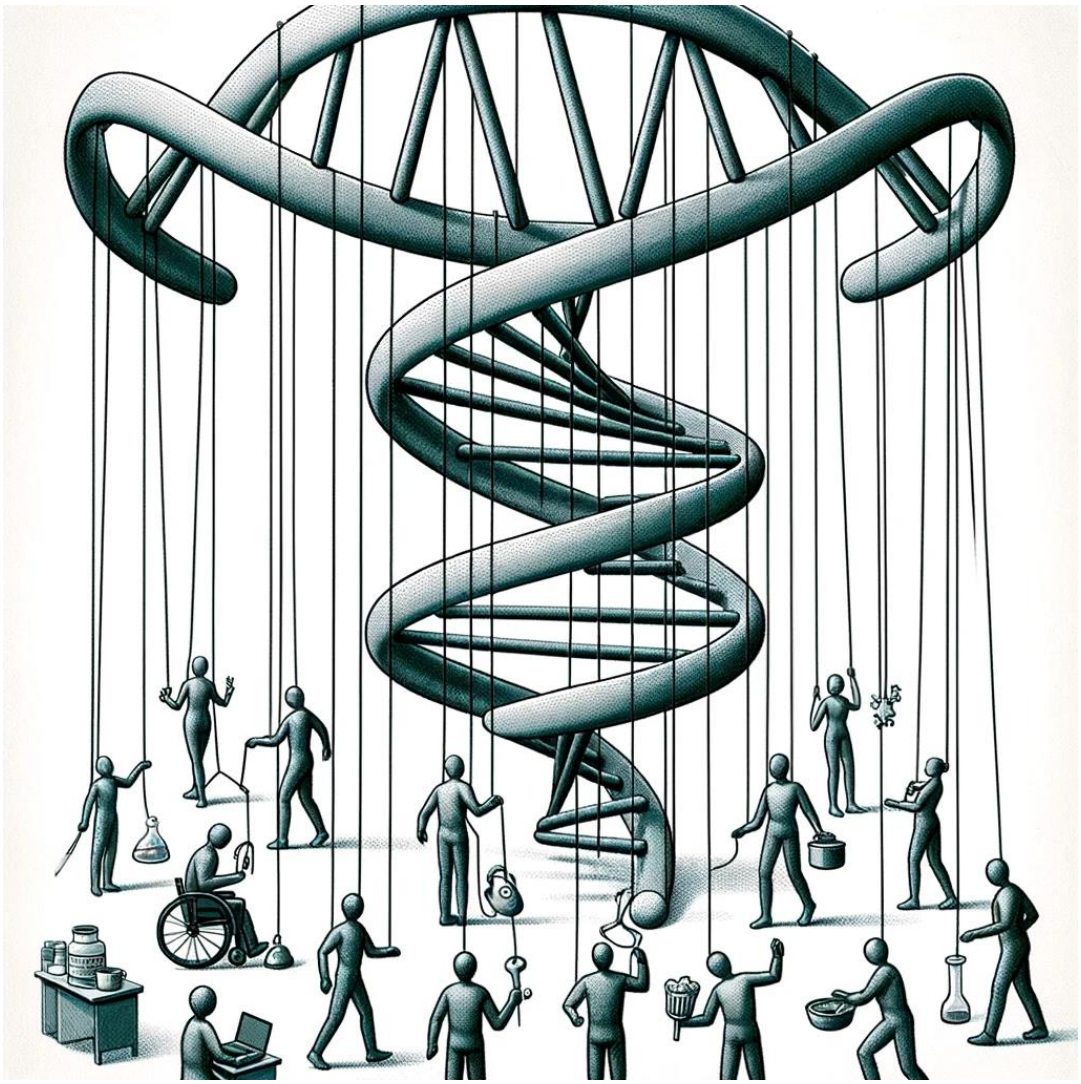
## Characterization of unknown Bacillus subtilis genes with transposon sequencing

Teng, Z.

