Research in Microbiology : Research Article. Revised manuscript

(Two tables, six figures; Suppl. Materials incl. three tables and two figures.)

Biofilm formation and cellulose expression by *Bordetella avium* 197N, the causative agent of bordetellosis in birds and an opportunistic respiratory pathogen in humans

Kimberley McLaughlin[1*], Ayorinde O. Folorunso[1*], Yusuf Y. Deeni[1], Dona Foster[2], Oksana Gorbatiuk[3], Simona M. Hapca[1], Corinna Immoor[1], Anna Koza[1,4], Ibrahim U. Mohammed[1], Olena Moshynets[3], Sergii Rogalsky[5], Kamil Zawadzki[1], and Andrew J. Spiers[1†]

[1] School of Engineering and Technology, Abertay University, Bell Street, Dundee DD1 1HG, UK.

[2] Department of Biological and Life Sciences, Oxford Brookes University, Gipsy Lane, Headington, Oxford OX3 0BP, UK.

[3] Institute of Molecular Biology and Genetics of the National Academy of Sciences of Ukraine, 150 Zabolotnoho Street, Kiev 03680, Ukraine.

[4] Current address : Technical University of Denmark, Novo Nordisk Foundation Center for Biosustainability, Kemitorvet, Building 220, 2800 Kgs. Lyngby, Denmark.

[5] Institute of Bioorganic Chemistry and Petrochemistry of the National Academy of Sciences of Ukraine, 50 Kharkivske Schose, Kiev 02160, Ukraine.

[*] Contributed equally to this work.

† Correspondence : a.spiers@abertay.ac.uk.

Abstract (194 / 200 words)

Although bacterial cellulose synthase (*bcs*) operons are widespread within the *Proteobacteria* phylum, subunits required for the partial-acetylation of the polymer appear to be restricted to a few γ-group soil, plant-associated and phytopathogenic pseudomonads including *P. fluorescens* SBW25 and several *P. syringae* pathovars. However, a *bcs* operon with acetylation subunits has also been annotated in the unrelated β-group respiratory pathogen, *Bordetella avium* 197N. Our comparison of subunit protein sequences and GC content analyses confirms the close similarity between the *Ba.* 197N and pseudomonad operons and suggests that in both cases the cellulose synthase and acetylation subunits were acquired as a single unit. Using static liquid microcosms, we can confirm that *Ba.* 197N expresses low levels of cellulose in air-liquid interface biofilms and that biofilm strength and attachment levels could be increased by elevating c-*di*-GMP levels like the pseudomonads, but cellulose was not required for biofilm-formation itself. The finding that *Ba.* 197N is capable of producing cellulose from a highly-conserved but relatively uncommon *bcs* operon raises the question of what functional role this modified polymer plays during the infection of the upper respiratory tract or survival between hosts, and what environmental signals control its production.

Keywords: Air-liquid interface, Biofilm, *Bordetella avium*, c-*di*-GMP, Cellulose, Microcosm.

1. Introduction

Bordetella avium is the causative agent of bordetellosis (tracheobronchitis), a highly contagious upper respiratory disease of domesticated and wild birds [1 - 3], as well an opportunistic human pathogen that may also be associated with cystic fibrosis [4 - 6]. It is phylogenetically distinct from *B. bronchiseptica*, *B. parapertussis,* and *B. pertussis* which are respiratory pathogens of mammals, though *B. avium* expresses similar virulence factors and also produces biofilms. However, whilst *Bordetella* polysaccharide (Bps) has been identified as the primary matrix component in *Bb.* RB50 and *B. pertussis* 536 biofilms [7, 8], it has not been reported for *B. avium* biofilms nor has a *bps*-like operon been found in the genome of the virulent turkey isolate *Ba.* 197N [9 - 11]. This raises the question of what extracellular polymeric substance (EPS) or substances are used as the biofilm matrix during host infection or the colonisation of other environments during transmission between hosts.

We have noted that the *Ba.* 197N genome [11] has been annotated with a potential bacterial cellulose synthase (*bcs*) operon (genes BAV2632 – 2623) (Fig. 1). *Bcs* operons have been identified in the genomes of a wide range of *Proteobacteria*, but cellulose expression *per se* has been reported for relatively few strains, including human gastrointestinal commensals and pathogens, and soil and plant-associated pseudomonads [12 - 14]. Although the functional role and fitness value of cellulose is poorly understood [13 - 15], it is associated with biofilm-formation in the intestinal epithelium and invasion of epithelial cells and macrophages [14], and provides a fitness advantage in soils and plant environments where it may help reduce water stress [16 - 19]. More generally cellulose may promote survival in natural environments during transmission between hosts [12].

We are interested in determining whether *Ba.* 197N is capable of expressing cellulose as part of a greater understanding of the functional role of cellulose in a range of

environments and life strategies. Unusually, the *Ba.* 197N *bcs* operon subunits show close amino acid sequence homology with those found in pseudomonads, particularly the soil and plant-associated *P. fluorescens* SBW25 (annotated as *wssA-J*, PFLU0300 – 0309 [20], Fig. 1), despite the fact that the *Bordetellae* and *Pseudomonas* are distantly-related genera.

The *Ba.* 197N and *Pf.* SBW25 *bcs* operons are further linked by the inclusion of cellulose acetylation-associated subunits originally described for *Pf.* SBW25 (WssF-I) which affects colony morphology and air-liquid (A-L) interface biofilms that develop in static microcosms [20, 21]. These subunits are rarely seen, but are present in a number of plant pathogens including *P. syringae* pv *tomato* DC3000, but not, for example, the closely-related soil-associated *P. putida* KT2440 [13, 21, 22] (Fig. 1). Cellulose is also expressed by *Pp.* KT2440 and *Ps.* DC3000 in similar conditions [19, 23], and static microcosms have been used to investigate biofilm-formation more widely amongst the pseudomonads [23, 24] where biofilms can be quantitatively measured by combined growth, strength and attachment assays [24, 25].

Ba. 197N has also been reported to produce limited biofilms at the meniscus of static liquids [26, 27]. We therefore decided to use a static microcosm approach to characterise *Ba.* 197N biofilm-formation more fully under a range of growth conditions. Although low levels of cellulose expression was observed in wild-type *Ba.* 197N biofilm samples and colonies, we constructed a cellulose-deficient (CD) mutant to demonstrate that cellulose was not an essential matrix component for biofilm-formation, suggesting that *Ba.* 197N utilises another but unidentified EPS for this purpose.

2. Materials and methods

2.1 Bioinformatics

Bacterial cellulose synthase subunit homologues were identified in *Proteobacteria* complete genomes using NCBI TBLASTN and *Pf.* SBW25 WssA – J protein query sequences (Accession numbers AAL71841 – AAL71850) following the selection criteria described in the Supplementary Methods. A WspR homologue was identified in the *Ba.* 197N genome using *Pf.* SBW25 WspR (AAL71852). Phylogenetic trees were produced using the NCBI COBALT multiple sequence alignment tool and the Neighbour Joining (NJ) method (see Suppl. Methods). Gene and whole genomic GC content data were obtained from the *Pseudomonas* Genome Database (www.pseudomonas.com) and NCBI Genbank. DNA dot plots were produced using the YASS Genomic Similarity Search Tool (bioinfo.lifl.fr).

2.2 Bacteria and plasmids

Bacteria and plasmids used in this work are listed in Table 1. All strains were kept at -80°C for long-term storage, and stocks produced by adding 15% (v/v) glycerol to over-night cultures. *Ba.* 197N was acquired directly from the American Type Culture Collection (ATCC BAA-1003). Antibiotic susceptibility was assessed using antibiotic disks (MAST, UK). Kanamycin (Km) sensitivity [10] was confirmed using LB plates containing 100 $\mu g.ml^{-1}$ Km incubated at 28°C for 48 hr.

2.3 Culturing conditions

Bacteria were cultured at 20, 28, 37, and 42°C using Brain Heart Infusion (BHI; Oxoid, UK), King's B (KB; 20 g Proteose Peptone No. 3 (BD Biosciences, UK), 10 g glycerol, 5 g K_2HPO_4, 1.5 g $MgSO_4$ per litre) and Luria-Bertani (LB; 10 g NaCl, 10 g Tryptone (Oxoid), and 5g Yeast extract (Merck, UK) per litre) media with 1.5% (w/v) Agar Technical (Oxoid) for plates and antibiotics added as required. Over-night shaken cultures

were used to provide inocula for experiments. Microcosms were 30 ml glass tubes containing 6 ml growth media [23]. Swimming motility was assessed using soft-agar LB plates (see Suppl. Methods).

2.4 Biofilm-formation in static microcosms

Microcosms inoculated with 100 µl aliquots of culture and incubated statically were inspected daily for up to 5 days for signs of A-L interface biofilm-formation, and then by pouring out into Petri dishes to assess strength and attachment [23]. Biofilms were quantitatively characterised using the combined biofilm assay in which replicate microcosms (n = 8) were tested for biofilm strength (grams) using small glass balls, biofilm attachment to the tube walls by staining with Crystal violet (A_{570}), and total growth by OD_{600} measurements [24, 25]. These assays were conducted for a number of strains, media, and temperatures, and data sets analysed using a General Linear Model (GLM) approach.

2.5 Cellulose expression

Cellulose was assessed using Congo red plates and by fluorescent microscopy. Congo red plates were BHI, KB, and LB (without NaCl) plus 0.001% (w/v) Congo red [21], and were drop-inoculated with 5 µl aliquots of culture and incubated for 2 – 3 days. Images were taken using a Nikon D3200 DSLR camera. Samples of colonies and biofilm material were stained with 10 $\mu g.ml^{-1}$ Calcofluor for 1 h before viewing at 10x – 40x magnification using a Leica DMR fluorescent microscope and imaging with an AxioCam MRc digital camera [21, 28]. Semi-purified biofilm matrix samples were investigated by ELISA using a cellulose binding domain–containing (SPA-CBD) fusion protein [29] as described in the Suppl. Methods. Biofilm samples were also investigated by Fourier transform infrared

spectroscopy (FTIR) and visualised by Scanning electron microscopy (SEM) (see Suppl. Methods).

2.6 Construction of pAS296 and the cellulose-deficient (CD) mutant

The suicide plasmid pAS296 was designed to disrupt *bcs* operons by homologous recombination of a *Pf.* SBW25 mini-transposon *wssB*::IS-Ω-Km/hah (Km^R) cassette (the insertion is located 1,693 bp from the start of the 2,219 bp *wssB* gene). This was integrated into the *Ba.* 197N chromosome following kanamycin selection to produce the cellulose-deficient (CD) mutant (see Suppl. Methods).

2.7 Statistical analyses and modelling

Quantitative assays were undertaken with replicates, and means with standard errors (SE) are shown where appropriate. Data were analysed using JMP v12 (SAS Institute Inc., USA) statistical software. A General Linear Model (GLM) approach was used to investigate the significance of strain, test media, growth, and attachment (effects) on biofilm strength (response) using the combined biofilm assay dataset. Attachment levels were also modelled as response. Significant effects were examined by LSMeans Differences Student's t-test and Tukey HSD tests, and associations examined by pairwise correlations. Differences between means were also tested by ANOVA, t-tests and Turkey-Kramer HSD tests.

