doi:10.1111/evo.12597

A generation-time effect on the rate of molecular evolution in bacteria

Cory Weller[1] and Martin Wu[1,2]

[1]*Department of Biology, University of Virginia, Charlottesville, Virginia 22904*

[2]*E-mail: mw4yv@virginia.edu*

Received September 13, 2014
Accepted December 17, 2014

Molecular evolutionary rate varies significantly among species and a strict global molecular clock has been rejected across the tree of life. Generation time is one primary life-history trait that influences the molecular evolutionary rate. Theory predicts that organisms with shorter generation times evolve faster because of the accumulation of more DNA replication errors per unit time. Although the generation-time effect has been demonstrated consistently in plants and animals, the evidence of its existence in bacteria is lacking. The bacterial phylum *Firmicutes* offers an excellent system for testing generation-time effect because some of its members can enter a dormant, nonreproductive endospore state in response to harsh environmental conditions. It follows that spore-forming bacteria would—with their longer generation times—evolve more slowly than their nonspore-forming relatives. It is therefore surprising that a previous study found no generation-time effect in *Firmicutes*. Using a phylogenetic comparative approach and leveraging on a large number of *Firmicutes* genomes, we found sporulation significantly reduces the genome-wide spontaneous DNA mutation rate and protein evolutionary rate. Contrary to the previous study, our results provide strong evidence that the evolutionary rates of bacteria, like those of plants and animals, are influenced by generation time.

KEY WORDS: Evolutionary rate, *Firmicutes*, sporulation.

The rate of molecular evolution varies significantly between species and much evidence has rejected the existence of a strict global molecular clock across the tree of life (Bousquet et al. 1992; Thomas et al. 2006; Welch et al. 2008; Kuo and Ochman 2009). It has been demonstrated that organism-level traits such as life history could influence the molecular evolutionary rate in eukaryotes, chief among them is the generation time. The generation-time hypothesis states that organisms with shorter generation time evolve faster, as they copy their genomes more frequently and therefore have more DNA replication errors per unit time.

The generation-time effect has been repeatedly observed in both animals and plants (Laird et al. 1969; Kohne 1970; Li and Tanimura 1987; Ohta 1993; Mooers and Harvey 1994; Bromham et al. 1996; Gaut et al. 1996, 1997; Laroche et al. 1997; Ainouche and Bayer 1999; Laroche and Bousquet 1999; Andreasen and Baldwin 2001; Bromham 2002; Nabholz et al. 2008; Smith and Donoghue 2008; Welch et al. 2008; Thomas et al. 2010). For example, molecular evolutionary rates in rodents were much faster than those of primates (Laird et al. 1969; Kohne 1970). Perennials, with longer generation times, have been shown to accumulate substitutions more slowly than rapidly maturing annual plants (Soria-Hernanz et al. 2008; Yue et al. 2010; Buschiazzo et al. 2012; but see Whittle and Johnston 2003).

It is less clear whether there is a generation-time effect on the rate of molecular evolution in bacteria. Testing the generation-time hypothesis in natural bacterial populations can be difficult. Unlike animals and plants that have relatively fixed generation time, bacteria populations of the same species can have highly variable generation times depending on the growth condition. It is therefore difficult to accurately estimate the generation time of bacteria in natural environments, as nutrients and other important environmental factors for growth (e.g., temperature, salinity) vary widely in time and space. Certain members of the bacteria phylum *Firmicutes* are capable of entering an encapsulated, dormant, and nonreproductive state known as an endospore. This state allows for bacteria to withstand and survive extreme conditions, such as ultraviolet radiation, desiccation, heat, and lack of nutrients. Spores can stay dormant for extended periods of time. When the

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

environment becomes favorable, spores can exit to the vegetative state. Revival of spores millions of years old has been reported (Cano and Borucki 1995). It is thought that spore-forming bacteria mostly exist as spores in nature (Priest and Grigorova 1990), therefore it is reasonable to assume that spore-forming *Firmicutes* (SFF) have longer generation times than their nonspore-forming relatives (NSFF). As such, *Firmicutes* represents an excellent system to test the generation-time hypothesis in bacteria.

Not replicating while dormant, SFF are expected to have fewer DNA replication errors per unit time, and thus a lower rate of evolution. It is therefore surprising that a previous study showed no differences in the evolutionary rates when comparing SFF and NSFF (Maughan 2007). However, the use of few *Firmicutes* genomes in that study may have prevented detection of differences between these groups, if such a difference exists. Leveraging on a substantially increased representation of the *Firmicutes* phylum and using a phylogenetic comparative approach, we revisited the relationship between spore formation and evolutionary rates in bacteria. We found strong evidence that rates of molecular evolution are correlated with generation time in bacteria.

Methods

RECONSTRUCTING A GENOME TREE OF *FIRMICUTES*

Using the Phyla-AMPHORA package (Wang and Wu 2013), protein sequences of 168 single-copy, "universal," phylum-level marker genes from 573 completed genomes of *Firmicutes* were identified, aligned, trimmed, and concatenated into a single alignment. To reduce the computational cost, a FastTree (Price et al. 2009) built from the concatenated alignment was used to select 200 representatives that maximized the phylogenetic diversity, using a greedy algorithm described in Steel (2005). A RAxML maximum-likelihood tree was made with 1000 bootstrap replicates using the LG + gamma model (Stamatakis 2014). Similarly, a bacterial genome tree containing 200 top representatives of major bacteria phyla was reconstructed, using the concatenated protein sequences of 31 "universal markers" (Wu and Scott 2012). The *Firmicutes* genome tree was then rooted using the bacterial genome tree as a guide.

