JOURNAL OF BACTERIOLOGY, Nov. 2008, p. 6983–6995 Vol. 190, No. 21
0021-9193/08/$08.00+0 doi:10.1128/JB.00722-08

The Origins of 168, W23, and Other *Bacillus subtilis* Legacy Strains[▽]†

Daniel R. Zeigler,[1]* Zoltán Prágai,[2] Sabrina Rodriguez,[2] Bastien Chevreux,[2] Andrea Muffler,[2] Thomas Albert,[3] Renyuan Bai,[2]‡ Markus Wyss,[2] and John B. Perkins[4]

Bacillus Genetic Stock Center, The Ohio State University, Columbus, Ohio[1]; DSM Nutritional Products, Kaiseraugst, Switzerland[2]; NimbleGen Systems Inc., Madison, Wisconsin[3]; and DSM Anti-Infectives, DAI Innovation, Delft, The Netherlands[4]

Received 21 May 2008/Accepted 13 August 2008

***Bacillus subtilis* is both a model organism for basic research and an industrial workhorse, yet there are major gaps in our understanding of the genomic heritage and provenance of many widely used strains. We analyzed 17 legacy strains dating to the early years of *B. subtilis* genetics. For three—NCIB 3610[T], PY79, and SMY—we performed comparative genome sequencing. For the remainder, we used conventional sequencing to sample genomic regions expected to show sequence heterogeneity. Sequence comparisons showed that 168, its siblings (122, 160, and 166), and the type strains NCIB 3610 and ATCC 6051 are highly similar and are likely descendants of the original Marburg strain, although the 168 lineage shows genetic evidence of early domestication. Strains 23, W23, and W23SR are identical in sequence to each other but only 94.6% identical to the Marburg group in the sequenced regions. Strain 23, the probable W23 parent, likely arose from a contaminant in the mutagenesis experiments that produced 168. The remaining strains are all genomic hybrids, showing one or more "W23 islands" in a 168 genomic backbone. Each traces its origin to transformations of 168 derivatives with DNA from 23 or W23. The common prototrophic lab strain PY79 possesses substantial W23 islands at its *trp* and *sac* loci, along with large deletions that have reduced its genome 4.3%. SMY, reputed to be the parent of 168, is actually a 168-W23 hybrid that likely shares a recent ancestor with PY79. These data provide greater insight into the genomic history of these *B. subtilis* legacy strains.**

Bacillus subtilis, a model organism for gram-positive bacteria, is the focus of diverse research interests in both academic and industrial settings (54, 63, 64). One primary emphasis is sporulation, an archetypical form of cell development (25, 58, 66). Of equal interest, however, is the production of commercially attractive levels of small metabolites and enzymes (46, 57, 61). Investigations with *B. subtilis* benefit from the ease of its genetic manipulation (32), the wealth of available physiological and biochemical data (64), and the accessibility of a well-annotated genome sequence (40). Associated technologies are leading to a growing understanding of the proteome (33, 39, 74), the transcriptome (43), the metabolome (51), and the metabolic flux patterns (26, 41, 59) of this organism.

B. subtilis strains used in virtually all academic research and many industrial processes derive from a single tryptophan-requiring auxotroph, strain 168. Despite its central importance to research, our knowledge of this strain's genomic heritage is incomplete. Strain 168 was isolated after *B. subtilis* Marburg was mutagenized with X-rays by two Yale University botanists, Paul Burkholder and Norman Giles (15). Their experiments showed that for both vegetative cells and spores of *B. subtilis*, sublethal doses of UV or X-rays caused high frequencies of auxotrophy among survivors. For many mutants, the auxotrophic requirement could be met by a single nutrient, allowing the "one gene-one enzyme" model previously developed for *Neurospora* (8) and *Escherichia coli* (42) to be extended to this gram-positive spore former as well. Soon afterwards, the Yale group abandoned its *B. subtilis* studies to pursue other research interests (24). Sadly, most of its *B. subtilis* collection, including the wild-type parent, was subsequently lost. At least five mutants, however—auxotrophs requiring threonine (strain 23), nicotinic acid (strain 122), or tryptophan (strains 160, 166, and 168)—were preserved and transferred to the possession of Charles Yanofsky. Nearly a decade later, Yanofsky provided the mutants to John Spizizen (65), who in a landmark publication demonstrated that three of them—122, 166, and 168—could be transformed to prototrophy when exposed to DNA from strain 23 (65). The highly transformable strain 168 became the subject of follow-up studies detailing this phenomenon (6, 79). Strain 168 was subsequently disseminated around the world. Researchers soon developed classical genetic methods—and later, recombinant and genomic technologies—to elucidate the physiology and spore development of this strain. By the mid-1970s, so many mutants had been developed from 168 that a centralized repository, the Bacillus Genetic Stock Center (BGSC), was established to maintain them (81).

Strain 168 and its siblings are not the only legacy strains surviving from the earliest years of *B. subtilis* genetics, however. By the early 1960s, strain W23 began appearing regularly in the literature (20, 69, 77, 78). Its origin has remained a mystery (34). A few researchers have explicitly stated, without citing evidence, that W23 is derived from Burkholder and Giles strain 23 (14). The connection is not intuitively obvious, however: strain 23 requires threonine (49, 65), while W23 is prototrophic and streptomycin resistant (69). Although W23 was initially considered a wild-type equivalent to 168, it was soon discovered that these strains differed significantly in bacterio-

* Corresponding author. Mailing address: Bacillus Genetic Stock Center, The Ohio State University, 484 W. 12th Ave., Columbus, OH 43210. Phone: (614) 292-5550. Fax: (614) 292-6773. E-mail: zeigler.1@osu.edu.

† Supplemental material for this article may be found at http://jb.asm.org/.

‡ Present address: The Johns Hopkins University, School of Medicine, Department of Neurosurgery, Baltimore, MD.

▽ Published ahead of print on 22 August 2008.

phage sensitivity (34, 68, 76), cell wall chemistry (17), prophage content (52), and transformability (34, 65). These phenotypic differences, combined with the observation that the strains showed only 60 to 70% relatedness in genomic hybridization studies, led to the formal description of two novel subspecies, *B. subtilis* subsp. *subtilis* for 168 and *B. subtilis* subsp. *spizizenii* for W23 (48).

Before these differences were understood, W23 and its derivatives were often used as DNA donors in strain construction experiments (6, 49, 65). When the BGSC was established, considerable effort was made to exclude hybrids from the collection (D. H. Dean, personal communication). Nevertheless, over 45% of the strain 168 derivatives maintained by the BGSC have a *trpC*$^+$ genotype. Because the 168 *trpC2* allele is nonreverting (4), these strains necessarily received their *trpC*$^+$ alleles from another source, possibly W23. "Contaminating" DNA in such strains is not necessarily limited to *trpC*. Genetic markers that are normally unlinked during transformation can nevertheless be cotransferred if the donor DNA concentration is saturating (5, 50). This phenomenon, called congression, is a consequence of the *B. subtilis* transformation mechanism. A competent cell has about 50 binding sites used for DNA uptake (22). Over one-half of any bound DNA fragments eventually integrate into homologous chromosomal sequences (10). When strains are constructed with saturating DNA, if the donor contains two unlinked genetic markers and only one of them is selected during transformation, it is common to find that 1 to 5% of the transformants have received the unlinked, unselected marker as well (19). Clearly, substantial W23 DNA could have entered the genomes of *B. subtilis* legacy strains undetected.

Such concerns are of more than historical interest. In recent years, researchers have addressed deep questions about *B. subtilis* physiology and development with novel "-omics" technologies. Each laboratory engaged in -omics analysis has used its own preferred strains. Use of a Trp$^+$ legacy strain is not uncommon (13, 43, 47, 62). If the *B. subtilis* community is to correlate data from different -omics projects and connect these results with biochemical and genetic analyses from previous decades, it is essential to first understand how these strains are related. Unfortunately, it is not always possible to reconstruct strain genealogies from published descriptions or even from archived laboratory notebooks. Even when a strain pedigree is understood, it is difficult to predict which parent contributed any given section of the genome. Undetected "contaminating" W23 DNA introduces a potentially significant uncontrolled component to comparing and correlating data obtained with Trp$^+$ strains. Furthermore, several industrial strains derived from 168, such as riboflavin producers, also have a Trp$^+$ phenotype (56). The introduction of divergent DNA sequences into an industrial strain could alter the activity of the encoded proteins, potentially affecting the overall performance of the process.

To address these concerns, we performed DNA sequence analysis of 17 *B. subtilis* legacy strains. We compared an 84-kb region surrounding *trpC* in the genome for 13 strains and smaller sequence samples from four of the Burkholder and Giles mutants. We also utilized the NimbleGen comparative genome sequencing method to compare the genomes of three well-known *B. subtilis* strains, NCIB 3610^T, PY79, and SMY. The data generated from this study provide several revealing insights into the early history of *B. subtilis* genetics as well as practical guidance about appropriate wild-type reference strains for comparison to strain 168.

MATERIALS AND METHODS

Bacterial strains and growth conditions. *B. subtilis* strains used in the study are listed in Table 1. All strains are available from the BGSC (The Ohio State University, Columbus, OH), except for ATCC 6051^T, which was obtained from the American Type Culture Collection. Growth and maintenance were done by standard methods (32).

Isolation of template DNA for PCR. Several colonies from an overnight culture on solid medium were suspended in 1 ml of buffer (10 mM Tris-HCl, pH 8.5), washed four times in buffer, and then suspended in 0.5 ml of fresh buffer. An equal volume of 0.1-mm glass beads (Scientific Industries, Inc.) was added to the tube, and the contents were mixed for 1 min in a bead beater apparatus (Disruptor Genie; Scientific Industries, Inc.). Finally, the tube was centrifuged at 21,000 $\times$ *g* for 5 min. Up to 200 μl of supernatant was withdrawn from the tube for further use. For PCR amplification, 0.5 μl of this cell lysate was used as template DNA without further purification. This DNA preparation could be stored at 4°C for up to several months without a loss of quality.

PCR amplification. The FastStart high-fidelity PCR system (Roche) was used for all amplifications, performed under the manufacturer's recommended reaction conditions. All primers (listed in Table S1 in the supplemental material) were synthesized at Operon Biotechnologies, Inc., making use of the genome sequence for *B. subtilis* 168 (SubtiList release R16.1 [http://genolist.pasteur.fr/SubtiList/]). Cycling parameters included an initial denaturation step (2 min at 95°C); 35 cycles of denaturation (30 s at 95°C), annealing (30 s at 57°C), and extension (60 s per kilobase at 72°C); and a final extension step (7 min at 72°C). Reaction products were purified using a High Pure PCR product purification kit (Roche) following the manufacturer's specifications.

PCR fragment cloning. In a few cases, PCR products were cloned prior to being sequenced, using the pCR4-TOPO vector and a Topo TA cloning kit for sequencing (Invitrogen) following the manufacturer's protocols. Recombinant plasmids were isolated with a High Pure plasmid isolation kit (Roche) following the manufacturer's protocol.

