

# SCIENTIFIC REPORTS



Correction: Publisher Correction

OPEN

## *Bacillus* SEVA siblings: A Golden Gate-based toolbox to create personalized integrative vectors for *Bacillus subtilis*

Jara Radeck<sup>1,2</sup>, Daniel Meyer<sup>2</sup>, Nina Lautenschläger<sup>1</sup> & Thorsten Mascher<sup>1</sup>

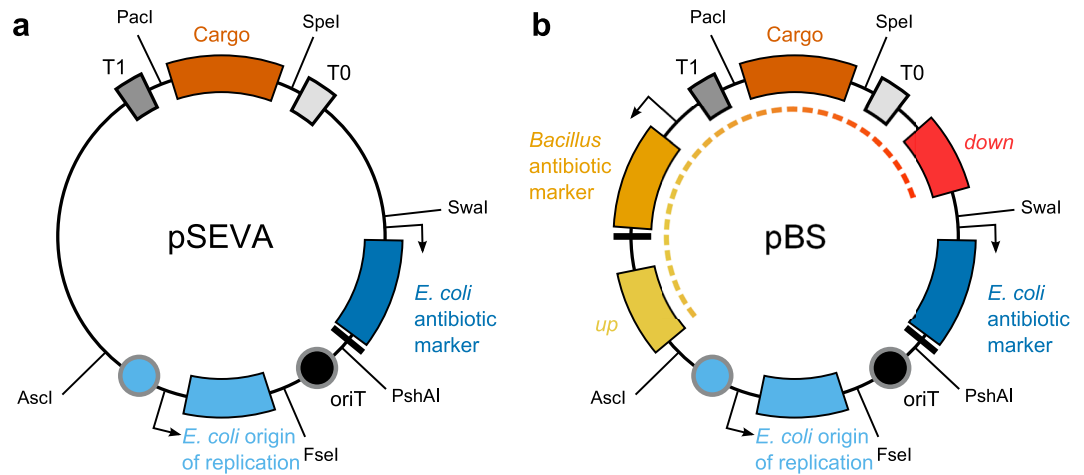
*Bacillus subtilis* combines natural competence for genetic transformation with highly efficient homologous recombination. These features allow using vectors that integrate into the genome via double homologous recombination. So far, their utilization is restricted by the fixed combination of resistance markers and integration loci, as well as species- or strain-specific regions of homology. To overcome these limitations, we developed a toolbox for the creation of personalized *Bacillus* vectors in a standardized manner with a focus on fast and easy adaptation of the sequences specifying the integration loci. We based our vector toolkit on the Standard European Vector Architecture (SEVA) to allow the usage of their vector parts. The *Bacillus* SEVA siblings are assembled via efficient one-pot Golden Gate reactions from four *entry* parts with the choice of four different enzymes. The toolbox contains seven *Bacillus* resistance markers, two *Escherichia coli* origins of replication, and a free choice of integration loci. Vectors can be customized with a cargo, before or after vector assembly, and could be used in different *B. subtilis* strains and potentially beyond. Our adaptation of the SEVA-standard provides a powerful and standardized toolkit for the convenient creation of personalized *Bacillus* vectors.

**The diversity of plasmid vectors in molecular cloning.** In molecular cloning, vectors are vehicles to transfer foreign nucleic acids into a living cell. In the context of this article, we restrict the term “vector” to “plasmid vectors”, small circular DNA molecules that originate from bacteria. They are easy to handle for inserts up to 10–15 kb, replicate independently of the bacterial chromosome and can be isolated in large amounts through standard plasmid preparation procedures. Over the last decades, vectors were increasingly modified to meet custom needs. For instance, they may contain promoters for ready-to-use gene expression, or reporter genes to measure transcription or translation rates, respectively. More elaborated vector types can contain biosafety features, internal measuring standards or may be used for clean chromosomal gene deletions or dual expression systems<sup>1–4</sup>. Synthetic biology strives to implement and extend the principles of engineering (standardization, decoupling, abstraction) into biology<sup>5</sup>. Especially standardization in vector and plasmid construction offers the clear advantage of comparability, compatibility, flexibility and reusability of single parts and whole vectors, as exemplified by the Standard European Vector Architecture (SEVA).

**The Standard European Vector Architecture provides standardized vectors for Gram-negative bacteria.** In 2013, the group of Victor de Lorenzo developed a standardized vector toolbox for the use in Gram-negative bacteria, with a special interest in *Pseudomonas putida*. With SEVA, they set the stage for a community-driven development platform and for evolving a standardized vector collection. This platform facilitates finding, creating, and naming of suitable vectors as well as their downstream handling<sup>6,7</sup>. Currently, the SEVA database lists 135 SEVA vectors and 49 SEVA siblings<sup>8</sup> (<http://seva.cnb.csic.es/>, March 2017). Subsequently, linker sequences were developed to make the vectors compatible to different cloning methods<sup>9</sup>.

