PNAS RESEARCH ARTICLE | EVOLUTION OPEN ACCESS

Evolutionary bioenergetics of sporulation

Canan Karakoç[a,b], William R. Shoemaker[c], and Jay T. Lennon[a,1]

Edited by Bo Barker Jorgensen, Aarhus University, Aarhus, Denmark; received September 1, 2025; accepted January 8, 2026

Energy is required for the expression and maintenance of complex traits. In many habitats, however, free energy available to support biosynthesis is in vanishingly short supply. As a result, many taxa have evolved persistence strategies that support survival in unfavorable environments. Among these is sporulation, an ancient bacterial program governed by a large genetic network that requires energy for both regulation and execution. Yet sporulation is a last resort, initiated when cellular energy is nearly exhausted. To resolve this paradox, we quantified the energetic cost of sporulation in units of ATP by integrating time-resolved genome, transcriptome, and proteome profiles. The full cost of the spore cycle, including both formation and revival, ranks among the most energy-intensive processes in the bacterial cell, requiring almost 10^{10} ATP and consuming about 10% of the total energy budget. The majority of this cost arises from translation, membrane synthesis, and protein turnover. Despite its considerable upfront investment, sporulation enables long-term survival and becomes optimal when harsh conditions extend over timescales of months or longer. This trade-off between immediate cost and delayed benefit helps explain when sporulation is maintained or replaced by alternative strategies. By incorporating our estimates into mechanistic models, we show how metabolic constraints shape sporulation efficiency, while genome-wide mutation accumulation data reveal that even modest energetic burdens can become visible to selection, influencing the evolutionary fate of this complex and widespread trait.

energy | mutation | development | selection | dormancy

Significance

Evolutionary bioenergetics examines how energetic constraints influence the origin, maintenance, and evolution of cellular components and organismal traits. It takes a bottom–up approach to quantifying the ATP required to assemble biological structures ranging from genes and membranes to virus particles, whole cells, and multicellular organisms. Here, we apply this framework to estimate the full energetic cost of making a bacterial endospore, one of the most abundant and persistent forms of life on Earth. By accounting for macromolecule synthesis, regulatory checkpoints, maternal investment, and subcellular coordination, we identify conditions under which sporulation is favored rather than alternative strategies and reveal the evolutionary forces that have driven the repeated loss of this complex trait across lineages over billions of years.

Life requires energy to support growth, homeostasis, and long-term persistence. Energy is stored and transferred by molecules such as adenosine triphosphate (ATP), which fuel catabolic and anabolic reactions that govern basal metabolism, biosynthesis, and cellular repair (1, 2). These energy carriers also power gene regulation and translation, processes essential for building and maintaining cellular structures and organismal traits. In the early history of life, energetic constraints shaped the emergence of self-replicating systems and the evolution of core molecular functions (3). Today, the availability of energy continues to influence both physiological and developmental processes, from relatively simple biochemical pathways to more elaborate features such as secretion systems, sensory networks, and multicellular differentiation (4–6). While it is generally assumed that traits are maintained when fitness benefits outweigh their energetic costs, such trade-offs are rarely quantified, particularly under conditions that limit metabolism (7, 8). As such, the role of bioenergetics in determining the evolutionary fate of complex traits remains largely unresolved.

In nature, organisms often inhabit environments where energy is insufficient to support sustained growth and reproduction (9). To endure these unfavorable conditions, many species have evolved mechanisms of persistence that buffer against demographic variance and environmental stochasticity (10). Dormancy is one such strategy, whereby individuals enter a reversible state of reduced metabolic activity. In some cases, dormancy arises passively as metabolism slows in response to environmental stress. In others, it is complex and actively regulated, requiring precise gene expression and cellular remodeling (11). One particularly elaborate and well-characterized form of dormancy is endospore formation, which is found among bacteria like *Bacillus* and *Clostridia*. Endospores are among the most metabolically inert and long-lived biological entities (12). They are remarkably resistant to harsh environmental conditions, including elevated temperatures, high doses of ionizing radiation, extreme energy limitation, and the vacuum of space (13, 14). Consequently, endospores are widely distributed across environmental, engineered, and host-associated ecosystems. With an estimated 10^{28} individuals in marine sediments alone (15), endospores rank among the most abundant cell types on Earth.

Despite its protective benefits, sporulation comes with an energetic cost. Building a spore demands sustained investment of cellular resources over a prolonged developmental

Author affiliations: [a]Department of Biology, Indiana University, Bloomington, IN 47405; [b]School of Biological Sciences and The Center for Microbial Dynamics and Infection, Georgia Institute of Technology, Atlanta, GA 30332; and [c]Quantitative Life Sciences, The Abdus Salam International Centre for Theoretical Physics, Trieste 34151, Italy

Author contributions: C.K., W.R.S., and J.T.L. designed research; performed research; analyzed data; and wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: lennonj@iu.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2524274123/-/DCSupplemental.

Published February 6, 2026.

timeline. Roughly, 5% of the *Bacillus subtilis* genome is devoted to this process, encoding genes involved in the synthesis of spore structures (e.g., cortex and coat proteins), DNA packaging and repair, and signaling pathways that guide cell fate decisions (16). Sporulation is orchestrated by a precisely timed cascade of sigma factors, a family of proteins that enable RNA polymerase to bind specific gene promoters and initiate transcription (17, 18). Each sigma factor is activated in one of two cellular compartments, directing stage-specific gene expression during development (18). This tightly regulated sequence is punctuated by checkpoints that ensure the mother cell adequately provisions the developing forespore (18). However, sporulation alone does not guarantee evolutionary success. To realize a fitness benefit, the resting cell must ultimately be revived through the processes of germination and outgrowth. This developmental transition is regulated by mechanisms that integrate prior germinant and nutrient exposures, producing a memory-like influence on environmental evaluation and decision-making (19, 20). Although germination and outgrowth are faster than sporulation, they entail substantial proteome remodeling (21), adding further costs to the full dormancy cycle.

The energetic demands of spore formation and revival may help explain the distribution and maintenance of this ancient trait. An ancestral feature within the phylum Bacillota, sporulation is thought to have originated nearly three billion years ago (22). Over this geological timespan, it has been conserved in some groups, partially retained in others, and entirely eliminated in major clades such as *Staphylococcaceae* and *Lactobacillaceae*, with consequences for molecular evolutionary dynamics (23). Although costly traits can be undermined by cheaters (24), the benefits of sporulation are largely privatized because each cell's investment directly secures its own survival, which may explain its stability in some lineages despite apparent energetic demands. These dynamics are reflected in phylogenomic comparisons showing that non-spore-forming bacteria diversify more rapidly than their spore-forming counterparts (25), while experimental evolution demonstrates that the loss of sporulation ultimately reduces genetic diversity within populations (26).

The evolutionary loss of sporulation may occur through neutral processes, in which random mutations accumulate simply because the trait is no longer maintained by purifying selection. Over time, this can lead to pseudogenization, frameshift mutations, or partial pathway degradation. Because the underlying sporulation network represents a large mutational target, models predict that neutral decay of this kind could lead to loss of sporulation within approximately 10^8 generations (27). In contrast, selective pressures can drive the loss of sporulation, favoring deletions that reduce the energetic costs of maintaining unused machinery. These adaptive changes may be reflected in genome streamlining and biased gene loss in nonexpressed regions (28). Distinguishing between neutral and selective mechanisms is crucial for determining whether trait loss reflects passive genetic drift or adaptive optimization to new ecological conditions, thereby illuminating how complex life-history traits are maintained or discarded over evolutionary timescales.

Accurately quantifying the energetic costs of complex traits may clarify the conditions that favor their maintenance or loss. However, standard measurements of energy demand, such as oxygen consumption, heat production, or metabolite turnover, often lack precision, use units that are difficult to interpret in cellular terms, and are challenging to compare across biological scales. To address these limitations, we used a quantitative bioenergetic framework grounded in ATP equivalents, the universal currency of energy transfer in all forms of life (29).

By integrating genomic information with temporally resolved transcriptomic and proteomic measurements, we estimated the energetic demands of sporulation and revival (germination and outgrowth). This approach enabled us to quantify both opportunity costs (P_O) and direct costs (P_D) associated with precursor synthesis, gene expression, and proteome turnover across developmental time. With these estimates, we compared the energetic costs of dormancy to other components of the total cellular budget, including baseline energy demands and other persistence traits that may confer advantages under fluctuating or suboptimal conditions. Finally, we incorporated the estimated costs into a mechanistic model to examine how energetic constraints influence sporulation efficiency. By combining these results with genome-scale mutation accumulation data, we tested predictions about how bioenergetics shape the relative contribution of neutral and selective processes to the long-term maintenance of sporulation.

Results and Discussion

We begin with a detailed accounting of the energetic costs underlying the spore life cycle of the model bacterium *B. subtilis*, resolving patterns of energy expenditure over time, across developmental transitions, and within subcellular compartments. We use an empirical, demand-based approach that is grounded in measured ATP fluxes, biomass synthesis rates, and proteome composition (29), as opposed to supply-limited frameworks that emphasize constraints imposed by respiratory membrane area and mitochondrial amplification (3) (*SI Appendix*, Table S4). Our analysis shows that the complete spore life cycle demands a major energetic investment, consuming nearly 10^{10} ATP, roughly 10% of the total cellular energy budget. Most of this cost arises from opportunity costs (P_O), including the diversion of metabolic precursors, along with the direct synthesis of spore-specific macromolecules. By incorporating these estimates into a mechanistic model, we identify thresholds beyond which dormancy is no longer advantageous and demonstrate how energetic constraints limit sporulation efficiency, defined as the proportion of cells that complete the transition to an endospore. Based on genome-scale comparisons, our findings indicate that the energetic burden of sporulation can generate positive selection for mutations that disable this pathway, acting alongside neutral decay to drive its evolutionary loss. The bioenergetic framework developed here provides a conceptual and quantitative foundation for understanding how resource allocation strategies shape the ecological and evolutionary dynamics of complex traits.