3. Results and Discussion

3.1 Bacterial cellulose synthase (bcs) operons within the Proteobacteria.

We identified over 100 bacterial whole genome sequences likely to contain fully-functional *bcs* operons based on amino acid sequence homologies to the core cellulose

synthase subunits of *Pf.* SBW25 (Fig. 1 and data not shown), and it is noteworthy that such a poorly-reported or tested phenotype such as the ability to express cellulose is so commonly annotated amongst the *Proteobacteria*. Within this phylum, species from one α-group, two β-group and ten γ-group genera were found to have convincing *bcs* operons (see Suppl. Table S1). As expected, the organisation of these operons, in terms of gene order and orientation, was highly variable [14], though phylogenetic-based clustering at the genus or family levels were seen in Neighbour-Joining (NJ) trees constructed using core subunit sequences. An example of one NJ tree generated using WssB homologues (see Suppl. Table S2) and rooted using α-group *Proteobacteria* is given in Fig. 2. The phylogenetic-based clustering of homologues seen here suggests that most sequence variation in core subunits is probably acquired through vertical transmission. However, the close branching of distantly-related genera such as the *Bordetellae* and *Pseudomonas* in these trees suggests that *bcs* operons may have been horizontally transmitted at an early stage in the radiation of the *Proteobacteria*.

We note the increasing number of plant pathogens including *Burkholderia* (β-group), *Dickeya*, and *Pseudomonas* (γ-group) spp., and animal pathogens including *Bordetella* (β-group), *Escherichia*, *Klebsiella*, *Salmonella*, and *Shigella* (γ-group) spp., encode *bcs* operons and are likely to express cellulose (Suppl. Table S1). Although these pathogens have diverse hosts and survive in a wide range of environments, cellulose might play a common role in single-cell attachment, micro-colony and biofilm-formation, and protection against predation, disturbance, or stress [13, 14].

3.2 The bcs operon in Bordetella avium 197N.

Following our survey of *bcs* operons within the *Proteobacteria*, we shifted our attention to the operon annotated in the *Ba.* 197N genome. The link between this avian

and opportunistic human respiratory pathogen [9 - 11] and the soil and plant-associated pseudomonads is intriguing, because of the high levels of *bcs* subunit homology with *Pf.* SBW25 and the presence of the rarely-seen cellulose acetylation-associated subunits apparently restricted to *Pf.* SBW25 and a number of related plant pathogens including *Ps.* DC3000 [21, 22].

The *Ba.* 197N *bcs* operon is also notable, as in NJ trees the WssBCDE homologues from *Ba.* 197N, *Pf.* SBW25 and *Ps.* DC3000 were always clustered together but distant from the *Pp.* KT2440 homologues (see also [22]). A representative NJ tree for WssB homologues is shown in Fig. 2 and a comparison of WssA – I homologues provided in Table 2.

However, the *Ba.* 197N *bcs* operon differs significantly from the *Pf.* SBW25, *Pp.* KT2440, and *Ps.* DC3000 operons by the duplication of WssC (annotated as WssC1 and WssC2), which challenges our presumption that this bacterium could express cellulose. Although the duplication is apparent in an alignment of the *Ba.* 197N and *Pf.* SBW25 genomic sequences, WssC1 and WssC2 are not direct copies of one another and share only 47% identity at the protein level (97% coverage, 0.0 E value) (see Suppl. Fig. S1). Furthermore, WssC1 and WssC2 branch separately in a NJ tree containing *Pf.* SBW25, *Pp.* KT2440 and *Ps.* DC3000 WssC homologues (Suppl. Fig. S1). The alignment of *Ba.* 197N and *Pf.* SBW25 sequences also indicates that the two genomes do not share significant levels of homology beyond the boundaries of the *bcs* operons, suggesting that one or the other acquired the operon by a limited lateral gene transfer event.

A comparison of the GC content of the *bcs* genes indicates that the *Ba.* 197N, *Pf.* SBW25, *Pp.* KT2440 and *Ps.* DC3000 operons are more similar to one another than they are to the *Pp.* KT2440 operon ($\alpha = 0.05$) (Fig. 3). The mean GC contents of the *bcs* genes of *Pf.* SBW25, *Pp.* KT2440 and *Ps.* DC3000 were also significantly different to their whole

genome GC content ($p \leq 0.0003$), suggesting that these bacteria had acquired the *bcs* operons relatively recently and before much amelioration of divergent sequences could occur. In contrast, no significant difference was observed between the *Ba.* 197N *bcs* genes and whole genome GC content ($p = 0.2033$), suggesting that the transfer event occurred much earlier in this bacterium or that the ancestral *bcs* donor was more similar to the *Bordetellae* than to the pseudomonads. As the GC content of intergenic regions and coding sequences are frequently correlated in bacterial genomes [30], a direct comparison between the GC content of the *bcs* genes with the GC content of all other chromosomally-encoded genes is unlikely to produce a substantially different result. However, we have not yet investigated the regions surrounding each of the *bcs* operons which may show significant differences to the rest of the chromosomes.

Finally, no significant differences were observed between the GC content of core synthase and acetylation-associated genes within operons ($p \geq 0.5085$), and no significant correlations found between GC content and gene order ($p > 0.05$). This suggests that the acetylation-associated genes had been transferred and maintained with the core synthase genes, rather than as subsequent and independently acquired functions.

3.3 Ba. 197N biofilm-formation in static microcosms

Given the close protein homologies and conserved operon structures seen between the *Ba.* 197N *bcs* operon and the *Pf.* SBW25, *Pp*. KT2440 and *Ps.* DC3000 *bcs* operons, we were interested to determine whether *Ba.* 197N expressed cellulose despite the WssC duplication, and whether it also played a role in A-L interface biofilm-formation as it does for some pseudomonads. As *Ba.* 197N has been reported to form limited biofilms at the meniscus of static liquids [26, 27], we decided to use the static microcosm approach we had previously developed to investigate A-L interface biofilm-formation by *Pf.* SBW25 and

other pseudomonads [13, 23, 24]. In these, competition for O_2 which is otherwise growth-limiting in the liquid column drives biofilm-formation at the liquid surface [31], even in the absence of host-derived factors or environmental signals that might normally regulate such activity including the expression of biofilm-associated EPS.

In order to determine whether *Ba.* 197N could form biofilms in static microcosms, a range of media and incubation temperatures (BHI, KB and LB; 20 – 42°C) were tested, as although *Ba.* 197N is routinely grown in BHI at 37°C, phenotypic differences have been noted under other conditions [9, 10, 32, 33]. We found clear visual evidence of growth (i.e. culture turbidity) and A-L interface biofilm-formation, including culture turbidity, films extending across the A-L interface, and sunken debris, within three days of incubation under all conditions tested (Fig. 4). The biofilms formed in KB microcosms were particularly obvious and robust, and were characterised as physically cohesive (PC)-class / air-liquid-solid surface (A-L-S) interface-type biofilms [23, 24]. In contrast, BHI biofilms were almost transparent with white flecks of material, and LB biofilms were very thin and fragile.

We recovered biofilm material from microcosms and spread onto plates to investigate colony morphologies. In *Pf.* SBW25 populations, radiation (i.e. mutation) results in significant numbers of biofilm-forming Wrinkly Spreaders identifiable by an altered colony morphology [34]. However, only wild-type colony morphologies were observed from 3 day-old *Ba.* 197N biofilm samples, and no other signs of radiation were seen in KB microcosms incubated for up to 15 days (e.g. in siderophore production, colouration, motility, etc.). This suggests that biofilm-formation we have observed in experimental static microcosms is the result of a normal physiological response by *Ba.* 197N rather than the activity of biofilm-producing mutants. Further testing using modified media containing avian tissue-derived signal compounds, cell cultures, tracheal rings or

lung tissue will be required to determine whether biofilm-formation is a behaviour more specifically associated with host infection or the colonisation of water, plants, and soils during transmission between hosts.

3.4 Quantitative characterisation of the Ba. 197N biofilm

In order to place *Ba.* 197N A-L interface biofilm-formation in static microcosms into context, we compared *Ba.* 197N with wild-type *Pf.* SBW25 and the *Pf.* SBW25 Wrinkly Spreader (WS) mutant using a range of media and a quantitative combined biofilm assay measuring microcosm growth, biofilm strength and attachment levels [24]. Wild-type *Pf.* SBW25 was included as it produces a poorly-attached, fragile biofilm referred to as the VM biofilm [35], whilst the WS produces a robust, well-attached biofilm in KB microcosms [20, 21, 28, 34].

We modelled biofilm strength using a GLM approach, and both strain and medium were found to be significant ($p \leq 0.0001$, see Suppl. Table S3, Model 1), and representative data for KB microcosms are shown in Fig. 5. Although the VM and WS biofilms were significantly different from one another, the *Ba.* 197N biofilms were of intermediate strength (Fig. 5B) and could not be differentiated from either the VM or WS biofilms ($\alpha = 0.05$). The strongest *Ba.* 197N biofilms were produced in KB microcosms, and across all three strains, KB biofilms could be differentiated from BHI and LB biofilms ($\alpha = 0.05$) which suggests that biofilm formation was influenced by media composition. Over-all growth and biofilm attachment levels were not found to be significant effects on biofilm strength ($p > 0.05$) across the three strains tested in these microcosms. However, in previous work the VM biofilm has been shown to have significantly lower attachment levels than the WS biofilm [35]. Finally, we also modelled biofilm attachment levels as biomass at the meniscus is often used as a measure of biofilm growth in microtiter plate-

based assays [36]. In this, the strain effect was found to be weakly insignificant ($p = 0.0501$, Suppl. Table S3, Model 2), with the *Ba.* 197N biofilm differentiated from the VM biofilm, but neither from the WS biofilm ($\alpha = 0.05$).

3.5 Cellulose expression by Ba. 197N

We used the cellulose-binding dye Calcofluor and fluorescent microscopy to demonstrate that cellulose fibres were present in *Ba.* 197N colonies and biofilm samples (Fig. 6A). However relatively little cellulose was observed compared to that reported for VM and WS biofilms [21, 35], and none was observed in samples from a cellulose-deficient (CD) mutant we produced by disrupting the *Ba.* 197N *bcs* operon with a *Pf.* SBW25 *wssB*::IS-**Ω**-Km/hah cassette. Additional evidence to support the presence of cellulose at low levels in biofilms was obtained by ELISA using a cellulose binding domain–containing (SPA-CBD) fusion protein [29]. This showed a 2.5x increase in cellulose levels in semi-purified biofilm matrix samples samples compare to the CD mutant, though this assay was un-replicated and may need further testing. However, cellulose levels were insufficient for detection by FTIR (see Suppl. Fig. S2). Despite this, significant differences in biofilm strength and attachment levels were found between the *Ba.* 197N and CD mutant biofilms ($\alpha = 0.05$) (Fig. 5). These changes may not necessarily be due to the loss of cellulose in the biofilm *per se*, but might result from the loss of the cellulose synthase complex which effects other regulatory systems involved in the expression of biofilm matrix components and attachment factors.

