



Article

New Insights on Metabolic Features of *Bacillus subtilis* Based on Multistrain Genome-Scale Metabolic Modeling

Blas Blázquez ^{1,2,†} , David San León ^{1,2,†} , Antonia Rojas ³, Marta Tortajada ³ and Juan Nogales ^{1,2,*}

¹ Department of Systems Biology, Centro Nacional de Biotecnología, CSIC, 28049 Madrid, Spain; bblazquez@cnb.csic.es (B.B.); dsanleon@cnb.csic.es (D.S.L.)

² Interdisciplinary Platform for Sustainable Plastics towards a Circular Economy-Spanish National Research Council (SusPlast-CSIC), 28040 Madrid, Spain

³ Archer Daniels Midland, Nutrition, Biopolis S.L. Parc Científic Universitat de València, Carrer del Catedrático Agustín Escardino Benlloch, 9, 46980 Paterna, Spain; antonia.rojas@amd.com (A.R.); marta.tortajada@adm.com (M.T.)

* Correspondence: jnogales@cnb.csic.es

† These authors contributed equality to this work.

Abstract: *Bacillus subtilis* is an effective workhorse for the production of many industrial products. The high interest aroused by *B. subtilis* has guided a large metabolic modeling effort of this species. Genome-scale metabolic models (GEMs) are powerful tools for predicting the metabolic capabilities of a given organism. However, high-quality GEMs are required in order to provide accurate predictions. In this work, we construct a high-quality, mostly manually curated genome-scale model for *B. subtilis* (iBB1018). The model was validated by means of growth performance and carbon flux distribution and provided significantly more accurate predictions than previous models. iBB1018 was able to predict carbon source utilization with great accuracy while identifying up to 28 metabolites as potential novel carbon sources. The constructed model was further used as a tool for the construction of the panphenome of *B. subtilis* as a species, by means of multistrain genome-scale reconstruction. The panphenome space was defined in the context of 183 GEMs representative of 183 *B. subtilis* strains and the array of carbon sources sustaining growth. Our analysis highlights the large metabolic versatility of the species and the important role of the accessory metabolism as a driver of the panphenome, at a species level.

Keywords: *Bacillus subtilis*; genome-scale metabolic model; flux balance analysis; multistrain modeling; panphenome



Citation: Blázquez, B.; San León, D.; Rojas, A.; Tortajada, M.; Nogales, J. New Insights on Metabolic Features of *Bacillus subtilis* Based on Multistrain Genome-Scale Metabolic Modeling. *Int. J. Mol. Sci.* **2023**, *24*, 7091. <https://doi.org/10.3390/ijms24087091>

Academic Editor: Rosa María Martínez-Espinoso

Received: 10 March 2023

Revised: 1 April 2023

Accepted: 10 April 2023

Published: 11 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Aside from being the model organism of Firmicutes clade, *Bacillus subtilis* is a biotechnological workhorse for the production of many industrial products, such as enzymes, vitamins, and antibiotics [1]. In recent years, a considerable effort has been aimed at engineering better-performing *B. subtilis* strains, producing high added value industrial products by means of metabolic engineering [2,3]. However, addressing such an ambitious goal often requires advanced approaches beyond classic metabolic engineering [4]. In this context, the complete understanding of microbial metabolic space, its optimization, and rational expansion require holistic approaches accounting for system-level properties. Since mathematical modeling can be used as accurate scaffolds towards this purpose, it is not surprising that modeling approaches have been developed and thoroughly used for optimizing microbial cell factories and to address important biological questions in the last decade [5,6]. Genome-scale metabolic models (GEMs) are mathematical representations of cellular metabolism accounting for the reactions' stoichiometries and reversibility under the assumption of a metabolic steady state [7]. Their intrinsic simplicity makes these models powerful tools to analyze biological networks at the genome scale. Overall, GEMs have

proven to be powerful computational tools for the computation of the metabolic capabilities and phenotype of a given organism, even under perturbations, and can also be used for strain optimization.

To date, *B. subtilis* has been subject to significant metabolic modeling efforts, and various genome-scale metabolic network models of *B. subtilis* are available: *iYO844* [8], *iBsu1103* [9], *iBsu1103v2* [2], *iBsu1147* [10], and *iBsu1144* [11]. The first GEM, *iYO844*, was constructed including phenotyping and gene essentiality data. More recently, this pioneer model has been updated by integrating 17 enzyme kinetics for the central carbon and related pathway reactions, significantly improving model predictions regarding growth rate and flux distribution [12]. Another series of GEMs of *B. subtilis* was constructed using the semi-automatic workflow developed by the SEED Project [9,13–15]. The initial model, *iBsu1103* (which also includes Gibbs free energy change values for 1403 of the reactions), has been updated several times in the last decade [2,10,11]. The second update, *iBsu1103v2*, analyzed and fixed the incorrect gene essentiality predictions of its initial version. Finally, the latest model updates (*iBsu1147* and *iBsu1144*) were aiming to optimize the production of relevant added-value compounds including riboflavin, Egl-237, isobutanol, butanediol, and serine protease. Despite the efforts in the reconstruction of *B. subtilis* metabolism, it has been recently shown that the current models are far from accurately predicting well-known metabolic traits of this strain [16]. Overall, any discrepancies with experimental data were mainly due to incorrect or incomplete annotations, missing reactions and/or pathways, and/or inaccurate formulation of the biomass reaction. Therefore, there is still much room for improvement in the metabolic modeling of this interesting microorganism.

On the other hand, the increasing number of genomes being sequenced is allowing researchers to address the variability across strains within the same species. As more genome sequences become available, substantial differences in genomic content and functions across strains can be identified [17]. Therefore, it is now possible to explore strain-specific variations using approaches such as pangenome analyses [18]. Pangenome studies have highlighted the existence of important differences among strains, thus explaining divergent phenotypes within a single species. However, despite the utility of pangenome-based functional analyses, they do not provide a mechanistic insight into phenotypic potential based on genetic and genomic variability within a species [19,20]. In this context, the possibility of addressing the metabolic modeling of the whole set of strains of one given species becomes, in an optimal computational framework, an opportunity to analyze, from a mechanistic point of view, the phenotypic differences in a bacterial species [21–28]. However, such multistrain modeling requires high-quality GEMs for each reference strain, which are not always available, thus limiting this approach to a low number of species [25,28–30].

In this work, we develop a new and high-quality genome-scale model for *B. subtilis* (*iBB1018*) based on the annotations of the SEED project [15]. *iBB1018* was built based on *iBsu1103v2* but subject to a thoroughly manual curation that included the available biochemical and physiological knowledge. The quality of *iBB1018* was further validated by means of growth performance and carbon flux distribution providing significantly more accurate predictions than previous models. Overall, *iBB1018* predicts the carbon sources tested with high precision (84%), while it identifies up to 28 metabolites as potential novel carbon sources, thus significantly expanding the known metabolic versatility of *B. subtilis*. Furthermore, this updated *B. subtilis* GEM provides an essential tool to study pan-metabolic capabilities, thus providing clues about the metabolic range of *Bacillus subtilis* as a species. Subsequently, we provide here a multistrain metabolic model of *B. subtilis*.

2. Results

2.1. Genome-Scale Metabolic Reconstruction

The genome-scale metabolic model *iBB1018* (Supplementary Additional Files S3 and S4) contains 1577, 1291, and 1018 unique reactions, metabolites, and genes, respectively. The model was created by updating the metabolic content derived from *iBsu1103v2* following a procedure that includes three iterative steps (Figure 1). Firstly, an initial draft model was

constructed from the SEED-based *iBsu1103v2* model. SEED nomenclature was then converted to BiGG annotations using the platform MetaNetX as an intermediate while genes ID were updated according to NCBI (BSU) nomenclature.