Energetics of Spore Formation. We estimate the total cost of producing a single spore to be $\sim 2.4 \times 10^9$ ATP (P_T). The majority of this cost (80%) is attributable to opportunity costs (P_O), which reflect the diversion of precursor metabolites away from growth. The remaining 20% consists of direct costs (P_D), representing ATP hydrolysis events required for the biosynthesis and polymerization of molecular building blocks. Across both P_O and P_D, we found that translation accounts for the largest share of energy expenditure (68%), followed by genome replication (17%), and transcription (3%). Upon initiation of sporulation, the cell must duplicate its entire genome (30). This requires 4×10^8 P_T, on the same order as the 7.8×10^7 P_T needed to transcribe spore-related genes. Approximately 20% of *B. subtilis* genes participate in or are expressed during sporulation,

contributing 18% (7.2×10^7 P_T) of the total replication cost. Although membrane synthesis has been hypothesized to impose a substantial energetic burden on cells (31), we estimate that the lipid required for the 1 μm septum that separates the mother cell from the forespore only amounts to 12% of P_T incurred during sporulation.

Our findings reveal that spore formation is characterized by a pronounced asymmetry in energetic investment, with the mother cell bearing the majority of the biosynthetic burden. We estimate that the mother cell accounts for roughly 87% of the total costs, while the forespore contributes the remaining 13%. Approximately 67% of the total energy expenditure, including the cost of genome replication, occurs within the first hour of development. This early demand drops off exponentially as development progresses, a pattern reminiscent of maternal investment during animal embryogenesis (32). Following initiation, septum formation creates two unequally sized cellular compartments, establishing the physical basis for asymmetric investment. More than one hundred genes are activated within the forespore to support differentiation and maintain coordination with the mother cell (18). Throughout development, the mother cell delivers precursors and recycled building blocks to the forespore through a tubular intercellular channel (33, 34), reducing the need for de novo biosynthesis. This structure also mediates continued signaling between compartments during intermediate and late stages of development (35). Ultimately, the mother cell lyses, releasing the mature spore into the environment (Fig. 1).

Spore formation poses a risk for cells navigating unpredictable environments. Development unfolds slowly, requiring 8 to 10 h to complete compared to the 0.5 to 1.0 h division cycle that is typical during vegetative growth. As a result, cells may engage in a costly process that offers no benefit if conditions improve. Our accounting shows that by the time the cell reaches the commitment point, ~2 h after initiation, 85% of the total transcriptional and translational costs of sporulation have already been incurred. If precursor molecules are fully recycled, then only 20% of these costs are truly nonrecoverable because polymerization and activation are accounted for in P_D, whereas full recycling is credited in P_O.

The initial phase of sporulation nonetheless represents a significant upfront energetic commitment. Before reaching the developmental checkpoint, cells can still abort the process in response to environmental cues, but such reversal incurs a partial energy loss (35). Once SpoIIE activates asymmetric septum formation and σ^F is engaged (36), development becomes irreversible (Fig. 1). Beyond this threshold, the cell continues to invest heavily in spore-specific proteins, including structural components of the coat, small acid-soluble proteins that protect DNA from damage, and enzymes such as proteases that are required for revival (37–39).

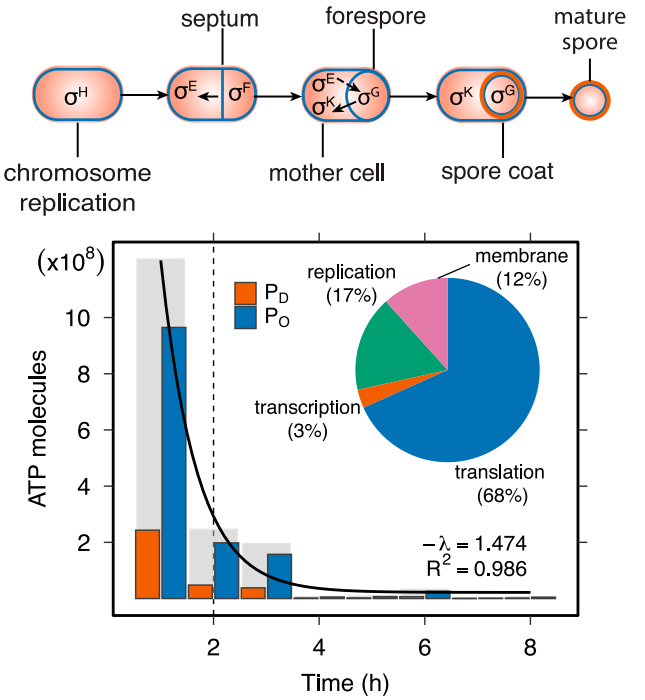

Fig. 1. Energetic cost of spore formation. *Top*: Major stages of endospore development. *Bottom*: Energetic cost of spore formation over time reflecting transcriptional and translational investments in units of ATP. Total costs (P_T) at each time interval (gray bars) represent the sum of opportunity costs (P_O), which are linked to biosynthesis (blue bars) and direct costs (P_D) associated with polymerization (orange bars). The total cost of spore formation declines exponentially over time. The embedded pie chart illustrates the proportional contribution (% of P_T) associated with key processes, including replication of the whole genome and septum synthesis, along with transcription and translation.

Energetics of Spore Revival. Spore revival, which includes germination and subsequent outgrowth, is essential to the success of dormancy. Our analyses reveal that revival is even more energetically costly than spore formation, with total expenditures of ~ 6.8×10^9 P_T (Fig. 2). Although germination (15 min) and outgrowth (3.5 h) proceed more rapidly than sporulation, they demand significant energy to reestablish vegetative growth, though the timing and extent of ATP accumulation during early spore revival remain debated (40). During the initial 15 min of revival, transcription and translation require ~7.6×10^8 P_T (1.5×10^8 P_D and 6×10^8 P_O), representing 12% of the total revival cost. The outgrowth phase, in contrast, accounts for the vast majority (88%) of the total energetic cost. Transcription and translation require ~5.4×10^9 P_T (1.1×10^9 P_D and 4.3×10^9 P_O), and membrane biogenesis adds another ~6.9×10^8 P_T (5.4×10^7 P_D and 6.4×10^8 P_O). Importantly, these estimates already account for the substantial recycling of spore membrane lipids during outgrowth: roughly one sixth of the membrane is reused, and only ~3% is newly synthesized at this stage (41, 42).

Our analysis reveals a highly uneven allocation of energy across core molecular processes during revival, with transcription accounting for just 4%, translation for 87%, and membrane assembly for 9% of total energy expenditure (Fig. 2). This elevated demand during revival reflects the rapid reactivation of biosynthetic pathways and reconstruction of the vegetative-cell proteome, in contrast to spore formation, which functionally relies on a smaller, specialized protein subset. Although revival-specific genes contribute an estimated ~4.6×10^7 (P_T), equivalent to approximately 11% of the energy required for genome replication, this cost is paid in advance at the onset of sporulation.

Despite their ability to rapidly reawaken, dormant spores contain remarkably little internal energy. Our bioenergetic accounting shows that a total of ~7.6×10^8 ATP are required for germination alone. However, ATP and GTP concentrations within *Bacillus* spores are extremely low, at around 2 nmol/g, as estimated using magnetic resonance spectroscopy with ^{31}P NMR (43). This corresponds to roughly 10^2 molecules per spore. Additional reserves including AMP (~400 nmol/g) and ADP (~100 nmol/g) (44), are also present in low

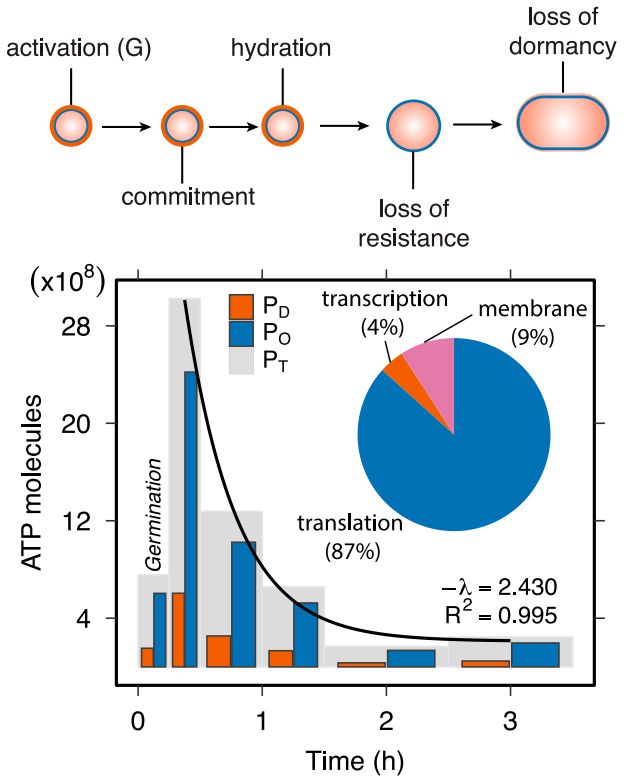

Fig. 2. Energetic cost of spore revival. *Top*: Major stages of spore revival, which include germination followed by outgrowth. *Bottom*: Energetic costs of spore revival over time reflecting transcriptional and translational investments units of ATP. Gray bars represent total costs (P_T) for each time interval. Total costs (P_T) at each time interval (gray bars) represent the sum of opportunity costs (P_O), which are linked to biosynthesis (blue bars) and direct costs (P_D) associated with polymerization (orange bars). The total cost of spore revival declines exponentially over time. The embedded pie chart illustrates the proportional contribution (% of P_T) associated with key processes, including transcription, translation, and membrane synthesis.

concentrations. Assuming a spore mass of ~200 femtograms (45), total endogenous stores amount to just 2×10^3 ATP equivalents. These estimates reveal that the energy required for germination exceeds internal reserves by nearly five orders of magnitude. While endogenous pools may be sufficient for initiating the early steps of germination, they are far from adequate to support the full transition from dormancy to active growth.