DNA sequencing, sequence assembly, and analysis. DNA sequencing was performed at the Plant-Microbe Genomics Facility of The Ohio State University, using an automated model 3730 DNA analyzer and BigDye Terminator (Applied Biosystems, Inc.) cycle sequencing chemistry. Most DNA sequences were obtained using purified PCR products as templates and the amplification primers as sequencing primers. Plasmid inserts were sequenced using standard primers. At least four separate determinations, including at least two pairs of primers, were performed for each residue in each DNA sequence. Contiguous sequences were assembled with the SeqMan II software package (DNAStar, Inc.). Multiple alignments were generated by ClustalW (36).

CGS. Comparative genome sequencing (CGS) is a two-step array-based mutation mapping and identification technique (3, 35, 45, 75). The first step compares the hybridization characteristics of a test genome against a previously sequenced reference genome, using array probes that are designed to "tile" densely across the reference sequence. This step maps the positions of single point mutations and of insertions and deletions of >5 bp to within 10 to 20 nucleotides. (Smaller insertions and deletions are often not identified, because single base changes appear to disrupt hybridization more strongly than small insertions and deletions do.) The second step uses directed array-based resequencing to identify the precise locations of the potential mutations detected in the first step. Resequencing depends on the differential hybridization of genomic fragments to perfect-match (PM) and mismatch oligonucleotides. Array oligonucleotide probes vary in length from 29 to 39 bases, depending on the probe melting temperature. Each nucleotide to be queried is located at a central position of each oligonucleotide probe. For each PM oligonucleotide, probes representing the three possible mismatch nucleotides at the query position were also synthesized on the array. The differences in hybridization signal intensities between sequences that bind strongly to the PM oligonucleotide and those that bind poorly to the corresponding mismatch oligonucleotides make it possible to discern the correct base at a given sequence position.

For this study, genomic DNA samples of 168 (obtained from the Pasteur Institute from the isolate used as the primary source of the published genome sequence), NCIB 3610^T, PY79, and SMY were prepared using a Qiagen genome 5-G isolation kit with a genomic DNA buffer set. These genomic DNA preparations were then separately hybridized against NimbleGen CGS resequencing microarrays that were manufactured using the published 168 sequence as the

TABLE 1. Bacterial strains used in this study

Strain	BGSC no.	Strain history[a]	Description as received	Construction/isolation	Reference[b]
168	1A1	BGSC←J. A. Shapiro←A. L. Sonenshein←P. Shaeffer←C. Anagnostopoulos	*trpC2*	X-ray or UV mutagenesis	15, 65
23	2A1	BGSC←J. Lederberg←J. Spizizen	*thr*	X-ray or UV mutagenesis	15, 65
122	1A410	BGSC←C. Yanofsky←J. Spizizen	*nic*	X-ray or UV mutagenesis	15, 65
160	3A12	BGSC←H. Callister←P. R. Burkholder	*trpB3*	X-ray or UV mutagenesis	15
166	1A65	BGSC←C. Yanofsky←J. Spizizen	*trpE26*	X-ray or UV mutagenesis	15, 65
W23	2A9	BGSC←NRRL B-14472←ATCC 23059←K. F. Bott←C. B. Thorne	Prototroph	Unknown	69
W23SR	2A3	BGSC←C. B. Thorne←M. S. Fox	*str* (prototroph)	Unknown	69
NCIB 3610^T	3A1^T	BGSC←NCIMB←R. S. Breed←H. J. Conn←University of Marburg	Wild type (Marburg)	Unknown	18
ATCC 6051^T		ATCC←H. J. Conn← University of Marburg	Wild type (Marburg)	Unknown	18
W168	1A308	BGSC←R. H. Doi←N. Sueoka	Prototroph	Transformation of 168 *his trpC2* with W23 DNA	67
PS832	1A757	BGSC←P. Setlow←D. Tipper	Prototroph	Unknown	44
GSY505	1A361	BGSC←C. Anagnostopoulos	*ilvA6*	Unknown	82
MU8U5U1	1A75	BGSC←N. C. Brown←D. A. Dubnau←N. Sueoka	*ilvA1 leuB8 metB5*	W168 exposed to three cycles of UV mutagenesis	77
PY79	1A747	BGSC←A. Driks←R. Losick←P. Youngman	Prototroph SPβ^s	Two transductions of CU1769 with PBS1 phage grown on 168	80
SB19	2A10	BGSC←J. Lederberg	Prototroph	Transformation of 168 to prototrophy with 23 DNA	49
SB491	1A2	BGSC←J. Lederberg←M. Shafer	Prototroph	Unknown	16
SMY	1A775	BGSC←T. Henkin←A. L. Sonenshein←J. Segall	Prototroph	Unknown	28

[a] BGSC, Bacillus Genetic Stock Center, Columbus, OH; NRRL, Agricultural Research Service Culture Collection, Peoria, IL; ATCC, American Type Culture Collection, Manassas, VA; NCIMB, NCIMB Ltd., Aberdeen, Scotland.

[b] Reference(s) describing the strain's provenance, if known; otherwise, earliest known publication(s) reporting its existence or use.

reference sequence. The sequence differences between each genome and 168 were extracted, recompiled in a visualization format, and sent to DSM Nutritional Products for analysis. If a sequence difference could be determined accurately by full resequencing microarrays, it was classified as a single nucleotide polymorphism (SNP). If a difference could not be determined accurately, it was reported as a region of interest (ROI). For comparison of strain 3610 against 168, common SNPs from two independent experiments are reported. Unpublished validation studies conducted at DSM Nutritional Products have estimated the accuracy of the NimbleGen CGS method. For two closely related strains from a well-characterized lineage, the SNP pattern determined by Sanger sequencing was compared to the SNP pattern generated by NimbleGen CGS. Results showed 80% accuracy for base pair changes (data not shown). Conversely, it was also found that there was an approximately 20% false-positive discovery rate. Therefore, the net effect is that the number of detected SNPs represents a good approximation of the sequence difference between strains.

Construction of His_6-tagged PanB expression cassettes. The *B. subtilis panB* gene was amplified from PY79 and 168 by PCR under standard conditions, using a 3′ primer containing a 15-bp spacer. The *panB* gene was then cloned into *E. coli*/pET-21b. This resulted in plasmids pPA304 and pPA306, which when expressed in *E. coli* BL21(D3) generated proteins with a spacer of five amino acids between the N-terminal end of PanB (ketopantoate hydroxymethyltransferase) and the His tag. Cultivating the cells from strains PA305 (pPA304) and PA307 (pPA306), with or without IPTG (isopropyl-β-D-thiogalactopyranoside) induction, the expected 29.6-kDa protein band was visualized by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Also, both plasmids were able to complement an *E. coli panB* mutant (SJ2) upon IPTG induction (data not shown). The *B. subtilis panE* (*ytbQ*) gene was PCR amplified from *B. subtilis* 168 by using primers based on the published genome sequence. The *panE* gene was then cloned with an N-terminal His_6 tag into the expression vector pQE80L and expressed in *E. coli* Fe8 cells under standard IPTG induction conditions. Ketopantoate reductase activity of the purified His_6-PanE protein was subsequently confirmed by the hydrolysis of NADPH in the presence of 0.5 mM α-ketopantoate; the V_{max} of the reaction was 100 μmol NADPH/min/mg at room temperature and pH 7 (data not shown).

PanB enzymatic assay. His_6-tagged PanB proteins from *B. subtilis* PY79 and 168 were purified with Ni-nitrilotriacetic acid columns by standard procedures. Enzymatic activity was measured by a coupled forward reaction. The assay mixture contained 0.5 mM α-ketoisovalerate (α-KIV), 0.4 mM 5,10-methylenetetrahydrofolate (MTHF), 0.25 mM NADPH, 5 mM $MgCl_2$, 0.25 μg His_6-PanE (*B. subtilis*), and 12.5 μg His_6-PanB. Hydrolysis of NADPH (μmol NADPH/min/mg protein) was measured at room temperature by monitoring the absorbance at 340 nm.

Nucleotide sequence accession numbers. Nucleotide sequences were submitted to GenBank under accession numbers EF191442 to EF191614, EF193037 to EF193038, EU081774, EU084745 to EU084749, and EU146074 to EU146104. Individual accession numbers are listed in Table S2 in the supplemental material.

RESULTS

Survey of the *trpC* region in 17 *B. subtilis* legacy strains. One likely site of genetic heterogeneity within the genomes of *B. subtilis* legacy strains is the *trpC* locus. Because the *trpC2* mutation in strain 168 is known to be a nonreverting 3-bp deletion (4), the presence of a trp^+ allele in a strain derived from 168 is presumptive evidence that nonisogenic DNA has been introduced at the locus. To assess the degree of genomic heterogeneity in 14 *B. subtilis* legacy strains, we sampled the DNA sequence from an 84-kb region of the chromosome encompassing *trpC*. Each sample included a contiguous 16.6-kb sequence flanking the *trp* operon and 12 shorter sequences dispersed throughout the 84-kb region, bringing the sample length to 29.4 kb. For four other strains, smaller sequence samples were obtained from the region immediately surrounding *trpC*. (A detailed list of strains, sequenced regions, and accession numbers is given in Table S2 in the supplemental material.)

Multiple sequence alignments (not shown) sorted the strains into three groups, as summarized in Fig. 1. The first group