Each SEVA vector contains at least three functional elements (Fig. 1a), (i) the origin of replication (*ori*) for vector replication in the host cell, (ii) a selectable marker, e.g. an antibiotic resistance cassette to select for vector

<sup>1</sup>Institute of Microbiology, Technische Universität (TU) Dresden, 01062, Dresden, Germany. <sup>2</sup>Department of Biology I, Ludwig-Maximilians-Universität (LMU) München, 82152, Planegg-Martinsried, Germany. Correspondence and requests for materials should be addressed to T.M. (email: [Thorsten.Mascher@tu-dresden.de](mailto:Thorsten.Mascher@tu-dresden.de))



**Figure 1.** Configuration of a SEVA vector and its *Bacillus* SEVA sibling. **(a)** The basic SEVA vector is composed of six parts, separated by defined endonuclease restriction sites. The transcriptional terminators T0 and T1 as well as the origin of transfer *oriT* are fixed, whereas the cargo, the ori and the antibiotic marker can be chosen freely from a pool of SEVA-compatible parts. **(b)** For genomic integration, three parts were added to the SEVA layout to create the *Bacillus* SEVA sibling pBS: Flanking homology regions *up* and *down*, as well as an antibiotic marker for *Bacillus*. Vector verification and propagation occurs in *E. coli* and only the part in between the homology regions (dashed line) will integrate into the genome. Vectors are drawn not to scale. Functional transcriptional units are indicated with an arrow (promoter) and black bar (terminator).

uptake and maintenance, and (iii) a multiple cloning site (MCS) for insertion of the DNA of interest. In the SEVA standard, the latter part is called cargo, irrespective of the size and nature of the insert. The cargo is isolated by flanking double transcriptional terminators (T1, T0), to avoid unwanted transcriptional read-through into other elements of the backbone. All parts are flanked by defined rare endonuclease restriction sites and assemble in a fixed order and orientation (Fig. 1a). They must not contain these and further restriction endonuclease recognition sites. Additionally, the transcriptional terminators and the origin of transfer (*oriT*, for plasmid conjugation) are predefined, whereas the selectable marker, ori and cargo can be chosen freely, as long as they adhere to the standard. An easy number-based nomenclature assures the fast determination of vector features from the vector's name<sup>6,7</sup>.

**Distinct features of vectors for *Bacillus subtilis*.** *Bacillus subtilis* is the best studied low-G + C Gram-positive bacterium (phylum *Firmicutes*) and one of the leading workhorses of the biotechnological industry<sup>10–12</sup>. Its ease of genetic manipulation is based on its natural competence, which includes the active uptake of (any) DNA and recombination of homologous regions into the chromosome<sup>13</sup>. These features enable to routinely use integrative instead of replicative vectors. This ensures genetic stability even without maintaining a constant selective pressure. Moreover, copy number effects are avoided, which can otherwise influence the promoter activity on replicative vectors, particularly at the single cell level<sup>14,15</sup>.

For cloning convenience, vectors designed for integration into the *B. subtilis* genome usually replicate in *E. coli* but do not contain an ori for *B. subtilis*. Instead, two homology regions of at least 400 bp, which define the insertion locus, flank the cargo and the resistance marker for *B. subtilis*. The plasmids are linearized in the *E. coli* part of the vector before *B. subtilis* transformation to avoid integration via single crossing over events. Consequently, only the DNA segment that is located between the homology regions will integrate into the chromosome, resulting in genetic stability of single copy number inserts even in the absence of selective pressure.

The currently available vectors combine known and effective selection cassettes with well-characterized integration loci<sup>16–20</sup>. This status quo limits the combination of resistance markers and integration sites, and does not allow targeting entirely new chromosomal regions. Moreover, existing vectors cannot be used in another *Bacillus* strain, which differs in the nucleotide sequence at the specified integration site.

Due to the genetic accessibility of *B. subtilis*, PCR-based methods, such as long-flanking homology (LFH)-PCR<sup>21</sup>, can be used to target new or strain-specific loci. While this approach is very convenient for generating knock-out mutants, it is restricted with respect to the cargo: fusing several PCR products, and obtaining large fragments, e.g. the *luxABCDE* bioluminescence reporter operon (5.7 kb) that allows online measurement<sup>17,19</sup>, with sufficient yield for efficient transformation (>1 µg<sup>22</sup>) can be tedious. In these cases, transformants have to be sequenced each time to ensure the integrity of the cargo (e.g. lack of mutations). In order to combine the advantages of vectors (stability and reusability) with those of PCR-based methods (flexibility) for the use in *B. subtilis* and related species, we designed, constructed and tested a new vector concept: the *Bacillus* SEVA siblings.

**Concept of *Bacillus* SEVA siblings as customized vectors.** Here, we describe a toolkit that was developed for the fast and easy generation of genomic integration vectors for *B. subtilis*, thereby overcoming the traditional limitations described above. Each vector contains a resistance marker, cargo and flanking homology regions of choice. To ensure compatibility and reusability of already existing parts for replication in *E. coli*, we

based our system on SEVA and added flanking homology regions as well as *Bacillus* resistance markers at defined positions (Fig. 1b). Our vectors will be named *Bacillus* SEVA siblings, which is in line with the current SEVA regulations<sup>6</sup>.

Our toolbox offers seven functional antibiotic resistance cassettes for selection in *B. subtilis* and high and medium copy number *E. coli* vectors modified to be assembled with your homology regions of interest in a one-pot Golden Gate assembly<sup>23,24</sup>. We tested the assembly efficiencies for five different enzymes and the *B. subtilis* transformability of every part provided. As proof of concept for efficient assembly and functionality, we analyzed the expression levels of the red fluorescing protein mKate2 under control of the xylose-inducible promoter  $P_{xylA}$  at different chromosomal insertion loci.

