candidate was taken. It was decided to maintain the D0 and D1 domains of *E. coli* FliC (UniProt - P04949) in order to preserve the export recognition sequence found in the D0 domain (11). Therefore, to understand if these outer domains are functionally required for flagellum

86 formation, the first *fliC* variants designed were the *fliC_ΔD3* and *fliC_ΔD2/3*, with deletions of residues
87 195-300 and 174-405, respectively. For the two-species and three-species FliC chimeras, five flagellin
88 amino acid sequences were used from across the bacterial tree of life, *Helicobacter mustelae*
89 (P50612), *Mesorhizobium* sp. ORS3359 (A0A090GIX0), *Pseudomonas aeruginosa* (P21184), *Collimonas*
90 *fungivorans* strain Ter331 (G0AIL6) and *Salmonella* Typhimurium strain LT2 (P06179). The D2 and/or
91 D3 domains of these flagellins were determined manually and used to replace the D2 and D3 domains
92 of *E. coli* FliC. In the case of *Salmonella* Typhimurium additional chimeras were made using the D2
93 domain of *Salmonella* Typhimurium and the D2 domains (in place of the D3 domains) of
94 *Mesorhizobium* sp. ORS3359, *P. aeruginosa*, *C. fungivorans* strain Ter331 were used. Structural
95 prediction for all domain deleted and chimeric FliC sequences was performed using AlphaFold2
96 (v1.5.5) using ColabFold (31).

97 **Cloning of FliC chimeras into expression vector**

98 Reverse translation of each FliC chimera from amino acid to nucleotide codon optimised sequence for
99 *E. coli* was performed in Benchling (Benchling Inc). Gene fragments for all chimeras and outer domain
100 deleted FliC were order using TWIST (Decode Bioscience) with EcoRI and NotI restriction sites
101 incorporated in the 5' and 3' regions, respectively. Digestion of gene fragments and pET21(+) (Merck)
102 expression vector was performed using EcoRI-HF and NotI-HF (NEB), followed by Antarctic
103 phosphatase (NEB) treatment of pET21(+), as per the manufacturer's instructions. Ligation was carried
104 out using T4 ligase (NEB) as per manufacturer's instructions and incubated at room temperature
105 overnight. After heat inactivation, the ligation mix was transformed into to chemically competent *E.*
106 *coli* 10-β (NEB) as per manufacturer's instructions. Transformations were incubated at 37°C for 1 hr
107 and plates on LB plates with AMP. Colonies were screened using pET21_F and pET21_R primers and
108 Q5 Polymerase (NEB) (See Table S4 for primer details and PCR conditions). Positive colonies were
109 grown overnight in LB media with AMP and plasmids were purified using a Qiagen Miniprep Kit,
110 following manufacturer's protocol. Purified plasmid was sent for Sanger sequencing performed by the
111 Ramaciotti Centre for Genomics (UNSW, Sydney). The confirmed plasmids were then transformed into
112 chemically competent *E. coli* SYC29 pDB108 and screened on LB plates containing AMP, CAM and TET.
113 Positive colonies were screened using the same PCR method as previously described and grown
114 overnight in LB media containing AMP, CAM and TET.

115 **Assessing the motility of chimeric FliC expressing *E. coli***

116 Swim plate (0.3% agar) motility assays were performed with required antibiotics and inducers as
117 previously described (32). Plates were inoculated with 1 µl induced overnight culture and incubated
118 at 30°C for 16 hrs. Plates were imaged using a Bio-rad Gel Doc XR with Gel documentation system with
119 analysis of Image Lab 6.0.1 (No filter, White Epi Illumination). Swim ring diameters were measured
120 using ImageJ (FIJI) (33). Data were analysed and displayed using GraphPad 2.0 (version 9.5.1).

121 **Free swimming assay for motile chimeric FliC constructs**

122 Constructs of interest were grown overnight in LB media containing appropriate antibiotics and
123 inducers at 30°C. A 1% inoculum was grown in TB media until OD_{600} 0.5. A tunnel slide was formed
124 using tape and a glass cover slip and 30 µl of sample was flown through, as previously described (34).
125 Cells were visualised using a phase contract microscope (Leica) and imaged using a Nikon camera. For
126 each sample 10 videos of 400 frames (20 frames/second) were recorded. Recordings were analysed
127 in a custom programme in LabVIEW (National Instruments) (32) to quantify the swim speeds of each
128 motile construct. GraphPad 2.0 (version 9.5.1) was used to analyse and display data.

129 **Flagella staining of FliC chimeric constructs**

130 Staining of flagellar filaments was performed using Remel ™ Flagella stain (Thermofisher) as per the
131 manufacturers protocol with minor changes. Constructs of interest were inoculated on 0.3% LB agar
132 plates, with appropriate antimicrobials and inducers, and incubated for 16 hrs at 30°C. To the edge of
133 each swim ring 30 µl of LB broth was added and incubated for two mins at room temperature after
134 which 10 µl was spotted on a glass slide. After this step the protocol was followed exactly. Samples
135 were visualised using a Leica DM750 microscope at 100x magnification and images captured using a
136 MEE-500 camera (Yegren, China) and the camera app on Windows 10 (v 10.0.1.19045). Filament
137 lengths were measured using ImageJ (FIJI) (33) using the freehand line tool. Data were analysed and
138 displayed using GraphPad 2.0 (version 9.5.1).

139 **Flagellin expression, precipitation and SDS-PAGE**

140 Overnight cultures were grown in LB media with appropriate antibiotics at 30°C. A 25 mL fresh culture
141 was grown to OD_{600} 0.4 in LB media at 30°C. IPTG and arabinose was added and incubated for a further
142 2 hrs. Cultures were heated at 65°C for 15 mins then centrifuged at 4000 x g at 4°C for 30 mins, the
143 supernatant was kept and filtered using a 0.22 µm filter (Merck). Ice-cold 100% trichloroacetic acid
144 was added (10% v/v) and incubated at 4°C overnight. Concentrated and precipitated proteins were
145 pelleted by centrifugation (4000 x g, 30 minutes, 4°C), washed with 20% volume ice-cold acetone for
146 2 hrs at -20°C, recentrifuged, supernatant removed then pellet dried until acetone was completely
147 evaporated. Pellet resuspended in 50 mM sodium acetate (pH 5.5) until fully dissolved. Quantification
148 was performed using a Bradfords reagent (Bio-Rad) as per manufacturer's instructions. 10 µg protein
149 was used per well in a 4-12% Bis-Tris gel (NuPage, ThermoFisher Scientific), after reducing each sample
150 as per manufacturer's instructions. The gel was run at 200 V for 35 mins and stained using SimplyBlue
151 SafeStain (ThermoFisher Scientific). Gels were imaged using a Bio-Rad Gel Doc XR with Gel
152 documentation system.

153 **Results**

154 **Flagellins cluster according to primarily phyla**

155 To study the structural diversity of flagellin outer domains, 211 representative flagellin amino acid
156 sequences from across the bacterial tree of life were selected, aligned and used in the generation of
157 a phylogenetic tree (Figure 1). Of the representative flagellin sequences selected 53% (112/211)
158 contained the outer D2 and D3 domains. The majority of the flagellins clustered according to the phyla
159 or class to which the bacteria encoding the respective flagellin gene belonged. In *Spirochaetota*
160 (12/12) and *Actinomycelotota* (10/10), all flagellins did not contain the outer domains (Figure 1, Red
161 branches). In *Bacillota,* all except one of the flagellins did not contain outer domains (53/54). The
162 classes *Epsilonproteobacteria* (20/20) and *Beta/Gammaproteobacteria* (66/67) all clustered within
163 their respective class and all but one flagellin contained the outer domains (Figure 1, Blue branches).
164 *Deltaproteobacteria* and the *Fibrobacteres-Chlorobi-Bacteroidetes* (FCB) superphylum had flagellin
165 sequences with and without outer domains and the clading of these was split on the presence or
166 absence of the outer domains. *Alphaproteobacteria* flagellins, in contrast, were the only group where
167 sequences with and without the outer domains where mixed in the one clade (10 with and 12 without,
168 respectively). In *Deltaproteobacteria* and the FCB superphylum, 64% (9/14) and 42% (5/12),
169 respectively, contained outer domains.