The apparent energy shortfall presents a fundamental problem concerning how a dormant spore generates enough ATP to complete germination. Our analysis suggests that spores overcome this limitation by mobilizing prepackaged molecular reserves that are metabolized immediately upon rehydration. Within 5 min of germination onset, essential components such as enzymes, ribosomes, amino acids, and nucleotides are activated to initiate core metabolic processes (44). A major source of amino acids is the rapid degradation of small acid-soluble proteins, which account for 10 to 20% of the spore's proteome (46). In addition, stored carbon sources, including 3-phosphoglyceric acid (~2,700 nmol/g) and malate (~3,000 nmol/g), fuel glycolysis and related metabolic pathways, yielding ~2.1×10^8 ATP in total (43, 47). These reserves are substantial but insufficient to cover the full germination demand; they likely bridge the earliest reactivation steps until exogenous carbon is taken up during outgrowth. However, full outgrowth and the return to vegetative growth requires additional energy from the external environment. Increased glucose uptake during outgrowth likely supports the elevated ATP production needed for biosynthesis and cell expansion (48). Because both spore formation and revival impose substantial energetic costs, their evolutionary persistence must be evaluated relative to alternative cellular stress responses, which we explore in the next section.

Head-to-Head Bioenergetics of Microbial Survival. Spore-forming bacteria like *Bacillus* deploy tiered stress responses that vary in reversibility, energy cost, and benefits over time. The early responses, coordinated by the sigma factor σ^B, involve low-cost physiological adjustments such as DNA repair, redox homeostasis, and osmoprotection (49, 50). These transient changes help cells buffer short-term environmental fluctuations without altering their developmental trajectory. If conditions worsen, cells may adopt facultative strategies such as motility (51), cannibalism (52), oligotrophic survival (53), or competence (54), which offer moderate costs and reversible outcomes. When these options are insufficient, cells initiate spore formation, a developmental program that requires a major upfront investment but offers long-term protection.

We define the total build cost as the ATP required for transcription and translation. By this measure, we estimate that the *Bacillus* spore life cycle requires about 7.8×10^9 ATP. When additional infrastructure costs are included, such as genome replication (~4.0×10^8 ATP), septum synthesis (~2.8×10^8 ATP), and membrane remodeling during revival (~6.9×10^8 ATP), the complete energetic investment rises to ~9.2×10^9 ATP, nearly 10 billion overall (Fig. 3).

To facilitate meaningful comparisons across different life history strategies, we express all costs on a per-generation basis. For vegetative growth, one generation corresponds to 1.16 h at 20° C, whereas a full spore life cycle, which includes formation, germination, and outgrowth, spans 11.5 h. For developmental programs such as biofilm formation (~12 h) and competence development (~3 h), we expressed total costs on a per generation basis by normalizing them to program duration. For other traits such as flagella and stress response proteins, we assumed complete resynthesis each generation, providing conservative upper bounds. In reality, many vegetative structures persist across generations (55), which further emphasizes the high relative cost of sporulation. Maintenance energy (~1.3×10^9 ATP) represents the per-generation operating cost of core housekeeping metabolism and is distinct from the build costs of homeostasis proteins, which are synthesized once to provide stress-response capacity. For reference, genome replication (~4.0×10^8 ATP) and membrane bilayer synthesis (~2.8×10^9 ATP) are shown on the same per-generation scale. Together, these comparisons indicate that the spore life cycle is among the most energy-intensive processes in the bacterial cell (Fig. 3).

Although seemingly expensive, spore formation is a front-loaded strategy whose benefits accumulate with time. We illustrate this by comparing a cell that sporulates with one that remains vegetative but does not grow. Using a maintenance rate of ~1.2×10^9 ATP h^{-1} at 20°C, the break-even time is

$$t^* = \frac{P_T^{\text{spore cycle}}}{C_M} \approx \frac{9.2 \times 10^9}{1.2 \times 10^9} \approx 7.7 \text{ h}.$$

If a cell can reduce its maintenance energy requirements tenfold to (~1.2×10^8 ATP h^{-1}), the break even time (t^*) becomes ≈ 77 h (3 d). With a hundredfold reduction to (~1.2×10^7 ATP h^{-1}), the break even time is ≈ 770 h (32 d). Thus, without committing to sporulation, *Bacillus* can remain viable for extended

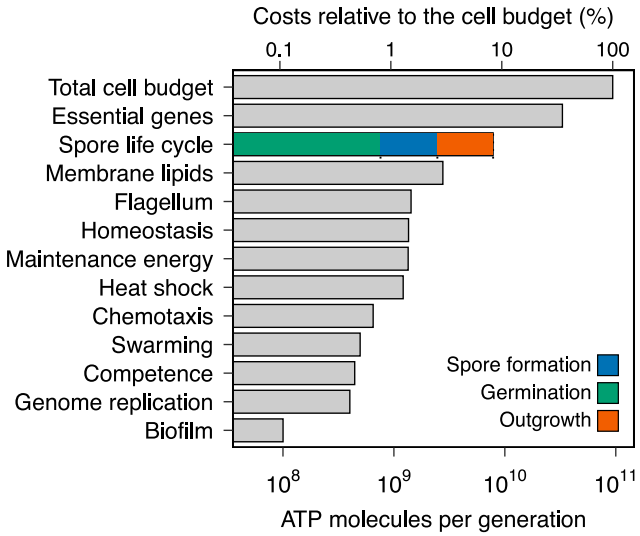

Fig. 3. Energetic costs of sporulation and other cellular processes and traits. Bars represent build costs (transcription + translation) in units of ATP molecules. The spore life cycle, with colored segments denoting germination, spore formation, and outgrowth, represents the total cost of one complete spore generation (11.5 h). All other bars show costs per vegetative generation (1.16 h at 20°C). For developmental programs (biofilm formation and competence), total costs were divided by program duration; for other traits, we assumed per-generation resynthesis. Genome replication and membrane lipid synthesis are included as reference benchmarks. The secondary axis shows costs as a percentage of the per-generation cellular budget ($C_G + t_{gen}C_M$). See *Materials and Methods* and *SI Appendix*, Table S3 for calculation details.

periods at low power draw (53). However, under prolonged or unpredictable stress, the one-time cost of dormancy becomes the energetically optimal investment. Note that maintenance and growth conditions are temperature dependent, so these values will change accordingly.

Although commonly viewed as a last resort, sporulation can represent a bioenergetically optimal strategy in fluctuating environments. Such conditions are widespread in nature, including energy-limited soils and sediments where resource inputs are both variable and vanishingly small (9). By comparing the energy costs of microbial stress responses in a common currency, our analysis reframes dormancy not as a passive fallback but as a strategic investment shaped by ecological, evolutionary, and developmental constraints.

Costs on Collective Outcomes: Sporulation Efficiency. Understanding the energetic costs of sporulation at the cellular level has the potential to explain how dormancy strategies scale up to shape population-level phenomena. Excluding transitional states, an individual *Bacillus* exists either as a metabolically active vegetative cell or as a dormant spore (Figs. 1 and 2). Yet within a given environment, populations rarely exhibit uniform behavior. Instead, the fraction of individuals that undergo sporulation, known as sporulation efficiency (φ), is highly variable.

In batch culture, where rapid physiological shifts and unbounded population growth eventually lead to resource exhaustion, our analysis of literature-reported data shows that the median sporulation efficiency is roughly 30%, with values ranging from 0 to 100% (Fig. 4*A*). Some of this variation likely reflects details of experimental design, including differences in strain background, resources, temperature, and sampling methodology. Incomplete sporulation has also been interpreted as the outcome of stochastic decision-making or bet-hedging strategies, in which subsets of cells commit to dormancy while others remain vegetative to exploit residual resources or potential improvements in environmental conditions. While such strategies are evolutionarily important, they modulate the long-term fitness consequences of existing costs but do not reduce the cost of an individual spore.

Because all cellular features require energy to build and maintain, we examined how our estimates of energetic costs influence sporulation efficiency. To investigate this relationship, we developed a population dynamics model in an environment where resources decline over time:

$$\frac{dN_v}{dt} = N_v\left[g_v(R) - f(R)\cdot g_s(R)\right], \quad [1]$$

$$\frac{dN_s}{dt} = N_v\cdot f(R)\cdot g_s(R), \quad [2]$$

$$\frac{dR}{dt} = -N_v\left[\frac{g_v(R)}{Y_v} + \frac{f(R)\cdot g_s(R)}{Y_s}\right]. \quad [3]$$

Here, N_v and N_s represent the concentrations of vegetative cells and spores, respectively. The functions $g_v(R)$, $g_s(R)$, and $f(R)$ represent the per capita rates of cellular growth, sporulation, and the initiation of sporulation, each dependent on the concentration of available resources R. The yields Y_v and Y_s describe the efficiency of resource use for producing cells and spores, respectively, with ATP used as the unit of energetic currency (see Methods and *SI Appendix*).

Our model predicts that sporulation efficiency declines sharply when energetic costs exceed empirically derived estimates (Fig. 4*C*). This outcome reflects a trade-off between resource availability and the energy required for spore formation, which limits the number of cells that can successfully sporulate. By integrating empirical measurements with mechanistic modeling, we show that the relative cost of producing a spore compared to a vegetative cell can constrain the success of dormancy as a survival strategy. As a point of comparison, we developed the chemostat equivalent of our model and found that steady-state sporulation efficiency does not depend on the cost of building a spore or a cell (*SI Appendix*).

Evolutionary Maintenance of an Energetically Costly Trait. Sporulation is an ancient survival strategy that originated billions of years ago. It confers resistance to harsh conditions, enables long-distance dispersal, and protects against viral infection (14, 56, 57). These advantages have contributed to the prevalence of spore-forming lineages across the globe. Yet despite its benefits, this complex form of dormancy remains susceptible to evolutionary decay. Repeated and independent losses of sporulation have occurred across the Bacillota phylogeny (26, 58), often attributed to relaxed selection in environments that favor continuous growth (58). In such settings, sporulation genes may no longer provide a fitness benefit and become prone to mutational degradation (59). In addition to neutral decay, the energetic burden of unused genetic material may drive the adaptive deletion of sporulation loci, as even nonexpressed genes and regulatory elements can impose fitness costs (60). The relative importance of neutral vs. selective forces in the loss of sporulation, however, remains unresolved.

