The Bacillus subtilis Minimal Genome Compendium

Michalik, Stephan; Reder, Alexander; Richts, Björn; Faßhauer, Patrick; Mäder, Ulrike; Pedreira, Tiago; Poehlein, Anja; van Heel, Auke J; van Tilburg, Amanda Y; Altenbuchner, Josef

Published in:
ACS Synthetic Biology

DOI:
10.1021/acssynbio.1c00339

ACS SyntheticBiology

pubs.acs.org/synthbio

Technical Note

The *Bacillus subtilis* Minimal Genome Compendium

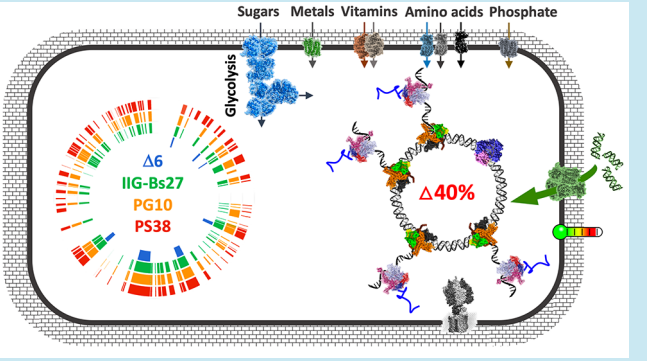
Stephan Michalik,[+] Alexander Reder,[+] Björn Richts,[+] Patrick Faßhauer,[+] Ulrike Mäder, Tiago Pedreira, Anja Poehlein, Auke J. van Heel, Amanda Y. van Tilburg, Josef Altenbuchner, Anika Klewing, Daniel R. Reuß, Rolf Daniel, Fabian M. Commichau, Oscar P. Kuipers, Leendert W. Hamoen, Uwe Völker,* and Jörg Stülke*

Cite This: https://doi.org/10.1021/acssynbio.1c00339

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: To better understand cellular life, it is essential to decipher the contribution of individual components and their interactions. Minimal genomes are an important tool to investigate these interactions. Here, we provide a database of 105 fully annotated genomes of a series of strains with sequential deletion steps of the industrially relevant model bacterium *Bacillus subtilis* starting with the laboratory wild type strain *B. subtilis* 168 and ending with *B. subtilis* PG38, which lacks approximately 40% of the original genome. The annotation is supported by sequencing of key intermediate strains as well as integration of literature knowledge for the annotation of the deletion scars and their potential effects. The strain compendium presented here represents a comprehensive genome library of the entire *MiniBacillus* project. This resource will facilitate the more effective application of the different strains in basic science as well as in biotechnology.



KEYWORDS: *MiniBacillus, genome reduction, Bacillus subtilis, essential genes, strain compendium*

INTRODUCTION

The *MiniBacillus* project was initiated to study fundamental processes of a cell, based on the well-studied model organism *Bacillus subtilis*.[1] To this end, it is critical to simplify the cell as much as possible without compromising its viability in the desired growth medium. The genome reduction can be achieved by bottom-up and top-down approaches.[2] Bottom-up approaches use chemically synthesized genomes to generate an organism with a defined gene set. This strategy has so far been applied to construct the artificial minimal organism *Mycoplasma mycoides* JCVI-syn3.0.[3] In contrast, the top-down approach used in the *MiniBacillus* project utilizes the naturally occurring genome sequence of a well described strain. Genome complexity is reduced by consecutive deletion of genes and genomic regions, which are not considered to have an essential function.[4] Although time-consuming, the major advantage of the top-down approach is its iterative nature, which allows direct causal inference from, for example, dead-end constructs and the affected genes, thus enabling adaptation of the deletion strategy and providing clues to the possible functions of the deleted genes. Thus, the genome-minimized strains cannot only serve as a promising basis for more efficient biotechnological applications but additionally provide valuable information for a deeper understanding of the physiology of the organism under study.

Here, we describe a compendium of 105 strains that have been generated in the *MiniBacillus* project.[5] This collection of strains is a valuable basis for further use in basic research as well as biotechnological utilization for the production of biogenic compounds. For example, the *MiniBacillus* strain PG10 has been proven superior for the production and secretion of difficult-to-express proteins and lantibiotics.[6,7] In order to make this collection of strains available to the scientific community and to provide the required information on the genome reduction steps, we here extend previous publications on the *MiniBacillus* project[5,8−11] and all deletion steps as well as their possible effects and polymorphisms of key strains in detail. This information is deposited as a complete genome compendium to make it available for wide use by the synthetic biology and industrial biotechnology communities.

