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

Article

Accepted Version

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

Fouad El Baidouri[1*], Chris Venditti[2] & Stuart Humphries[1]

Author affiliation:

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

shumphries@lincoln.ac.uk

[1] School of Life Sciences University of Lincoln, Joseph Banks Laboratories, Green Lane, Lincoln, LN6 7DL, UK.

Chris Venditti

c.d.venditti@reading.ac.uk

[2] School of Biological Sciences, University of Reading, Reading RG6 6BX, UK.

* **Corresponding author:** Fouad El Baidouri

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

Tel: +(44) 01522887420

felbaidouri@lincoln.ac.uk

ORCID ID: 0000-0001-5204-6244

Abstract

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

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

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

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

habitat[14]. While the role of flagella as a virulence factor in pathogenesis is well documented[16], the function of bacterial cell morphology remains elusive.

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

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

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

Results

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

Shape and motility evolve independently

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

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

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

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

Motility is strongly associated with lifestyle

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

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

Transition rates for shape and motility

In agreement with previous observations from phylogenies[24,25] our estimated transition rates for shape provided no evidence for transitions from coccoid (C) to rod (R) (i.e., qCR = 0) in the group (log-BF = -3.46, Fig. 1c), indicating the transitions from rod to coccoid are probably irreversible (Discussion). In contrast, transition rate estimates for motility indicated that this is a labile character where both loss and regain occur, and with the transition rate from motile to non-motile approximately six times that of the reverse (qMN = 0.6 ± 0.124, Z = 0 %; qNM = 0.12 ± 0.055, Z = 0 %, Figure 1c). However, while the rate of transitions to motile from non-motile forms (qNM) was low, it was significantly different from a rate of zero (log-BF = 8.32). We suggest (Discussion) that lability of flagellar motility is most likely explained through instances of flagellar resurrection[32] or horizontal gene transfer (HGT)[33–38].

It has been previously posited that bacteria have evolved from a rod shaped ancestor[24,39,40]. Here we provide statistical support for this hypothesis in this particular group of bacteria as indicated by a transition model (root posterior probability (rod) = 0.99±0.001, Fig. 1a). This model also indicates that this group probably derived from an ancestral motile bacterium (root posterior probability (motile) = 0.99±0.003), (Fig. 1a).

In agreement with work suggesting that functional traits resulting from complex genetic machineries are conserved in prokaryotes[41], we found strong phylogenetic signal in both 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).

Discussion

Shape and motility evolved independently

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

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

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

Motility is associated with lifestyle

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

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

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

Conclusions

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

Methods

Phylogenetic tree and species selection

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

Data collection

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

Phenotypic characterization

Cell morphology

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

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

Motility

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

Habitat and pathogenicity

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

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

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

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

Phylogenetic comparative methods

Probit model

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

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

Assessing robustness of multiple regression results using MCMCglmm

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

Phylogenetic signal

We used the estimated posterior heritability (h^2) of our models as a measure of the degree of the phylogenetic signal in our data, a parameter that is equivalent to λ[54] in phylogenetic generalised least-squares models[55]. We used a Bayesian approach to take into account the uncertainty in model parameter estimation and calculated the posterior heritability across the entire posterior distribution of model variances.

Transition model

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

the character state of the other. For an effective estimate of the marginal likelihoods we used three independent chains run for 5,000,000 generations after discarding the first 10% as burning period and the stepping stone sampling procedure (1,000 stones, each sampled for 20,000 iterations) implemented in BayesTraits v2. Chain convergence was assessed using Tracer v1.6.03[56]. The models were evaluated by two methods. First by comparing the marginal likelihood of the two models using Bayes factor (BF). Second, given that the number of visits to the dependent or independent model is propositional to the posterior probability of the model, support for correlated evolution was evaluated by comparing the ratio of prior and posterior odds for visits of the two models during the chains1. For both methods a log-BF < 2 was considered as weak evidence for correlated evolution.

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

References

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

pathogenicity and host immune responses in mucosa. *Trends Microbiol.* **12,** 509–517 (2004).

13. Cullender, T. C. *et al.* Innate and adaptive immunity interact to quench microbiome flagellar motility in the gut. *Cell Host Microbe* **14,** 571–581 (2013).

14. Lovewell, R. R. *et al.* Step-wise loss of bacterial flagellar torsion confers progressive phagocytic evasion. *PLoS Pathog.* **7,** (2011).