**Publication date**
2024

# Chapter 1

General introduction



# General introduction

Zihao Teng[1],

Leendert Hamoen[1]

[1]Swammerdam Institute for Life Sciences, University of Amsterdam, The Netherlands

**ABSTRACT**

To understand what is necessary for a bacterial cell to live, and to know what is necessary to build new bacterial cells, an understanding of the composition of the minimal working genome is crucial. A first step is the identification of the core genome of species that lend themselves as useful blueprints for a minimal synthetic bacterial cell. With this knowledge it is then possible to build a minimal genome. There are two paths leading towards this goal, a top-down reduction of existing genomes or a bottom-up construction using DNA synthesis and assembly of a synthetic genome. Unfortunately, despite decades of research the knowledge of what is required to build a minimal bacterium is still limited, even in case of the best studied model bacteria *Escherichia coli* and *Bacillus subtilis*. Transposon sequencing (Tn-seq) provides a powerful tool to pinpoint the importance of uncharacterized genes for bacterial growth, and is a promising technique to fill in the missing knowledge of uncharacterized genes that are necessary for bacterial life and for the building of a useful synthetic bacterial cell.

**Core and pangenome**

The advent of high-throughput DNA sequencing technologies has catalyzed a paradigm shift in bacterial research, offering unprecedented opportunities to explore genetic diversity and evolutionary dynamics within microbial species. For genome comparison, two concepts are distinguished, the core genome and the pangenome (Yuvaraj *et al.*, 2017). The core genome relates to the set of genes shared by all individuals within a bacterial species, (Medini *et al.*, 2005, Ciccarelli *et al.*, 2006), and these genes are essential for the basic cellular functions and play a pivotal role in the fundamental traits and physiological characteristics of the species (Tettelin *et al.*, 2008, Grazziotin *et al.*, 2015). The addition of new genomes to the family does not change the size of the core genome.

In addition to the core genome, the rest of the genome encompasses genes that exhibit variability across different strains within the same species, reflecting their adaptation to diverse ecological niches and environmental challenges (Darmon & Leach, 2014). The pangenome encompasses both the core genome and the accessory genome, thus the complete gene repertoire within an entire bacterial species (Brynildsrud *et al.*, 2016). Generally, with the addition of new genomes the pangenome increases in size (Ku *et al.*, 2015).

The pangenome shows the genomic plasticity, and reflects the capacity of horizontal gene transfer, which facilitates the acquisition of new genetic material (Soucy *et al.*, 2015). In

epidemiological investigations, core genomes serve as valuable tools for tracing the origin and transmission routes of bacterial infections, while pangenome analysis aids in identifying genes associated with virulence, antibiotic resistance, and other adaptive traits (Yang *et al.*, 2023a). Moreover, the pangenome concept has played a pivotal role in advancing novel biotechnological applications, such as the identification of novel enzymes, potential targets for drug discovery and vaccine development (Ismail *et al.*, 2022, Swetha *et al.*, 2022, Zhu *et al.*, 2020).

**Core genome sizes**

A key question in biology is what is necessary for life. Since bacteria comprise the simplest life forms this question can be condensed to what genes are necessary for a bacterial cell to live. The core genome can provide answers to this question. The genomes of the Gram-negative and -positive model bacteria *E. coli* and *B. subtilis* typically harbor around 4500 genes. However, their core genomes encompass approximately 990 and 810 genes, respectively (Lukjancenko *et al.*, 2010, Alcaraz *et al.*, 2010). If we focus at simple bacterial species with small genomes, such as the cell wall-less *Mycoplasma* species, we find for *e.g. Mycoplasma genitalium* a core genome of about 150 genes, of which almost a third code for ribosomal proteins. There are bacterial species with even less genes, such as *Buchnera* species, which are endosymbiont of aphids, and contain genomes with fewer than 200 genes (McCutcheon & Moran, 2011, Wernegreen, 2002, Shigenobu *et al.*, 2000). However, these obligate endosymbionts cannot grow independently outside their hosts, and they are therefore less useful for defining the core set of genes necessary for a functional minimal bacterial cell.