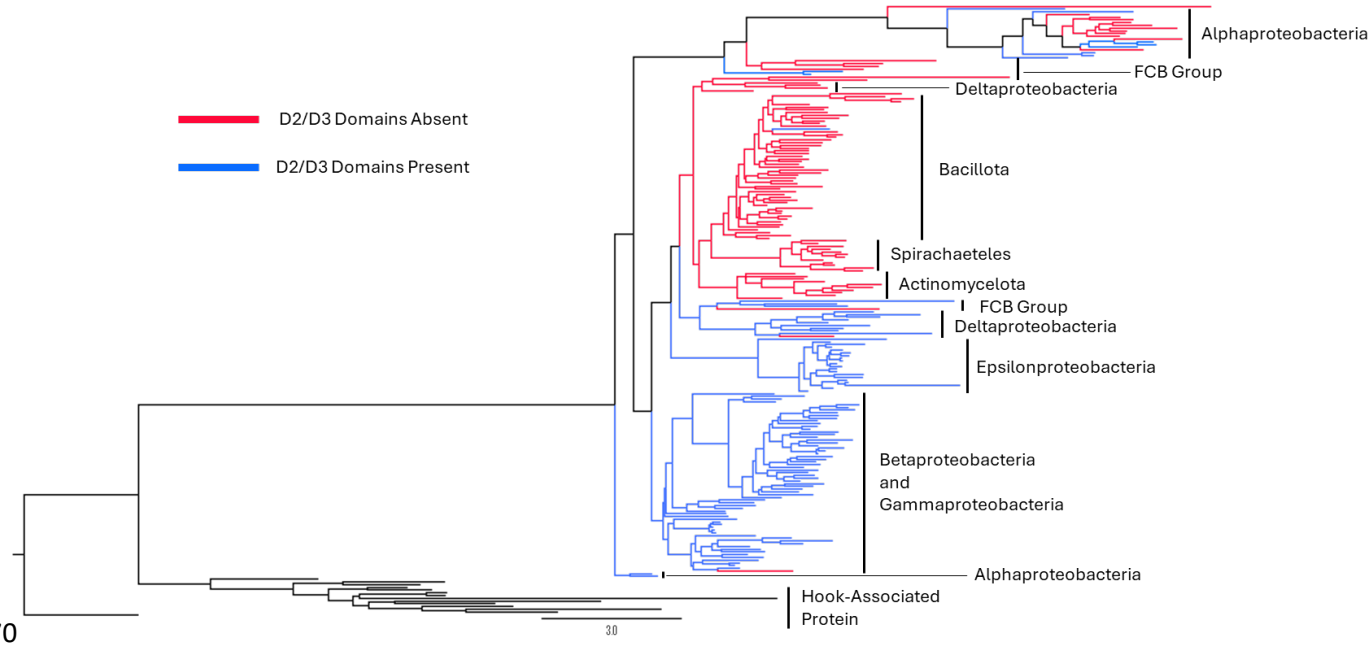


Figure 1. Phylogeny of 211 representative flagellin sequences, with 14 flagellar hook associated proteins as an outgroup. Branches are coloured based on the presence (Blue) or absence (red) of the D2 and D3 domains.

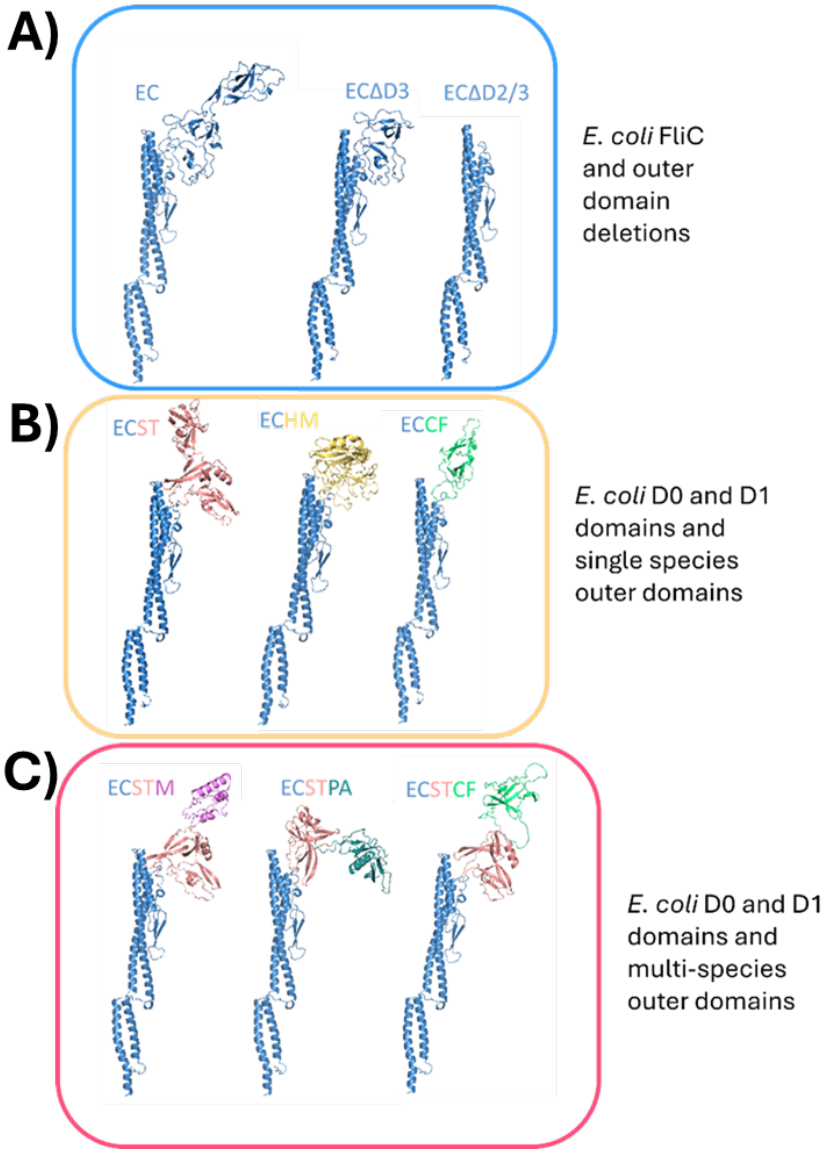
Analysis of the isolation sources for the bacteria from which the flagellin sequences were identified showed that there was a higher percentage of flagellins containing the D2 and D3 domains which were isolated from animals (40% to 24%) (Figure S1). However, flagellins lacking the outer domains were found predominately in aquatic (30%) and soil (24%) environments compared with flagellins that contained these domains (26% and 10%, respectively). Within each phyla-based cluster of the flagellins there was minimal association with a particular environment, with the exception of *Epsilonproteobactera*, 88% of which were isolated from animal sources and *Alphaproteobacteria*, which were all either from aquatic or plant associated environments.

Outer domain deleted and Chimeric flagellin can rescue motility in *E. coli*

First, to investigate the potential redundancy of the outer domains for bacterial motility we deleted the outer domains from *E. coli fliC.* Structural prediction of these outer domain deleted constructs showed that sequential removal of these domains, to generate constructs ECΔD3 and ECΔD2/3 had no effect on the structure of the core domains (Figure 2A). We then designed chimeric *fliC* constructs using the D0 and D1 domains of *E. coli* and the outer domains of *Salmonella* Typhimurium, *H. mustelae* and *C. fungivorans*. The D0 and D1 domains were maintained as they are in *E. coli* to maintain the export recognition site found between residues 26-47, given our use of *E. coli* proteins for all other elements of the motor, including flagellin export (11). *Salmonella* Typhimurium (*Gammaproteobacteria*) (ECST) and *C. fungivorans* (*Betaproteobacteria*) (ECCF) were selected as members of the *Pseudomonadota* phylum, which clustered in the same clade as *E. coli* FliC (Figure 1). The outer domains of *H. mustelae* (ECHM), belonging to *Epsilonproteobacteria*, were selected due to being in a cluster where all members contained the outer domains (Figure 1). The structures of these chimeric FliC variants were predicted and these structural predictions showed that the tertiary structure of the core domains was conserved (Figure 2B).

In addition to direct replacement of the outer domain replacement with a domain from a single species, we decided to engineer three-species FliC chimeras. This was to assess the inter-compatibility

199 of the outermost D3 domain. The D2 domain of *Salmonella* Typhimurium was used for all D2 domains
200 in these three-species FliC constructs (Figure S2). The D3 domain was replaced with the outer domain
201 from *C. fungivorans* strain Ter331 (ECSTCF), *P. aeruginosa* (ECSTPA), *Mesorhizobium* sp. ORS 3359
202 (ECSTM) (Figure 2C). The D3 domain from *C. fungivorans* strain Ter331 was chosen due to its FliC
203 phylogenetic clading with the FliC from *E. coli*. Alternately, the D3 domain from *P. aeruginosa* and
204 *Mesorhizobium* sp. ORS 3359 were chosen as previous work showed that the outer domains of *P.*
205 *aeruginosa* PA01 and *Sinorhizobium meliloti* formed ridged filaments formed by interaction between
206 the flagellin outer domains (Figure 2C) (20,21). Overall, chimeric and outer domain deleted FliC ranged
207 in size from 266 to 535 residues in length, with the outer D2 and D3 domains ranging from 107 to 269
208 amino acids in length (Table S4).



209

210 **Figure 2.** AlphaFold2 predictions of chimeric and outer domain deleted FliC monomers designed for
211 this study. A) Predicted structures of wild type, D3 deleted and D2/3 deleted flagellin. B) Predicted
212 structure of two-species flagellin chimeras. C) Predicted structures of three-species flagellin chimeras.
213 *E. coli* K-12 (Blue), *H. mustelae* (Yellow), *C. fungivorans* (Light Green), *Salmonella* Typhimurium (Pink),
214 *Mesorhizobium* sp. OR3359 (Purple), *P. aeruginosa* (Dark Green).