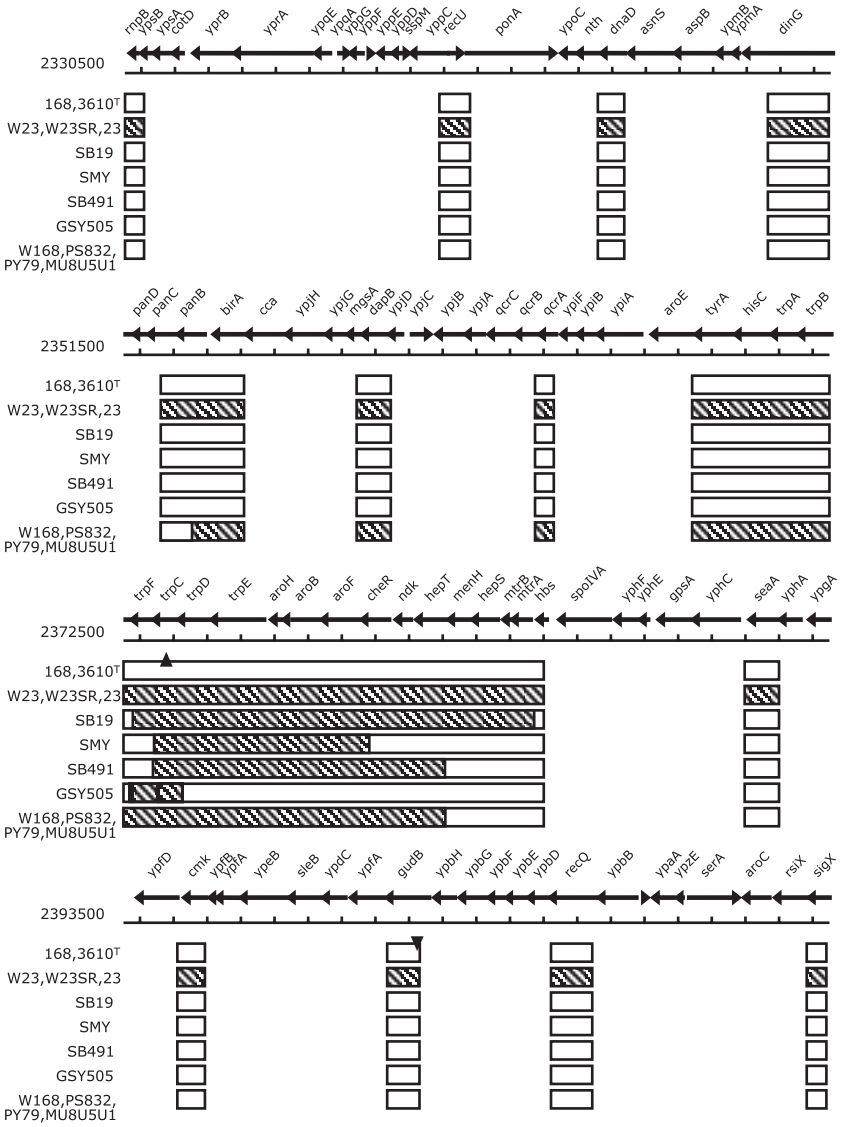


FIG. 1. DNA sequence heterogeneity in the *trp* region of *B. subtilis* legacy strains. The arrows show reading frames in the *B. subtilis* 168 genome, while the axis indicates the corresponding position in the genome sequence. For each strain, boxes indicate the sizes and positions of contiguous DNA sequences determined in this study. Open boxes show sequences that are essentially identical to that of strain 168, while hatched boxes show sequences identical to that of W23. The upward-facing triangle (▲) shows the position of the *trpC2* deletion, while the inverted triangle (▼) shows the position of the *gudB* duplication. The strain group labeled "168,3610^T" also includes ATCC 6051^T and the Burkholder and Giles mutants 122, 160, and 166.

included the Marburg strains NCIB 3610^T and ATCC 6051^T and the Burkholder and Giles mutants 122, 160, 166, and 168. These strains were nearly identical in sequence and showed a high level of agreement with the published genome sequence for strain 168 (40). The second group, including strains W23 and W23SR and Burkholder and Giles strain 23, were identical in sequence to each other but only 94.3% identical to the 168-like strains. The final group displayed hybrid genomes, with W23-like sequences for *trpC* and immediate flanking regions but 168-like sequences for more distant regions of the chromosome. As shown in Fig. 1, the sizes of these "W23 islands" ranged from 1.6 to 28.6 kb, depending on the strain tested. Four strains—W168, PY79, PS832, and Mu8u5u1—showed an identical 28.6-kb "island."

Within the first group, the 168-like strains, a SNP was detected within the *aroH* coding sequence. The sequence of codon 112 in the two Marburg type strains is GCT, encoding alanine, while the same codon in the Burkholder and Giles mutants is GTT, encoding valine. An NCBI BLAST search showed that alanine is very highly conserved at this position, occurring in chorismate mutases from genera as diverse as *Clostridium*, *Geobacillus*, and *Listeria* as well as in taxa far removed from the gram-positive lineage (not shown).

The *gudB* locus also showed two alleles. The coding sequence from 168 is known to have a 9-bp repeat that renders the GudB protein inactive in this strain (9). We observed this allele in Burkholder and Giles strains 160 and 166 as well as in each of the genomic hybrid strains. A spontaneous deletion

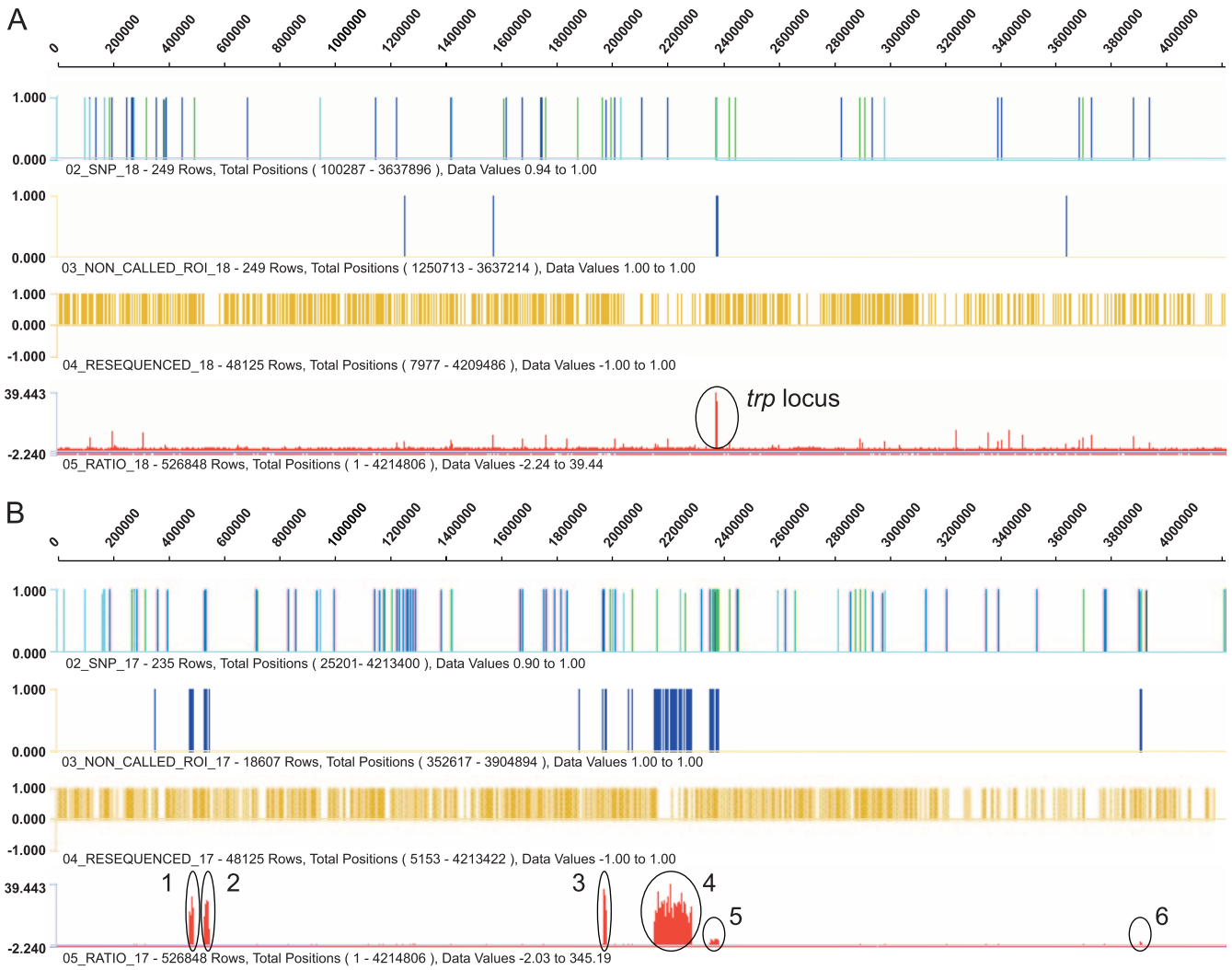


FIG. 2. Visualization of DNA sequence heterogeneity revealed by genome resequencing of *B. subtilis* strains SMY (A) and PY79 (B). For each figure, four panels are shown aligned to the position in the *B. subtilis* genome sequence. Bar lines in the first row indicate called SNPs, as follows: black bars, nonsense SNPs; blue bars, missense SNPs; light blue bars, silent (no amino acid change) SNPs; turquoise bars, SNPs within intergenic region. Blue bars in the second row show uncalled ROI. Gold lines in the third row show fully resequenced residues, and red lines in the fourth row show the net difference (arbitrary units) in hybridization intensity between the two tested strains, with the indicated "peaks" indicating large hybridization differences. For PY79, the positions of areas 1 to 6 are marked.

that removes the repeat and restores GudB activity to strain 168 has been described (9). Our sequences showed that this functional allele occurs naturally in the two wild Marburg strains as well as in Burkholder and Giles strain 122.

Legacy strain alleles affecting swarming behavior. Unlike NCBI 3610^T and other wild strains, *B. subtilis* 168 does not swarm on motility agar plates. Two genes required for swarming, *sfp* and *swrA*, are functional in NCBI 3610^T but are each inactivated in strain 168 due to single base duplications that disrupt their reading frames (38). To further probe the relationship between the Marburg strains ATCC 6051^T and NCBI 3610^T and the Burkholder and Giles strains, we determined the sequences of their *sfp* and *swrA* alleles. Our results confirmed that ATCC 6051^T, like NCBI 3610^T, has the functional alleles for both *sfp* (GenBank accession no. EU146075 to EU146076) and *swrA* (GenBank accession no. EU146080 to EU146081). In contrast, all four of the Burkholder and Giles strains share the frameshift mutations noted for strain 168 (GenBank accession no. EU146074, EU146077 to EU146079, and EU146082 to EU146085). Clearly, the nonswarming alleles of *sfp* and *swrA* were already present in the Burkholder and Giles parent strain, unlike the existing Marburg strains.

Genome "resequencing" of NCIB 3610^T, SMY, and PY79. The genomes of three of the strains in the legacy strain set, NCIB 3610^T, SMY, and PY79, were selected for comparative genome hybridization, or "resequencing," using Nimblegen technology. First, the hybridization of each genomic DNA sample to reference tiling microarrays manufactured based on the published 168 genome sequence was compared to the hybridization of 168 genomic DNA. Next, positions showing likely polymorphisms were then analyzed by full resequencing microarrays. Results for SMY and PY79 are shown in Fig. 2. A complete list of individual SNPs detected in the resequenced genomes is provided in Table S3 in the supplemental material.

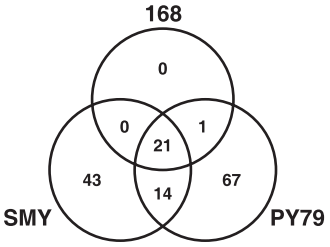


FIG. 3. SNPs shared among *B. subtilis* 168, SMY, and PY79. Numbers indicate which SNPs are unique to a strain, which are shared between two strains, and which are shared among all three strains. The wild-type Marburg strain NCIB3610^T was chosen as the root sequence for this comparison. Residues falling within "W23 islands" were excluded from the comparison.

A summary of these results, tallying the SNPs that are unique to each strain and those that are shared by two or more strains, is presented in Fig. 3.