IDENTIFYING SPORULATION GENES IN *FIRMICUTES* GENOMES

A total of 199 genes whose annotation contained keywords "spore," "sporulation," or "germination" were downloaded from *Bacillus subtilis* database SubtiList (http://genolist.pasteur.fr/SubtiList/), and combined with 175 *B. subtilis* sporulation genes used in Wu et al. (2005). Genes with highly conserved domains (e.g., kinases, phosphatases, ATPases) were removed from the list, resulting in a set of 163 sporulation genes, which were used as query sequences to BLASTP search against 573 complete *Firmicutes* genomes. The mutual best hits of each sporulation gene in each *Firmicutes* genome were identified as orthologous genes. The *e*-value cutoff was 1×10^{-3} in both the forward and reverse BLASTP searches. The same procedure was also used to identify sporulation genes in representatives of non-*Firmicutes* bacteria.

PREDICTING SPORULATION POTENTIAL WITH PHYLOGENETIC PROFILING

The distribution patterns of the 163 sporulation genes in each of the 200 *Firmicutes* representatives were represented with a binary matrix, where each column represented one species and each row represented one gene. "1" and "0" denoted the presence and absence of genes, respectively. Species and genes were then grouped by the gene distribution patterns (phylogenetic profiles) with the CLUSTER program (http://rana.lbl.gov/EisenSoftware.htm), using the absolute correlation (centered) as the similarity metric and complete linkage as the clustering method.

ANCESTOR STATE RECONSTRUCTION

Having predicted sporulation potential by phylogenetic profiling, we reconstructed the ancestral state with the R package "ape" (Paradis et al. 2004). A binary categorical trait of "spore-forming" and "nonspore-forming" was mapped onto the tips of the *Firmicutes* genome tree. Transition probabilities for trait gain and loss were set to be equal. Results were similar when transition probabilities of trait gain to loss were set to be 1:2, 1:5, and 1:10.

PAML ANALYSES

The CDS (coding DNA sequences) of the 168 protein-coding genes were aligned using their protein sequence alignments as a guide. Using the alignment, the likelihood of the corresponding *Firmicutes* genome tree was calculated in PAML (Yang 2007) using three molecular clock models: no molecular clock, in which the rate of every branch varied freely; a global clock, in which every branch shared the same evolutionary rate; and a local clock, where spore-forming and nonspore-forming branches had separate rates. The Akaike information criterion (AIC) was calculated for each model to determine which was the best. From the codon alignments, the synonymous substitution rate, *dS*, was calculated in PAML (Yang 2007) using a neutral site, no molecular clock model.

PHYLOGENETIC INDEPENDENT CONTRAST ANALYSIS OF EVOLUTIONARY RATES

Independent contrasts were performed using the R package "caper" (Orme 2013). Discrete contrasts were made using the predicted sporulation potential (spore-forming or nonspore-forming). Continuous contrasts were made using the numeric count of identified sporulation genes. Contrasts with studentized residuals

having absolute values greater than 3 were excluded from our analyses as potential outliers.

ESTIMATION OF CODON BIAS

For each of the 197 *Firmicutes* representatives, CDS of all protein-coding genes in the genome were concatenated into a single nucleotide sequence. The codon bias index (CBI) was calculated for each genome in CodonW (http://codonw.sourceforge.net) based on the usage of a subset of *B. subtilis* optimal codons.

Results

PREDICTION OF A SPECIES' SPORULATION POTENTIAL

Sporulation is a highly complex but one of the best-studied biological processes. For example, in *B. subtilis*, the model organism for studying sporulation, sporulation involves a cascade of expression of at least 200 genes (Piggot and Losick 2002). For a nonmodel organism, classifying whether it can sporulate or not based on laboratory observation is unreliable because we do not know exactly what environmental factors might trigger sporulation in the organism. When a species loses its ability to sporulate (e.g., by losing a key sporulation gene), it is expected to eventually lose most of its sporulation genes because there is no longer selective pressure to maintain them, unless they also perform other biological functions. The tight correlation between the sporulation process and the number of sporulation genes has been used to predict a species' potential to form spores (Onyenwoke et al. 2004; Maughan 2007).

Gene content has been used to group species (Huson and Steel 2004) and in theory could be used to predict phenotypes associated with the genes as well. If we can partition *Firmicutes* into spore-forming and nonspore-forming groups by the distribution patterns of sporulation genes, then we should be able to predict the spore-forming ability of a species based on its membership in the groups. To predict the sporulation potential of *Firmicutes*, we carried out a phylogenetic profiling analysis of sporulation genes in all complete *Firmicutes* genomes. Using a mutual best hit approach, we identified orthologs of 163 known *B. subtilis* sporulation genes in other *Firmicutes* species (Table S1). The distribution patterns of sporulation genes in 200 representatives of *Firmicutes* are shown in Figure 1. Based on the presence/absence of sporulation genes, *Firmicutes* were partitioned into two major groups at the root of the tree (Fig. 1). The first group contained Staphylococcus, Streptococcus, and Lactobacillus that are unable to sporulate, whereas the second group contained well-known spore-forming genera such as Bacillus and Clostridium (Fig. 1). We therefore predicted that members of the first and second groups are NSFF and SFF, respectively. Consistent with our predictions, not a single species within the first group is known to sporulate. The average number of sporulation genes in the genomes of the first group (95% CI: 15.05 $\pm$ 0.39 genes per genome) is significantly smaller than that of the second group (95% CI: 85.2 $\pm$ 4.42, $P < 0.0001$). Accordingly, when plotting the number of sporulation genes against the genome size, SFF form a cluster that is well separated from NSFF and other bacterial phyla (Fig. S1). However, we note that one butyrate-producing bacterium SM4/1 (NCBI taxonid 245012) in group 2 is missing 146 of 163 sporulation genes including the master switch gene *spo0A*. We therefore predicted it to be nonspore-forming.

