

# *Independent evolution of shape and motility allows evolutionary flexibility in Firmicutes bacteria*

Article

Accepted Version

El Baidouri, F., Venditti, C. and Humphries, S. (2016) Independent evolution of shape and motility allows evolutionary flexibility in Firmicutes bacteria. *Nature Ecology & Evolution*, 1 (1). 0009. ISSN 2397-334X doi: <https://doi.org/10.1038/s41559-016-0009> Available at <https://centaur.reading.ac.uk/68268/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://dx.doi.org/10.1038/s41559-016-0009>

To link to this article DOI: <http://dx.doi.org/10.1038/s41559-016-0009>

Publisher: Nature

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

1 **Title: Independent evolution of shape and motility allows evolutionary flexibility in**  
2 ***Firmicutes* bacteria.**

3 Fouad El Baidouri<sup>1\*</sup>, Chris Venditti<sup>2</sup> & Stuart Humphries<sup>1</sup>

4 **Author affiliation:**

5 Stuart Humphries, ORCID ID: 0000-0001-9766-6404

6 [shumphries@lincoln.ac.uk](mailto:shumphries@lincoln.ac.uk)

7 <sup>1</sup> School of Life Sciences University of Lincoln, Joseph Banks Laboratories, Green Lane,  
8 Lincoln, LN6 7DL, UK.

9 Chris Venditti

10 [c.d.venditti@reading.ac.uk](mailto:c.d.venditti@reading.ac.uk)

11 <sup>2</sup> School of Biological Sciences, University of Reading, Reading RG6 6BX, UK.

12 \* **Corresponding author:** Fouad El Baidouri

13 School of Life Sciences University of Lincoln, Joseph Banks Laboratories, Green Lane,  
14 Lincoln, LN6 7DL, UK.

15 Tel: +(44) 01522887420

16 [felbaidouri@lincoln.ac.uk](mailto:felbaidouri@lincoln.ac.uk)

17 ORCID ID: 0000-0001-5204-6244

18

19 **Abstract**

20

21 Functional morphological adaptation is an implicit assumption across many ecological  
22 studies. However, despite a few pioneering attempts to link bacterial form and function,  
23 functional morphology is largely unstudied in prokaryotes. One intriguing candidate for  
24 analysis is bacterial shape, as multiple lines of theory indicate that cell shape and motility  
25 should be strongly correlated. Here we present a large-scale use of modern phylogenetic  
26 comparative methods to explore this relationship across 325 species of the phylum  
27 *Firmicutes*. In contrast to clear predictions from theory, we show that cell shape and motility  
28 are not coupled, and that transitions to and from flagellar motility are common and strongly  
29 associated with lifestyle (free-living or host-associated). We find no association between  
30 shape and lifestyle, and contrary to recent evidence, no indication that shape is associated  
31 with pathogenicity. Our results suggest that the independent evolution of shape and motility  
32 in this group might allow a greater evolutionary flexibility.

33

34 Studies of functional morphology are commonplace in eukaryotes, with an implicit  
35 understanding that form and function are generally correlated<sup>1-3</sup>. However, such  
36 morphological functional adaptation is largely unstudied in prokaryotes<sup>4</sup>. While explanations  
37 for the functions of rod and coccoid forms have been posited based on scaling and  
38 hydrodynamic arguments<sup>5</sup>, and we know of functions for a limited subset of species-specific  
39 bacterial morphologies from detailed experimental work<sup>6</sup>, we are realistically no closer to a  
40 more general understanding of prokaryote functional morphology than we were a decade  
41 ago<sup>7,8</sup>.

42 One clear and recurring prediction for form and function in microorganisms in general is that  
43 shape and motility are correlated. In bacteria, where the majority of motile species use  
44 flagella to propel themselves, these two traits are commonly thought to co-vary tightly<sup>7,9,10</sup>.  
45 Mathematical modelling of optimal shapes for efficient swimming suggests that flagellar  
46 motility in particular should be an important driving force for bacterial shape evolution.  
47 Specifically, motility imposes substantial physical and energetic constraints<sup>5,9-11</sup> that favour  
48 bacterial cells with ellipsoid or rod-like morphologies within a narrow aspect ratio  
49 (length/width) range to reduce drag<sup>7,9,10</sup>. However, while mathematical models predict a  
50 strong relationship between shape and motility, explicit experimental tests are rare and,  
51 surprisingly, analyses of this relationship in an evolutionary context are lacking entirely.

52 In addition to the requirement of efficient motility for host invasion and colonization in  
53 pathogenic species<sup>12</sup>, flagellar motility is also known to activate strong immune system  
54 responses in mammalian hosts through recognition of flagellar components and movement<sup>13-</sup>  
55 <sup>15</sup>. Selective pressures exerted by host immune responses against flagella have been  
56 suggested to have led many bacteria to lose their ability to move during adaptation to the host

57 habitat<sup>14</sup>. While the role of flagella as a virulence factor in pathogenesis is well  
58 documented<sup>16</sup>, the function of bacterial cell morphology remains elusive.

59 Bacterial shape has been linked to immune evasion and virulence in some disease-causing  
60 groups<sup>7,17-19</sup>. A small number of experimental studies suggest that the host-immune system  
61 can recognise shape and size of artificial particles and could therefore be a strong selective  
62 force acting on cell shape in pathogenic species<sup>20,21</sup>. Recent work has led to the prediction  
63 that the coccoid form is adaptive in a pathogenic context, owing to reduction in cell surface  
64 area exposed to immune attack<sup>22</sup>.

