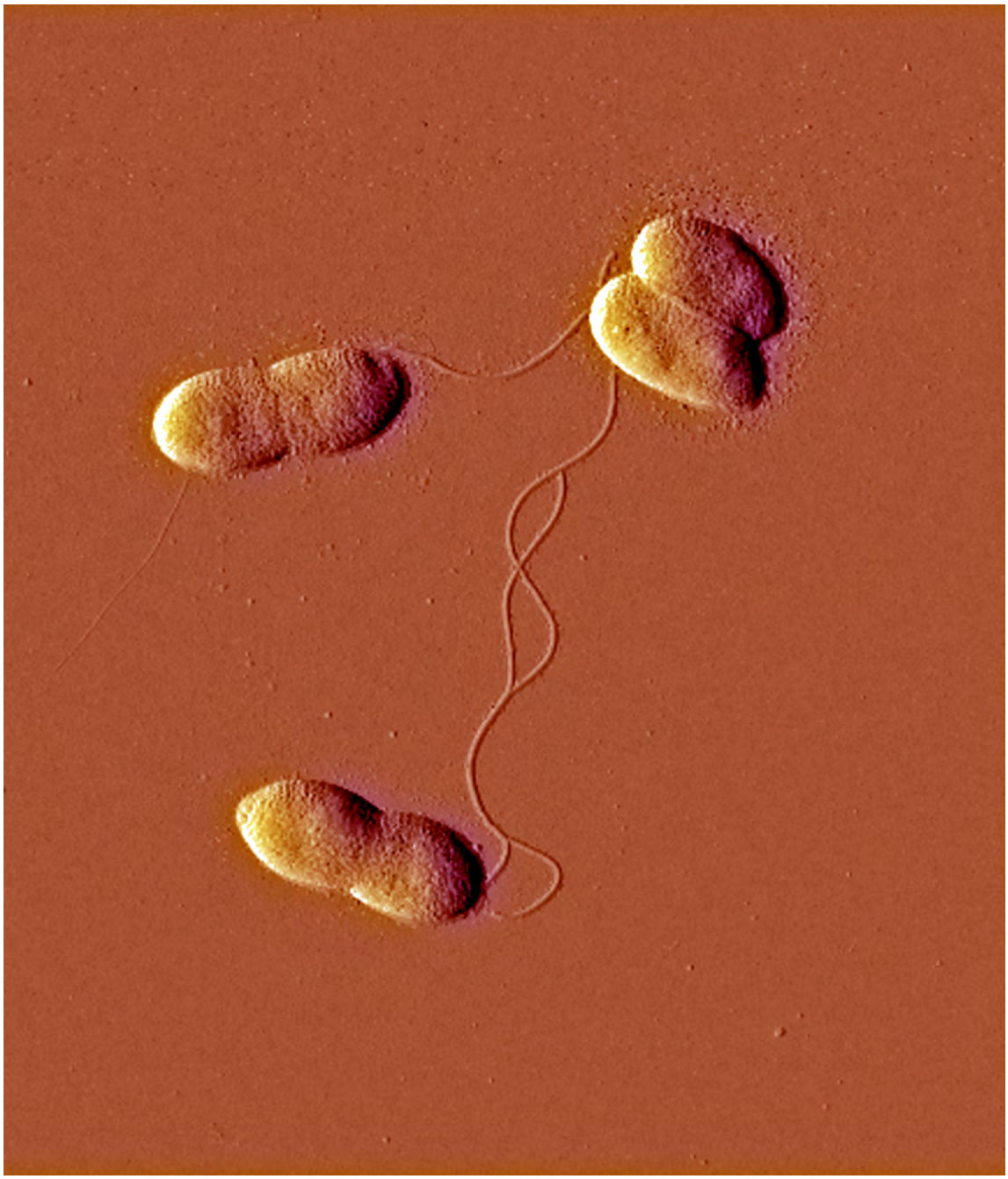

December 2021
Volume 203
Number 23
Published Twice Monthly

AMERICAN SOCIETY FOR MICROBIOLOGY

JB

Journal of Bacteriology

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology®

RESEARCH ARTICLE

Multiple Copies of *flhDC* in *Paraburkholderia unamae* Regulate Flagellar Gene Expression, Motility, and Biofilm Formation

Shelley N.-M. Thai,[a] Michelle R. Lum,[b] Jeanine Naegle,[a] Michael Onofre,[b] Hassan Abdulla,[b] Allison Garcia,[a] Andreh Fiterz,[a] Ashley Arnell,[b] Thuthiri T. Lwin,[b] Aaron Kavanaugh,[c] Zade Hikmat,[a] Nora Garabedian,[a] Ryan Toan Ngo,[a] Brenda Dimaya,[b] Adan Escamilla,[b] Luiza Barseghyan,[a] Maria Shibatsuji,[b] Salma Soltani,[b] Luke Butcher,[a] Firas Hikmat,[a] Dro Amirian,[a] Artin Bazikyan,[a] Nathan Brandt,[a] Mary Sarkisian,[a] Xavier Munoz,[a] Andrew Ovakimyan,[a] Emily Burnett,[a] Jennifer Ngoc Pham,[a] Ania Shirvanian,[a] Roberto Hernandez,[a] Maria Vardapetyan,[a] Matthew Wada,[a] Cuauhtemoc Ramirez,[a] Martin Zakarian,[a] Fabrizio Billi[c]

[a]Division of Biology, Glendale Community College, Glendale, California, USA
[b]Department of Biology, Loyola Marymount University, Los Angeles, California, USA
[c]OIC-Department of Orthopaedic Surgery, David Geffen School of Medicine, University of California, Los Angeles, California, USA

Shelley N.-M. Thai and Michelle R. Lum contributed equally to this work. Author order was determined by reverse alphabetical order.

ABSTRACT FlhDC is a heterohexameric complex that acts as a master regulator of flagellar biosynthesis genes in numerous bacteria. Previous studies have identified a single *flhDC* operon encoding this complex. However, we found that two *flhDC* loci are present throughout *Paraburkholderia*, and two additional *flhC* copies are also present in *Paraburkholderia unamae*. Systematic deletion analysis in *P. unamae* of the different *flhDC* copies showed that one of the operons, *flhDC1*, plays the predominant role, with deletion of its genes resulting in a severe inhibition of motility and biofilm formation. Expression analysis using promoter-*lacZ* fusions and real-time quantitative PCR support the primary role of *flhDC1* in flagellar gene regulation, with *flhDC2* a secondary contributor. Phylogenetic analysis shows the presence of the *flhDC1* and *flhDC2* operons throughout *Paraburkholderia*. In contrast, *Burkholderia* and other bacteria only carry the copy syntenous with *flhDC2*. The variations in impact each copy of *flhDC* has on downstream processes indicate that regulation of FlhDC in *P. unamae*, and likely other *Paraburkholderia* species, is regulated at least in part by the presence of multiple copies of these genes.

IMPORTANCE Motility is important in the colonization of plant roots by beneficial and pathogenic bacteria, with flagella playing essential roles in host cell adhesion, entrance, and biofilm formation. Flagellar biosynthesis is energetically expensive. Its complex regulation by the FlhDC master regulator is well studied in peritrichous flagella expressing enterics. We report the unique presence throughout *Paraburkholderia* of multiple copies of *flhDC*. In *P. unamae*, the *flhDC1* copy showed higher expression and a greater effect on swim motility, flagellar development, and regulation of downstream genes, than the *flhDC2* copy that is syntenous to *flhDC* in *Escherichia coli* and pathogenic *Burkholderia* spp. The *flhDC* genes have evolved differently in these plant-growth-promoting bacteria, giving an additional layer of complexity in gene regulation by FlhDC.

KEYWORDS *Paraburkholderia*, RT-qPCR, biofilm, flagellar gene regulation, *flhDC*, *fliA*, *fliC*, motility

The bacterial flagellum is a complex rotary engine required by many bacteria for swimming motility. Motility plays a key role in the colonization of plant roots, by both beneficial and pathogenic bacteria, with the flagellum allowing for chemotaxis toward root exudates (1–3). The flagellum can also be important for adhesion to host

Citation Thai SN-M, Lum MR, Naegle J, Onofre M, Abdulla H, Garcia A, Fiterz A, Arnell A, Lwin TT, Kavanaugh A, Hikmat Z, Garabedian N, Ngo RT, Dimaya B, Escamilla A, Barseghyan L, Shibatsuji M, Soltani S, Butcher L, Hikmat F, Amirian D, Bazikyan A, Brandt N, Sarkisian M, Munoz X, Ovakimyan A, Burnett E, Pham JN, Shirvanian A, Hernandez R, Vardapetyan M, Wada M, Ramirez C, Zakarian M, Billi F. 2021. Multiple copies of *flhDC* in *Paraburkholderia unamae* regulate flagellar gene expression, motility, and biofilm formation. J Bacteriol 203: e00293-21. https://doi.org/10.1128/JB.00293-21.

Editor Michael Y. Galperin, NCBI, NLM, National Institutes of Health

Address correspondence to Michelle R. Lum, michelle.lum@lmu.edu.

Received 30 May 2021
Accepted 10 September 2021
Accepted manuscript posted online 20 September 2021
Published 5 November 2021

cells, host cell entrance, and biofilm formation (4–6). Thus, the flagellum is often a critical component facilitating bacterial association with the host, in both beneficial and pathogenic associations.

In *Escherichia coli*, more than 50 genes are required for the biosynthesis and function of the flagellum (7–9). Thus, it is not surprising that motility is an energetically expensive process (10) and is regulated by a hierarchical level of gene expression. At the top of this hierarchy is the class I *flhDC* master operon, whose gene products form a heterohexameric complex ($FlhD_4C_2$) that binds upstream of class II promoters and recruits RNA polymerase to promote transcription (11–13). The class II genes encode the basal body and flagellar export system as well as FliA (σ^{28}). FliA is a transcriptional regulator that controls expression of the class III genes, such as *fliM*, which encodes the flagellar motor switch protein (14, 15). FlhDC is also implicated in the regulation of other pathways, including anaerobic metabolism and the Entner-Doudoroff pathway (8, 16).

The *flhDC* operon itself is transcriptionally controlled by numerous transcription factors that allow regulation by environmental cues, including quorum sensing (QseBC), temperature, osmolarity (OmpR) (17, 18), catabolite repression (CRP), and pH (19–21). Additionally, in *E. coli* and *Erwinia amylovora*, it was found that there is posttranscriptional regulation by small RNAs (sRNAs) (22–25) and RNA-binding proteins (9, 26, 27). $FlhD_4C_2$ binding to the DNA is also regulated by anti-$FlhD_4C_2$ factors through their interaction with either the FlhD or FlhC subunit (28–30).

The $FlhD_4C_2$ activator is generally thought to control the regulation of lateral (peritrichous) flagella, and has been characterized in enterics, such as *E. coli* and *Salmonella enterica* (31), as well as some *Betaproteobacteria*, such as *Ralstonia* (32). Only one copy of the *flhDC* operon is typically present in the organisms in which these genes have been studied and published. Bacteria with a polar flagellum, such as *Pseudomonas* and *Vibrio*, generally use the σ^{54}-dependent NtrC family of transcriptional activators (33, 34). However, *Burkholderia glumae*, a plant pathogen that is polarly flagellated, controls flagellar genes through an FlhDC system (35). *B. glumae* mutants in *flhD* are nonmotile and avirulent on rice (35), again demonstrating the importance of motility for the successful colonization of the plant.

Burkholderia sensu lato encompasses a wide range of pathogenic and environmental bacteria, including *Paraburkholderia* spp., which are plant-associative beneficial and environmental (PBE) bacteria that were recently split from the pathogenic *Burkholderia* group (36–38). *Paraburkholderia unamae* is a soil bacterium found in the rhizosphere of diverse plant species, including maize, coffee, sugarcane, and tomato (39–42). It is interesting agriculturally for its potential as a bioinoculant, as it can reside endophytically and promote plant growth, with plant-growth-promoting properties including N_2 fixation, ACC deaminase activity, siderophore production, and phosphate solubilization (40, 43). In addition, *P. unamae* can utilize and degrade aromatic compounds, so may be useful for bioremediation of organic pollutants (40).