Secondly, we carried out an exhaustive revision and manual curation of all the reactions based on the analysis of each Gene–Protein Reaction (GPR) association included in the draft. Therefore, each individual GPR was evaluated based on physiological evidence and information allocated on several databases, such as KEGG [31], MetaCyc [32], and BRENDA [33]. During the manual curation, multiple inconsistencies were found in the draft, which mainly fell into four categories: (1) 5.4% of the reactions initially included in the draft were incorrectly described, in terms of stoichiometry and/or cofactor dependency; (2) 10.8% of the reactions of the initial model were mass or charge unbalanced and were adjusted accordingly; (3) 16.1% of the reactions were modified to change orientation and/or reversibility according to KEGG and MetaCyc databases. Many of these incorrect reactions yielded an unrealistic production of ATP which was not associated with metabolic costs, e.g., pyruvate carboxylase was modified to be irreversible; and (4) 39 new reactions, mostly transporter-related, were added to the model based on physiological and/or genomic data.

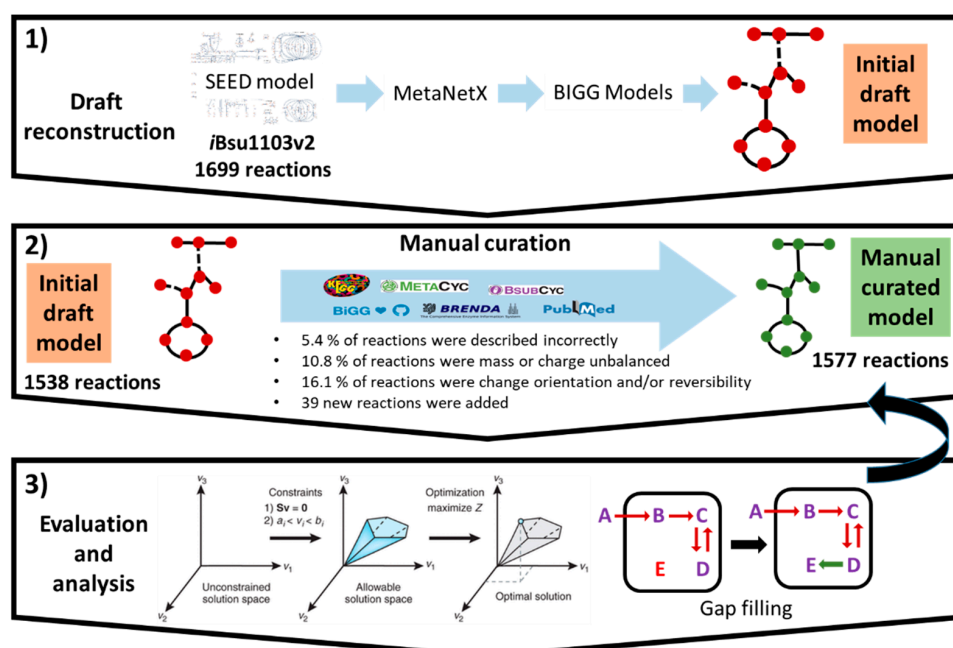


Figure 1. Reconstruction process of *iBB1018*. (1) The initial model was reconstructed using *iBsu1103v2* as a template within the framework of MetaNetX and BiGG model databases. (2) An exhaustive manual curation was carried out using bibliographic, biochemical, and metabolic databases. (3) Biomass production was selected as an objective function for gap-filling analysis.

Finally, in a third step, we applied a manual gap-filling analysis in order to complete several biosynthetic pathways required to synthesize elements from the biomass reaction. This process was profusely guided by the information available in biological databases and the scientific literature. For instance, it was necessary to add the transporter reactions for D-arabinose, galactitol, and sucrose due to a lack of genes associated with these reactions in the initial draft model.

The metabolic expansion accounted for by *iBB1018* in comparison with the highly curated model *iYO844* was highlighted by means of a Clusters of Orthologous Genes (COG) analysis (Figure 2, Table S2). We observed an increase in the number of reactions in almost all the different COGs categories. A noteworthy feature was the significant increase of reactions on carbohydrates transport and metabolism, which led to an increase in the number of new carbon sources used by the model, thus driving a better understanding of the actual metabolic capabilities of *B. subtilis*.

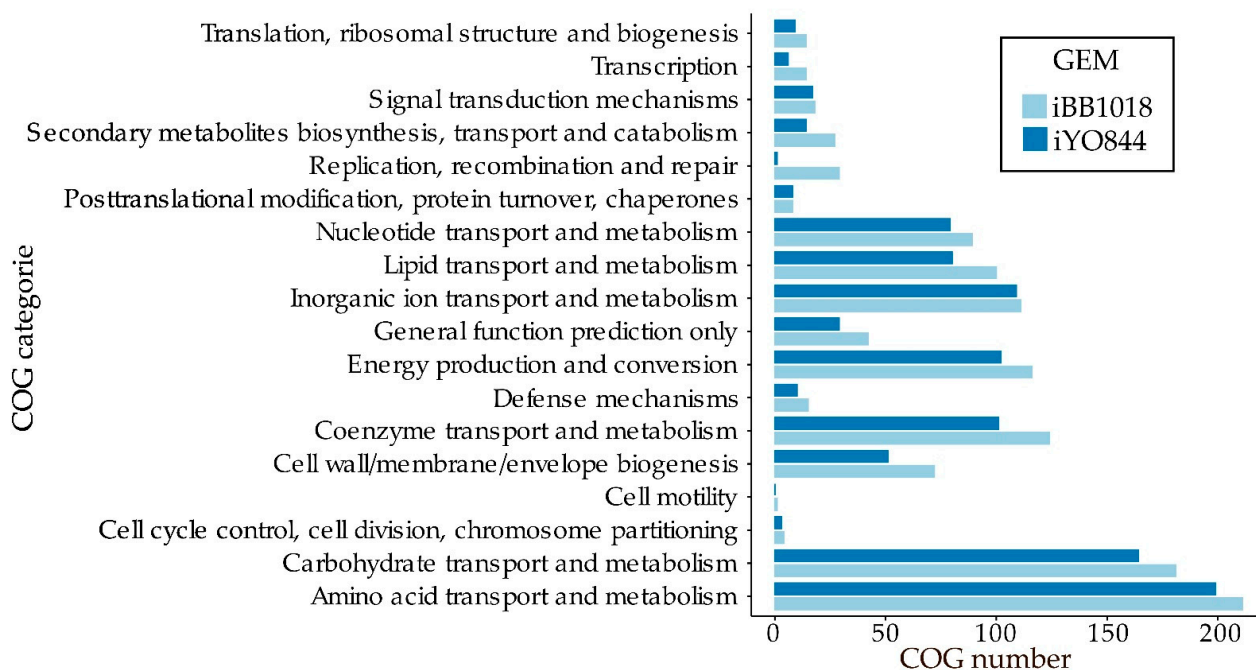


Figure 2. COG comparison between *iBB1018* and *iYO844* models. The bars indicate the number of genes associated with a specific COG category.

In order to ensure the quality of the *iBB1018* model, we used the MEMOTE tool [34], a well-known standardized quality control tool for GEMs. The quality evaluation included the missing annotations test, format structuration, searching imbalanced reactions, and searching of errors in stoichiometries. These tests allow the evaluation of the consistency of the metabolic network. The model reached an overall score of 81.09%, which indicates a good model completeness (Figure S1). The consistency of the model scored was 99.8%, which represents the accuracy in stoichiometry, mass and charge balance, metabolite connectivity, and reaction cycles. This analysis confirms that *iBB1018* is a complete and detailed model. A comparison between the models *iBB1018*, *iBsu1103v2*, and *iYO844* by MEMOTE showed high scores for *iBB1018* and *iYO844*, with 81% and 80%, respectively. On the other hand, the *iBsu1103v2* model had the lowest score, 36% (Figure S1, Supplementary Additional File S2).

2.2. Growth Rate Performance of *iBB1018* Is in Agreement with Several In Vivo Nutritional Scenarios

Large completeness and consistency of a model does not always mean a high-quality model and GEMs need to be validated by assessing their ability to compute physiological states [35]. To validate *iBB1018*, we evaluated its capability predicting well-known metabolic traits of *B. subtilis*, with such growth performance in glucose as the sole carbon source [36]. Aerobic growth was simulated allowing only the uptake of glucose while the upper bound of the rest of the exchange reactions remained unconstrained (Table 1). The performance of *iBB1018* was further compared with that from previous models including *iYO844* and the reference SEED model *iBsu1103v2*. This analysis shows the higher accuracy of *iBB1018* when compared with previous models. In fact, *iBB1018* not only predicts the closest growth rate to that reported in vivo, but also a more accurate acetate production. It is noteworthy that while *iYO844* slightly overestimated acetate secretion, resulting in a significantly lower growth rate, *iBsu1103v2* was unable to predict acetate secretion at all, thus overestimating the growth rate.