## Results

**Vector layout for double cross-over homologous recombination in *B. subtilis*.** We designed and constructed our vector building toolkit with the main goal to easily exchange the target loci for double homologous integration into the genome. This is important for two major reasons: Current vector collections do not allow the free combination of integration loci and selectable markers. Moreover, they omit the use of *Bacillus* strains if they differ from the reference lab strains in their nucleotide sequence, so that the integration sites of standard vectors are not compatible. For our collection of customizable vectors, we focus on allelic replacement via double recombination, where the DNA-sequence to be integrated into the chromosome (here termed *integration part*) is flanked by two regions of homology (Fig. 1b). For *B. subtilis*, approximately 400 bp of identical sequence are required for efficient integration. Shorter sequences (70 bp) can sometimes be sufficient, but are not recommended due to the significantly reduced efficiency<sup>25</sup>. For ease of construction, the *integration part* is combined with an ori and selectable marker for *E. coli* (here termed *replication part*). Before *B. subtilis* transformation, the vector is linearized in the *replication part* to avoid single cross-over events and ensure double cross-over integration.

Although our toolkit was optimized for the exchange of integration sites, it can be customized in manifold ways as will be outlined in the discussion. A quick user manual with helpful information for vector construction can be found in the supplemental material (Text S1).

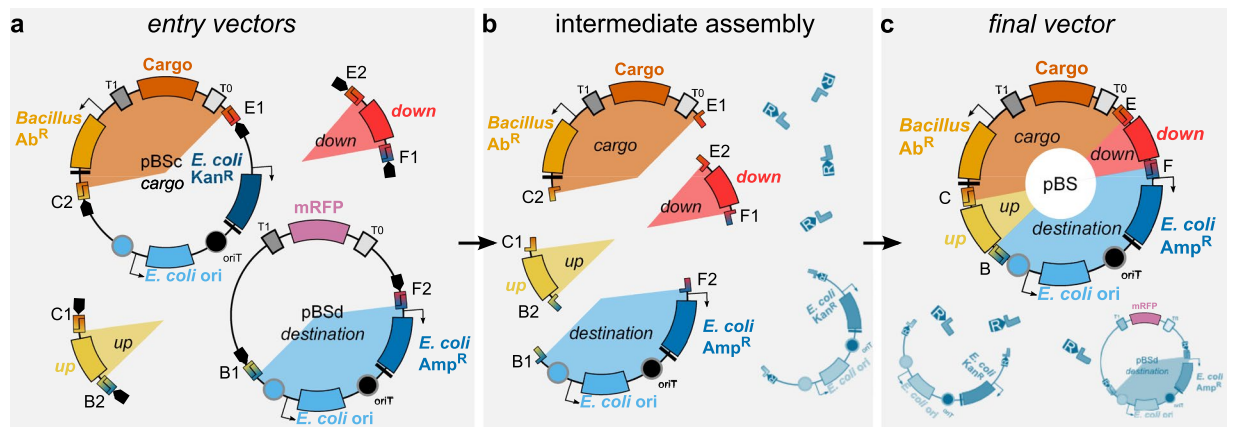
**The customizable vector collection is based on SEVA to allow re-use of *E. coli* parts.** Currently, most vectors and plasmids used for genetic manipulation of *B. subtilis* are propagated in *E. coli*. As demonstrated for NarK and  $\beta$ -carotene production, the *E. coli* vector propagation strongly depends on the ori and resistance markers in ways that cannot be foreseen yet<sup>9</sup>. Consequently, we designed our vectors to be flexible not only with respect to their cargo, but also their ori and *E. coli* resistance marker. For this purpose, we based our vector toolkit on SEVA. This standard was designed for vectors to allow exchanging of the cargo, resistance marker and ori, using defined rare type II restriction enzymes<sup>7</sup>. The accompanying vector collection was very well-received by scientists and is still growing. It can be used to adapt the vectors of our toolbox, which are therefore named *Bacillus* SEVA siblings. SEVA restriction sites were removed from critical parts of the *entry vectors*, so that the *final vectors* adhere to the SEVA standard, if no forbidden restriction sites are present in the customized parts.

### ***Bacillus* SEVA siblings are assembled *de novo* from multiple fragments via Golden Gate cloning.**

As depicted in Fig. 1b, the *integration* and *replication parts* of an integration vector are separated by the regions of homology, called *up* and *down*. Consequently, the exchange of both *up* and *down* fragments with standard cloning techniques would involve two cloning and verification steps without a selectable marker. To avoid laborious stepwise cloning, advanced cloning techniques for the easy, fast, efficient and directed one-pot assembly of multiple fragments are available, such as Gibson assembly<sup>26,27</sup> or Golden Gate cloning<sup>23,24</sup>. The latter is based on a ligase and type IIS endonucleases (e.g. BsaI), which cut outside (next to) their recognition site. Consequently, the restriction site can be designed according to need and separated from the recognition site during the cloning procedure. The reaction mix can include linear and circular DNA, containing restriction sites for the same enzyme but different overhangs. This allows the easy and efficient assembly of up to ten parts in the correct order and loss of the recognition sites in the *final vector*<sup>23</sup>. For our vector toolbox, we chose Golden Gate assembly for the following reasons: (i) Gibson assembly usually must be established in a lab to run smoothly, whereas restriction based cloning is more robust. (ii) The *up* and *down* regions need to be PCR-generated and Gibson assembly asks for longer overhangs thus increasing the primer costs.