EVOLUTION OF SPORULATION IN *FIRMICUTES*

To test whether SFF and NSFF have different evolutionary rates, it is a prerequisite to obtain a robust species tree of *Firmicutes*. First of all, the tree is required for estimating the overall species evolutionary rates. Second, the tree provides a framework of the natural relationships between species, so correlation due to common ancestry can be removed (Felsenstein 1985; Harvey and Purvis 1991) when we test the effect of sporulation on the evolutionary rates. Using the concatenated protein sequence alignment of 168 single-copy genes that are universally present in all *Firmicutes*, we reconstructed a maximum-likelihood tree of 200 *Firmicutes* representatives (Fig. 2). The tree was rooted with the last common ancestor (LCA) of *Ammonifex* and *Sulfobacillus* as the deepest lineage, based on the topology of a bacterial species tree containing all major bacterial phyla. The tree is fully resolved and as typical of genome trees, the nodes in the tree are all highly supported with few exceptions.

Consistent with the previous study (Maughan 2007), our ancestral state reconstructions indicate that sporulation arose only once, in the LCA of *Firmicutes*, and subsequently has been lost many times (Fig. 2). Interestingly, our ancestral state reconstruction shows that once the ability to form spores was lost, it was never subsequently regained—even with the relaxed assumption that trait gain and loss are equally likely. This is consistent with Dollo's law that complex traits, once lost, may be difficult to re-evolve (Marshall et al. 1994).

Some losses were ancient, for example, in the LCA of *Staphylococcus* and *Lactobacillus*, leading to the diversification of large clades of *Firmicutes* that are all unable to form spores. Other losses were more recent, one striking example being *Clostridiales genomospecies* BVAB3. A previous study (Galperin et al. 2012) and we both predicted that BVAB3 is an NSFF as it encodes merely eight sporulation genes. Consistently, endospore formation has not been observed for this species. Our ancestral state reconstructions predicted that BVAB3 lost the ability to sporulate after it split off from its close relatives. These relatives, *Clostridium clariflavum*, *Ruminiclostridium thermocellum*, *C. cellulolyticum*, and *C. stercorarium* have all been observed to form spores in the laboratory (Freier et al. 1988; Gehin et al. 1995;

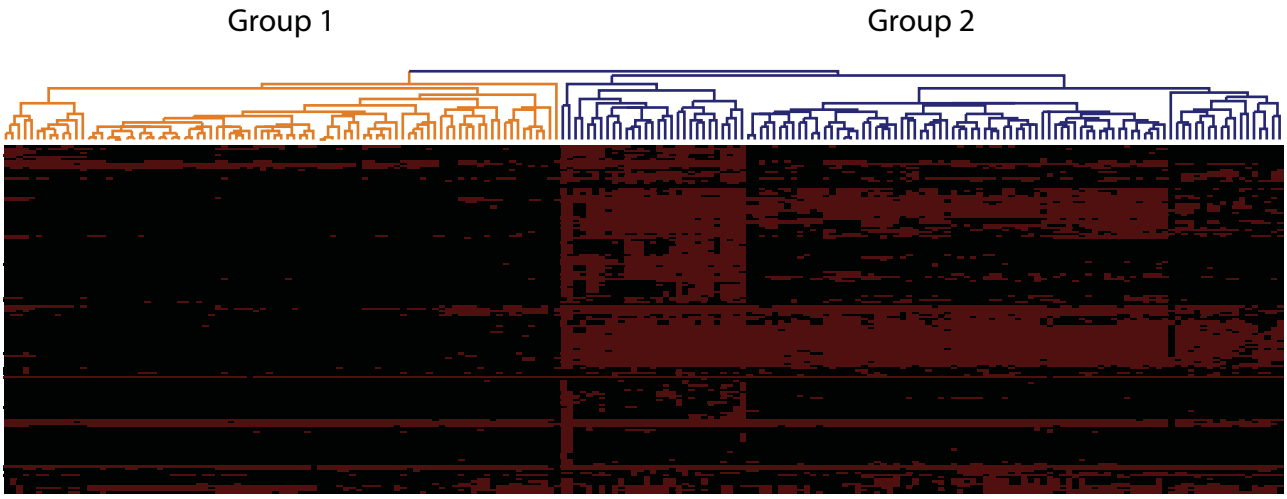

Figure 1. Phylogenetic profile analysis of 163 sporulation genes in genomes of 200 *Firmicutes* representatives. Each row represents one sporulation gene and each column represents one genome. The presence and absence of genes are indicated in red and black, respectively. *Firmicutes* were clustered into two groups by the similarities of the distribution profiles of sporulation genes in the genomes. The phylogenetic profiles are also described in Table S1.

Shiratori et al. 2009; Dumitrache et al. 2013). Consistent with our prediction that NSFF evolve more quickly, the branch length of BVAB3 to its last spore-forming ancestor is twice of those of its closest relatives—a clear demonstration of more rapid evolution coinciding with the loss of spore-forming ability.

Similar losses were predicted in species for which sporulation has not been reported: *B. selenitireducens* (Afkar et al. 2003), the LCA of *Finegoldia magna* (formerly *Peptostreptococcus magnus*, Jassem et al. 1996) and *Anaerococcus prevotii* (LaButti et al. 2009), and the LCA of *Acetobacterium woodii* (Balch et al. 1977) and *Eubacterium limosum* (Genthner et al. 1981), among others. In each of these instances, the loss of endospore formation appears to coincide with longer branch lengths, compared to their close relatives.

EVOLUTIONARY RATE VARIES BETWEEN LINEAGES

In comparing branch lengths of *Firmicutes* possessing various counts of sporulation genes, it is clear that evolutionary rate is not uniform (Fig. 3, Table S2). To determine if variation does in fact exist between lineages, various molecular clock models were tested in PAML. Our results show that a global molecular clock is the poorest model compared to both local and no-clock models (Table S3), supporting the previous finding that protein evolutionary rates vary between *Firmicutes* lineages (Maughan 2007). In the local clock model, SFF and NSFF were each described by a separate evolutionary rate. The local model fits the data better than the global clock, showing that some variation in evolutionary rate can be explained by spore-forming ability. Because the no-clock model fits the data best, there exists additional variation in evolutionary rate that cannot be explained by spore-forming ability alone. As the local clock does not take differences due to phylogeny into consideration, this is to be expected.