In addition to these SNPs, the hybridization data detected several ROI, that is, hybridization differences that could not be correlated with specific sequence differences on full resequencing microarrays. This kind of hybridization difference may correspond either to a false-positive result or to a locus where more than one SNP is located within a 28-bp region. Comparison of NCIB 3610^T and SMY to 168 revealed fewer than 10 ROI. Comparison of PY79 to 168 revealed a larger number of ROI, but all corresponded to deletions or W23 "islands" (Fig. 2; see below).

One limitation of comparative genome resequencing technology is the difficulty in detecting small insertions or deletions. For this reason, several of the differences between 168 and NCIB 3610^T that we demonstrated by conventional sequencing—such as the duplications in *gudB*, *swrA*, and *sfp* and the deletion in *trpC*—were not observed with the microarray hybridization data. The SNP in *aroH* was observed, however. In all, these genomes were distinguished by only 22 SNPs, indicating an extremely close relationship between them. The 168 and SMY genomes were measurably more divergent. The "W23 island" surrounding the *trpC* locus of SMY was clearly visible in the hybridization data (Fig. 2A). Outside this island, an additional 57 SNPs distinguished SMY from 168 (Fig. 3).

Among the resequenced genomes, PY79 showed the greatest divergence from 168. The "W23 island" surrounding *trpC* in the PY79 genome was clearly seen in the data, as was a second region of sequence heterogeneity near the *sacA* locus. Four large deletions in PY79 (discussed below) were revealed by long regions of significant hybridization differences compared to 168 (Fig. 2B). A smaller deletion was also detected when we used conventional Sanger sequencing to examine a cluster of SNPs lying within the *rrnIHG* locus. (These SNPs are visible as vertical bars in the first row of Fig. 2B, at about 162 to 170 kb in the genome sequence.) Analysis of this locus in PY79 (GenBank accession no. EU081774) revealed that a complex rearrangement between *rrnH* and *rrnG* had resulted in the net loss of one rRNA operon from the PY79 genome. Similar deletions of rRNA operons have been detected in other laboratory strains of *B. subtilis* (73). Besides these large-scale changes in the genome, PY79 differed from the 168 reference sequence by 81 additional SNPs, 14 of which it also had in common with SMY (Fig. 3).

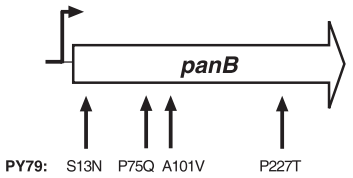
Sequence of the *sacA* region in PY79, 168, and W23. To confirm the presence of a second "W23 island" in the PY79 genome, we sequenced a 5.7-kb region stretching from *sacA* to *ywcI* in PY79, 168, and W23. Multiple alignments of these sequences (not shown) confirmed that in PY79, a contiguous 3.9-kb stretch of W23-like DNA has replaced the orthologous 168-like sequences. The "left" endpoint of the island lies between residues 3902561 and 3902583, and the "right" endpoint lies between residues 3906406 and 3906471. As a result, PY79 has W23-like sequences for *sacP*, *ywcJ*, *sacT*, and *ywcI* and a hybrid sequence for *sacA*. A smaller sample sequence (not shown) indicated that strains W168, PS832, and Mu8u5u1 are also W23-like in their *ywcI* genes.

Determination of endpoints for the four large deletions in PY79. The resequencing data provided strong evidence that four large deletions exist in the genome of PY79 and suggested the approximate locations of their endpoints. Using this information, we designed primers to amplify and sequence each deletion end-join fragment (Table 2). We found that the two largest deletions, "area 2" (20 kb) and "area 4" (134 kb), correspond to the loss of the ICEBs1 mobile genetic element and the SPβ prophage, respectively. PY79 is known to be cured of SPβ (80), and the excision and curing of ICEBs1 from other *B. subtilis* strains have been reported (7). The "area 3" deletion was caused by a recombination between homologous sequences in the plipastatin synthetase operon (*ppsABCDE*). In strain 168, there are several long stretches of near sequence identity lying in tandem within the operon. ClustalW alignments (not shown) between 168 and PY79 show that in PY79 a recombination event occurred between residues 1967024 and 1967085 in *ppsD* and residues 1977806 and 1977867 in *ppsA*, deleting 10,782 bp from the genome in the process. The "area 1" deletion in PY79 was found to extend from residues 475358 to 491807 in the genome sequence. The deletion removed 16,448 bp from the PY79 genome, encompassing all or part of 13 open reading frames. The PY79 sequence suggested no obvious model for the deletion; there are no significant regions of sequence homology near these endpoints, nor do there appear to be any prophages or mobile elements in the region. In total, these four deletions removed 180 kb from the PY79 genome, reducing it 4.3% in total size. Diagnostic PCRs showed that among the legacy strains in our test set, the "area 4" deletion (curing the SPβ prophage) is fixed only in the

TABLE 2. Large areas of sequence difference between PY79 and 168

Area	Size (kb)	Type of genome alteration[a]	Loci
1	17	Δ	*ydzA* to *ydaQ*
2	20	Δ	*ydcL* to *yddM*
3	9	Δ	*ppsD* and *ppsC*
4	134	Δ	SPβ
5	29	ΩW23	*panB* to *hepT*
6	4	ΩW23	*sacAP*, *ywcJ*, *sacT*, *ywcI*

[a] Δ, large deletion; ΩW23, replacement of 168-like sequences with orthologous W23-like sequences.



His-PanB	Specific Activity (μmol NADPH/min/mg)	$K_{m/a}$/KIV (μM)	K_m/MTHF (μM)	K_i/pantoate (μM)
168	1.15 ± 0.07	100	100	100
PY79	0.38 ± 0.03	300	200	90

FIG. 4. Correlation between sequence divergence and enzyme activity of PanB. The top panel shows the positions of the amino acid differences between the PanB enzymes from 168 and PY79; the bent arrow represents the wild-type sigma A promoter that controls transcription of *panB*. The bottom panel compares the activities of the His_6-tagged PanB enzymes isolated from *B. subtilis* 168 and PY79 (with the latter containing the 29-kb *B. subtilis* W23-derived DNA island surrounding the *trp* locus). The values are means for three independent experiments.

genomes of PY79 and SB491, while the other three deletions are fixed in PY79 alone (see Fig. S1 in the supplemental material). These diagnostic PCRs revealed that in the remaining legacy strains, the area 2, 3, and 4 deletions are not fixed in the genome but may exist in a minority of cells within a culture (data not shown).

Comparison of PanB from 168 and PY79. DNA sequence divergence between identical genes has the potential to alter the biochemical properties of the encoded proteins or enzymes. To test this possibility, the enzyme properties of the PanB enzymes encoded by 168 and the "W23 island" in PY79 were characterized. PanB (ketopantoate hydroxymethyl transferase) transfers a hydroxymethyl group from MTHF to α-KIV to form ketopantoate. The activity of this enzyme is under negative feedback control by pantoate. Inspection of the DNA sequences from the two strains revealed four amino acid differences, namely, S13N, P75Q, A101V, and P227T (Fig. 4). We prepared His_6-tagged versions of the 168 and W23 (PY79) PanB enzymes (referred to as $PanB_{168}$ and $PanB_{W23}$, respectively) as described in Materials and Methods and determined the enzymatic activities by using a PanE-coupled assay measuring NADPH-to-$NADP^+$ conversion at 340 nm. As shown in Fig. 4, $PanB_{168}$ displayed a threefold higher specific activity than that of $PanB_{W23}$. Moreover, the K_m values for α-KIV and MTHF were similarly higher for $PanB_{W23}$ than for $PanB_{168}$, whereas the K_i values for pantoate were similar between the two enzymes. Interestingly, an intermediate level of activity was observed when a His_6-tagged PanB protein was prepared from a recombinant *panB* gene containing only three of the four SNPs (S13N, P75Q, and A101V) (data not shown). Finally, this difference in PanB activities directly resulted in similar differences in pantothenate overproduction when the three *panB* genes were separately introduced into an integrated single-copy engineered *panBCD* operon driven by a constitutive promoter that had been integrated in single copy into the chromosome (data not shown).

Construction of an isogenic Trp^+ derivative of *B. subtilis* 168. All of the Trp^+ derivatives of strain 168 examined in this study contained substantial amounts of "contaminating" DNA sequences derived from strain W23. While related wild-type isolates are available, there are no known Trp^+ strains that are completely isogenic to 168. For this reason, we transformed strain 168 to prototrophy with an 810-bp fragment amplified from the NCIB 3610^T *trpC* gene. The *trpC* sequence of the transformant was confirmed to be completely 168-like, except for the correction of the 3-bp *trpC2* deletion. This transformant was deposited into the publicly available BGSC collection as strain 1A900.

DISCUSSION

Although they have been used widely in basic and applied research for 5 decades, the genomic heritage of many *Bacillus subtilis* legacy strains is poorly understood. The whole-genome and locus-specific DNA sequences we report here offer a clearer picture of the relationships among these strains and the genetic events that produced them. A model describing the deduced genomic heritage of these strains is presented in Fig. 5 and detailed below. Every DNA sequence analyzed during this study could be explained as deriving ultimately from either a 168-like or W23-like parent genome. Although we cannot rule out the possibility that further sequencing could detect additional independent parents, we have no evidence for such strains and so have ignored this possibility in our model.

Origin of *B. subtilis* 168 and its siblings (Fig. 5A and B). The provenance of *B. subtilis* 168 is well known: it was isolated as a Trp^- mutant by Burkholder and Giles at Yale University in the early years of the postwar era (15). But the identity of its parent strain, now presumed lost, has remained something of a mystery. Although the Yale researchers explicitly called their organism the "Marburg strain" of *B. subtilis*, surviving Marburg isolates, such as ATCC 6051^T and NCIB 3610^T, seem much "wilder" than strain 168. They are capable of complex multicellular behaviors, forming swarming colonies that develop a variety of intricate architectures. In contrast, strain 168 lacks these abilities (2). Despite these differences, however, our sequence data establish that 168 is indeed a derivative of *B. subtilis* Marburg.

Harold J. Conn, then in residence at the Cornell Geneva station, published his description of *B. subtilis* Marburg in 1930 (18). That year, Conn deposited the strain in the ATCC under accession no. 6051. Two decades later, the British NCIB received the Marburg strain from Robert S. Breed, chief bacteriologist at the Geneva station, and accessioned it as strain 3610. Despite the divergent histories of these strains, our data failed to uncover even a single base difference between ATCC 6051^T and NCIB 3610^T. While a more complete characterization of the strains could still uncover differences, it seems likely that both are authentic representations of Conn's Marburg strain. Our NCIB 3610^T genome "resequencing" experiment uncovered only 22 SNPs relative to strain 168, despite the latter strain's history of mutagenesis. Whether Burkholder and Giles received their Marburg strain directly from the Geneva station or indirectly from the ATCC or another intermediate source remains unknown, but the identification of the strain as *B. subtilis* Marburg is strongly supported.