RESULTS AND DISCUSSION

Resequencing of Key *MiniBacillus* Strains. Important strains in the *MiniBacillus* project such as *B. subtilis* Δ6[8] or *B.*

Received: July 21, 2021

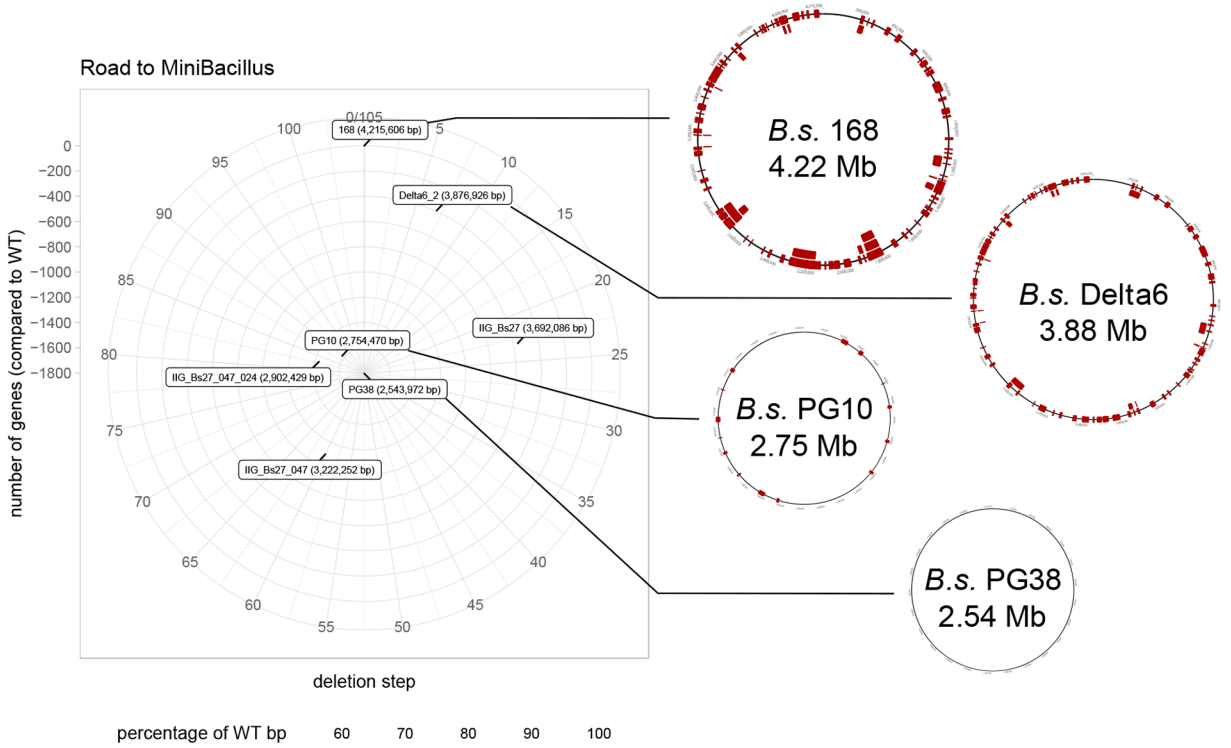


Figure 1. *MiniBacillus* deletion steps are shown linked to the respective genome reduction in percent. Gray dots indicate individual deletion steps, whereas black dots depict key strains. The genome sizes of key strains and the deleted genomic regions (red) are depicted by the circular genome plots.

Table 1. Key Summary of Key Deletion Strains That Where Resequenced with a Hybrid Sequencing Approach

sequence name	deletion step	strain	Mb	genome reduction in %	CDS	rRNA	tRNA	misc_RNA	ncRNA	S-feature
000_Bs168_Refseq	0	168	4.22	0.00	4325	30	86	93	2	1583
007_Delta6_2	7	Delta6_2	3.88	8.03	3941	30	86	86	1	1482
090_PG10	90	PG10	2.75	34.66	2773	27	86	71	0	1101
105_PG38	105	PG38	2.54	39.65	2560	25	80	67	0	1030