We suggest that the low levels of cellulose seen in *Ba.* 197N biofilms produced under the conditions tested here may be the result of either the unusual *wssC* duplication, which might reduce *wss* operon transcription or cellulose synthase complex assembly, or the result of leaky transcriptional or synthase regulation in the absence of the normal

318 signals required to express cellulose. It is possible that host-specific factors are required to
319 induce cellulose expression to higher levels during invasion of avian tissues, or that other
320 environmental signals are required for expression during transmission between hosts for
321 survival in water, soils or plants. Future analysis will require transcriptional analyses of the
322 *wss* operon plus quantitative measurements of cellulose production to better understand
323 how and when cellulose production is controlled by this bacterium.

324 It remains unclear what EPS is being utilised by *Ba.* 197N as the primary biofilm
325 matrix in these microcosms, as *Ba.* 197N lacks a *bps*-like operon required to express Bps
326 polymer produced by other *Bordetellae* [7, 8, 11]. SEM imaging of wild-type *Ba.* 197N and
327 CD mutant biofilm samples appear very similar (Fig. 6B) and are like those produced by *B.*
328 *bronchiseptica* RB50 and *B. pertussis* Bp536 for which Bps is required for biofilm-
329 formation [26, 27]. *Ba.* 197N biofilms may utilise a similar polysaccharide, eDNA, or a
330 proteinaceous attachment factor or adhesin as the primary biofilm matrix component, but
331 further biochemical and genetic analyses will be required to identify the nature of this EPS.

332

333 *3.6 Involvement of ci-di-GMP in Ba. 197N in colony morphology and biofilms*

334 Biofilm formation and cellulose expression in *Pf.* SBW25 is primarily controlled by c-*di*-
335 GMP levels regulated by WspR and other diguanylate cyclases (DGC) [34]. DGCs are
336 found in other pseudomonads and *Bordetellae*, and we have noted a close homologue in
337 *Ba.* 197N (WP_012416712; 336 aa, with 63% ID and an E value of 1e-149). In order to
338 determine whether c-*di*-GMP levels may also play a role in *Ba.* 197N biofilm-formation or
339 cellulose expression, we have tested the effect of the over-expression of a constitutively-
340 active and dominant WspR19 mutant that induces a phenotype indistinguishable from the
341 WS in *Pf.* SBW25 [28, 37]. As WspR19 is functional in other pseudomonads and in more
342 distantly related bacteria such as *Caulobacter crescentus* CB15 [23, 38, 39], it might also

have an impact on the phenotype of *Ba.* 197N regardless of the activity of WP_012416712.

We first tested whether WspR19 affected *Ba.* 197N swimming motility [10], as the expression of a similar DGC in *B. bronchiseptica* 9.73 was found to repress swimming motility through the elevation of c-*di*-GMP levels [40]. WspR19 was found to partially repress swimming to 11 – 18%% of the levels shown by *Ba.* 197N carrying the control plasmid ($p \leq 0.0001$). We then investigated whether WspR and WspR mutants affected colony morphology, including the up-take of Congo red stain used to assess cellulose and attachment-factor expression [20, 21, 28, 35]. WspR19 was found to have a consistent impact on *Ba.* 197N colony morphology, resulting in larger and slightly rougher colonies with higher levels of Congo red staining compared to wild-type *Ba.* 197N and the CD mutant (Fig. 6C). In comparison, *Ba.* 197N colonies expressing the wild-type WspR protein (WspR12) [28, 41] were only slightly stained by Congo red, and colonies expressing the dominant-negative WspR9 mutant [41] were indistinguishable from *Ba.* 197N, CD mutant, and the plasmid control colonies. Although more cellulose could be observed by fluorescent microscopy in colony material produced by *Ba.* 197N expressing WspR19 compared to either *Ba.* 197N with the control plasmid or the CD mutant (data not shown), it is unclear whether the changes in colony morphology are the result of minor increases in cellulose levels or changes in other EPS or attachment factors which also bind Congo red.

The expression of WspR mutants in *Ba.* 197N also affected biofilm strength, with strain, medium and growth all found to be significant effects ($p \leq 0.0001$, Suppl. Table S3, Model 3). In particular, strains could be differentiated into three groups with Ba. 197N carrying the control plasmid or expressing WspR9 having the least effect on strength, WspR12 having an intermediate effect, and WspR19 having the greatest effect on strength ($\alpha = 0.05$). The strongest biofilms were produced in KB microcosms by *Ba.* 197N

368 expressing WspR19 which were significantly stronger than the plasmid control (2.5 x, p =
369 0.0353), though growth was significantly reduced (0.6 x, p = 0.0002). This suggests that
370 WspR19 has a negative pleiotropic effect when expressed in *Ba.* 197N, as it does in *Pf.*
371 SBW25 [37]. Although strain was weakly insignificant (p < 0.0727) when attachment was
372 modelled (Suppl. Table S3, Model 4), it seems likely that biofilm-formation by *Ba.* 197N in
373 static microcosms under the conditions tested here is regulated to some degree by c-*di*-
374 GMP levels.
375

376 **4. Conclusions**

377 Bacterial cellulose synthase *bcs* operons are common amongst the *Proteobacteria*,
378 though acetylation-associated genes are rare and have only been identified in a select
379 group of pseudomonads and a few other species including *Bordetella avium* 197N. In this
380 work, we confirm that *Ba.* 197N is able to express cellulose at low levels, though it was not
381 necessary for the formation of A-L interface biofilms in static microcosms, and further
382 investigation is required to identify the primary matrix components utilised by *Ba.* 197N
383 under the conditions tested here and to identify the environmental signals and regulatory
384 mechanisms which control expression under natural conditions. Although cellulose might
385 be used during the infection of the upper respiratory tract of hosts, it may also be important
386 for survival between hosts in water, soil or on plant surfaces, and this might explain why
387 *bcs* operons are carried by environmental bacteria as well as pathogenic and commensal
388 intestinal bacteria.
389

390 **Conflict of Interests**

391 The Authors declare that there is no conflict of interests regarding the publication of
392 this paper.

Author's contributions

KMcL and AOF contributed equally to this work as first co-authors. KMcL, OM and AJS designed the experiments. AOF, KZ, DF and AJS carried out the bioinformatics analyses. KMcL, CI, AK, IUM, OM and SR carried out the experiments. KMcL, SMH and AJS analysed the data. AJS and OM prepared the manuscript with comments from all authors. AJS, YD and OM were responsible for the management of this project, with AJS leading.

Supplementary Material

Methods	Construction of pAS296 and the cellulose-deficient (CD) mutant; Fourier transform infrared spectroscopy (FTIR); Plasmid isolation, electroporation and strain manipulation; Scanning electron microscopy (SEM); and Swimming motility.
Fig. S1	The *Ba.* 197N *bcs* operon shows DNA-level similarities with those found in key pseudomonads.
Fig. S2	FTIR does not differentiate between wild-type *Ba.* 197N and cellulose-deficient (CD) mutant biofilm samples.
Table S1	*Proteobacteria* having convincing bacterial cellulose synthase (*bsc*) operons.
Table S2	Representative *Proteobacteria* having WssB homologues.
Table S3	General Linear Modelling of biofilm strength and attachment levels.

References

1. Hinz KH, Glünder G. Occurrence of *Bordetella avium* sp. nov. and *Bordetella bronchiseptica* in birds. [In German]. Berl Munch Tierarztl Wochenschr 1985;98:369-73.

2. Raffel TR, Register KB, Marks SA, Temple L. Prevalence of *Bordetella avium* infection in selected wild and domesticated birds in the eastern USA. J Wildl Dis 2002;38:40-46.

3. Grespan A, Camera O, Knöbl T, Gomes CR, Felizardo MR, Ferreira TSP, et al. Virulence and molecular aspects of *Bordetella avium* isolated from cockatiel chicks (*Nymphicus hollandicus*) in Brazil. Vet Microbiol 2012;160:530-34.

4. Spilker T, Liwienski AA, LiPuma JJ. Identification of *Bordetella* spp. in respiratory specimens from individuals with cystic fibrosis. Clin Microbiol Infect 2008;14:504-6.

5. Harrington AT, Castellanos JA, Ziedalski TM, Clarridge JE, Cookson BT. Isolation of *Bordetella avium* and novel *Bordetella* strain from patients with respiratory disease. Emerg Infect Dis 2009;15:72-4.

6. Bos AC, Beemsterboer P, Wolfs TFW, Versteegh FGA, Arets HGM. *Bordetella* species in children with cystic fibrosis: what do we know?: The role in acute exacerbations and chronic course. J Cyst Fibros 2011;10:307-12.

7. Sloan GP, Love CF, Sukumar N, Mishra M, Deora R. The *Bordetella* Bps polysaccharide is critical for biofilm development in the mouse respiratory tract. J Bacteriol 2007;189:8270-82.

8. Conover MS, Sloan GP, Love CF, Sukumar N, Deora R. The Bps polysaccharide of *Bordetella pertussis* promotes colonization and biofilm formation in the nose by functioning as an adhesion. Mol Microbiol 2010;77:1439-55.

9. Gentry-Weeks CR, Cookson BT, Goldman WE, Rimler RB, Porter SB, Curtiss R. Dermonecrotic toxin and tracheal cytotoxin, putative virulence factors of *Bordetella avium*. Infect Immunol 1988;56:1698-1707.

10. Temple LA, Weiss AA, Walker KE, Barnes HJ, Christensen VL, Miyamoto DM, et al. *Bordetella avium* virulence measured *in vivo* and *in vitro*. Infect Immunity 1998;66:5244-51.

11. Sebaihia M, Preston A, Maskell DJ, Kuzmiak H, Connell TD, King ND, et al. Comparison of the genome sequence of the poultry pathogen *Bordetella avium* with those of *B. bronchiseptica*, *B. pertussis*, and *B. parapertussis* reveals extensive diversity in surface structures associated with host interaction. J Bacteriol 2006;188:6002-15.

12. Römling U. Molecular biology of cellulose production in bacteria. Res Microbiol 2002;153:205-12.

13. Spiers AJ, Deeni YY, Folorunso AO, Koza A, Moshynets O, Zawadzki K. Cellulose expression in *Pseudomonas fluorescens* SBW25 and other environmental pseudomonads. In: Cellulose. Van De Ven TGM, Godbout L, editors. Rijeka, Croatia, InTech; 2013.

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

15. Augimeri R, Varley AJ, Strap JL. Establishing a role for bacterial cellulose in environmental interactions: lessons learned from diverse biofilm-producing *Proteobacteria*. Front Microbiol 2015;6:1282.

