Review

Ecology and genomics of *Bacillus subtilis*

Ashlee M. Earl[1], Richard Losick[2] and Roberto Kolter[1]

[1] Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA 02115, USA
[2] Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA

Bacillus subtilis **is a remarkably diverse bacterial species that is capable of growth within many environments. Recent microarray-based comparative genomic analyses have revealed that members of this species also exhibit considerable genomic diversity. The identification of strain-specific genes might explain how *B. subtilis* has become so broadly adapted. The goal of identifying ecologically adaptive genes could soon be realized with the imminent release of several new *B. subtilis* genome sequences. As we embark upon this exciting new era of *B. subtilis* comparative genomics we review what is currently known about the ecology and evolution of this species.**

Where do we find *Bacillus subtilis*?

B. subtilis can be isolated from many environments – terrestrial and aquatic – making it seem that this species is ubiquitous and broadly adapted to grow in diverse settings within the biosphere. However, like all members of the genus *Bacillus*, *B. subtilis* can form highly resistant dormant endospores in response to nutrient deprivation and other environmental stresses [1,2]. These spores are easily made airborne and dispersed by wind [3,4]. Thus, spores might migrate long distances, land in a given environment but never germinate there. Considering that the traditional methods for isolating *B. subtilis* require that the organism be in its spore form, there is no guarantee that when a strain is isolated from a particular environment it was actually growing at that location. Thus, the question of where *B. subtilis* grows has not been so simple to answer [5,6].

B. subtilis is often referred to as a 'soil dweller'. Does *B. subtilis* actually grow in soil or is this a place where spores accumulate until they encounter conditions suitable for their germination and proliferation? Over 30 years ago, the use of fluorescent antibodies to distinguish vegetative and spore forms of *B. subtilis* in diverse soil samples [7] revealed that the organism was most often in its vegetative form when associated with decaying organic material [8]. Although this early study is the only one to date that has directly examined growth in natural soils, further support for the idea that *B. subtilis* can lead a saprophytic lifestyle comes from recent experiments in which spores were inoculated into artificial soil microcosms saturated with filter-sterilized soluble organic matter extracted from soil [9]. Under these conditions the spores not only germinated but the vegetative cells proliferated for several days until they again sporulated, probably in response to nutrient depletion. Soon after germination the cells formed bundled chains that moved on the surface in a flagella-independent fashion [9]. Interestingly, a similar transition to growth as bundled chains is observed during the early stages of biofilm development under laboratory conditions [10] (Box 1).

B. subtilis can also grow in close association with plant root surfaces. In the laboratory, when *B. subtilis* was inoculated on the roots of *Arabidopsis thaliana,* growth of biofilms was observed [11,12]. In addition, *B. subtilis* can be isolated in greater numbers than most other spore-forming bacteria from the rhizosphere of a variety of plants [13–15]. There is evidence that through these associations *B. subtilis* can promote plant growth [15]. Possible explanations for this growth promotion are that: (i) *B. subtilis* outcompetes other microbes that would otherwise adversely affect the plant; (ii) *B. subtilis* activates the host defense system so that the plant is poised to resist potential pathogens; or (iii) *B. subtilis* makes certain nutrients more readily available to the plant (e.g. phosphorus and nitrogen) [16].

Considering that *B. subtilis* is found on and around plants and that many animals consume plants, it is not surprising that this bacterium is often found in feces [6,17–19]. Passage of *B. subtilis* through animal gastrointestinal (GI) tracts might not be without effects; the idea that *B. subtilis* has an active role within the GI tract has had anecdotal support for years. In fact, *B. subtilis* has been touted as a probiotic that when ingested has 'beneficial' effects, probably by helping to maintain or restore 'healthy' bacterial communities in the body [20]. *B. subtilis* is also found in several commercially available fermented food products, including soybeans fermented with *B. subtilis natto*, which is popular in Japan and has long been thought to confer health benefits [21]. But like its role in plant growth promotion, it is not fully understood how *B. subtilis* imparts its probiotic effect.

Work in recent years has transformed our view of what *B. subtilis* can do within the GI tract of animals. Formerly, *B. subtilis* was thought to be an obligate aerobe that simply traveled through the mostly anaerobic GI tract as a spore. Therefore, any benefit derived from its consumption was thought to be due to some intrinsic property of the spore. Recent evidence, however, indicates that *B. subtilis* can complete its entire life cycle within the GI tract, going from

Corresponding author: Kolter, R. (rkolter@hms.harvard.edu).

0966-842X/$ – see front matter doi:10.1016/j.tim.2008.03.004

Box 1. Biofilm formation by *Bacillus subtilis*

It is now broadly recognized that bacteria have the widespread capacity to form surface-associated multicellular aggregates, commonly referred to as biofilms [73,74]. Within biofilms the constituent cells are embedded in matrices composed of diverse extracellular polymeric substances [73]. Although it is not entirely known how microbes benefit from life within these structures, biofilm-associated microbes exhibit marked metabolic and physiological differences compared with their planktonic brethren, including decreased susceptibility to antimicrobial agents [75].