subtilis PG10[5] have been sequenced previously using short-read sequencing technology (Illumina) and assembled using a mapping approach against the *B. subtilis* 168 sequence.[12] However, a short-read based genome assembly is error-prone in *de novo* assemblies with respect to genome segments containing repetitive highly homologous regions and often results in incomplete genome assemblies with multiple contigs per genome.[13] Therefore, hybrid whole genome sequencing (WGS) approaches that use a combination of long and short reads are advantageous because genomes can be fully assembled and are better resolved regarding repetitive regions in the assembly process.[14] Here, we used Illumina and Oxford Nanopore (ONT) sequencing for resequencing of *B. subtilis* 168 (Göttingen) and the Δ6 and PG10 strains, as well as for first sequencing of the so far terminal *MiniBacillus* strain *B. subtilis* PG38. Thus, identical techniques based on a combination of *de novo* long-read assembly approaches and high-accuracy short read polishing were used to determine and compare the genome sequences of the *B. subtilis* 168 Gö wild type strain and the three key *MiniBacillus* strains Δ6, PG10, and PG38 (see Supplementary Figure S1). With both approaches of ONT and Illumina sequencing, a depth of more than 100-fold coverage could be achieved, with error rates of 5.21% ± 0.013 for ONT sequencing and 0.625% ± 0.261 for Illumina sequencing (see Supplemental Table S1). The read coverage was almost 100% in all cases, and the percentage of reads mapped to the final complete genome assemblies was 99.73% ± 0.29 for Illumina and 96.53% ± 3.61 for ONT sequencing.

In order to provide a complete breakdown of the deletion steps in the *MiniBacillus* project and to make the individual deletion steps easily comprehensible for the community, the *B. subtilis* 168 sequence (NC_000964.3) was used for the uniform annotation of all genomes. In a first instance, the *B. subtilis* 168 sequence annotation was curated for misassigned CDS stop codons (N = 19) (Supplemental Table S2). In addition, core promoter elements as well as transcribed regions (S_features) defined by the up- and down-shifts detected in a large-scale transcriptome analysis[15] were added to the annotations.

This reference wild-type sequence allows for the assessment of possible artificially introduced transcriptional effects, such as transcriptional couplings or antisense transcripts, and served as the basis for the annotations of all 105 *MiniBacillus* strains.

In total, a genome reduction of 39.65% was achieved from the *B. subtilis* 168 to PG38 strain (see Figure 1, Table 1, Supplemental Table S3). This corresponds to a loss of 1765 coding sequences, 5 rRNAs, 6 tRNAs, 26 misc_RNAs, 2 ncRNAs, and 553 S-features.

https://doi.org/10.1021/acssynbio.1c00339
ACS Synth. Biol. XXXX, XXX, XXX−XXX

Genome Alterations in Addition to the Introduced Deletions. As expected, the genome sequences of the strains revealed the absence of all deleted regions in the strains. However, we observed additional alterations of the coding genome that were caused by spontaneous events or by the assembly of the remaining genome regions. These genomic alterations are briefly mentioned below and are detailed in the Supporting Information S1.

In PG10 and PG38, we observed additional deletions of 4.8 and 5.8 kb, respectively. In both cases, the deletions affect a genomic region that contains the *rrnI*, *rrnH*, and *rrnG* rRNA operons. These operons are virtually identical to each other and are therefore a hotspot for genomic rearrangements (both deletions and amplifications). In PG10, the central *rrnH* operon is spontaneously deleted, and in PG38 a stretch of tRNA genes as well as the *rrnH* and *rrnG* operons are missing. The loss of these rRNA operons may also contribute to the slower growth rate of PG10 and PG38 as compared to the Δ6 strain.[5]

The top-down approach applied in the *MiniBacillus* project resulted in the accumulation of deletion scars. These scars may generate new expression signals or put genes under the control of noncognate expression signals; moreover, the scars may result in the emergence of new coding sequences as well as in the truncation of original protein-coding genes.

Further modifications of the deletion strains were introduced on purpose such as the repair of the tryptophan auxotrophy and the integration of an inducible *comKS* cassette[16] and the *slrR* and *flgM* genes[5,17] to prevent loss of genetic competence and cell separation, respectively. Finally, some sequence polymorphisms emerged during the construction process. These unintended changes are important and should be considered for future optimization and omics data analysis of the individual strains.

Availability of Genome-Reduced Strains and Genome Information. The fully annotated Geneious files for the genomes of each deletion step can be accessed *via* a download option on *Subti*Wiki[18] (http://www.subtiwiki.uni-goettingen.de/v4/minibacillus). All strains are available from L.W.H.