16. Gal M, Preston GM, Massey RC, Spiers AJ, Rainey PB. Genes encoding a cellulosic polymer contribute toward the ecological success of *Pseudomonas fluorescens* SBW25 on plant surfaces. Mol Ecol 2003;12:3109-21.

17. White AP, Gibson DL, Kim W, Kay WW, Surette MG. Thin aggregative fimbriae and cellulose enhance long-term survival and persistence of *Salmonella*. J Bacteriol 2006;188:3219-27.

18. Gualdi L, Tagliabue L, Bertagnoli S, Ieranò T, De Castro C, Landini P. Cellulose modulates biofilm formation by counteracting curli-mediated colonization of solid surfaces in *Escherichia coli*. Microbiology 2008; 154:2017-24.

19. Nielsen L, Li X, Halverson LJ. Cell–cell and cell–surface interactions mediated by cellulose and a novel exopolysaccharide contribute to *Pseudomonas putida* biofilm formation and fitness under water-limiting conditions. Environ Microbiol 2011;13:1342-56.

20. Spiers AJ, Kahn SG, Bohannon J, Travisano M, Rainey PB. Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. I. Genetic and phenotypic bases of Wrinkly Spreader fitness. Genetics 2002;161:33-46.

21. Spiers AJ, Bohannon B, Gehrig SM, Rainey PB. Biofilm formation at the air–liquid interface by the *Pseudomonas fluorescens* SBW25 Wrinkly Spreader requires an acetylated form of cellulose. Mol Microbiol 2003;50:15-27.

22. Arrebola E, Carrión VJ, Gutiérrez-Barranquero JA, Pérez-García A, Rodríguez-Palenzuela P, Cazorla FM, et al. Cellulose production in *Pseudomonas syringae* pv. *syringae*: a compromise between epiphytic and pathogenic lifestyles. FEMS Microbiol Ecol 2015;91:fiv071.

23. Ude S, Arnold DL, Moon CD, Timms-Wilson T, Spiers AJ. Biofilm formation and cellulose expression among diverse environmental *Pseudomonas* isolates. Environ Microbiol 2006;8:1997-2011.

24. Robertson M, Hapca SM, Moshynets O, Spiers AJ. Air-liquid interface biofilm formation by psychrotrophic pseudomonads recovered from spoilt meat. Antonie van Leeuwenhoek 2013;103:251-9.

25. Udall YC, Deeni Y, Hapca SM, Raikes D, Spiers AJ. The evolution of biofilm-forming Wrinkly Spreaders in static microcosms and drip-fed columns selects for subtle differences in wrinkleality and fitness. FEMS Microbiol Ecol 2015;91:fiv057.
26. Mishra M, Parise G, Jackson KD, Wozniak DJ, Deora R. The BvgAS signal transduction system regulates biofilm development in *Bordetella*. J Bacteriol 2005;187:1474-84.
27. Parise G, Mishra M, Itoh Y, Romeo T, Deora R. Role of a putative polysaccharide locus in *Bordetella* biofilm development. J Bacteriol 2007;198:750-60.
28. Spiers AJ, Rainey PB. The *Pseudomonas fluorescens* SBW25 Wrinkly Spreader biofilm requires attachment factor, cellulose fibre and LPS interactions to maintain strength and integrity. Microbiology 2005;151:2829-39.
29. Gorbatiuk OB, Tsapenko MV, Pavlova MV, Okunev OV, Kordium VA. Bioaffinity sorbent based on immobilized protein A *Staphylococcus aureus*: development and application. Biopolym Cell 2012;28:141-8.
30. Brocchieri L. The GC content of bacterial genomes. J. Phylogen Evolution Biol 2014; 2:1.
31. Koza A, Moshynets O, Otten W, Spiers AJ. Environmental modification and niche construction: developing O_2 gradients drive the evolution of the Wrinkly Spreader. ISME J 2011; 5:665-73.
32. Loker SB, Temple LM, Preston A. The *Bordetella avium BAV1965-1962* fimbrial locus is regulated by temperature and produces fimbriae involved in adherence to Turkey tracheal tissue. Infect Immunity 2011;79:2423-9.
33. Beach NM, Thompson S, Mutnick R, Brown L, Kettig G, Puffenbarger R, et al. *Bordetella avium* antibiotic resistance, novel enrichment culture, and antigenic characterization. Vet Microbiol 2012;160:189-96.

34. Spiers AJ. A mechanistic explanation linking adaptive mutation, niche change and fitness advantage for the Wrinkly Spreader. Int J Evolutionary Biol 2014;Article ID 675432.

35. Koza A, Hallett PD, Moon CD, Spiers AJ. Characterization of a novel air–liquid interface biofilm of *Pseudomonas fluorescens* SBW25. Microbiology 2009; 155:1397-1406.

36. O'Toole GA, Kolter R. Initiation of biofilm formation in *Pseudomonas fluorescens* WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. Mol Microbiol 1998;28:449-61.

37. Goymer P, Kahn SG, Malone JG, Gehrig SM, Spiers AJ, Rainey PB. Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. II. Role of the GGDEF regulator WspR in evolution and development of the Wrinkly Spreader phenotype. Genetics 2006; 173:515-26.

38. D'Argenio DA, Calfee MW, Rainey PB, Everett C, Pesci EC. Autolysis and autoaggregation in *Pseudomonas aeruginosa* colony morphology mutants. J Bacteriol 2002;184:6481-9.

39. Aldridge P, Paul R, Goymer P, Rainey P, Jenal U. Role of the GGDEF regulator PleD in polar development of *Caulobacter crescentus*. Mol Microbiol 2003;47:1695-1708.

40. Sisti F, Ha DG, O'Toole GA, Hozbor D, Fernández J. Cyclic-di-GMP signalling regulates motility and biofilm formation in *Bordetella bronchiseptica*. Microbiology 2013; 159:869-79.

41. Goymer PJ. The role of the WspR response regulator in the adaptive evolution of experimental populations of *Pseudomonas fluorescens* SBW25. D.Phil. thesis. Oxford, University of Oxford; 2002.

42. Rainey PB, Bailey MJ. Physical and genetic map of the *Pseudomonas fluorescens* SBW25 chromosome. Mol Microbiol 1996;19:521-33.

Figure Legends

Figure 1. **Bacterial cellulose synthase (*bcs*) operon structures.** *Bcs* operons have been identified in the whole-genome sequences of a wide range of bacteria. In *Pf.* SBW25 the genes are annotated as *wssA* – *J* and are predicted to form the core cellulose synthase complex (dark blue) including an endoglucanase (light blue), associated acetylation subunits (green), and be involved in the localization of the complex at the periplasmic membrane (mauve) (top section). Shown also are the *bcs* operons of *Pp.* KT2440, *Ps.* DC3000, and *Ba.* 197N which also contains a duplication of a truncated version of *wssC* (blue stripes). The scale bar indicates 1 kb.

Figure 2. **Phylogenetic analyses of representative WssB homologues.** Shown here is a phylogenetic tree showing the relationships between WssB proteins from the key strains *Ba.* 197N, *Pf.* SBW25, *Pp.* KT2440 and *Ps.* DC3000 (bold) as well as a number of other bacteria in which *wss*-like *bcs* operons have been identified. A number of *Komagataeibacter* spp. sequences were used to root the tree (indicated by the black square), and these α-group sequences represented by *Gluconacetobacter xylinus* E25 (light blue) were found to cluster separately from the other β and γ-group sequences. Most γ-group *Pseudomonas* spp. sequences clustered together (light green). Although the β-group sequences root at the base of the *Pseudomonas* spp. cluster, the *Bordetella* spp. sequences including *Ba.* 197N (purple) clustered separately from the

Burkholderia spp. (light orange). This Neighbour joining (NJ) tree is based on COBALT multiple alignments of the coding sequences listed in Table 1. Some sequences have not been labelled in the tree for clarity. The distance bar indicates 0.1 units.

Figure 3. **GC content of the *bcs* operons.** An analysis of the *bcs* operons from *Ba.* 197N, *Pf.* SBW25, *Pp.* KT2440 and *Ps.* DC3000 reveals substantial variation in gene GC content within operons. The GC contents of individual genes (grey circles), the mean with SE bars for each operon (squares), and each genome (triangles) are shown. See the main results section for statistical comparisons.

Figure 4. ***Ba.* 197N forms biofilms at the air-liquid (A-L) interface of static microcosms**. Shown are representative images of *Ba.* 197N biofilms *in situ* in BHI, KB and LB microcosms (A – C, top row) and of biofilm material after pouring into petri dishes (bottom row). Biofilms become progressively more robust with longer incubation periods, but tend to detach from the meniscus region and sink to the bottom of the vials. Microcosms were incubated at 20°C for three days before imaging.

Figure 5. **The *Ba.* 197N biofilm is intermediate in strength but with high levels of attachment.** Biofilm-formation in static microcosms can be quantified by using a combined biofilm assay in which microcosm growth (A), biofilm strength (B) and attachment levels (C) are determined. Shown here are data for *Pf.* SBW25 which produces the VM biofilm (white bars), the Wrinkly Spreader which produces the WS biofilm (black), *B. avium* 197N (indicated by ‘Wt’ for wild-type,

dark grey), and the *Ba.* 197N cellulose-deficient (CD) mutant (light grey) incubated in KB microcosms at 20°C for three days before assay. Means and SE are shown. Significant differences between means were observed for growth ($p < 0.0001$), strength ($p < 0.0001$) and attachment ($p = 0.0005$); means within assays not connected by the same letter are significantly different ($\alpha = 0.05$).

Figure 6. **Cellulose is expressed by *Ba.* 197N but is not essential for biofilms.** Cellulose fibres forming the matrix of the *Pf.* SBW25 Wrinkly Spreader mutant are readily visualised by fluorescent microscopy after staining with Calcofluor (A). In comparison, only low levels of cellulose expression can be detected in *Ba.* 197N biofilm samples and none in the cellulose-deficient (CD) mutant even at higher magnification. Qualitative comparison of scanning electron microscope images of *Ba.* 197N and CD mutant biofilm samples show no obvious differences in structure (B), suggesting that the presence of cellulose has little apparent impact on biofilm structure. Congo red, which also binds cellulose and other compounds including proteinaceous attachment factors, stains colonies of *Ba.* 197N expressing the constitutively-active diguanylate cyclase WspR19 mutant *in trans* (C). However, little staining is apparent in colonies of wild-type *Ba.* 197N, *Ba.* 197N transformed with pVSP61 (the control plasmid), pVSP61-WspR9, pVSP61-WspR12, or the CD mutant.