SFF HAVE LOWER PROTEIN EVOLUTIONARY RATES

The presence of many loss-of-trait events in the species tree of *Firmicutes* provides a powerful framework for us to test the effect of spore formation on evolutionary rate. We first tested whether the amino acid substitution rate variation correlates with spore-forming ability. To remove correlation due to shared evolutionary history, we conducted phylogenetically independent contrast analyses. Independent contrasts performed using 200 *Firmicutes* representatives showed that there is a strong association between the amino acid substitution rate (measured by the tree branch length of the tree in Fig. 2) and spore-forming ability ($P < 0.001$, $df = 14$, Table S4). One contrast involved three species (*Thermodesulfobium narugense*, *Coprothermobacter proteolyticus*, and *Natranaerobius thermophilus*) that were suspected to be clustered together due to long-branch attraction. These three species were therefore removed, resulting in a 197-species dataset. Independent contrasts conducted with this reduced dataset also showed a strong association between spore-forming ability and evolutionary rate ($P < 0.0001$, $df = 12$, Table 1).

To ensure that our conclusions are not solely dependent on our predictions of spore-forming ability, as a precautionary measure we removed any SFF not corroborated by Galperin et al. (2012). This resulted in removal of 60 taxa predicted to be spore-forming in our study. We retained species that we classified as NSFF as these classifications are most likely accurate due to the low number of sporulation genes and the lack of master sporulation gene *spo0A* (Bueche et al. 2013).

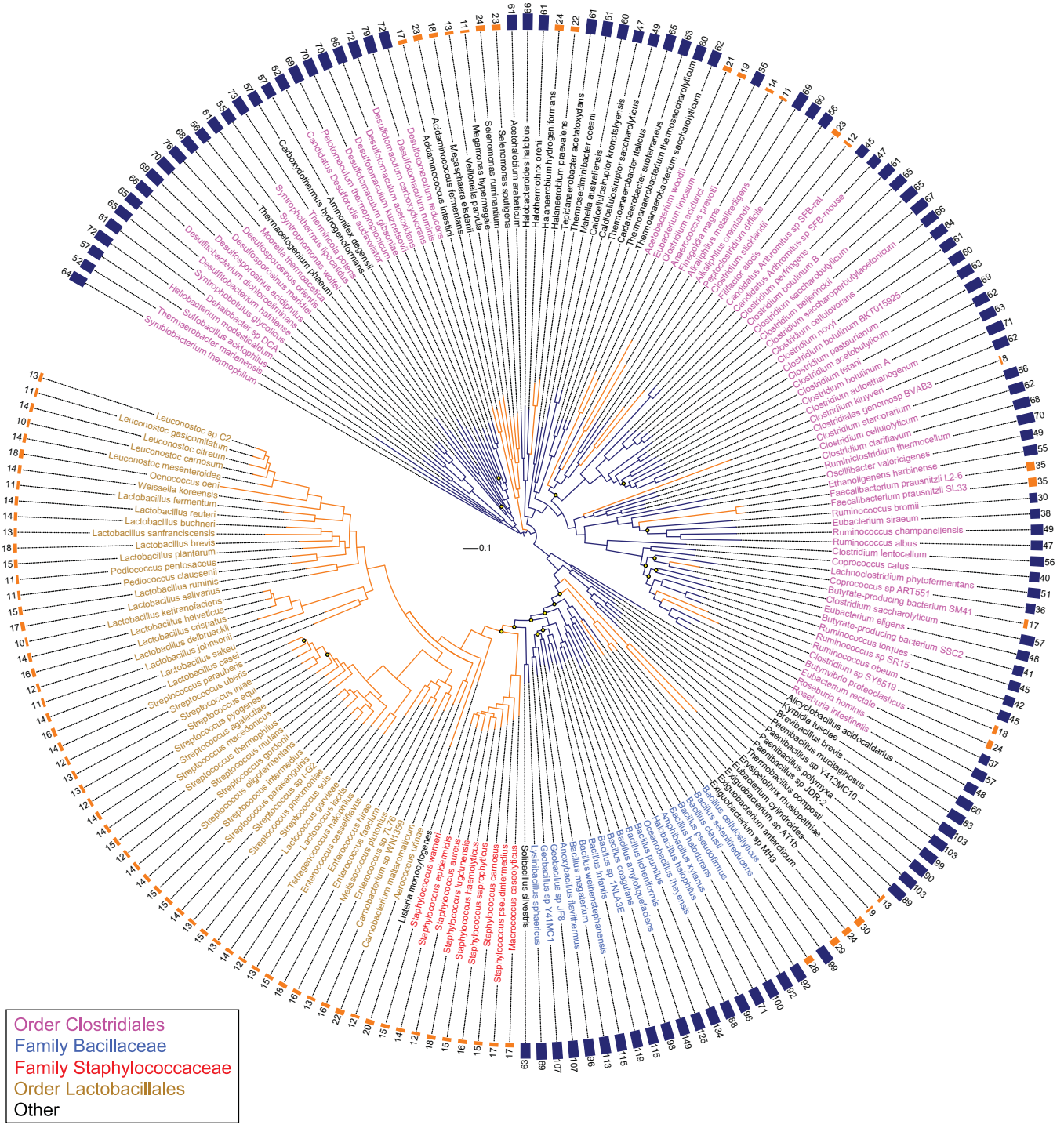

Figure 2. **A maximum-likelihood genome tree of 197 *Firmicutes* representatives based on 168 shared marker genes. The size of the bar on the outer circle represents the number of identified sporulation genes in the genome. The bars were colored by the predicted spore-forming potential of their corresponding genomes (blue: spore-forming; orange: nonspore-forming). The branches were highlighted using the same color scheme based on the maximum-likelihood ancestral state reconstruction of spore-forming ability of the lineages. Bootstrap support for nodes is ≥80% except for those indicated with a yellow circle.**