We report here that multiple copies of the *flhD* and *flhC* genes are present in *P. unamae* as well as other beneficial plant-associative *Paraburkholderia* spp. We investigate the biological role of the *flhD* and *flhC* genes in *P. unamae* to establish a general framework to understand why these gene duplications may be useful. We show through systematic mutant analysis that the different copies of *flhD* and *flhC* impact swimming motility as well as play a role in biofilm formation. Furthermore, we show that the level of expression of the different copies correlates with how these mutations affect motility.

RESULTS

Multiple copies of *flcD* and *flhC* are present in *Paraburkholderia* species. We had previously identified a *P. unamae* transposon mutant of *flhC*, MO384, that is greatly reduced, though not completely inhibited, in swimming motility, and determined the transposon to reside within the *flhC* gene (M. Onofre and M. R. Lum, unpublished data). The *flhC* gene in other organisms is required for motility and resides in an operon downstream of *flhD*. Intrigued by the reduction versus absence of motility in the MO384 mutant,

we performed a BLASTP search of the *P. unamae* MTI-641^T genome using Integrated Microbial Genomes (IMG) tools and found four copies of *flhC*. Two of the *P. unamae flhC* copies reside in operons with *flhD*, which was supported by Operon-Mapper (44). We designated the copy with the transposon insertion as *flhC1* and the operon it resides in as *flhDC1*. The other operon was designated *flhDC2*. The *flhDC1* operon is in a locus with a manganese transporter and multidrug efflux pump (Fig. 1A). The *flhDC2* operon resides upstream of the *motA* and *motB* genes (Fig. 1A), which is similar in location to the single *flhDC* operon in *E. coli* and *B. glumae*. The third copy of *flhC*, *flhC3*, resides alone. A fourth putative copy, *flhC4*, is in an operon with hypothetical proteins (Fig. 1A), with the predicted protein having an E value to *E. coli* FlhC just under e−05 (Table 1), suggesting it could be homologous. There are only two copies of *flhD* in the *P. unamae* genome.

Protein sequence comparison between FlhD1 and FlhD2 shows that they share 58.09% identity and 64.76% similarity to each other (Fig. 1B and Table 1). FlhC1 shows the greatest protein sequence identity to FlhC2 (58.33%), followed by FlhC3 (Fig. 1C and Table 1). *flhC4* encodes a longer protein (406 amino acids [aa]), in contrast to the 180- to 200-aa proteins encoded by the other *flhC* copies, and it is the C-terminus half of the sequence that aligns with the other FlhC sequences. However, the identity is much less than that seen between other copies, with FlhC4 showing the greatest identity to FlhC2 (22.22% identity and 33.88% similarity) (Fig. 1C; Table 1). The single FlhDC operon of the betaproteobacterium *B. glumae* has greatest identity to FlhDC2 (Fig. 1B and C; Table 1). Thus, it is FlhDC2 that shows the greatest similarity to the FlhDC previously identified in other organisms.

FlhDC has been characterized in numerous organisms because of its critical role in motility and other processes. However, in all these organisms, only a single copy of these genes is present. Intrigued by the multiple copies present in *P. unamae*, we investigated whether multiple copies are present in other species. We carried out a BLASTP analysis of FlhC or FlhD against genomes in IMG. We found multiple copies of both *flhD* and *flhC* in all species of *Paraburkholderia* and *Caballeronia* (see Table S3 in the supplemental material) (data not shown), which are both genera that reside within *Burkholderia sensu lato* and were recently separated from *Burkholderia* (45). In contrast, few species of *Burkholderia* contain multiple copies of *flhD* or *flhC* (Table S3). A single copy of *flhDC* is typically present in genera outside of *Burkholderia sensu lato*, although duplications occur on occasion—for instance, in *Cuprividus taiwanensis* and specific strains of *Erwinia amylovora* (Table S3).

We used the IMG gene neighborhood tool to compare the loci for the different *flhD* and *flhC* gene copies. We found that the locus containing the *flhDC2* operon upstream of the *motA* and *motB* genes is syntenous not only in *E. coli* and *B. glumae* but also in other organisms with *flhDC*. However, there are exceptions, such as with *Paraburkholderia sacchari*, in which *flhC2* is deleted (see Fig. S1 and Table S3 in the supplemental material). All *Paraburkholderia* and *Caballeronia* species we analyzed additionally contained a genetically syntenous *flhDC1* operon (Fig. S1; Table S3). Using IMG tools, as well as tBLASTn against the NCBI nucleotide database, we could not find any organism outside of *Paraburkholderia* and *Caballeronia* containing the *flhDC1* locus. The *flhC3* and *flhC4* genes were in loci syntenous in other genomes, but in a small subset of *Paraburkholderia* (Fig. S1). We also identified *flhC4* copies in some strains of *Burkholderia* as well as in *C. taiwanensis*. Additional copies of *flhD* or *flhC*, including those in *Burkholderia* and *Cupriavidus*, appear to be unique, and the neighboring regions lack synteny with each other or that of *flhDC1* or *flhDC2* (Fig. S1).

We then carried out phylogenetic analyses to determine the evolutionary relationship between *flhD* and *flhC* in different species. As the numbers of copies of *flhD* and *flhC* are often different, we generated a phylogram for each *flhD* and *flhC*. High bootstrap values support the distinct clades that are formed for each of the gene copies, which correspond to the distinct genetic loci of the various gene copies. The *flhD* copies grouped into two major clades, with all members of *Paraburkholderia* represented in each of the two groups (Fig. 2A). The *flhD1* group contains only *Paraburkholderia* and *Caballeronia*. In contrast, the *flhD2* group includes all organisms that contain

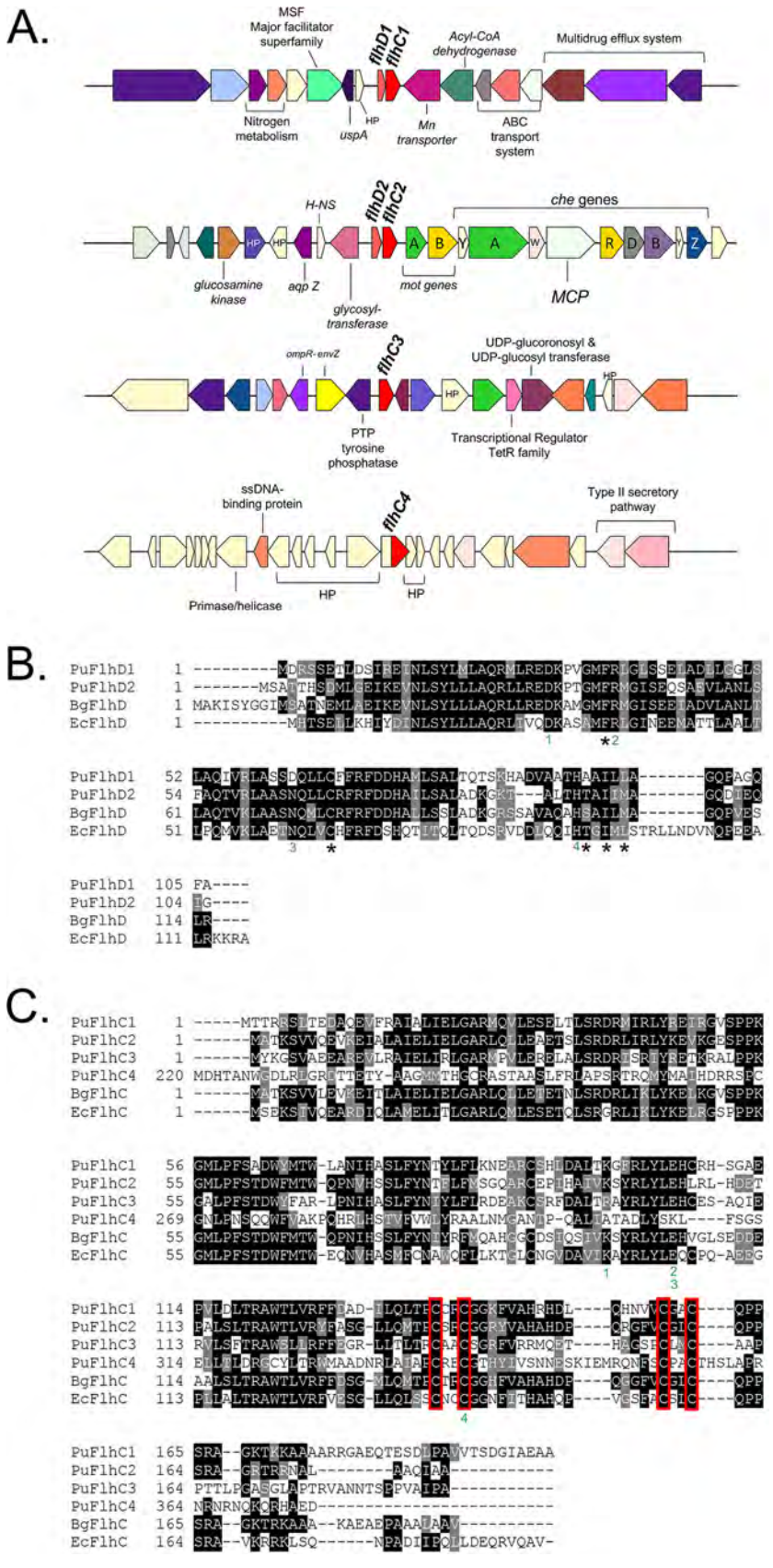

FIG 1 (A) Genetic maps of the *P. unamae flhDC* regions, showing two *flhDC* operons and two additional *flhC* genes. (B and C) Alignment of the different *P. unamae* (Pu) FlhC and FlhD proteins and those of

(Continued on next page)

TABLE 1 Percentages of identity and similarity of FlhC and FlhD proteins

	% of identity and similarity to protein shown[a]					
Protein	**PuFlhD1**	**PuFlhD2**	**BgFlhD**	**EcFlhD**		
PuFlhD1	100 [2e−73]					
PuFlhD2	58.09 (64.76) [1e−38]	100 [7e−75]				
BgFlhD	56.6 (65.09) [4e−42]	70.47 (81.9) [9e−52]	100 [2e−81]			
EcFlhD	44.33 (56.6) [4e−29]	41.9 (55.23) [1e−25]	46.08 (54.78) [8e−34]	100 [9e−84]		
	PuFlhC1	**PuFlhC2**	**PuFlhC3**	**PuFlhC4**	**BgFlhC**	**EcFlhC**
PuFlhC1	100 [7e−152]					
PuFlhC2	58.33 (71.11) [2e−82]	100 [8e−136]				
PuFlhC3	49.47 (59.47) [2e−63]	44.44 (55.0) [1e−53]	100 [3e−139]			
PuFlhC4	21.5 (29.5) [2e−05]	22.22 (33.88) [2e−05]	16.31 (26.31) [0.002]	100 [0.0]		
BgFlhC	56.91 (65.42) [2e−77]	76.66 (83.88) [2e−102]	43.61 (53.19) [7e−54]	20.75 (32.44) [4e−08]	100 [3e−143]	
EcFlhC	47.87 (57.81) [2e−65]	56.11 (65.55) [8e−77]	38.94 (49.47) [7e−49]	20.31 (27.6) [6e−06]	55.31 (63.82) [6e−74]	100 [6e−145]

[a]The percentages of identity of the different *P. unamae* (Pu) FlhC and FlhD copies are given relative to each other and relative to *B. glumae* (Bg) and *E. coli* (Ec) FlhC and FlhD, and the percentages of similarity are indicated in parentheses. E values from a BLAST comparison are shown in brackets.

flhDC. The *flhC* phylogram shows four distinct clades corresponding to *flhC1*, *flhC2*, *flhC3*, and *flhC4*, which again correspond to groups of genetically syntenous loci (Fig. 2B). The *flhC1* and *flhC2* clades were largely congruent with those found for *flhD1* and *flhD2*. The *flhC3* clade is evolutionarily closer to *flhC1*. The analysis also shows that the additional copies of *flhD* or *flhC* found in some species of *Paraburkholderia* were most similar to *flhDC1*. Overall, the phylograms confirm that distinct lineages of *flhD* and *flhC* exist for *flhDC1*, *flhDC2*, *flhC3*, and *flhC4*. Although some *Burkholderia* species have multiple copies of the genes, none of the copies shared a neighborhood with *flhDC1*.