When growth is favored, sporulation genes are typically not expressed, rendering their transcriptional and translational costs negligible. However, DNA replication still requires energy and

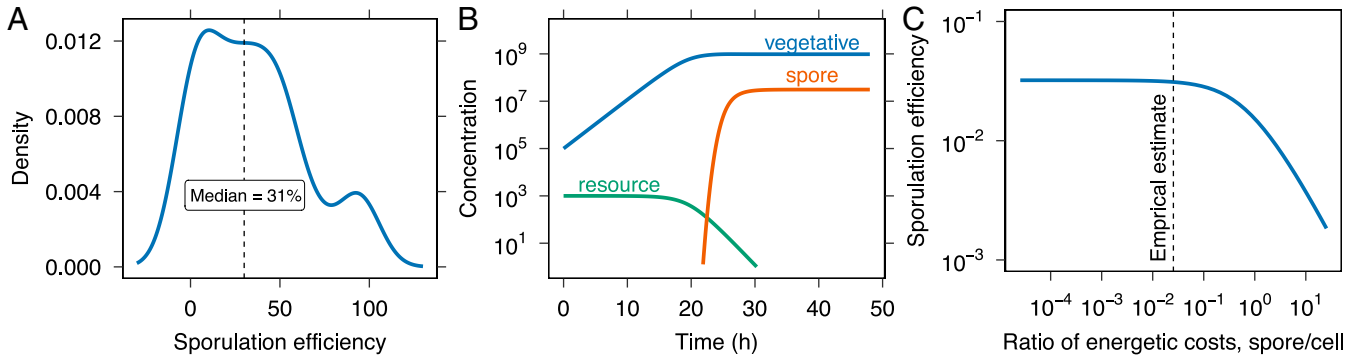

Fig. 4. Energetic constraints on sporulation efficiency. (*A*) The frequency of sporulation efficiency (fraction of cells that sporulate in a population) derived from publicly available data (*SI Appendix*, Table S2). (*B*) In a dynamical model, resource depletion reduces cell growth and triggers spore formation. (*C*) An increase in spore costs relative to cell costs results in a monotonic decline in sporulation efficiency. The dashed line represents the empirical estimate of the spore-to-cell cost ratio. The sharp decrease in efficiency as costs rise suggests that sporulation becomes energetically unfavorable with increasing costs.

resources. If even small costs are visible to selection, sporulation genes may be lost through the fixation of beneficial deletions. This contrasts with trait loss driven by neutral processes, in which mutations accumulate gradually through nucleotide substitutions. In *Bacillus*, the estimated rate of deletions is 1.2×10^{-10} per site per generation, while the rate of nucleotide substitution is 3.35×10^{-10} per site per generation, making substitutions nearly three times more frequent (61). Whether the energetic cost of maintaining unused sporulation genes is sufficient to favor deletions over neutral decay depends on the strength of selection and the size of a deletion. To examine this, we compared the fixation rate of a beneficial deletion to that of a neutral substitution:

$$\frac{d_{\text{del}}}{d_{\text{sub}}} = \frac{2N_e s}{1 - e^{-2N_e s}} \times \frac{U_{\text{del}}}{U_{\text{sub}}}, \qquad [4]$$

where N_e is the effective population size, s is the selection coefficient for a deletion of a given size, U_{del} is the rate of deletions, and U_{sub} is the rate of substitutions.

To examine how the relative fixation rate depends on deletion size, we combined our energetic cost estimates with published values for N_e and the empirical distribution of deletion sizes from mutation accumulation experiments (61, 62) (Fig. 5*A*). Although deletions occur less frequently than substitutions, selection acting on large deletions (~1 kb) is strong enough to outweigh the rate of neutral substitutions. This length corresponds to the typical size of a gene, providing evidence that selection can drive the loss of sporulation in bacterial populations (Fig. 5*B*). While previous studies have emphasized mutational degradation under relaxed selection (58), our analysis shows that the energetic burden of maintaining unused sporulation genes can be sufficient to favor their adaptive deletion. Even small metabolic costs, if persistent, may become visible to selection (27, 63). Observations from host-associated Bacillota, where sporulation loss often coincides with genome streamlining and metabolic specialization, support the idea that energy-driven processes shape trait loss in natural populations under real-world conditions (64, 65).

Theoretical treatments of sporulation and dormancy often emphasize their role as bet hedging strategies that promote survival under environmental uncertainty (66). Other modeling approaches have cast sporulation as an optimal switching problem (67), dissected the genetic circuitry that governs commitment (68, 69), and explored trade-offs between spore quality and quantity (70). These perspectives illuminate the conditions under which sporulation is expected to be advantageous, but they generally treat its costs only abstractly or implicitly. By quantifying the energetic demands of the spore life cycle, our analysis shows how these costs influence the selective conditions under which sporulation is maintained, and importantly, the circumstances in which it can be lost.

The loss of complex traits is a recurring theme in evolution, often attributed to relaxed selection and the accumulation of neutral mutations (71, 72). Yet even in classic examples such as eye reduction in cave-dwelling fish or flight loss in birds, growing evidence suggests that the regression of biological function can be favored by selection, especially when a trait imposes physiological burdens or energetic costs (73, 74). Our study provides a first-principles demonstration of how the energetic cost of a complex trait can influence its evolutionary fate. More broadly, the trade-offs we identify may apply to other metabolically expensive traits across the tree of life, including bioluminescence, motility, and secondary metabolite production in microbes, symbiotic nodulation and inducible defenses in plants, and elaborate courtship displays or environmentally contingent morphologies in animals. Although our model focuses on sporulation, a widespread form of dormancy that supports the persistence and dispersal of a globally abundant group of bacteria, the broader approach of quantifying energetic investment and linking it to evolutionary dynamics provides a scalable framework for understanding the distribution and loss of complex traits in nature.

Materials and Methods

Bioenergetic Accounting: *Definitions and Assumptions*. We estimated bioenergetic costs using glucose as the sole carbon and energy source. In bacteria, glucose is metabolized through the tricarboxylic acid (TCA) cycle, yielding ATP. The hydrolysis of ATP to ADP and inorganic phosphate (P_i) releases approximately 30 kJ/mol of free energy (ΔG) under standard conditions. While some ATP is used to power energy-consuming processes, other hydrolysis events support the biosynthesis of macromolecular building blocks. Following established conventions (7, 29), we estimate cellular energy expenditure in ATP (or ATP-equivalents) and use *P* to denote one hydrolysis event (GTP, etc., counted 1:1 as *P*). Following refs. 29 and 75, we partition costs into *direct* (P_D) and *opportunity* (P_O). Direct costs capture ATP-powered steps such as monomer activation/charging and polymerization (e.g., aminoacyl-tRNA charging, chain elongation). Opportunity costs capture the energetic value of precursor synthesis (e.g., NADPH-consuming reductions) that forgoes growth if diverted to a trait. Total cost is $P_T = P_D + P_O$. We report P_T unless stated otherwise, because it best reflects resource allocation under balanced growth. However, P_D is useful when considering heat dissipation and instantaneous power, and P_O is useful

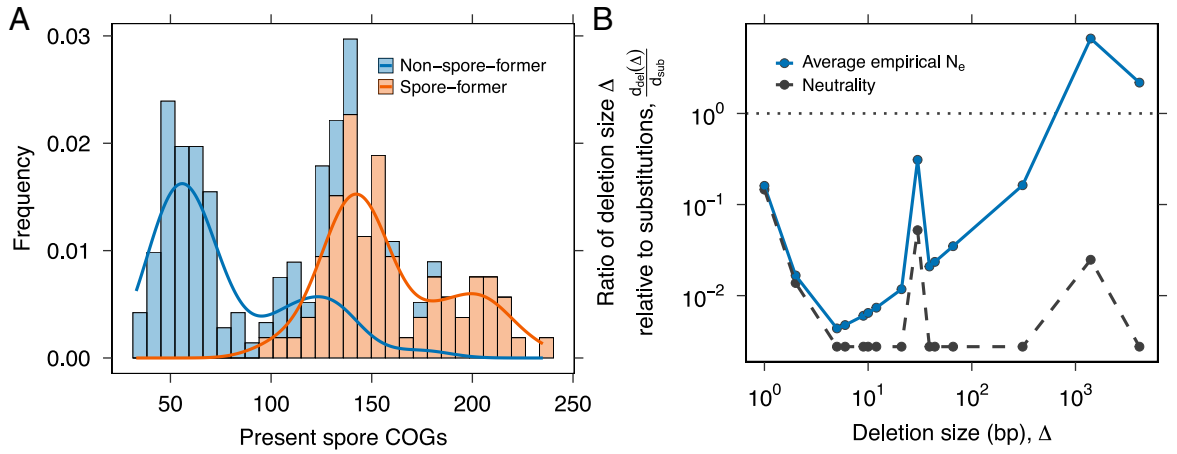

Fig. 5. Evolutionary maintenance of sporulation. (*A*) Conservation of sporulation-related genes across spore-forming and non-spore-forming lineages in the Bacillota phylum based on the frequency of COGs (Clusters of Orthologous Genes). There are 237 COGs in total (23). (*B*) The loss of sporulation genes in *Bacillus* can be examined by combining energetic cost estimates, which inform the strength of selection (*s*), with empirical mutation rates (61). Here, we calculate the fixation rate of beneficial deletions relative to that of neutral substitutions (analogous to dN/dS in the population genetics literature; Supporting Information). The black dashed line shows the neutral expectation ($s = 0$), where the ratio is determined solely by the mutation rates of deletions and substitutions. The solid blue line shows the average nonneutral case using published estimates of N_e (61, 62). For deletion sizes where the ratio is greater than 1, the energetic cost of maintaining sporulation genes is sufficient for selection to favor their loss. The sharp peaks reflect the empirical distribution of deletion sizes.

when considering that some monomers can be recycled during processes. We assume 20°C for budget references and use the growth (C_G) and maintenance (C_M) entries of ref. 29 where indicated.

It is important to note that biosynthetic costs vary across species due to differences in metabolic pathways and environmental factors such as resource availability (76). Our analysis focused on cellular processes with sufficient quantitative data, including those related to the central dogma and membrane synthesis. We excluded minor costs below approximately 10^5 ATP, such as posttranslational modifications, and protein folding (29), although even these small expenditures may be evolutionarily significant in certain regulatory contexts (60).