Instead of re-using and modifying preexisting vectors, each new vector will be freshly assembled by combining four fragments via Golden Gate assembly (Fig. 2): the *up* and *down* fragments, the *replication part*, and the *cargo* which includes an antibiotic marker for selection in *Bacillus* and a MCS. Therefore, all vectors and sequences offered through the *Bacillus* Genetic Stock Center (BGSC) and the SEVA collection (see Table 1 and supplementary data S2) are *entry vectors* to allow customized assembly, but no *final Bacillus* vectors.

**Architecture of *Bacillus* SEVA siblings.** Figure 1 compares a *final Bacillus* SEVA sibling (pBS), with a standard SEVA vector for *E. coli*. SEVA suggests a designated location for the insertion of special features that are not part of the cargo. They are positioned at the terminator sequences and next to, but not obstructing the AscI and SwaI recognition sites which allow the exchange of SEVA vector parts. In line with this regulation, the homology regions and resistance marker are located at both terminator sequences. The respective *Bacillus* resistance cassette is placed between the cargo and the T0-terminator and directed counter-clockwise in order to not interfere with the transcription of the cargo. SEVA-vectors are named according to their features in a number-based code, see SEVA 2.0 for a comprehensive description<sup>6</sup>. We suggest the naming of *final vectors* to be based on the SEVA-standard, in which the *E. coli* features are specified in three digits: the first digit describes the *E. coli* resistance marker, e.g. 1 for ampicillin resistance or 2 for kanamycin. The second digit indicates the ori, e.g.



**Figure 2.** Assembly of a *Bacillus* SEVA sibling pBS. (a) Collection of *entry* parts needed for the assembly of a pBS vector: one *cargo* vector, one *destination* vector, one *up* and one *down* flanking homology fragment. The latter two are depicted as PCR fragments, but can also be located on a vector. Each of the desired fragments is flanked by IIS-restriction sites where the recognition site (R) is located outside the desired fragment. The compatibility of the resulting overhangs is indicated with letters and a color gradient, e.g. E1 and E2 overhangs can anneal. (b) Intermediate stage of the Golden Gate assembly, showing the desired fragments and some of the by-products (grey). (c) Creation of the *final* vector, including some possible by-products (grey). Only the *destination* vector and the *final* vector carry the ampicillin resistance marker and will be selected for after transformation of the reaction mix into *E. coli*. The *destination* vector will be counter-selected by a red/white screen based on an mRFP-marker.

4 for pRO1600/ColE1 or 9 for pBR322/ROP. The third digit encodes the cargo, e.g. 1 for the default MCS or 3 for the *lacZ* $\alpha$ -pUC18 MCS. pSEVA243 consequently is a vector mediating kanamycin resistance [2] with a high copy number [4 = pRO1600/ColE1] that carries a MCS for blue/white screening [3 = *lacZ* $\alpha$ -pUC18]. *Bacillus*-specific features should be added behind, so pBS143K-amyE is a vector with an *E. coli* ampicillin resistance marker [1], pRO1600/ColE1 ori [4] and *lacZ* $\alpha$ -pUC18 MCS [3] that carries a kanamycin *Bacillus* resistance marker [K] and integrates into the *amyE*-locus [amyE].

We conceptualized the assembly so that SEVA-parts can be exchanged before or after the assembly of the *final* vectors, e.g. to accommodate the presence of either SEVA- or Golden Gate-forbidden restriction sites.

**Golden Gate shuffling to assemble *Bacillus* SEVA siblings with type IIS restriction enzymes.** In our current set-up, four different fragments are needed to assemble the *final* vector: the *up* and *down* homology fragments, the *cargo* with the *Bacillus* resistance cassette and the part for replication and selection in *E. coli*, which we call *destination* vector. These *entry* parts can be combined using Golden Gate assembly as detailed in Fig. 2. The *up* and *down* homology fragments can be used either as PCR products (as depicted) or as cargo of the specialized vectors containing the *up* (pSEVA243X) and *down* (pSEVA243Y) fragments, respectively. Each *entry* part is flanked by recognition sites for a type IIS restriction endonuclease, which creates overhangs that allow for the directional assembly of all *entry* parts. Compatible overhangs are named with the same capital letter in Fig. 2a.

As necessary for Golden Gate assembly, the recognition sites are located “outside” of the part desired for the assembly, so that correctly assembled parts cannot be re-cleaved – in contrast to the re-ligation products of *entry* vectors. By this means, assembly of mostly correct *final* vectors is ensured. The *destination* vector carries a different antibiotic marker than all other *entry* vectors to select for vectors carrying the correct backbone. An mRFP1-cassette present on the original *destination* vector is used for red/white screening. This part is removed during assembly of the *final* vector so that colonies carrying the original vector appear red and colonies carrying the correct *final* vector appear white.

Classic Golden Gate assembly uses BsaI and BbsI (=BpiI) as type IIS restriction enzymes<sup>23,28</sup>, but BsmBI and BtgZI were recently found to also be suitable<sup>29,30</sup>. Our *Bacillus* SEVA siblings toolbox accommodates all four of them for assembly to circumvent compatibility issues with the desired genomic region, e.g. the presence of one or more recognition sites in the *up* and *down* fragments. In addition to BsaI, BbsI, BsmBI and BtgZI, we also included a fifth restriction enzyme (AarI) – not reported previously for its use in Golden Gate assembly.