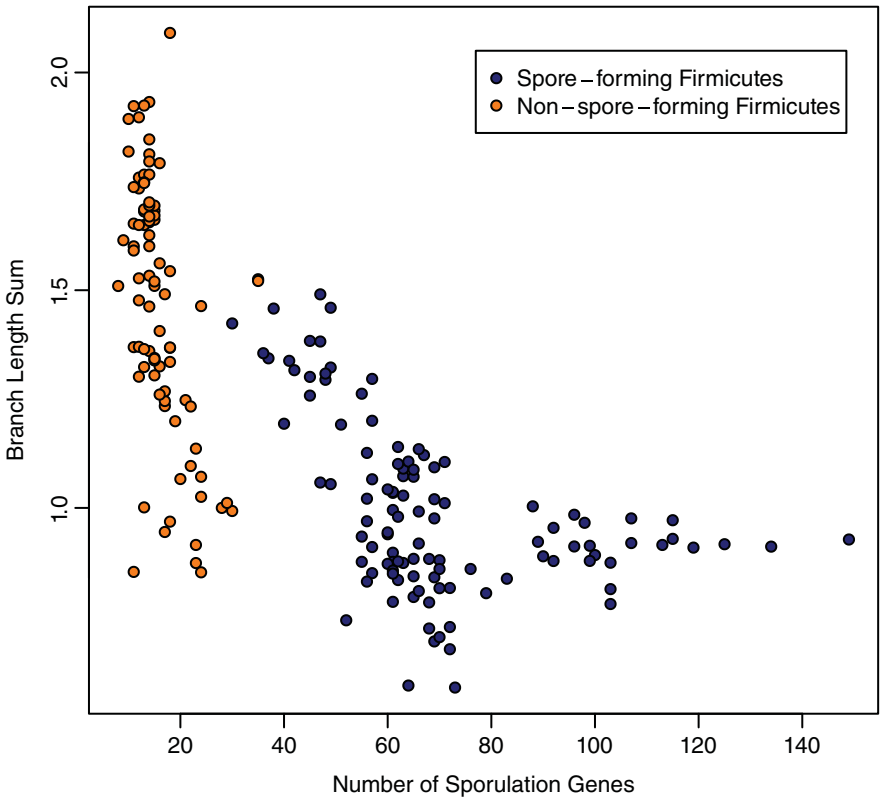

Figure 3. **Relationship between sporulation gene count and the branch length (from the tip to the root) of 197 *Firmicutes* representatives. Each dot represents a genome and is colored by the genome's predicted sporulation potential.**

Table 1. **Independent contrasts analysis of evolutionary rates of 197 *Firmicutes* representatives.**

Sporulation variable	Evolutionary rate	Contrasts	*P*-value (exact)	*P*-value	R^2
Discrete	Amino acid substitution	13	0.0008876	<0.001	0.6158
Continuous	Amino acid substitution	192	1.29×10^{-13}	<0.0001	0.2502
Discrete	*dS*	14	0.02487	<0.05	0.3309
Continuous	*dS*	194	1.41×10^{-11}	<0.0001	0.2111

Independent contrasts yielded results similar to the 197 and 200 species dataset (Table S5).

Because there is uncertainty associated with sporulation prediction based on phylogenetic profiles, we also performed phylogenetic contrast analysis using the nominal count of sporulation genes as a continuous variable. These continuous independent contrasts conducted with the 200 and 197 species datasets corroborated the results of discrete contrasts ($P < 0.0001$, $df = 199$, Table S4; $P < 0.0001$, $df = 192$, Table 1) without having presumed any spore-forming ability, showing that species with greater counts of sporulation genes evolve more slowly.

To evaluate the effect of the robustness of the tree topology on our results, we removed three contrasts with less than 80% bootstrap support. Specifically, this eliminated 85 species descending from the LCA of *Exiguobacterium* spp. and Lactobacillales. Independent contrasts for this 112-species tree yielded results similar to the 197 and 200 species datasets (Table S6).

SFF HAVE LOWER SPONTANEOUS MUTATION RATES

Because the amino acid substitution rate is the compound result of spontaneous mutation, selection, and genetic drift, next we determined whether sporulation affects the spontaneous mutation rate alone. Spontaneous mutation is caused by DNA polymerase errors and other molecular processes that introduce errors during the transmission of genetic information, and therefore directly correlates with the generation time. We expect SFF with longer generation time to have lower spontaneous mutation rates than NSFF.

We measured the spontaneous mutation rate using the synonymous substitution rate (*dS*). Because synonymous mutations in general do not alter the phenotype, they are assumed to be

selectively neutral and have been used to estimate the spontaneous mutation rate (Kimura and Ohta 1971, Ohta 1993). *dS* were calculated with codeml of the PAML package using aligned DNA sequences of the 168 genes shared among *Firmicutes* (Table S2). Both discrete and continuous *dS*-based phylogenetic contrasts indicate that spore-forming ability is significantly associated with the spontaneous mutation rates (Table 1).

NO EVIDENCE OF CORRELATION BETWEEN POPULATION SIZE AND SPORULATION

If the population size correlates with spore-forming ability, then genetic drift resulting from small population size could also explain differences in amino acid substitution rates. To test this potential confounding factor, we calculated genome-wide codon bias indices—a proxy for population size—for *Firmicutes* used in our analyses (Table S2). Independent contrasts performed on CBIs revealed no association with either predicted spore-forming ability ($df = 13$, $P > 0.05$, $R^2 = 0.007$) or sporulation gene count ($df = 192$, $P > 0.05$, $R^2 = 0.01154$).