**Essential genes**

The core genes are not necessarily the genes essential for growth. For example, a systematic gene inactivation study in *E. coli* revealed that only 248 genes are essential (Goodall *et al.*, 2018), and a comparable study in *B. subtilis* identified 271 essential genes (Kobayashi *et al.*, 2003). In later CRISPRi and gene deletion studies this number was reduced to 257 genes (Koo *et al.*, 2017, Peters *et al.*, 2016a). Most of these genes are involved in critical cellular processes, including protein synthesis, metabolism, cell envelope and division, DNA replication and maintenance, and RNA synthesis. This set of 257 essential genes was based on growth at 37 °C in nutrient rich Lysogeny broth (LB) medium (Kobayashi *et al.*, 2003, Koo *et al.*, 2017), therefore many genes involved in amino acid synthesis and other important building blocks are missing in this list. Some of the essential genes are unique to *B. subtilis*, such as *yqcF*, which encodes an antitoxin protein found exclusively in *Bacillaceae* (Holberger

*et al.*, 2012), or the PBSX prophage repressor *xre* (Koo *et al.*, 2017). Importantly, the absence of a gene from this set of 257 essential genes does not necessarily imply that it is non-essential, since there are many redundant genes with overlapping functions. For example, there are two proteins that anchor the key cell division protein FtsZ to the cell membrane, FtsA and SepF, both of which are non-essential, whereas they cannot be simultaneously removed without compromising cell viability (Duman *et al.*, 2013).

Currently, there are approximately 1800 genes in the *B. subtilis* genome encoding proteins of unknown function, the so called y-genes. Among these, three genes have been designated as essential: *ylaN*, *yneF*, and *yqeG* (Pedreira *et al.*, 2022b). The *ylaN* gene encodes a constitutively expressed cytosolic protein that has been shown to interact with Fur, an iron homeostasis regulator, suggesting its role as an antagonist of Fur (de Jong *et al.*, 2021, Pi & Helmann, 2018, Steingard & Helmann, 2023). The YneF protein has been linked to sulfur metabolism and potential tRNA modification (Engelen *et al.*, 2012), and its conserved diproline residue is predicted to be involved in acylating long-chain fatty acids relevant to membrane proteins (Danchin & Fang, 2016). The *yqeG* gene encodes a protein with suggestive similarities to NagD (16 %), belonging to the haloacid dehalogenase superfamily (Plumbridge, 1989). It has been demonstrated that the *yqeG* gene is necessary for normal growth on solid medium and can be induced by oxidative stress (Terakawa *et al.*, 2016).

**Top-down genome reduction**

The easiest way to genome reduction is the top-down approach, whereby unnecessary genes are removed from a known genome. This approach has been undertaken with both *E. coli* and *B. subtills*. One of the first attempts reduced the genome of *E. coli* by 30 %. However, this strain, called Δ16, grows much slower than the wild type mother strain, resulting in an increase in doubling time from 26 to 45 min (Hashimoto *et al.*, 2005), despite the fact that no known important genes were removed. A later genome reduction attempt resulted in *E. coli* strain DGF-298 with a 35 % reduced genome. This strain showed no auxotrophy and demonstrated better growth fitness in both rich and minimal media compared to the wild type strain. (Hirokawa *et al.*, 2013, Mizoguchi *et al.*, 2007). Transcriptome analyses revealed the down-regulation of chaperone and protease-encoding genes, a phenotype that was linked to the enhanced fitness of this minimal *E. coli*.

One of the first top down approaches with *B. subtilis* was the removal of all prophages and a large polyketide synthesis gene cluster, resulting in strain Δ6 with a 8 % reduced genome (Westers

*et al.*, 2003). Later this reduction was extended, resulting in the minimal genome strain MBG874 that lacks 20 % of the original genome. This strain exhibited significantly enhanced capabilities in the secretion of heterologous enzymes compared to the wild-type strain, including alkaline cellulase and proteases, demonstrating the potential industrial value of minimal strains (Morimoto *et al.*, 2008, Manabe *et al.*, 2011, Manabe *et al.*, 2013). The Δ6 strain formed the basis for another genome reduction project, culminating in strain IIG-Bs20-4 with a 14 % reduced genome (Wenzel & Altenbuchner, 2015). This strain demonstrated nearly identical growth rates to the wild type *B. subtilis* 168 mother strain in both rich and minimal media. Further deletions were hampered by reduced natural genetic competence, which was used to genetically delete genes. The reason for this is the fact that activation of genetic competence is a carefully regulated processes that includes medium, growth phase and quorum sensing dependent control mechanisms, and the accumulation of genetic deletions can easily hamper the optimal regulation of genetic competence. To overcome this problem, expression of the competence transcription factor encoding gene *comK* was placed under control of a mannitol-inducible promoter (Rahmer *et al.*, 2015b, Ara *et al.*, 2007). Subsequent genome deletions using this artificial inducible genetic competence system lead to the minimal genome strains PG10 and PS38 that each lack 36 % of the genome (Aguilar Suarez *et al.*, 2019, Reuss *et al.*, 2017). Phenotypic analyses revealed that these strains had longer cell lengths, grew slower than the wild type strain, and were no longer able to grow on minimal medium. mRNA expression levels were reduced due to changes in the transcriptional network, and central metabolic pathways were also affected to varying degrees. Nevertheless, for PG10 it was shown that the production of certain heterologous proteins was significantly improved compare to the wild type mother strain (Aguilar Suarez *et al.*, 2019). Interestingly, the function of approximately 22 % of genes of PG10 and PS38 are still unknown.