**Golden Gate assembly restriction sites are arranged in special MCS-IIS.** All four *entry* parts are flanked by MCSs that contain recognition sites for all five type IIS restriction enzymes (MCS-IIS). Inside each MCS-IIS, recognition and restriction sites are designed so that the same overhang sequence is created, independent of the enzyme used. The overhangs are non-palindromic and differ in at least two nucleotides to ensure the correct assembly of the desired vector. For ease of understanding, fusion sites were named with capital letters B, C, E, and F as indicated in Fig. 2. Figure 3 shows the annotated MCS-IIS C2, in which all five enzymes create the overhang GCGA. For the detailed sequence of all 8 MCS-IIS, see Fig. S1.

To compare the assembly efficiency as a function of the enzyme of choice, we used one set of *entry* vectors to construct *final* vectors with identical features. However, nucleotide sequences differ at the assembly scar sites, due to the MCS-IIS. We used pSEVA23X-amyE (*up*), pSEVA243Y-amyE (*down*), pBSc243M and pBSd141R as *entry*

BGSC <sup>†</sup>	Name <sup>‡</sup>	Description <sup>‡</sup>	Resistance in <i>E. coli</i> / <i>B. subtilis</i>	Source
<b>Vectors for default assembly</b>				
<b>Destination vectors</b>				
ECE701	pBSd141R	<i>mRFP1</i> , MCS-IIS F2, <i>bla</i> , ori pRO1600/ColE1, MCS-IIS B1	Amp <sup>r</sup> /—	This study
ECE702	pBSd191R	<i>mRFP1</i> , MCS-IIS F2, <i>bla</i> , ori pBR322/ROP, MCS-IIS B1	Amp <sup>r</sup> /—	This study
<b>Vectors for flanking homology regions<sup>§</sup></b>				
	pSEVA243	<i>lacZα</i> -pUC18 MCS, <i>neo</i> , ori pRO1600/ColE1	Kan <sup>r</sup> /—	SEVA <sup>7</sup>
ECE703	pSEVA243X	<i>lacZα</i> ** <sup>¶</sup> -pUC18 MCS incl. MCS-IIS B2 + C1 for <i>up</i> , <i>neo</i> , ori pRO1600/ColE1	Kan <sup>r</sup> /—	This study
ECE704	pSEVA243Y	<i>lacZα</i> ** <sup>¶</sup> -pUC18 MCS incl. MCS-IIS E2 + F1 for <i>down</i> , <i>neo</i> , ori pRO1600/ColE1	Kan <sup>r</sup> /—	This study
<b>Cargo-Resistance vectors</b>				
ECE706	pBSc241B	MCS-default, MCS-IIS E1, <i>neo</i> , ori pRO1600/ColE1, MCS-IIS C2, <i>bleO</i>	Kan <sup>r</sup> /ble <sup>r</sup>	This study
ECE707	pBSc241C	MCS-default, MCS-IIS E1, <i>neo</i> , ori pRO1600/ColE1, MCS-IIS C2, <i>cat</i>	Kan <sup>r</sup> /cm <sup>r</sup>	This study
ECE708	pBSc241M	MCS-default, MCS-IIS E1, <i>neo</i> , ori pRO1600/ColE1, MCS-IIS C2, <i>ermC</i>	Kan <sup>r</sup> /MLS <sup>r</sup>	This study
ECE709	pBSc241S	MCS-default, MCS-IIS E1, <i>neo</i> , ori pRO1600/ColE1, MCS-IIS C2, <i>aad(9)</i>	Kan <sup>r</sup> /spc <sup>r</sup>	This study
ECE710	pBSc241T	MCS-default, MCS-IIS E1, <i>neo</i> , ori pRO1600/ColE1, MCS-IIS C2, <i>tetL</i>	Kan <sup>r</sup> /tet <sup>r</sup>	This study
ECE711	pBSc241Z	MCS-default, MCS-IIS E1, <i>neo</i> , ori pRO1600/ColE1, MCS-IIS C2, <i>ble-Sh</i>	Kan <sup>r</sup> /zeo <sup>r</sup>	This study
ECE720	pBSc291K	MCS-default, MCS-IIS E1, <i>neo</i> , ori pBR322/ROP, MCS-IIS C2, <i>aph(3)IIIa</i>	Kan <sup>r</sup> /kan <sup>r</sup>	This study
ECE713	pBSc243B	<i>lacZα</i> * <sup>¶</sup> -pUC18 MCS, MCS-IIS E1, <i>neo</i> , ori pRO1600/ColE1, MCS-IIS C2, <i>bleO</i>	Kan <sup>r</sup> /ble <sup>r</sup>	This study
ECE714	pBSc243C	<i>lacZα</i> * <sup>¶</sup> -pUC18 MCS, MCS-IIS E1, <i>neo</i> , ori pRO1600/ColE1, MCS-IIS C2, <i>cat</i>	Kan <sup>r</sup> /cm <sup>r</sup>	This study
ECE715	pBSc243M	<i>lacZα</i> * <sup>¶</sup> -pUC18 MCS, MCS-IIS E1, <i>neo</i> , ori pRO1600/ColE1, MCS-IIS C2, <i>ermC</i>	Kan <sup>r</sup> /MLS <sup>r</sup>	This study
ECE716	pBSc243S	<i>lacZα</i> * <sup>¶</sup> -pUC18 MCS, MCS-IIS E1, <i>neo</i> , ori pRO1600/ColE1, MCS-IIS C2, <i>aad(9)</i>	Kan <sup>r</sup> /spc <sup>r</sup>	This study
ECE717	pBSc243T	<i>lacZα</i> * <sup>¶</sup> -pUC18 MCS, MCS-IIS E1, <i>neo</i> , ori pRO1600/ColE1, MCS-IIS C2, <i>tetL</i>	Kan <sup>r</sup> /tet <sup>r</sup>	This study
ECE718	pBSc243Z	<i>lacZα</i> * <sup>¶</sup> -pUC18 MCS, MCS-IIS E1, <i>neo</i> , ori pRO1600/ColE1, MCS-IIS C2, <i>ble-Sh</i>	Kan <sup>r</sup> /zeo <sup>r</sup>	This study
ECE721	pBSc293K	<i>lacZα</i> * <sup>¶</sup> -pUC18 MCS, MCS-IIS E1, <i>neo</i> , ori pBR322/ROP, MCS-IIS C2, <i>aph(3)IIIa</i>	Kan <sup>r</sup> /kan <sup>r</sup>	This study
<b>Vectors for customizable resistance markers</b>				
ECE705	pBSc241	MCS-default, MCS-IIS E1, <i>neo</i> , ori pRO1600/ColE1, MCS-IIS C2	Kan <sup>r</sup> /—	This study
ECE712	pBSc243	<i>lacZα</i> * <sup>¶</sup> -pUC18 MCS, MCS-IIS E1, <i>neo</i> , ori pRO1600/ColE1, MCS-IIS C2	Kan <sup>r</sup> /—	This study
ECE719	pBSc291	MCS-default, MCS-IIS E1, <i>neo</i> , ori pBR322/ROP, MCS-IIS C2	Kan <sup>r</sup> /—	This study
ECE725	pBSc293	<i>lacZα</i> * <sup>¶</sup> -pUC18 MCS, MCS-IIS E1, <i>neo</i> , ori pBR322/ROP, MCS-IIS C2	Kan <sup>r</sup> /—	This study
ECE722	pBSc391	MCS-default, MCS-IIS E1, <i>cat</i> , ori pBR322/ROP, MCS-IIS C2	Cm <sup>r</sup> /—	This study
ECE726	pBSc393	<i>lacZα</i> * <sup>¶</sup> -pUC18 MCS, MCS-IIS E1, <i>cat</i> , ori pBR322/ROP, MCS-IIS C2	Cm <sup>r</sup> /—	This study
<b>Plasmids for activity-test</b>				
	pSEVA243X-amyE	pSEVA243X-derivative carrying a 550 bp <i>amyE up</i> fragment	Kan <sup>r</sup> /—	This study
	pSEVA243Y-amyE	pSEVA243Y-derivative carrying a 580 bp <i>amyE down</i> fragment	Kan <sup>r</sup> /—	This study
	pBSc241M- <i>P<sub>xyIA</sub>-mkate2</i> (2195)	pBSc241M-derivative carrying <i>mkate2</i> under the control of xylose-inducible <i>P<sub>xyIA</sub></i>	Kan <sup>r</sup> /—	This study
	pBS141M-amyE- <i>mkate2</i>	<i>P<sub>xyIA</sub>-mkate2</i> , <i>amyE</i> , <i>bla</i> , ori pRO1600/ColE1, <i>amyE</i> , <i>erm</i>	Amp <sup>r</sup> /MLS <sup>r</sup>	This study
	pBS141M-ypqP- <i>mkate2</i>	<i>P<sub>xyIA</sub>-mkate2</i> , <i>ypqP</i> , <i>bla</i> , ori pRO1600/ColE1, <i>ypqP</i> , <i>erm</i>	Amp <sup>r</sup> /MLS <sup>r</sup>	This study
	pBS141M-ykoS- <i>mkate2</i>	<i>P<sub>xyIA</sub>-mkate2</i> , <i>ykoS</i> , <i>bla</i> , ori pRO1600/ColE1, <i>ykoS</i> , <i>erm</i>	Amp <sup>r</sup> /MLS <sup>r</sup>	This study
Continued				