Nevertheless, our data suggest that before Burkholder and Giles isolated strain 168, its parent was subjected to at least

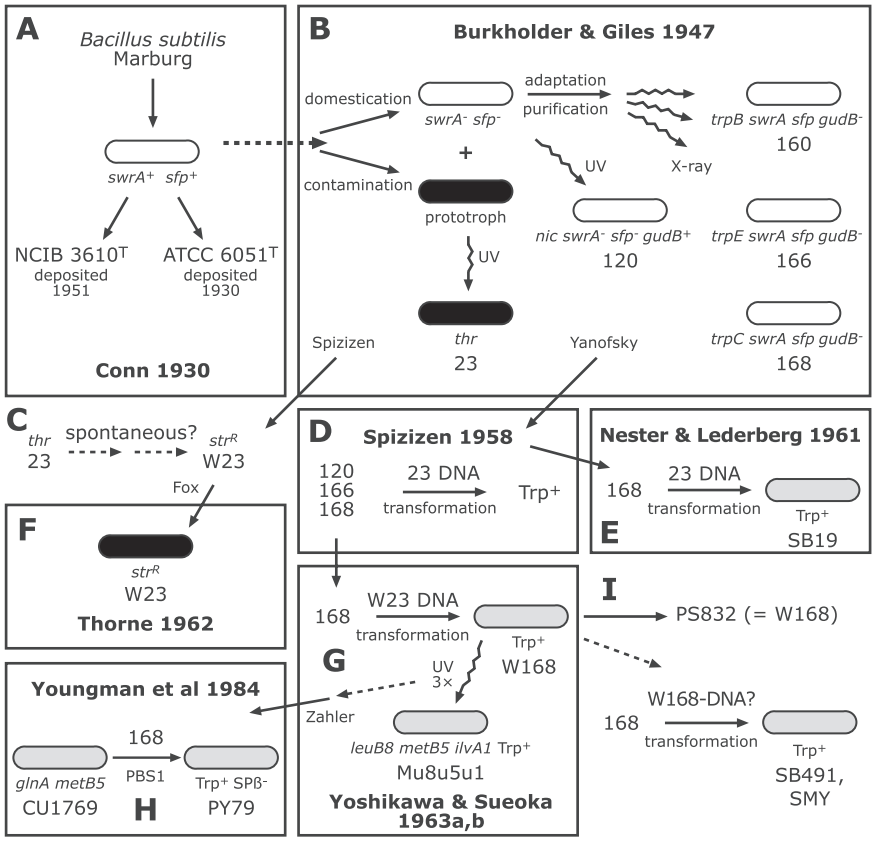


FIG. 5. Model describing the history of early *B. subtilis* Trp^+ legacy strains. The genomic heritage of individual strains is represented by shading, as follows: white cells, 168-like genomes; black cells, W23-like genomes; gray cells, 168-W23 hybrid genomes. The event leading to the isolation of a legacy strain is represented by a straight arrow (transformation, transduction, domestication, contamination, or renaming), zigzag arrow (mutagenesis through radiation), or dashed arrow (inferred but undocumented event). Rectangular boxes surround isolation events that are documented in the publications indicated.

two phases of domestication, that is, adaptation to life in the laboratory. Because the Geneva station supplied wild Marburg cultures to the ATCC in the 1930s and to the NCIB in the 1950s, this adaptation occurred elsewhere, probably at Yale. From our sequence data, we can infer that the first domestication took place before the mutagenesis experiments began. Strain 168 and its known siblings—122, 160, and 166—share at least three DNA sequence polymorphisms that distinguish them from wild Marburg strains. One, the *aroH* SNP documented in our sequences, has no known relationship to domestication and should perhaps be regarded merely as a sequence signature of the Yale strain. The other two polymorphisms, however—the single base duplications that inactivate *sfp* and *swrA*—are each related to a loss of swarming and multicellularity. These mutations make *B. subtilis* behave like the classical "Koch's postulate" model of a uniform bacterial culture composed of pure single cells (2), perhaps making it easier to use in experiments typically performed in the early days of bacterial genetics. Because all four siblings show these three mutations, the simplest explanation is that they were already present in the parental strain.

A second phase of domestication is documented in the Burkholder and Giles publication itself. After they completed their UV mutagenesis experiments but before they commenced X-ray mutagenesis, the researchers selected for a variant of their Marburg strain that was capable of faster growth on glucose-ammonia minimal medium (15). We propose that this variant corresponds to another sequence polymorphism observed in our data, the nine-base duplication that inactivates *gudB* in strains 160, 166, and 168. Interestingly, this mutation has been shown by others to enhance the growth rate of *B. subtilis* on media quite similar to those employed by Burkholder and Giles (9). Strain 122, in contrast to its siblings, has a functional *gudB* allele identical to what is seen in the wild Marburg strains. A possible explanation is that 122 arose by UV mutagenesis before the isolation of the faster-growing variant, while the other three siblings arose from the X-ray mutagenesis experiments conducted afterwards. The available data are consistent with this hypothesis. Our sequence of the strain 160 *trpB* locus (GenBank accession no. EU146100) reveals a two-base duplication that disrupts the reading frame. Our sequences also confirm that the 168 *trpC* locus is inactivated by a three-base deletion, as suggested by others (4). The 166 *trpE* locus is known to be inactivated by a large translocation (70). Each of these lesions is entirely consistent with X-ray mutagenesis. In contrast, the strain 122 *nic* mutation probably corresponds to a SNP in its *nadB* promoter, with the only difference we observed located in a 5-kb region encompassing the NAD biosynthetic cluster in strains 122 (GenBank accession no. EU146091) and 168 (GenBank accession no. EU146090). If

indeed this mutation corresponds to the 122 *nic* lesion, then it could easily have arisen from UV mutagenesis. Regardless of whether our exact interpretations are correct, our data strongly indicate that strain 168 and its siblings arose from an already domesticated version of the Marburg strain.

Origin of *B. subtilis* W23 (Fig. 5B, C, and F). When John Spizizen commenced his bacterial transformation experiments in the late 1950s, he found that strains 160, 166, and 168 each developed genetic competence naturally during stationary phase, while under the same conditions a fifth Burkholder and Giles auxotroph, the threonine-requiring strain 23, was nontransformable (65). As a result, strain 23 was a logical choice as a DNA donor for the three tryptophan-requiring mutants. Our DNA sequences establish, however, that strains 23 and 168 cannot be siblings but are instead independent isolates. The two strains show only 94.3% identity over nearly 30 kb of sequence. We speculate that Burkholder and Giles inadvertently worked with a mixed culture of *B. subtilis* during the early stages of their experiments and that strain 23 arose from the contaminating strain. If so, the same experiments that isolated a faster-growing variant of their Marburg strain would also have purified it from the contaminant prior to the isolation of strains 160, 166, and 168. While we have yet to determine the lesion causing threonine auxotrophy in strain 23, it does readily revert to prototrophy (E. W. Nester and our unpublished observations), consistent with the hypothesis that it was generated by UV mutagenesis during the first stage of the Yale experiments. Regardless, our sequences confirm the anecdotal evidence that strain 23 is the ancestor of strain W23 as well as strain W23SR. We detected no DNA sequence polymorphisms among the three strains in a 29.3-kb sequence sample.

The exact details of how strain W23 arose from 23 are still somewhat unclear, and our model at this point is admittedly speculative, though consistent with strain descriptions in the literature. In 1961, Curtis Thorne was at Fort Detrick, MD, isolating transducing phages for *B. subtilis* by using a strain described as "W-23-Sr, a streptomycin-resistant mutant of wild-type W-23," obtained from Maurice Fox, then at Rockefeller University (Fig. 5F) (69). This strain corresponds to BGSC strain 2A3, and perhaps to strain 2A9 as well. In 1965, Brodetsky and Romig of UCLA, who had also studied *Bacillus* phages for several years, described a strain W23Sr as "a prototrophic derivative of the threonine-requiring strain 23 (Spizizen, 1958) [that] is streptomycin resistant," which they acknowledged receiving from Thorne. Our data are consistent with this explanation for the origins of W23. It could have arisen in two steps from strain 23, following spontaneous reversion to prototrophy and mutation to streptomycin resistance. Interestingly, the genomes of strains 23 and 168, which were isolated independently at Yale in the 1940s, were to be mingled in hybrid strains decades later in several labs around the world.

Introduction of W23 DNA into the genomes of *B. subtilis* 168 Trp$^+$ legacy strains. Because the *trpC2* deletion of strain 168 cannot revert to the wild type, one would expect that all strains descended from it would be tryptophan auxotrophs. The fact that nearly one-half of the legacy strains maintained at the BGSC are Trp$^+$ requires an explanation, especially since many of them to this day remain important subjects for both academic and industrial research. We analyzed a representative group of *B. subtilis* strains in the hope that we could infer their parentage from the DNA sequences of their *trpC* regions. Our test group included strains that were constructed quite deliberately from crosses between W23 (or 23) and 168, another strain that was reputedly a wild-type "revertant" of 168 (strain SB491), still other strains with incomplete or missing historical records, and at least one strain alleged to be the original parent of 168 (strain SMY). Interestingly, our sample sequences demonstrated that each of these eight strains is in fact a genomic hybrid, with at least one sequence "island" derived from strain W23 integrated into the 168 genome. The extent of these *trpC* islands is strain specific, with their size ranging from 1.6 to 28.6 kb. There were apparently several independent genetic crosses leading to the legacy strains in our group, but in each case the donor was W23 (or 23) and the recipient was 168 or a strain derived from it.

The W23 sequence islands can be accounted for by the standard model for *B. subtilis* transformation (21). During each strain construction, W23 donor DNA first bound to receptors on the 168 recipient cell surface. The DNA was processed by nuclease cleavage into a few shorter double-stranded fragments that were subsequently converted into single strands. One strand from each fragment was degraded while the other was imported into the cell, where it integrated into the chromosome by displacing a homologous sequence. Integrated fragments of donor DNA have been estimated to be 8.5 to 12 kb in length (22, 29). Each of the smaller islands seen in Fig. 1, then, can be explained by the integration of a single fragment of donor DNA. The largest island (28.6 kb) would have required the recipient to import at least three contiguous fragments processed from a larger DNA molecule at a single receptor. Because there are no large gaps of 168-like sequence in the island, each of these imported fragments must have arisen by chance from the same strand of the donor DNA helix. This type of clustering of independent integration events into a contiguous region of the chromosome has in fact been observed experimentally for *B. subtilis* (22, 29).