Table 1. Parameters of *B. subtilis* 168 growing in glucose in vivo and in silico. Constraints used to feed the model are underlined.

	Glucose Uptake mmol g ⁻¹ h ⁻¹	Oxygen Uptake mmol g ⁻¹ h ⁻¹	Growth Rate h ⁻¹	Acetate Production mmol g ⁻¹ h ⁻¹
In vivo [36]	8.71 ± 0.64	18	0.67 ± 0.02	4.28 ± 0.29
<i>i</i> BB1018	<u>8.71</u>	<u>18</u>	0.69	4.03
<i>i</i> Bsu1103v2	<u>8.71</u>	<u>18</u>	1.27	0
<i>i</i> YO844	<u>8.71</u>	<u>18</u>	0.61	5.53

Since *B. subtilis* is able to grow on additional carbon sources besides glucose, we decided to test the performance of *i*BB1018 on alternative nutritional scenarios in order to complete the model validation. Specifically, we monitored the growth rate of *i*BB1018 using single carbon sources such as citrate, gluconate, and glucose [37], while fixing the riboflavin secretion to experimental data [38,39]. Overall, we found large agreement between in vivo and in silico values, despite model-based predictions slightly overestimating in vivo growth rates in all the carbon sources simulated (Table 2). These discrepancies are within an acceptable error and are often found even in high-quality GEMs [29,30]. They might be explained by: (i) a still incomplete formulation of biomass function, (ii) higher condition-specific energy maintenance requirements not accounted for in the current reconstruction, and (iii) missing adaptation to these alternative compounds as primary carbon sources [35]. Furthermore, these discrepancies may also refer to an incomplete in vivo phenotyping of the strain, for example, not having taken into account possible metabolic byproducts such lactate and pyruvate or the secretion of partially oxidized metabolites.

Table 2. Validation of growth performance of *i*BB1018 under various carbon source and production scenarios. Constraints used to feed the model are underlined.

Carbon Source Substrate ^a Uptake (mmol g ⁻¹ h ⁻¹)	In Vivo Riboflavin Secretion (mmol g ⁻¹ h ⁻¹)	In Vivo Growth Rate (h ⁻¹)	In Silico Growth Rate (h ⁻¹)
$v_{glc} = \underline{1.55}; v_{cit} = 0.5$	<u>0.0173</u>	0.10	0.12
$v_{glc} = \underline{1.55}; v_{glcn_D} = \underline{0.6}$	<u>0.0258</u>	0.12	0.13
$v_{glc} = \underline{1.7}$	<u>0.0181</u>	0.10	0.10
$v_{glc} = \underline{3.2}$	<u>0.0210</u>	0.20	0.24
$v_{glc} = \underline{4.7}$	<u>0.0231</u>	0.30	0.39
$v_{glc} = \underline{6.2}$	<u>0.0255</u>	0.40	0.54

^a glc: D-glucose; cit: citrate; glcn_D: D-gluconate.

Once *i*BB1018 was validated applying experimental data from the literature, the capability of the in silico model *i*BB1018 to predict aerobic growth on various carbon sources was determined (Figure 3, Table S3 in Supplementary Additional File S1). The in silico computations were compared against experimental data from the literature [8,40–42]. Growth on different substrates was simulated by fixing their specific uptake rates to 10 mmol g⁻¹ h⁻¹ under aerobic conditions in minimal media. *i*BB1018 was able to correctly predict the growth of the bacterium in 80 carbon sources. The model also suggests 28 new compounds as potential carbon sources, including D-xylose or stachyose. This was possible thanks to the expansion of the network, with manual gap filling of reactions involved in the transport of metabolites. Network gap analysis found metabolites that were either consumed only or produced only, and therefore broke the material balance on the metabolites. The candidate metabolites were identified, and reactions were analyzed and filled. Although in silico predictions generally agreed with experimental data, this model also erroneously indicated growth for four carbon sources (4-aminobutanoate, glycine, L-isoleucine, and L-valine), as

previous *iBsu1103v2* and *iYO844* models did (Figure 3). Such extended inaccurate predictions argue in favor of a lack of adaptation of *B. subtilis* to these alternative compounds as primary carbon sources, as has been discussed previously [28].

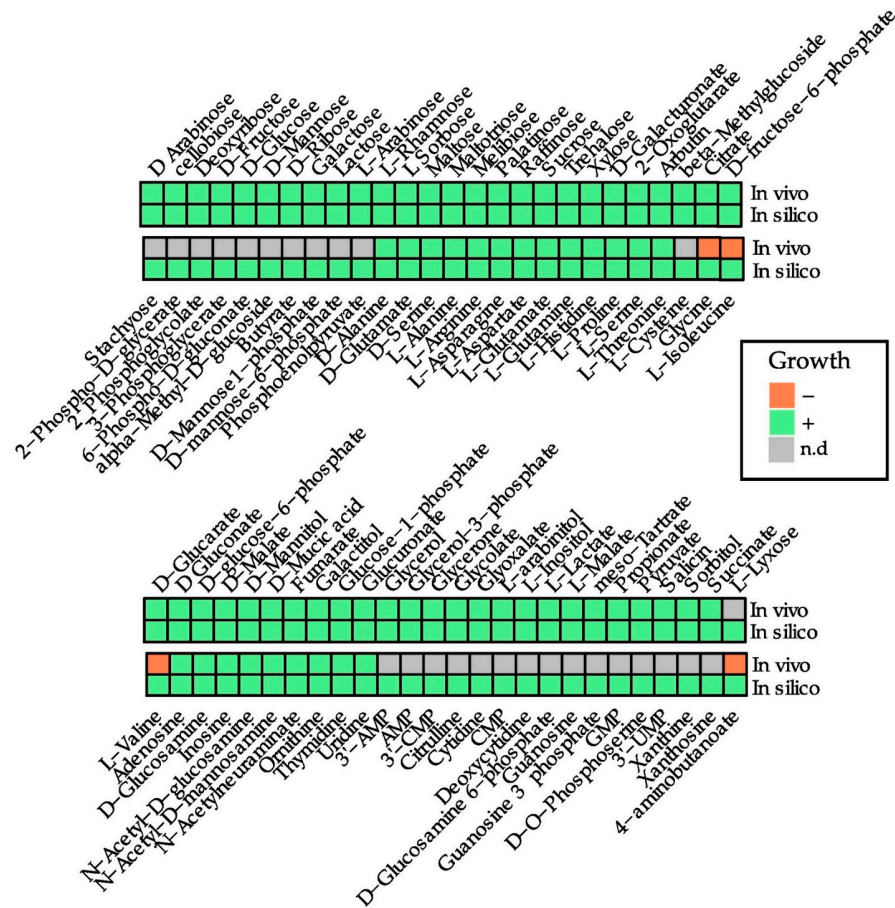


Figure 3. In vivo growth rates based on Biolog data from the literature [8,40–42] and in silico carbon sources predicted for *B. subtilis* using the *iBB1018* model under aerobic conditions. Only carbon sources with positive growth in vivo or in silico are included. The uptake of each potential carbon source was limited to 10 mmol/gDW.h. n.d.: not determinate.

2.3. *iBB1018* Exhibits Superior Performance Predicting Carbon Flux Distribution Than Previous Models

To further quantitatively evaluate the predictive capabilities of *iBB1018*, we performed a careful analysis of carbon flux distribution using glucose as a single carbon and energy source under aerobic conditions. Subsequently, we compared the fluxes predicted with those experimentally reported [43]. Fluxes prediction using *iYO844* and *iBsu1103v2* were included in the comparison in order to benchmark the performance of these previous models. Overall, this analysis highlights that the manually curated models (*iYO844* and *iBB1018*) can predict experimentally reported metabolic fluxes with more accuracy (Figure S2, Table S1). In contrast, the automatically constructed *iBsu1103v2* yields significantly less accurate fluxes' predictions.