65 Almost everything we understand about the evolution of bacterial shape is based on  
66 qualitative descriptions of morphologies ‘mapped’ on to phylogenetic trees<sup>23,24</sup>. For example,  
67 Siefert and Fox<sup>24</sup> observed that the coccoid form has evolved repeatedly and independently,  
68 as a persistent end-state morphology, in several distinct bacterial groups. Tamames *et al.*<sup>25</sup>,  
69 reported the same observation using arrangement of genes involved in division and cell-wall  
70 synthesis, suggesting that transitions back to a rod shape are unlikely. However, a new  
71 generation of comparative phylogenetic methods now exist, meaning a robust statistical  
72 approach can now be brought to bear on such questions.

73 Here we draw on this new generation of methods to assess the link between form and  
74 function in a large monophyletic group within the *Firmicutes*, a phylum of considerable  
75 environmental, medical, and biotechnological importance (e.g.<sup>26-28</sup>). More specifically, we  
76 investigate the evolutionary associations between shape and motility and whether transitions  
77 from a free-living to a host-associated mode of life were accompanied by a coordinated  
78 change in both traits. As these transitions represent a steep evolutionary hill, coordinated  
79 morphological changes in this context most likely represent adaptations to counter immune  
80 responses, competition, and new nutrient sources.

## 81 **Results**

82 To assess the correlations between different bacterial traits (see Figure 1a and Methods) we  
83 used a recently developed probit model<sup>29</sup> that accommodates binary response variables while  
84 simultaneously accounting for shared ancestry (Methods). The presence of shared ancestry  
85 often biases visual interpretation<sup>30</sup> and accounting for it forms the basis of comparative  
86 phylogenetic methods<sup>31</sup>.

### 87 **Shape and motility evolve independently**

88 Owing to physical and energetic constraints imposed on cell shape by flagellar motility, these  
89 two traits are predicted to co-vary tightly in bacteria<sup>7,9,11</sup>. In contrast to this prediction we  
90 found no evidence for an association between shape and either motility or *mode of life* (free-  
91 living vs non-free living species, see Methods) based on the probit model, despite the  
92 preponderance of rod-shaped motile bacteria (pMCMC<sup>29</sup> = 0.17 and pMCMC = 0.98  
93 respectively). pMCMC is the proportion of coefficients in the posterior distribution estimates  
94 that are  $\neq 0$ , multiplied by two (for a two-tailed test), and is analogous to a frequentist p-  
95 value. This result is robust to a resampling procedure that examines the effect of species  
96 composition in our dataset (pMCMC values  $> 0.05$ , Methods).

97 Given that *mode of life* was not a significant predictor of shape, we used a model with  
98 motility and the subset of lifestyle which constitutes only host-associated species and those  
99 that are free-living as predictors (see Fig. 1a and Methods). This model also provided no  
100 support for an association with shape despite the fact that the host habitat is often assumed to  
101 exert selective pressure on this trait<sup>7</sup> (pMCMC<sub>motility</sub> = 0.31 and pMCMC<sub>lifestyle</sub> = 0.58, Table  
102 1). This result too was robust to our resampling procedure (95% of pMCMC values  $> 0.05$ ,  
103 Methods), and a transition model that models discrete character evolution as a continuous-

104 time Markov process (Methods and Supplementary Table S1) also lends support in that  
105 transitions between free-living and host-associated lifestyles and between a free-living and  
106 non-free living mode of life were not accompanied by change in shape (*mode of life*: BF = -  
107 1.81; *lifestyle*: BF = -1.75, Methods and Supplementary Table S1).

108 An expectation of strong selective pressures exerted by immune system responses on cell  
109 morphology, leads to a prediction of an association between shape and pathogenicity<sup>7,22</sup>.  
110 However, the probit model provided no evidence for any association (pMCMC = 0.32, Table  
111 1), a result robust to our resampling procedure (pMCMC values > 0.05, Methods). Despite  
112 the selective pressures due to habitat and the immune system on shape<sup>7</sup>, our analyses suggest  
113 that selection for shape in this group of bacteria is driven by factors other than the simple  
114 ecological pressures previously assumed.

### 115 **Motility is strongly associated with lifestyle**

116 To investigate whether evolutionary transitions from a free-living to a host-associated  
117 lifestyle were accompanied by a change in motility status, we assessed the association  
118 between motility and our two classifications of habitat (Methods). Our results indicate that  
119 motility is not associated with *mode of life per se* but that motility loss is linked to a host-  
120 associated lifestyle (pMCMC<sub>*lifestyle*</sub> = 0.014, Table 1). This is supported by the strong  
121 rejection of a transition model in which motility and *lifestyle* are assumed to evolve  
122 independently, in favour of a dependent model (log-BF = 9.29, Fig. 1b, Supplementary Table  
123 S1). The most likely transition model thus suggests that transitions from a free-living to a  
124 host-associated lifestyle are often accompanied by a loss of motility (Fig. 1b), and we infer  
125 that selective pressures within the host are likely to have selected against flagellar motility in  
126 host-associated bacteria (Discussion). However, the pMCMC values < 0.05 for the model  
127 with motility as the response and *lifestyle* as predictor after sampling 50% and 75% of the



128 data were 77.4% and 86.8% respectively. This result indicates that the relationship between  
129 motility and lifestyle is not significant, probably due to a reduction of the statistical power in  
130 comparison with the model using all of the data.