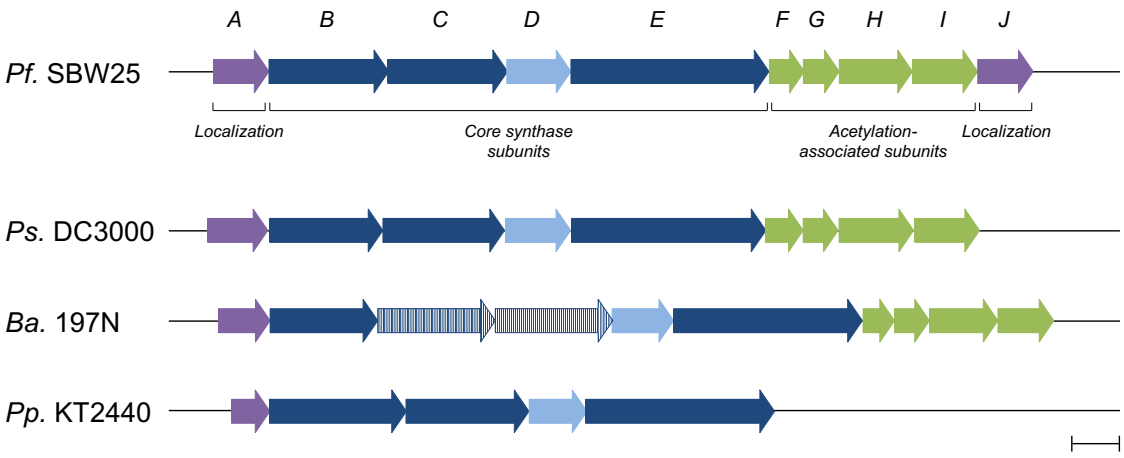

A
B
C
D
E
F
G
H
I
J
Pf. SBW25
Localization
Core synthase subunits
Acetylation-associated subunits
Localization
Ps. DC3000
Ba. 197N
Pp. KT2440

Figure 1

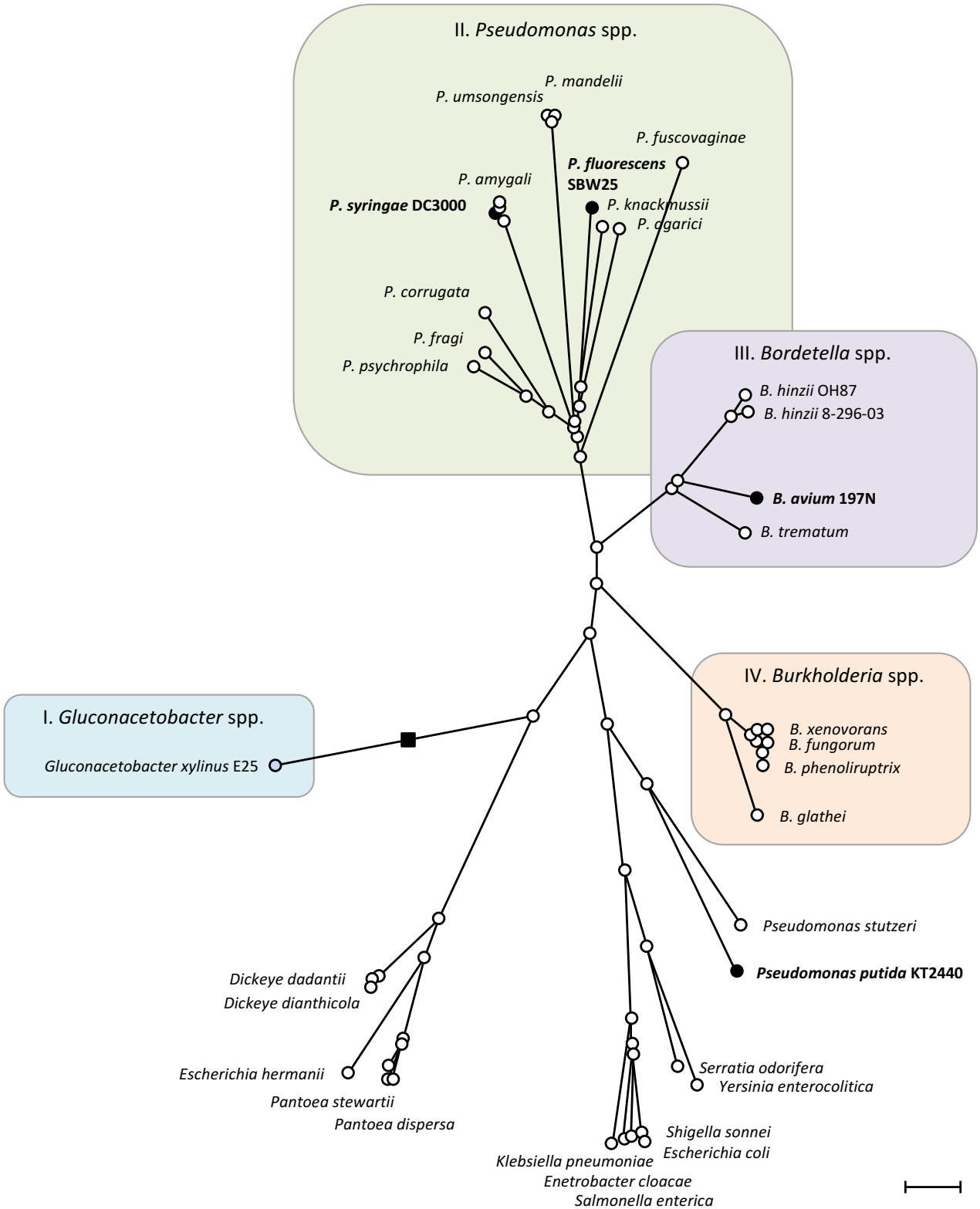

II. Pseudomonas spp.
P. mandelii
P. umsongensis
P. fuscovaginae
P. fluorescens SBW25
P. amygali
P. syringae DC3000
P. knackmussii
P. agarici
P. corrugata
P. fragi
P. psychrophila
III. Bordetella spp.
B. hinzii OH87
B. hinzii 8-296-03
B. avium 197N
B. trematum
IV. Burkholderia spp.
B. xenovorans
B. fungorum
B. phenoliruptrix
B. glathei
I. Gluconacetobacter spp.
Gluconacetobacter xylinus E25
Pseudomonas stutzeri
Pseudomonas putida KT2440
Dickeye dadantii
Dickeye dianthicola
Escherichia hermanii
Pantoea stewartii
Pantoea dispersa
Serratia odorifera
Yersinia enterocolitica
Shigella sonnei
Escherichia coli
Klebsiella pneumoniae
Enetrobacter cloacae
Salmonella enterica

Figure 2

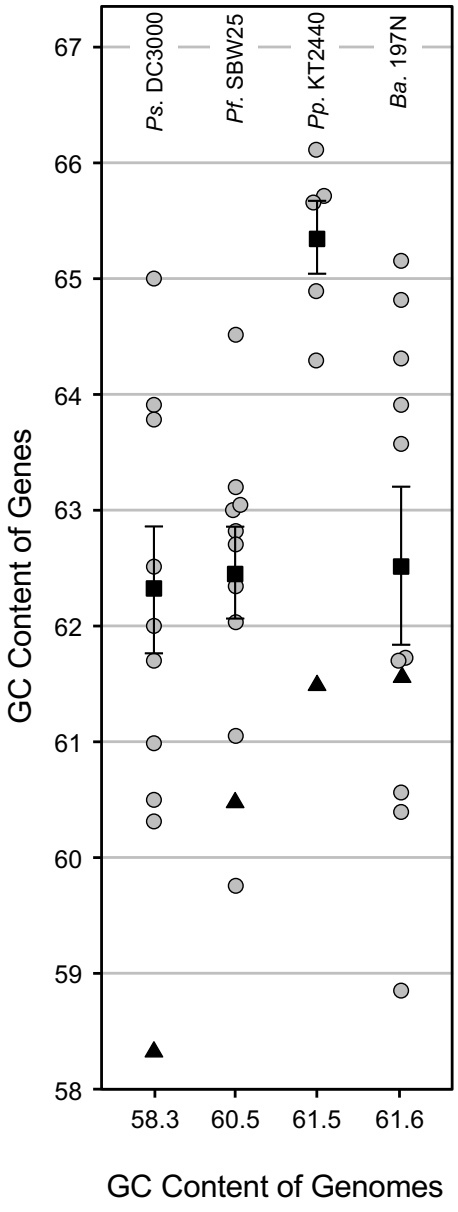

Ps. DC3000
Pf. SBW25
Pp. KT2440
Ba. 197N
67
66
65
64
63
62
61
60
59
58
GC Content of Genes
58.3
60.5
61.5
61.6
GC Content of Genomes