Bacillus subtilis has been a model organism for the study of biofilm formation [10,51,76–78]. Through a combination of genetic and biochemical approaches, both the structural and regulatory components necessary for the development of these multicellular communities have been identified in strain NCIB3610 [10,77–80] (Figure I, left), a possible progenitor of the sequenced *B. subtilis* 168 [43] (Figure I, right). In short, the extracellular matrix that holds the biofilm together is composed of both a protein and polysaccharide component encoded by the *yqxM* and *eps* operons, respectively [10,77,80]. Although not essential in *B. subtilis* NCIB3610, γ-polyglutamate has also been shown to participate as a component of the extracellular matrix in biofilms of some strains of *B. subtilis* [51,52]. Expression of the *yqxM* and *eps* operons is controlled by a complex regulatory circuit involving a core regulatory duo, SinI and SinR, whose activities are subject to upstream regulators including Spo0A. The ability to make complex biofilms like that of *B. subtilis* NCIB3610 seems to be conserved among wild strains of this species. All of the biofilm-related genes identified in NCIB3610 also seem to be conserved among strains examined by microarray-based comparative genomic hybridization [43].

Figure I. *B. subtilis* biofilm morphologies. *B. subtilis* NCIB3610 (left) and *B. subtilis* 168 (right) spotted onto MSgg agar, a biofilm-inducing medium commonly used in the laboratory to study biofilm development [10]. Note the difference in contoured texture of the surface of each strain; *B. subtilis* NCIB3610 exhibits much more complex biofilm architecture than does the sequenced and domesticated strain, *B. subtilis* 168. Scale bar = 1 cm.

spore to vegetative cell and sporulating again [18–20,22]. In fact, growth of *B. subtilis* in the GI tract is robust enough for this organism to outcompete pathogens like *Escherichia coli* in poultry GI tracts when it is administered orally [22].

B. subtilis has also been isolated repeatedly from aquatic environments [23–25]. However, there is no published account that directly demonstrates the growth of *B. subtilis* in natural waters. Although growth in marine water might occur, the abundance of *B. subtilis* in these environments might also be explained by its observed association with the GI tract of marine organisms [26] and other biotic surfaces [25].

In summary, current data indicate that the apparent ubiquity of *B. subtilis* is not solely a consequence of spore persistence in these environments. Instead, *B. subtilis* seems to grow in diverse environments including soils, on plant roots, and within the GI tract of animals.

What can genomics teach us about *B. subtilis* ecology?

Today we find ourselves in a golden age of genomics thanks to increasingly facile methods for generating, assembling and analyzing large amounts of sequence information [27]. We no longer need to rely solely on isolation geography, behaviors in the laboratory, or anecdotal reports to characterize the ecology of a species. In addition, we can investigate the genes present or absent in any strain of interest. The identity of the proteins predicted to be encoded by the genome of an organism can reveal much about the lifestyle of that organism and the habitats where it resides.

The genome sequence of *B. subtilis* 168 has provided many insights into the lifestyles of the organism [28]. Consistent with the view that the bacterium is not a pathogen, no genes encoding known virulence factors were found. Interestingly, the genome encoded components of numerous pathways for the use of plant-derived molecules, bolstering the idea that this species associates intimately with plants [28]. One observation challenged the long-held belief that *B. subtilis* was an obligate aerobe: genes encoding a putative respiratory nitrate reductase were found [28]. This indicated that *B. subtilis* should be able to grow anaerobically using nitrate instead of oxygen as an electron acceptor. Anaerobic growth of *B. subtilis* in the presence of nitrate has since been demonstrated experimentally [29]. The discovery that *B. subtilis* can grow anaerobically further supports the idea that vegetative life within the mostly anaerobic GI tract of animals is feasible.

The genome sequence also revealed that *B. subtilis* has dedicated a relatively large portion of its genome (~4%) to making secondary metabolites. Some of these compounds are potent inhibitors of fungi and bacteria and probably enable *B. subtilis* to compete in the natural environment [16,17,30], promote plant growth and serve as a probiotic.

The limitations of a genome sequence from a single laboratory strain

The genome of *B. subtilis* 168 was chosen for sequencing because the laboratory strain had been the workhorse for molecular genetic studies for several decades. The popularity of *B. subtilis* 168 stemmed from the ease with which this strain could be genetically manipulated; its increased genetic competence was probably brought about by X- and UV-irradiation of its parent. The mutagenic consequences of irradiation coupled with repeated growth under artificial settings caused *B. subtilis* 168 to 'evolve' in ways that improved fitness in the laboratory, a process commonly referred to as domestication [10]. But this domestication came at a cost. We now recognize that *B. subtilis* 168 is deficient in several traits that are characteristic of wild strains. Among these are surface swarming and the ability to form architecturally complex biofilms [10,31] (see Figure I in Box 1). Conversely, *B. subtilis* 168 produces a much greater proportion of cells in the state of genetic competence than do wild strains.

At the same time that investigators began to recognize strain domestication as a common laboratory phenomenon the genomic era delivered a surprise. In some cases the genomes of different strains of a single species were greatly

conserved whereas in others the genetic variability was enormous. In the case of *E. coli*, although different strains possess identical 16S rRNA gene sequences, strains can harbor more than 1000 strain-specific genes [32]. There seems to be a trend that the amount of differences in gene content observed within a given species correlates with certain features of the ecology of that species [33]. Bacterial species with little genome variability seem to occupy few habitats whereas those with more genomic diversity within strains seem to colonize diverse environments [33].

Where does *B. subtilis* lie in the spectrum of genomic diversity? Does the genome of *B. subtilis* 168 tell the full tale of the biology and ecology of this species? Is there genomic variation among members of this species? And, if so, could this variation explain differences in strain ecology?