### 131 **Transition rates for shape and motility**

132 In agreement with previous observations from phylogenies<sup>24,25</sup> our estimated transition rates  
133 for shape provided no evidence for transitions from coccoid (C) to rod (R) (i.e.,  $q_{CR} = 0$ ) in  
134 the group ( $\log\text{-BF} = -3.46$ , Fig. 1c), indicating the transitions from rod to coccoid are  
135 probably irreversible (Discussion). In contrast, transition rate estimates for motility indicated  
136 that this is a labile character where both loss and regain occur, and with the transition rate  
137 from motile to non-motile approximately six times that of the reverse ( $q_{MN} = 0.6 \pm 0.124$ ,  $Z$   
138  $= 0$  %;  $q_{NM} = 0.12 \pm 0.055$ ,  $Z = 0$  %, Figure 1c). However, while the rate of transitions to  
139 motile from non-motile forms ( $q_{NM}$ ) was low, it was significantly different from a rate of  
140 zero ( $\log\text{-BF} = 8.32$ ). We suggest (Discussion) that lability of flagellar motility is most likely  
141 explained through instances of flagellar resurrection<sup>32</sup> or horizontal gene transfer (HGT)<sup>33-38</sup>.

142 It has been previously posited that bacteria have evolved from a rod shaped ancestor<sup>24,39,40</sup>.  
143 Here we provide statistical support for this hypothesis in this particular group of bacteria as  
144 indicated by a transition model (root posterior probability (rod) =  $0.99 \pm 0.001$ , Fig. 1a). This  
145 model also indicates that this group probably derived from an ancestral motile bacterium  
146 (root posterior probability (motile) =  $0.99 \pm 0.003$ ), (Fig. 1a).

147

148 In agreement with work suggesting that functional traits resulting from complex genetic  
149 machineries are conserved in prokaryotes<sup>41</sup>, we found strong phylogenetic signal in both  
150 shape (mean  $h^2 = 0.84$  with 95 % probability of lying between 0.74 and 0.92, Supplementary

151 Figure S5 and motility (mean  $h^2 = 0.64$  with 95 % probability of lying between 0.40 and 0.86,  
152 Figure S6).

153 **Discussion**

154 **Shape and motility evolved independently**

155 Cell shape and motility are often thought to have important adaptive functions in bacteria<sup>7</sup>.  
156 Based mainly on fluid dynamic arguments, it has been suggested that these two traits co-vary  
157 tightly because of the physical and energetic constraints imposed on cell shape by flagellar  
158 motility<sup>5,7,9</sup>. However, in this study we demonstrate that shape and motility are not  
159 statistically coupled. The lines of evidence we present here suggest that the independent  
160 evolution of motility and shape in this group of bacteria provides a mechanism to allow  
161 greater evolutionary flexibility. Here we draw parallels with analysis of leaf economics and  
162 hydraulic traits in higher plants<sup>42</sup>, where decoupling of suites of traits from each other  
163 suggests that independent trait dimensions can exist. For subtropical forests a leaf economics  
164 dimension corresponding to light capture and tissue longevity, and a hydraulic dimension  
165 corresponding to water-use and leaf temperature maintenance were identified. We suggest  
166 that in the same way that the independent evolution of leaf economics and hydraulic traits  
167 allows more possible plant trait combinations, so independence of shape and motility in  
168 bacteria may allow adaptation to distinct niches. However, in the case of these bacteria there  
169 is a difference in that we observe an evolutionarily irreversible character state (the coccoid  
170 form). The existence of this ‘dead end’ state reduces trait dimensionality somewhat, while the  
171 independent evolution of shape from motility allows at least partial release from this  
172 constraint.

173 The true morphological diversity in the prokaryotes is larger than the simple rod or coccoid  
174 dichotomy used here<sup>7</sup>, with shape complexity that belies a widespread perception that there is  
175 limited morphological variation in groups such as bacteria (e.g.<sup>7</sup>). Given this variation, we  
176 expect morphology in prokaryotes to be finely tuned to function where selection pressures are

177 high, in line with many studies on individual species. Here we provide evolutionary  
178 arguments suggesting that shape is under selective pressure in this monophyletic group. We  
179 provide, to our knowledge, the first statistical support for a rod shaped ancestor of the group  
180 (root posterior distribution =  $99 \pm 0.001$ ) and, as suggested by two previous studies<sup>24,25</sup>,  
181 statistical support for the coccoid shape being a derived end state. This progressive  
182 development of the coccoid shape implies that selective forces are operating<sup>8</sup>. Also, our  
183 analysis suggests that the coccoid shape has evolved several times independently. This  
184 convergence indicates that similar selective forces have led to similar responses across the  
185 group<sup>43</sup>. The complex biochemical machinery, cellular mechanisms and mechanical  
186 constraints involved in rod morphogenesis<sup>44-46</sup> support the idea that transitions from rod to  
187 coccoid are irreversible as our results suggest.

#### 188 **Motility is associated with lifestyle**

189 In contrast to the irreversibility we observe for cell morphology, our results suggest that  
190 flagellar motility in this group is a highly labile character. Although the rate of motility regain  
191 was much lower than that of motility loss, it was still significantly different from zero. Thus,  
192 despite the complex regulatory system involved in flagellar assembly, flagellar motility has  
193 been regained several times, providing complementary evidence for flagellar resurrection<sup>32</sup>,  
194 or horizontal gene transfer (HGT),<sup>33-37</sup> *in natura*.