15. Patankar, Y. R. *et al.* Flagellar motility is a key determinant of the magnitude of the inflammasome response to Pseudomonas aeruginosa. *Infect. Immun.* **81,** 2043–2052 (2013).

16. Chaban, B., Hughes, H. V. & Beeby, M. The flagellum in bacterial pathogens: for motility and a whole lot more. *Semin. Cell Dev. Biol.* **46,** 91–103 (2015).

17. Dalia, A. B. & Weiser, J. N. Minimization of bacterial size Allows for complement evasion and Is overcome by the agglutinating effect of antibody. *Cell Host Microbe* **10,** 486–496 (2011).

18. Frirdich, E. & Gaynor, E. C. Peptidoglycan hydrolases, bacterial shape, and pathogenesis. *Curr. Opin. Microbiol.* **16,** 767–778 (2013).

19. Sycuro, L. K. *et al.* Multiple peptidoglycan modification networks modulate helicobacter pylori's cell shape, motility, and colonization potential. *PLoS Pathog.* **8,** (2012).

20. Champion, J. a & Mitragotri, S. Role of target geometry in phagocytosis. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 4930–4934 (2006).

21. Doshi, N. & Mitragotri, S. Macrophages recognize size and shape of their targets. *PLoS One* **5,** 1–6 (2010).

22. Veyrier, F. J. *et al.* Common Cell Shape Evolution of Two Nasopharyngeal Pathogens. *PLOS Genet.* **11,** e1005338 (2015).

23. Stackebrandt, E. & Woese, C. R. A phylogenetic dissection of the family micrococcaceae. *Curr. Microbiol.* **2,** 317–322 (1979).

24. Siefert, J. L. & Fox, G. E. Phylogenetic mapping of bacterial morphology. *Microbiology* **144,** 2803–2808 (1998).

25. Tamames, J., González-Moreno, M., Mingorance, J., Valencia, a. & Vicente, M. Bringing gene order into bacterial shape. *Trends Genet.* **17,** 124–126 (2001).

26. Ley, R., Turnbaugh, P., Klein, S. & Gordon, J. Microbial ecology: human gut microbes associated with obesity. *Nature* **444,** 1022–3 (2006).

27. Wrighton, K. C. *et al.* A novel ecological role of the Firmicutes identified in thermophilic microbial fuel cells. *ISME J.* **2,** 1146–1156 (2008).

28. Sharmin, F., Wakelin, S., Huygens, F. & Hargreaves, M. Firmicutes dominate the bacterial taxa within sugar-cane processing plants. *Sci. Rep.* **3,** 3107 (2013).

29. Hadfield, J. D. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *J. Stat. Softw.* **33,** 1–22 (2010).

30. Maddison, W. P. & FitzJohn, R. G. The Unsolved Challenge to Phylogenetic Correlation Tests for Categorical Characters. *Syst. Biol.* **64,** 127–136 (2015).

31. Pagel, M. Detecting Correlated Evolution on Phylogenies: A General Method for the Comparative Analysis of Discrete Characters. *Proc. R. Soc. London B Biol. Sci.* **255,** 37–45 (1994).

32. Taylor, T. B. *et al.* Evolutionary resurrection of flagellar motility via rewiring of the nitrogen regulation system. *Science (80-.).* **347,** 1014–1017 (2015).

33. Chiara, M. *et al.* Comparative genomics of Listeria *sensu lato* : genus-wide differences in evolutionary dynamics and the progressive gain of complex, potentially pathogenicity-related traits through lateral gene transfer. *Genome Biol. Evol.* **7,** evv131 (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. & Forneya, 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 Socity 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

categorical characters. *J. Evol. Biol.* **23,** 494–508 (2010).

56. Rambaut, A. & Drummond, A. J. Tracer v1.6. *Available from http://tree.bio.ed.ac.uk/software/tracer/* (2013).

57. Kass, R. E. & Raftery, A. E. Bayes Factors. *J. Am. Stat. Assoc.* **90430,** 773–795 (1995).

Corresponding author: felbaidouri@lincoln.ac.uk

Acknowledgment:

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

Author contributions:

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

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

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 q12, q21 and q13 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 (qRC
516 and qCR) and between being motile and non-motile (qMN and qNM).
517

518 **Table 1.**

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

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[53]

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