Replication costs. Costs associated with the replication of the mother cell genome are incurred early in the spore cycle, before the commitment stage ~2 h. Although DNA replication is a complex process involving unwinding, primer synthesis, Okazaki fragment ligation, and proofreading, most of the energy is spent on synthesizing nucleotide building blocks and polymerizing them into DNA.

Let L_g be the genome length (bp). With $\bar{c}_s$ the *direct* synthesis cost of a DNA nucleotide (excluding polymerization; $\approx$ 11 P_D), c_p the polymerization cost (2 P_D/nucleotide), c_{hel} = 1 P_D/bp for helicase unwinding, and $c_{\text{prim}} = 0.32\ P_D$/bp for lagging-strand primers (29, 75), we write:

$$C_{R_D} = 2L_g(\bar{c}_s + c_p) + L_g(c_{\text{hel}} + c_{\text{prim}}). \quad [5]$$

The *opportunity* cost uses the average opportunity per DNA nucleotide, $\bar{c}_o \approx 34\ P_O$ (precursor synthesis ~ 33 P_O plus conversion), giving

$$C_{R_O} = 2L_g\,\bar{c}_o. \quad [6]$$

Ligation and proofreading are $< 10^{-4}$ of C_{R_D} and are neglected.

Transcription costs. Transcription is also affected by complex mechanisms such as activation, initiation, termination, proofreading, and RNA processing. From a bioenergetic perspective, however, most of the transcription budget is dedicated to nucleotide synthesis and polymerization (29). Therefore, the expression costs of genes can be estimated as the sum of the costs associated with protein-coding genes involved in spore formation and spore revival. Like replication, transcription costs have two components: direct costs associated with the polymerization and the opportunity costs, which are the energy needed to synthesize the ribonucleotide building blocks. Assuming that the length of genes and mRNAs (premature and mature) are approximately the same as in other bacteria (29), we estimate the direct costs of transcription as follows:

i) One-time costs: We assume efficient nucleotide recycling, so nucleotide synthesis for each transcript is charged once. Let $L_{\text{mRNA},j}$ be the length of transcript j, $N_{\text{mRNA},j}$ the number of transcripts produced over the program, and $\bar{c}_{sr}$ the direct synthesis cost per RNA nucleotide (i.e., RNA direct cost minus polymerization; $\bar{c}_{sr} \approx \bar{c}_{r,\text{dir}} - 2\ P_D$). The corresponding opportunity cost uses $\bar{c}_{or}$ per nucleotide:

$$C_{T_D}^{(\text{synth})} = \bar{c}_{sr} \sum_{j=1}^{k} \left(L_{\text{mRNA},j}\, N_{\text{mRNA},j}\right) \quad [7]$$

$$C_{T_O} = \bar{c}_{or} \sum_{j=1}^{k} \left(L_{\text{mRNA},j}\, N_{\text{mRNA},j}\right) \quad [8]$$

Here $\bar{c}_{sr}$ is the average direct costs of synthesizing an RNA nucleotide (10 P_D), and $\bar{c}_o$ is the average opportunity cost of synthesizing an RNA nucleotide, which is $\approx$ 31.5 P_O.

ii) Time-resolved polymerization: Polymerization is paid whenever a transcript is made. We distribute the transcript production of gene j across hours t using weights $w_{j,t}$ proportional to the measured expression at hour t, with $\sum_t w_{j,t} = 1$. If $R_j(t) = N_{\text{mRNA},j} w_{j,t}$ transcripts are made at hour t, the hourly direct cost is

$$C_{T_D}(t) = \sum_j L_{\text{mRNA},j}\, R_j(t)\, (c_{pr} + c_{sc}), \quad [9]$$

where $c_{pr} = 2\ P_D$ per nucleotide is the RNA polymerase elongation cost, and c_{sc} is a per nucleotide small direct overhead for relieving transcription-induced supercoils, motivated by the twin-domain model of Liu and Wang (77) and the mechanochemistry of DNA gyrase ($\approx$ 2 ATP per negative supercoil); we parameterize this as $c_{sc} = 0.1\ P_D$/nucleotide, typically <5% of polymerization. The totals over the program are $C_{T_D} = C_{T_D}^{(\text{synth})} + \sum_t C_{T_D}(t)$ and C_{T_O} from Eq. **8**.

To connect to protein demand, we estimate $N_{\text{mRNA},j}$ from protein abundances (parts per million, ppm) and a yield Y of proteins per mRNA lifetime: $P_j = (\text{ppm}_j/10^6)\ P_{\text{total}}$, $N_{\text{mRNA},j} = P_j/Y$. We use $P_{\text{total}} \approx$ 1,774,445 proteins per cell (78) and a representative Y = 100. Because this yield-based accounting already fixes the number and timing of transcript synthesis via $w_{j,t}$, we do not introduce an explicit mRNA degradation rate in the cost formulas to avoid double-counting.

Translational costs. Although mRNA has a shorter half-life and ribonucleotide polymerization is relatively expensive, proteins are 100 to 1,000 times more abundant than transcripts (29). In addition, nucleotides are already activated molecules, whereas tRNA must be charged with amino acids, making chain elongation more energetically demanding (29). As a result, most of the energetic cost of protein synthesis is attributed to translation, while processes such as initiation, termination, and posttranslational modification contribute minimally to total expenditure (29). We estimated the per-cell translation costs as follows:

$$C_{Pr_D} = \sum_{j=1}^{k} LC_j^{(D)} N_{\mathrm{pr},j}, \quad [10]$$

$$C_{Pr_O} = \sum_{j=1}^{k} LC_j^{(O)} N_{\mathrm{pr},j}, \quad [11]$$

where $LC_j^{(D)}$ includes $\sim 4\,P_D$ per incorporated amino acid for aminoacylation and elongation, and $LC_j^{(O)}$ uses amino acid opportunity costs [bacterial mean ($\sim 25\,P_O$)] (29). Protein degradation during the sporulation window is slow relative to synthesis, so we neglect turnover terms.

Beyond elongation and aminoacylation, we add a constant $2\,P_D$ per completed protein to account for one GTP at initiation (IF2) and one GTP during termination/recycling (RF3), consistent with standard bacterial translation cycles (79). Other small terms (e.g., proofreading, chaperone cycles) were not included, consistent with our $< 10^5$ ATP cutoff.

Membrane synthesis and remodeling. We estimated the energetic cost of synthesizing and remodeling lipid bilayers for the cell envelope and sporulation-specific structures. The number of lipids is the bilayer area divided by the headgroup area $a_1 = 0.65\,\mathrm{nm}^2$ (80–82). We modeled the vegetative cell as a *spherocylinder* (cylinder + hemispherical caps) with length $L = 2.5\,\mu\mathrm{m}$ and diameter $D = 1\,\mu\mathrm{m}$ (41). Writing $a = L/2$ and $b = D/2$, and using bilayer thickness $h = 4\,\mathrm{nm}$ (3.0 nm hydrophobic core + 0.5 nm headgroup radius per leaflet) (83, 84), the outer and inner leaflet areas are

$$A_{\mathrm{outer}} = 4\pi ab, \qquad A_{\mathrm{inner}} = 4\pi(a-h)(b-h). \quad [12]$$

Assuming a membrane protein occupancy of $\phi_{\mathrm{prot}} = 0.5$, the lipid count is

$$N_{\mathrm{lipid}} = \frac{A_{\mathrm{outer}} + A_{\mathrm{inner}}}{a_1}\left(1 - \phi_{\mathrm{prot}}\right). \quad [13]$$

We take unit lipid costs as $18\,P_D$ and $212\,P_O$ (75), giving

$$P_D^{\mathrm{mem}} = 18\,N_{\mathrm{lipid}}, \qquad P_O^{\mathrm{mem}} = 212\,N_{\mathrm{lipid}}. \quad [14]$$

The septum is modeled as a flat circular bilayer (thickness effects negligible relative to b), with total bilayer area:

$$A_{\mathrm{sept}} = 2\pi b^2. \quad [15]$$

The costs follow analogously with the same ϕ_{prot} and unit costs.

During revival (0 to 3.5 h), we assume that only a fraction f_{rev} of the full vegetative membrane is newly synthesized and that a fraction $f_{\mathrm{recycle}} = 1/6$ of lipid cost is offset by recycling. Net revival membrane cost is

$$P_T^{\mathrm{mem,rev}} = f_{\mathrm{rev}}\left(1 - f_{\mathrm{recycle}}\right) P_T^{\mathrm{mem}}. \quad [16]$$

We use $f_{\mathrm{rev}} = 0.30$ in the main analysis.

Empirical Data Used for Bioenergetic Accounting. We used various publicly available cellular and molecular information to support this theoretical framework.

Gene sets and lengths. Gene and protein metadata (including lengths and SubtiWiki symbols) were obtained from SubtiWiki (85). To unify identifiers, we normalized gene tokens and propagated locus tags across synonym lists.

Reference genome and counting conventions. All genome-level quantities were referenced to *B. subtilis* subsp. *subtilis* strain 168 (NCBI RefSeq NC_000964.3; assembly GCF_000009045.1, ASM904v1). The genome length used in our calculations is $L_g = 4{,}215{,}606$ bp. Unless stated otherwise, gene counts refer to protein-coding loci (CDS); the total number used here is $n_{\mathrm{CDS}} = 4{,}237$. Noncoding RNAs were excluded from gene-count totals but may appear in expression sets where relevant. All percentages "of the genome" and "of genes" are computed relative to these values.

Protein sequences and per-protein costs. Reviewed UniProt sequences for *B. subtilis* were mapped deterministically to locus tags via BSU identifiers in name/gene fields, then through unambiguous SubtiWiki primary/alias tokens. Per-protein direct and opportunity costs ($LC_j^{(D)}$, $LC_j^{(O)}$) were computed by summing amino acid costs from ref. 75 over each sequence.