Foreshadowing *B. subtilis* genomic diversity

For many years most of the available evidence concerning genotypic variation among different *B. subtilis* isolates came from the assessment of phenotypic variation and cell wall chemistry [34,35]. Only in the 1990s were sequences from loci other than 16S rRNA genes examined among multiple strains [36–38]. These studies revealed that *B. subtilis* was not nearly as genetically monomorphic as its pathogenic relative, *B. anthracis* [39]. One such survey used restriction fragment length polymorphisms (RFLPs) of three housekeeping genes as markers for genetic diversity among strains isolated from geographically distant locations [36]. The results revealed that these strains were clearly phylogenetically separate from other recognized species of the genus *Bacillus*, yet fell into two distinct phylogenetic groups [36]. This robust phylogenetic separation called into question the assignment of *B. subtilis* as a single species. In other words, did the strains from both phylogenetic clusters belong to the species *B. subtilis* or was there enough variation to reclassify one of these groups as a distinct species within the genus *Bacillus*? Using 'classical' methods for bacterial species assignments [40], including DNA reassociation analysis, it was concluded that the two groups exhibited sufficient 'relatedness' to be kept within the same species, but were different enough to warrant classification as subspecies [41]. Thus, strains of *B. subtilis* were divided into subspecies *B. subtilis* subsp. *subtilis*, containing the sequenced strain *B. subtilis* 168, and *B. subtilis* subsp. *spizizenii* [41].

Analyses involving DNA reassociation kinetics also indicated that the genetic diversity among members of this species exceeded that found by nucleotide variation at conserved sites [41]. The results indicated that a large percentage of the genomic DNA of each strain was strain specific. However, the identities of these strain-specific regions were entirely unknown. Could the identities of genes within these variable regions inform our view of the ecology of *B. subtilis*?

Microarray-based comparative genomic hybridization analyses

Ideally, to begin to answer the foregoing question one would seek to identify and compare all of the genes harbored by each strain. But although whole-genome sequencing has become an increasingly feasible option for such an analysis, it is still not a quick or inexpensive undertaking. However, the available *B. subtilis* 168 genome sequence did provide an opportunity to explore genome variation among strains at much lower cost. Using an oligonucleotide microarray designed to represent each of the predicted coding sequences of *B. subtilis* 168, it was possible to query closely related strains for variation in each *B. subtilis* gene. This technique, called microarray-based comparative genomic hybridization (M-CGH), is simply a DNA reassociation method that provides more detailed information about which genes are contributing to lowered reassociation values [42]. In this case, DNA from strains that either lack or possess a divergent copy of a *B. subtilis* 168 gene will not hybridize as well as the DNA from *B. subtilis* 168 to that gene-specific oligonucleotide. The relative hybridization of the DNA of a strain can be easily assessed by measuring variation in fluorescence intensity at each gene spot when the genomes of *B. subtilis* 168 and of the test strains are differentially labeled with fluorescent nucleotides [43].

Such analyses were performed using a collection of diverse strains from both subspecies groups [43]. The results from this study revealed that 30% of the predicted coding sequences of *B. subtilis* 168 were cumulatively absent or divergent in the strains tested [43]. Not surprisingly, strains that were more closely related to *B. subtilis* 168 (within the *subtilis* subspecies) exhibited less total gene diversity relative to those in the other subspecies, consistent with the RFLP and DNA reassociation data.

Where is genome diversity localized? To answer this question, knowledge of the extent of synteny, or conservation in gene order, among strains is needed. Underscoring limitations of M-CGH analyses [42,44], it is not possible to predict whether the 'conserved' set of genes are also similarly ordered relative to *B. subtilis* 168 among the strains examined. Although there was only one *B. subtilis* genome sequence available, a considerable degree of synteny among *B. subtilis* strains is to be expected given the observed synteny between the *B. subtilis* 168 sequence and the recently published *B. licheniformis* ATCC 14580 genome sequence [45,46]. Assuming that synteny among *B. subtilis* strains is great, it seems that genomic diversity among this species is not localized to only a few areas within the genome. Rather, it is distributed along the entire genome. In summary, based on the M-CGH analyses there are few large stretches of genomic DNA that do not have some possibility of variation.

M-CGH analysis reveals regions of variability among wild strains of *B. subtilis*

Within these distributed regions of diversity were some genes that, given previous phenotypic and biochemical observations, came as no surprise. These included genes that encode components involved in the synthesis of secondary metabolites [30,47], teichoic acid [48], and the adaptive response to alkylation DNA damage [49]. The M-CGH analysis revealed that there was also variability in nearly all 'functional' categories of genes, some of which could prove ecologically relevant by changing (expanding or limiting) the environments in which these strains can

live. Divergence was observed in genes that encode proteins involved in the uptake and breakdown of carbohydrates (e.g. xylose) and amino acids (e.g. glutamine) in addition to several cell surface-associated proteins, including those involved in environmental sensing [43]. The observed variability among these loci, and others like them, indicates that certain metabolic and environmental-monitoring capabilities might not be required for the life of *B. subtilis* in all environments.

It is equally informative to determine which genes exhibit limited or no variability. Presumably these greatly conserved loci would encode proteins that are selected for in all environments inhabited by the species. As expected, nearly all of the genes that had previously been shown to be essential under laboratory conditions in *B. subtilis* 168 [50] were invariable among the *B. subtilis* strains examined [43]. Also, a large fraction of the sporulation genes were conserved. This is unsurprising given that all of the strains from the M-CGH study were originally isolated as spores. It is interesting to note, however, that many of the germination genes of *B. subtilis* 168 exhibited divergence. Hence, the cues for reinitiating growth might not be the same in all environments.