BGSC <sup>†</sup>	Name <sup>‡</sup>	Description <sup>‡</sup>	Resistance in <i>E. coli</i> / <i>B. subtilis</i>	Source
pBS191M-ndk_mkate2		<i>P<sub>xyIA</sub>-mkate2, ndk, bla, ori pBR322/ROP, ndk', erm</i>	Amp <sup>r</sup> /MLS <sup>r</sup>	This study
pBS141M-thrC_mkate2		<i>P<sub>xyIA</sub>-mkate2, thrC, bla, ori pRO1600/ColE1, thrC', erm</i>	Amp <sup>r</sup> /MLS <sup>r</sup>	This study
<b>Vectors used for vector construction (templates for PCR reactions)</b>				
pDG148		<i>bla, kan, ble/phle, P<sub>spac</sub></i>	Amp <sup>r</sup> /kan <sup>r</sup>	37,40
pDG780		pBluescriptKS+, <i>kan</i>	Kan <sup>r</sup> /kan <sup>r</sup>	41
pBS3Clux		pAH328 derivative; <i>sacA'...sacA, luxABCDE, bla, cat</i>	Amp <sup>r</sup> /cm <sup>r</sup>	17
pBS4S		pDG1731 derivative; <i>thrC'...thrC, 'hom, thrB', spc, bla</i>	Amp <sup>r</sup> /spc <sup>r</sup>	17
pSB1A3-mkate-B0014		<i>bla, Bba_K823051 (Bsu codon-adapted red fluorescing protein mkate2, terminator B0014), pMB1 ori (high copy number)</i>	Amp <sup>r</sup> /—	Lab stock