Protein abundances. PaxDb abundances (ppm) (86, 87) were tokenized and mapped to locus tags using the same symbol map. When a PaxDb row matched multiple tags, ppm was split evenly (mass-conserving). Proteins without a match received a tiny floor (0.1 ppm), then all ppm were renormalized to the known total protein count $P_{\mathrm{total}} = 1{,}774{,}445$ (78) to obtain copies per cell.

Sporulation time series. Expression heatmaps from SporeWeb (88, 89) were converted to long format (t_0–t_8). For each gene j, hourly weights $w_{j,t} \propto$ normalized expression were used to i) assign one-time transcript synthesis at first appearance and ii) distribute polymerization across hours (*Transcription Costs*).

Revival (germination and outgrowth). Newly synthesized proteins from Swarge et al. (21) were converted to interval scores as successive differences (H0.25, H0.5, H1, H1.5, H2.5, H3.5). We excluded H5.5 (210 to 330 min) to avoid vegetative growth. For each protein, interval weights (normalized positive scores) distribute copies and thereby transcription/translation costs over time. Revival membrane remodeling used $f_{\mathrm{rev}} = 0.30$ with $f_{\mathrm{recycle}} = 1/6$ as above.

Handling of missing fields. When gene or protein length was unavailable in SubtiWiki or the mapped UniProt entry, we substituted the dataset median (gene length and protein length medians computed over the available entries). If a protein lacked sequence-derived per-protein cost totals, we estimated them as (length × median per-amino acid cost) computed across proteins; if length was also missing, we used protein-level medians. For proteins without a PaxDb match we assigned a small abundance floor (0.1 ppm) and then renormalized all ppm values to the total protein count, ensuring mass conservation.

Head-to-head bioenergetics. To compare energetic costs of sporulation with other cellular processes, we expressed all trait costs on a per-generation basis. For vegetative traits, one generation corresponds to the division time of approximately 1.16 h at 20°C. For the spore life cycle, one generation corresponds to the complete developmental program (spore formation, germination, and outgrowth), which spans ~11.5 h. This framework enables direct comparison of the metabolic burden imposed by different life history strategies operating on distinct timescales.

We define build costs as the sum of ATP expenditures for transcription and translation required to synthesize the proteins associated with each trait, excluding maintenance, protein turnover, and operational costs. For the spore life cycle, build costs were calculated by integrating time-resolved gene expression data for sporulation (88) and revival (21) over the developmental program. For other cellular traits, build costs were estimated from steady-state protein abundances during growth (86, 87).

Gene sets for biofilm formation, motility, chemotaxis, competence, and stress responses were curated from SubtiWiki (85), mapped to locus tags, and merged with protein abundances. Transcription costs were calculated assuming steady-state transcript abundance given by a protein-to-mRNA yield ratio of $Y = 100$. Translation costs were computed as protein copy number multiplied by per-amino acid synthesis costs (direct + opportunity) (75).

For developmental programs operating over multiple vegetative generations (biofilm formation, competence development), total build costs were divided by program duration to obtain per-generation equivalents. For other traits, we assumed complete per-generation resynthesis (i.e., costs calculated directly from steady-state abundances represent per-generation values), providing conservative upper bounds given that some cellular structures may persist across cell divisions. For traits exhibiting population heterogeneity (e.g., biofilm

formation, competence), bulk proteomic measurements yield population-averaged costs. These values represent the mean energetic burden across all cells, including both producers and nonproducers, and may underestimate per-producer investment. However, population-averaged costs remain appropriate for evolutionary analyses, as selection shapes population-level fitness.

All costs are expressed relative to the per-generation budget $C_{\text{total}} = C_G + t_{\text{gen}} C_M \approx 9.4 \times 10^{10}$ ATP at 20°C (29), where C_G is the growth cost and C_M is the maintenance energy. Detailed calculation methods, data sources, and assumptions for each trait are provided in *SI Appendix*, Table S3.

Bioenergetic Model of Spore Efficiency. To investigate how energy limitation influences sporulation efficiency, we developed a population dynamics model of *Bacillus* that incorporates the metabolic demands of endospore formation and vegetative cell production. We used bioenergetic yield parameters to capture how efficiently resources are converted into either spores or cells. The yields of vegetative cells (Y_v) and spores (Y_s) are defined as

$$[Y_v] = \frac{\text{cell}}{\mu g \text{ resource}} = \frac{\text{ATP}}{\mu g \text{ resource}} \cdot \frac{\text{cell}}{\text{ATP}} \quad [17]$$

$$[Y_s] = \frac{\text{spore}}{\mu g \text{ resource}} = \frac{\text{ATP}}{\mu g \text{ resource}} \cdot \frac{\text{spore}}{\text{ATP}}. \quad [18]$$

Using our ATP estimates, we can parameterize these yields:

$$Y_v = \frac{\epsilon}{C_v}, \qquad Y_s = \frac{\epsilon}{C_s},$$

where ϵ represents the ATP produced per (μg) of resource, reflecting the efficiency of energy extraction from a given substrate. These expressions link the yields of cells and spores, reducing the degrees of freedom in our model by one. Given our definition of yield and the focus of our study on bioenergetics, it is important to explicitly account for the consumption of a limiting resource (R). To do so, we apply a minimal model of microbial growth in batch culture, defined as follows:

$$\frac{dN_v}{dt} = N_v \left[\underbrace{g_v(R)}_{\text{Cell growth}} - \underbrace{f(R) \cdot g_s(R)}_{\text{Spore formation}} \right], \quad [19]$$

$$\frac{dN_s}{dt} = N_v \left[\underbrace{f(R) \cdot g_s(R)}_{\text{Spore formation}} \right], \quad [20]$$

$$\frac{dR}{dt} = -N_v \left[\underbrace{\frac{g_v(R)}{Y_v}}_{\text{Consumption}} + \underbrace{\frac{f(R) \cdot g_s(R)}{Y_s}}_{\text{Consumption}} \right], \quad [21]$$

$$g_v(R) = r_{\max}^{(v)} \frac{R}{R + K_v}, \quad [22]$$

$$g_s(R) = r_{\max}^{(s)} \frac{R}{R + K_s}, \quad [23]$$

where r_{max} is the maximum growth rate, and K is the half-saturation constant, specifying the value of R where $g(R)$ is half of r_{max}. The rate of spore initiation is modeled as the following function:

$$f(R) = \frac{1}{1 + e^{\sigma(R - R_{\min})}}. \quad [24]$$

The sporulation initiation function exhibits intuitive properties. The sensitivity parameter σ governs the steepness of the transition from a low to high rate of spore formation. When $\sigma \gg 1$, the function approximates a Heaviside step function (90):

$$f(R) = \begin{cases} 1 & \text{for } R(t) \le R_{\min} \\ 0 & \text{for } R(t) > R_{\min}. \end{cases}$$

The inclusion of a spore germination term is unnecessary for a deterministic system where the resource concentration only decreases over time.

Evolutionary Maintenance of Sporulation. In this section, we use a minimal model of natural selection to illustrate the selective advantage conferred by the loss of a single nonfunctional nucleotide via a mutation-induced deletion. We focus on a scenario in which a population of spore-forming *Bacillus* inhabits an environment that supports continuous growth, making sporulation unnecessary. In this context, sporulation genes are not expressed and become effectively nonfunctional. As a result, they fall under relaxed selection (91), allowing mutations to accumulate over time, including those that would be deleterious in environments where sporulation is required.

However, nonfunctional DNA can still impose a metabolic cost. A cell that loses a single nucleotide through a mutation-induced deletion may gain a fitness advantage relative to its ancestor due to having been relieved of this slight, though existent, metabolic burden. This advantage can be calculated as the number of ATP required to build nucleotides ($c_n \approx 50$ ATP), expressed relative to the cell's total energetic budget (29). In the case of *Bacillus*, the selective advantage of the mutant can be calculated as

$$s = \frac{c_n}{C_{\text{cell}}}, \qquad C_{\text{cell}} = C_G + t_{\text{gen}} C_M. \quad [25]$$

Under the strong-selection, weak-mutation evolutionary limit, the probability that a mutation of frequency f will fix can be derived as (92):

$$P_{\text{fix}}(s, f) = \frac{1 - e^{-2Nsf}}{1 - e^{-2Ns}}. \quad [26]$$

A de novo mutation has an initial frequency of $\frac{1}{N}$, from which we obtain for $s \ll 1$:

$$P_{\text{fix}}\left(s, \frac{1}{N}\right) = \frac{2s}{1 - e^{-2Ns}}, \quad [27]$$

where N is the effective population size. Deletions are then successfully fixed at the following rate:

$$\lambda_{\text{del}} = N \cdot U_{\text{del}} \cdot P_{\text{fix}}\left(s, \frac{1}{N}\right) = \frac{2NsU_{\text{del}}}{1 - e^{-2Ns}}, \quad [28]$$

$$d_{\text{del}} \approx \lambda_{\text{del}} \cdot t = \frac{2NsU_{\text{del}}t}{1 - e^{-2Ns}}, \quad [29]$$

We have identified how positive selection can reduce the energetic demands imposed by a nonfunctional region of the genome. The function d_{del}, however, is somewhat unwieldy because it depends on the number of elapsed generations. To simplify interpretation, it is useful to normalize this function by the rate of putatively neutral substitutions. Since deletions have been defined in terms of fitness gains, it is reasonable to treat substitutions at the same sites as neutral. Therefore, we consider only substitutions within the region encoding endospore formation as neutral. The rate of substitution in this region can be defined as:

$$d_{\text{sub}} = U_{\text{sub}} t. \quad [30]$$

By taking the ratio, we can acquire a time-independent quantity:

$$\frac{d_{del}}{d_{sub}} = \frac{2Ns}{1 - e^{-2Ns}} \cdot \frac{U_{del}}{U_{sub}}. \quad [31]$$

This equation is effectively the classic ratio of nonsynonymous to synonymous substitutions, adjusted for differing mutation rates. This adjustment is important because U_{del} is typically lower than U_{sub} (61, 93–95). In *Bacillus*, mutation accumulation experiments estimated the rate of insertions and deletions at 1.20×10^{-10} per event, per individual, per generation, while the rate of substitution was 3.35×10^{-10}, nearly three times higher (61). By counting the fraction of indels that were identified as deletions in each reported mutation accumulation line, we can estimate a deletion rate of 6.0×10^{-11}.