Genes involved in biofilm formation were greatly conserved [43]. Life within matrix-associated multicellular communities seems to be a universally important ecological trait for this species. However, there are reports of strain-to-strain variation among conserved loci that can affect the outcome of this developmental process [51,52]. This strain-to-strain variation was not detected in the M-CGH analyses because it involved minor sequence changes in conserved genes and regulatory regions. The noted allelic variation might thus have been the consequence of laboratory domestication alone and not necessarily reflective of variation among wild isolates.

As mentioned, *B. subtilis* is also noted for its ability to become naturally competent for transformation, that is, the ability to take up and recombine extracellular DNA into its genome [53]. The M-CGH analyses revealed that the competence machinery identified in *B. subtilis* 168 was greatly conserved except for one operon. The three-gene *comPQX* operon is involved in the synthesis, processing and recognition of an extracellular signal that is required for the initiation of competence [54]. *comPQX* had been previously recognized as variable among strains of *B. subtilis* [55,56]. It was further demonstrated that the observed genetic variation also resulted in functional variation such that different strains produced and recognized different variants of the extracellular signal to the exclusion of others [57]. Considering that competence signal recognition is a population-density-dependent phenomenon, *B. subtilis* strains probably become competent only when their own numbers are elevated. This would indicate that genetic transfer by transformation would occur most often with DNA from 'self'.

What are the drivers of diversity and evolution in this species?

How does genomic diversity arise? Mutagens in addition to DNA replication and repair errors can introduce mutations into a genome. If a mutation is neutral or confers an advantage for life in a given environment, that mutation can become fixed within a population and eventually predominate. Although this mechanism for genetic change unquestionably occurs in nature, it is not the primary driver of evolution among bacterial species [58]. Instead, horizontal gene transfer (HGT), through transduction, conjugation or transformation, is thought to have the most important role in this process [58]. Consistent with this notion, the *B. subtilis* 168 genome sequence revealed that a large portion of the genome of this strain might have arisen by HGT [59,60]. And, perhaps not surprisingly, many of the divergent genes (~40%) among the strains examined by M-CGH were located in these regions [43].

Among the genes predicted to have been horizontally transferred, many are clustered, exhibit atypical genetic signatures and encode for phage-related proteins, indicating that these regions were gained by transduction. A recent study reported that phage integration could account for as much as 16% of the predicted HGT regions in the genome of *B. subtilis* 168 [60]. This indicates that, as in many other bacterial species examined to date [58,61], phages have a role in the evolution of this species. Whether phages are actually shaping the ecology of *B. subtilis* by introducing novel loci that could be used in the exploration of or residence in different environments is yet to be conclusively determined. However, the presence of genes that encode components for antibiotic synthesis and detoxification within these phage elements strongly indicates that they could serve such a purpose.

Plasmids mediate gene transfer through conjugation and thus have a role in bacterial evolution [58]. Plasmid transfer by conjugation has been reported to occur between strains of *B. subtilis* in soil microcosms, indicating that the movement of plasmids commonly occurs among members of this species [62]. However, a survey of plasmid diversity among 50 natural isolates of *B. subtilis* estimated that only ~10% of strains harbor these extrachromosomal elements [38]. All of the plasmids identified seemed to be greatly homologous, probably all sharing the same basic replicon [38]. This is different from the account of plasmid diversity in other bacterial species, such as *E. coli* [63,64]. There is no evidence to indicate that *B. subtilis* plasmids confer any benefit, perhaps explaining their low occurrence among natural *B. subtilis* populations and their genetic homogeneity [38].

Transfer of a chromosomally encoded conjugal element has also been reported [65]. Integrative and conjugative element *Bs1* (ICE*Bs1*), encoded on the *B. subtilis* 168 genome, is a conjugative transposon that has been shown to transfer to other species of *Bacillus* and to species of *Listeria* by conjugation under laboratory conditions [65]. Consistent with other known transposable elements, ICE*Bs1* is activated by DNA damage, but its activation is also influenced by external concentrations of an ICE*Bs1*-encoded extracellular peptide, PhrI. When the cell senses that PhrI levels are low (i.e. the number of non-ICE*Bs1*-containing cells is raised), ICE*Bs1* is activated and transferred to naive cells by conjugation [65]. Although ICE*Bs1* is the only known transposable element encoded on the *B. subtilis* 168 genome, additional transposable elements (not necessarily conjugative) have been reported in other

strains of the species [66]. Ultimately, conjugation, whether mediating the movement of plasmids or chromosomally encoded genetic elements, could represent an important driver in the evolution of this species.

Finally, it seems that transformation could help drive the evolution of *B. subtilis*. Under laboratory conditions, strains can take up and recombine exogenously added genomic DNA from relatives [67]. This can occur even between subspecies although the number of recombinants declines as relatedness decreases, a phenomenon termed sexual isolation [68]. Also, early experiments using sterilized soil microcosms monitored what happened when differentially 'marked' variants of strains were mixed [69]. Such exchange was observed even between different species, for example, *B. subtilis* and *B. licheniformis* [70]. However, the results observed were probably biased by the choice of strains, because both laboratory strains used are known to be much more readily transformable than wild strains. The 'hybrid species' recombinants were also unstable, indicating that the results might not be relevant to the situation in nature. It does seem, however, that wild populations of *B. subtilis* recombine their genes in nature [36]. How this exchange is mediated – by transformation, transduction or conjugation – is yet to be determined.