When focused on glycolysis, the three models exhibit good agreement with experimental data, despite *iBsu1103v2*'s slightly underestimated flux through the late pathway. In the case of pentose phosphate pathways, the flux predictions are underestimated, even though the performance of the manually curated models is slightly better in terms of flow correlation. The detailed analysis of the TCA cycle shows that *iBB1018* and *iYO844* predicted net flux through all the reactions of TCA. In contrast, *iBsu1103v2* significantly underutilizes TCA, predicting null flux through five of the eight reactions involved in

the cycle. In addition, the only three active reactions presented underestimated fluxes. A correlation plot was used to compare the models (Figure 4), showing low values for all the models and presenting the model *iBB1018* with the highest correlation ($R = 0.49$, p -value 0.2). The *iBsu1103v2* model shows, as expected, the lowest correlation values (Figure 4A).

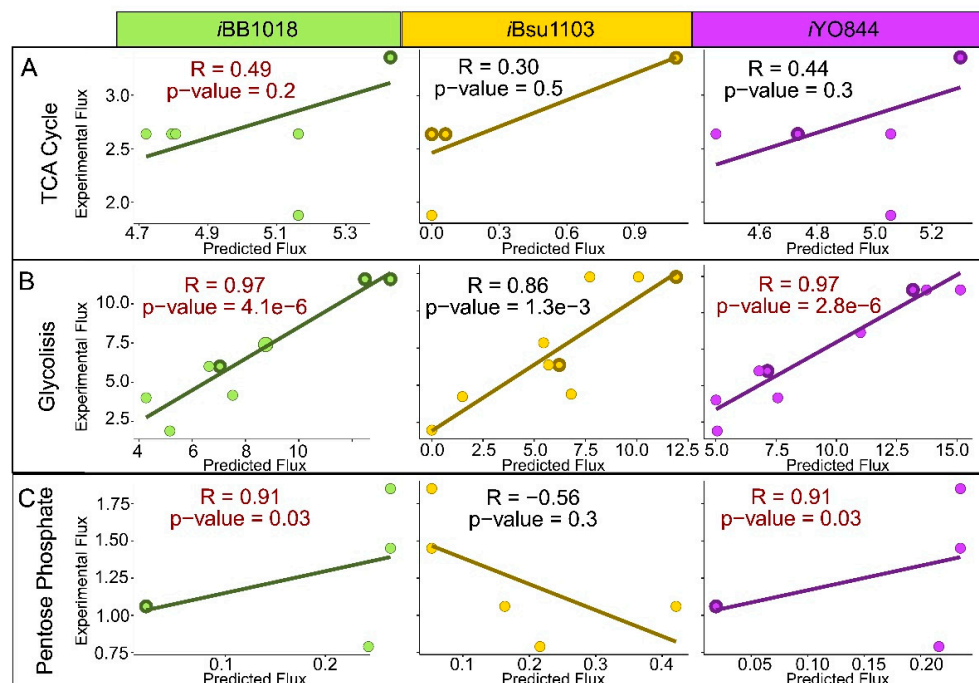


Figure 4. Correlation plot between experimental fluxes from the literature [43] and in silico fluxes predicted by *iBB1018*, *iBsu1103v2*, and *iYO844* models for TCA cycle (A), glycolysis (B), and pentose phosphate pathway (C). All the represented fluxes are measured in $\text{mmol gDW}^{-1} \text{h}^{-1}$. The dots with wider borders contain overlapping values.

2.4. Multistrain Modeling of *B. subtilis* as Species

Multistrain genome-scale metabolic models can be used to study metabolic diversity and speciation [44]. In order to analyze the phenotypic potential of *B. subtilis*, as species, we first computed the pangenome of *B. subtilis* and, by using the high-quality model *iBB1018* as a template, we further addressed the multistrain metabolic modeling of the whole species [28].

The pangenome of 184 *B. subtilis* strains was addressed by performing a homology analysis in order to identify orthologous genes among the analyzed strains. We noted a large variability, in terms of number of genes, among the different genomes, ranging from 2623 (CP029052, *B. subtilis* BS155) to 7044 (AP012495, *B. subtilis* BEST7613) (Figure S3, Table S4). Interestingly, the second-largest genome including 4191 genes from *B. subtilis* BEST7613 was identified as a chimera genome harboring an important fragment of the genome of the cyanobacterium *Synechocystis* strain PCC6803 [45]; thus, it was discarded in order to avoid potential errors in the downstream analysis.

Pangenomes make reference to the entire set of genes from all the strains within a clade. Therefore, we defined the core pangenome as that including those genes present in a high percentage of strains (>95%), while the accessory pangenome was defined as that representing unique or shared functions by a reduced group of strains (Figures 5 and S3). While the pangenome is useful in terms of comparative analysis, for modeling purposes, the panphenome is more useful since it is the set of all the in silico metabolic phenotypes displayed by the collection of *B. subtilis* GEMs constructed. Subsequently, we used the estimated pangenome to compute the panphenome of *B. subtilis*. For this, we filtered metabolic genes using the *iBB1018* and BiGG database as a source of metabolic functions

(Methods) and, subsequently, the list of metabolic genes for each strain was reduced to those with any known metabolic function. Interestingly, we found a more homogeneous number of genes in each strain (between 900 and 1200 genes), which strongly suggests that metabolism is highly conserved in contrast to other functionalities. The computed panphenome was finally used as basic metabolic information to construct strain-specific GEMs for each strain analyzed. The GEMs were automatically gap-filled followed by semi-manual curation to include metabolic information other than that included in the core panphenome. A total of 183 strain-specific GEMs were constructed, covering 2324 of the 13,398 gene families present in the calculated pangenome (17.3%). Most of the sequences not included into the GEMs were not annotated or failed to show any known function (35%).

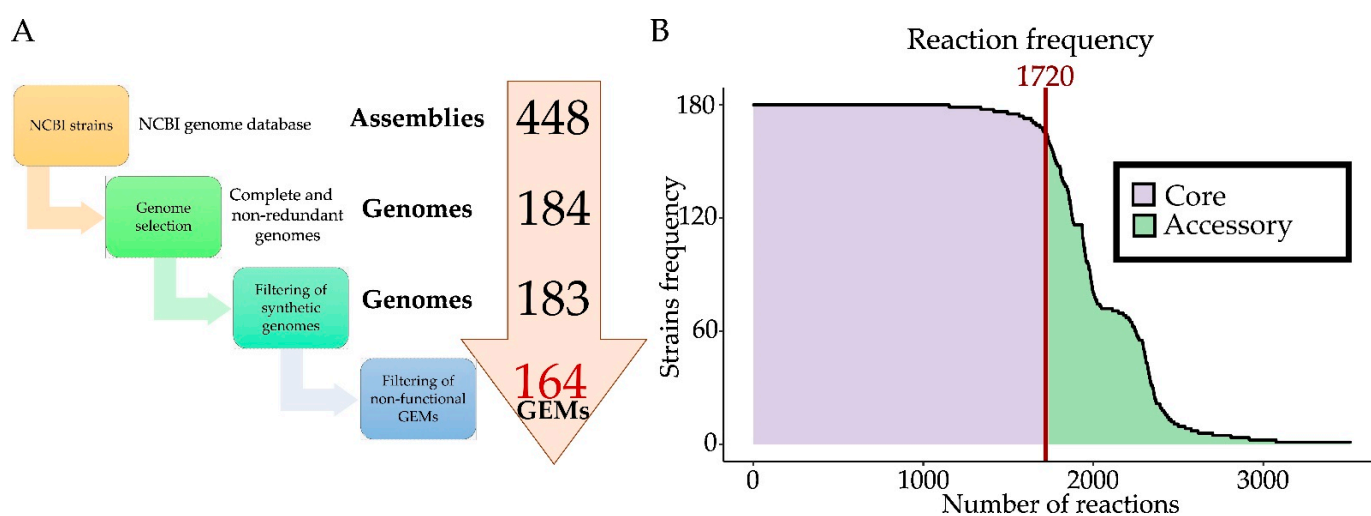


Figure 5. (A) Schematic representation of the multiple filtering performed to calculate the final GEMs. (B) Reaction frequency among the selected *B. subtilis* strains. The purple area represents core metabolism and green represents accessory metabolism.