**Table 1.** Vectors of the *Bacillus* SEVA siblings toolbox. <sup>†</sup>*Bacillus* Genetic Stock Center (BGSC, <http://www.bgsc.org/>). <sup>‡</sup>The vector names act as identifiers for the SEVA or SEVA siblings collection. **p**, Plasmid. **BS**, *Bacillus* SEVA sibling. **d**, destination vector. **c**, cargo vector. **f**, final vector. Numbers according to SEVA standard: 1<sup>st</sup> position, resistance marker (1, amp. 2, kan). 2<sup>nd</sup> position, origin of replication (4, pRO1600/ColE1, a narrow-host-range ori with high copy number in *E. coli* and varying copy number in *Pseudomonas aeruginosa* and close relatives<sup>7</sup>). 9, pBR322/ROP (medium copy number ori in *E. coli* and few other bacteria<sup>42</sup>). 3<sup>rd</sup> position, cargo (1, MCS default. 3, *lacZα*-pUC18 MCS which allows for blue-white screening with X-Gal). <sup>‡</sup>*lacZα*\*, premature stop codon (C202A, Q68Stop). *lacZα*\*\*\*, premature stop codon (pSEVA243X: G390A, W130Stop, pSEVA243Y: G399A, W133Stop). Both variants are still suitable for blue-white screening. Genes encoding antibiotic resistance markers are explained in detail in Table 3. <sup>§</sup>Flanking homology regions can be stored in those vectors. If the PCR fragment contains the restriction sites needed for assembly, it can be ligated blunt end into pSEVA243 via SmaI. If restriction sites should be added for all enzymes, the PCR fragments can be ligated blunt end into pSEVA243X (for up) or pSEVA243Y (for down) via EcoRV. In this case, the correct orientation needs to be verified by sequencing.

Enzyme	% Blue <sup>§</sup>	% White	% Red	Total colonies	Corr. test digest <sup>‡</sup>	Corr. Sequencing	<i>B. subtilis</i> transf. <sup>§</sup>
AarI	36.8 (±26.1)	0.9 (±0.9)	62.2 (±25.5)	3017 (±1292)	14/18	3/3	++ 4/4
BbsI	41.5 (±34.2)	0.7 (±0.1)	57.7 (±34.1)	3568 (±1084)	15/18	3/5	++ 4/4
BsaI	52.8 (±13.1)	6.1 (±2.2)	41.1 (±13.9)	4004 (±1443)	15/18	3/3	++ 4/4
BsmBI	49.8 (±11.0)	3.7 (±0.5)	46.5 (±10.9)	2000 (±412)	13/18	3/3	++ 4/4
BtgZI	0	0	100	182	0/6	0/3	n.a. n.a.

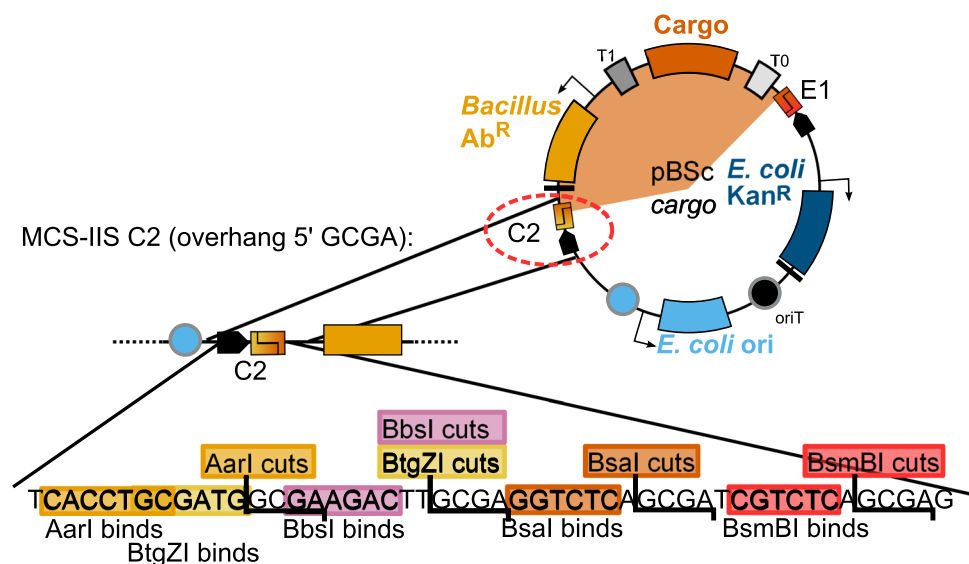
**Table 2.** Assembly efficiencies<sup>‡</sup> depending on restriction enzyme. <sup>†</sup>*Entry* parts: pBSc243M, pBSd141R, pSEVA243X-amyE (up) and pSEVA243Y-amyE (down). <sup>§</sup>Colonies appear blue if carrying the correct final vector and red if the destination vector is unchanged. If available, only blue colonies were used for test digest and correct test digests were used for sequencing. Data is shown for 3 independent Golden Gate assemblies with AarI, BbsI, BsaI or BsmBI. The total number of colonies as well as their apparent color is shown as average and standard deviation. BtgZI did not lead to correct assemblies of final vector and data is derived from a single experiment only. <sup>‡</sup>Test digests, sequencing and transformation results are given in total. <sup>§</sup>++Indicates > 1000 colonies per 100 µl of *B. subtilis* W168 transformation mixture. The correct insertion locus was verified with a starch test.