Concluding remarks and future perspectives

In summary, *B. subtilis* is a widely adapted bacterial species, capable of growing within many environments including soil, plant roots and the GI tracts of animals. The *B. subtilis* 168 genome sequence has been an important tool in aiding our understanding of how growth within some of these environments is possible. It is now clear, however, that the *B. subtilis* 168 genome does not tell the entire story. Recent M-CGH analyses have revealed great variability among the genes of different members of the species [43].

Although intriguing, the results of M-CGH give us an incomplete picture. For instance, M-CGH cannot forecast what, if any, genes are present within regions of divergence that are not already found within the *B. subtilis* 168 genome [44]. However, we are poised to answer this question. Whole genome sequences from selected representatives of both *B. subtilis* subspecies will soon be available (Figure 1), revealing the identities of genes within these regions of divergence. It will be interesting to see whether some of these genes prove to be ecologically significant and whether they broaden our view of the habitats of *B. subtilis* and the adaptations it has acquired to propagate in diverse environments. From an evolutionary perspective, newly sequenced *B. subtilis* genomes, especially in combination with the genomes of the closest relatives of *B. subtilis* (Figure 1), will undoubtedly provide a wealth of information regarding the mechanisms by which this group has evolved.

This is an exciting time for the study of bacterial ecology and evolution – due largely to recent rapid advances in genome sequencing platforms, such as 454 from Roche (www.454.com), Solexa from Illumina (www.illumina.com) and SOLiD from ABI (www3.appliedbiosystems.com). It is now possible to sequence the genome of an organism much faster and more cheaply compared with more traditional methods. However, it is not yet possible to solve a genome sequence *de novo* because of limitations in sequence read length. Although current and forthcoming sequencing technologies, including SMRT Technology from Pacific Biosciences (www.pacificbiosciences.com), are attempting

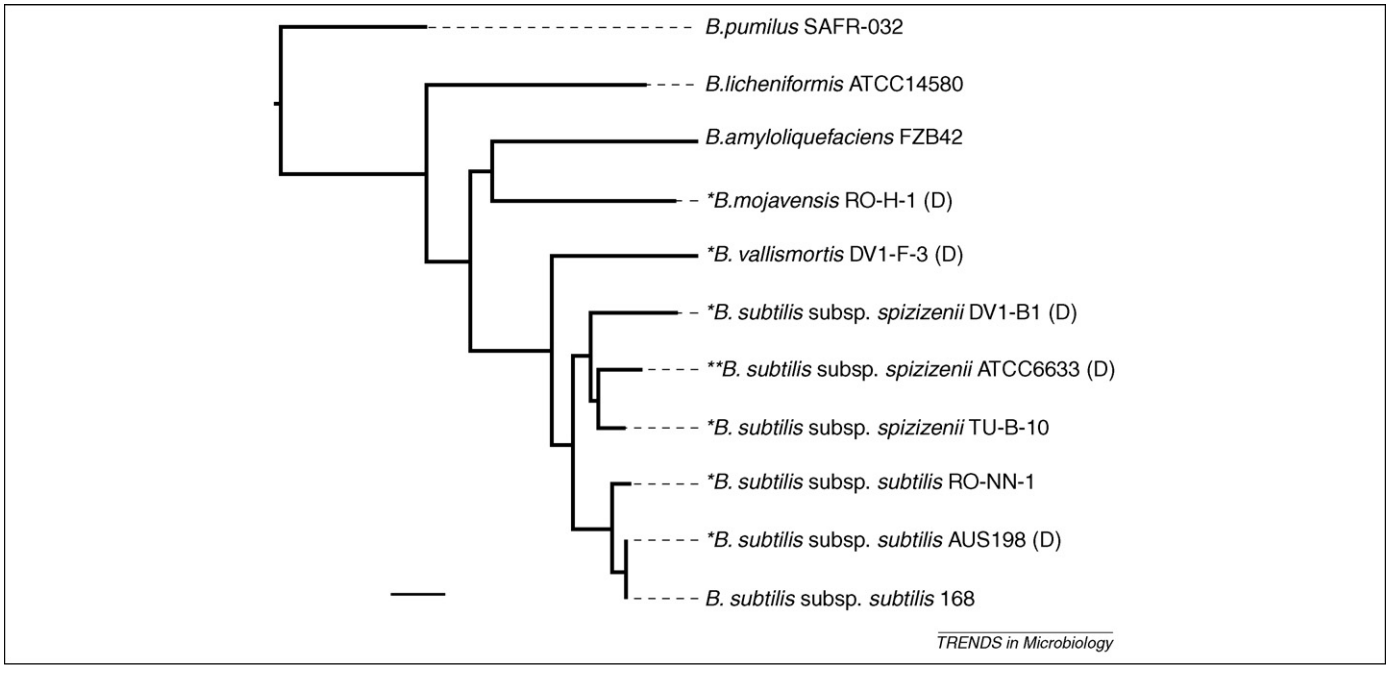

Figure 1. Neighbor-joining tree representing the relationship among sequenced strains of *B. subtilis* and close relatives. This tree was generated from the alignment of partial *gyrA* sequences from each strain. Strain names that are not preceded by an asterisk are those whose genome sequences have already been published [28,45,46,71,72]. Asterisk(s) denotes strains whose genomes are being sequenced, but are not yet published. The symbol (D) denotes strains for which only draft or incomplete genome sequences will be available upon completion of the project. Scale bar = 0.02 nucleotide substitutions per site. Genome sequencing performed at the Institute for Genome Sciences, University of Maryland School of Medicine (*) and University of Illinois, Urbana-Champaign (**). Sequence information can be found at www.bacillusgenomics.org/bsubtilis.

to improve read length, this limitation can presently be overcome when there is a closely related genome sequence available. The short sequence reads from a new strain can be readily assembled using the genome from a close relative as a scaffold. Given the distribution of genome sequences that will soon be available among *B. subtilis* and its close relatives (Figure 1) the *Bacillus* community is poised to take advantage of these faster and cheaper methods for whole genome sequencing.