195 Motility has been suggested to play important roles in dispersal, niche colonization,  
196 predation, desiccation, and chemotaxis under natural conditions<sup>7</sup>, and perhaps unsurprisingly,  
197 it is clearly associated with a free-living lifestyle in this group. Linked to the lability of  
198 flagellar motility we also provide evidence for an association between this trait and habitat  
199 (Table 1), where loss of motility is associated with transitions to a host-associated life-style  
200 (Figure 1b). In agreement with common observations, 83.3 % of free-living bacteria in our

201 dataset were motile while only 8.8 % of those species associated with a host were. It has been  
202 suggested that in the *Staphylococcaceae*, the transition from a free-living mode to a host-  
203 associated habitat coincided with a loss of motility<sup>38</sup>. Flagellar molecular machinery is known  
204 to be targeted by the mammalian immune system via Toll-like receptor 5 (TLR5) and the  
205 membrane spanning protein FLS2 in plants<sup>12,13</sup>, and there are likely to be homologous  
206 immune responses in other animal groups. Such selective pressures are, we suggest, highly  
207 likely to have selected against flagellar motility in host-associated bacteria. In contrast, while  
208 adoption of a coccoid form may confer increased resistance to the host's immune system by  
209 reducing the size of bacterial cells<sup>22</sup>, and the coccoid form has also been suggested to play a  
210 crucial role in pathogenesis<sup>47</sup>, we found no correlation between cell morphology and  
211 pathogenicity (pMCMC = 0.32).

## 212 **Conclusions**

213 We now have strong evidence that shape and motility are not correlated in this large  
214 monophyletic group within the *Firmicutes*, allowing this group to overcome perceived  
215 constraints imposed by irreversible transitions from rod to coccoid morphologies. While we  
216 find a general lack of correlation between shape and lifestyle, there is also little support for  
217 the idea that shape in microorganisms may be untouched by selection, *sensu* Bonner<sup>48</sup>. We  
218 provide evidence that flagellar motility is a highly labile character in the wild, and suggest  
219 that the independent evolution of shape and motility in this group may allow an increase in  
220 bacterial trait dimensions. We think it is likely that such trait independency could be a general  
221 pattern in bacteria as well as for leaf economics and hydraulic traits in plants.

## 222 **Methods**

### 223 **Phylogenetic tree and species selection**

224 In order to account for shared ancestry in our statistical treatment, we used the monophyletic  
225 *Firmicutes* (*Bacilli* and *Erysipelotrichia*) section of the phylogenetic tree of Chai *et al.*<sup>49</sup>,  
226 based on 14,727 prokaryotic genomes (see Fig. 1a).

### 227 **Data collection**

228 We collected phenotypic data on shape, motility, pathogenicity and lifestyle type for 325  
229 species of the *Firmicutes* from *Bergey's Manual of Systematic Bacteriology*<sup>50</sup>. Data for  
230 species described after the manual was published were collated from the primary literature  
231 (Supplementary Table S3). Data were not reported for taxa without a species description (e.g.  
232 *Bacillus* sp. 1NLA3E, *Streptococcus* sp. GMD2S). Data for outlier strains (i.e. potentially  
233 misclassified species), were not included in the analysis (e.g. *Clostridium difficile* strain P28  
234 did not cluster with members of the genus *Clostridium*). The tree and phylogenetic  
235 distribution of phenotypic data are shown in Figure 1a.

### 236 **Phenotypic characterization**

#### 237 **Cell morphology**

238 Shape characterization in the species description section from *Bergey's Manual of Systematic*  
239 *Bacteriology* and the primary literature is generally subjective and not geometrically precise.  
240 To provide a more reliable description for the purpose of our study we classified shape based  
241 mainly on size measurements of individual cells. When cell size was not available we used a  
242 simplified classification.

243 We first reported cell length and width (diameter for coccoid cells) and calculated the aspect  
244 ratio (AR) as length divided by width. AR of species for which a range of width and length  
245 was provided was calculated as the average length divided by the average width. We defined  
246 as rod-shaped any cylindrical cell with an  $AR > 1$  (blue tips in Fig. 1a) and as coccoid any  
247 cell with  $AR = 1$  (red tips in Fig. 1a). Pleomorphic species for which width and length data  
248 were provided (seven species with  $AR > 1$  and two with  $AR = 1$ ) were considered as missing  
249 for shape as these data were reported only for the cells that were rod or coccoid among other  
250 morphologies. Species for which length and/or width was missing were classified as being  
251 either rod or coccoid based on a qualitative description. The descriptions of shape in Bergey's  
252 Manual and the primary literature were usually words and phrases such as "rods", "rods with  
253 rounded ends", "straight rods", "curved rods", "slightly curved rods" and "rods with tapered  
254 ends", "cocci", "spherical", "coccoid", "and ovoid" or "ovococcoid". Based on these  
255 descriptions, we recorded the shape as coccoid for "cocci", "spherical", "coccoid", "ovoid"  
256 and "ovococcoid", and as rod for the remaining categories. We did not consider the curvature  
257 or end type for the rod categories. Ovococcoid species with  $AR > 1$  were excluded from the  
258 analyses. Species with ambiguous shapes (e.g. "ovoid or rods", "cocci or rods") were  
259 excluded from the analysis.

## 260 **Motility**

261 Species were classified as motile or non-motile regardless of motility type (e.g. swarming or  
262 swimming motility). Species exhibiting changes in motility status depending on growth  
263 conditions were recorded as being motile only if the presence of flagella was reported. Motile  
264 and non-motile species are coloured in light green and orange respectively on the inner ring  
265 in Fig. 1a.

## 266 **Habitat and pathogenicity**

267 For the purpose of this study and due to limited information on microenvironments we used a  
268 broad categorization (based on macro-environment descriptions) of habitat types (i.e. the  
269 different locations where the organism naturally lives and grows and from which it could be  
270 recovered and isolated). When the habitat was not known, the first isolation site was used  
271 (e.g. human tissue, soil, etc.). This categorization was a simple division between free-living  
272 (i.e. bacteria living independently in the environment) and non-free-living species (*mode of*  
273 *life* dataset and middle ring in Fig. 1a). Bacteria living in soil, water, lake, sea or sediment,  
274 for instance, were considered as free-living. Species associated with plant, animal or insect  
275 organisms and species living in confined environments (e.g. food production and  
276 fermentation processes) were recorded as non-free living.