**Bottom-up construction of genomes**

In the first successful bottom-up genome construction approach, the 1.08 Mbp genome of *Mycoplasma mycoides* was synthesized and completely assembled in the yeast host *Saccharomyces cerevisiae*. Subsequently, this genome was transplanted to a *Mycoplasma capricolum* cell, resulting in the synthetic *Mycoplasma mycoides* strain JCVI-syn1.0 (Gibson *et al.*, 2010). This remarkable feat showed that constructing and kickstarting artificial genomes is possible, and subsequent rounds of in silico reduction steps and genome synthesis and transplantation resulted in *M. mycoides* JCVI-syn3.0 with only 473 genes, half the size of the original genome. Despite this significant reduction,

JCVI-syn3.0 boasts a respectable doubling time of 180 minutes, approximately three times longer than the wild-type Mycoplasma mycoides, but five times faster than the smallest naturally occurring mycoplasma, *M. genitalium*, which contains 525 genes (Hutchison *et al.*, 2016, Jewett & Forster, 2010, Juhas *et al.*, 2012). It appears that JCVI-syn3.0 is very close to achieving the goal of a minimal bacterial genome. However, the function of 149 of its genes is still unknown and homologous are only present in other *Mycoplasma*'s, and therefore, we still lack a full understanding of what is required for bacterial growth. This is also illustrated by attempts to create a functional synthetic minimal *E. coli* genome, MGE-syn1.0, with 449 genes comprising essential and 267 important genes, (Zhou *et al.*, 2016, Hutchison *et al.*, 2016, Jewett & Forster, 2010, Juhas *et al.*, 2012). Although this minimal genome was successfully assembled in *S. cerevisiae*, it has not been possible to replace the wild type *E. coli* genome by this synthetic construct, underlining the fact that crucial information about bacterial life is still lacking.

**Transposon insertion sequencing**

In the process of genome minimization, a common problem is the observation that genes which were deemed non-essential become important for optimal growth after a genome reduction step. As a result, further gene deletions generally result in a slowly but steadily reduction in the growth rate (Reuss *et al.*, 2017, Zhou *et al.*, 2016). Therefore, it is important to have a real-time understanding of the fitness of individual genes during the genome reduction process. Transposon insertion sequencing (Tn-seq) is a powerful molecular genetic technique that offers this option (van Opijnen *et al.*, 2009). This method uses the random insertion of transposons in genes, followed by growth selection and genome-wide mapping of transposon insertion site by next generation sequencing, to determine the importance of genes under a particular growth and/or genetic condition (Poulsen *et al.*, 2022, Matern *et al.*, 2020, Klein *et al.*, 2015). In this thesis, we used Tn-seq to identify novel essential genes in the genome reduced *B. subtilis* strain PG10.

## Outline of this thesis

**Chapter 1** provides a short general review of genome minimization attempts in bacteria. In **Chapter 2** the results of a Tn-seq experiments with the minimal *B. subtilis* strain PG10 is described. The Tn-seq analysis yielded 133 new essential genes of which 67 of unknown function, including the conserved genes *ytiB*, *ywnA and ywgA*, which are further analyzed in subsequent chapters. **Chapter 3** describes the characterization of *ytiB*, and shows that the deletion of this gene activates the