Next we investigated whether chimeric FliC could form functional filaments and therefore rescue motility in filament-disrupted strains (*fliC::tetRA*). Plasmids expressing FliC mutants or flagellar stators were co-transformed into an *E. coli ΔmotAB fliC::tetRA* strain. Motility of outer domain deleted and chimeric FliC expressing constructs was confirmed using swim plate motility assays. Both outer domain deletions produced swim rings (Figure 3A). The average swim ring diameter when both outer domains (ECΔD2/3) was larger, 14.5 ± 2.5 mm, than that with only the D3 domain deleted, 11.9 ± 2.5 mm, however this difference was not statistically significant (Figure 3B).

Likewise, motility of the two-species outer domain FliC mutants, ECST, ECHM and ECCF, was confirmed using a swim plate motility assay. Only ECST was motile, with a swim ring diameter of 9.9 ± 2.1 mm (Figure 3A and B). ECCF and ECHM produced no swim rings (Figure 3A and B).

All three of the three-species chimeras were observed to be motile. Both ECSTCF (41.1 ± 5.3 mm) and ECSTPA (44.4 ± 12.0 mm) expressing constructs produced a similar swim ring diameter when compared to wild type FliC expressed from pET21 (39.1 ± 2.5 mm) (Figure 3B). The FliC chimera that contained the D3 domain from *Mesorhizobium* sp. OR3359 consistently produced smaller swim ring diameters (10.6 ± 0.3 mm) compared to the other three-species FliC chimeras. S3

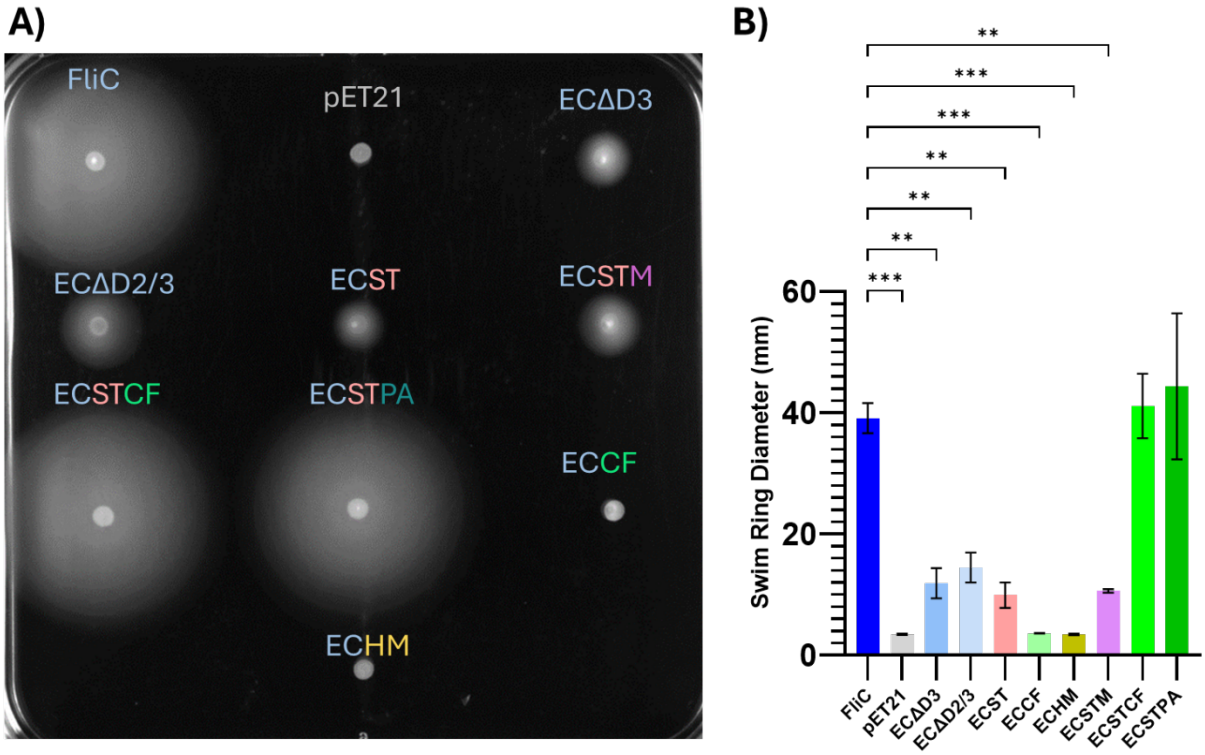
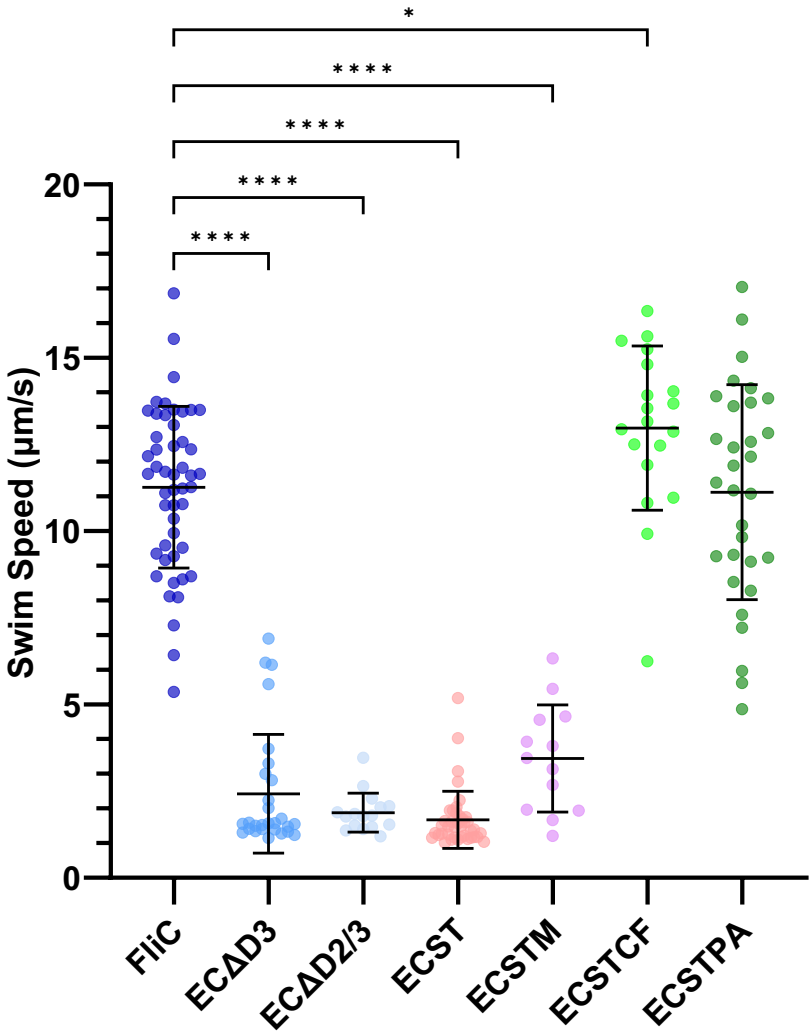


Figure 3. A) Representative 0.3% swim plate (120 mm) of all FliC chimeras. *E. coli ΔmotAB fliC::tetRA* (SYC29) pDB108 pET21_*fliC* and *E. coli* SYC29 pDB108 pET21 were used as positive and negative controls, respectively. Plates were incubated at 30°C for 16 hrs. B) Quantification of the mean swim ring diameter of each chimeric and outer domain deleted FliC recorded from three independent 0.3% swim plates incubated at 30°C for 16 hrs. Error bars indicate standard error of the mean. A two-way ANOVA comparing each column to the FliC value was used to generate the p values (** = ≤ 0.01, *** = ≤ 0.001).

FliC chimeras that produced the largest swim rings had the fastest swim speeds

239 Free swimming assays were used to measure the swimming speeds for each of the motile FliC
240 constructs that were generated as part of this work. Constructs expressing FliC (11.3 ± 2.3 µm/s),
241 ECSTCF (13.0 ± 2.7 µm/s) and ECSTPA (11.1 ± 3.1 µm/s) all had swim speeds significantly faster than
242 the other motile variants (p value < 0.0001) (Figure 4). The ECSTCF expressing construct had a
243 significantly faster free-swimming speed compared to the FliC expressing construct (p value < 0.05).
244 The ECΔD3 (2.4 ± 1.7 µm/s), ECΔD2/3 (1.9 ± 0.6 µm/s), ECST (1.7 ± 0.8 µm/s) and ECSTM (3.44 ± 1.5
245 µm/s) constructs had lower swim speeds.