277 To investigate whether host-associated species exhibit a particular morphology and motility  
278 status in comparison to free-living species, we used a subset of our data containing free-living  
279 species and only those non-free-living species associated with a host (*lifestyle* dataset, n =  
280 145 species and outer ring in Fig. 1a). Host-associated species were defined as those living  
281 within a plant, animal or insect hosts while species associated with food production and  
282 fermentation processes were not considered in this classification. Species living in multiple  
283 environments (e.g. human tissues, food and soil) were recorded as missing for both datasets.

284 To test for a correlation between shape and pathogenicity we took the host-associated species  
285 and classified them as either pathogenic or non-pathogenic. We defined pathogenicity as the  
286 capacity to cause disease. Opportunistic and obligate plant, animal and insect pathogens were  
287 considered as pathogenic while commensal species and those not yet reported as being  
288 involved in host infections were considered as non-pathogenic. Species for which  
289 pathogenicity information was not available were not included in the analysis. As data on



290 pathogenesis were only available for two of 12 motile host-associated species we did not  
291 include motility in this analysis.

## 292 **Phylogenetic comparative methods**

### 293 **Probit model**

294 We modelled the probability of a correlation between our response variable (shape or  
295 motility) and our predictors using phylogenetic generalised linear mixed models in a  
296 Bayesian framework<sup>29</sup>. We used this type of model as it allows testing models with binary  
297 response variables while accounting for shared ancestry as implied by the phylogeny. We  
298 also used the more familiar Markov transition model developed by Pagel<sup>31</sup> in a Bayesian  
299 framework<sup>51</sup>.

300 Shape, motility and lifestyle data were coded as discrete binary characters (rod, motile and  
301 free-living as 1 and coccoid, non-motile and non-free living or host-associated as 0). We used  
302 a probit model in MCMCglmm<sup>29</sup> with largely uninformative priors (normal distribution with  
303 a mean of zero and a variance of  $10^8$ ) for our fixed factor predictors, and a  $\chi^2$  prior for the  
304 phylogeny treated as a random factor as this best approximates a uniform distribution<sup>29,52</sup>. As  
305 binary response variables do not provide sufficient information for estimating the residual  
306 variance, we fixed the residual variance to 1<sup>29,52</sup>. The MCMC (Markov chain Monte Carlo)  
307 chains were run for 5 million iterations with an additional burn-in of 300,000 iterations and a  
308 sampling interval of 1000 iterations. Chain convergence and mixing were assessed visually  
309 (Supplementary Fig. S1-S4) as well as by ensuring that the effective sample sizes for all  
310 estimated parameters were  $> 1000$ . To assess the autocorrelation for the sampling factor we  
311 checked that all correlation between samples after lag zero was less than 0.1<sup>29</sup>.

312

### 313 **Assessing robustness of multiple regression results using MCMCglmm**

314 To test whether our results from multiple regressions were robust, we applied a cross-  
315 validation test. We ran 500 independent chains by sampling 50% of the data in each run for  
316 all the models. For regression model with motility and *lifestyle* we also performed an  
317 additional run of 500 independent chains by sampling 75% of the data due to a decrease in  
318 the statistical power when sampling only 50% of the data. Chains mixing was assessed  
319 visually and percentage of pMCMC values below 0.05 among 500 samples for each model is  
320 reported (Results section). To account for multiple testing for our hypotheses regarding  
321 motility we performed a False Discovery Rate (FDR) test<sup>53</sup>.

322

### 323 **Phylogenetic signal**

324 We used the estimated posterior heritability ( $h^2$ ) of our models as a measure of the degree of  
325 the phylogenetic signal in our data, a parameter that is equivalent to  $\lambda$ <sup>54</sup> in phylogenetic  
326 generalised least-squares models<sup>55</sup>. We used a Bayesian approach to take into account the  
327 uncertainty in model parameter estimation and calculated the posterior heritability across the  
328 entire posterior distribution of model variances.

### 329 **Transition model**

330 To assess whether transitions from a free-living lifestyle to being host-adapted were  
331 associated with a change in shape or motility status we used a transition model under a  
332 reversible-jump MCMC approach as implemented in BayesTraits v2<sup>51</sup>, by comparing two  
333 competing models. The first (independent) model assumes that two characters evolve  
334 independently while the second (dependent) model allows one character to vary depending on

335 the character state of the other. For an effective estimate of the marginal likelihoods we used  
336 three independent chains run for 5,000,000 generations after discarding the first 10% as  
337 burning period and the stepping stone sampling procedure (1,000 stones, each sampled for  
338 20,000 iterations ) implemented in BayesTraits v2. Chain convergence was assessed using  
339 Tracer v1.6.03<sup>56</sup>. The models were evaluated by two methods. First by comparing the  
340 marginal likelihood of the two models using Bayes factor (BF). Second, given that the  
341 number of visits to the dependent or independent model is proportional to the posterior  
342 probability of the model, support for correlated evolution was evaluated by comparing the  
343 ratio of prior and posterior odds for visits of the two models during the chains<sup>1</sup>. For both  
344 methods a  $\log\text{-BF} < 2$  was considered as weak evidence for correlated evolution.