mannitol promoter. In PG10, this results in the artificial and lethal induction of genetic competence. **Chapter 4** investigates the function of *ywnA*. Amino acid sequence comparison revealed the presence of a helix-turn-helix motif found in the Rrf2 transcription regulator family. Transcriptome analysis showed that deletion of *ywnA* results in the overexpression of the downstream located gene *ywnB*, which turned out to be lethal in PG10. **Chapter 5** describes the characterization of *ywgA* and its neighboring gene *ywfO*. We could show that YwgA is essential in PG10 because it controls the activity off YwfO, which we show is a deoxynucleotide triphosphate triphosphohydrolase. Interestingly, microscopic analyses revealed that YwfO exhibits characteristics of liquid-liquid phase separation. In the final experimental chapter, **Chapter 6**, we describe a finding that we encountered during the transcriptome studies of several deletion mutants, and that is the unexpected effect on transcription by the presence of the erythromycin resistance marker, which is commonly used in *B. subtilis* deletion and transcriptome studies (Pietiainen *et al.*, 2009, Jain *et al.*, 2019).

**REFERENCE**

Aguilar Suarez, R., Stulke, J., and van Dijl, J.M. (2019) Less Is More: Toward a Genome-Reduced Bacillus Cell Factory for "Difficult Proteins". *ACS Synth Biol* **8**: 99-108.

Alcaraz, L.D., Moreno-Hagelsieb, G., Eguiarte, L.E., Souza, V., Herrera-Estrella, L., and Olmedo, G. (2010) Understanding the evolutionary relationships and major traits of Bacillus through comparative genomics. *Bmc Genomics* **11**: 332.

Ara, K., Ozaki, K., Nakamura, K., Yamane, K., Sekiguchi, J., and Ogasawara, N. (2007) Bacillus minimum genome factory: effective utilization of microbial genome information. *Biotechnol Appl Biochem* **46**: 169-178.

Brynildsrud, O., Bohlin, J., Scheffer, L., and Eldholm, V. (2016) Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome Biol* **17**: 238.

Ciccarelli, F.D., Doerks, T., von Mering, C., Creevey, C.J., Snel, B., and Bork, P. (2006) Toward automatic reconstruction of a highly resolved tree of life. *Science* **311**: 1283-1287.

Danchin, A., and Fang, G. (2016) Unknown unknowns: essential genes in quest for function. *Microb Biotechnol* **9**: 530-540.

Darmon, E., and Leach, D.R. (2014) Bacterial genome instability. *Microbiol Mol Biol Rev* **78**: 1-39.

de Jong, L., Roseboom, W., and Kramer, G. (2021) A composite filter for low FDR of protein-protein interactions detected by in vivo cross-linking. *J Proteomics* **230**: 103987.

Duman, R., Ishikawa, S., Celik, I., Strahl, H., Ogasawara, N., Troc, P., Lowe, J., and Hamoen, L.W. (2013) Structural and genetic analyses reveal the protein SepF as a new membrane anchor for the Z ring. *Proc Natl Acad Sci U S A* **110**: E4601-4610.

Engelen, S., Vallenet, D., Medigue, C., and Danchin, A. (2012) Distinct co-evolution patterns of genes associated to DNA polymerase III DnaE and PolC. *Bmc Genomics* **13**: 69.

Gibson, D.G., Glass, J.I., Lartigue, C., Noskov, V.N., Chuang, R.Y., Algire, M.A., Benders, G.A., Montague, M.G., Ma, L., Moodie, M.M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E.A., Young, L., Qi, Z.Q., Segall-Shapiro, T.H., Calvey, C.H., Parmar, P.P., Hutchison, C.A., 3rd, Smith, H.O., and Venter, J.C. (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**: 52-56.

Goodall, E.C.A., Robinson, A., Johnston, I.G., Jabbari, S., Turner, K.A., Cunningham, A.F., Lund, P.A., Cole, J.A., and Henderson, I.R. (2018) The Essential Genome of Escherichia coli K-12. *Mbio* **9**.

Grazziotin, A.L., Vidal, N.M., and Venancio, T.M. (2015) Uncovering major genomic features of essential genes in Bacteria and a methanogenic Archaea. *Febs J* **282**: 3395-3411.