Figure 3

(A) BHI microcosms (B) KB microcosms (C) LB microcosms

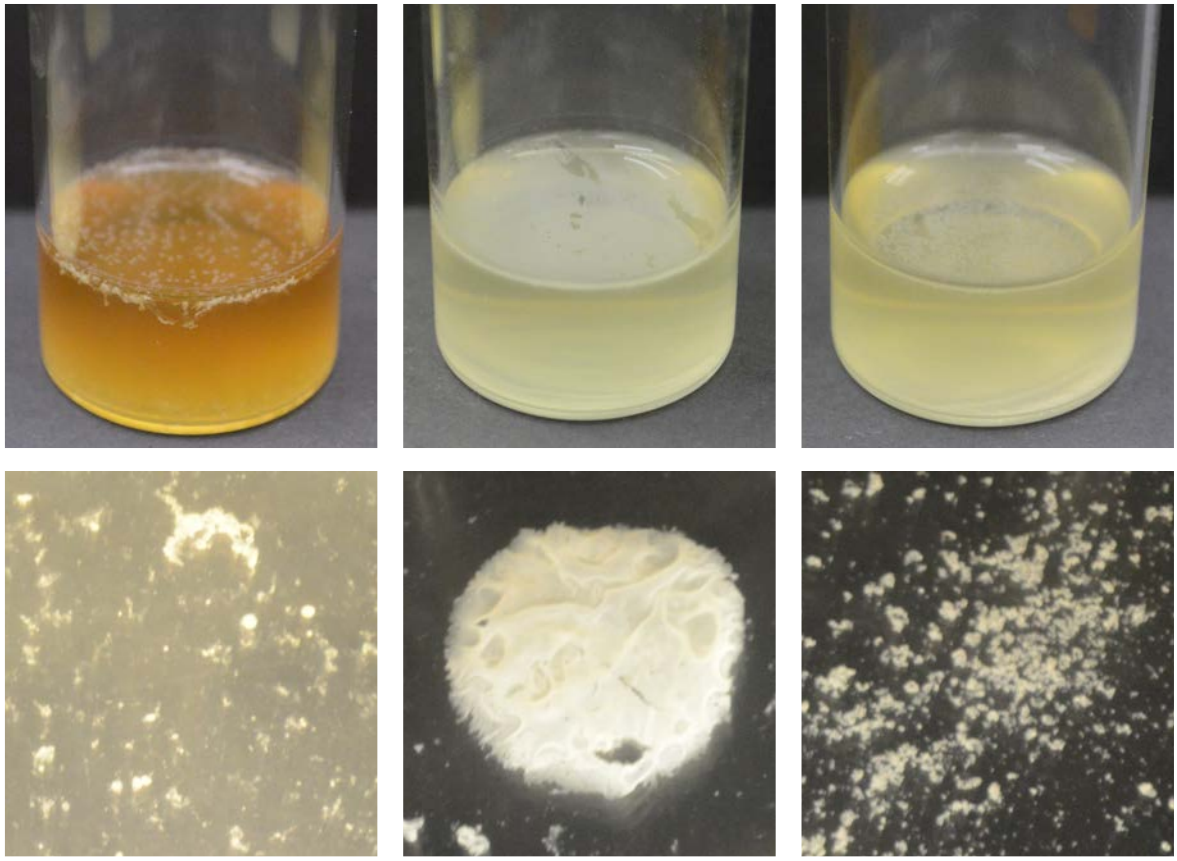


Figure 4

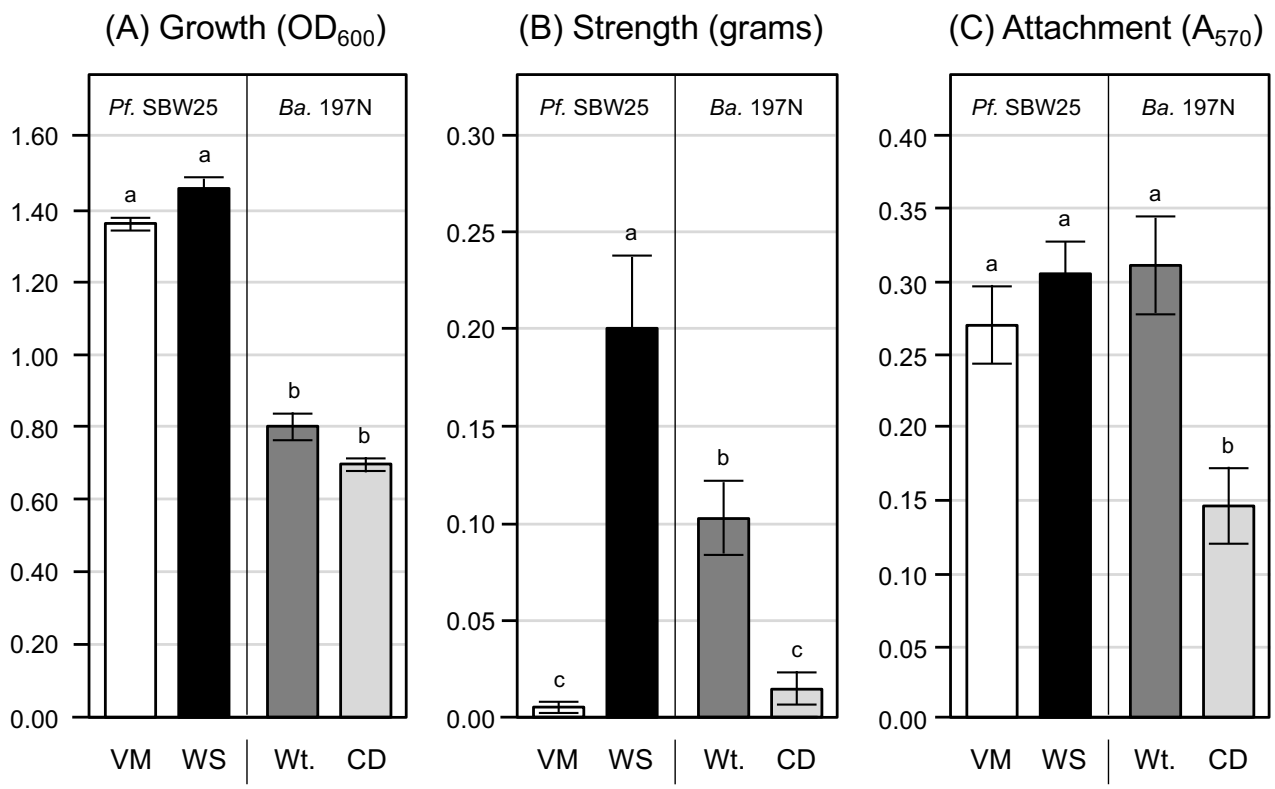
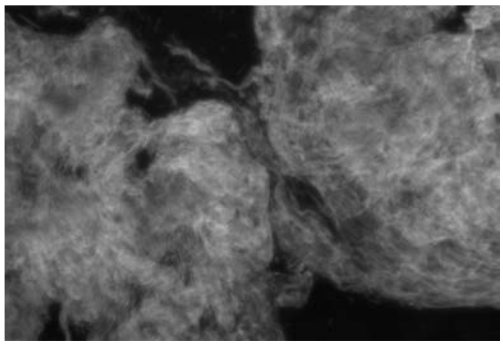
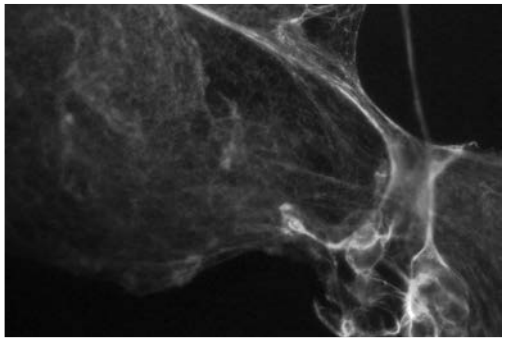

(A) Growth (OD_{600})
(B) Strength (grams)
(C) Attachment (A_{570})
Pf. SBW25
Ba. 197N
VM
WS
Wt.
CD
a
b
c

Figure 5

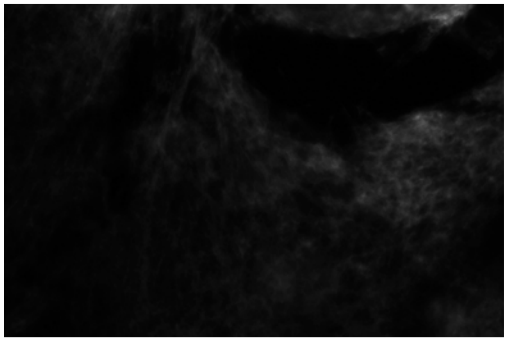
(A) Fluorescent microscopy



Pf. SBW25 Wrinkly Spreader (10x)

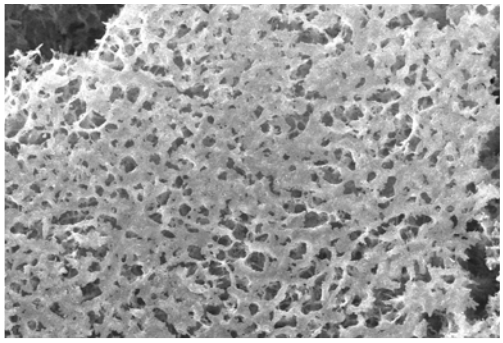


Ba. 197N (40x)

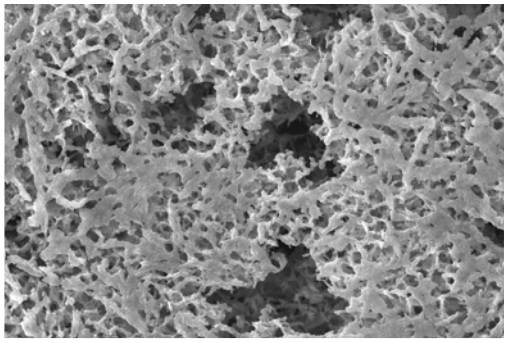


CD mutant (40x)

(B) Scanning electron microscopy

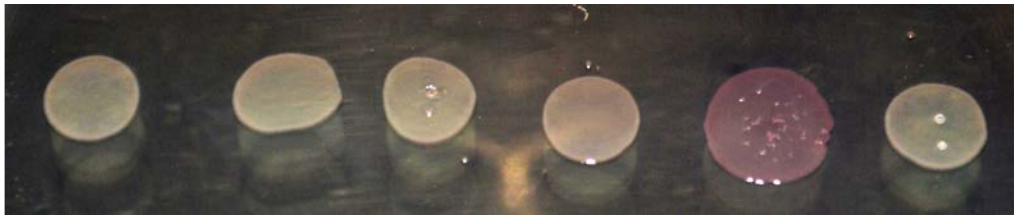


Ba. 197N (500x)



CD mutant (500x)

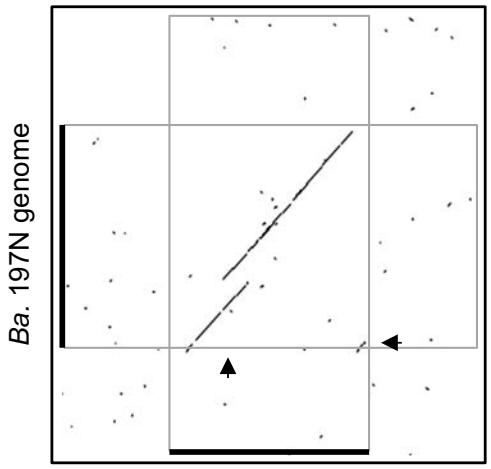
(C) Congo red plates



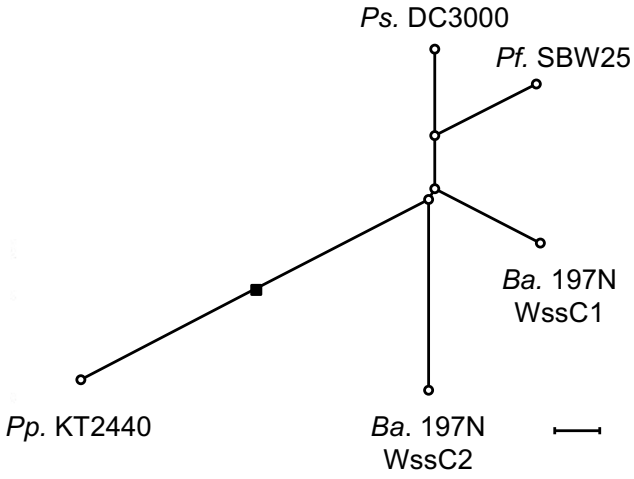
Ba. 197N pVSP61 WspR9 WspR12 WspR19 CD

Figure 6

(A) DNA sequence similarity



(B) NJ Tree of WssC homologues



Suppl. Materials Figure S1

Table 1. Bacteria and plasmids.