345 To estimate the transition rates for discrete phenotypic characters and to assess whether the  
346 rates were asymmetric we modelled discrete character evolution as a continuous-time  
347 Markov process using the multistate method in BayesTraits v2. All models were run for  
348 5,000,000 iterations (sampled every 1,000 iterations) with all priors set to an exponential with  
349 a mean of 10. Marginal likelihoods were obtained from the harmonic mean estimates of the  
350 model. Where strong asymmetry was detected, we then compared a constrained with a full  
351 model in order to assess whether low transition rates differed significantly from zero rates. In  
352 the constrained model, the transition rate from state 0 to 1 (reversal) was fixed to zero ( $q_{01} =$   
353  $0$ ), while the full model estimated both parameters simultaneously ( $q_{01} \neq q_{10}$ ). To identify the  
354 best-fitting model, we compared the log marginal likelihoods obtained from estimates for the  
355 two models using BF. A  $\log\text{-BF} < 2$  was considered as a weak support<sup>57</sup> for the model where  
356 the rates are different ( $q_{01} \neq q_{10}$ ).

357

358 **References**

- 359 1. Díaz, S. *et al.* The global spectrum of plant form and function. *Nature* **529**, 167–171  
360 (2016).
- 361 2. Hale, M. S. & Mitchell, J. G. Functional morphology of diatom frustule  
362 microstructures: Hydrodynamic control of brownian particle diffusion and advection.  
363 *Aquat. Microb. Ecol.* **24**, 287–295 (2001).
- 364 3. Wainwright, P. C. Functional Versus Morphological Diversity in Macroevolution.  
365 *Annu. Rev. Ecol. Evol. Syst.* **38**, 381–401 (2007).
- 366 4. Martiny, J. B. H., Jones, S. E., Lennon, J. T. & Martiny, A. C. Microbiomes in light of  
367 traits: A phylogenetic perspective. *Science (80-. )*. **350**, aac9323 (2015).
- 368 5. Dusenbery, D. B. *Living at micro scale: the unexpected physics of being small*.  
369 (Harvard University Press, 2009).
- 370 6. Persat, A., Stone, H. a & Gitai, Z. The curved shape of *Caulobacter crescentus*  
371 enhances surface colonization in flow. *Nat. Commun.* **5**, 3824 (2014).
- 372 7. Young, K. D. The selective value of bacterial shape. *Microbiol. Mol. Biol. Rev.* **70**,  
373 660–703 (2006).
- 374 8. Young, K. D. Bacterial morphology: why have different shapes? *Curr. Opin.*  
375 *Microbiol.* **10**, 596–600 (2007).
- 376 9. Mitchell, J. G. The energetics and scaling of search strategies in bacteria. *Am. Nat.*  
377 **160**, 727–740 (2002).
- 378 10. Cooper, S. & Denny, M. W. A conjecture on the relationship of bacterial shape to  
379 motility in rod-shaped bacteria. *FEMS Microbiol. Lett.* **148**, 227–231 (1997).
- 380 11. Dusenbery, D. B. Fitness landscapes for effects of shape on chemotaxis and other  
381 behaviors of bacteria. *J. Bacteriol.* **180**, 5978–5983 (1998).
- 382 12. Ramos, H. C., Rumbo, M. & Sirard, J. C. Bacterial flagellins: Mediators of

- 383 pathogenicity and host immune responses in mucosa. *Trends Microbiol.* **12**, 509–517  
384 (2004).
- 385 13. Cullender, T. C. *et al.* Innate and adaptive immunity interact to quench microbiome  
386 flagellar motility in the gut. *Cell Host Microbe* **14**, 571–581 (2013).
- 387 14. Lovewell, R. R. *et al.* Step-wise loss of bacterial flagellar torsion confers progressive  
388 phagocytic evasion. *PLoS Pathog.* **7**, (2011).
- 389 15. Patankar, Y. R. *et al.* Flagellar motility is a key determinant of the magnitude of the  
390 inflammasome response to *Pseudomonas aeruginosa*. *Infect. Immun.* **81**, 2043–2052  
391 (2013).
- 392 16. Chaban, B., Hughes, H. V. & Beeby, M. The flagellum in bacterial pathogens: for  
393 motility and a whole lot more. *Semin. Cell Dev. Biol.* **46**, 91–103 (2015).
- 394 17. Dalia, A. B. & Weiser, J. N. Minimization of bacterial size Allows for complement  
395 evasion and Is overcome by the agglutinating effect of antibody. *Cell Host Microbe*  
396 **10**, 486–496 (2011).
- 397 18. Fridrich, E. & Gaynor, E. C. Peptidoglycan hydrolases, bacterial shape, and  
398 pathogenesis. *Curr. Opin. Microbiol.* **16**, 767–778 (2013).
- 399 19. Sycuro, L. K. *et al.* Multiple peptidoglycan modification networks modulate  
400 helicobacter pylori's cell shape, motility, and colonization potential. *PLoS Pathog.* **8**,  
401 (2012).
- 402 20. Champion, J. a & Mitragotri, S. Role of target geometry in phagocytosis. *Proc. Natl.*  
403 *Acad. Sci. U. S. A.* **103**, 4930–4934 (2006).
- 404 21. Doshi, N. & Mitragotri, S. Macrophages recognize size and shape of their targets.  
405 *PLoS One* **5**, 1–6 (2010).
- 406 22. Veyrier, F. J. *et al.* Common Cell Shape Evolution of Two Nasopharyngeal Pathogens.  
407 *PLOS Genet.* **11**, e1005338 (2015).