Hashimoto, M., Ichimura, T., Mizoguchi, H., Tanaka, K., Fujimitsu, K., Keyamura, K., Ote, T., Yamakawa, T., Yamazaki, Y., Mori, H., Katayama, T., and Kato, J. (2005) Cell size and nucleoid organization of engineered Escherichia coli cells with a reduced genome. *Mol Microbiol* **55**: 137-149.

Hirokawa, Y., Kawano, H., Tanaka-Masuda, K., Nakamura, N., Nakagawa, A., Ito, M., Mori, H., Oshima, T., and Ogasawara, N. (2013) Genetic manipulations restored the growth fitness of reduced-genome Escherichia coli. *J Biosci Bioeng* **116**: 52-58.

Holberger, L.E., Garza-Sanchez, F., Lamoureux, J., Low, D.A., and Hayes, C.S. (2012) A novel family of toxin/antitoxin proteins in Bacillus species. *FEBS Lett* **586**: 132-136.

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., and Venter, J.C. (2016) Design and synthesis of a minimal bacterial genome. *Science* **351**: aad6253.

Ismail, S., Alsowayeh, N., Abbasi, H.W., Albutti, A., Tahir Ul Qamar, M., Ahmad, S., Raza, R.Z., Sadia, K., and Abbasi, S.W. (2022) Pan-Genome-Assisted Computational Design of a Multi-

Epitopes-Based Vaccine Candidate against Helicobacter cinaedi. *Int J Environ Res Public Health* **19**.

Jain, S., Chang, A.M., Singh, M., McLean, J.S., Coats, S.R., Kramer, R.W., and Darveau, R.P. (2019) Identification of PGN_1123 as the Gene Encoding Lipid A Deacylase, an Enzyme Required for Toll-Like Receptor 4 Evasion, in Porphyromonas gingivalis. *J Bacteriol* **201**.

Jewett, M.C., and Forster, A.C. (2010) Update on designing and building minimal cells. *Curr Opin Biotechnol* **21**: 697-703.

Juhas, M., Eberl, L., and Church, G.M. (2012) Essential genes as antimicrobial targets and cornerstones of synthetic biology. *Trends Biotechnol* **30**: 601-607.

Klein, B.A., Duncan, M.J., and Hu, L.T. (2015) Defining essential genes and identifying virulence factors of Porphyromonas gingivalis by massively parallel sequencing of transposon libraries (Tn-seq). *Methods Mol Biol* **1279**: 25-43.

Kobayashi, K., Ehrlich, S.D., Albertini, A., Amati, G., Andersen, K.K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S.C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L.C., Danchin, A., Debarbouille, M., Dervyn, E., Deuerling, E., Devine, K., Devine, S.K., Dreesen, O., Errington, J., Fillinger, S., Foster, S.J., Fujita, Y., Galizzi, A., Gardan, R., Eschevins, C., Fukushima, T., Haga, K., Harwood, C.R., Hecker, M., Hosoya, D., Hullo, M.F., Kakeshita, H., Karamata, D., Kasahara, Y., Kawamura, F., Koga, K., Koski, P., Kuwana, R., Imamura, D., Ishimaru, M., Ishikawa, S., Ishio, I., Le Coq, D., Masson, A., Mauel, C., Meima, R., Mellado, R.P., Moir, A., Moriya, S., Nagakawa, E., Nanamiya, H., Nakai, S., Nygaard, P., Ogura, M., Ohanan, T., O'Reilly, M., O'Rourke, M., Pragai, Z., Pooley, H.M., Rapoport, G., Rawlins, J.P., Rivas, L.A., Rivolta, C., Sadaie, A., Sadaie, Y., Sarvas, M., Sato, T., Saxild, H.H., Scanlan, E., Schumann, W., Seegers, J.F., Sekiguchi, J., Sekowska, A., Seror, S.J., Simon, M., Stragier, P., Studer, R., Takamatsu, H., Tanaka, T., Takeuchi, M., Thomaides, H.B., Vagner, V., van Dijl, J.M., Watabe, K., Wipat, A., Yamamoto, H., Yamamoto, M., Yamamoto, Y., Yamane, K., Yata, K., Yoshida, K., Yoshikawa, H., Zuber, U., and Ogasawara, N. (2003) Essential Bacillus subtilis genes. *Proc Natl Acad Sci U S A* **100**: 4678-4683.