FlhD and FlhC retain conserved amino acids. Alanine scanning mutagenesis of FlhD in *E. coli* has identified a number of critical residues for binding to FlhC (46), and likewise, residues of FlhC have been determined that interact with FlhD (13). In addition, in FlhC, four cysteine residues make up a zinc-binding site and may be important for FlhC binding to DNA (13). We used this information to extrapolate the potential functional importance of certain residues in FlhD and FlhC in *P. unamae*. Alignment of FlhD protein sequences revealed that key amino acid residues important for FlhD function in *E. coli* are conserved in *Paraburkholderia* copies, including *P. unamae* (Fig. 1B) (data not shown). Specifically, Cys-65 is important for the FlhD dimerization, and Asp-28, Phe-34, Arg-35, His-91, and Ile-94 are important for FlhD to form a complex with FlhC in *E. coli* (Fig. 1B). These residues are present in both FlhD1 and FlhD2 of all species analyzed (Fig. 1B) (data not shown). Interestingly, the alignment also showed that at certain positions, FlhD1 and FlhD2 have different conserved residues. Asn-61 (*E. coli*) is present in FlhD2, but it is Asp in all FlhD1 sequences we analyzed (Fig. 1B) (data not shown). Similarly, at the residue that aligned with Thr-92 in *E. coli*, it is either a Thr or Ser. However, in all species analyzed, it is an Ala in FlhD1. At Leu-96, it is either predominantly Met or Leu in FlhD2, but it is Leu in FlhD1 (Fig. 1B) (data not shown). Asn-61, Thr-92, and Leu-96 are also important in forming the FlhDC complex. Residues in FlhC known to interact with FlhD in *E. coli* are well conserved in FlhC1, FlhC2, and FlhC3 but not in FlhC4 (Fig. 1C). However, all copies of FlhC, including FlhC4, have the four cysteines involved in zinc binding. Overall, it appears that the various FlhC and FlhD copies retain many of the critical residues found in *E. coli* that are important for forming the FlhDC complex and binding to DNA.

FlhD and FlhC copies have different impacts on motility. The multiple copies of *flhD* and *flhC* suggest there may be some functional redundancy in the duplicate cop-

FIG 1 Legend (Continued)
B. glumae (Bg) and *E. coli* (Ec). Black boxes indicate identities, and gray boxes indicate similarities. Asterisks and numbers below the alignment designate residues in *E. coli* known to be involved in dimerization and binding, with numbers corresponding to the amino acids that interact between FlhD and FlhC. Positions in red boxes designate the cysteines in *E. coli* FlhC that are involved in zinc binding.

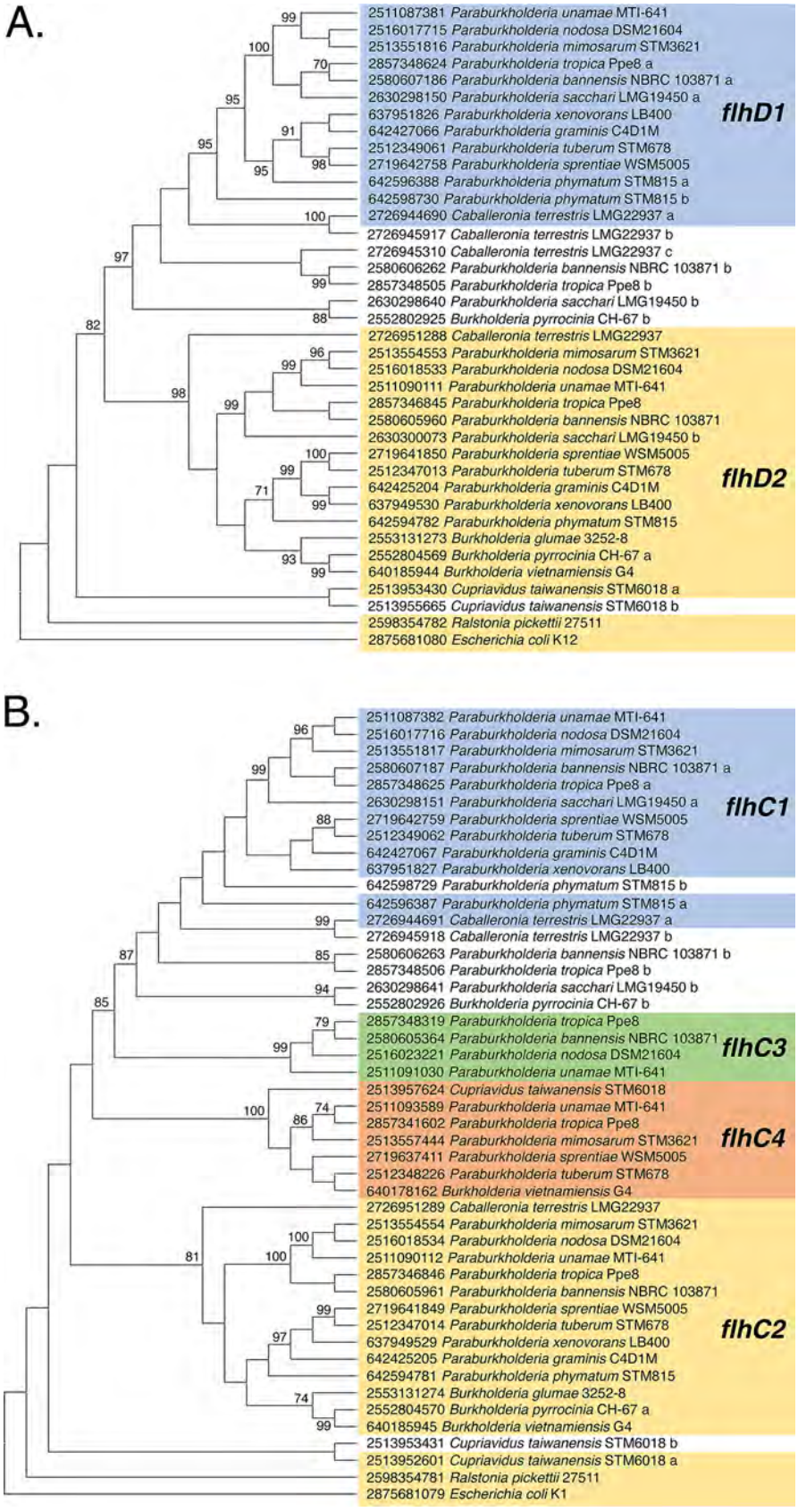

FIG 2 Phylogenetic analysis of *flhD* and *flhC* DNA sequences by maximum likelihood analysis. (A) Analysis based on the *flhD* gene. (B) Analysis based on the *flhC* gene. Bootstrap values greater than 70% are shown

(Continued on next page)

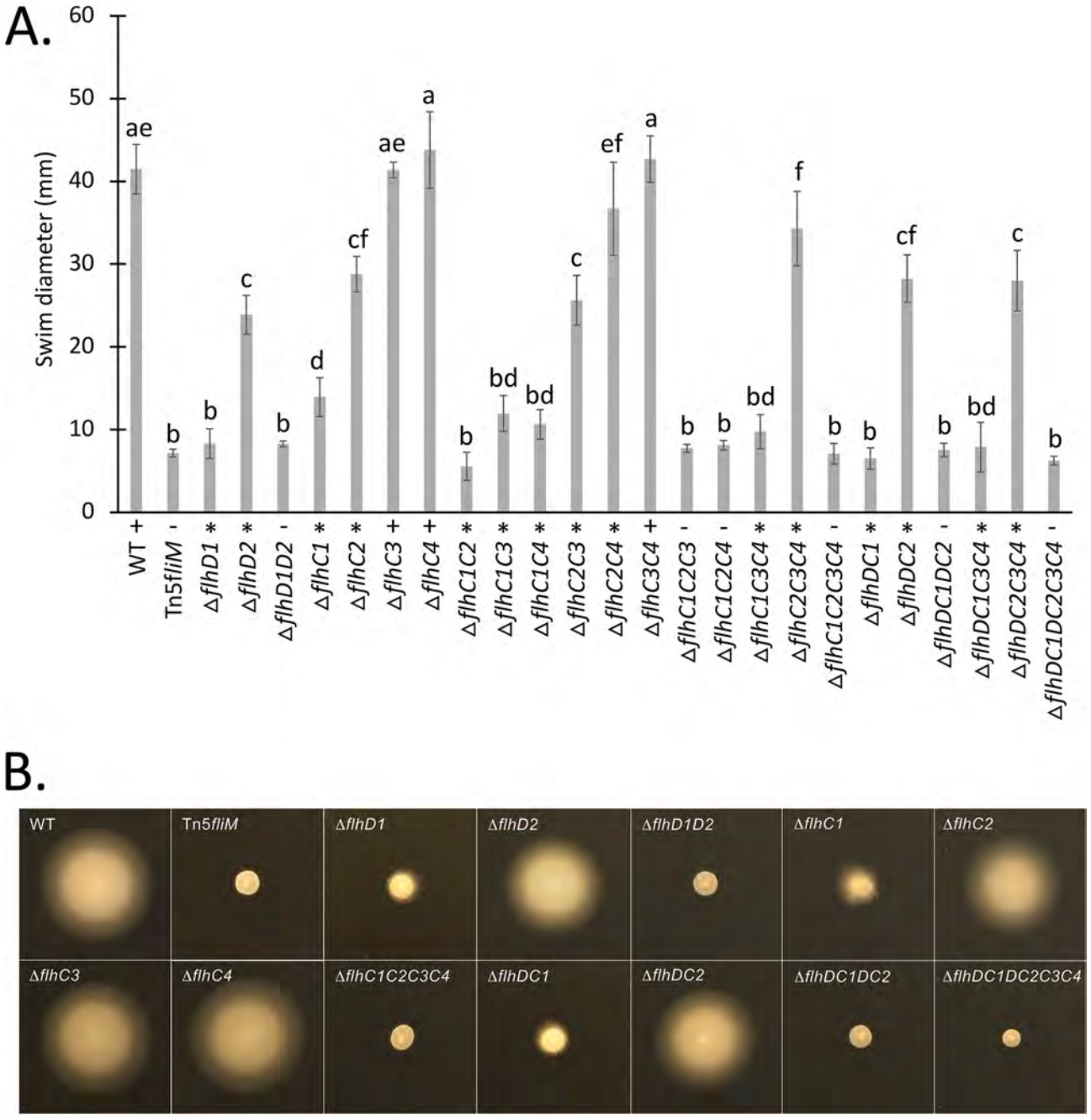

FIG 3 Swimming motility phenotype of the wild type, an *fliM* mutant, and *flhD* and *flhC* mutants of *P. unamae*. (A) Average swim halo diameters on motility agar after 48 h. Error bars represent standard deviation of results from at least three replicates. Organisms are noted as fully motile (+), partially motile (*), or nonmotile (−) based on swimming motility assay and visual observation. Different letters indicate a statistically significant difference as determined by one-way ANOVA and *post hoc* Tukey's test ($P < 0.05$). (B) Swimming halos on motility agar after 48 h.

ies. In-frame deletion mutants were generated for each copy as well as in combination. We performed swimming motility assays on soft agar with all the mutants and measured swim diameters in order to compare motilities between mutants and the wild type. Mutants lacking *flhD1* and/or *flhC1* showed severely impaired motility (Fig. 3A

FIG 2 Legend (Continued)
at the nodes. Sequences highlighted in the same color share genetic synteny. Letters following the species name indicate additional copies of the gene, with those labeled with b and c not having genetic synteny. IMG accession numbers precede the species name.