Ultimately, the *B. subtilis* field is blessed with a treasure trove of information on the genetics, molecular biology and physiology of the laboratory strain 168 from over a century of research. This historical underpinning together with the emerging availability of whole genome sequences for wild isolates promises to usher in a new era of exciting discoveries about the ecology and evolution of this remarkable bacterium.

Acknowledgements

We would like to thank the anonymous reviewers for their insight and helpful comments. We also thank Hera Vlamakis for her comments and critical reading of this review. Work in our laboratories on *B. subtilis* biofilms and genome diversity is funded by grants from the National Institutes of Health GM18568 to R.L. and GM58213 to R.K. A.M.E. was the recipient of a postdoctoral fellowship from the National Institutes of Health (GM072393).

References

1 Sonenshein, A.L. *et al.*, eds (2002) *Bacillus subtilis and Its Closest Relatives: From Genes to Cells*, ASM Press
2 Ricca, E. *et al.*, eds (2004) *Bacterial Spore Formers: Probiotics and Emerging Applications*, Horizon Scientific Press
3 Merrill, L. *et al.* (2006) Composition of Bacillus species in aerosols from 11 U.S. cities. *J. Forensic Sci.* 51, 559–565
4 Jaenicke, R. (2005) Abundance of cellular material and proteins in the atmosphere. *Science* 308, 73
5 Felske, A.D.M. (2004) Ecology of Bacillus species in soil. In *Bacterial Spore Formers: Probiotics and Emerging Applications* (Ricca, E. *et al.*, eds), pp. 35–44, Horizon Scientific Press
6 Nicholson, W.L. (2004) Ubiquity, longevity, and ecological roles of Bacillus spores. In *Bacterial Spore Formers: Probiotics and Emerging Applications* (Ricca, E. *et al.*, eds), pp. 1–15, Horizon Scientific Press
7 Norris, J.R. and Wolf, J. (1961) A study of antigens of the aerobic spore-forming bacteria. *J. Appl. Bacteriol.* 24, 42–56
8 Siala, A. *et al.* (1974) Populations of spore-forming bacteria in an acid forest soil, with special reference to B. subtilis. *J. Gen. Microbiol.* 81, 183–190
9 Vilain, S. *et al.* (2006) Analysis of the life cycle of the soil saprophyte *Bacillus cereus* in liquid soil extract and in soil. *Appl. Environ. Microbiol.* 72, 4970–4977
10 Branda, S.S. *et al.* (2001) Fruiting body formation by *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* 98, 11621–11626
11 Bais, H.P. *et al.* (2004) Biocontrol of *Bacillus subtilis* against infection of *Arabidopsis roots* by *Pseudomonas syringae* is facilitated by biofilm formation and surfactin production. *Plant Physiol.* 134, 307–319
12 Rudrappa, T. *et al.* (2007) A degradation product of the salicylic acid pathway triggers oxidative stress resulting in down-regulation of *Bacillus subtilis* biofilm formation on *Arabidopsis thaliana* roots. *Planta* 226, 283–297
13 Fall, R. *et al.* (2004) A simple method to isolate biofilm-forming *Bacillus subtilis* and related species from plant roots. *Syst. Appl. Microbiol.* 27, 372–379
14 Pandey, A. and Palni, L.M. (1997) *Bacillus* species: the dominant bacteria of the rhizosphere of established tea bushes. *Microbiol. Res.* 152, 359–365
15 Cazorla, F.M. *et al.* (2007) Isolation and characterization of antagonistic *Bacillus subtilis* strains from the avocado rhizoplane displaying biocontrol activity. *J. Appl. Microbiol.* 103, 1950–1959
16 Nagorska, K. *et al.* (2007) Multicellular behaviour and production of a wide variety of toxic substances support usage of *Bacillus subtilis* as a powerful biocontrol agent. *Acta Biochim. Pol.* 54, 495–508
17 Barbosa, T.M. *et al.* (2005) Screening for bacillus isolates in the broiler gastrointestinal tract. *Appl. Environ. Microbiol.* 71, 968–978
18 Tam, N.K. *et al.* (2006) The intestinal life cycle of *Bacillus subtilis* and close relatives. *J. Bacteriol.* 188, 2692–2700