Discussion

Using a large dataset of 200 *Firmicutes* species, we found sporulation significantly reduces the rates of molecular evolution in bacteria. This is in sharp contrast to a previous study that found rates of molecular evolution in bacteria are relatively constant despite spore dormancy (Maughan 2007). One possible reason that the study by Maughan (2007) failed to detect significant difference in evolutionary rates between SFF and NSFF is the lack of power in the data: its conclusion was based on only two phylogenetic contrasts. In comparison, we had at least 13 contrasts because of the large number of species included in our study. Another and maybe more important reason is the likely misclassification of sporulation ability of some key *Firmicutes* species in the previous study. One of the two contrasts in Maughan (2007) study involved *Thermoanaerobacter tengcongensis*, which was classified as nonspore forming and was contrasted against three spore-forming *Clostridium* spp. We think the contrast is likely wrong. *Thermoanaerobacter tengcongensis* encodes 63 sporulation genes (Fig. 2) and therefore was predicted to be able to sporulate in this and a previous study (Traag et al. 2013). Although *T. tengcongensis* has not been observed to sporulate when it was first identified (Xue et al. 2001), RNA-seq has shown that a cold-shock protein induces expression of sporulation genes in response to low temperatures (Liu et al. 2014). We note that misclassification of SFF or NSFF will reduce the power to detect generation-time effect. The fact we were able to detect generation-time effect suggests that it is real and might be even stronger.

In this study, we observed a generation-time effect in both synonymous and nonsynonymous substitution rates of protein-coding genes. In theory, synonymous changes are not affected by selection and genetic drift and therefore are expected to reflect the underlying mutation rate. The generation-time effect on evolutionary rate was originally found to be most evident at synonymous sites (Li and Tanimura 1987; Ohta 1993). This is because on top of mutation rate, nonsynonymous substitutions are also affected by selection and genetic drift, which can obscure the generation-time effect in amino acid substitutions. For example, mammals with longer generation time tend to have smaller population size, which causes accelerated substitution rates for slightly deleterious mutations. This explains the failure to detect a generation-time effect in nonsynonymous substitution rates in mammals: it is cancelled out by the genetic drift. In bacteria, population size varies between species and is known to have a large impact on nonsynonymous substitution rates. For example, obligate intracellular bacteria usually have much smaller population sizes than free-living bacteria. The increased genetic drift in obligate intracellular bacteria contributes to their highly accelerated amino acid substitution rates (Ochman et al. 1999). The extent to which selection and genetic drift differ between SFF and NSFF is not well known. The fact that we were able to detect a generation-time effect in nonsynonymous substitution rates suggests that neither has obscured the signal in *Firmicutes*. Consistent with this, using the codon usage bias as a proxy for estimating the effective population size, we did not detect significant correlation between population size and sporulation, indicating that genetic drift is not the main cause of rate variation in *Firmicutes*. In fact, in our study we found even a stronger evidence of a generation-time effect in nonsynonymous substitution rate (Table 1). We think this is because the rate differences in synonymous substitution rate were underestimated due to saturated synonymous substitutions in our dataset. Indeed, the average *dS* in our phylogenetic contrast analysis approached 2.66 substitutions per site, a clear sign of saturated substitutions. In addition, it is possible that endospores experience less selection at nonsynonymous sites because of the dormancy, resulting in fewer substitutions in SFF and contributing to the greater disparity.

Another source of mutations in bacteria is replication-independent mutations such as these caused by unrepaired DNA damages. Because they are independent of DNA replication, they are unaffected by generation time and therefore can obscure the generation-time effect. Spores have tremendous resistance to DNA damage, however, due to a combination of multiple factors including DNA packing assisted by small acid-soluble spore proteins, dehydration, the high concentration of dipicolinic acid (Desnous et al. 2010), and spore photoproduct lyase that repairs UV damage (Van Wang and Rupert 1977). Again, the fact that we observed generation-time effect in *dS* suggests that replication-independent mutations do not play a major role in the evolution of *Firmicutes*.

Concordant with Abecasis et al. (2013), our ancestral reconstruction suggests that sporulation evolved only once in the LCA of *Firmicutes* and has since been lost many times in nonspore-forming lineages. Assuming that once the sporulation trait is lost so will most of the sporulation genes, the sporulation gene content can be used not only for predicting the sporulation ability but more importantly the time a species has spent as a spore-former in its evolutionary history. Using sporulation gene number in phylogenetic contrast analysis therefore circumvents the uncertainty associated with the binary trait prediction. Regardless of whether we use a binary (sporulation ability) or continuous variable (number of sporulation genes) in our phylogenetic contrast analyses, we detected significant generation-time effect in *Firmicutes*.

The generation-time effect appears to be genome wide as our results were based on the analysis of 168 genes common among all *Firmicutes*. Our analyses were not based on laboratory observations, but instead on genomes—the products of bacteria reproducing and evolving over millions of years. As such, they provide strong evidence of generation-time effect in natural bacterial populations throughout an extended evolutionary history.

ACKNOWLEDGMENT

The authors declare they have no competing interests.

DATA ARCHIVING

The doi for our data is 10.5061/dryad.n855t.

LITERATURE CITED

Abecasis, A. B., M. Serrano, R. Alves, L. Quintais, J. B. Pereira-Leal, and A. O. Henriques. 2013. A genomic signature and the identification of new sporulation genes. J. Bacteriol. 195:2101–2115.

Afkar, E., J. Lisak, C. Saltikov, P. Basu, R. S. Oremland, and J. F. Stolz. 2003. The respiratory arsenate reductase from *Bacillus selenitireducens* strain MLS10. FEMS Microbiol. Lett. 226:107–112.

Ainouche, A.-K., and R. J. Bayer. 1999. Phylogenetic relationships in *Lupinus* (Fabaceae: Papilionoideae) based on internal transcribed spacer sequences (ITS) of nuclear ribosomal DNA. Am. J. Bot. 86:590–607.

Andreasen, K., and B. G. Baldwin. 2001. Unequal evolutionary rates between annual and perennial lineages of checker mallows (Sidalcea, Malvaceae): evidence from 18S-26S rDNA internal and external transcribed spacers. Mol. Biol. Evol. 18:936–944.

Balch, W. E., S. Schoberth, R. S. Tanner, and R. S. Wolfe. 1977. *Acetobacterium*, a new genus of hydrogen-oxidizing, carbon dioxide-reducing, Anaerobic Bacteria. Int. J. Syst. Bacteriol. 27:355–361.