Four of the strains—W168, PS832, Mu8u5u1, and PY79—possess a 28.6-kb W23 "island" that stretches from *panB* to *menH*. At least 16.6 kb is contiguous W23-derived sequence; indeed, flanking sequence samples are consistent with the notion that the entire island is derived from W23. Strain W168 (short for "wild-type 168"), an important early strain from the Noboru Sueoka lab at Princeton, is an example of a legacy strain with a well-documented provenance. It was produced by transforming a 168 *trpC his* double auxotroph to prototrophy with W23 DNA (67). Both the *hisC* and *trpC* loci are included within the 28.6-kb island. Strain Mu8u5u1 (Fig. 5G) was derived from W168 following three consecutive cycles of UV mutagenesis that introduced first *leuB8*, then *metB5*, and finally *ilvA1* into the genome (77). Our sequence samples for W168 and Mu8u5u1 are identical except for two SNPs (not shown). Strain Mu8u5u1 and its close relatives provided an entry for unsuspected "contaminating" W23 DNA into the genomes of many other Trp$^+$ legacy strains. In addition to Mu8u5u1, the BGSC maintains 116 strains with a genotype including one or more of the Mu8u5u1 genetic markers in a *trpC*$^+$ background, strongly suggesting that the genomic heritage of these strains can also be traced ultimately to W168. A further 50 strains have at least one of the markers in a *trpC2* background, sug-

gesting that they were produced by a backcross of a W168-derived hybrid strain into a 168 background. If so, then many of these strains would likely contain W23-derived sequences in their genomes as well.

Strain PY79 (Fig. 5H) has been used widely in a variety of *B. subtilis* genetic studies and has in recent years become one of the most commonly used strains in large-scale -omics projects (13, 43, 47, 62). PY79 was isolated when the auxotrophic markers of strain CU1769 (*glnA100* and *metB5*) were removed by two cycles of PBS1-mediated transduction, using lysates grown on strain 168 (80). As discussed above, the *metB5* marker first appeared in a strain with a W168 genetic background during mutant isolation experiments in the Sueoka lab (77). It is likely, then, that CU1769 traces its genomic heritage to W168. Others have already noted that the PY79 *birA* gene—which is located within the 28.6-kb region—shows significant sequence divergence from its 168 ortholog (12). Our sequences make it clear that the PBS1 transductions failed to remove or reduce the 28.6-kb W23 "island" in the PY79 genome. Interestingly, the PY79 and Mu8u5u1 genomes share a SNP within their *panB* coding sequences, just outside the endpoint of the W23 sequence island (not shown). Since this SNP is not found in any of the other strains included in this study, it may represent a relic from the UV mutagenesis experiments of Sueoka. One possible ancestor of PY79 is the Sueoka strain Mu8u5u5 (*leuB8 metB5 thrA5*), since *su*+3, a suppressor strain derived from it, is known to be cured of the SPβ prophage (72), as is PY79 (80).

PS832, a widely used *B. subtilis* strain, was originally acquired by the Setlow laboratory at the University of Connecticut from the Tipper lab at the University of Massachusetts (71). All records describing the origins of PS832 were lost in a mishap during a relocation of the Tipper lab many years ago (D. J. Tipper, personal communication). PS832 has since been distributed to researchers around the world, either directly from the Setlow lab or via the BGSC (as strain 1A757). In publications, the strain has typically been referenced as a "prototrophic derivative" or even a "wild-type, *trp*$^+$ revertant" of strain 168. It has been assumed to be equivalent to and interchangeable with strain 168 in genome-wide mutational screening and functional genomic studies (1, 53). Our studies reveal that PS832 has the same 28.6-kb "island" of W23 DNA as does W168. We have been unable to find any SNPs distinguishing the sequences of these two strains. Like PY79, PS832 has a second W23 island in the *sacA* region (see below). It is highly likely, then, that PS832 is simply a synonym of W168 (Fig. 5I).

One intriguing prototrophic legacy strain is SB491 (Fig. 5I). Like several other such strains, it came to the BGSC from the Lederberg lab at Stanford University. According to the Nester strain notebook archived at the BGSC, SB491 was a "full revertant" of 168, obtained in the summer of 1962 from M. Schafer, a member of the lab. It was deposited into the BGSC collection immediately after 168, receiving the code 1A2. Clearly, several researchers in the *Bacillus* community have come to regard SB491 as a prototrophic equivalent of 168. Although not a nomenclatural type strain, it has been used as a taxonomic reference for *Bacillus subtilis* subsp. *subtilis* in DNA hybridization studies underpinning the descriptions of two novel *Bacillus* species (55, 60). Our sequence analysis reveals that SB491 cannot be a genetic revertant of 168. Instead, it contains a W23 sequence "island" that stretches 8.7 kb, from *trpC* to *menH* (Fig. 1). Interestingly, the *menH* endpoint is identical to that of the 28.6-kb island found in the W168-like strains. One plausible scenario for the construction of SB491, then, would be the transformation of strain 168 with DNA from a W168-like donor, perhaps in a conscious attempt to obtain a Trp$^+$ reference strain with higher isogenicity to 168.

Strain SB19 arose from a very early attempt by the Lederberg laboratory to construct a prototrophic reference *Bacillus subtilis* strain (Fig. 5E). It was a spontaneous streptomycin-resistant mutant obtained from a prototrophic transformant of 168 with DNA from strain 23 (49). Lab notebooks from Lederberg and Nester archived at the BGSC state that the parental prototroph came from Spizizen. Our DNA sequence analysis revealed that SB19 contains a W23 sequence island of 11.9 kb (Fig. 1). Although this island is smaller than that seen in the W168-like strains, it extends several kilobases farther in the direction of the replication terminus, consistent with the documentary evidence that SB19 and W168 arose from independent strain constructions. SB19 has, in turn, served as a parent for further construction of other mutants (23). SB19, then, constitutes another entry point for W23-like DNA into the genomes of other 168-derived legacy strains.

Many of the Trp$^+$ legacy strains in the BGSC collection have been deposited with a careful genotypic description but with a completely undocumented pedigree. One representative of this group of strains is GSY505, another of the earliest accessions in the BGSC collection. GSY505 comes from one of the first *B. subtilis* laboratories, that of Constantine Anagnostopoulos at Gif-sur-Yvette, France, and its strain number suggests that it was constructed in the mid-1960s. As Fig. 1 indicates, GSY505 contains a small, 1.6-kb "island" of W23-like sequence that removed the *trpC2* lesion and restored prototrophy. Closer inspection reveals that this small island is itself a mosaic composed of three W23-like segments interrupted by two short blocks of 168-like sequence. Such a structure could have arisen from a single strand of W23 DNA that was incorporated into the genome with multiple crossovers—at least six within a stretch of only 1,600 bp. Perhaps it is more likely that GSY505 is several steps removed from a first-generation W23–168 genomic hybrid. DNA from this hybrid was used to transform another recipient to tryptophan prototrophy, and the process was repeated. At each step, recombination could have shortened the W23 "island" or replaced segments of it with 168-like sequences. If so, it is possible that many other Trp$^+$ legacy strains with complicated construction histories possess complex mosaic islands in their genomes.

One longstanding theory is that strain SMY ("subtilis Marburg Yale") is in fact the lost parental strain used by Burkholder and Giles. In 1976, Pierre Schaeffer coauthored a paper calling SMY "the Marburg strain of *B. subtilis*" and 168 "its indole-requiring mutant" and noting that both strains had been "maintained in this laboratory for many years" (11). By the early 1970s, Schaeffer had supplied SMY to laboratories in the United States (27, 31, 37). North American researchers took note of subtle differences between SMY and 168 in colony morphology, transformability, and surface stickiness properties but hypothesized that these strain-specific peculiarities could be explained by the mutagenesis-induced changes in 168 (A. L. Sonenshein, personal communication). To test this hypothesis,

we sequenced the *trpC* region from two lab strains of SMY, i.e., BGSC strain 1A757, obtained from Tina Henkin at The Ohio State University, and a stock of SMY stored in the A. L. Sonenshein lab at Tufts University since 1979. Both isolates yielded identical results. SMY contains a contiguous 6.4-kb "island" of W23 DNA in its genome, with endpoints lying within the *trpC* and *cheR* coding sequences (Fig. 1). These data alone cannot distinguish whether SMY arose by a primary transformation of 168 with 23 or W23 DNA or by a retransformation of 168 with DNA from a different primary transformant, such as W168. Regardless, it is obvious that SMY is not the parent of 168, but rather a hybrid of 168 and W23 (Fig. 5I).

Comparison of the 168, NCIB 3610^T, PY79, and SMY genomes. The availability of genome resequencing data for three of the legacy strains fills in several gaps in our understanding of their genomic heritage. Figure 3 compares the sequence polymorphisms that distinguish strains 168, NCIB 3610^T, PY79, and SMY. Any pedigree connecting the strains will have to take the following observations into account. Resequencing detected 22 SNPs that distinguish strain 168 from NCIB 3610^T. At each of the detected positions, SMY and PY79 agree with 168 against NCIB 3610^T, while at no point do they agree with NCIB 3610^T against 168. There is only a single residue where PY79 and 168 agree against the other two strains. There are, however, 14 SNPs where SMY and PY79 agree with each other against both NCIB 3610^T and 168. Finally, there are 67 SNPs that belong uniquely to PY79 and 43 SNPs that belong uniquely to SMY. We can confidently conclude, therefore, that strains 168, SMY, and PY79 all share a common ancestor more recent than NCIB 3610^T; that ancestor would logically be strain 168 itself. We can also infer that strains SMY and PY79 share a common ancestor more recent than 168, since they share far more SNPs than could be explained by random mutations. We do not know the identity of this ancestor, but we do know that both SMY and PY79 are W23–168 hybrid strains and that PY79 was ultimately derived from W168. The last common ancestor of SMY and PY79, then, would have been either W168 or a strain derived from it. A model in which SMY was produced from a cross under congressing conditions between a W168-like donor and a 168-like recipient accounts economically for several observations, including the common SNPs seen in SMY and PY79, the reduced W23 island in the SMY *trpC* region, and the absence of a W23 island in the SMY *sac* region. Finally, it appears that there was deliberate mutagenesis in the pedigrees of both PY79 and SMY after their lineages diverged. PY79 was constructed from a parent bearing *metB5*, a marker isolated from a W168-derived strain after multiple rounds of mutagenesis. We do not know the circumstances under which SMY or a recent ancestor encountered mutagenesis.

Implication of W23 islands in the physiology of *B. subtilis* legacy strains. A question arising from this work is whether the presence of W23 sequences in the genomes of 168 strains results in a significant change in the physiology of the cell. Certainly from a gross phenotype perspective the answer is no. Other than the presence of the *trpC2* allele in 168, which causes tryptophan auxotrophy, no significant differences in growth characteristics and competence are reported between 168 and the Trp$^+$ legacy strains. Yet PY79 contains four large deletions, two W23 islands, and at least 81 other SNPs relative to 168, differences which appear to have arisen during laboratory handling and mutant construction over the years (30, 80). The full impact of these genetic differences on physiology is unknown. Our findings may, however, be relevant to metabolic pathway engineering of *B. subtilis*, where the maximum efficiency of enzyme activities is critically important. We show that for at least one gene, *panB* of the pantothenate biosynthetic pathway, sequence divergence between 168 and PY79 results in enzymes with significant differences in catalytic activity and feedback regulation. Under normal laboratory conditions, this difference in PanB activity should have little impact on cell growth or viability. However, in using these genes in the construction of engineered strains to overproduce pantothenate, a significant difference in the output of the vitamin has been observed. Selecting a host for isolation of specific genes thus requires careful attention to its genomic heritage. In practice, Trp$^+$ derivatives of 168 have been used as hosts for large-scale industrial production of enzymes or other small molecules (56, 57). Such strains clearly contain some or all of the W23 DNA sequences we identified. It remains to be seen whether the cumulative effect of W23 sequences has a significant impact on the performance of these strains compared to 168-derived strains, especially where optimal performance of a strain under stressful growth conditions is required for maximum product formation. Similar questions accompany the use of hybrid legacy strains, such as PY79, in large-scale -omics research. Genome differences of the magnitude observed between 168 and PY79 might well translate into significantly different patterns of global gene regulation and cell physiology. Utilization of 168 Trp$^+$ strains in which only the *trpC2* allele is replaced by the wild-type allele by use of PCR-derived fragments, such as BGSC strain 1A900, could be an option to circumvent these issues.