Using these rates, we evaluated the likelihood that a beneficial deletion in spore formation and revival genes becomes fixed relative to a neutral substitution, given the energetic cost of a single nucleotide. Under strict neutrality, this ratio is expected to be less than one due to the lower mutation rate of deletions (*SI Appendix*, Fig. S3). To examine how this ratio scales with deletion size, we applied two published estimates of the effective population size for *Bacillus*: 6.119×10^7 from ref. 61 and 3.224×10^8 from ref. 62.

To better capture how metabolic costs influence the evolution of nonfunctional DNA in *Bacillus*, additional features can be incorporated into the model. For example, deletions do not always remove a single nucleotide. Large segments of the bacterial genome can be lost through individual deletion events (96). We can define the strength of selection on a deletion of size Δ as the product of the deletion size and the relative metabolic burden of a single nucleotide, $s_\Delta = s \cdot \Delta$. However, these larger fitness gains may be rare due to the lower probability of observing large deletions (97).

We can examine the empirical probability distribution of Δ in *Bacillus*. Although it does not capture all possible deletion sizes, it reflects a general pattern supported by data (*SI Appendix*, Fig. S3). Smaller deletions are more frequent than larger ones. We incorporated this distribution into our calculation of the ratio d_{del}/d_{sub}, defining a size-specific deletion function:

$$\frac{d_{del}(\Delta)}{d_{sub}} = \frac{2Ns \cdot \Delta \cdot p(\Delta)}{1 - e^{-2Ns \cdot \Delta}} \cdot \frac{U_{del}}{U_{sub}}, \quad [32]$$

where $p(\Delta)$ is the empirical probability of acquiring a deletion of size Δ.

Data, Materials, and Software Availability. All data and code have been deposited in GitHub (https://github.com/LennonLab/cost_of_spore) (98) and Zenodo (https://zenodo.org/records/18332422) (99).

ACKNOWLEDGMENTS. We acknowledge support from the US NSF (DEB-1934554 and DBI-2022049 to J.T.L.), the US Army Research Office Grant (W911NF2210014 to J.T.L.), the NASA (80NSSC20K0618 to J.T.L.), and the Alexander von Humboldt Foundation (to J.T.L.). We thank D.A. Schwartz, B.K. Lehmkuhl, and J. Ahmed for valuable discussions and technical assistance.

1. S. J. Pirt, The maintenance energy of bacteria in growing cultures. *Proc. R. Soc. London. Ser. B, Biol. Sci.* **163**, 224–231 (1965).
2. C. P. Kempes, L. Wang, J. P. Amend, J. Doyle, T. Hoehler, Evolutionary tradeoffs in cellular composition across diverse bacteria. *ISME J.* **10**, 2145–2157 (2016).
3. N. Lane, W. Martin, The energetics of genome complexity. *Nature* **467**, 929–934 (2010).
4. J. E. Niven, S. B. Laughlin, Energy limitation as a selective pressure on the evolution of sensory systems. *J. Exp. Biol.* **211**, 1792–1804 (2008).
5. R. Unni, K. L. Pintor, A. Diepold, D. Unterweger, Presence and absence of type VI secretion systems in bacteria. *Microbiology* **168**, 001151 (2022).
6. M. Lynch, The bioenergetic cost of building a metazoan. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2414742121 (2024).
7. A. Wagner, Energy constraints on the evolution of gene expression. *Mol. Biol. Evol.* **22**, 1365–1374 (2005).
8. W. R. Shoemaker, J. T. Lennon, Evolution with a seed bank: The population genetic consequences of microbial dormancy. *Evol. Appl.* **11**, 60–75 (2018).
9. M. A. Lever *et al*., Life under extreme energy limitation: A synthesis of laboratory- and field-based investigations. *FEMS Microbiol. Rev.* **39**, 688–728 (2015).
10. K. D. Webster, J. T. Lennon, Dormancy in the origin, evolution and persistence of life on earth. *Proc. R. Soc. B* **292**, 20242035 (2025).
11. J. T. Lennon, F. den Hollander, M. Wilke-Berenguer, J. Blath, Principles of seed banks and the emergence of complexity from dormancy. *Nat. Commun.* **12**, 1–16 (2021).
12. P. Setlow, Spore resistance properties. *Microbiol. Spectr.* **2**, 1–14 (2014).
13. P. Setlow, Spores of *Bacillus subtilis*: Their resistance to and killing by radiation, heat and chemicals. *J. Appl. Microbiol.* **101**, 514–525 (2006).
14. W. L. Nicholson, N. Munakata, G. Horneck, H. J. Melosh, P. Setlow, Resistance of *Bacillus* endospores to extreme terrestrial and extraterrestrial environments. *Microbiol. Mol. Biol. Rev.* **64**, 548–572 (2000).
15. L. Wörmer *et al*., Microbial dormancy in the marine subsurface: Global endospore abundance and response to burial. *Sci. Adv.* **5**, eaav1024 (2019).
16. C. R. Harwood, S. Pohl, W. Smith, A. Wipat, "Chapter 4–*Bacillus subtilis*: Model gram-positive synthetic biology chassis" in *Methods in Microbiology, Microbial Synthetic Biology*, C. Harwood, A. Wipat, Eds. (Academic Press, 2013), vol. 40, pp. 87–117.
17. A. L. Sonenshein, Control of sporulation initiation in *Bacillus subtilis*. *Curr. Opin. Microbiol.* **3**, 561–566 (2000).
18. S. T. Wang *et al*., The forespore line of gene expression in *Bacillus subtilis*. *J. Mol. Biol.* **358**, 16–37 (2006).
19. K. Kikuchi *et al*., Electrochemical potential enables dormant spores to integrate environmental signals. *Science* **378**, 43–49 (2022).
20. Yq. Li *et al*., Thioflavin-t does not report on electrochemical potential and memory of dormant or germinating bacterial spores. *mBio* **14**, e02220-23 (2023).
21. B. Swarge *et al*., Integrative analysis of proteome and transcriptome dynamics during *Bacillus subtilis* spore revival. *mSphere* **5**, e00463-20 (2020).
22. F. U. Battistuzzi, A. Feijao, S. B. Hedges, A genomic timescale of prokaryote evolution: Insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol. Biol.* **4**, 44 (2004).
23. M. Y. Galperin, N. Yutin, Y. I. Wolf, R. Vera Alvarez, Conservation and evolution of the sporulation gene set in diverse members of the firmicutes. *J. Bacteriol.* **204**, e00079-22 (2022).
24. S. A. West, A. S. Griffin, A. Gardner, S. P. Diggle, Social evolution theory for microorganisms. *Nat. Rev. Microbiol.* **4**, 597–607 (2006).
25. C. Weller, M. Wu, A generation-time effect on the rate of molecular evolution in bacteria. *Evolution* **69**, 643–652 (2015).
26. W. R. Shoemaker, E. Polezhaeva, K. B. Givens, J. T. Lennon, Seed banks alter the molecular evolutionary dynamics of *Bacillus subtilis*. *Genetics* **221**, iyac071 (2022).
27. J. Masel, O. D. King, H. Maughan, The loss of adaptive plasticity during long periods of environmental stasis. *Am. Nat.* **169**, 38–46 (2007).
28. C. J. Field, K. L. Bowerman, P. Hugenholtz, Multiple independent losses of sporulation and peptidoglycan in the Mycoplasmatales and related orders of the class Bacilli. *Microb. Genomics* **10**, 001176 (2024).
29. M. Lynch, G. K. Marinov, The bioenergetic costs of a gene. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15690–15695 (2015).
30. J. W. Veening, H. Murray, J. Errington, A mechanism for cell cycle regulation of sporulation initiation in *Bacillus subtilis*. *Genes Dev.* **23**, 1959–1970 (2009).
31. M. Lynch, G. K. Marinov, Membranes, energetics, and evolution across the prokaryote–eukaryote divide. *eLife* **6**, e20437 (2017).
32. S. Ghosh, A. Körte, G. Serafini, V. Yadav, J. Rodenfels, Developmental energetics: Energy expenditure, budgets and metabolism during animal embryogenesis. *Semin. Cell Dev. Biol.* **138**, 83–93 (2023).
33. A. H. Camp, R. Losick, A feeding tube model for activation of a cell-specific transcription factor during sporulation in *Bacillus subtilis*. *Genes Dev.* **23**, 1014–1024 (2009).
34. C. R. Harwood, Y. Kikuchi, The ins and outs of Bacillus proteases: Activities, functions and commercial significance. *FEMS Microbiol. Rev.* **46**, fuab046 (2021).
35. D. W. Hilbert, P. J. Piggot, Compartmentalization of gene expression during *Bacillus subtilis* spore formation. *Microbiol. Mol. Biol. Rev.* **68**, 234–262 (2004).
36. D. Higgins, J. Dworkin, Recent progress in *Bacillus subtilis* sporulation. *FEMS Microbiol. Rev.* **36**, 131–148 (2012).
37. B. Setlow, P. Setlow, Small, acid-soluble proteins bound to DNA protect *Bacillus subtilis* spores from killing by dry heat. *Appl. Environ. Microbiol.* **61**, 2787–2790 (1995).
38. P. Setlow, Spore germination. *Curr. Opin. Microbiol.* **6**, 550–556 (2003).
39. A. O. Henriques, C. P. Moran Jr, Structure, assembly, and function of the spore surface layers. *Annu. Rev. Microbiol.* **61**, 555–588 (2007).
40. P. Setlow, G. Christie, If ATP and macromolecular synthesis are needed for dormant spores of *Bacillota* species to trigger spore germination, where do the energy and precursors come from? *J. Appl. Microbiol.* **136**, lxaf177 (2025).
41. I. Barák, K. Muchová, The positioning of the asymmetric septum during sporulation in *Bacillus subtilis*. *PLoS ONE* **13**, e0201979 (2018).
42. J. Reith, C. Mayer, Peptidoglycan turnover and recycling in Gram-positive bacteria. *Appl. Microbiol. Biotechnol.* **92**, 1–11 (2011).
43. S. Ghosh, G. Korza, M. Maciejewski, P. Setlow, Analysis of metabolism in dormant spores of *Bacillus* species by 31P nuclear magnetic resonance analysis of low-molecular-weight compounds. *J. Bacteriol.* **197**, 992–1001 (2015).
44. P. Setlow, A. Kornberg, Biochemical studies of bacterial sporulation and germination. *J. Biol. Chem.* **245**, 3637–3644 (1970).
45. M. Carrera, R. Zandomeni, J. L. Sagripanti, Wet and dry density of *Bacillus anthracis* and other *Bacillus* species. *J. Appl. Microbiol.* **105**, 68–77 (2008).
46. P. Setlow, Small, acid-soluble spore proteins of *Bacillus* species: Structure, synthesis, genetics, function, and degradation. *Annu. Rev. Microbiol.* **42**, 319–338 (1988).
47. F. M. Meyer, J. Stülke, Malate metabolism in *Bacillus subtilis*: Distinct roles for three classes of malate-oxidizing enzymes. *FEMS Microbiol. Lett.* **339**, 17–22 (2013).
48. B. J. F. Keijser *et al*., Analysis of temporal gene expression during *Bacillus subtilis* spore germination and outgrowth. *J. Bacteriol.* **189**, 3624–3634 (2007).
49. A. Petersohn *et al*., Global analysis of the general stress response of *Bacillus subtilis*. *J. Bacteriol.* **183**, 5617–5631 (2001).