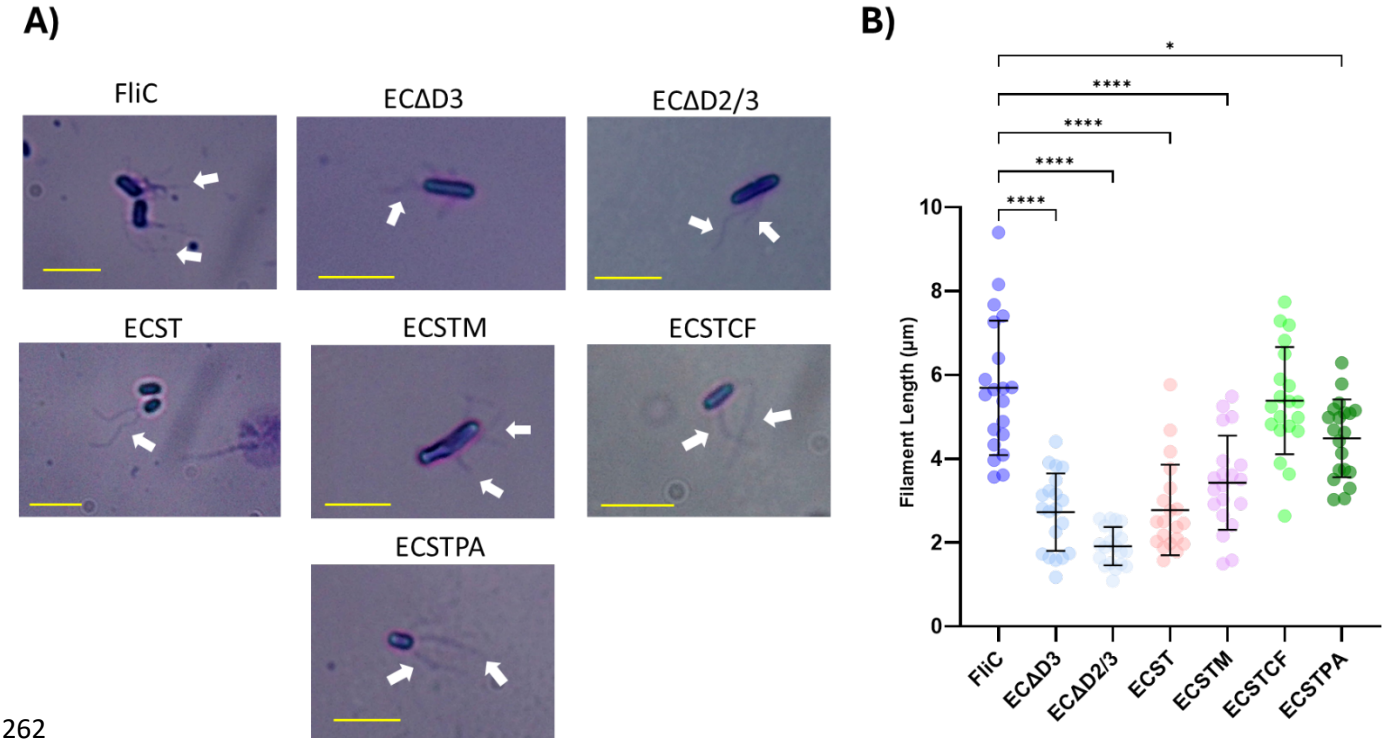
246

247 **Figure 4.** Free-swimming speeds for motile chimeric and truncated FliC in µm/s. Error bars indicate
248 standard deviation. A two-way ANOVA comparing each column to FliC was used to generate the p
249 values. FliC, STCF and STPA constructs are significantly different to all other FliC variants, with the
250 swim speed of STCF being significantly greater than FliC (p value ≤ 0.05) (* = ≤ 0.05, **** = ≤ 0.0001).

251 **Crystal violet staining of filaments shows formation of chimeric filaments**

252 Staining of flagella confirmed export of flagellin and the formation of filaments. Filaments were
253 observed in all motile constructs (Figure 5). This provided confirmation that all motile constructs were
254 able to export chimeric flagellin. For non-motile constructs ECHM and ECCF, there were no observable
255 filaments. An SDS-PAGE analysis showed that there was no flagellin export into the supernatant which
256 explained the lack of motility observed (Figure S3). Filament length was determined for all motile
257 constructs with FliC, ECSTCF and ECSTPA producing the longest filaments at 5.7 ± 1.6 µm, 5.4 ± 1.3 µm

258 and 4.5 ± 0.9 µm, respectively (Figure 6). The outer domain deleted contructs ECΔD3 and ECΔD2/3
259 produced significantly shorter filaments at 2.7 ± 0.9 µm and 1.9 ± 0.5 µm, respectively. Shorter
260 filaments were also formed in the two-species chimera ECST (2.8 ± 1.1 µm) and in the three-species
261 chimera ECSTM (3.4 ± 1.1 µm) (Figure 6).



262

263 **Figure 5.** A) Confirmation of filament export and assembley. Reprentative images of motile constructs
264 after flagella staining. Scale bar (Yellow) represents 5 µm. White arrows indicate presence of flagella.
265 B) Lengths of FliC chimeraic filaments. Error bars represent standard deviation. A two-way ANOVA
266 comparing each column to FliC was used to generate the p values (**** = ≤0.0001, * = ≤ 0.05).

267 Discussion

268 In this study, we aimed to investigate whether deletion or replacement of the outer domains of *E. coli*
269 FliC through could rescue motility in filament-deleted constructs. We observed that flagellin outer
270 domains were not required for motility in *E. coli* and that two and three-species FliC outer domain
271 chimeras are able to form functional filaments.

272 FliC outer domains can affect bacterial motility

273 Although originally thought that the outer domains have no interaction with the core domains or each
274 other (8,12), recent studies have shown that in some cases there are inter-domain interactions which
275 help to provide filament stability (20). In Kreutzberger *et al.*, this phenomenon was identified in
276 *Achromobacter* filaments where the outer domains exist in four conformations and can form a
277 tetrameric sheath which surrounds the filament (21). The authors also observed outer domain
278 interactions in *Sinorhizobium meliloti* forming a dimeric outer domain interaction in two
279 conformations (21). Similarly, *P. aeruginosa* PAO1 flagellin has been shown to form a ridged filament
280 due to interactions of its outer domains (20). This illustrates that along with the well-studied role of
281 immune activation of the flagellin outer domains there is also a role that these domains have on
282 bacterial swim speed in different environments (35). In *Salmonella* Typhimurium and *E. coli* it has been
283 previously shown that partial outer domain deleted FliC mutants are motile (36,37), however, full D2/3

deletions were non-motile, whereas our results show that full deletion of D2/D3 in *E. coli* is still motile. We also observed the highest swim speeds and swim ring diameters in constructs with the largest outer domains (ECSTCF and ECSTPA). Future work will aim to investigate any outer domain interactions are present in these mutant filaments.

Bacterial swim speed can vary with *E. coli* swim speeds ranging from 25-66 µm/s, *Thiovulum majus* at approximately 600 µm/s and *Bdellovibrio bacteriovorus* at 100 µm/s (38–40). The former two bacterial species containing flagellin outer domains whereas the latter does not. We observed lower swim speeds and shorter filaments for ECST, ECSTM, ECΔD3 and ECΔD2/3 constructs. This could be due to weaker filaments that are more prone to shearing, a difference in the filament export rate or the number of filaments present.

Outer domain chimeras require *Salmonella* Typhimurium D2 domain to form filaments

We observed that six out of eight FliC constructs rescued motility, with the remainder being two-species chimeras ECHM and ECCF that were non-motile. In work by Nedeljkocić *et al*., 13 flagellin chimeras, with combinations of all domains, were generated using the FliC from *Salmonella* Typhimurium and *P. aeruginosa* PAO1 and only one chimera produced a functional filament (20). In our constructs containing outer domains, motility was only rescued in chimeras that contained the *Salmonella* Typhimurium D2 domain. This may indicate that there are key structural elements or residues found in this domain that allow of export and formation of a functional filament. The two non-motile constructs also showed no export of flagellin, indicating that incompatible outer domains can inhibit export.

Conclusion

Engineering of the flagellin outer domains can not only increase our understanding of filament formation but could also provide biotechnological benefits. Flagellin is already used to display antigens from other organisms and stimulate the immune system and has been engineered to contain fluorescent and enzymatic domains (35,41,42). Here we sought to develop chimeras for the outer domains and we have shown that three-species chimeras can form functional filaments. Future work will extend the investigation on how hypervariable FliC domains impact the adaptation of motility to environmental conditions and establishing the FliC scaffold as a synthetic biology platform for biotechnological applications.

Author Information

1. School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, Australia
   Jacob Scadden, Divyagi Pandit, Pietro Ridone, Matthew AB Baker
2. Department of Frontier Bioscience, Hosei University, Tokyo, Japan
   Yoshiyuki Sowa
3. Research Center for Micro-Nano Technology, Hosei University, Tokyo, Japan
   Yoshiyuki Sowa

Corresponding Author

Correspondence to Matthew AB Baker
matthew.baker@unsw.edu.au

Funding

MABB is supported by US Navy Office of Naval Research, Research grant no. N62909-22-1-2051 and HFSP Project Grant RGY0072/21 and Australian Research Council Discovery Project DP24. YS is supported by JSPS KAKENHI grant numbers JP23H04429 and JP23K05701.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Miyata M, Robinson RC, Uyeda TQ, Fukumori Y, Fukushima S ichi, Haruta S, et al. Tree of motility–A proposed history of motility systems in the tree of life. Genes to Cells. 2020;25(1):6–21.