To validate this multistrain modeling and gain further insights in the phenotypic features of *B. subtilis*, strain-specific GEMs were used to compute those carbon sources supporting growth in the analyzed strains. (Figure 6, Table S5). This analysis unveiled a small set of 18 incomplete and non-functional models which still need additional gap filling or manual curation steps; therefore, only 165 GEMs were further analyzed. Overall, the functional analysis of *B. subtilis* strains based on their cognate GEMs allows the classification of the analyzed strains in two main groups: (i) High Metabolically Versatile (HMV) strains and (ii) Low Metabolically Versatile (LMV) strains (Figure 6). HMV strains show the ability for growth on a large variety of carbon sources, including mono and polysaccharides, amino acids, nucleotides, organic acids, and TCA intermediate, among others. In contrast, LMV strains exhibit a reduced catabolic space, being limited to mono and polysaccharides and a few nucleotides. It is noteworthy that two of the strains identified as LMV, i.e., *B. subtilis* strain PS38 (CP016789) and PG10 (CP016788), were subject to synthetic genome reduction in the context of minimal genome identification [46], therefore explaining, to a great extent, the low metabolic versatility computed. HMV strains could be additionally classified into two subgroups, including HMV α which exhibits growth efficiently in ascorbic acid, 2-dehydro-D-gluconate, Cis-aconitate, and L-alanine-L-glutamate but not on L-arabinitol, and HMV β which exhibits the opposite phenotype.

Regarding carbon sources, we identified a group of selected monosaccharides, oligosaccharides, and polysaccharides, such as xylose, fructose, and starch, as universal carbon sources (>90% strains with growth) that were able to sustain the growth of most of the *B. subtilis* strains. Interestingly, such carbon sources provided the highest growth rate

computed, strongly suggesting optimal carbon sources for *B. subtilis* irrespective of the particular strain being analyzed. A second group of carbon sources that included additional mono and polysaccharides, amino acids, nucleotides, and organic acids were used for a large number of *B. subtilis* strains, constituting a second core carbon source for this species. Finally, we identified an important variety of carbon sources providing growth to only a low number of strains (less than 50); therefore, we consider these compounds as strain-specific carbon sources.

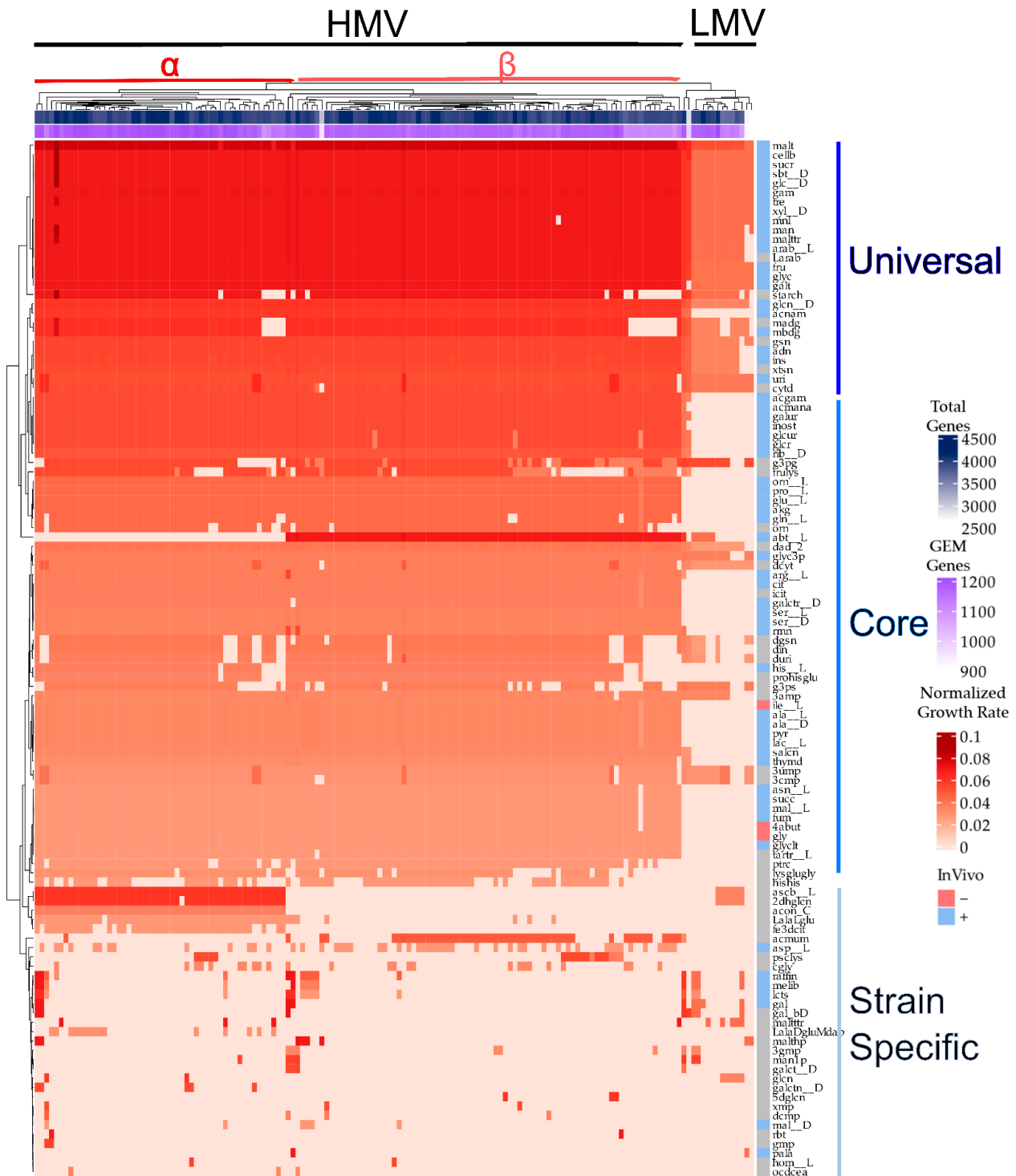


Figure 6. Heatmap with the panphenome growth rates across multiple carbon sources. The top color bars indicate the total number of genes per genome (blue) and the total genes included in the GEMs

(purple). The right bar indicates if the carbon source was tested in Biolog with *B. subtilis* 168 and the growth result in vivo. The data were hierarchically clustered with the “manhattan” distance method. The α and β symbols represent HMV subgroups. The universal class contains carbon sources with more than 90% of the strains with growth, the core class contains between 30% and 90%, and the strain-specific class contains less than 30%. The *B. subtilis* strain names are not shown, and the carbon sources are identified with the reduced BiGG namespace IDs to facilitate the visualization. Growth rates were normalized with the number of carbons of each carbon source metabolite.

3. Discussion

Bacillus subtilis is an important microorganism used for biotechnological processes which have triggered a large effort of metabolic modeling in order to construct accurate and predictive computational platforms towards biotechnology endeavors. However, available in silico GEMs of *B. subtilis* are far away, in term of completeness and predictive potential, from equivalent models of other well-known and widely used microbial cell factories such as *Escherichia coli* [29], *Pseudomonas putida* [30], and *Saccharomyces cerevisiae* [47]. In this study, we addressed this challenge and presented an upgraded genome-scale model for *B. subtilis* (*iBB1018*). *iBB1018* presents higher metabolic completeness and predictive capabilities than previously published *B. subtilis* GEMs, and scores as the highest-quality GEM when analyzed with the MEMOTE tool [34]. The final model was highly validated against reported data and was used for an in-depth analysis of *B. subtilis* metabolism exhibiting a high grade of accuracy in terms of carbon source utilization, growth rate, and flux distribution predictions under various nutritional scenarios.