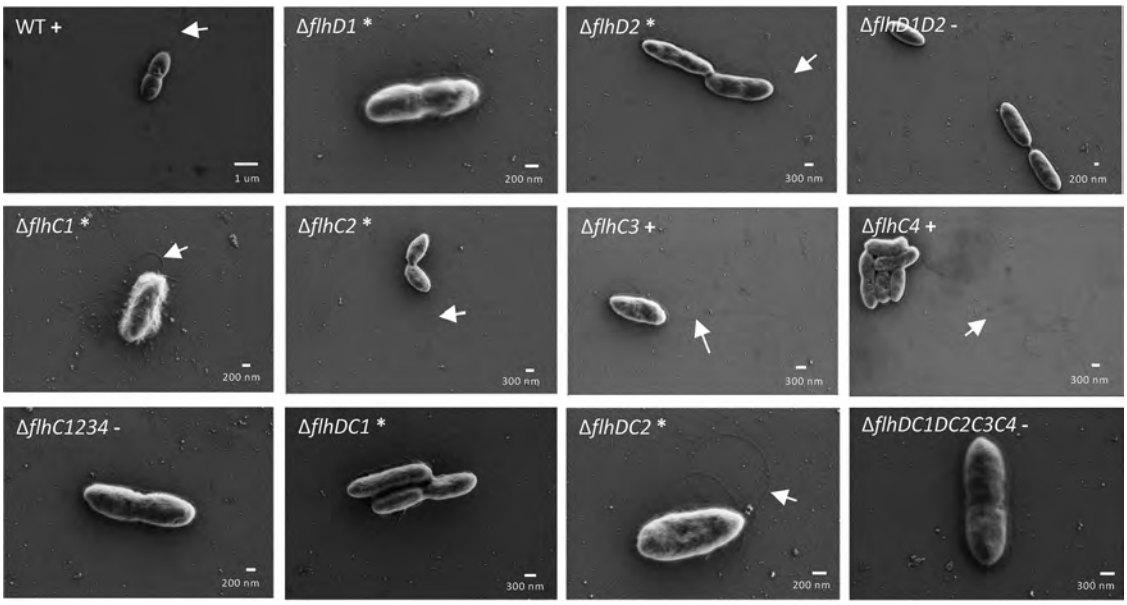

FIG 4 SEM depicting cellular morphology of *P. unamae* strains. The cells depicted represent typical morphology observed from biological replicates. Arrows indicate the presence of flagella in strains that are fully (+) or partially (*) motile in the swimming motility assay; − indicates nonmotile strains.

and B). The *flhD1* deletion mutant showed a more severe motility defect than the *flhC1* mutant. Deletion of *flhD2*, *flhC2*, or *flhDC2* also reduced motility, but to a lesser extent (Fig. 3A and B). Therefore, both operons regulate motility, but the *flhDC1* operon plays the primary role under the conditions tested. Deletion of both *flhD* genes completely abolished swimming, consistent with the known requirement of *flhD* in other organisms for motility. An *flhC1C2* double deletion mutant still visually retained some motility compared to the completely nonmotile *fliM* mutant (Fig. 3A and B), suggesting a role for *flhC3* and/or *flhC4*. However, motility did not seem to be affected in *flhC3* and *flhC4* mutants, although it was completely absent in the *flhC1C2C3* and *flhC1C2C4* triple mutants (Fig. 3A and B). Therefore, both *flhC3* and *flhC4* play a role, albeit minor, in motility. As expected, mutants lacking both *flhDC* operons or all *flhC* copies lacked motility, similar to the *fliM* mutant (Fig. 3A and B).

Scanning electron microscopy (SEM) analysis was performed to determine if mutant strains with defective motility are still able to produce flagella. Wild-type *P. unamae* strains are diplobacilli and lophotrichous, with multiple flagella expressed at one pole (Fig. 4). We were not able to detect the flagellar structure in mutant strains with the *flhD1* gene removed. These strains demonstrated severely impaired motility (Δ*flhD1* and Δ*flhDC1*) or were nonmotile (Δ*flhD1D2* and Δ*flhDC1DC2C3C4*) in the swimming motility assay (Fig. 3). In contrast, mutant strains with observable swim halos expressed polar flagella. Multiple flagella were detected in *flhC3* and *flhC4* mutants, which displayed swim halos similar to that of the wild type. The *flhC1* mutants, which had shown over 50% reduction in swim diameter, had flagella that appeared shorter and reduced in number compared to those of other mutants and the wild type (Fig. 4).

FlhD and FlhC are involved in biofilm formation. Previous studies have shown that in some organisms, the flagella and FlhDC can be involved in biofilm formation (47–49). To determine if mutations in *flhD* or *flhC* affect the ability of *P. unamae* to form biofilms on abiotic surfaces, biofilm assays on polyvinyl chloride (PVC) plates were performed. FliM mutants of *Listeria monocytogenes* lack flagella and are nonmotile (50). When we looked at a nonmotile *P. unamae fliM* mutant, we found no decrease in biofilm formation compared to the wild type (Fig. 5A and B). This indicates that motility is not required for biofilm formation in *P. unamae*. However, any mutant with deletion in *flhC1* or *flhD1* showed a severe inhibition of biofilm formation. Deletion of other *flhC* or

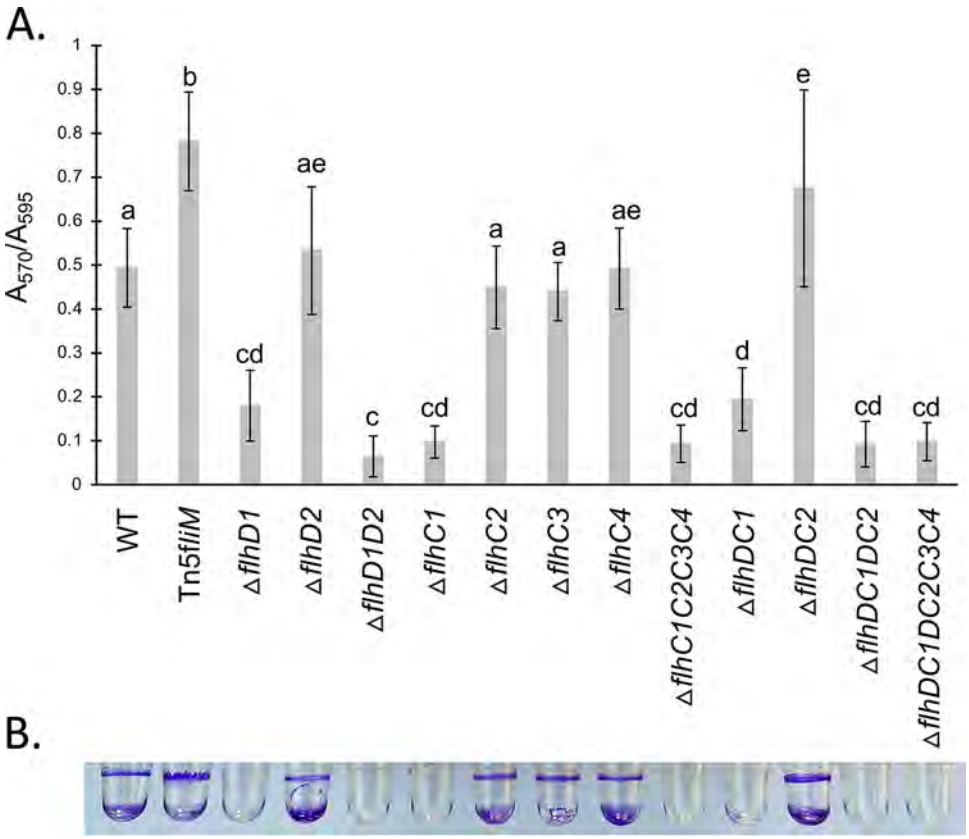

FIG 5 Effect of *flhD* and *flhC* deletions in *P. unamae* on biofilm formation. (A) Mean crystal violet absorbance in a 48-h biofilm. (B) Crystal violet staining of biofilm in 96-well plates. Error bars represent the standard deviation of results from three independent experiments, and different letters indicate a statistically significant difference between means as found by analysis with one-way ANOVA and *post hoc* Tukey's test ($P < 0.05$).

flhD copies did not cause a decrease (Fig. 5A and B). Therefore, the *flhDC1* operon in *P. unamae* plays a significant role in promoting biofilm formation on abiotic surfaces, but this appears independent of flagellar biosynthesis.

***flhDC1* is expressed at higher levels than other copies.** The different phenotypes of the *flhC* and *flhD* mutants, especially with regard to motility and biofilm formation, suggest that although there is some redundancy in the function of the genes, they are regulated differently. To address this question, transcriptional promoter-*lacZ* fusions were made to each of the four *flhC* genes, and β-galactosidase activity from each promoter was quantified after the organisms were grown in TY (tryptone yeast extract) medium to mid-exponential phase. Of the different copies, the *flhDC1* operon is expressed the highest, followed by the *flhDC2* operon. *flhC3* and *flhC4* showed little expression, if any (Fig. 6A). In case we had missed regulatory elements in the promoters fused to *lacZ*, we also carried out quantitative PCR (qPCR) for each of the loci. The qPCR results were consistent with that of the *lacZ* fusions, with *flhDC1* showing the greatest expression (Fig. 6B). As observed with the *lacZ* fusions, *flhDC2* expression was less than that of *flhDC1*, and *flhC3* and *flhC4* showed very little expression. The different levels of expression observed under this growth condition correlate with the varying degrees of effect on motility and biofilm formation when the respective genes were deleted.

***flhDC* regulates the expression of flagellar class II and class III genes.** *fliA* is a class II gene that is regulated by the master regulator of flagellar biosynthesis, $FlhD_4C_2$. It codes for a sigma factor (σ^{28}) and regulates the expression of *fliC*, a class III gene in the flagellum biosynthesis hierarchy that codes for the structural component, flagellin, to form the filament of the flagellum. To determine if the *flhD* and *flhC* genes in *P. unamae* behave the same as in *E. coli* and if they have different effects on the class II and

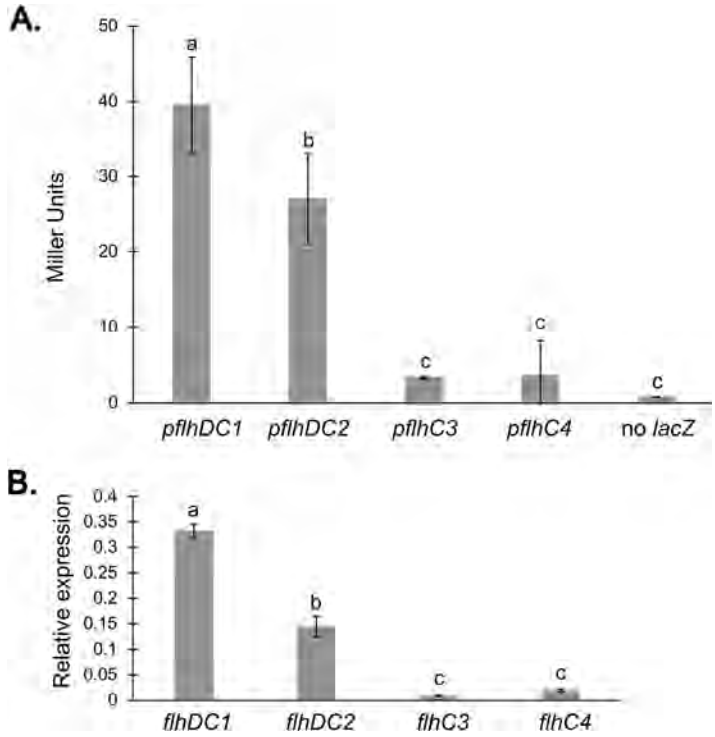

FIG 6 Levels of *flhD* and *flhC* gene expression in wild-type *P. unamae*. (A) β-Galactosidase activity (Miller units) produced by *lacZ* transcriptional fusions to promoters of the different copies of *flhDC* or *flhC*. Error bars represent the standard deviation of results from at least three independent replicates. (B) Levels of *flhD* and *flhC* transcripts relative to *rpoB* transcripts as determined by quantitative PCR. Error bars represent the standard error of results from three biological replicates, and different letters indicate a statistically significant difference found by analysis with one-way ANOVA and *post hoc* Tukey's test ($P < 0.05$).