2. Sowa Y, Berry RM. Bacterial flagellar motor. Quarterly reviews of biophysics. 2008;41(2):103–32.

3. Nedeljković M, Sastre DE, Sundberg EJ. Bacterial Flagellar Filament: A Supramolecular Multifunctional Nanostructure. Int J Mol Sci. 2021 Jul 14;22(14).

4. Bulieris PV, Shaikh NH, Freddolino PL, Samatey FA. Structure of FlgK reveals the divergence of the bacterial Hook-Filament Junction of Campylobacter. Scientific reports. 2017;7(1):15743.

5. Hong HJ, Kim TH, Song WS, Ko HJ, Lee GS, Kang SG, et al. Crystal structure of FlgL and its implications for flagellar assembly. Scientific Reports. 2018;8(1):14307.

6. Postel S, Deredge D, Bonsor DA, Yu X, Diederichs K, Helmsing S, et al. Bacterial flagellar capping proteins adopt diverse oligomeric states. Elife. 2016;5:e18857.

7. O'Brien E, Bennett PM. Structure of straight flagella from a mutant Salmonella. Journal of molecular biology. 1972;70(1):133–52.

8. Samatey FA, Imada K, Nagashima S, Vonderviszt F, Kumasaka T, Yamamoto M, et al. Structure of the bacterial flagellar protofilament and implications for a switch for supercoiling. Nature. 2001;410(6826):331–7.

9. Yonekura K, Maki-Yonekura S, Namba K. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. Nature. 2003;424(6949):643–50.

10. Beatson SA, Minamino T, Pallen MJ. Variation in bacterial flagellins: from sequence to structure. Trends Microbiol. 2006 Apr;14(4):151–5.

11. Végh BM, Gál P, Dobó J, Závodszky P, Vonderviszt F. Localization of the flagellum-specific secretion signal in Salmonella flagellin. Biochemical and biophysical research communications. 2006;345(1):93–8.

12. Yamaguchi T, Toma S, Terahara N, Miyata T, Ashihara M, Minamino T, et al. Structural and Functional Comparison of Salmonella Flagellar Filaments Composed of FljB and FliC. Biomolecules. 2020 Feb 6;10(2).

13. Il Kim M, Lee C, Park J, Jeon BY, Hong M. Crystal structure of Bacillus cereus flagellin and structure-guided fusion-protein designs. Sci Rep. 2018 Apr 11;8(1):5814.

361 14. Mohari B, Thompson MA, Trinidad JC, Setayeshgar S, Fuqua C. Multiple flagellin proteins have
362 distinct and synergistic roles in Agrobacterium tumefaciens motility. Journal of Bacteriology.
363 2018;200(23):10–1128.

364 15. Wang F, Burrage AM, Postel S, Clark RE, Orlova A, Sundberg EJ, et al. A structural model of
365 flagellar filament switching across multiple bacterial species. Nature communications.
366 2017;8(1):960.

367 16. Blum TB, Filippidou S, Fatton M, Junier P, Abrahams JP. The wild-type flagellar filament of the
368 Firmicute Kurthia at 2.8 Å resolution in vivo. Scientific reports. 2019;9(1):14948.

369 17. San Martin F, Fule L, Iraola G, Buschiazzo A, Picardeau M. Diving into the complexity of the
370 spirochetal endoflagellum. Trends in microbiology. 2023;31(3):294–307.

371 18. Hu D, Reeves PR. The Remarkable Dual-Level Diversity of Prokaryotic Flagellins. Zhaxybayeva O,
372 editor. mSystems. 2020 Feb 11;5(1):e00705-19.

373 19. Fields JL, Zhang H, Bellis NF, Petersen HA, Halder SK, Rich-New ST, et al. Structural diversity and
374 clustering of bacterial flagellar outer domains. Nat Commun. 2024 Nov 3;15:9500.

375 20. Nedeljković M, Kreutzberger MA, Postel S, Bonsor D, Xing Y, Jacob N, et al. An unbroken network
376 of interactions connecting flagellin domains is required for motility in viscous environments.
377 PLoS Pathogens. 2023;19(5):e1010979.

378 21. Kreutzberger MA, Sobe RC, Sauder AB, Chatterjee S, Peña A, Wang F, et al. Flagellin outer
379 domain dimerization modulates motility in pathogenic and soil bacteria from viscous
380 environments. Nature communications. 2022;13(1):1422.

381 22. Esteves NC, Bigham DN, Scharf BE. Phages on filaments: A genetic screen elucidates the complex
382 interactions between Salmonella enterica flagellin and bacteriophage Chi. PLoS Pathogens.
383 2023;19(8):e1011537.

384 23. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in Escherichia coli K-12
385 using PCR products. Proceedings of the National Academy of Sciences. 2000 Jun 6;97(12):6640–
386 5.

387 24. Kinosita Y, Ishida T, Yoshida M, Ito R, Morimoto YV, Goto K, et al. Distinct chemotactic behavior
388 in the original Escherichia coli K-12 depending on forward-and-backward swimming, not on run-
389 tumble movements. Sci Rep. 2020 Sep 28;10(1):15887.

390 25. UniProt: the universal protein knowledgebase in 2023. Nucleic acids research.
391 2023;51(D1):D523–31.

392 26. Suzek BE, Wang Y, Huang H, McGarvey PB, Wu CH, UniProt Consortium. UniRef clusters: a
393 comprehensive and scalable alternative for improving sequence similarity searches.
394 Bioinformatics. 2015;31(6):926–32.

395 27. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large
396 phylogenetic trees. In Ieee; 2010. p. 1–8.

397 28. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
398 Nucleic acids research. 2004;32(5):1792–7.

29. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics. 2014;30(22):3276–8.

30. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular biology and evolution. 2015;32(1):268–74.

31. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. Nature methods. 2022;19(6):679–82.

32. Ishida T, Ito R, Clark J, Matzke NJ, Sowa Y, Baker MA. Sodium-powered stators of the bacterial flagellar motor can generate torque in the presence of phenamil with mutations near the peptidoglycan-binding region. Molecular Microbiology. 2019;111(6):1689–99.

33. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nature methods. 2012;9(7):676–82.

34. Ridone P, Winter DL, Baker MAB. Tuning the stator subunit of the flagellar motor with coiled-coil engineering. Protein Science. 2023;32(12):e4811.

35. Hajam IA, Dar PA, Shahnawaz I, Jaume JC, Lee JH. Bacterial flagellin—a potent immunomodulatory agent. Experimental & molecular medicine. 2017;49(9):e373–e373.

36. Malapaka RRV, Adebayo LO, Tripp BC. A deletion variant study of the functional role of the Salmonella flagellin hypervariable domain region in motility. Journal of molecular biology. 2007;365(4):1102–16.

37. Kuwajima G. Construction of a minimum-size functional flagellin of Escherichia coli. J Bacteriol. 1988 Jul;170(7):3305–9.

38. Herzog B, Wirth R. Swimming Behavior of Selected Species of Archaea. Appl Environ Microbiol. 2012 Mar 15;78(6):1670–4.

39. Berg HC, Brown DA. Chemotaxis in Escherichia coli analysed by three-dimensional tracking. nature. 1972;239(5374):500–4.

40. Fenchel T. Motility and chemosensory behaviour of the sulphur bacterium Thiovulum majus. Microbiology. 1994 Nov 1;140(11):3109–16.

41. Klein A, Tóth B, Jankovics H, Muskotál A, Vonderviszt F. A polymerizable GFP variant. Protein Engineering, Design & Selection. 2012;25(4):153–7.

42. Klein Á, Szabó V, Kovács M, Patkó D, Tóth B, Vonderviszt F. Xylan-Degrading Catalytic Flagellar Nanorods. Mol Biotechnol. 2015 Sep;57(9):814–9.

432 **Supplementary Materials**

Strain	**Relevant Characteristics**	**Reference**
E. coli NEB 10-β	DH10B derivative. Shuttle vector cloning strain	New England Biolabs (C3019)
E. coli SYC29	*E. coli* RP437, Δ*motAB fliC*::*tetRA*, TET^R	This study
Plasmid	**Relevant Characteristics**	**Reference**
pET21(+)	Expression vector, T7 promoter, Amp^R, IPTG induction	Merck (69740)
pDB108	pKG116 backbone, motAB, CAM^R, arabinose induction	David Blair* (University of Utah, USA)
pET21_*fliC*	Wild type E. coli *fliC*, pET21a backbone	This study
pET21_*fliC_ΔD3*	*E. coli fliC* with deletion of 195-300, pET21a backbone	This study
pET21_*fliC_ΔD2/3*	*fliC* with deletion of 174-405, pET21a backbone	This study
pET21_*fliC_ST.D2/3*	*E. coli fliC* with D2/3 domains replaced by *Salmonella* Typhimurium strain LT2 D2/3 domains, pET21a backbone	This study
pET21_*fliC_STM.D2/3*	*E. coli fliC* with D2/3 domains replaced by *Salmonella* Typhimurium strain LT2 D2 domain and the *Mesorhizobium* sp. ORS 3359 D3 domain, pET21a backbone	This study
pET21_*fliC_STPA.D2/3*	*E. coli fliC* with D2/3 domains replaced by *Salmonella* Typhimurium strain LT2 D2 domain and the *Pseudomonas aeruginosa* D3 domain, pET21a backbone	This study
pET21_*fliC_STCF.D2/3*	*E. coli fliC* with D2/3 domains replaced by *Salmonella* Typhimurium strain LT2 D2 domain and the *Collimonas fungivorans* strain Ter331 D3 domain, pET21a backbone	This study
pET21_*fliC_HM.D2/3*	*E. coli fliC* with D2/3 domains replaced by the *Helicobacter mustelae* D2/3 domains, pET21a backbone	This study
pET21_*fliC_CF.D2/3*	*E. coli fliC* with D2/3 domains replaced by the *Collimonas fungivorans* strain Ter331 D2/3 domains, pET21a backbone	This study

433

434 **Table S1**. Bacterial strains and plasmids used in this study.