- 408 23. Stackebrandt, E. & Woese, C. R. A phylogenetic dissection of the family  
409 micrococcaceae. *Curr. Microbiol.* **2**, 317–322 (1979).
- 410 24. Siefert, J. L. & Fox, G. E. Phylogenetic mapping of bacterial morphology.  
411 *Microbiology* **144**, 2803–2808 (1998).
- 412 25. Tamames, J., González-Moreno, M., Mingorance, J., Valencia, a. & Vicente, M.  
413 Bringing gene order into bacterial shape. *Trends Genet.* **17**, 124–126 (2001).
- 414 26. Ley, R., Turnbaugh, P., Klein, S. & Gordon, J. Microbial ecology: human gut  
415 microbes associated with obesity. *Nature* **444**, 1022–3 (2006).
- 416 27. Wrighton, K. C. *et al.* A novel ecological role of the Firmicutes identified in  
417 thermophilic microbial fuel cells. *ISME J.* **2**, 1146–1156 (2008).
- 418 28. Sharmin, F., Wakelin, S., Huygens, F. & Hargreaves, M. Firmicutes dominate the  
419 bacterial taxa within sugar-cane processing plants. *Sci. Rep.* **3**, 3107 (2013).
- 420 29. Hadfield, J. D. MCMC Methods for Multi-Response Generalized Linear Mixed  
421 Models: The MCMCglmm R Package. *J. Stat. Softw.* **33**, 1–22 (2010).
- 422 30. Maddison, W. P. & FitzJohn, R. G. The Unsolved Challenge to Phylogenetic  
423 Correlation Tests for Categorical Characters. *Syst. Biol.* **64**, 127–136 (2015).
- 424 31. Pagel, M. Detecting Correlated Evolution on Phylogenies: A General Method for the  
425 Comparative Analysis of Discrete Characters. *Proc. R. Soc. London B Biol. Sci.* **255**,  
426 37–45 (1994).
- 427 32. Taylor, T. B. *et al.* Evolutionary resurrection of flagellar motility via rewiring of the  
428 nitrogen regulation system. *Science (80-. )*. **347**, 1014–1017 (2015).
- 429 33. Chiara, M. *et al.* Comparative genomics of *Listeria sensu lato* : genus-wide differences  
430 in evolutionary dynamics and the progressive gain of complex, potentially  
431 pathogenicity-related traits through lateral gene transfer. *Genome Biol. Evol.* **7**, evv131  
432 (2015).

- 433 34. Cousin, F. J. *et al.* Detection and Genomic Characterization of Motility in  
434 *Lactobacillus curvatus*: Confirmation of Motility in a Species outside the *Lactobacillus*  
435 *salivarius* Clade. *Appl. Environ. Microbiol.* **81**, 1297–1308 (2015).
- 436 35. Palmer, K. L., Schaik, W. Van, Willems, R. J. L. & Gilmore, M. S. Enterococcal  
437 Genomics. *E-Book* (2014). doi:NBK190425 [bookaccession]
- 438 36. Mendes-Soares, H., Suzuki, H., Hickey, R. J. & Forney, L. J. Comparative functional  
439 genomics of *Lactobacillus* spp. reveals possible mechanisms for specialization of  
440 vaginal lactobacilli to their environment. *J. Bacteriol.* **196**, 1458–1470 (2014).
- 441 37. Poggio, S. *et al.* A complete set of flagellar genes acquired by horizontal transfer  
442 coexists with the endogenous flagellar system in *Rhodobacter sphaeroides*. *J.*  
443 *Bacteriol.* **189**, 3208–3216 (2007).
- 444 38. Shah, N. *et al.* Reductive evolution and the loss of PDC/PAS domains from the genus  
445 *Staphylococcus*. *BMC Genomics* **14**, 524 (2013).
- 446 39. Koch, A. L. Were Gram-positive rods the first bacteria? *Trends Microbiol.* **11**, 166–  
447 170 (2003).
- 448 40. Errington, J. L-form bacteria, cell walls and the origins of life. *Open Biol.* **3**, 120143  
449 (2013).
- 450 41. Martiny, A. C., Treseder, K. & Pusch, G. Phylogenetic conservatism of functional  
451 traits in microorganisms. *ISME J.* **7**, 830–8 (2013).
- 452 42. Li, L. *et al.* Leaf economics and hydraulic traits are decoupled in five species-rich  
453 tropical-subtropical forests. *Ecol. Lett.* n/a–n/a (2015). doi:10.1111/ele.12466
- 454 43. Pagel, M. in *Phylogenetics and Ecology* (eds. Eggleton, P. & Richard, V.-W.) 29–51  
455 (Linnean Society Symposium Series, 1994).
- 456 44. Chang, F. & Huang, K. C. How and why cells grow as rods. *BMC Biol.* **12**, 54 (2014).
- 457 45. Jiang, C., Caccamo, P. D. & Brun, Y. V. Mechanisms of bacterial morphogenesis:

- 458 Evolutionary cell biology approaches provide new insights. *BioEssays* n/a–n/a (2015).  
459 doi:10.1002/bies.201400098
- 460 46. Randich, A. M. & Brun, Y. V. Molecular mechanisms for the evolution of bacterial  
461 morphologies and growth modes. *Front. Microbiol.* **6**, 1–13 (2015).
- 462 47. Dworkin, J. Form equals function? Bacterial shape and its consequences for  
463 pathogenesis. *Mol. Microbiol.* **78**, 792–795 (2010).
- 464 48. Bonner, J. T. *Randomness in evolution*. (Princeton University Press, 2013).
- 465 49. Chai, J., Kora, G., Ahn, T.-H., Hyatt, D. & Pan, C. Functional phylogenomics analysis  
466 of bacteria and archaea using consistent genome annotation with UniFam. *BMC Evol.*  
467 *Biol.* **14**, 1–13 (2014).
- 468 50. Vos, P. *et al.* *Bergey's Manual of Systematic Bacteriology - Vol 3: The Firmicutes*.  
469 *Springer-Verlag New York Inc.* (2009). doi:10.1007/b92997
- 470 51. Pagel, M., Meade, A. & Barker, D. Bayesian estimation of ancestral character states on  
471 phylogenies. *Syst. Biol.* **53**, 673–684 (2004).
- 472 52. de Villemereuil, P., Gimenez, O. & Doligez, B. Comparing parent-offspring regression  
473 with frequentist and Bayesian animal models to estimate heritability in wild  
474 populations: a simulation study for Gaussian and binary traits. *Methods Ecol. Evol.* **4**,  
475 260–275 (2013).
- 476 53. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and  
477 powerful approach to multiple testing. *Journal of the Royal Statistical Society* **57**, 289–  
478 300 (1995).
- 479 54. Pagel, M. Inferring the historical patterns of biological evolution. *Nature* **401**, 877–884  
480 (1999).
- 481 55. Hadfield, J. D. & Nakagawa, S. General quantitative genetic methods for comparative  
482 biology: Phylogenies, taxonomies and multi-trait models for continuous and



- 483 categorical characters. *J. Evol. Biol.* **23**, 494–508 (2010).
- 484 56. Rambaut, A. & Drummond, A. J. Tracer v1.6. Available from  
485 <http://tree.bio.ed.ac.uk/software/tracer/> (2013).
- 486 57. Kass, R. E. & Raftery, A. E. Bayes Factors. *J. Am. Stat. Assoc.* **90**, 773–795  
487 (1995).
- 488
- 489

490 **Corresponding author: [felbaidouri@lincoln.ac.uk](mailto:felbaidouri@lincoln.ac.uk)**

491 **Acknowledgment:**

492 We thank Mark Pagel, Oscar Guadayol Roig, Rudi Schuech, Joanna Baker and Andrew  
493 Meade for helpful comments and advice. This work is financially supported by The  
494 Leverhulme Trust project RLA RL-2012-022, "Form and function in a microbial world",  
495 granted to SH. CV was supported by a Leverhulme Trust Research Project Grant (RPG-2013-  
496 185).

497 **Author contributions:**

498 SH, CV and FE designed the study; FE and SH developed the protocol for the data collection;  
499 FE collected the data; FE, CV and SH analysed the data; FE wrote the first draft of the  
500 manuscript, and all authors contributed substantially to revisions.

501 **Competing financial interests:** The authors declare no competing financial interests.

502

503 **Figure Legends**

504 **Figure 1.** (a) Phylogenetic tree and distribution of traits. The inner and middle rings are  
505 colour coded according to motility status and *mode of life* respectively, while the outer ring is  
506 coded according to *lifestyle*. Histograms show the posterior distribution of probability at the  
507 root to be a rod (top) and motile (bottom). (b) Transition rate estimates between motility and  
508 *lifestyle*. The  $q_{ij}$  transition rates denote changes in one trait that are dependent on the state of  
509 the other trait.  $i$  and  $j$  take trait values 1 to 4 corresponding to 1 = non-motile, 2 = motile, 3 =  
510 host-associated and 4 = free-living. Histograms on the arrows indicate the posterior  
511 distribution of transition rates under a reversible-jump MCMC model for a given transition  
512 from one state to another (arrow). On the left is a model where all transition rates were  
513 estimated, on the right one where transition rates  $q_{12}$ ,  $q_{21}$  and  $q_{13}$  were set to zero (First  
514 model in Supplementary Table S2). (c) Transition rates for shape and motility. Histograms  
515 indicate the posterior distribution of transition rates estimates between rod and coccoid ( $q_{RC}$   
516 and  $q_{CR}$ ) and between being motile and non-motile ( $q_{MN}$  and  $q_{NM}$ ).

517

518 **Table 1.**

Model	Posterior mean	Lower 95 % CI*	Upper 95 % CI	pMCMC	HPD† Lower	HPD Upper
<b>Response: Shape</b>					0.74	0.92
(Intercept)	6.22	0.68	11.97	<b>0.016</b>		
Motility	2.08	-0.93	5.26	0.17		
<i>mode of life</i>	-0.06	-2.98	2.98	0.98		
<b>Response: Shape</b>					0.79	0.93
(Intercept)	4.96	-0.1	10.31	<b>0.04</b>		
Motility	1.54	-1.8	4.42	0.31		
<i>lifestyle</i>	0.77	-2.25	3.73	0.58		
<b>Response: Shape</b>					0.5	0.88
(Intercept)	3.64	-1.1	9.03	0.11		
Pathogenicity	-2.7	-6.67	1.15	0.146		
<b>Response: Motility</b>					0.53	0.88
(Intercept)	0.99	-1.77	3.82	0.45		
<i>mode of life</i>	1.1	-0.17	2.28	0.09		
<b>Response: Motility</b>					0.4	0.86
(Intercept)	0.21	-2.38	2.62	0.89		
<i>lifestyle</i>	1.9	0.35	3.4	<b>0.028</b> <b>(0.014)</b>		

519

520

521 \* CI: Credible interval

522 † HPD: 95% credible interval for heritability

523 Our conclusions were not affected by multiple testing by using the false discovery rate (FDR) control test<sup>53</sup>

524 (Methods). Corrected pMCMC values are given with originals in brackets only where a significant result was

525 influenced.

526

527 **Table 1. MCMCglmm results for the different models.** For each model we report the

528 posterior mean, the 95 % credible interval, the pMCMC values and the 95 % credible interval

529 for heritability. Significant (< 0.05) pMCMC values are in bold.

530