19 Leser, T.D. *et al.* (2008) Germination and outgrowth of *Bacillus subtilis* and *Bacillus licheniformis* spores in the gastrointestinal tract of pigs. *J. Appl. Microbiol.* 104, 1025–1033
20 Hong, H.A. *et al.* (2005) The use of bacterial spore formers as probiotics. *FEMS Microbiol. Rev.* 29, 813–835
21 Inatsu, Y. *et al.* (2006) Characterization of *Bacillus subtilis* strains in Thua nao, a traditional fermented soybean food in northern Thailand. *Lett. Appl. Microbiol.* 43, 237–242
22 Hong, H.A. and Duc le, H. (2004) The fate of ingested spores. In *Bacterial Spore Formers: Probiotics and Emerging Applications* (Ricca, E. *et al.*, eds), pp. 107–112, Horizon Scientific Press
23 Miranda, C.A. *et al.* (2008) Species-level identification of *Bacillus* strains isolates from marine sediments by conventional biochemical, 16S rRNA gene sequencing and inter-tRNA gene sequence lengths analysis. *Antonie Van Leeuwenhoek* 93, 297–304
24 El-Helow, E.R. (2001) Identification and molecular characterization of a novel *Bacillus* strain capable of degrading Tween-80. *FEMS Microbiol. Lett.* 196, 119–122
25 Ivanova, E.P. *et al.* (1999) Characterization of *Bacillus* strains of marine origin. *Int. Microbiol.* 2, 267–271
26 Newaj-Fyzul, A. *et al.* (2007) *Bacillus subtilis* AB1 controls *Aeromonas* infection in rainbow trout (*Oncorhynchus mykiss*, Walbaum). *J. Appl. Microbiol.* 103, 1699–1706
27 Hall, N. (2007) Advanced sequencing technologies and their wider impact in microbiology. *J. Exp. Biol.* 210, 1518–1525
28 Kunst, F. *et al.* (1997) The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* 390, 249–256
29 Nakano, M.M. *et al.* (1997) Characterization of anaerobic fermentative growth of *Bacillus subtilis*: identification of fermentation end products and genes required for growth. *J. Bacteriol.* 179, 6749–6755
30 Stein, T. (2005) *Bacillus subtilis* antibiotics: structures, syntheses and specific functions. *Mol. Microbiol.* 56, 845–857
31 Kearns, D.B. and Losick, R. (2003) Swarming motility in undomesticated *Bacillus subtilis*. *Mol. Microbiol.* 49, 581–590
32 Welch, R.A. *et al.* (2002) Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 99, 17020–17024
33 Konstantinidis, K.T. *et al.* (2006) The bacterial species definition in the genomic era. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361, 1929–1940
34 Gordon, R.E. *et al.* (1973) *The Genus Bacillus*, Agricultural Research Service, U.S. Dept. of Agriculture
35 Sneath, P.H.A. and Bergey, D.H. (1986) *Bergey's Manual of Systematic Bacteriologyf* Lippincott Williams & Wilkins
36 Roberts, M.S. and Cohan, F.M. (1995) Recombination and migration rates in natural populations of *Bacillus subtilis* and *Bacillus mojavensis*. *Evolution Int. J. Org. Evolution* 49, 1081–1094
37 Istock, C.A. *et al.* (1992) Sexuality in a natural population of bacteria – *Bacillus subtilis* challenges the clonal paradigm. *Mol. Ecol.* 1, 95–103
38 Zawadzki, P. *et al.* (1996) Homology among nearly all plasmids infecting three *Bacillus species*. *J. Bacteriol.* 178, 191–198
39 Van Ert, M.N. *et al.* (2007) Global genetic population structure of *Bacillus anthracis*. *PLoS ONE* 2, e461
40 Wayne, L.G. *et al.* (1987) Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int. J. Syst. Bacteriol.* 37, 463–464
41 Nakamura, L.K. *et al.* (1999) Relationship of *Bacillus subtilis* clades associated with strains 168 and W23: a proposal for *Bacillus subtilis* subsp. *subtilis* subsp. nov. and *Bacillus subtilis* subsp. *spizizenii* subsp. nov. *Int. J. Syst. Bacteriol.* 49, 1211–1215
42 Taboada, E.N. *et al.* (2007) Studying bacterial genome dynamics using microarray-based comparative genomic hybridization. In *Comparative Genomics* (Vol. 2) (Bergman, N.H., ed.), pp. 223–253, Humana Press
43 Earl, A.M. *et al.* (2007) *Bacillus subtilis* genome diversity. *J. Bacteriol.* 189, 1163–1170