CONCLUSIONS

With this work, we provide a collection of more than 100 *B. subtilis* strains that have undergone consecutive genome reduction. Moreover, we provide the fully annotated, highly detailed genome history of the sequential deletion steps generated in the *MiniBacillus* project. The polymorphisms that were detected in this study will certainly be of relevance when working with the genome-reduced strains. The usefulness of such strains has been proven by the successful application of the *MiniBacillus* strain GP10 for the production and secretion of *Staphylococcus aureus* antigens that are difficult to express in other strains as well as of lantibiotics. In both applications, strain PG10 was found to be superior as compared to other expression hosts. With the complete library of *MiniBacillus* strains available, it will be possible to select the strain that is best suited for the planned application. All genome data can be freely downloaded on *Subti*Wiki and are thus available to the broad scientific and biotechnological communities.

MATERIALS AND METHODS

Detailed protocols for the extraction of genomic DNA, DNA sequencing, and genome assembly as well as for the mapping of genomic features and the analysis of altered sequence features are provided in Supporting Information S2.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.1c00339.

Supplementary Figure S1, schematic overview of the sequencing of the genome-reduced *B. subtilis* strains and the annotation of deletions; Supplementary Figure S2, instable rRNA/tRNA cluster revealed by hybrid sequencing approach; Supplementary Figure S3, schematic overview of deletion scar events; Supplementary Figure S4, selected sequence exchange/insertion; Supplementary Figure S5, location of detected polymorphisms (*B. subtilis* Δ6, PG10, PG38) mapped on the genome of *B. subtilis* 168; Supporting Information S1, genomic alterations in the genome-reduced strains; and Supporting Information S2, detailed protocols for the extraction of genomic DNA, DNA sequencing, and genome assembly as well as for the mapping of genomic features (PDF)

Supplemental Table S7, SNPs of Δ6, PG10, PG38 (XLSX)

AUTHOR INFORMATION

Corresponding Authors

Uwe Völker − *C_FunGene, Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, 17487 Greifswald, Germany*; Email: voelker@uni-greifswald.de

Jörg Stülke − *Institute of Microbiology and Genetics, GZMB, Georg-August University Göttingen, 37077 Göttingen, Germany*; orcid.org/0000-0001-5881-5390; Email: jstuelk@gwdg.de

Authors

Stephan Michalik − *C_FunGene, Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, 17487 Greifswald, Germany*; orcid.org/0000-0002-6550-0132

Alexander Reder − *C_FunGene, Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, 17487 Greifswald, Germany*

Björn Richts − *Institute of Microbiology and Genetics, GZMB, Georg-August University Göttingen, 37077 Göttingen, Germany*

Patrick Faßhauer − *Institute of Microbiology and Genetics, GZMB, Georg-August University Göttingen, 37077 Göttingen, Germany*

Ulrike Mäder − *C_FunGene, Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, 17487 Greifswald, Germany*

Tiago Pedreira − *Institute of Microbiology and Genetics, GZMB, Georg-August University Göttingen, 37077 Göttingen, Germany*

https://doi.org/10.1021/acssynbio.1c00339
ACS Synth. Biol. XXXX, XXX, XXX−XXX

Anja Poehlein − *Institute of Microbiology and Genetics, GZMB, Georg-August University Göttingen, 37077 Göttingen, Germany*
Auke J. van Heel − *Department of Molecular Genetics, University of Groningen, 9747 AG Groningen, The Netherlands*
Amanda Y. van Tilburg − *Department of Molecular Genetics, University of Groningen, 9747 AG Groningen, The Netherlands*
Josef Altenbuchner − *Institute for Industrial Genetics, University of Stuttgart, 70569 Stuttgart, Germany*
Anika Klewing − *Institute of Microbiology and Genetics, GZMB, Georg-August University Göttingen, 37077 Göttingen, Germany*
Daniel R. Reuß − *Institute of Microbiology and Genetics, GZMB, Georg-August University Göttingen, 37077 Göttingen, Germany*
Rolf Daniel − *Institute of Microbiology and Genetics, GZMB, Georg-August University Göttingen, 37077 Göttingen, Germany*
Fabian M. Commichau − *Institute of Microbiology and Genetics, GZMB, Georg-August University Göttingen, 37077 Göttingen, Germany; FG Synthetic Microbiology, Brandenburg University of Technology, 01958 Senftenberg, Germany*
Oscar P. Kuipers − *Department of Molecular Genetics, University of Groningen, 9747 AG Groningen, The Netherlands;* orcid.org/0000-0001-5596-7735
Leendert W. Hamoen − *Swammerdam Institute for Life Sciences, University of Amsterdam, 1090 GE Amsterdam, The Netherlands;* orcid.org/0000-0001-9251-1403