435 Kindly gifted from the *Blair Lab

436

UniProt ID	Bacteria	Protein ID	Length (Amino Acids)	D2/3 Domains Present
P04949	*E. coli* (strain K12)	Flagellin	498	Yes
A0A1H7N2K9	*Blastococcus* sp. DSM 46786	Flagellin	277	No
P13118	*Rhizobium meliloti*	Flagellin A	395	D2 Only
Q06064	*Bordetella bronchiseptica* (strain ATCC BAA-588)	Flagellin	391	D2 Only
Q6AJQ8	*Desulfotalea psychrophile* (DSM 12343)	Flagellin	857	Yes
P0A0S1	*Helicobacter pylori* (strain ATCC 700392)	Flagellin A	510	Yes
P80583	*Clostridium tyrobutyricum*	Flagellin	377	D2 Only
A0A1N6WG11	*Alkalispirochaeta americana*	Flagellin	284	No
P24216	***E. coli* (strain K12)**	**Flagellar hook-associated protein**	**468**	**N/A**

437

438 **Table S2**. Flagellin reference sequences used in the BLASTP search. The final row, FliD, (**Bold**) was used
439 as the outgroup (P24216).

IQ-Tree on ACCESS - Parameters	
abayes_test_	FALSE
datatype_	PROTEIN
extract_datefile_	FALSE
fix_branchlengths_	FALSE
lbp_test_	FALSE
median_approximation_	FALSE
more_memory2_	FALSE
more_memory_	FALSE
no_mlpairwise_	FALSE
optimize_weights_	FALSE
parametrical_test_	FALSE
per_sitefile_	FALSE
per_sitewplfile_	FALSE
per_sitewslfile_	FALSE
per_sitewslmfile_	FALSE
per_sitewslmrfile_	FALSE
per_sitewslrfile_	FALSE
per_sitewspmfile_	FALSE
per_sitewspmrfile_	FALSE
per_sitewsprfile_	FALSE
print_sitestats_	FALSE
runtime_	0.5
sh_test_	FALSE
slower_NNI_	FALSE
specify_fulltreesearch_	FALSE

specify_numpatterns_	921
specify_prefix_	output
specify_radius_	6
specify_safe_	FALSE
specify_symtype_	pval
thorough_estimation_	FALSE
unbiased_test_	FALSE
use_fasttreesearch_	FALSE
use_nj_	FALSE
use_random_	FALSE
use_symmtest_	FALSE
which_iqtree_	232
write_ancestralseqs_	FALSE
write_locally_optimal_	FALSE
write_loglikelihoods_	FALSE
write_sitelikelihoods_	FALSE

Table S3. Parameters used for the generation of the flagellin phylogeny using IQTree on ACCESS.

FliC Variant	Domains Present	Amino Acid Sequence	Length
FliC	D0, D1, D2, D3	MAQVINTNSLSLITQNNINKNQSALSSSIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQAARNANDGISVAQTTEGALSEINNNLQRVRELTVQATTGTNSESDLSSIQDEIKSRLDEIDRVSGQTQFNGVNVLAKNGSMKIQVGANDNQTITIDLKQIDAKTLGLDGFSVKNNDTVTTSAPVTAFGATTTNNIKLTGITLSTEAATDTGGTNPASIEGVYTDNGNDYYAKITGGDNDGKYYAVTVANDGTVTMATGATANATVTDANTTKATTITSGGTPVQIDNTAGSATANLGAVSLVKLQDSKGNDTDTYALKDTNGNLYAADVNETTGAVSVKTITYTDSSGAASSPTAVKLGGDDGKTEVVDIDGKTYDSADLNGGNLQTGLTAGGEALTAVANGKTTDPLKALDDAIASVDKFRSSLGAVQNRLDSAVTNLNNTTTNLSEAQSRIQDADYATEVSNMSKAQIIQQAGNSVLAKANQVPQQVLSLLQG	498
FliC ΔD3	D0, D1, D2	MAQVINTNSLSLITQNNINKNQSALSSSIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQAARNANDGISVAQTTEGALSEINNNLQRVRELTVQATTGTNSESDLSSIQDEIKSRLDEIDRVSGQTQFNGVNVLAKNGSMKIQVGANDNQTITIDLKQIDAKTLGLDGFSVKNNDTVTTSAPVTAFGATTAVSLVKLQDSKGNDTDTYALKDTNGNLYAADVNETTGAVSVKTITYTDSSGAASSPTAVKLGGDDGKTEVVDIDGKTYDSADLNGGNLQTGLTAGGEALTAVANGKTTDPLKALDDAIASVDKFRSSLGAVQNRLDSAVTNLNNTTTNLSEAQSRIQDADYATEVSNMSKAQIIQQAGNSVLAKANQVPQQVLSLLQG	393
FliC ΔD2/3	D0, D1	MAQVINTNSLSLITQNNINKNQSALSSSIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQAARNANDGISVAQTTEGALSEINNNLQRVRELTVQATTGTNSESDLSSIQDEIKSRLDEIDRVSGQTQFNGVNVLAKNGSMKIQVGANDNQTITIDLKQIDAKTLGLDGFKTTDPLKALDDAIASVDKFRSSLGAVQNRLDSAVTNLNNTTTNLSEAQSRIQDADYATEVSNMSKAQIIQQAGNSVLAKANQVPQQVLSLLQG	266
HM	D0, D1, D2, D3	MAQVINTNSLSLITQNNINKNQSALSSSIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQAARNANDGISVAQTTEGALSEINNNLQRVRELTVQATTGTNSESDLSSIQDEIKSRLDEIDRVSGQTQFNGVNVLAKNGSMKIQVGANDNQTITIDLKQIDAKTLGLDGFQVRINTGAMITAASEATLTFKQINGGGTSPLEGVKISHSVGTGLGVLAEVINKNSDKTGIRAKASVETTSDKEIMSGNLKNLTINDVNIGNIVDIKKGDADGRLVQAINALTSSTGVEASTDSKGRLNLRSVDGRGIVLKADASEDNGDGKSAPMAIDAVNGGQSITDGEGAANYGRLSLVRLDARDIVLTSSDKPDENKFSAIGFGDNNVAMATVNLRDVLGKFDASVKSASGANYNAVIASGNSNLKTTDPLKALDDAIASVDKFRSSLGAVQNRLDSAVTNLNNTTTNLSEAQSRIQDADYATEVSNMSKAQIIQQAGNSVLAKANQVPQQVLSLLQG	514
CF	D0, D1, D2	MAQVINTNSLSLITQNNINKNQSALSSSIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQAARNANDGISVAQTTEGALSEINNNLQRVRELTVQATTGTNSESDLSSIQDEIKSRLDEIDRVSGQTQFNGVNVL	373