III genes *fliA* and *fliC*, respectively, reverse transcription-quantitative PCR (RT-qPCR) was performed on mid-exponential growth cultures. Compared to wild-type *P. unamae*, *fliA* and *fliC* levels were reduced in a Δ*flhD2* mutant and barely detectable in a Δ*flhD1* mutant. The same decrease was observed when both *flhD* copies were deleted (Fig. 7A). Removal of the *flhC1* gene also greatly reduced *fliA* and *fliC* expression. The Δ*flhC2* mutant showed a trend in reduced expression of these genes, although the trend was not statistically significant, whereas Δ*flhC3* and Δ*flhC4* mutants did not show a change (Fig. 7B). As expected, removal of the *flhDC1* or both *flhDC* operons severely reduced flagellar gene expression and resulted in the same phenotype as Δ*flhD1*, Δ*flhD1D2*, and Δ*flhC1C2C3C4* mutants (Fig. 7C). Deletion of the *flhDC2* operon resulted in some reduction, with a greater impact on *fliC* expression. Consistent with the motility and SEM results, mutants with a motility defect and no detection of the flagellum in SEM also showed significant reduction in *fliA* and *fliC* expression. Thus, the *flhDC1* operon plays a significant role and the major role in regulating the expression of class II and III genes.

DISCUSSION

Regulation of the *flhD* and *flhC* genes and their gene products has been well studied in numerous bacteria because of the critical role the $FlhD_4C_2$ master regulator plays in flagellar biosynthesis and as a global regulator of many nonflagellar genes (51, 52). This regulation is known to occur at the transcriptional level as well as numerous ways posttranscriptionally, and involves only a single copy of *flhDC* in organisms where it has been studied. We report here the presence of an *flhDC* gene family in *P. unamae* and other *Paraburkholderia* species and show that these copies have various roles in regulating gene expression, motility, and biofilm formation in *P. unamae*. Our findings

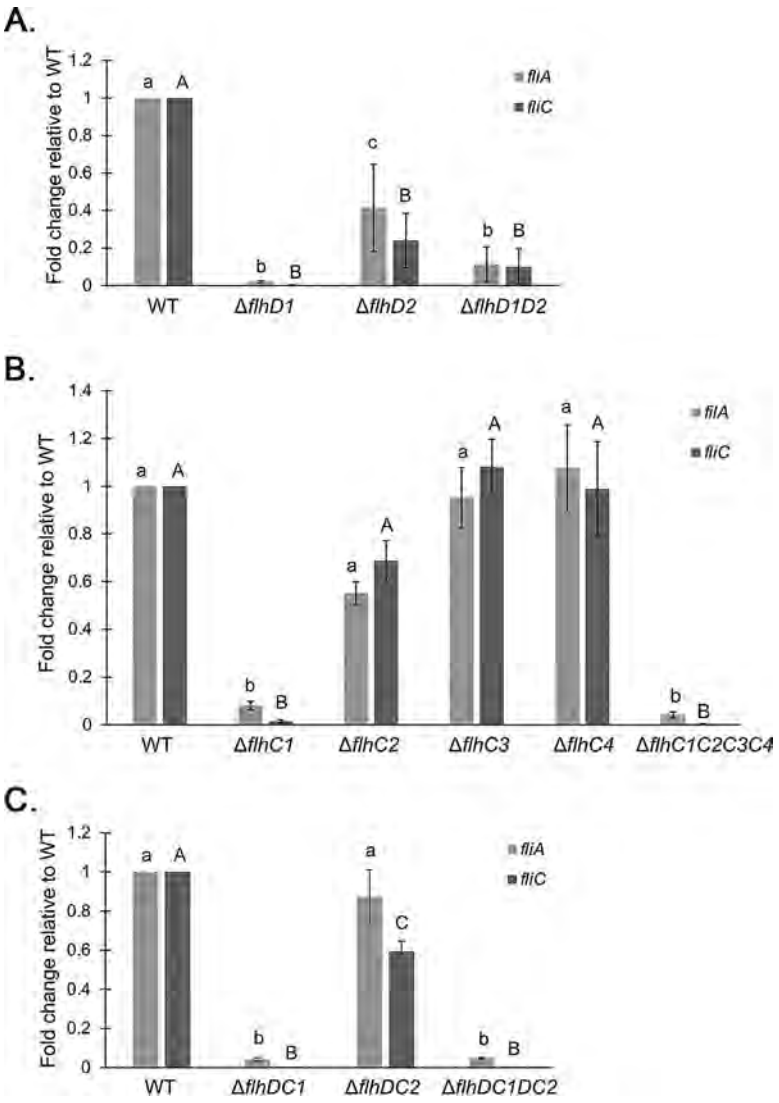

FIG 7 Expression levels of *fliA* and *fliC* in *P. unamae* mutant strains compared to the wild type as determined by RT-qPCR. Shown is the fold change relative to wild-type expression by (A) *flhD* deletion mutants, (B) *flhC* deletion mutants, and (C) *flhDC* operon mutants. Results were calculated from biological triplicates, and error bars represent standard errors. *rpoB* was used as the internal control. Different letters indicate a statistically significant difference using one-way ANOVA and *post hoc* Tukey's test ($P < 0.05$). Light gray bars show *fliA* expression, and dark gray bars show *fliC* expression.

indicate that in *Paraburkholderia*, an additional layer of *flhDC* regulation occurs by the use of multiple *flhDC* loci.

The presence of an *flhDC* gene family in *P. unamae* is in marked contrast to the single copy of these genes typically found outside the genus, including well-studied organisms such as *E. coli* and *S. enterica*. The presence of the *flhDC* gene family throughout *Paraburkholderia* suggests that these genes have evolved with the genus and confer some evolutionary adaptation. Gene duplication resulting in the expansion of protein families is associated with increased genome size (53). The expansion of protein families can result in novel functions, novel regulatory structures, and overall system robustness (54–56). Gene duplication can have an evolutionary advantage in the short term, allowing adaptation of organisms to various environmental conditions, such as thermal stress or starvation (57, 58). Prokaryotes cope with proteins that may no longer function appropriately (e.g., do not fold properly) in a changing environment

by having multiple copies of genes that can function under different environmental conditions (59). Members of *Burkholderia sensu lato* are known for their large genome sizes (3.75 to 11.5 Mbp) and genomic plasticity, which is associated with their metabolic diversity and ability to thrive in a wide range of environments (60, 61). Members of *Burkholderia* display increased numbers of genes related to virulence, including multiple copies of secretion systems, which are not found in *Paraburkholderia* (62). Our analysis of the *flhDC* operon shows that the genes have undergone frequent duplication and rearrangements, especially in *Paraburkholderia*, which has a minimum of two conserved *flhDC* loci, *flhDC1* and *flhDC2*, throughout the genus. The presence of the *flhDC1* locus in only *Paraburkholderia* and *Caballeronia* further differentiates these genera from the pathogenic *Burkholderia* and other members of *Burkholderia sensu lato*.

The intricacies of FlhDC regulation have not been studied in *Paraburkholderia*. However, analysis of numerous other bacterial species reveals that regulation of FlhDC is complex and occurs through a number of different mechanisms. Transcriptional regulation of *flhDC* is controlled by a variety of transcription factors in *E. coli*, conferring sensitivity of the operon to environmental factors (52). There are RNA binding proteins that bind to the 5′ untranslated region of *flhDC* mRNA, thus negatively regulating protein expression (23, 63). Additional posttranscriptional control by sRNAs and proteins occurs from direct interaction with FlhDC and affects downstream gene regulation (9, 22–24). Anti-FlhDC factors identified in *S. enterica* and *E. coli* inhibit FlhDC activity by direct protein-protein interactions in response to changing nutrient conditions, altering DNA binding or preventing recruitment of RNA polymerase (28–30, 64, 65). For example, the anti-$FlhD_4C_2$ factor FliT binds FlhC and prevents the FlhDC complex from binding to DNA (28). Furthermore, even between closely related organisms, there may be marked differences in regulation of FlhDC. For instance, *flhDC* expression in the enterobacteria *E. coli* and *S. enterica* shows contrasting responses to nutrient concentrations and temperature (30, 52). Complementation analyses of these organisms also showed that although FlhD is functionally similar, FlhC from *E. coli* is less effective at binding *S. enterica* FlhDC-regulated promoters (85). The intricacies of FlhDC regulation seem to reflect the critical role FlhDC has in mediating flagellum production as well as acting as a global regulator for numerous nonflagellar genes. Our finding of multiple functional copies of each *flhC* and *flhD* in *P. unamae* adds additional layers of regulatory complexity, as each copy may have its own transcriptional and posttranscriptional mechanisms in play.

Under the experimental conditions in this study, *flhDC1* had the highest expression and strongest effects on motility and biofilm formation. Deletion of *flhD1*, *flhC1*, or *flhDC1* in *P. unamae* dramatically reduced motility. Deletion of *flhC2*, *flhD2*, or *flhDC2*, the copies syntenous with *E. coli flhDC*, resulted in a less dramatic reduction in swim diameter. This is further supported by our expression analysis, which shows that *flhDC1* is expressed at the highest levels, followed by *flhDC2*. Likewise, deletion of *flhDC1* had the greatest impact on expression of the flagellar genes *fliA* and *fliC*. Thus, it appears that *flhDC1*, at least in *P. unamae* under the conditions we tested, has acquired the primary role in regulating flagellar production and, thus, swimming motility, although *flhDC2* plays a significant role as well. Similarly, *flhDC1* plays a greater role in biofilm formation in *P. unamae*, although we found motility itself is not a requirement in the formation of biofilms. Motility and biofilm formation are generally considered to be inversely regulated, as bacteria in the biofilm are nonmotile. In some organisms, flagella are important for the initiation of the biofilm, with flagellar mutants of some bacteria showing reduced biofilm production (66). FlhDC is known to influence biofilm formation (49). However, mutation of *flhD* or *flhC* resulted in an increase in biofilm formation in *S. enterica* (48), but a decrease in *E. coli* (47), indicating the varied role of *flhDC* in biofilm formation.

The role of *flhC3* and *flhC4* is more difficult to define. They also form distinct *flhC* phylogenetic clusters of orthologous loci, but are present in only a subset of *Paraburkholderia*, including *P. unamae*. Our expression assays using real-time qPCR and *lacZ* promoter

fusions both showed little, if any, expression of either gene. In addition, deletion of either gene alone did not impact *fliA* or *fliC* expression under the conditions studied. However, it was only when either *flhC3* or *flhC4* was deleted in the *flhC1C2* deletion background that motility was completely abolished, indicating that *flhC3* and *flhC4* must play some role in regulating motility. The *flhC3* gene, which resides alone, is in a chromosomal region near *ompR* and *envZ*. Interestingly, in organisms such as *E. coli* and *Yersinia enterocolitica*, OmpR regulates *flhDC* expression (17, 18). The role of FlhC4 appears complex. FlhC4 is double the length of the typical FlhC protein, and it is the C-terminal half that aligns with other FlhC sequences. The C-terminal half shows only 22.22% identity and 33.88% similarity to FlhC2, and it is not clear how this larger protein would interact in an FlhDC complex. FlhC4 lacks the conserved residues known to interact with FlhD but retains the four cysteines that form the zinc-binding site involved in DNA binding within promoters.