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.1c00339

Author Contributions

D.R.R., J.A., F.M.C., O.P.K., L.W.H., U.V., and J.S. designed the experiments. S.M., A.R., B.R., P.F., U.M., A.P., A.J.v.H., A.Y.v.T., J.A., A.K., D.R.R., and R.D. performed and analyzed the experiments. T.P. set up the database. S.M., A.R, U.V., and J.S. wrote the manuscript. S.M. prepared the figures and tables. All authors contributed to the interpretation of results and manuscript editing. All authors approved the final version of the manuscript.

Author Contributions

[+]S.M., A.R., B.R., and P.F. contributed equally to this work.

Funding

This work was supported by a grant from the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 720776).

Notes

The authors declare no competing financial interest.
Genome sequences of the genome-reduced strains are available in *Subti*Wiki (http://www.subtiwiki.uni-goettingen.de/v4/minibacillus). All strains are available from L.W.H. Any other data are available from the corresponding authors upon request.

■ ACKNOWLEDGMENTS

We are grateful to Daniel López for organizing the European Rafts for Biotech network.

■ ABBREVIATIONS

gDNA, genomic DNA; CDS, coding sequence sequence; ONT, Oxford Nanopore Technologies; orf, open reading frame; refseq, reference sequence; SNP, single nucleotide polymorphism

■ REFERENCES

(1) Errington, J.; van der Aart, L. T. Microbe profile: *Bacillus subtilis*: model organism for cellular development, and industrial workhorse. *Microbiology* **2020**, *166*, 425−427.
(2) Juhas, M.; Reuß, D. R.; Zhu, B.; Commichau, F. M. *Bacillus subtilis* and *Escherichia coli* essential genes and minimal cell factories after one decade of genome engineering. *Microbiology* **2014**, *160*, 2341−2351.
(3) Hutchison, C. A., 3rd; Chuang, R. Y.; Noskov, V. N.; Assad-Garcia, N.; Deerinck, T. J.; Ellisman, M. H.; Gill, J.; Kannan, K.; Karas, B. J.; Ma, L.; Pelletier, J. F.; Qi, Z. Q.; Richter, R. A.; Strychalski, E. A.; Sun, L.; Suzuki, Y.; Tsvetanova, B.; Wise, K. S.; Smith, H. O.; Glass, J. I.; Merryman, C.; Gibson, D. G.; Venter, J. C. Design and synthesis of a minimal bacterial genome. *Science* **2016**, *351*, No. aad6253.
(4) Reuß, D. R.; Commichau, F. M.; Gundlach, J.; Zhu, B.; Stülke, J. The blueprint of a minimal cell: *MiniBacillus*. *Microbiol. Mol. Biol. Rev.* **2016**, *80*, 955−987.
(5) Reuß, D. R.; Altenbuchner, J.; Mäder, U.; Rath, H.; Ischebeck, T.; Sappa, P. K.; Thürmer, A.; Guérin, C.; Nicolas, P.; Steil, L.; Zhu, B.; Feussner, I.; Klumpp, S.; Daniel, R.; Commichau, F. M.; Völker, U.; Stülke, J. Large-scale reduction of the *Bacillus subtilis* genome: Consequences for the transcriptional network, resource allocation, and metabolism. *Genome Res.* **2017**, *27*, 289−299.
(6) van Tilburg, A. Y.; van Heel, A. J.; Stulke, J.; de Kok, N. A. W.; Rueff, A.-S.; Kuipers, O. P. *MiniBacillus* PG10 as a convenient and effective production host for lantibiotics. *ACS Synth. Biol.* **2020**, *9*, 1833−1842.
(7) Aguilar Suarez, R.; Stulke, J.; van Dijl, J. M. Less is more: Toward a genome-reduced *Bacillus* cell factory for "difficult proteins. *ACS Synth. Biol.* **2019**, *8*, 99−108.
(8) Reuß, D. R.; Thürmer, A.; Daniel, R.; Quax, W. J.; Stülke, J. Complete genome sequence of *Bacillus subtilis* subsp. subtilis strain −6. *Genome Announc.* **2016**, *4*, No. e00759-16.
(9) Tanaka, K.; Henry, C. S.; Zinner, J. F.; Jolivet, E.; Cohoon, M. P.; Xia, F.; Bidnenko, V.; Ehrlich, S. D.; Stevens, R. L.; Noirot, P. Building the repertoire of dispensable chromosome regions in *Bacillus subtilis* entails major refinement of cognate large-scale metabolic model. *Nucleic Acids Res.* **2013**, *41*, 687−699.
(10) Westers, H.; Dorenbos, R.; van Dijl, J. M.; Kabel, J.; Flanagan, T.; Devine, K. M.; Jude, F.; Séror, S. J.; Beekman, A. C.; Darmon, E.; Eschevins, C.; de Jong, A.; Bron, S.; Kuipers, O. P.; Albertini, A. M.; Antelmann, H.; Hecker, M.; Zamboni, N.; Sauer, U.; Bruand, C.; Ehrlich, D. S.; Alonso, J. C.; Salas, M.; Quax, W. J. Genome engineering reveals large dispensable regions in *Bacillus subtilis*. *Mol. Biol. Evol.* **2003**, *20*, 2076−2090.
(11) Commichau, F. M.; Pietack, N.; Stülke, J. Essential genes in *Bacillus subtilis*: A re-evaluation after ten years. *Mol. BioSyst.* **2013**, *9*, 1068−1075.
(12) Barbe, V.; Cruveiller, S.; Kunst, F.; Lenoble, P.; Meurice, G.; Sekowska, A.; Vallenet, D.; Wang, T.; Moszer, I.; Médigue, C.; Danchin, A. From a consortium sequence to a unified sequence: The *Bacillus subtilis* 168 reference genome a decade later. *Microbiology* **2009**, *155*, 1758−1775.
(13) Loman, N. J.; Misra, R. V.; Dallman, T. J.; Constantinidou, C.; Gharbia, S. E.; Wain, J.; Pallen, M. J. Performance Comparison of Benchtop High-Throughput Sequencing Platforms. *Nat. Biotechnol.* **2012**, *30* (5), 434−439.
(14) Wick, R. R.; Judd, L. M.; Gorrie, C. L.; Holt, K. E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **2017**, *13*, No. e1005595.