Bousquet, J., S. H. Strauss, A. H. Doerksen, and R. A. Price. 1992. Extensive variation in evolutionary rate of rbcL gene sequences among seed plants. Proc. Natl. Acad. Sci. 89:7844–7848.

Bromham, L. 2002. Molecular clocks in reptiles: life history influences rate of molecular evolution. Mol. Biol. Evol. 19:302–309.

Bromham, L., A. Rambaut, and P. H. Harvey. 1996. Determinants of rate variation in mammalian DNA sequence evolution. J. Mol. Evol. 43:610–621.

Bueche, M., T. Wunderlin, L. Roussel-Delif, T. Junier, L. Sauvain, N. Jeanneret, and P. Junier. 2013. Quantification of endospore-forming *Firmicutes* by quantitative PCR with the functional gene *spo0A*. Appl. Environ. Microbiol. 79:5302–5312.

Buschiazzo, E., C. Ritland, J. Bohlmann, and K. Ritland. 2012. Slow but not low: genomic comparisons reveal slower evolutionary rate and higher dN/dS in conifers compared to angiosperms. BMC Evol. Biol. 12:8.

Cano, R. J., and M. K. Borucki. 1995. Revival and identification of bacterial spores in 25- to 40-million-year-old Dominican amber. Science 268:1060–1064.

Desnous, C., D. Guillaume, and P. Clivio. 2010. Spore photoproduct: a key to bacterial eternal life. Chem. Rev. 110:1213–1232.

Dumitrache, A., G. Wolfaardt, G. Allen, S. N. Liss, and L. R. Lynd. 2013. Form and function of *Clostridium thermocellum* biofilms. Appl. Environ. Microbiol. 79:231–239.

Felsenstein, J. 1985. Phylogenies and the comparative method. Am. Nat. 125:1–15.

Freier, D., C. P. Mothershed, and J. Wiegel. 1988. Characterization of *Clostridium thermocellum* JW20. Appl. Environ. Microbiol. 54:204–211.

Galperin, M. Y., S. L. Mekhedov, P. Puigbo, S. Smirnov, Y. I. Wolf, and D. J. Rigden. 2012. Genomic determinants of sporulation in Bacilli and Clostridia: towards the minimal set of sporulation-specific genes. Environ. Microbiol. 14:2870–2890.

Gaut, B. S., B. R. Morton, B. C. McCaig, and M. T. Clegg. 1996. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene rbcL. Proc. Natl. Acad. Sci. 93:10274–10279.

Gaut, B. S., L. G. Clark, J. F. Wendel, and S. V Muse. 1997. Comparisons of the molecular evolutionary process at rbcL and ndhF in the grass family (Poaceae). Mol. Biol. Evol. 14:769–777.

Gehin, A., E. Gelhaye, G. Raval, and H. Petitdemange. 1995. Clostridium cellulolyticum viability and sporulation under cellobiose starvation conditions. Appl. Environ. Microbiol. 61:868–871.

Genthner, B. R., C. L. Davis, and M. P. Bryant. 1981. Features of rumen and sewage sludge strains of *Eubacterium limosum*, a methanol- and H_2-CO_2-utilizing species. Appl. Environ. Microbiol. 42:12–19.

Harvey, P. H., and A. Purvis. 1991. Comparative methods for explaining adaptations. Nature 351:619–624.

Huson, D. H., and M. Steel. 2004. Phylogenetic trees based on gene content. Bioinformatics 20:2044–2049.

Jassem, E., A. Kedzia, M. Rek, L. Wolska-Goszka, and K. Szelezyński. 1996. Occurrence of non-spore forming anaerobic bacteria in the upper airways of patients with chronic obstructive pulmonary disease. Med. Dosw. Mikrobiol. 48:49–54.

Kohne, D. E. 1970. Evolution of higher-organism DNA. Q. Rev. Biophys. 3:327–375.

Kuo, C.-H., and H. Ochman. 2009. Inferring clocks when lacking rocks: the variable rates of molecular evolution in bacteria. Biol. Direct 4:35.

Labutti, K., R. Pukall, K. Steenblock, T. Glavina Del Rio, H. Tice, A. Copeland, J.-F. Cheng, S. Lucas, F. Chen, M. Nolan, et al. 2009. Complete genome sequence of *Anaerococcus prevotii* type strain (PC1). Stand. Genomic Sci. 1:159–165.

Laird, C. D., B. L. Mcconaughy, and B. J. Mccarthy. 1969. Rate of fixation of nucleotide substitutions in evolution. Nature 224:149–154.

Laroche, J., and J. Bousquet. 1999. Evolution of the mitochondrial rps3 intron in perennial and annual angiosperms and homology to nad5 intron 1. Mol. Biol. Evol. 16:441–452.

Laroche, J., P. Li, L. Maggia, and J. Bousquet. 1997. Molecular evolution of angiosperm mitochondrial introns and exons. Proc. Natl. Acad. Sci. 94:5722–5727.

Li, W. H., and M. Tanimura. 1987. The molecular clock runs more slowly in man than in apes and monkeys. Nature 326:93–96.

Liu, B., Y. Zhang, and W. Zhang. 2014. RNA-Seq-based analysis of cold shock response in *Thermoanaerobacter tengcongensis*, a bacterium harboring a single cold shock protein encoding gene. PLoS One 9:e93289.

Marshall, C. R., E. C. Raff, and R. A. Raff. 1994. Dollo's law and the death and resurrection of genes. Proc. Natl. Acad. Sci. USA 91:12283–12287.

Maughan, H. 2007. Rates of molecular evolution in bacteria are relatively constant despite spore dormancy. Evolution 61:280–288.

Mooers, A. O., and P. H. Harvey. 1994. Metabolic rate, generation time, and the rate of molecular evolution in birds. Mol. Phylogenet. Evol. 3:344–350.