Studies have shown that the $FlhD_2$ homodimer can bind to DNA by itself; however, it is $FlhD_4$ that gives the $FlhD_4C_2$ complex specificity and guides it to the FlhDC-regulated flagellar genes (67). In *E. coli*, several residues in FlhD were found important for interaction with FlhC, and site-directed mutagenesis of those positions to alanine resulted in defective motility. For example, Cys-65 is important for the FlhD dimerization and is highly conserved in other *Gammaproteobacteria* as well as *Betaproteobacteria* (46). Our analysis found that these residues were also well conserved in *P. unamae*. Furthermore, residues important in *E. coli* for FlhC binding to FlhD were well conserved in three FlhC copies in *P. unamae*, FlhC1, FlhC2 and FlhC3, suggesting they all interact with FlhD. Although FlhC4 did not contain these residues, it does contain the four cysteine amino acids important for zinc binding in *E. coli* (13). The zinc-binding site is important for FlhC binding to the DNA; thus, it is possible that FlhC4 in *P. unamae* may retain the ability to bind the DNA.

In *E. coli*, which has a single copy of *flhDC*, the gene products homodimerize and form a heterohexameric complex (11, 13). We found *flhDC* in *P. unamae* to play a role in motility and gene expression similar to that seen in *E. coli* and other organisms; therefore, it seems likely an analogous FlhDC heteromeric complex is formed. Heteromeric transcription factors, in which the subunits are made up of different genes, are not common in bacteria. However, they account for as much as 10% of gene regulation in *E. coli*. Some, such as the integration host factor IHF-$\alpha\beta$ and RcsBA are global regulators (68, 69). Others, such as RelBE, can act as a RelB homodimer that derepresses transcription or a heteromeric RelBE complex that acts as a repressor, with complex formation regulated by the concentration of the different subunits to meet a particular stochiometric ratio (70). With multiple copies of the *flhD* and *flhC* genes in *Paraburkholderia*, a question raised is whether the FlhD and FlhC subunits from different operons might be interchangeable and act as an additional regulatory mechanism. It will be of interest to determine how the FlhDC complex is formed in *Paraburkholderia*. In particular, is the complex formed from the assembly of the FlhD and FlhC proteins encoded from the same operon or is the complex formed from a combination of proteins encoded by the other gene copies? Perhaps stochiometric regulation driving the assembly between different FlhD and FlhC subunits may contribute to the regulation of this complex and downstream effects. Future studies on protein interactions of the different copies of the *flhD* and *flhC* gene products will help to further elucidate their function. Future analyses can also reveal how these different copies respond to environmental signals and influence downstream gene expression. The presence of multiple genetically distinct *flhDC* operons throughout *Paraburkholderia* indicates an additional layer of complexity to what is known about the regulation of gene expression by FlhDC.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The strains used in the study are listed in Table S1 in the supplemental material. *E. coli* strains were maintained on Luria-Bertani agar at 37°C with 50 mg $liter^{-1}$ kanamycin. A rifampin-resistant strain of *P. unamae* MTI-641^T (42) was selected for and used for all subsequent studies. *P. unamae* strains were grown at 30°C on tryptone yeast extract (TY) (71) or yeast mannitol agar (YMA) (72) with 20 mg $liter^{-1}$ rifampin and 50 mg $liter^{-1}$ kanamycin as needed.

Generation of mutant and transcriptional *lacZ*-fusion strains. In-frame deletion mutants in each copy of *flhD* and *flhC* were made by allelic exchange using the vector pK18*mobsacB* (73). In brief, regions bordering the gene to be deleted were amplified and connected by fusion PCR, then cloned into the pK18*mobsacB* plasmid to generate the vector for allelic exchange. Constructs were sequenced to verify absence of mutation, then transformed into *E. coli* S17 using the TransformAid bacterial transformation kit (Thermo Scientific, Inc.) and selected on LB with kanamycin (50 mg liter^{-1}). The vector was introduced into *P. unamae* by a biparental mating strategy, selecting for kanamycin resistance for plasmid integration and then using sucrose sensitivity to select for subsequent excision of the plasmid. PCR on individual colonies was performed to identify strains where the gene of interest had been excised.

Multiple gene deletions were generated by using sequential rounds of this strategy. *P. unamae* mutants were conjugated with the S17 strain carrying the construct with the additional gene to be removed. The mutant strains generated and vectors used are listed in Table S1.

The *lacZ*-promoter fusions were generated by PCR amplification of the region encompassing the promoter for each *flhDC* or *flhC* copy using primers designed with 5′ end restriction sites. PCR products and the pVIK112 vector (74) were digested with the appropriate restriction enzymes and ligated together. Plasmids were transformed into *E. coli* DH5α λpir and confirmed by sequencing. The pVIK112-promoter-*lacZ* fusion constructs were introduced into *P. unamae* MTI-641^{T} by triparental mating with *E. coli* (pRK2013) (75). Selection for integration of the constructs by Campbell insertion was done using rifampin and kanamycin. All primers used are listed in Table S2 in the supplemental material.

Swimming motility assay. We found that carrying out soft agar motility assays using bacteria from fresh colonies was more consistent than with an inoculum from liquid cultures. Therefore, cultures of bacteria were grown on TY agar at 30°C for 48 h and immediately used in motility assays. An inoculating needle was used to obtain bacteria from the center of the colony and stabbed into TY plates containing 0.25% agar (Sigma). Motility plates were incubated at 30°C, and halo diameters were measured after 48 h. Three independent trials were performed, with each trial analyzing at least three individual colonies of each strain.

Biofilm assay. *P. unamae* strains were grown for 24 h in liquid LB medium without NaCl, as this medium was found to promote the best biofilm formation by *P. unamae* on 96-well polyvinyl chloride (PVC) plates (A. Escamilla and M. R. Lum, unpublished data). Cultures were then diluted to an optical density at 600 nm (OD_{600}) of 0.05, and 150 μl was placed into each well of a column in a 96-well PVC microplate, covered with breathable film (AeraSeal; Genesee Scientific), and incubated at 30°C for 48 h in a container with a wet paper towel to maintain humidity. Equivalent growth between wells was checked by measurement at OD_{595} using a microplate reader. Medium was aspirated, and plates were gently rinsed with water, stained with 0.1% crystal violet for 20 min, rinsed three times with water, and dried, and the bound crystal violet was resuspended in 175 μl of 95% ethanol. Crystal violet was quantified using a microplate reader at OD_{570}. Average absorbance was obtained for at least six wells for each trial, and three independent trials were carried out. Wells were imaged prior to resuspension in ethanol for qualitative demonstration of biofilm formation.

β-Galactosidase assay. Preliminary tests indicated a higher level of expression of the *flhDC* genes during the mid-exponential phase (S. N.-M. Thai, unpublished data). *P. unamae* strains carrying the promoter-*lacZ* fusions were grown overnight at 30°C with shaking in liquid TY, diluted to an OD_{600} of 0.02, and then grown to mid-exponential phase (OD_{600} of 0.8 to 0.9). β-Galactosidase assays were performed, and Miller units were calculated as described by Miller et al. (76).

Scanning electron microscopy. *P. unamae* strains were streaked onto TY agar with rifampin and incubated for 48 to 72 h at 30°C, and motility assays were then performed with the fresh cultures. A sterile spatula was used to excise a small piece of agar containing bacterial growth at the edge of the swim diameter from biological replicate samples and submerged in ultrafiltered water for 5 min to allow the cells to swim out of the agar. Cells were then fixed in 4% paraformaldehyde (PFA) and transferred to a cleaned silicon wafer (Ted Pella, Inc.) coated with marine mussel glue (Cell-Tak; BD Bioscience) to attach the cells. Preparation of wafers and imaging were performed as outlined by Billi et al. (77). The wafers were then washed with ultrafiltered water, mounted onto aluminum stubs, and allowed to air dry before imaging. The wafers were scanned for the presence of cells, and their morphology was then imaged with a field emission scanning electron microscope (Supra VP-40; Zeiss) at a voltage of 1 kV with a 3-mm working distance. Multiple images were taken from each sample that represents the typical cell morphology of the strains observed after scanning the majority of the wafer surface.

RNA extraction and RT-qPCR. To extract total RNA from respective test organisms for RT-qPCR, overnight cell cultures were diluted to OD_{600} of 0.01 and then grown for 10 h in TY with rifampin. Samples were prepared in biological triplicates. RNAprotect cell reagent (Qiagen, Inc.) was added to the cell cultures per the manufacturer's recommendation to provide immediate stabilization of the RNA prior to cell collection. Approximately 10^9 cells were pelleted and quickly frozen and stored at −80°C until RNA extraction was performed using the RNeasy minikit (Qiagen, Inc.) per the manufacturer's suggestion. Purified RNA samples (6.5 μg each) were treated with the Turbo DNA-free kit (Invitrogen, Inc.), and 240 ng of the DNase-treated total RNA was synthesized to cDNA with the iScript cDNA synthesis kit (Bio-Rad, Inc). For the qPCR, this cDNA reaction mixture was used at 1/20 per reaction with iTaq Universal SYBR green Supermix (Bio-Rad, Inc.) on a 7500 fast real-time PCR system (Applied Biosystems, Inc.) for 40 cycles with an annealing temperature of 60°C. Each qPCR mixture contained 500 nM forward and reverse primers. The primers used in the qPCR are listed in Table S2 in the supplemental material. The melt curves showed a single peak for the first derivative of fluorescence versus temperature, indicating a single product. The primer efficiencies for targeted genes and housekeeping genes were similar. Analysis was done using the relative standard curve method (ABI User Bulletin 2). Serial dilutions of

P. unamae genomic DNA were used to generate a standard curve for each primer pair to calculate the relative amount of transcript in each sample. Levels of gene expression were normalized to *rpoB*.

Sequence analysis and phylogenetic tree generation. The *E. coli* FlhC1 and FlhD1 sequences were used to carry out a BLASTP search (78) in IMG (79) (to identify FlhC and FlhD sequences in *P. unamae*, using default parameters with the BLOSUM62 matrix and a cutoff of 1e−05. Each copy of *P. unamae* FlhC and FlhD was then used for BLASTP searches of the IMG bacterial genomes and NCBI databases. The corresponding DNA and protein sequences were downloaded for selected type strains. MUSCLE was used to construct multiple sequence alignments (80) and MEGAX (81) to generate phylogenetic trees using the maximum likelihood (ML) algorithm. Bootstrap analysis with 1,000 resamplings was used to provide statistical support for the trees (82). The best model to use for each tree was determined by MEGAX, with the *flhC* phylogeny done using the general time-reversible (GTR) model (83) with a gamma distribution (G) and the *flhD* phylogeny using the Tamura 3-parameter (T92) model (84) with gamma-distributed selection and invariant sites (G+I). A MUSCLE alignment of the protein sequences was used as input into the SIAS server to obtain the sequence identities and similarities (http://imed.med.ucm.es/Tools/sias.html). FlhC and FlhD alignments were shaded using Boxshade 3.21 (https://embnet.vital-it.ch). To determine whether the genes were in operons, DNA regions encompassing the *P. unamae flhDC1* and *flhDC2* regions were entered into the operon predicting tool Operon-Mapper (44).

Statistical analysis. Statistical analyses were carried out using Excel or SPSS and one-way analysis of variance (ANOVA) with *post hoc* Tukey's test.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 0.6 MB.

ACKNOWLEDGMENTS

This research was supported in part by the William F. McLaughlin Chair in Biology and a Kadner-Pitts Research Grant awarded to M.R.L. and by a U.S. Department of Education HSI-STEM (P031C110180) grant awarded to Glendale Community College (GCC). It was also partially supported by the GCC Foundation and Associated Students of GCC. M.O. received support through an American Society of Microbiology (ASM) Undergraduate Research Fellowship.