Strains and plasmids	*Comment*	*Source / Reference*
Bordetella avium 197N	Wild-type strain (ATCC BAA-1003), Km^S, Tc^S.*	ATCC [9, 10]
Ba. 197N cellulose deficient (CD) mutant	*Ba.* 197N::*wssB*::ISΩ-Km/*hah* constructed using pAS296, Km^R.	This work
Escherichia coli K12 DH5α	Standard host for maintaining plasmids.	Gibco-BRL, UK
Pseudomonas fluorescens SBW25	Wild-type strain expressing low-levels of cellulose and producing a weak, viscous mass (VM)-class biofilm.	[42]
Pf. SBW25 Wrinkly Spreader (WS) mutant	*Pf.* SBW25 *wssF* (*S301R*) over-expressing cellulose and producing a robust, physically cohesive (PC)-class biofilm.	[20]
pAS296	Suicide plasmid designed to disrupt *Pf.* SBW25 *wss*-like *bcs* operons derived from pNR9.3, Km^R.	This work
pNR9.3	A self-replicating *wssB*::IS-**Ω**-Km/hah mini-transposon construct, Km^R.	S. Giddens
pVSP61-Tc	A modified version of the pVSP61 broad host-range cloning vector, Km^R and Tc^R.	[28, 41]
pVSP61-WspR9-Tc	pVSP61-Tc expressing the dominant-negative diguanylate cyclase mutant, WspR9 (*G296R*), Km^R and Tc^R.	[41]
pVSP61-WspR12-Tc	pVSP61-Tc expressing the wild-type diguanylate cyclase, WspR12, Km^R and Tc^R.	[28, 41]
pVSP61-WspR19-Tc	pVSP61-Tc expressing the constitutively-active diguanylate cyclase mutant, WspR19 (*R129C*), Km^R and Tc^R.	[28, 41]

All plasmids were maintained in *Ec.* K12. ATCC, American Type Culture Collection. Km and Tc were used at 50 and 12.5 $\mu g.ml^{-1}$, respectively, to select for plasmids in *Ec.* K12, and at 100 and 12.5 $\mu g.ml^{-1}$ for *Ba.* 197N strains, respectively. *, Antibiotic susceptibility using MASTRING M13, M14, and M51 disks was determined for *Ba.* 197N (S, Sensitive; R, Resistant): Ampicillin (R), Cephalexin (R), Cephalthin (R), Chloramphenicol (S), Ciprofloxacin (R), Coliston sulphate (S), Cotrimoxazole (R), Erythromycin (R), Fusidic acid (R), Gentamycin (S), Nalidixic acid (R), Nitrofurantoin (R), Novobiocin (R), Oxacillin (R), Penicillin (R), Streptomycin (S), Tetracycline (S), Trimethoprim (R) (these are in agreement with [33].

Biofilm formation and cellulose expression by *Bordetella avium* 197N, the causative agent of bordetellosis in birds and an opportunistic respiratory pathogen in humans (McLaughlin *et al*.)

Supplementary Material

This contains supplementary methods, figures and tables.

S2.1 Bioinformatics

Protein homologues were identified using the Basic Local Alignment Search Tool (BLAST) and the International Nucleotide Sequence Database Collaboration (www.ncbi.nlm.nih.gov/blast). Bacterial cellulose synthase subunit homologues were identified in *Proteobacteria* complete genomes using TBLASTN and *Pf.* SBW25 WssA – J protein query sequences (Accession numbers AAL71841 – AAL71850) with default parameters selected, requiring ≥ 25% similarity and ≥ 80% coverage. Genomes were further sorted by requiring at least three homologues from WssB, C, D or E, and the annotation of each homologue confirmed using Inter-ProScan (www.ebi.ac.u/interpro). The annotated genome sequences were obtained from GenBank and imported into Genious v5.6.4. (Biomatters Ltd, NZ) to visualise operon structures. The *Pseudomonas* Genome Database (www.pseudomonas.com) was also used for this purpose. WssB homologues were also identified in the non-redundant (NR) protein database using BLASTP to complement those identified in the complete genome search. A WspR homologue was similarly identified in the *Ba.* 197N genome using the WspR protein query sequence (AAL71852).

Phylogenetic trees were produced using the COBALT multiple sequence alignment tool (www.ncbi.nlm.nih.gov/tools/cobalt). Protein sequences were aligned and viewed in Phylogenetic Tree View using the Neighbour Joining (NJ) method, with a maximum sequence difference of 0.85 and distances calculated according to the Grishin General (Protein) method. The trees were rendered using the radial method and WssB trees rooted using α-group *Proteobacteria* sequences. For the WssC tree, the root was automatically positioned mid-way along the longest branch.

Gene and genomic GC content data were obtained from the *Pseudomonas* Genome Database and Genbank. The YASS Genomic Similarity Search Tool (bioinfo.lifl.fr) with default parameters selected was used to produce dot plots showing the DNA-level homology between *Ba.* 197N and *Pf.* SBW25 sequences.

S2.2 Construction of pAS296 and the cellulose-deficient (CD) mutant

The suicide plasmid pAS296 was designed to disrupt *wss*-like *bcs* operons by homologous recombination of a mini-transposon *wssB*::IS-Ω-Km/hah (Km^R) cassette cassette. This was based on pNR9.3, a circularised *Kpn*I chromosomal fragment from a *Pf.* SBW25 Wrinkly Spreader mutant containing a IS-Ω-Km/hah (Km^R) mini-transposon insertion in *wssB* (the insertion is located 1,693 bp from the start of the 2,219 bp *wssB* gene). pNR9.3 was reduced in size by digesting with *Eco*RV (NEB) and re-ligation using T4 DNA Ligase (NEB) to produce pAS296. This was integrated into the *Ba.* 197N chromosome, following electroporation and selection with kanamycin, to produce the cellulose-deficient (CD) mutant. pAS296 plasmid DNA could not be re-isolated from the CD mutant and used to transform *Ec.* K12 DH5α, indicating that the homologous recombinant was stable. Furthermore, inspection of colony and biofilm samples produced by the CD mutant by fluorescent microscopy after staining with Calcofluor showed no signs of cellulose production, confirming that the *Ba.* 197N *bcs* operon had been disrupted by the cassette.

S2.3 Fourier transform infrared spectroscopy (FTIR)

Biofilm samples (n = 3) were first drained to remove excess media, and then prepared by washing 3 times in de-ionised water over a period of three days at room temperature. The samples were then freeze-dried for 48 hours at -40°C. A control sample was prepared using powdered cellulose (SIGMA-ALDRICH, UK) and similarly freeze-dried. The vibrational properties of the dried samples were determined using a Vertex-70 Fourier Transform Infrared (FT-IR) spectrometer equipped with a DTGS detector (Bruker, Germany). The samples were placed into contact with the single reflection diamond ATR crystal and the

spectra were collected over the range of 400 – 4000 cm^{-1} at a resolution of 1 cm^{-1}.

S2.4 Plasmid isolation, electroporation and strain manipulation

Plasmid DNA was isolated from over-night cultures using Wizard *Plus* SV (Promega, UK) and GeneJET (Thermo Scientific, UK) plasmid miniprep kits, and DNA checked by TBE-agarose gel electrophoresis. Plasmid DNA was dialysed against de-ionised water using Millipore 0.2 µm HA MF-membrane filters for 1 h before electroporation using an Eppendorf Electroporator 25109 set at 1.75 V and 200 Ω (with a time constant of ≥ 4) and *Ba.* 197N competent cells prepared using 10% (v/v) glycerol, 1 mM HEPES (after [S1]). Electroporated cells were spread onto selective plates and appropriate colonies re-streaked and used to prepare -80°C stocks for further use. The ability of pVSP61-Tc and pVSP61-Tc–based plasmids (Tc^R) to replicate autonomously in *Ba.* 197N was confirmed by re-isolating plasmid DNA and demonstrating that it could successfully transform $CaCl_2$-treated competent *Ec.* K12 DH5α cells using the heat-shock method [S1]. The presence of plasmid DNA in *Ec.* K12 DH5α transformants was further confirmed by isolation, *Bam*HI and *Eco*RI (NEB, UK) digestion and gel electrophoresis.

S2.5 Scanning electron microscopy (SEM)

Biofilm samples were first drained to remove excess media. They were then placed on specimen mounts, frozen in liquid nitrogen for 15 minutes, then transferred to a cooled vacuum freezer and freeze-dried for 48 hours at -40°C. Dried samples were covered with gold and biofilms samples imaged using a JSM-6060 LA microscope (JEOL, Japan) and representative images compared qualitatively visually.

S2.6 SPA-CBD ELISA

Cellulose was detected in semi-purified biofilm matrix samples by ELISA using a *Staphylococcus aureus* Protein A-Cellulose Binding Domain fusion protein (SPA-CBD) [S2]. Biofilm samples were prepared by washing 10 times in de-ionised water over a period of three days. Samples were then incubated in

1.5 ml 8M urea for 1 hr followed by 1.5 ml in PBS (2.7mM KCl, 137mM NaCl, 10mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) for 1 hr, before sonication in 100 mM Tris-HCl (pH 8.0), 6M guanidine hydrochloride (GdnHCl), 200 mM NaCl to solubilize cell debris and proteins. After 1 hr incubation at room temperature, the matrix sample was collected by centrifugation at 7,000 xg for 5 min and the sonication procedure repeated two further times. The matrix sample was washed twice with 0.5 ml TBS buffer (1 mM EDTA, 500 mM NaCl, 20 mM Tris-HCl, pH 8.0) to release the GdnHCl and left in fresh TBS buffer.

A 0.5 mg/ml solution of SPA-CBD was prepared in TBS buffer. 200 µl of SPA-CBD was added to the matrix sample in 20 µl of TBS and incubated for 1 hr at room temperature. The sample was then washed to remove unbound SPA-CBD three times with TBS buffer containing 0.5 M NaCl and then in PBST (PBS with 150 mM NaCl, 0.1 % (w/v) Tween 20) to block further non-specific binding. 0.2 ml 0.2 mg/ml purified rabbit IgG was then added and incubated for 1 hr. 0.2 ml of 1:10,000 Anti-rabbit IgG – Alkaline Phosphatase conjugate (Sigma, USA) and 0.2 ml 10 mg/ml solution of p-nitrophenylphosphate (Sigma, USA) in buffer (0.14 M NaCl, 0,015 M $MgSO_4$, 50 mM Tris-HCl, pH 9.5) were added and absorbance measured after 30 min at 405 nm using a Multiscan MCC/340 multichannel photometer (Titertek, USA).

Supplementary References

S1. Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning: A Laboratory Manual. 2nd edition. Cold Spring Harbor, CSH Laboratory Press; 1989.

S2. Gorbatiuk OB, Tsapenko MV, Pavlova MV, Okunev OV, Kordium VA. Bioaffinity sorbent based on immobilized protein A *Staphylococcus aureus*: development and application. Biopolym Cell 2012;28:141-8.