44 Taboada, E.N. *et al.* (2005) A new approach for the analysis of bacterial microarray-based comparative genomic hybridization: insights from an empirical study. *BMC Genomics* 6, 78
45 Rey, M.W. *et al.* (2004) Complete genome sequence of the industrial bacterium *Bacillus licheniformis* and comparisons with closely related *Bacillus species*. *Genome Biol.* 5, R77
46 Veith, B. *et al.* (2004) The complete genome sequence of *Bacillus licheniformis* DSM13, an organism with great industrial potential. *J. Mol. Microbiol. Biotechnol.* 7, 204–211
47 Duitman, E.H. *et al.* (1999) The mycosubtilin synthetase of *Bacillus subtilis* ATCC6633: a multifunctional hybrid between a peptide synthetase, an amino transferase, and a fatty acid synthase. *Proc. Natl. Acad. Sci. U. S. A.* 96, 13294–13299
48 Lazarevic, V. *et al.* (2002) Comparison of ribitol and glycerol teichoic acid genes in *Bacillus subtilis* W23 and 168: identical function, similar divergent organization, but different regulation. *Microbiology* 148, 815–824
49 Morhoshi, F. and Munakata, N. (1995) Diverse capacities for the adaptive response to DNA alkylation in *Bacillus* species and strains. *Mutat. Res.* 337, 97–110
50 Kobayashi, K. *et al.* (2003) Essential *Bacillus subtilis* genes. *Proc. Natl. Acad. Sci. U. S. A.* 100, 4678–4683
51 Stanley, N.R. and Lazazzera, B.A. (2005) Defining the genetic differences between wild and domestic strains of *Bacillus subtilis* that affect poly-gamma-dl-glutamic acid production and biofilm formation. *Mol. Microbiol.* 57, 1143–1158
52 Morikawa, M. *et al.* (2006) Biofilm formation by a *Bacillus subtilis* strain that produces γ-polyglutamate. *Microbiology* 152, 2801–2807
53 Chen, I. and Dubnau, D. (2004) DNA uptake during bacterial transformation. *Nat. Rev. Microbiol.* 2, 241–249
54 Magnuson, R. *et al.* (1994) Biochemical and genetic characterization of a competence pheromone from *B. subtilis*. *Cell* 77, 207–216
55 Tortosa, P. *et al.* (2001) Specificity and genetic polymorphism of the *Bacillus* competence quorum-sensing system. *J. Bacteriol.* 183, 451–460
56 Tran, L.S. *et al.* (2000) Divergent structure of the ComQXPA quorum-sensing components: molecular basis of strain-specific communication mechanism in *Bacillus subtilis*. *Mol. Microbiol.* 37, 1159–1171
57 Ansaldi, M. *et al.* (2002) Specific activation of the Bacillus quorum-sensing systems by isoprenylated pheromone variants. *Mol. Microbiol.* 44, 1561–1573
58 Binnewies, T.T. *et al.* (2006) Ten years of bacterial genome sequencing: comparative-genomics-based discoveries. *Funct. Integr. Genomics* 6, 165–185
59 Garcia-Vallve, S. *et al.* (2003) HGT-DB: a database of putative horizontally transferred genes in prokaryotic complete genomes. *Nucleic Acids Res.* 31, 187–189
60 Dufraigne, C. *et al.* (2005) Detection and characterization of horizontal transfers in prokaryotes using genomic signature. *Nucleic Acids Res.* 33, e6
61 Weinbauer, M.G. and Rassoulzadegan, F. (2004) Are viruses driving microbial diversification and diversity? *Environ. Microbiol.* 6, 1–11
62 Lotareva, O.V. and Prozorov, A.A. (2005) The conjugated transfer of plasmid pUB110 in *Bacillus subtilis* in soils from different natural landscapes. *Mikrobiologiia* 74, 132–134
63 Boyd, E.F. *et al.* (1996) Mosaic structure of plasmids from natural populations of *Escherichia coli*. *Genetics* 143, 1091–1100
64 Thomas, C.M. (ed.) (2000) *The Horizontal Gene Pool. Bacterial Plasmids and Gene Spread*, Harwood Academic Publishers
65 Auchtung, J.M. *et al.* (2005) Regulation of a *Bacillus subtilis* mobile genetic element by intercellular signaling and the global DNA damage response. *Proc. Natl. Acad. Sci. U. S. A.* 102, 12554–12559
66 Nagai, T. *et al.* (2000) A new IS4 family insertion sequence, IS4Bsu1, responsible for genetic instability of poly-gamma-glutamic acid production in *Bacillus subtilis*. *J. Bacteriol.* 182, 2387–2392
67 Tortosa, P. and Dubnau, D. (1999) Competence for transformation: a matter of taste. *Curr. Opin. Microbiol.* 2, 588–592
68 Zawadzki, P. *et al.* (1995) The log-linear relationship between sexual isolation and sequence divergence in *Bacillus* transformation is robust. *Genetics* 140, 917–932
69 Graham, J.B. and Istock, C.A. (1978) Genetic exchange in *Bacillus subtilis* in soil. *Mol. Gen. Genet.* 166, 287–290
70 Duncan, K.E. *et al.* (1989) Genetic exchange between *Bacillus subtilis* and *Bacillus licheniformis*: variable hybrid stability and the nature of bacterial species. *Evolution Int. J. Org. Evolution* 43, 1585–1609
71 Chen, X.H. *et al.* (2007) Comparative analysis of the complete genome sequence of the plant growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. *Nat. Biotechnol.* 25, 1007–1014
72 Gioia, J. *et al.* (2007) Paradoxical DNA repair and peroxide resistance gene conservation in *Bacillus pumilus* SAFR-032. *PLoS ONE* 2, e928
73 Branda, S.S. *et al.* (2005) Biofilms: the matrix revisited. *Trends Microbiol.* 13, 20–26
74 O'Toole, G. *et al.* (2000) Biofilm formation as microbial development. *Annu. Rev. Microbiol.* 54, 49–79
75 Otto, M. (2006) Bacterial evasion of antimicrobial peptides by biofilm formation. *Curr. Top. Microbiol. Immunol.* 306, 251–258
76 Hamon, M.A. and Lazazzera, B.A. (2001) The sporulation transcription factor Spo0A is required for biofilm development in *Bacillus subtilis*. *Mol. Microbiol.* 42, 1199–1209
77 Branda, S.S. *et al.* (2006) A major protein component of the *Bacillus subtilis* biofilm matrix. *Mol. Microbiol.* 59, 1229–1238
78 Chu, F. *et al.* (2006) Targets of the master regulator of biofilm formation in *Bacillus subtilis*. *Mol. Microbiol.* 59, 1216–1228
79 Kearns, D.B. *et al.* (2005) A master regulator for biofilm formation by *Bacillus subtilis*. *Mol. Microbiol.* 55, 739–749
80 Branda, S.S. *et al.* (2004) Genes involved in formation of structured multicellular communities by *Bacillus subtilis*. *J. Bacteriol.* 186, 3970–3979