Once a high-quality GEM exists, its contents (e.g., genes, metabolites, and reactions) can be straightforwardly mapped onto a closely related strain in a species. Following this, a multistrain approach can be integrated with genome-scale modeling in order to gain novel insights underlying the variability of phenotypes. This approach enables panphenome analysis, thus empowering species-wide comparative systems biology that has been applied to well-known species in a variety of studies, mainly involving the assessment of metabolic potential. *E. coli* was the first species subject to multistrain modeling following its high number of GEMs. An initial study by Monk and colleagues addressed the modeling of 55 fully sequenced *E. coli* strains identifying strain-specific auxotrophies and substrate preferences among the set of strains [21]. Extending this analysis, a new multistrain effort included 1200 *E. coli* strains and demonstrated a large variability both in gene content and function within the species. The computed panphenome was further used to construct a robust classification tree for determination between extra-intestinal and intra-intestinal pathogens [29]. Following the *E. coli* example, multistrain modeling has been applied to compute the panphenome of other bacteria having high-quality GEMs available. For instance, 450 *Salmonella* strains were modeled from various serovars to show that metabolic capabilities can be used to discriminate these serovars [25]. Additionally, and taking advantage of high-quality GEMs, environmental and biotechnologically relevant bacteria such as *P. putida* were subject to multistrain modeling, showing that the large metabolic versatility described for many *P. putida* strains is a common feature of the whole species [30].

Despite the large interest in *B. subtilis*, its multistrain modeling has remained elusive due to the lack of high-quality GEMs to be used as guide. Here, by using *iBB1018* as a high-quality template, we have addressed the modeling of 184 *B. subtilis* strains whose genomes were available at the time this study started. Overall, we were able to construct a total of 165 functional GEMs which define a large panphenome space in terms of a carbon source supporting growth (Figure 6). Overall, we found a large phenotypic homogeneity. Most of the strains could be classified as high metabolically versatile according to the number of carbon sources they were capable of using. With some exceptions, we found that 95% of the strains were able to use up to 72% of the carbon source analyzed as a sole carbon and energy source. Interestingly, we identified that a large number of carbon sources (28%) were only used by a few strains (less than 50), highlighting the role of the accessory metabolism as the driving force of the panphenome at a species level [48]. In

addition to the intrinsic value of the panphenome identified in this study, the collection of *B. subtilis* GEMs provided here constitutes an intangible value which can be leveraged as high-quality guidance in future metabolic studies of this valuable bacteria species.

4. Materials and Methods

4.1. Model Reconstruction and Analysis

*i*BB1018 was reconstructed by refining and updating the previous genome-scale model for *B. subtilis*. The metabolic network was reconstructed using standardized protocols for metabolic reconstruction [35,49]. An initial draft reconstruction was generated from the annotated genome of *B. subtilis* str. 168 (GenBank number: AL009126.3) and the metabolic content from model *i*Bsu1103v2 based on SEED annotations. The draft model was converted to BiGG annotations through MetaNetX [50]. The final draft was subjected to manual curation of the metabolic information in order to remove potential inconsistencies and to connect the network as much as possible, using metabolic and GEMs databases including KEGG [31], MetaCyc [32], BRENDA [33], and BiGG [51]. Transport reactions were added by using the TransportDB database [52]. To improve the functionality of the model, gap filling was performed using the RAVEN function *fillGaps* to ensure that the model could grow in silico. Moreover, a manual gap-filling step was performed in order to connect the network as much as possible and remove potential inconsistencies or close death-end metabolites. Finally, it was confirmed that all reactions were balanced, and all metabolites were annotated with their charge and location (*_c*, cytosol; *_e*, extracellular space).

A robust correlation method, biweight midcorrelation [53] implemented in the WGCNA R package [54], was used to compare the samples. This non-parametric, median-based method was chosen because the datasets do not follow a normal distribution and it is less sensitive to outliers. The plots were generated with the R package *ggpubr* [55].

To compare the performance between the three models, MEMOTE [34] was executed with the function “diff”.

For the comparison of COGs between models, the genes from the *i*BB1018 and *i*YO844 models were linked to COGs by using the NCBI COG database version 2020 (<https://www.ncbi.nlm.nih.gov/research/cog>, accessed on 22 December 2022) (Table S2, Supplementary Additional File S1). The genes were aggregated into COG categories to identify the improvements in high-level systems and in the case associated with multiple COG categories, they were counted multiple times.

4.2. Flux Balance Analysis (FBA)

Flux Balance Analysis (FBA) was used to evaluate the biomass production and to predict flux distributions once the biomass reaction was fixed as the objective function (BOF, Biomass Objective Function). The foundations and applications of FBA have been reviewed elsewhere [56]. The growth prediction when executing FBA was the growth rate (h^{-1}) predicted under the minimal M9 media, as previously described [30]. All the exchange reactions were sequentially tested in minimal media as potential carbon and nitrogen sources, using an uptake of $-10 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ to avoid limitations. Regarding carbon sources, growth rate was further normalized considering the number of carbons of each compound in both oxic and anoxic conditions. The COBRA ToolBox v.3.0 [57] running on the MatLab platform was used for FBA simulations.

4.3. Multistrain

All assemblies of diverse strains of *B. subtilis* were downloaded from the NCBI database (488 downloaded assemblies). Only the complete assemblies (NCBI completeness class) with gene annotation and proteome availability were used for this analysis. To reduce redundancy, only non-redundant genomes at the proteome level were selected, with 184 strains in total. The final genome dataset contained 184 genomes of *B. subtilis*. To perform multistrain models, the model developed for this work was used as a template

and, following the protocol designed by Norsigian [28], the models of other strains were created. To identify the multiple carbon sources used by the strains, in-house python scripts were used.

Multi-gap filling was run over the single-strain models without growth in any carbon source to identify essential genes from the metabolic information, with the results combined with the homology results. For this analysis, the minimum percentage of homology used to mark a gene as an ortholog of *B. subtilis* 168 was 70%, but this threshold was relaxed when the number of genes of strains with a percentage of homology over 40% was present in more than 75% of strains. In this case, the threshold decreased to 40% of homology.

The genes without homology against the template were compared to the total BiGG gene datasets to identify possible homologs in other species and models. The involved reactions were included in the final models.

The carbon source analysis over the *B. subtilis* strains GEMs were driven by the FBA implementation of COBRApy v0.20. The uptakes were normalized based on the number of carbons. The code and the multistrain GEMs can be accessed in the following GitHub repository (https://github.com/SBGlab/Bacillus_Subtilis_multistrain_GEM).

Supplementary Materials: The supplementary materials are available online at <https://www.mdpi.com/article/10.3390/ijms24087091/s1>.

Author Contributions: Conceptualization, B.B. and J.N.; methodology B.B., D.S.L. and J.N.; software, D.S.L.; formal analysis, B.B., D.S.L. and A.R.; data curation, B.B., D.S.L. and J.N.; writing—original draft preparation, B.B., D.S.L. and J.N.; writing—review and editing, B.B., D.S.L., J.N., A.R. and M.T.; supervision, J.N.; project administration, B.B.; funding acquisition, J.N. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement no. 814650 (SynBio4Flav). Funding was likewise provided by the Spanish Ministry of Science and Innovation under the RobExplode project: PID2019-108458RB-I00 (AEI/10.13039/501100011033) and CSIC’s Interdisciplinary Platform for Sustainable Plastics towards a Circular Economy+ (PTI-SusPlast+).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: https://github.com/SBGlab/Bacillus_Subtilis_multistrain_GEM.

Acknowledgments: The authors wish to thank Jose Miguel Seoane for his critical reading of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of this study; in the collection, analyses, or interpretation of data; in the writing of this manuscript, or in the decision to publish the results.