		AKNGSMKIQVGANDNQTITIDLKQIDAKTLGLDGFNTLNLAGPSGATAAMTPTGKDNTGASVAVAA GGFTAVNRDAAAGFTATVADLVTDARGNTYVKNGANFYAATVTVADATQGATAATPTATVTYDSS KSLTAASGSPLKTTDPLKALDDAIASVDKFRSSLGAVQNRLDSAVTNLNNTTTNLSEAQSRIQDADYAT EVSNMSKAQIIQQAGNSVLAKANQVPQQVLSLLQG	
ST	D0, D1, D2, D3	MAQVINTNSLSLITQNNINKNQSALSSSIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQAARNAN DGISVAQTTEGALSEINNNLQRVRELTVQATTGTNSESDLSSIQDEIKSRLDEIDRVSGQTQFNGVNVL AKNGSMKIQVGANDNQTITIDLKQIDAKTLGLDGFQKYKVSDTAATVTGYADTTIALDNSTFKASATG LGGTDQKIDGDLKFDDTTGKYYAKVTVTGGTGKDGYYEVSVDKTNGEVTLAGGATSPLTGGLPATAT EDVKNVQVANADLTEAKAALTAAGVTGTASVVKMSYTDNNGKTIDGGLAVKVGDDYYSATQNKDG SISINTTKYTADDGTSKTALNKLGGADGKTEVVSIGGKTYAASKAEGHNFKAQPDLAEAAATTTENPL KTTDPLKALDDAIASVDKFRSSLGAVQNRLDSAVTNLNNTTTNLSEAQSRIQDADYATEVSNMSKAQI IQQAGNSVLAKANQVPQQVLSLLQG	499
STM	D0, D1, D2, D2	MAQVINTNSLSLITQNNINKNQSALSSSIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQAARNAN DGISVAQTTEGALSEINNNLQRVRELTVQATTGTNSESDLSSIQDEIKSRLDEIDRVSGQTQFNGVNVL AKNGSMKIQVGANDNQTITIDLKQIDAKTLGLDGFQKYKVSDTAATVTINVESTKLFDSGLSTAVKNQ GILDRKTSIYSTAADQSAYDAAYATAIAGGATDIAANTAGQAAVAGTPVKIDNVSAFNLDITATGVTD DIITQMVTNADLTEAKAALTAAGVTGTASVVKMSYTDNNGKTIDGGLAVKVGDDYYSATQNKDGSI SINTTKYTADDGTSKTALNKLGGADGKTEVVSIGGKTYAASKAEGHNFKAQPDLAEAAATTTENPLKT TDPLKALDDAIASVDKFRSSLGAVQNRLDSAVTNLNNTTTNLSEAQSRIQDADYATEVSNMSKAQIIQ QAGNSVLAKANQVPQQVLSLLQG	499
STCF	D0, D1, D2, D2	MAQVINTNSLSLITQNNINKNQSALSSSIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQAARNAN DGISVAQTTEGALSEINNNLQRVRELTVQATTGTNSESDLSSIQDEIKSRLDEIDRVSGQTQFNGVNVL AKNGSMKIQVGANDNQTITIDLKQIDAKTLGLDGFQKYKVSDTAATVTNTLNLAGPSGATAAMTPT GKDNTGASVAVAAGGFTAVNRDAAAGFTATVADLVTDARGNTYVKNGANFYAATVTVADATQGA TAATPTATVTYDSSKSLTAASGSPLNADLTEAKAALTAAGVTGTASVVKMSYTDNNGKTIDGGLAVK VGDDYYSATQNKDGSISINTTKYTADDGTSKTALNKLGGADGKTEVVSIGGKTYAASKAEGHNFKAQ PDLAEAAATTTENPLKTTDPLKALDDAIASVDKFRSSLGAVQNRLDSAVTNLNNTTTNLSEAQSRIQD ADYATEVSNMSKAQIIQQAGNSVLAKANQVPQQVLSLLQG	510
STPA	D0, D1, D2, D2	MAQVINTNSLSLITQNNINKNQSALSSSIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQAARNAN DGISVAQTTEGALSEINNNLQRVRELTVQATTGTNSESDLSSIQDEIKSRLDEIDRVSGQTQFNGVNVL AKNGSMKIQVGANDNQTITIDLKQIDAKTLGLDGFQKYKVSDTAATVTNGTYFKADGGGAVTAATA SGTVDIAIGITGGSAVNVKVDMKGNETAEQAAAKIAAAVNDANVGIGAFSDGDTISYVSKAGKDGS GAITSAVSGVVIADTGSTGVGTAAGVTPSATAFAKTNDTVAKIDISTANADLTEAKAALTAAGVTGTA SVVKMSYTDNNGKTIDGGLAVKVGDDYYSATQNKDGSISINTTKYTADDGTSKTALNKLGGADGKTE VVSIGGKTYAASKAEGHNFKAQPDLAEAAATTTENPLKTTDPLKALDDAIASVDKFRSSLGAVQNRLD SAVTNLNNTTTNLSEAQSRIQDADYATEVSNMSKAQIIQQAGNSVLAKANQVPQQVLSLLQG	535

Table S4. Sequences of FliC chimeras and outer domain deleted FliC. Sequences are coloured based on domains incorporated into the outer domains for *E. coli* FliC.

Primer Name	Template Sequence	5' -> 3' Sequence	Tm (°C)	Length
pET21_F	pET21(+)	TTCCTTTCGGGCTTTGTTAGCAG	66.4	23
pET21_R	pET21(+)	GGATCGAGATCTCGATCCCGC	66.9	21

Table S5. Primers and PCR conditions used in this study to confirm chimeric and outer domain deleted *fliC* into pET21(+).

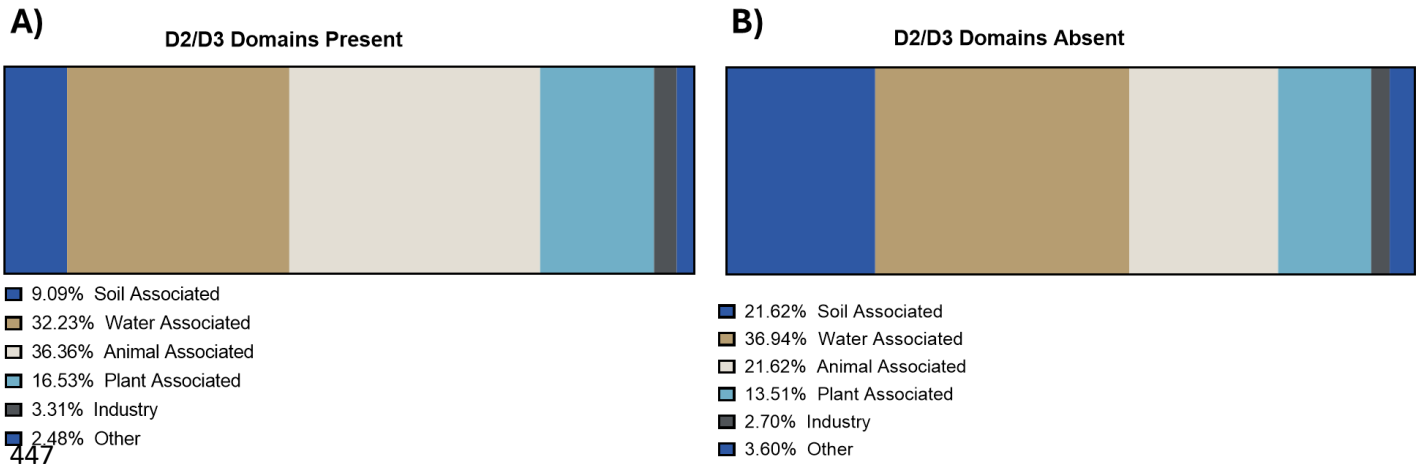
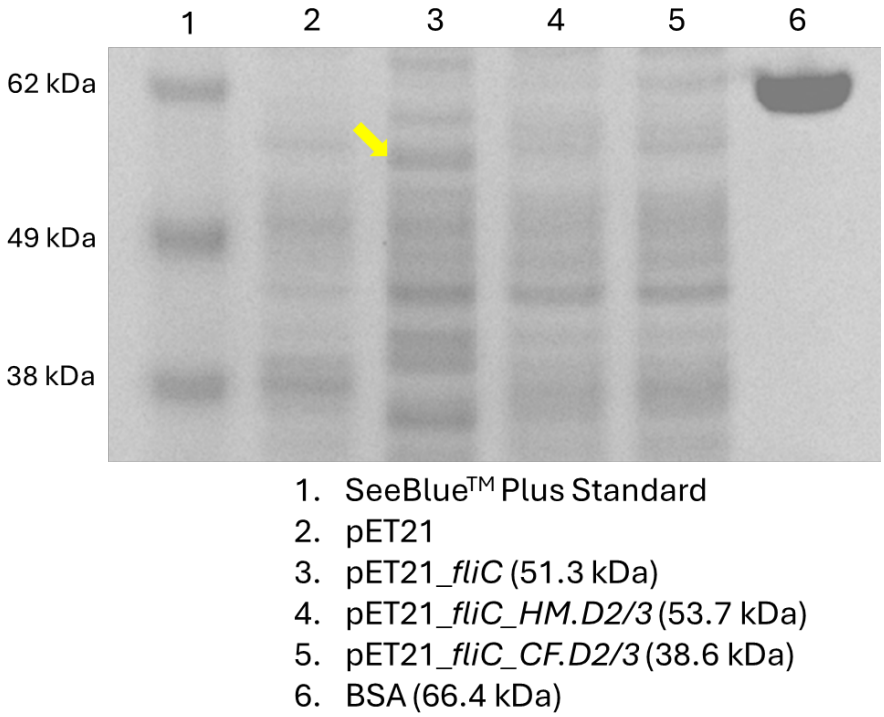


Figure S1. Isolation sources of bacteria used to make the flagellin phylogeny. Isolation sources are group in six major clusters; Soil (Blue), Aquatic (Dark Brown), Animal Associated (Light Brown), Plant Associated (Light Blue), Industrial (Grey) and Other (Dark Blue).