We thank members of Dianne Newman's lab at Caltech, Megan Bergkessel and Melanie Spero, for valuable feedback and technical support and C. Thomas Payne for comments on the manuscript. We thank Joseph Soriano for technical assistance and Ann M. Hirsch for providing the original *P. unamae* MTI-641 strain.

REFERENCES

1. de Weert S, Vermeiren H, Mulders IHM, Kuiper I, Hendrickx N, Bloemberg GV, Vanderleyden J, De Mot R, Lugtenberg BJJ. 2002. Flagella-driven chemotaxis towards exudate components is an important trait for tomato toot colonization by *Pseudomonas fluorescens*. Mol Plant Microbe Interact 15:1173–1180. https://doi.org/10.1094/MPMI.2002.15.11.1173.
2. Scharf BE, Hynes MF, Alexandre GM. 2016. Chemotaxis signaling systems in model beneficial plant-bacteria associations. Plant Mol Biol 90:549–559. https://doi.org/10.1007/s11103-016-0432-4.
3. Zhang N, Wang D, Liu Y, Li S, Shen Q, Zhang R. 2014. Effects of different plant root exudates and their organic acid components on chemotaxis, biofilm formation and colonization by beneficial rhizosphere-associated bacterial strains. Plant Soil 374:689–700. https://doi.org/10.1007/s11104-013-1915-6.
4. Nougayrede J-P, Fernandes PJ, Donnenberg MS. 2003. Adhesion of enteropathogenic *Escherichia coli* to host cells. Cell Microbiol 5:359–372. https://doi.org/10.1046/j.1462-5822.2003.00281.x.
5. Gardel CL, Mekalanos JJ. 1996. Alterations in *Vibrio cholerae* motility phenotypes correlate with changes in virulence factor expression. Infect Immun 64:2246–2255. https://doi.org/10.1128/iai.64.6.2246-2255.1996.
6. Eaton KA, Suerbaum S, Josenhans C, Krakowka S. 1996. Colonization of gnotobiotic piglets by *Helicobacter pylori* deficient in two flagellin genes. Infect Immun 64:2445–2448. https://doi.org/10.1128/iai.64.7.2445-2448.1996.
7. Chevance FFV, Hughes KT. 2008. Coordinating assembly of a bacterial macromolecular machine. Nat Rev Microbiol 6:455–465. https://doi.org/10.1038/nrmicro1887.
8. Fitzgerald DM, Bonocora RP, Wade JT. 2014. Comprehensive mapping of the *Escherichia coli* flagellar regulatory network. PLoS Genet 10:e1005456. https://doi.org/10.1371/journal.pgen.1004649.
9. Yakhnin AV, Baker CS, Vakulskas CA, Yakhnin H, Berezin I, Romeo T, Babitzke P. 2013. CsrA activates *flhDC* expression by protecting *flhDC* mRNA from RNase E-mediated cleavage: mechanism of CsrA-mediated activation of *flhDC*. Mol Microbiol 87:851–866. https://doi.org/10.1111/mmi.12136.
10. Macnab RM. 1996. Flagella and motility, p 123–145. *In* Neidhardt FC, Curtiss R, III, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umbarger HE (ed), *Escherichia coli* and *Salmonella:* cellular and molecular biology, 2nd ed. ASM Press, Washington, DC.
11. Liu X, Matsumura P. 1994. The FlhD/FlhC complex, a transcriptional activator of the *Escherichia coli* flagellar class II Operons. J Bacteriol 176:7345–7351. https://doi.org/10.1128/jb.176.23.7345-7351.1994.
12. Yanagihara S, Iyoda S, Ohnishi K, Iino T, Kutsukake K. 1999. Structure and transcriptional control of the flagellar master operon of *Salmonella typhimurium*. Genes Genet Syst 74:105–111. https://doi.org/10.1266/ggs.74.105.
13. Wang S, Fleming RT, Westbrook EM, Matsumura P, McKay DB. 2006. Structure of the *Escherichia coli* FlhDC complex, a prokaryotic heteromeric regulator of transcription. J Mol Biol 355:798–808. https://doi.org/10.1016/j.jmb.2005.11.020.
14. Komeda Y. 1986. Transcriptional control of flagellar genes in *Escherichia coli* K-12. J Bacteriol 168:1315–1318. https://doi.org/10.1128/jb.168.3.1315-1318.1986.
15. Ohnishi K, Kutsukake K, Suzuki H, Iino T. 1990. Gene *fliA* encodes an alternative sigma factor specific for flagellar operons in *Salmonella typhimurium*. Mol Gen Genet 221:139–147. https://doi.org/10.1007/BF00261713.
16. Prüss BM, Campbell JW, Van Dyk TK, Zhu C, Kogan Y, Matsumura P. 2003. FlhD/FlhC is a regulator of anaerobic respiration and the Entner-Doudoroff pathway through induction of the methyl-accepting chemotaxis

protein Aer. J Bacteriol 185:534–543. https://doi.org/10.1128/JB.185.2.534-543.2003.
17. Shin S, Park C. 1995. Modulation of flagellar expression in *Escherichia coli* by acetyl phosphate and the osmoregulator OmpR. J Bacteriol 177:4696–4702. https://doi.org/10.1128/jb.177.16.4696-4702.1995.
18. Raczkowska A, Skorek K, Bielecki J, Brzostek K. 2011. OmpR controls *Yersinia enterocolitica* motility by positive regulation of *flhDC* expression. Antonie Van Leeuwenhoek 99:381–394. https://doi.org/10.1007/s10482-010-9503-8.
19. Li C, Louise CJ, Shi W, Adler J. 1993. Adverse conditions which cause lack of flagella in *Escherichia coli*. J Bacteriol 175:2229–2235. https://doi.org/10.1128/jb.175.8.2229-2235.1993.
20. Shi W, Li C, Louise CJ, Adler J. 1993. Mechanism of adverse conditions causing lack of flagella in *Escherichia coli*. J Bacteriol 175:2236–2240. https://doi.org/10.1128/jb.175.8.2236-2240.1993.
21. Soutourina OA, Krin E, Laurent-Winter C, Hommais F, Danchin A, Bertin PN. 2002. Regulation of bacterial motility in response to low pH in *Escherichia coli*: the role of H-NS protein. Microbiology 148:1543–1551. https://doi.org/10.1099/00221287-148-5-1543.
22. Mika F, Hengge R. 2013. Small regulatory RNAs in the control of motility and biofilm formation in *E. coli* and *Salmonella*. Int J Mol Sci 14:4560–4579. https://doi.org/10.3390/ijms14034560.
23. De Lay N, Gottesman S. 2012. A complex network of small non-coding RNAs regulate motility in *Escherichia coli*: sRNA regulation of motility in *Escherichia coli*. Mol Microbiol 86:524–538. https://doi.org/10.1111/j.1365-2958.2012.08209.x.
24. Guillier M, Gottesman S. 2006. Remodeling of the *Escherichia coli* outer membrane by two small regulatory RNAs. Mol Microbiol 59:231–247. https://doi.org/10.1111/j.1365-2958.2005.04929.x.
25. Schachterle JK, Zeng Q, Sundin GW. 2019. Three Hfq-dependent small RNAs regulate flagellar motility in the fire blight pathogen *Erwinia amylovora*. Mol Microbiol 111:1476–1492. https://doi.org/10.1111/mmi.14232.
26. Wei BL, Brun-Zinkernagel A-M, Simecka JW, Prüß BM, Babitzke P, Romeo T. 2001. Positive regulation of motility and *flhDC* expression by the RNA-binding protein CsrA of *Escherichia coli*: effects of CsrA on *flhDC* expression. Mol Microbiol 40:245–256. https://doi.org/10.1046/j.1365-2958.2001.02380.x.
27. Chatterjee A, Cui Y, Chakrabarty P, Chatterjee AK. 2010. Regulation of motility in *Erwinia carotovora* subsp. *carotovora*: quorum-sensing signal controls FlhDC, the global regulator of flagellar and exoprotein genes, by modulating the production of RsmA, an RNA-binding protein. Mol Plant Microbe Interact 23:1316–1323. https://doi.org/10.1094/MPMI-01-10-0017.
28. Yamamoto S, Kutsukake K. 2006. FliT acts as an anti-FlhD2C2 factor in the transcriptional control of the flagellar regulon in *Salmonella enterica Serovar Typhimurium*. J Bacteriol 188:6703–6708. https://doi.org/10.1128/JB.00799-06.
29. Saini S, Brown JD, Aldridge PD, Rao CV. 2008. FliZ is a posttranslational activator of FlhD4C2-dependent flagellar gene expression. J Bacteriol 190:4979–4988. https://doi.org/10.1128/JB.01996-07.
30. Wada T, Hatamoto Y, Kutsukake K. 2012. Functional and expressional analyses of the anti-FlhD4C2 factor gene *ydiV* in *Escherichia coli*. Microbiology (Reading) 158:1533–1542. https://doi.org/10.1099/mic.0.056036-0.
31. Kutsukake K, Ohya Y, Iino T. 1990. Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*. J Bacteriol 172:741–747. https://doi.org/10.1128/jb.172.2.741-747.1990.
32. Tans-Kersten J, Brown D, Allen C. 2004. Swimming motility, a virulence trait of *Ralstonia solanacearum*, is regulated by FlhDC and the plant host environment. Mol Plant Microbe Interact 17:686–695. https://doi.org/10.1094/MPMI.2004.17.6.686.
33. Ritchings BW, Almira EC, Lory S, Ramphal R. 1995. Cloning and phenotypic characterization of *fleS* and *fleR*, new response regulators of *Pseudomonas aeruginosa* which regulate motility and adhesion to mucin. Infect Immun 63:4868–4876. https://doi.org/10.1128/iai.63.12.4868-4876.1995.
34. Prouty MG, Correa NE, Klose KE. 2001. The novel sigma54- and sigma28-dependent flagellar gene transcription hierarchy of *Vibrio cholerae*. Mol Microbiol 39:1595–1609. https://doi.org/10.1046/j.1365-2958.2001.02348.x.
35. Kim J, Kang Y, Choi O, Jeong Y, Jeong J-E, Lim JY, Kim M, Moon JS, Suga H, Hwang I. 2007. Regulation of polar flagellum genes is mediated by quorum sensing and FlhDC in *Burkholderia glumae*: regulation of flagellum genes in *Burkholderia glumae*. Mol Microbiol 64:165–179. https://doi.org/10.1111/j.1365-2958.2007.05646.x.
36. Sawana A, Adeolu M, Gupta RS. 2014. Molecular signatures and phylogenomic analysis of the genus *Burkholderia*: proposal for division of this genus into the emended genus *Burkholderia* containing pathogenic organisms and a new genus *Paraburkholderia* gen. nov. harboring environmental species. Front Genet 5:429. https://doi.org/10.3389/fgene.2014.00429.
37. Beukes CW, Palmer M, Manyaka P, Chan WY, Avontuur JR, van Zyl E, Huntemann M, Clum A, Pillay M, Palaniappan K, Varghese N, Mikhailova N, Stamatis D, Reddy TBK, Daum C, Shapiro N, Markowitz V, Ivanova N, Kyrpides N, Woyke T, Blom J, Whitman WB, Venter SN, Steenkamp ET. 2017. Genome data provides high support for generic boundaries in *Burkholderia* sensu lato. Front Microbiol 8:1154. https://doi.org/10.3389/fmicb.2017.01154.
38. Estrada-de los Santos P, Palmer M, Chávez-Ramírez B, Beukes C, Steenkamp E, Briscoe L, Khan N, Maluk M, Lafos M, Humm E, Arrabit M, Crook M, Gross E, Simon M, dos Reis Junior F, Whitman W, Shapiro N, Poole P, Hirsch A, Venter S, James E. 2018. Whole genome analyses suggests that *Burkholderia* sensu lato contains two additional novel genera (*Mycetohabitans* gen. nov., and *Trinickia* gen. nov.): implications for the evolution of diazotrophy and nodulation in the *Burkholderiaceae*. Genes (Basel) 9:389. https://doi.org/10.3390/genes9080389.
39. Perin L, Martínez-Aguilar L, Castro-González R, Estrada-de los Santos P, Cabellos-Avelar T, Guedes HV, Reis VM, Caballero-Mellado J. 2006. Diazotrophic *Burkholderia* species associated with field-grown maize and sugarcane. Appl Environ Microbiol 72:3103–3110. https://doi.org/10.1128/AEM.72.5.3103-3110.2006.
40. Estrada-De Los Santos P, Bustillos-Cristales R, Caballero-Mellado J. 2001. *Burkholderia*, a genus rich in plant-associated nitrogen fixers with wide environmental and geographic distribution. Appl Environ Microbiol 67:2790–2798. https://doi.org/10.1128/AEM.67.6.2790-2798.2001.
41. Caballero-Mellado J, Onofre-Lemus J, Estrada-de los Santos P, Martínez-Aguilar L. 2007. The tomato rhizosphere, an environment rich in nitrogen-fixing *Burkholderia* species with capabilities of interest for agriculture and bioremediation. Appl Environ Microbiol 73:5308–5319. https://doi.org/10.1128/AEM.00324-07.
42. Caballero-Mellado J, Martínez-Aguilar L, Paredes-Valdez G, los Santos PE. 2004. *Burkholderia unamae* sp. nov., an N_2-fixing rhizospheric and endophytic species. Int J Syst Evol Microbiol 54:1165–1172. https://doi.org/10.1099/ijs.0.02951-0.
43. Onofre-Lemus J, Hernández-Lucas I, Girard L, Caballero-Mellado J. 2009. ACC (1-aminocyclopropane-1-carboxylate) deaminase activity, a widespread trait in *Burkholderia* species, and its growth-promoting effect on tomato plants. Appl Environ Microbiol 75:6581–6590. https://doi.org/10.1128/AEM.01240-09.
44. Taboada B, Estrada K, Ciria R, Merino E. 2018. Operon-mapper: a web server for precise operon identification in bacterial and archaeal genomes. Bioinformatics 34:4118–4120. https://doi.org/10.1093/bioinformatics/bty496.
45. Dobritsa AP, Samadpour M. 2016. Transfer of eleven species of the genus *Burkholderia* to the genus *Paraburkholderia* and proposal of *Caballeronia* gen. nov. to accommodate twelve species of the genera *Burkholderia* and *Paraburkholderia*. Int J Syst Evol Microbiol 66:2836–2846. https://doi.org/10.1099/ijsem.0.001065.
46. Campos A, Matsumura P. 2001. Extensive alanine scanning reveals protein-protein and protein-DNA interaction surfaces in the global regulator FlhD from *Escherichia coli*. Mol Microbiol 39:581–594. https://doi.org/10.1046/j.1365-2958.2001.02248.x.
47. Pratt LA, Kolter R. 1998. Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. Mol Microbiol 30:285–293. https://doi.org/10.1046/j.1365-2958.1998.01061.x.
48. Teplitski M, Al-Agely A, Ahmer BMM. 2006. Contribution of the SirA regulon to biofilm formation in *Salmonella enterica serovar Typhimurium*. Microbiology (Reading) 152:3411–3424. https://doi.org/10.1099/mic.0.29118-0.
49. Sule P, Horne SM, Logue CM, Prüss BM. 2011. Regulation of cell division, biofilm formation, and virulence by FlhC in *Escherichia coli* O157:H7 grown on meat. Appl Environ Microbiol 77:3653–3662. https://doi.org/10.1128/AEM.00069-11.
50. Cheng C, Wang H, Ma T, Han X, Yang Y, Sun J, Chen Z, Yu H, Hang Y, Liu F, Fang W, Jiang L, Cai C, Song H. 2018. Flagellar basal body structural proteins FlhB, FliM, and FliY are required for flagellar-associated protein expression in *Listeria monocytogenes*. Front Microbiol 9:208. https://doi.org/10.3389/fmicb.2018.00208.
51. Komeda Y, Kutsukake K, Iino T. 1980. Definition of additional flagellar genes in *Escherichia coli* K12. Genetics 94:277–290. https://doi.org/10.1093/genetics/94.2.277.
52. Soutourina O, Kolb A, Krin E, Laurent-Winter C, Rimsky S, Danchin A, Bertin P. 1999. Multiple control of flagellum biosynthesis in *Escherichia coli*: role of H-NS protein and the cyclic AMP-catabolite activator protein