References

1. Harwood, C.R.; Cranenburgh, R. *Bacillus* protein secretion: An unfolding story. *Trends Microbiol.* **2008**, *16*, 73–79. [[CrossRef](#)] [[PubMed](#)]
2. 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. [[CrossRef](#)] [[PubMed](#)]
3. Kang, Z.; Yang, S.; Du, G.; Chen, J. Molecular engineering of secretory machinery components for high-level secretion of proteins in *Bacillus* species. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 1599–1607. [[CrossRef](#)] [[PubMed](#)]
4. Long, M.R.; Ong, W.K.; Reed, J.L. Computational methods in metabolic engineering for strain design. *Curr. Opin. Biotechnol.* **2015**, *34*, 135–141. [[CrossRef](#)]
5. Gu, C.; Kim, G.B.; Kim, W.J.; Kim, H.U.; Lee, S.Y. Current status and applications of genome-scale metabolic models. *Genome Biol.* **2019**, *20*, 121. [[CrossRef](#)]

6. Gudmundsson, S.; Nogales, J. Recent advances in model-assisted metabolic engineering. *Curr. Opin. Syst. Biol.* **2021**, *28*, 100392. [[CrossRef](#)]
7. Fang, X.; Lloyd, C.J.; Palsson, B.O. Reconstructing organisms in silico: Genome-scale models and their emerging applications. *Nat. Rev. Microbiol.* **2020**, *18*, 731–743. [[CrossRef](#)]
8. Oh, Y.K.; Palsson, B.O.; Park, S.M.; Schilling, C.H.; Mahadevan, R. Genome-scale reconstruction of metabolic network in *Bacillus subtilis* based on high-throughput phenotyping and gene essentiality data. *J. Biol. Chem.* **2007**, *282*, 28791–28799. [[CrossRef](#)]
9. Henry, C.S.; Zinner, J.F.; Cohoon, M.P.; Stevens, R.L. iBsu1103: A new genome-scale metabolic model of *Bacillus subtilis* based on SEED annotations. *Genome Biol.* **2009**, *10*, R69. [[CrossRef](#)]
10. Hao, T.; Han, B.; Ma, H.; Fu, J.; Wang, H.; Wang, Z.; Tang, B.; Chen, T.; Zhao, X. In silico metabolic engineering of *Bacillus subtilis* for improved production of riboflavin, Egl-237, (R,R)-2,3-butanediol and isobutanol. *Mol. Biosyst.* **2013**, *9*, 2034–2044. [[CrossRef](#)]
11. Kocabaş, P.; Çalık, P.; Çalık, G.; Özdamar, T.H. Analyses of extracellular protein production in *Bacillus subtilis*—I: Genome-scale metabolic model reconstruction based on updated gene-enzyme-reaction data. *Biochem. Eng. J.* **2017**, *127*, 229–241. [[CrossRef](#)]
12. Massaiu, I.; Pasotti, L.; Sonnenschein, N.; Rama, E.; Cavaletti, M.; Magni, P.; Calvio, C.; Herrgård, M.J. Integration of enzymatic data in *Bacillus subtilis* genome-scale metabolic model improves phenotype predictions and enables in silico design of poly- γ -glutamic acid production strains. *Microb. Cell Fact.* **2019**, *18*, 3. [[CrossRef](#)] [[PubMed](#)]
13. Overbeek, R.; Dlsz, T.; Stevens, R. The seed: A peer-to-peer environment for genome annotation. *Commun. ACM* **2004**, *47*, 46–50. [[CrossRef](#)]
14. Aziz, R.K.; Bartels, D.; Best, A.; DeJongh, M.; Disz, T.; Edwards, R.A.; Formsma, K.; Gerdes, S.; Glass, E.M.; Kubal, M.; et al. The RAST Server: Rapid annotations using subsystems technology. *BMC Genom.* **2008**, *9*, 75. [[CrossRef](#)]
15. Overbeek, R.; Olson, R.; Pusch, G.D.; Olsen, G.J.; Davis, J.J.; Disz, T.; Edwards, R.A.; Gerdes, S.; Parrello, B.; Shukla, M.; et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* **2014**, *42*, D206. [[CrossRef](#)]
16. Ghasemi-Kahrizsangi, T.; Marashi, S.-A.; Hosseini, Z. Genome-scale metabolic network models of *Bacillus* species suggest that model improvement is necessary for biotechnological applications. *Iran. J. Biotechnol.* **2018**, *16*, e1684. [[CrossRef](#)]
17. Brockhurst, M.A.; Harrison, E.; Hall, J.P.J.; Richards, T.; McNally, A.; MacLean, C. The ecology and evolution of pangenomes. *Curr. Biol.* **2019**, *29*, R1094–R1103. [[CrossRef](#)]
18. Vernikos, G.S. A review of pangenome tools and recent studies. In *The Pangenome: Diversity, Dynamics and Evolution of Genomes*; Springer International Publishing: Berlin/Heidelberg, Germany, 2020; pp. 89–112. ISBN 9783030382810.
19. Norsigian, C.J.; Fang, X.; Palsson, B.O.; Monk, J.M. Pangenome flux balance analysis toward panphenomes. In *The Pangenome: Diversity, Dynamics and Evolution of Genomes*; Springer: Cham, Switzerland, 2020; pp. 219–232. [[CrossRef](#)]
20. Hyun, J.C.; Monk, J.M.; Palsson, B.O. Comparative pangenomics: Analysis of 12 microbial pathogen pangenomes reveals conserved global structures of genetic and functional diversity. *BMC Genom.* **2022**, *23*, 7. [[CrossRef](#)]
21. Monk, J.M.; Charusanti, P.; Aziz, R.K.; Lerman, J.A.; Premyodhin, N.; Orth, J.D.; Feist, A.M.; Palsson, B. Genome-scale metabolic reconstructions of multiple *Escherichia coli* strains highlight strain-specific adaptations to nutritional environments. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 20338–20343. [[CrossRef](#)]
22. Bosi, E.; Monk, J.M.; Aziz, R.K.; Fondi, M.; Nizet, V.; Palsson, B. Comparative genome-scale modelling of *Staphylococcus aureus* strains identifies strain-specific metabolic capabilities linked to pathogenicity. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E3801–E3809. [[CrossRef](#)]
23. Fouts, D.E.; Matthias, M.A.; Adhikarla, H.; Adler, B.; Amorim-Santos, L.; Berg, D.E.; Bulach, D.; Buschiazzo, A.; Chang, Y.F.; Galloway, R.L.; et al. What makes a bacterial species pathogenic?: Comparative genomic analysis of the genus *Leptospira*. *PLoS Negl. Trop. Dis.* **2016**, *10*, e0004403. [[CrossRef](#)]
24. Norsigian, C.J.; Kavvas, E.; Seif, Y.; Palsson, B.O.; Monk, J.M. iCN718, an updated and improved genome-scale metabolic network reconstruction of *Acinetobacter baumannii* AYE. *Front. Genet.* **2018**, *9*, 121. [[CrossRef](#)] [[PubMed](#)]
25. Seif, Y.; Kavvas, E.; Lachance, J.C.; Yurkovich, J.T.; Nuccio, S.P.; Fang, X.; Catoiu, E.; Raffatellu, M.; Palsson, B.O.; Monk, J.M. Genome-scale metabolic reconstructions of multiple *Salmonella* strains reveal serovar-specific metabolic traits. *Nat. Commun.* **2018**, *9*, 3771. [[CrossRef](#)]
26. Fang, X.; Monk, J.M.; Mih, N.; Du, B.; Sastry, A.V.; Kavvas, E.; Seif, Y.; Smarr, L.; Palsson, B.O. *Escherichia coli* B2 strains prevalent in inflammatory bowel disease patients have distinct metabolic capabilities that enable colonization of intestinal mucosa. *BMC Syst. Biol.* **2018**, *12*, 66. [[CrossRef](#)] [[PubMed](#)]
27. Norsigian, C.J.; Pusarla, N.; McConn, J.L.; Yurkovich, J.T.; Dräger, A.; Palsson, B.O.; King, Z. BiGG Models 2020: Multi-strain genome-scale models and expansion across the phylogenetic tree. *Nucleic Acids Res.* **2020**, *48*, D402–D406. [[CrossRef](#)]
28. Norsigian, C.J.; Fang, X.; Seif, Y.; Monk, J.M.; Palsson, B.O. A workflow for generating multi-strain genome-scale metabolic models of prokaryotes. *Nat. Protoc.* **2020**, *15*, 1–14. [[CrossRef](#)]
29. Monk, J.M.; Lloyd, C.J.; Brunk, E.; Mih, N.; Sastry, A.; King, Z.; Takeuchi, R.; Nomura, W.; Zhang, Z.; Mori, H.; et al. iML1515, a knowledgebase that computes *Escherichia coli* traits. *Nat. Biotechnol.* **2017**, *35*, 904–908. [[CrossRef](#)] [[PubMed](#)]