ACKNOWLEDGMENTS

The BGSC is funded in part by a grant from the National Sciences Foundation (0234214).

We thank the following scientists who patiently answered our queries about the history of *Bacillus subtilis* legacy strains: Mary Case, Donald H. Dean, Maurice S. Fox, Matthew C. Frank, Norman H. Giles, John D. Helmann, Tina M. Henkin, James A. Hoch, Sidney R. Kushner, Joshua Lederberg, Richard Losick, Eugene Nester, David P. Penney, Kathleen Sandman, Peter Setlow, Abraham L. Sonenshein, Donald J. Tipper, Charles Yanofsky, Philip Youngman, and Stanley A. Zahler. We also thank J. Wang for sharing SNP data for *B. subtilis* strains before publication.

REFERENCES

1. **Aertsen, A., I. Van Opstal, S. C. Vanmuysen, E. Y. Wuytack, and C. W. Michiels.** 2005. Screening for *Bacillus subtilis* mutants deficient in pressure induced spore germination: identification of *ykvU* as a novel germination gene. FEMS Microbiol. Lett. **243:**385–391.
2. **Aguilar, C., H. Vlamakis, R. Losick, and R. Kolter.** 2007. Thinking about *Bacillus subtilis* as a multicellular organism. Curr. Opin. Microbiol. **10:**638–643.
3. **Albert, T. J., D. Dailidiene, G. Dailide, J. E. Norton, A. Kalia, T. A. Richmond, M. Molla, J. Singh, R. D. Green, and D. E. Berg.** 2005. Mutation discovery in bacterial genomes: metronidazole resistance in *Helicobacter pylori*. Nat. Methods **2:**951–953.
4. **Albertini, A. M., and A. Galizzi.** 1999. The sequence of the *trp* operon of Bacillus subtilis 168 (*trpC2*) revisited. Microbiology **145:**3319–3320.
5. **Anagnostopoulos, C., and I. P. Crawford.** 1961. Transformation studies on the linkage of markers in the tryptophan pathway in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **47:**378–390.
6. **Anagnostopoulos, C., and J. Spizizen.** 1961. Requirements for transformation in *Bacillus subtilis*. J. Bacteriol. **81:**741–746.
7. **Auchtung, J. M., C. A. Lee, R. E. Monson, A. P. Lehman, and A. D. Grossman.**

2005. Regulation of a *Bacillus subtilis* mobile genetic element by intercellular signaling and the global DNA damage response. Proc. Natl. Acad. Sci. **102:** 12554–12559.
8. **Beadle, G. W., and E. L. Tatum.** 1941. Genetic control of biochemical reactions in *Neurospora*. Proc. Natl. Acad. Sci. USA **27:**499–506.
9. **Belitsky, B. R., and A. L. Sonenshein.** 1998. Role and regulation of *Bacillus subtilis* glutamate dehydrogenase genes. J. Bacteriol. **180:**6298–6305.
10. **Bodmer, W. F.** 1966. Integration of deoxyribonuclease-treated DNA in *Bacillus subtilis* transformation. J. Gen. Physiol. **49:**233–258.
11. **Bohin, J. P., D. Rigomier, and P. Schaeffer.** 1976. Ethanol sensitivity of sporulation in *Bacillus subtilis*: a new tool for the analysis of the sporulation process. J. Bacteriol. **127:**934–940.
12. **Bower, S., J. Perkins, R. R. Yocum, P. Serror, A. Sorokin, P. Rahaim, C. L. Howitt, N. Prasad, S. D. Ehrlich, and J. Pero.** 1995. Cloning and characterization of the *Bacillus subtilis birA* gene encoding a repressor of the biotin operon. J. Bacteriol. **177:**2572–2575.
13. **Britton, R. A., P. Eichenberger, J. E. Gonzalez-Pastor, P. Fawcett, R. Monson, R. Losick, and A. D. Grossman.** 2002. Genome-wide analysis of the stationary-phase sigma factor (sigma-H) regulon of *Bacillus subtilis*. J. Bacteriol. **184:**4881–4890.
14. **Brodetsky, A. M., and W. R. Romig.** 1965. Characterization of *Bacillus subtilis* bacteriophages. J. Bacteriol. **90:**1655–1663.
15. **Burkholder, P. R., and N. H. Giles.** 1947. Induced biochemical mutations in *Bacillus subtilis*. Am. J. Bot. **34:**345–348.
16. **Carlton, B. C.** 1967. Transformation mapping of the genes controlling tryptophan biosynthesis in *Bacillus subtilis*. J. Bacteriol. **94:**660–665.
17. **Chin, T., M. M. Burger, and L. Glaser.** 1966. Synthesis of teichoic acids. VI. The formation of multiple wall polymers in *Bacillus subtilis* W-23. Arch. Biochem. Biophys. **116:**358–367.
18. **Conn, H. J.** 1930. The identity of *Bacillus subtilis*. J. Infect. Dis. **46:**341–350.
19. **Cutting, S. M., and P. B. Vander Horn.** 1990. Genetic analysis, p. 27–74. *In* C. R. Harwood and S. M. Cutting (ed.), Molecular biological methods for *Bacillus*. John Wiley & Sons Ltd., Chichester, United Kingdom.
20. **Doi, R. H., and R. T. Igarashi.** 1964. Ribonucleic acids of *Bacillus subtilis* spores and sporulating cells. J. Bacteriol. **87:**323–328.
21. **Dubnau, D.** 1991. Genetic competence in *Bacillus subtilis*. Microbiol. Rev. **55:**395–424.
22. **Dubnau, D., and C. Cirigliano.** 1972. Fate of transforming DNA following uptake by competent *Bacillus subtilis*. IV. The endwise attachment and uptake of transforming DNA. J. Mol. Biol. **64:**31–46.
23. **Dubnau, D., I. Smith, and J. Marmur.** 1965. Gene conservation in *Bacillus* species. II. The location of genes concerned with the synthesis of ribosomal components and soluble RNA. Proc. Natl. Acad. Sci. USA **54:**724–730.
24. **Ehrlich, J., Q. R. Bartz, R. M. Smith, D. A. Joslyn, and P. R. Burkholder.** 1947. Chloromycetin, a new antibiotic from a soil actinomycete. Science **106:**417.
25. **Errington, J.** 2003. Regulation of endospore formation in *Bacillus subtilis*. Nat. Rev. Microbiol. **1:**117–126.
26. **Fischer, E., and U. Sauer.** 2005. Large-scale in vivo flux analysis shows rigidity and suboptimal performance of *Bacillus subtilis* metabolism. Nat. Genet. **37:**636–640.
27. **Fisher, S., L. Buxbaum, K. Toth, E. Eisenstadt, and S. Silver.** 1973. Regulation of manganese accumulation and exchange in *Bacillus subtilis* W23. J. Bacteriol. **113:**1373–1380.
28. **Fisher, S. H., and A. L. Sonenshein.** 1977. Glutamine-requiring mutants of *Bacillus subtilis*. Biochem. Biophys. Res. Commun. **79:**987–995.
29. **Fornili, S. L., and M. S. Fox.** 1977. Electron microscope visualization of the products of *Bacillus subtilis* transformation. J. Mol. Biol. **113:**181–191.
30. **Georgopoulos, C. P.** 1969. Suppressor system in *Bacillus subtilis* 168. J. Bacteriol. **97:**1397–1402.
31. **Greenleaf, A. L., and R. Losick.** 1973. Appearance of a ribonucleic acid polymerase-binding protein in asporogenous mutants of *Bacillus subtilis*. J. Bacteriol. **116:**290–294.
32. **Harwood, C., and S. M. Cutting (ed.).** 1990. Molecular biological methods for *Bacillus.* John Wiley & Sons, Ltd., Chichester, United Kingdom.
33. **Hecker, M., and U. Volker.** 2004. Towards a comprehensive understanding of *Bacillus subtilis* cell physiology by physiological proteomics. Proteomics **4:**3727–3750.
34. **Hemphill, H. E., and H. R. Whiteley.** 1975. Bacteriophages of *Bacillus subtilis*. Bacteriol. Rev. **39:**257–315.
35. **Herring, C. D., A. Raghunathan, C. Honisch, T. Patel, M. K. Applebee, A. R. Joyce, T. J. Albert, F. R. Blattner, D. van den Bloom, C. R. Cantor, and B. O. Palsson.** 2006. Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. Nat. Genet. **38:**1406–1412.
36. **Higgins, D., J. Thompson, T. Gibson, J. D. Thompson, D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
37. **Imae, Y., and J. L. Strominger.** 1976. Cortex content of asporogenous mutants of *Bacillus subtilis*. J. Bacteriol. **126:**914–918.
38. **Kearns, D. B., F. Chu, R. Rudner, and R. Losick.** 2004. Genes governing swarming in *Bacillus subtilis* and evidence for a phase variation mechanism controlling surface motility. Mol. Microbiol. **52:**357–369.
39. **Kobayashi, K., S. D. Ehrlich, A. Albertini, G. Amati, K. K. Andersen, M. Arnaud, K. Asai, S. Ashikaga, S. Aymerich, P. Bessieres, F. Boland, S. C. Brignell, S. Bron, K. Bunai, J. Chapuis, L. C. Christiansen, A. Danchin, M. Debarbouille, E. Dervyn, E. Deuerling, K. Devine, S. K. Devine, O. Dreesen, J. Errington, S. Fillinger, S. J. Foster, Y. Fujita, A. Galizzi, R. Gardan, C. Eschevins, T. Fukushima, K. Haga, C. R. Harwood, M. Hecker, D. Hosoya, M. Hullo, H. Kakeshita, D. Karamata, Y. Kasahara, F. Kawamura, K. Koga, P. Koski, R. Kuwana, D. Imamura, M. Ishimaru, S. Ishikawa, I. Ishio, C. Le, D. A. Masson, C. Mauel, R. Meima, R. P. Mellado, A. Moir, S. Moriya, E. Nagakawa, H. Nanamiya, S. Nakai, P. Nygaard, M. Ogura, T. Ohanan, M. O'Reilly, M. O'Rourke, Z. Pragai, H. M. Pooley, G. Rapoport, J. P. Rawlins, L. A. Rivas, C. Rivolta, A. Sadaie, Y. Sadaie, M. Sarvas, T. Sato, H. H. Saxild, E. Scanlan, W. Schumann, J. Seegers, J. Sekiguchi, A. Sekowska, S. J. Seror, M. Simon, P. Stragier, R. Studer, H. Takamatsu, T. Tanaka, M. Takeuchi, H. B. Thomaides, V. Vagner, J. M. van Dijl, K. Watabe, A. Wipat, H. Yamamoto, M. Yamamoto, Y. Yamamoto, K. Yamane, K. Yata, K. Yoshida, H. Yoshikawa, U. Zuber, and N. Ogasawara.** 2003. Essential *Bacillus subtilis* genes. Proc. Natl. Acad. Sci. USA **100:**4678–4683.
40. **Kunst, F., N. Ogasawara, I. Moszer, A. Albertini, G. Alloni, V. Azevedo, M. Bertero, P. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S. C. Brignell, S. Bron, S. Brouillet, C. V. Bruschi, B. Caldwell, V. Capuano, N. M. Carter, S. K. Choi, J. J. Codani, I. F. Connerton, A. Danchin, et al.** 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature **390:**249–256.
41. **Kurata, H., Q. Zhao, and K. Shimizu.** 2007. Integration of enzyme activities into metabolic flux distributions by elementary mode analysis. BMC Syst. Biol. **1:**31–44.
42. **Lederberg, J., and E. L. Tatum.** 1946. Gene recombination in *Escherichia coli*. Nature **158:**558.
43. **Lee, J. M., S. H. Zhang, S. Saha, S. S. Anna, C. Jiang, and J. Perkins.** 2001. RNA expression analysis using an antisense *Bacillus subtilis* genome array. J. Bacteriol. **183:**7371–7380.
44. **Loshon, C. A., P. C. Genest, B. Setlow, and P. Setlow.** 1999. Formaldehyde kills spores of *Bacillus subtilis* by DNA damage and small, acid-soluble spore proteins of the α/β-type protect spores against this DNA damage. J. Appl. Microbiol. **87:**8–14.
45. **Manjunatha, U. H., H. Boshoff, C. S. Dowd, L. Zhang, T. J. Albert, J. E. Norton, L. Daniels, T. Dick, S. S. Pang, and C. E. Barry III.** 2006. Identification of a nitroimidazo-oxazine specific protein involved in PA-824 resistance in *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. USA **103:**431–436.
46. **Meima, R. B., J. M. van Dijl, S. Holsappel, and S. Bron.** 2004. Expression systems in *Bacillus*, p. 199–252. *In* F. Baneyx (ed.), Protein expression technologies: current status and future trends. Horizon Bioscience, Wymondham, United Kingdom.
47. **Molle, V., M. Fujita, S. T. Jensen, P. Eichenberger, J. E. González-Pastor, J. S. Liu, and R. Losick.** 2003. The Spo0A regulon of *Bacillus subtilis*. Mol. Microbiol. **50:**1683–1701.
48. **Nakamura, L. K., M. S. Roberts, and F. M. Cohan.** 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.
49. **Nester, E. W., and J. Lederberg.** 1961. Linkage of genetic units of *Bacillus subtilis* in DNA transformation. Proc. Natl. Acad. Sci. USA **47:**52–55.
50. **Nester, E. W., M. Schafer, and J. Lederberg.** 1963. Gene linkage in DNA transfer: a cluster of genes concerned with aromatic biosynthesis in *Bacillus subtilis*. Genetics **48:**529–554.
51. **Oh, Y. K., B. O. Palsson, S. M. Park, C. H. Schilling, and R. Mahadevan.** 2007. Genome-scale reconstruction of metabolic network in *Bacillus subtilis* based on high-throughput phenotyping and gene essentiality data. J. Biol. Chem. **282:**28791–28799.
52. **Okamoto, K., J. A. Mudd, J. Mangan, W. M. Huang, T.-V. Subbaiah, and J. Marmur.** 1968. Properties of the defective phage of *Bacillus subtilis*. J. Mol. Biol. **34:**413–428.
53. **Oomes, S. J. C. M., and S. Brul.** 2006. The effect of metal ions commonly present in food on gene expression of sporulating *Bacillus subtilis* cells in relation to spore wet heat resistance. Innovative Food Sci. Emerg. Technol. **2004:**307–316.
54. **Outtrup, H., and S. T. Jorgensen.** 2002. The importance of *Bacillus* species in the production of industrial enzymes, p. 206–218. *In* R. C. W. Berkley (ed.), Applications and systems of *Bacillus* and relatives. Blackwell Science Inc., Malden, MA.
55. **Palmisano, M. M., L. K. Nakamura, K. E. Duncan, C. A. Istock, and F. M. Cohan.** 2001. *Bacillus sonorensis* sp. nov., a close relative of *Bacillus licheniformis*, isolated from soil in the Sonoran Desert, Arizona. Int. J. Syst. Evol. Microbiol. **51:**1671–1679.
56. **Perkins, J. B., A. Sloma, T. Hermann, K. Theriault, E. Zachgo, T. Erdenberger, and N. Hannett.** 1999. Genetic engineering of *Bacillus subtilis* for the commercial production of riboflavin. J. Ind. Microbiol. Biotechnol. **22:**8–18.
57. **Perkins, J. B., M. Wyss, U. Sauer, and H.-P. Hohmann.** Metabolic engineer-