D0 D1 D2 D3

```
Mesorhizobium(ORS3359)_A0A090GIX    -MASIMTNAAALTALQSLNATNKSLEQTQARISTGYRVSEASDNAAYWSIATTMRSDNSA
H.mustelae_P50612                   MAFQVNTNINALTT--SAGATQLGLKNSLEKLSSGLRINKAADDASGMTISDSLRSQASA
P.aeruginosa_P21184                 MALTVNTNIASLNTQRNLNNSSASLNTSLQRLSTGSRINSAKDDAAGLQIANRLTSQVNG
S.Typhimurium(strain_LT2)_P06179    MAQVINTNSLSLLTQNNLNKSQSALGTAIERLSSGLRINSAKDDAAGQAIANRFTANIKG
C.fungivorans(strain_Ter331)_G0A    -MSVINTNILSLVAQNNLTKSASSLSSAIQRLSSGMRINSAKDDAAGQAIANRFTANITG
E.coli(strain_K12)_P04949           MAQVINTNSLSLITQNNINKNQSALSSSIERLSSGLRINSAKDDAAGQAIANRFTSNIKG
                                        : **  :* :  .   .  .*  :  .:*:* *:..* *:*:   *:  : :: ..

Mesorhizobium(ORS3359)_A0A090GIX    LSTVQDALGLGASKVDTAYTGMNNVLDTIGKIKTKLLSAV--GQSDANKAKTQTEITALQ
H.mustelae_P50612                   LGQAISNANDGIGIIQVADKAMDEQLKILDTIKVKATQAAQDGQSLESRKAIQSDIIRLI
P.aeruginosa_P21184                 LNVATKNANDGISLAQTAEGALQQSTNILQRMRDLSLQSANGSNSDSERTALNGEVKQLQ
S.Typhimurium(strain_LT2)_P06179    LTQASRNANDGISIAQTTEGALNEINNNLQRVRELAVQSANSTNSQSDLDSIQAEITQRL
C.fungivorans(strain_Ter331)_G0A    LTQASRNANDGISLAQTTEGALTEVNNNLQRIRELSVQAASGTNSSSDLQSIQDEVTQRL
E.coli(strain_K12)_P04949           LTQAARNANDGISVAQTTEGALSEINNNLQRVRELTVQATTGTNSESDLSSIQDEIKSRL
                                    *  .    . * .  :.:  .: :  . :  :.    .:.   :*  .    : ::

Mesorhizobium(ORS3359)_A0A090GIX    AQMKSFADAATFSGSNYL----SVTSKQV-AAANDGVQPDAQIVSS--------------
H.mustelae_P50612                   QGLDNIGNTTSYNGQSLLSGQWTNKEFQIGTYSNQSIKVSVGSTTSDKIGQVRINTGAMI
P.aeruginosa_P21184                 KELDRISNTTTFGGRKLLDGSFGVASFQVGSAANEIISVGIDEMSAES------------
S.Typhimurium(strain_LT2)_P06179    NEIDRVSGQTQFNGVKVLAQD-NTLTIQVGANDGETIDIDLKQINSQTLGLDTLNVQQKY
C.fungivorans(strain_Ter331)_G0A    KEIDRISSQTDFNGVKVLAAN-QTLNIQVGANDGQTIGIDLQKIDSTALKLNTLN-----
E.coli(strain_K12)_P04949           DEIDRVSGQTQFNGVNVLAKN-GSMKIQVGANDNQTITIDLKQIDAKTLGLDGFSVKNND
                                      :. ... : :.* . *         *: :  .: :  .     :

Mesorhizobium(ORS3359)_A0A090GIX    -------------------------FNRSSSGAISL------------------------
H.mustelae_P50612                   TAASEATLTFKQINGGGTSPLEGVKISHSVGTGLGVLAEVINKNSDKTGIRAKASVETTS
P.aeruginosa_P21184                 --------------------LNGTYFKA----DGGG------------------------
S.Typhimurium(strain_LT2)_P06179    KVSDTAATVTGYADTTIA--LDNSTFKASATGLGGT-----DQKIDGDLKFDDTTGKYYA
C.fungivorans(strain_Ter331)_G0A    --------------------LAGPSGATAAMTPTGK------------------------
E.coli(strain_K12)_P04949           TVTTSAPVTAFGATTTNNIKLTGITLSTEAATDTGG-----TNPASIEGVYTDNGNDYYA
                                                                            .

Mesorhizobium(ORS3359)_A0A090GIX    ----------------------GTIDINVESTKLFDSGLSTAVKNQ--------------
H.mustelae_P50612                   DKEIMSGNLKNL--TINDVNIGNIVDIKKGDAD----GRLVQAINALTSSTGVEAS-TDS
P.aeruginosa_P21184                 --------------AVTAATASGTVDIAIGITG----GSAVNVKVDMK------GNETAE
S.Typhimurium(strain_LT2)_P06179    KVTVTGGTGKDGYYEVSVDKTNGEVTLAGGATSPLTGGLPATATEDVKNVQVANADLTEA
C.fungivorans(strain_Ter331)_G0A    ------------------DNTGASVAVAAG-------GFTAVNRD---------------
E.coli(strain_K12)_P04949           K--ITGGDNDGKYYAVTVAN-DGTVTMATGATA----NATVTDANTTKATTITSGGTPVQ
                                                                 : :          .  .

Mesorhizobium(ORS3359)_A0A090GIX    -------GILDRKTSIYSTAADQSAYDAAYATAIAG---GATDIAAN-------------
H.mustelae_P50612                   KGRLNLRSVDGRGIVLKADASEDNGDGKSAPMAIDAVNGGQSITDGEGAANYGRLSLVRL
P.aeruginosa_P21184                 QAAAKIAAAVNDANVGIGAFSD--GDTI-SYVS-KA---GKDGSGAITSAVSGVVIA---
S.Typhimurium(strain_LT2)_P06179    KAALTAAGVTGTASVVKMSYTDNNGKTIDGGLAVKV---GDDYYSAT-QNKDGSISI---
C.fungivorans(strain_Ter331)_G0A    -------AAAGFTATVADLVTDARGNTY-----VKN---GANFYAAT-------------
E.coli(strain_K12)_P04949           IDNTAGSATANLGAVSLVKLQDSKGNDTDTYALKDT---NGNLYAADVNETTGAVSV---
                                          .  .          :  .              .     .

Mesorhizobium(ORS3359)_A0A090GIX    ---TAGQAAVAGTPVKIDNVSAFN------------------------------------
H.mustelae_P50612                   DARDIVLTSSDKPDENKFSAIGFGDNNVAMATVNLRDVLGKFDASVKSASGANYNAVIAS
P.aeruginosa_P21184                 ---DTGSTGVGTAAGVTPSATAFA-------------K--TNDTV---------------
S.Typhimurium(strain_LT2)_P06179    ---NTTKYTADDGTSKT-ALNKLGGADGKTEVVSIGGK--TYAAS--KAEGHNFKAQ---
C.fungivorans(strain_Ter331)_G0A    ---VTVADATQGATAATPTATV------------------TYDSS---------------
E.coli(strain_K12)_P04949           ---KTITYTDSSGAASSPTAVKLGGDDGKTEVVDIDGK--TYDSA--DLNGGNLQTGLTA

Mesorhizobium(ORS3359)_A0A090GIX    --LDITATGVTDDIITQMVTKIDNVMSQLTDAATVLGAAKSSIDLQKTFTQSLMDSIDRG
H.mustelae_P50612                   GNSNLGAGVTTLVGAMLVMDIADSARKTLDKIRSDLGSVQGQMVSTVNNISVTQVNVKAA
P.aeruginosa_P21184                 --AKIDISTAK--GAQSAVLVIDEAIKQIDAQRADLGAVQNRFDNTINNLKNIGENVSAA
S.Typhimurium(strain_LT2)_P06179    --PDLAEAAAT--TTENPLQKIDAALAQVDTLRSDLGAVQNRFNSAITNLGNTVNNLTSA
C.fungivorans(strain_Ter331)_G0A    --KSLTAASGS------PLTAMDSALSQVDSLRSSLGAVQNRFASTIANLGTTVTNLSSS
E.coli(strain_K12)_P04949           GGEALTAVANG--KTTDPLKALDDAIASVDKFRSSLGAVQNRLDSAVTNLNNTTTNLSEA
                                        :             :   * .   :    : **:.:. :            .:  .

Mesorhizobium(ORS3359)_A0A090GIX    VGQLVDADMNKESTRLQALQVQQQLGIQSLSIANSSSQSILQLFKNG
H.mustelae_P50612                   ESRMREVDFAAESAEFNKYNILAQ-----------------------
P.aeruginosa_P21184                 RGRIEDTDFAAETANLTKNQVLQQAGTAILAQANQLPQSVLSLLR--
S.Typhimurium(strain_LT2)_P06179    RSRIEDSDYATEVSNMSRAQILQQAGTSVLAQANQVPQNVLSLLR--
C.fungivorans(strain_Ter331)_G0A    RSGIQDADYATEVSNMTRSQILQQAGTSVLAKANQSTQSVLTLLQ--
E.coli(strain_K12)_P04949           QSRIQDADYATEVSNMSKAQIIQQAGNSVLAKANQVPQQVLSLLQG-
                                     . : : *   * : :   ::  *
```

453 **Figure S2**. Alignment of flagellins used to make two and three species FliC chimeras. Colours indicate
454 flagellin domains (Blue = D0, Orange = D1, Green = D2, Purple = D3).



455

456 **Figure S3**. SDS-PAGE gel of non-motile FliC chimera constructs (ECHM and ECCF) heat treated
457 supernatants. Empty vector pET21 and pET21_*fliC* were used as a negative and positive control,
458 respectively. BSA was used as a marker control.