(15) Nicolas, P.; Mader, U.; Dervyn, E.; Rochat, T.; Leduc, A.; Pigeonneau, N.; Bidnenko, E.; Marchadier, E.; Hoebeke, M.; Aymerich, S.; Becher, D.; Bisicchia, P.; Botella, E.; Delumeau, O.; Doherty, G.; Denham, E. L.; Fogg, M. J.; Fromion, V.; Goelzer, A.; Hansen, A.; Hartig, E.; Harwood, C. R.; Homuth, G.; Jarmer, H.; Jules, M.; Klipp, E.; Le Chat, L.; Lecointe, F.; Lewis, P.; Liebermeister, W.; March, A.; Mars, R. A. T.; Nannapaneni, P.; Noone, D.; Pohl, S.; Rinn, B.; Rugheimer, F.; Sappa, P. K.; Samson, F.; Schaffer, M.; Schwikowski, B.; Steil, L.; Stulke, J.; Wiegert, T.; Devine, K. M.; Wilkinson, A. J.; Maarten van Dijl, J.; Hecker, M.; Volker, U.; Bessieres, P.; Noirot, P. Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science* **2012**, *335*, 1103−1106.

(16) Rahmer, R.; Morabbi Heravi, K.; Altenbuchner, J. Construction of a super-competent *Bacillus subtilis* 168 using the P_{mtlA}-*comKS* inducible cassette. *Front. Microbiol.* **2015**, *6*, 1431.

(17) van Tilburg, A. Y.; Fülleborn, J. A.; Reder, A.; Völker, U.; Stülke, J.; van Heel, A. J.; Kuipers, O. P. Unchaining mini*Bacillus* PG10: relief of FlgM-mediated repression of autolysin genes. *Appl. Environ. Microbiol.* **2021**, *87*, No. e01123-21.

(18) Michna, R. H.; Commichau, F. M.; Tödter, D.; Zschiedrich, C. P.; Stülke, J. *Subti*Wiki - a database for the model organism *Bacillus subtilis* that links pathway, interaction and expression information. *Nucleic Acids Res.* **2014**, *42*, D692−D698.

https://doi.org/10.1021/acssynbio.1c00339
ACS Synth. Biol. XXXX, XXX, XXX−XXX