Koo, B.M., Kritikos, G., Farelli, J.D., Todor, H., Tong, K., Kimsey, H., Wapinski, I., Galardini, M., Cabal, A., Peters, J.M., Hachmann, A.B., Rudner, D.Z., Allen, K.N., Typas, A., and Gross, C.A. (2017) Construction and Analysis of Two Genome-Scale Deletion Libraries for Bacillus subtilis. *Cell Syst* **4**: 291-305 e297.

Ku, C., Nelson-Sathi, S., Roettger, M., Garg, S., Hazkani-Covo, E., and Martin, W.F. (2015) Endosymbiotic gene transfer from prokaryotic pangenomes: Inherited chimerism in eukaryotes. *Proc Natl Acad Sci U S A* **112**: 10139-10146.

Lukjancenko, O., Wassenaar, T.M., and Ussery, D.W. (2010) Comparison of 61 sequenced Escherichia coli genomes. *Microb Ecol* **60**: 708-720.

Manabe, K., Kageyama, Y., Morimoto, T., Ozawa, T., Sawada, K., Endo, K., Tohata, M., Ara, K., Ozaki, K., and Ogasawara, N. (2011) Combined effect of improved cell yield and increased specific productivity enhances recombinant enzyme production in genome-reduced Bacillus subtilis strain MGB874. *Appl Environ Microbiol* **77**: 8370-8381.

Manabe, K., Kageyama, Y., Morimoto, T., Shimizu, E., Takahashi, H., Kanaya, S., Ara, K., Ozaki, K., and Ogasawara, N. (2013) Improved production of secreted heterologous enzyme in Bacillus subtilis strain MGB874 via modification of glutamate metabolism and growth conditions. *Microb Cell Fact* **12**: 18.

Matern, W.M., Jenquin, R.L., Bader, J.S., and Karakousis, P.C. (2020) Identifying the essential genes of Mycobacterium avium subsp. hominissuis with Tn-Seq using a rank-based filter procedure. *Sci Rep* **10**: 1095.

McCutcheon, J.P., and Moran, N.A. (2011) Extreme genome reduction in symbiotic bacteria. *Nat Rev Microbiol* **10**: 13-26.

Medini, D., Donati, C., Tettelin, H., Masignani, V., and Rappuoli, R. (2005) The microbial pan-genome. *Curr Opin Genet Dev* **15**: 589-594.

Mizoguchi, H., Mori, H., and Fujio, T. (2007) Escherichia coli minimum genome factory. *Biotechnol Appl Biochem* **46**: 157-167.

Morimoto, T., Kadoya, R., Endo, K., Tohata, M., Sawada, K., Liu, S., Ozawa, T., Kodama, T., Kakeshita, H., Kageyama, Y., Manabe, K., Kanaya, S., Ara, K., Ozaki, K., and Ogasawara, N. (2008) Enhanced recombinant protein productivity by genome reduction in Bacillus subtilis. *DNA Res* **15**: 73-81.

Pedreira, T., Elfmann, C., and Stulke, J. (2022) The current state of SubtiWiki, the database for the model organism Bacillus subtilis. *Nucleic Acids Res* **50**: D875-D882.

Peters, J.M., Colavin, A., Shi, H., Czarny, T.L., Larson, M.H., Wong, S., Hawkins, J.S., Lu, C.H.S., Koo, B.M., Marta, E., Shiver, A.L., Whitehead, E.H., Weissman, J.S., Brown, E.D., Qi, L.S., Huang, K.C., and Gross, C.A. (2016) A Comprehensive, CRISPR-based Functional Analysis of Essential Genes in Bacteria. *Cell* **165**: 1493-1506.

Pi, H., and Helmann, J.D. (2018) Genome-Wide Characterization of the Fur Regulatory Network Reveals a Link between Catechol Degradation and Bacillibactin Metabolism in Bacillus subtilis. *Mbio* **9**.

Pietiainen, M., Francois, P., Hyyrylainen, H.L., Tangomo, M., Sass, V., Sahl, H.G., Schrenzel, J., and Kontinen, V.P. (2009) Transcriptome analysis of the responses of Staphylococcus aureus to antimicrobial peptides and characterization of the roles of vraDE and vraSR in antimicrobial resistance. *Bmc Genomics* **10**: 429.