Supplementary Figures

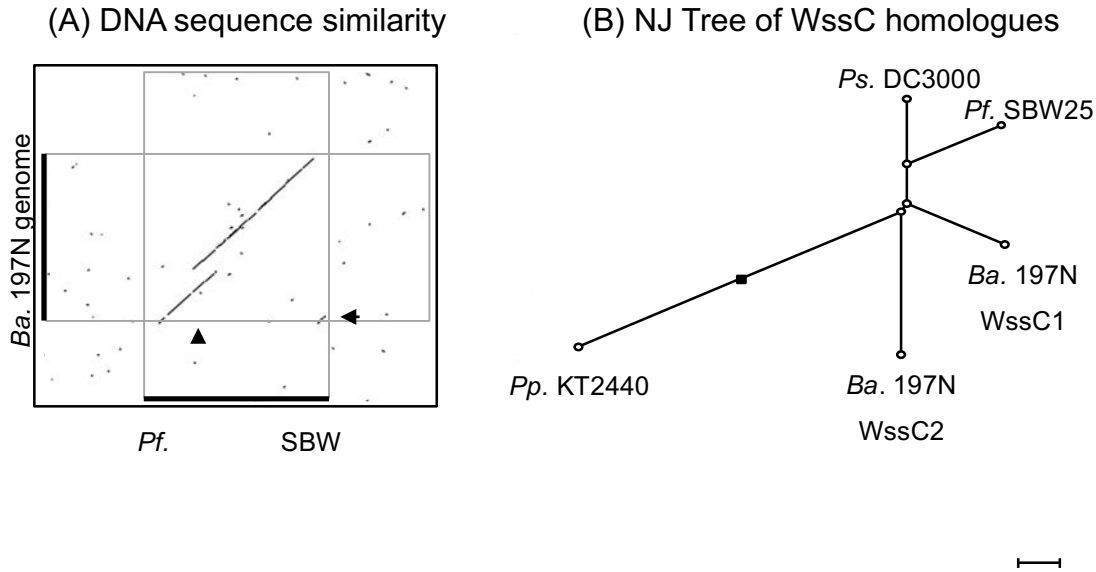


Figure S1. **The *Ba.* 197N *bcs* operon shows DNA-level similarities with those found in key pseudomonads.** The *Ba.* 197N *bcs* operon shows DNA-level similarity when compared to the archetypal *Pf.* SBW25 *wss* operon in a dot plot (A). However, no significant sequence homology is seen outside the two operons indicated by the two black bars. The vertical arrow indicates the duplication of *wssC* in *Ba.* 197N, and the horizontal arrow indicates the high level of homology between *wssA* and *wssJ* in *Pf.* SBW25. The YASS dot plot is approximately 56 Kbp wide and high. The phylogenetic tree (B) shows the relationship between the duplicated WssC proteins of *Ba.* 197N (WssC1, CAJ50239, and WssC2, CAJ50240), and the WssC proteins of *Pf.* SBW25 (CAY46579), *Pp.* KT2440 (NP_744780), and *Ps.* DC3000 (AAO54559.1). This Neighbour joining (NJ) tree is based on COBALT multiple alignments of the WssC sequences; the root is marked by the square, and the distance bar indicates 0.2 units.

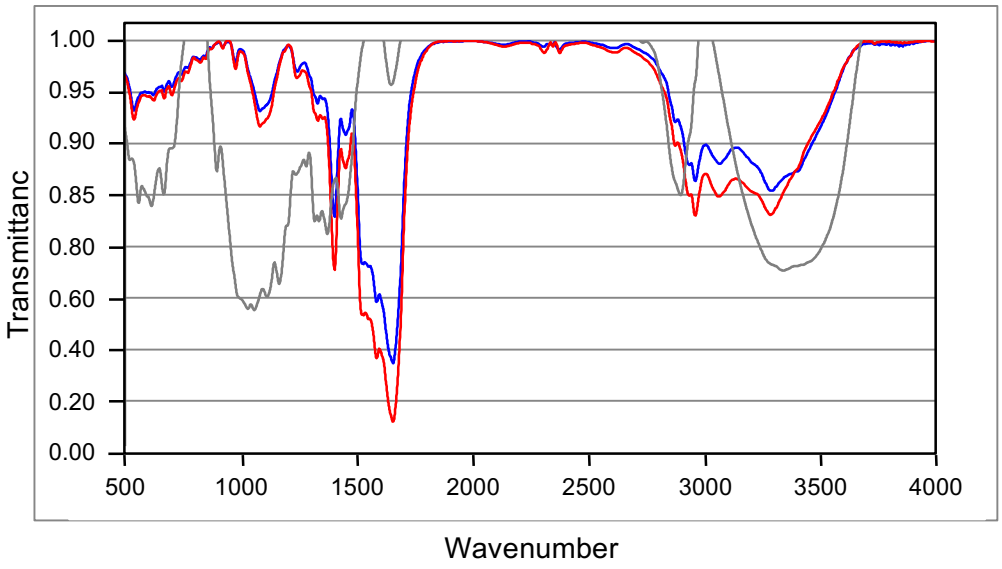


Figure S2. **FTIR does not differentiate between wild-type *Ba.* 197N and cellulose-deficient (CD) mutant biofilm samples.** Fourier transform infrared spectroscopy (FTIR) can differentiate samples based on chemical properties. However, FTIR traces obtained from wild-type *Ba.* 197N (red) and the CD mutant (blue) show no obvious differences, suggesting that cellulose must be a minor component of the *Ba.* 197N biofilm. The FTIR trace for purified cellulose (grey) is provided for comparison and transmittance re-scaled to match the biofilm traces. As transmittance is inversely proportional to chemical concentration, troughs in the cellulose trace indicate regions of the spectra that are characteristic of the polymer. Mean traces (n = 3) are shown and SE omitted for clarity.

Table S1. *Proteobacteria* having convincing bacterial cellulose synthase (*bsc*) operons.

Class	*Genus*	*Characteristic*
α	*Komagataeibacter*	Environmental (rotting fallen fruit) species.
β	*Bordetella*	Animal (including human) pathogens.
	Burkholderia	Animal (including human) and plant pathogens, environmental (soil and plant-associated) species.
γ	*Enterobacter*	Opportunistic human (respiratory and urinary tract) pathogen.
	Escherichia	Animal pathogen and gastrointestinal tract commensal species.
	Dickeya	Plant pathogens.
	Klebsiella	Opportunistic pathogens of animals (including humans), gastrointestinal tract commensal, environmental (water, soil and plant-associated) species.
	Morganella	Gastrointestinal tract commensal species.
	Pantoea	Human and plant pathogens.
	Pseudomonas	Mushroom and plant pathogens, opportunistic human pathogens, environmental (soil and plant-associated) species.
	Salmonella	Animal (including human) pathogens.
	Serratia	Human (respiratory and urinary tract) pathogens.
	Yersinia	Human pathogens.

This list is not exhaustive but includes those genera represented by the WssB homologues listed in Table S2 and shown in Figure 2.

Table S2. Representative *Proteobacteria* having WssB homologues.

Class	*Genus*	*Species*	*WssB Accession*
α	*Komagataeibacter* [A]	*Gluconacetobacter rhaeticus* AF1	KDU95596
		Gluconacetobacter xylinus E25	AHI26282
		Komagataeibacter oboediens	WP_029329219
		Komagataeibacter xylinus	WP_025438893
β	*Bordetella* [B]	*Bordetella avium* 197N	WP_012418273
		Bordetella hinzii 4161	KCB44674
		Bordetella hinzii 8-296-03	KCB37643
		Bordetella hinzii L60	WP_029577393
		Bordetella hinzii OH87 BAL007II	KCB24853
		Bordetella trematum	WP_025518087
	Burkholderia [C]	*Burkholderia bryophila*	WP_020066409
		Burkholderia caledonica	WP_027775993
		Burkholderia fungorum	WP_028200609
		Burkholderia glathei	KDR42276
		Burkholderia phenoliruptrix BR3459a	WP_014973068
		Burkholderia phytofirmans PsJN	WP_012427681
		Burkholderia xenovorans LB400	WP_011491297
γ	*Enterobacteria* [D]	*Dickeya dadantii 3937*	YP_003880884
		Dickeya dianthicola	WP_024104083
		Enterobacter cloacae UCI 30	EUL69702
		Escherichia coli K12	WP_020236356
		Escherichia hermannii	WP_002434320
		Klebsiella pneumoniae CIP 52.145	CDO12002
		Morganella morganii	WP_024249582
		Pantoea dispersa	WP_021507015
		Pantoea sp. At-9b	WP_013507310
		Pantoea stewartii	WP_006121414
		Salmonella enterica subsp. *enterica serovar* Typhimurium	CAC44015
		Serratia odorifera DSM 4582	WP_004957272
		Yersinia enterocolitica	WP_025376533
	Pseudomonas [E]	*Pseudomonas agarici*	WP_017132712

Pseudomonas amygdali	WP_005733961
Pseudomonas amygdali	WP_005746495
Pseudomonas corrugata	WP_024776313
Pseudomonas fluorescens SBW25	AAL71842
Pseudomonas fragi	WP_016780387
Pseudomonas fuscovaginae	WP_019360253
Pseudomonas knackmussii B13	CDF84393
Pseudomonas plecoglossicida	WP_028625022
Pseudomonas psychrophila	WP_019411348
Pseudomonas putida BIRD-1	YP_005930869
Pseudomonas putida DOT-T1E	YP_006531049
Pseudomonas putida GB-1	YP_001669425
Pseudomonas putida KT2440	WP_010953555
Pseudomonas putida NBRC 14164	YP_008114258
Pseudomonas putida ND6	YP_006384055
Pseudomonas putida S12	ESP85389
Pseudomonas savastanoi pv. *savastanoi* NCPPB 3335	EFI00895
Pseudomonas stutzeri	WP_014595547
Pseudomonas stutzeri A1501	YP_001170829
Pseudomonas stutzeri ATCC 17588	YP_004712680
Pseudomonas stutzeri DSM 4166	YP_005936994
Pseudomonas syringae pv *tomato* DC3000	AAO54558

A, Representative *Komagataeibacter* spp. used to root the trees. B, All *Bordetella* spp. having BcsA / WssB homologues. C, Representative *Burkholderia* spp.. D, Representative *Enterobacteria* spp. including human and plant pathogens (with many species highly represented in the databases not sampled further). E, Representative *Pseudomonas* spp. with additional *P. putida* and *P. stutzeri* strains selected to investigate apparently mono-phyletic branches (with many species highly represented in the databases not sampled further).

Table S3. General Linear Modelling of biofilm strength and attachment levels.

Model	*1*	*2*	*3*	*4*
(A) Model Fit and Analysis of Variance				
RSquare	0.54	0.40	0.66	0.45
df	6, 65	5, 66	7, 84	6, 86
F	12.78	8.65	23.25	11.55
p value	<0.0001	<0.0001	<0.0001	<0.0001
(B) Response and Effects (p values)				
Response	Strength	Attachment	Strength	Attachment
Medium	<0.0001	<0.0001	<0.0001	<0.0001
Strain	0.0001 [A]	0.0501 [A]	<0.0001 [B]	0.0727 [B]
Growth	0.1797	<0.0001	<0.0001	0.0710
Attachment	0.1222	–	0.1354	–

The combined biofilm assay dataset was used to model biofilm strength as response with various effects using a GLM approach. Assay temperature was 20°C. Media included BHI, KB, and LB. A, *Ba.* 197N, *Pf.* SBW25 (VM), and WS. B, *Ba.* 197N carrying pVSP61-Tc, pVSP61-WspR9-Tc, pVSP61-WspR12-Tc, and pVSP61-WspR19-Tc. –, Effect not included in model.