50. F. Rodriguez Ayala, M. Bartolini, R. Grau, The stress-responsive alternative sigma factor SigB of *Bacillus subtilis* and its relatives: An old friend with new functions. *Front. Microbiol.* **11**, 1761 (2020).
51. R. Colin, B. Ni, L. Laganenka, V. Sourjik, Multiple functions of flagellar motility and chemotaxis in bacterial physiology. *FEMS Microbiol. Rev.* **45**, fuab038 (2021).
52. C. Höfler *et al*., Cannibalism stress response in *Bacillus subtilis*. *Microbiology (Reading, England)* **162**, 164–176 (2016).
53. D. A. Gray *et al*., Extreme slow growth as alternative strategy to survive deep starvation in bacteria. *Nat. Commun.* **10**, 890 (2019).
54. D. Dubnau, Genetic competence in *Bacillus subtilis*. *Microbiol. Rev.* **55**, 395–424 (1991).
55. S. Mukherjee, D. B. Kearns, The structure and regulation of flagella in *Bacillus subtilis*. *Annu. Rev. Genet.* **48**, 319–340 (2014).
56. N. I. Wisnoski, M. A. Leibold, J. T. Lennon, Dormancy in metacommunities. *Am. Nat.* **194**, 135–151 (2019).
57. D. A. Schwartz, W. R. Shoemaker, A. Măgălie, J. S. Weitz, J. T. Lennon, Bacteria-phage coevolution with a seed bank. *ISME J.* **17**, 1315–1325 (2023).
58. H. Maughan, J. Masel, C. W. Birky, W. L. Nicholson, The roles of mutation accumulation and selection in loss of sporulation in experimental populations of *Bacillus subtilis*. *Genetics* **177**, 937–948 (2007).
59. H. Maughan, C. W. Birky, W. L. Nicholson, Transcriptome divergence and the loss of plasticity in *Bacillus subtilis* after 6,000 generations of evolution under relaxed selection for sporulation. *J. Bacteriol.* **191**, 428–433 (2009).
60. E. Ilker, M. Hinczewski, Bioenergetic costs and the evolution of noise regulation by microRNAs. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2308796121 (2024).
61. W. Sung *et al*., Evolution of the insertion-deletion mutation rate across the tree of life. *G3* **6**, 2583–2591 (2016).
62. L. M. Bobay, H. Ochman, Factors driving effective population size and pan-genome evolution in bacteria. *BMC Evol. Biol.* **18**, 153 (2018).
63. C. J. Murren *et al*., Constraints on the evolution of phenotypic plasticity: Limits and costs of phenotype and plasticity. *Heredity* **115**, 293–301 (2015).
64. S. J. Giovannoni, J. Cameron Thrash, B. Temperton, Implications of streamlining theory for microbial ecology. *ISME J.* **8**, 1553–1565 (2014).
65. H. P. Browne *et al*., Host adaptation in gut Firmicutes is associated with sporulation loss and altered transmission cycle. *Genome Biol.* **22**, 204 (2021).
66. E. Kussell, R. Kishony, N. Q. Balaban, S. Leibler, Bacterial persistence. *Genetics* **169**, 1807–1814 (2005).
67. D. Schultz, P. G. Wolynes, E. B. Jacob, J. N. Onuchic, Deciding fate in adverse times: Sporulation and competence in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 21027–21034 (2009).
68. J. Narula, S. N. Devi, M. Fujita, O. A. Igoshin, Ultrasensitivity of the *Bacillus subtilis* sporulation decision. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E3513–E3522 (2012).
69. E. Ben-Jacob, M. Lu, D. Schultz, J. N. Onuchic, The physics of bacterial decision making. *Front. Cell. Infect. Microbiol.* **4**, 154 (2014).
70. A. Mutlu, C. Kaspar, N. Becker, I. B. Bischofs, A spore quality-quantity tradeoff favors diverse sporulation strategies in *Bacillus subtilis*. *ISME J.* **14**, 2703–2714 (2020).
71. R. Albalat, C. Cañestro, Evolution by gene loss. *Nat. Rev. Genet.* **17**, 379–391 (2016).
72. D. C. Lahti *et al*., Relaxed selection in the wild. *Trends Ecol. Evol.* **24**, 487–496 (2009).
73. J. Krishnan, N. Rohner, Cavefish and the basis for eye loss. *Philos. Trans. R. Soc. B: Biol. Sci.* **372**, 20150487 (2017).
74. N. A. Wright, D. W. Steadman, C. C. Witt, Predictable evolution toward flightlessness in volant island birds. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4765–4770 (2016).
75. G. Mahmoudabadi, R. Milo, R. Phillips, Energetic cost of building a virus. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E4324–E4333 (2017).
76. R. Milo, R. Phillips, *Cell Biology by the Numbers* (Garland Science, New York, NY, ed. 1, 2015).
77. L. F. Liu, J. C. Wang, Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7024–7027 (1987).
78. S. Maass *et al*., Efficient, global-scale quantification of absolute protein amounts by integration of targeted mass spectrometry and two-dimensional gel-based proteomics. *Anal. Chem.* **83**, 2677–2684 (2011).
79. M. V. Rodnina, Translation in prokaryotes. *Cold Spring Harb. Perspect. Biol.* **10**, a032664 (2018).
80. J. F. Nagle, S. Tristram-Nagle, Lipid bilayer structure. *Curr. Opin. Struct. Biol.* **10**, 474–480 (2000).
81. H. I. Petrache, S. W. Dodd, M. F. Brown, Area per lipid and acyl length distributions in fluid phosphatidylcholines determined by 2H NMR spectroscopy. *Biophys. J.* **79**, 3172–3192 (2000).
82. N. Kučerka, M. P. Nieh, J. Katsaras, Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochim. Biophys. Acta* **1808**, 2761–2771 (2011).
83. B. A. Lewis, D. M. Engelman, Lipid bilayer thickness varies linearly with acyl chain length in fluid phosphatidylcholine vesicles. *J. Mol. Biol.* **166**, 211–217 (1983).
84. K. Mitra, I. Ubarretxena-Belandia, T. Taguchi, G. Warren, D. M. Engelman, Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 4083–4088 (2004).
85. B. Zhu, J. Stülke, SubtiWiki in 2018: From genes and proteins to functional network annotation of the model organism *Bacillus subtilis*. *Nucleic Acids Res.* **46**, D743–D748 (2018).
86. M. Wang *et al*., PaxDb, a database of protein abundance averages across all three domains of life*. *Mol. Cell. Proteomics* **11**, 492–500 (2012).
87. M. Wang, C. J. Herrmann, M. Simonovic, D. Szklarczyk, Cv Mering, Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. *Proteomics* **15**, 3163 (2015).
88. R. T. Eijlander, A. de Jong, A. O. Krawczyk, S. Holsappel, O. P. Kuipers, SporeWeb: An interactive journey through the complete sporulation cycle of *Bacillus subtilis*. *Nucleic Acids Res.* **42**, D685–D691 (2014).
89. P. Nicolas *et al*., Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science* **335**, 1103–1106 (2012).
90. M. Abramowitz, I. A. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables* (Courier Corporation, 1965).
91. D. A. Drummond, J. D. Bloom, C. Adami, C. O. Wilke, F. H. Arnold, Why highly expressed proteins evolve slowly. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 14338–14343 (2005).
92. J. Crow, M. Kimura, *An Introduction to Population Genetics Theory* (Blackburn Press, 2009).
93. H. Long *et al*., Evolutionary determinants of genome-wide nucleotide composition. *Nat. Ecol. Evol.* **2**, 237–240 (2018).
94. M. Lynch *et al*., Genetic drift, selection and the evolution of the mutation rate. *Nat. Rev. Genet.* **17**, 704–714 (2016).
95. H. Long, S. F. Miller, E. Williams, M. Lynch, Specificity of the DNA mismatch repair system (MMR) and mutagenesis bias in bacteria. *Mol. Biol. Evol.* **35**, 2414–2421 (2018).
96. A. I. Nilsson *et al*., Bacterial genome size reduction by experimental evolution. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12112–12116 (2005).
97. E. Wygoda *et al*., Statistical framework to determine indel-length distribution. *Bioinformatics* **40**, btae043 (2024).
98. C. Karakoç, W. R. Shoemaker, J. T. Lennona, cost_of_spore. GitHub. https://github.com/LennonLab/cost_of_spore. Deposited 22 January 2026.
99. C. Karakoç, W. R. Shoemaker, J. T. Lennona, Evolutionary bioenergetics of sporulation. Zenodo. https://zenodo.org/records/18332422. Deposited 22 January 2026.