Nabholz, B., S. Glémin, and N. Galtier. 2008. Strong variations of mitochondrial mutation rate across mammals—the longevity hypothesis. Mol. Biol. Evol. 25:120–130.

Ochman, H., S. Elwyn, and N. A. Moran. 1999. Calibrating bacterial evolution. Proc. Natl. Acad. Sci. USA 96:12638–12643.

Ohta, T. 1993. An examination of the generation-time effect on molecular evolution. Proc. Natl. Acad. Sci. U. S. A. 90:10676–10680.

Ohta, T., and M. Kimura. 1971. On the constancy of the evolutionary rate of cistrons. J. Mol. Evol. 1:18–25.

Onyenwoke, R. U., J. A. Brill, K. Farahi, and J. Wiegel. 2004. Sporulation genes in members of the low G+C Gram-type-positive phylogenetic branch (Firmicutes). Arch. Microbiol. 182:182–192.

Orme, D. 2013. The caper package: comparative analysis of phylogenetics and evolution in R.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290.

Piggot, P. J., and R. Losick. 2002. Sporulation genes and intercompartmental regulation. Pp. 483–518 *in* A. L. Sonenshein, J. A. Hoch, and R. Losick, eds. Bacillus subtilis and its closest relatives. American Society for Micriobiology, Washington, DC.

Price, M. N., P. S. Dehal, and A. P. Arkin. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. 26:1641–1650.

Priest, F. G., and R. Grigorova. 1990. 18 methods for studying the ecology of endospore-forming bacteria. Methods Microbiol. 22:565–591.

Shiratori, H., K. Sasaya, H. Ohiwa, H. Ikeno, S. Ayame, N. Kataoka, A. Miya, T. Beppu, and K. Ueda. 2009. *Clostridium clariflavum* sp. nov. and *Clostridium caenicola* sp. nov., moderately thermophilic, cellulose-/cellobiose-digesting bacteria isolated from methanogenic sludge. Int. J. Syst. Evol. Microbiol. 59:1764–1770.

Smith, S. A., and M. J. Donoghue. 2008. Rates of molecular evolution are linked to life history in flowering plants. Science 322:86–89.

Soria-Hernanz, D. F., O. Fiz-Palacios, J. M. Braverman, and M. B. Hamilton. 2008. Reconsidering the generation time hypothesis based on nuclear ribosomal ITS sequence comparisons in annual and perennial angiosperms. BMC Evol. Biol. 8:344.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Steel, M. 2005. Phylogenetic diversity and the greedy algorithm. Syst. Biol. 54:527–529.

Thomas, J. A., J. J. Welch, M. Woolfit, and L. Bromham. 2006. There is no universal molecular clock for invertebrates, but rate variation does not scale with body size. Proc. Natl. Acad. Sci. USA 103:7366–7371.

Thomas, J. A., J. J. Welch, R. Lanfear, and L. Bromham. 2010. A generation time effect on the rate of molecular evolution in invertebrates. Mol. Biol. Evol. 27:1173–1180.

Traag, B. A., A. Pugliese, J. A. Eisen, and R. Losick. 2013. Gene conservation among endospore-forming bacteria reveals additional sporulation genes in *Bacillus subtilis*. J. Bacteriol. 195:253–260.

Van Wang, T. C., and C. S. Rupert. 1977. Evidence for the monomerization of spore photoproduct to two thymines by the light-independent "spore repair" process in *Bacillus subtilis*. Photochem. Photobiol. 25:123–127.

Wang, Z., and M. Wu. 2013. A phylum-level bacterial phylogenetic marker database. Mol. Biol. Evol. 30:1258–1262.

Welch, J. J., O. R. P. Bininda-Emonds, and L. Bromham. 2008. Correlates of substitution rate variation in mammalian protein-coding sequences. BMC Evol. Biol. 8:53.

Whittle, C.-A., and M. O. Johnston. 2003. Broad-scale analysis contradicts the theory that generation time affects molecular evolutionary rates in plants. J. Mol. Evol. 56:223–233.

Wu, M., and A. J. Scott. 2012. Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. Bioinformatics 28:1033–1034.

Wu, M., Q. Ren, A. S. Durkin, S. C. Daugherty, L. M. Brinkac, R. J. Dodson, R. Madupu, S. A. Sullivan, J. F. Kolonay, D. H. Haft, et al. 2005. Life in hot carbon monoxide: the complete genome sequence of Carboxydothermus hydrogenoformans Z-2901. PLoS Genet. 1:e65.

Xue, Y., Y. Xu, Y. Liu, Y. Ma, and P. Zhou. 2001. *Thermoanaerobacter tengcongensis* sp. nov., a novel anaerobic, saccharolytic, thermophilic bacterium isolated from a hot spring in Tengcong, China. Int. J. Syst. Evol. Microbiol. 51:1335–1341.

Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24:1586–1591.

Yue, J.-X., J. Li, D. Wang, H. Araki, D. Tian, and S. Yang. 2010. Genome-wide investigation reveals high evolutionary rates in annual model plants. BMC Plant Biol. 10:242.

Associate Editor: A. D. Visser
Handling Editor: T. Lenormand

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Figure S1. The number of sporulation genes plotted against the number of protein coding genes in the genomes.
Table S1. Phylogenetic profiles of sporulation genes in genomes of 200 *Firmicutes* representatives.
Table S2. List of molecular evolutionary rates and codon bias index for each of the 197 *Firmicutes* species analyzed in this study.
Table S3. Log-likelihood and AIC values of different clock models.
Table S4. Independent contrast analysis of protein evolutionary rates of 200 *Firmicutes* representatives.
Table S5. Independent contrast analysis of evolutionary rates in 137 *Firmicutes* representatives after removing 60 species not positively identified as spore-forming in Galperin et al. (2012).
Table S6. Independent contrast analysis of evolutionary rates in 112 *Firmicutes* representatives after removing three contrast nodes with less than 80% bootstrap support from the 197-species dataset.