complex in transcription of the *flhDC* master operon. J Bacteriol 181: 7500–7508. https://doi.org/10.1128/JB.181.24.7500-7508.1999.
53. Ohno S. 1970. Evolution by gene duplication. Springer, Berlin, Germany.
54. Pasek S, Risler J-L, Brézellec P. 2006. The role of domain redundancy in genetic robustness against null mutations. J Mol Biol 362:184–191. https://doi.org/10.1016/j.jmb.2006.07.033.
55. Pereira-Leal JB, Levy ED, Kamp C, Teichmann SA. 2007. Evolution of protein complexes by duplication of homomeric interactions. Genome Biol 8: R51. https://doi.org/10.1186/gb-2007-8-4-r51.
56. Wagner A. 2008. Gene duplications, robustness and evolutionary innovations. Bioessays 30:367–373. https://doi.org/10.1002/bies.20728.
57. Riehle MM, Bennett AF, Long AD. 2001. Genetic architecture of thermal adaptation in Escherichia coli. Proc Natl Acad Sci U S A 98:525–530. https://doi.org/10.1073/pnas.021448998.
58. Yamanaka K, Fang L, Inouye M. 1998. The CspA family in *Escherichia coli*: multiple gene duplication for stress adaptation. Mol Microbiol 27:247–255. https://doi.org/10.1046/j.1365-2958.1998.00683.x.
59. Sanchez-Perez G, Mira A, Nyiro G, Pasić L, Rodriguez-Valera F. 2008. Adapting to environmental changes using specialized paralogs. Trends Genet 24: 154–158. https://doi.org/10.1016/j.tig.2008.01.002.
60. Compant S, Nowak J, Coenye T, Clément C, Ait Barka E. 2008. Diversity and occurrence of *Burkholderia* spp. in the natural environment. FEMS Microbiol Rev 32:607–626. https://doi.org/10.1111/j.1574-6976.2008.00113.x.
61. Coenye T, Vandamme P. 2003. Diversity and significance of *Burkholderia* species occupying diverse ecological niches. Environ Microbiol 5: 719–729. https://doi.org/10.1046/j.1462-2920.2003.00471.x.
62. Angus AA, Agapakis CM, Fong S, Yerrapragada S, Estrada-de los Santos P, Yang P, Song N, Kano S, Caballero-Mellado J, de Faria SM, Dakora FD, Weinstock G, Hirsch AM. 2014. Plant-associated symbiotic *Burkholderia* species lack hallmark strategies required in mammalian pathogenesis. PLoS One 9:e83779. https://doi.org/10.1371/journal.pone.0083779.
63. Thomason MK, Fontaine F, De Lay N, Storz G. 2012. A small RNA that regulates motility and biofilm formation in response to changes in nutrient availability in *Escherichia coli*: sRNA regulator of motility and biofilm formation. Mol Microbiol 84:17–35. https://doi.org/10.1111/j.1365-2958.2012.07965.x.
64. Li B, Yue Y, Yuan Z, Zhang F, Li P, Song N, Lin W, Liu Y, Yang Y, Li Z, Gu L. 2017. *Salmonella* STM1697 coordinates flagella biogenesis and virulence by restricting flagellar master protein $FlhD_4C_2$ from recruiting RNA polymerase. Nucleic Acids Res 45:9976–9989. https://doi.org/10.1093/nar/gkx656.
65. Chatterjee A, Cui Y, Chatterjee AK. 2009. RsmC of *Erwinia carotovora* subsp. *carotovora* negatively controls motility, extracellular protein production, and virulence by binding FlhD and modulating transcriptional activity of the master regulator. J Bacteriol 191:4582–4593. https://doi.org/10.1128/JB.00154-09.
66. Guttenplan SB, Kearns DB. 2013. Regulation of flagellar motility during biofilm formation. FEMS Microbiol Rev 37:849–871. https://doi.org/10.1111/1574-6976.12018.
67. Claret L, Hughes C. 2000. Functions of the subunits in the $FlhD_2C_2$ transcriptional master regulator of bacterial flagellum biogenesis and swarming. J Mol Biol 303:467–478. https://doi.org/10.1006/jmbi.2000.4149.
68. Fang FC, Rimsky S. 2008. New insights into transcriptional regulation by H-NS. Curr Opin Microbiol 11:113–120. https://doi.org/10.1016/j.mib.2008.02.011.
69. Francez-Charlot A, Laugel B, Van Gemert A, Dubarry N, Wiorowski F, Castanié-Cornet M-P, Gutierrez C, Cam K. 2003. RcsCDB His-Asp phosphorelay system negatively regulates the *flhDC* operon in *Escherichia coli*: regulation of *flhDC* by RcsB. Mol Microbiol 49:823–832. https://doi.org/10.1046/j.1365-2958.2003.03601.x.
70. Cataudella I, Sneppen K, Gerdes K, Mitarai N. 2013. Conditional cooperativity of toxin-antitoxin regulation can mediate bistability between growth and dormancy. PLoS Comput Biol 9:e1003174. https://doi.org/10.1371/journal.pcbi.1003174.
71. Beringer JE. 1974. R factor transfer in *Rhizobium leguminosarum*. Microbiology 84:188–198. https://doi.org/10.1099/00221287-84-1-188.
72. Vincent JM. 1970. A manual for the practical study of root-nodule bacteria. International Biological Programme, handbook no. 15. Blackwell Scientific, Oxford, United Kingdom.
73. Schäfer A, Tauch A, Jäger W, Kalinowski J, Thierbach G, Pühler A. 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene 145:69–73. https://doi.org/10.1016/0378-1119(94)90324-7.
74. Kalogeraki VS, Winans SC. 1997. Suicide plasmids containing promoterless reporter genes can simultaneously disrupt and create fusions to target genes of diverse bacteria. Gene 188:69–75. https://doi.org/10.1016/S0378-1119(96)00778-0.
75. Figurski DH, Helinski DR. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc Natl Acad Sci U S A 76:1648–1652. https://doi.org/10.1073/pnas.76.4.1648.
76. Miller JH. 1992. A short course in bacterial genetics: a laboratory manual and handbook for *Escherichia coli* and related bacteria. Cold Spring Harbor Laboratory Press, Plainview, NY.
77. Billi F, Benya P, Kavanaugh A, Adams J, McKellop H, Ebramzadeh E. 2012. The John Charnley Award: an accurate and extremely sensitive method to separate, display, and characterize wear debris part 2: metal and ceramic particles. Clin Orthop 470:339–350. https://doi.org/10.1007/s11999-011-2058-9.
78. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389.
79. Markowitz VM, Mavromatis K, Ivanova NN, Chen I-MA, Chu K, Kyrpides NC. 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics 25:2271–2278. https://doi.org/10.1093/bioinformatics/btp393.
80. Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113. https://doi.org/10.1186/1471-2105-5-113.
81. Stecher G, Tamura K, Kumar S. 2020. Molecular evolutionary genetics analysis (MEGA) for macOS. Mol Biol Evol 37:1237–1239. https://doi.org/10.1093/molbev/msz312.
82. Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x.
83. Nei M, Kumar S. 2000. Molecular evolution and phylogenetics. Oxford University Press, Oxford, United Kingdom.
84. Tamura K. 1992. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. Mol Bio Evol 9:678–687.
85. Albanna A, Sim M, Hoskisson PA, Gillespie C, Rao CV, Aldridge PD. 2018. Driving the expression of the *Salmonella enterica* sv Typhimurium flagellum using *flhDC* from *Escherichia coli* results in key regulatory and cellular differences. Sci Rep 8:16705. https://doi.org/10.1038/s41598-018-35005-2.