30. Nogales, J.; Mueller, J.; Gudmundsson, S.; Canalejo, F.J.; Duque, E.; Monk, J.; Feist, A.M.; Ramos, J.L.; Niu, W.; Palsson, B.O. High-quality genome-scale metabolic modelling of *Pseudomonas putida* highlights its broad metabolic capabilities. *Environ. Microbiol.* **2020**, *22*, 255–269. [[CrossRef](#)]
31. Ogata, H.; Goto, S.; Sato, K.; Fujibuchi, W.; Bono, H.; Kanehisa, M. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **1999**, *27*, 29–34. [[CrossRef](#)]
32. Caspi, R.; Altman, T.; Billington, R.; Dreher, K.; Foerster, H.; Fulcher, C.A.; Holland, T.A.; Keseler, I.M.; Kothari, A.; Kubo, A.; et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res.* **2014**, *42*, D459–D471. [[CrossRef](#)]
33. Placzek, S.; Schomburg, I.; Chang, A.; Jeske, L.; Ulbrich, M.; Tillack, J.; Schomburg, D. BRENDA in 2017: New perspectives and new tools in BRENDA. *Nucleic Acids Res.* **2017**, *45*, D380–D388. [[CrossRef](#)]
34. Lieven, C.; Beber, M.E.; Olivier, B.G.; Bergmann, F.T.; Ataman, M.; Babaei, P.; Bartell, J.A.; Blank, L.M.; Chauhan, S.; Correia, K.; et al. MEMOTE for standardized genome-scale metabolic model testing. *Nat. Biotechnol.* **2020**, *38*, 272–276. [[CrossRef](#)]
35. Thiele, I.; Palsson, B. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* **2010**, *5*, 93–121. [[CrossRef](#)] [[PubMed](#)]
36. Tännler, S.; Decasper, S.; Sauer, U. Maintenance metabolism and carbon fluxes in *Bacillus* species. *Microb. Cell Fact.* **2008**, *7*, 19. [[CrossRef](#)] [[PubMed](#)]
37. Belda, E.; Sekowska, A.; Le Fèvre, F.; Morgat, A.; Mornico, D.; Ouzounis, C.; Vallenet, D.; Médigue, C.; Danchin, A. An updated metabolic view of the *Bacillus subtilis* 168 genome. *Microbiology* **2013**, *159*, 757–770. [[CrossRef](#)] [[PubMed](#)]
38. Dauner, M.; Sauer, U. Stoichiometric growth model for riboflavin-producing *Bacillus subtilis*. *Biotechnol. Bioeng.* **2001**, *76*, 132–143. [[CrossRef](#)] [[PubMed](#)]
39. Dauner, M.; Sonderegger, M.; Hochuli, M.; Szyperski, T.; Wüthrich, K.; Hohmann, H.P.; Sauer, U.; Bailey, J.E. Intracellular carbon fluxes in riboflavin-producing *Bacillus subtilis* during growth on two-carbon substrate mixtures. *Appl. Environ. Microbiol.* **2002**, *68*, 1760–1771. [[CrossRef](#)]
40. Chaudhuri, A.; Mishra, A.K.; Nanda, G. Variation of antimetabolite sensitivity with different carbon sources in *Bacillus subtilis*. *Folia Microbiol.* **1982**, *27*, 73–75. [[CrossRef](#)] [[PubMed](#)]
41. Sonenshein, A.L.; Hoch, J.A.; Losick, R. *Bacillus subtilis*: From cells to genes and from genes to cells. In *Bacillus subtilis and Its Closest Relatives*; ASM Press: Washington, DC, USA, 2001.
42. Yoshida, K.I.; Yamamoto, Y.; Omae, K.; Yamamoto, M.; Fujita, Y. Identification of two myo-inositol transporter genes of *Bacillus subtilis*. *J. Bacteriol.* **2002**, *184*, 983–991. [[CrossRef](#)]
43. Chubukov, V.; Uhr, M.; Le Chat, L.; Kleijn, R.J.; Jules, M.; Link, H.; Aymerich, S.; Stelling, J.; Sauer, U. Transcriptional regulation is insufficient to explain substrate-induced flux changes in *Bacillus subtilis*. *Mol. Syst. Biol.* **2013**, *9*, 709. [[CrossRef](#)]
44. Monk, J.M. Genome-scale metabolic network reconstructions of diverse *Escherichia* strains reveal strain-specific adaptations. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2022**, *377*, 20210236. [[CrossRef](#)]
45. Itaya, M.; Tsuge, K.; Koizumi, M.; Fujita, K. Combining two genomes in one cell: Stable cloning of the *Synechocystis* PCC6803 genome in the *Bacillus subtilis* 168 genome. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15971–15976. [[CrossRef](#)] [[PubMed](#)]
46. Reuß, D.R.; Altenbuchner, J.; Mäder, U.; Rath, H.; Ischebeck, T.; Sappa, P.K.; Thürmer, A.; Guérin, C.; Nicolas, P.; Steil, L.; et al. Large-scale reduction of the *Bacillus subtilis* genome: Consequences for the transcriptional network, resource allocation, and metabolism. *Genome Res.* **2017**, *27*, 289–299. [[CrossRef](#)] [[PubMed](#)]
47. Oftadeh, O.; Salvy, P.; Masid, M.; Curvat, M.; Miskovic, L.; Hatzimanikatis, V. A genome-scale metabolic model of *Saccharomyces cerevisiae* that integrates expression constraints and reaction thermodynamics. *Nat. Commun.* **2021**, *12*, 4790. [[CrossRef](#)] [[PubMed](#)]
48. Belbahri, L.; Bouket, A.C.; Rekik, I.; Alenezi, F.N.; Vallat, A.; Luptakova, L.; Petrovova, E.; Oszako, T.; Cherrad, S.; Vacher, S.; et al. Comparative genomics of *Bacillus amyloliquefaciens* strains reveals a core genome with traits for habitat adaptation and a secondary metabolites rich accessory genome. *Front. Microbiol.* **2017**, *8*, 1438. [[CrossRef](#)]
49. Monk, J.; Nogales, J.; Palsson, B.O. Optimizing genome-scale network reconstructions. *Nat. Biotechnol.* **2014**, *32*, 447–452. [[CrossRef](#)]
50. Moretti, S.; Tran, V.D.T.; Mehl, F.; Ibberson, M.; Pagni, M. MetaNetX/MNXref: Unified namespace for metabolites and biochemical reactions in the context of metabolic models. *Nucleic Acids Res.* **2021**, *49*, D570–D574. [[CrossRef](#)]
51. King, Z.A.; Lu, J.; Dräger, A.; Miller, P.; Federowicz, S.; Lerman, J.A.; Ebrahim, A.; Palsson, B.O.; Lewis, N.E. BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.* **2016**, *44*, D515–D522. [[CrossRef](#)]
52. Elbourne, L.D.H.; Tetu, S.G.; Hassan, K.A.; Paulsen, I.T. TransportDB 2.0: A database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Res.* **2017**, *45*, D320–D324. [[CrossRef](#)]
53. Song, L.; Langfelder, P.; Horvath, S. Comparison of co-expression measures: Mutual information, correlation, and model based indices. *BMC Bioinform.* **2012**, *13*, 328. [[CrossRef](#)]
54. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [[CrossRef](#)] [[PubMed](#)]
55. Kassambara, A. ggpubr: Publication Ready Plots—Articles—STHDA. Available online: <http://www.sthda.com/english/articles/24-ggpubr-publication-ready-plots/> (accessed on 24 January 2022).

56. Orth, J.D.; Thiele, I.; Palsson, B.O. What is flux balance analysis? *Nat. Biotechnol.* **2010**, *28*, 245–248. [[CrossRef](#)] [[PubMed](#)]
57. Heirendt, L.; Arreckx, S.; Pfau, T.; Mendoza, S.N.; Richelle, A.; Heinken, A.; Haraldsdóttir, H.S.; Wachowiak, J.; Keating, S.M.; Vlasov, V.; et al. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat. Protoc.* **2019**, *14*, 639–702. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.