Plumbridge, J.A. (1989) Sequence of the nagBACD operon in Escherichia coli K12 and pattern of transcription within the nag regulon. *Mol Microbiol* **3**: 505-515.

Poulsen, B.E., Clatworthy, A.E., and Hung, D.T. (2022) The Use of Tn-Seq and the FiTnEss Analysis to Define the Core Essential Genome of Pseudomonas aeruginosa. *Methods Mol Biol* **2377**: 179-197.

Rahmer, R., Morabbi Heravi, K., and Altenbuchner, J. (2015) Construction of a Super-Competent Bacillus subtilis 168 Using the P mtlA -comKS Inducible Cassette. *Front Microbiol* **6**: 1431.

Reuss, D.R., Altenbuchner, J., Mader, U., Rath, H., Ischebeck, T., Sappa, P.K., Thurmer, A., Guerin, C., Nicolas, P., Steil, L., Zhu, B., Feussner, I., Klumpp, S., Daniel, R., Commichau, F.M., Volker, U., and Stulke, J. (2017) Large-scale reduction of the Bacillus subtilis genome: consequences for the transcriptional network, resource allocation, and metabolism. *Genome Res* **27**: 289-299.

Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y., and Ishikawa, H. (2000) Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS. *Nature* **407**: 81-86.

Soucy, S.M., Huang, J., and Gogarten, J.P. (2015) Horizontal gene transfer: building the web of life. *Nat Rev Genet* **16**: 472-482.

Steingard, C.H., and Helmann, J.D. (2023) Meddling with Metal Sensors: Fur-Family Proteins as Signaling Hubs. *J Bacteriol* **205**: e0002223.

Swetha, R.G., Basu, S., Ramaiah, S., and Anbarasu, A. (2022) Multi-Epitope Vaccine for Monkeypox Using Pan-Genome and Reverse Vaccinology Approaches. *Viruses* **14**.

Terakawa, A., Natsume, A., Okada, A., Nishihata, S., Kuse, J., Tanaka, K., Takenaka, S., Ishikawa, S., and Yoshida, K.I. (2016) Bacillus subtilis 5'-nucleotidases with various functions and substrate specificities. *BMC Microbiol* **16**: 249.

Tettelin, H., Riley, D., Cattuto, C., and Medini, D. (2008) Comparative genomics: the bacterial pan-genome. *Curr Opin Microbiol* **11**: 472-477.

van Opijnen, T., Bodi, K.L., and Camilli, A. (2009) Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat Methods* **6**: 767-772.

Wenzel, M., and Altenbuchner, J. (2015) Development of a markerless gene deletion system for Bacillus subtilis based on the mannose phosphoenolpyruvate-dependent phosphotransferase system. *Microbiology* **161**: 1942-1949.

Wernegreen, J.J. (2002) Genome evolution in bacterial endosymbionts of insects. *Nat Rev Genet* **3**: 850-861.

Westers, H., Dorenbos, R., van Dijl, J.M., Kabel, J., Flanagan, T., Devine, K.M., Jude, F., Seror, 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., and Quax, W.J. (2003) Genome engineering reveals large dispensable regions in Bacillus subtilis. *Mol Biol Evol* **20**: 2076-2090.

Yang, M.R., Su, S.F., and Wu, Y.W. (2023) Using bacterial pan-genome-based feature selection approach to improve the prediction of minimum inhibitory concentration (MIC). *Front Genet* **14**: 1054032.

Yuvaraj, I., Sridhar, J., Michael, D., and Sekar, K. (2017) PanGeT: Pan-genomics tool. *Gene* **600**: 77-84.

Zhou, J., Wu, R., Xue, X., and Qin, Z. (2016) CasHRA (Cas9-facilitated Homologous Recombination Assembly) method of constructing megabase-sized DNA. *Nucleic Acids Res* **44**: e124.

Zhu, D., Yang, Z., Xu, J., Wang, M., Jia, R., Chen, S., Liu, M., Zhao, X., Yang, Q., Wu, Y., Zhang, S., Liu, Y., Zhang, L., Yu, Y., Chen, X., and Cheng, A. (2020) Pan-genome analysis of Riemerella anatipestifer reveals its genomic diversity and acquired antibiotic resistance associated with genomic islands. *Funct Integr Genomics* **20**: 307-320.