ing of *Bacillus subtilis. In* C. D. Smolke (ed.), Metabolic engineering handbook, in press. CRC Press, Boca Raton, FL.
58. **Piggot, P. J., and D. W. Hilbert.** 2004. Sporulation of *Bacillus subtilis*. Curr. Opin. Microbiol. **7:**579–586.
59. **Rahman, S. A., and D. Schomburg.** 2006. Observing local and global properties of metabolic pathways: 'load points' and 'choke-points' in the metabolic networks. Bioinformatics **22:**1767–1774.
60. **Roberts, M. S., L. K. Nakamura, and F. M. Cohan.** 1996. *Bacillus vallismortis* sp. nov., a close relative of *Bacillus subtilis*, isolated from soil in Death Valley, California. Int. J. Syst. Bacteriol. **46:**470–475.
61. **Schallmey, M., A. Singh, and O. P. Ward.** 2004. Developments in the use of *Bacillus* species for industrial production. Can. J. Microbiol. **50:**1–17.
62. **Silvaggi, J. M., J. B. Perkins, and R. Losick.** 2006. Genes for small, noncoding RNAs under sporulation control in *Bacillus subtilis*. J. Bacteriol. **188:** 532–541.
63. **Sonenshein, A. L., J. A. Hoch, and R. Losick (ed.).** 2001. *Bacillus subtilis* and its closest relatives: from genes to cells. ASM Press, Washington, DC.
64. **Sonenshein, A. L., J. A. Hoch, and R. Losick (ed.).** 1993. *Bacillus subtilis* and other gram-positive bacteria: biochemistry, physiology, and molecular genetics. American Society for Microbiology, Washington, DC.
65. **Spizizen, J.** 1958. Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. Proc. Natl. Acad. Sci. USA **44:**1072–1078.
66. **Stragier, P., and R. Losick.** 1996. Molecular genetics of sporulation in *Bacillus subtilis*. Annu. Rev. Genet. **30:**297–341.
67. **Sueoka, N., and H. Yoshikawa.** 1963. Regulation of chromosome replication in *Bacillus subtilis*. Cold Spring Harbor Symp. Quant. Biol. **28:**47–54.
68. **Taylor, M. J., and C. B. Thorne.** 1963. Transduction of *Bacillus licheniformis* and *Bacillus subtilis* by each of two phages. J. Bacteriol. **86:**452–461.
69. **Thorne, C. B.** 1962. Transduction in *Bacillus subtilis*. J. Bacteriol. **83:**106–111.
70. **Trowsdale, J., and C. Anagnostopoulos.** 1975. Evidence for the translocation of a chromosome segment in *Bacillus subtilis* strains carrying the *trpE26* mutation. J. Bacteriol. **122:**886–896.
71. **Wahome, P. G., and P. Setlow.** 2007. Growth, osmotic downshock resistance and differentiation of *Bacillus subtilis* strains lacking mechanosensitive channels. Arch. Microbiol. **189:**49–58.
72. **Warner, F. D., G. A. Kitos, M. P. Romano, and H. E. Hemphill.** 1977. Characterization of SPβ: a temperate bacteriophage from *Bacillus subtilis* 168M. Can. J. Microbiol. **23:**45–55.
73. **Widom, R. L., E. D. Jarvis, G. LaFauci, and R. Rudner.** 1988. Instability of rRNA operons in *Bacillus subtilis*. J. Bacteriol. **170:**605–610.
74. **Wolff, S., H. Antelmann, D. Albrecht, D. Becher, J. Bernhardt, S. Bron, K. Buttner, J. M. van Dijl, C. Eymann, A. Otto, L. T. Tam, and M. Hecker.** 2007. Towards the entire proteome of the model bacterium *Bacillus subtilis* by gel-based and gel-free approaches. J. Chromatogr. B **849:**129–140.
75. **Wong, C. W., T. J. Albert, V. B. Vega, J. E. Norton, D. J. Cutler, T. A. Richmond, L. W. Stanton, E. T. Lui, and L. D. Miller.** 2004. Tracking the evolution of the SARS coronavirus using high-throughput, high-density resequencing arrays. Genome Res. **14:**398–405.
76. **Yasbin, R. E., V. C. Maino, and F. E. Young.** 1976. Bacteriophage resistance in *Bacillus subtilis* 168, W23, and interstrain transformants. J. Bacteriol. **125:**1120–1126.
77. **Yoshikawa, H., and N. Sueoka.** 1963. Sequential replication of *Bacillus subtilis* chromosome. I. Comparison of marker frequencies in exponential and stationary growth phases. Proc. Natl. Acad. Sci. USA **49:**559–566.
78. **Yoshikawa, H., and N. Sueoka.** 1963. Sequential replication of the *Bacillus subtilis* chromosome. II. Isotopic transfer experiments. Proc. Natl. Acad. Sci. USA **49:**806–813.
79. **Young, F. E., and J. Spizizen.** 1961. Physiological and genetic factors affecting transformation of *Bacillus subtilis*. J. Bacteriol. **81:**823–829.
80. **Youngman, P., J. Perkins, and R. Losick.** 1984. Construction of a cloning site near one end of Tn917 into which foreign DNA may be inserted without affecting transposition in *Bacillus subtilis* or expression of the transposon-borne erm gene. Plasmid **12:**1–9.
81. **Zeigler, D. R.** 1990. The Bacillus Genetic Stock Center (BGSC), p. 553–556. *In* C. R. Harwood and S. M. Cutting (ed.), Molecular biological methods for *Bacillus*. J. Wiley & Sons, Chichester, United Kingdom.
82. **Zeigler, D. R.** 2000. Bacillus Genetic Stock Center catalog of strains, 7th ed., vol. 1. *Bacillus subtilis* 168. Bacillus Genetic Stock Center, Columbus, OH.