vectors. In this case, the cargo carries a *lacZα*-fragment that allows for blue-white screening and colonies carrying the correct vector appear blue in the presence of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal). Red colonies carry the original destination vector and white colonies an incorrect vector. We checked colony color, test digest and sequencing results and found good assembly efficiencies for four enzymes: AarI, BbsI, BsaI and BsmBI (Table 2). For BtgZI however, we failed in finding conditions allowing the correct assembly of the final vector. This was surprising, since its use in Golden Gate assembly has been described previously in combination with BsmBI<sup>30</sup>. Even if the desired fragments were digested and gel purified separately, the ligation yielded no correct vector. But in principle, BtgZI can be used for assembly once suitable conditions are found. In contrast, BsaI, BpiI and BsmBI were particularly well-suited with efficiencies of >95% in routinely vector assembly (data not shown). The optimized conditions we used for each enzyme are given briefly in the Methods section and are described in more detail in the Supplemental Text S1.

Taken together, *Bacillus* SEVA siblings vectors can be efficiently assembled using one of four type IIS restriction endonucleases. In addition to the established enzymes (BsaI, BbsI and BsmBI), AarI was also found to be suitable for Golden Gate assembly. Because of its 7 bp recognition site, it should be found less frequently in genomic sequences compared to the “classical” enzymes with a 6 bp recognition site.

Abbr.	Gene	Description*	Antibiotic	Conc. [ $\mu\text{g ml}^{-1}$ ]	Source	
	amp <sup>r</sup>	<i>bla</i>	$\beta$ -lactamase (Eco)	Ampicillin	100	pSEVA143 <sup>7</sup>
	kan <sup>r</sup>	<i>neo</i>	neomycin-kanamycin phosphotransferase type I (Eco)	Kanamycin	50	pSEVA241 <sup>7</sup>
	cm <sup>r</sup>	<i>cat</i>	chloramphenicol <i>O</i> -acetyltransferase (Eco)	Chloramphenicol	35 <sup>5</sup>	pSEVA341 <sup>7</sup>
B	ble <sup>r</sup>	<i>bleO</i> <sup>#</sup>	bleomycin binding protein (phleomycin D)	Phleomycin D1	100	pDG148 <sup>40</sup>
C	cm <sup>r</sup>	<i>cat</i>	chloramphenicol <i>O</i> -acetyltransferase	Chloramphenicol	5	pBS3Clux <sup>17,19</sup>
K	kan <sup>r</sup>	<i>aph(3')IIIa</i>	aminoglycoside <i>O</i> -phosphotransferase APH(3')-IIIa	Kanamycin	10	pDG780 <sup>41</sup>
M	MLS <sup>r</sup>	<i>ermC</i>	23 S rRNA (adenine(2058)-N(6))-methyltransferase	Erythromycin & Lincomycin	1 & 25	pDG647 <sup>41</sup>
S	spc <sup>r</sup>	<i>aad(9)</i>	aminoglycoside nucleotidyltransferase ANT9	Spectinomycin	200	pBS4S <sup>16,17</sup>
T	tet <sup>r</sup>	<i>tetL</i>	tetracycline efflux MFS transporter	Tetracycline	12.5	pDG1513 <sup>41</sup>
Z	zeo <sup>r</sup>	<i>ble-Sh</i> <sup>#</sup>	phleomycin/bleomycin binding protein (codon-optimized for Bsu)	Zeocin	100	This study

**Table 3.** Description of antibiotic markers used in this study. \*Eco, *E. coli*; Bsu, *B. subtilis*. <sup>5</sup>5  $\mu\text{g ml}^{-1}$  were used for medium copy number vectors. <sup>#</sup>*bleO* and *ble-Sh* both mediate resistance against phleomycin or zeocin (both from the bleomycin family) by binding to the antibiotic, respectively. There are differences in amino acid sequence of the encoded proteins and in the properties of the antibiotics, but mediation of cross-resistance cannot be excluded.



**Figure 3.** Architecture of the MCS-IIS C2. This DNA-sequence is located on the cargo vector between the *E. coli* ori and the *Bacillus* antibiotic marker. The recognition sites for five type IIS restriction enzymes (AarI, BtgZI, BbsI, BsaI, BsmBI), each designed to create a 5' GCGA-overhang are encoded on the DNA stretch. Architecture of all MCS-IIS can be found in Fig. S1.

**Our collection of entry parts for the assembly of *Bacillus* SEVA siblings.** For the assembly of *Bacillus* SEVA siblings vectors, four different categories of entry parts are needed: cargo and resistance, up, down, and the replication part. They will be described in more detail below. The entry vectors offered with our toolbox are depicted in Fig. 4 and listed with detailed descriptions in Table 1.

**pBSc: Cargo and resistance.** Since for most experiments the same cargo will usually be combined with the same resistance marker, both are located on the cargo vector pBSc. As a result, only four instead of five parts have to be assembled, thereby increasing cloning efficiency. For our collection of cargo vectors, we used well-established antibiotic resistance markers to enable selection on phleomycin D, chloramphenicol, kanamycin, macrolide and streptogramin B antibiotics (MLS), spectinomycin, tetracycline, and zeocin (Table 3). Forbidden restriction sites were removed and transcriptional terminators were added where necessary. Transcription occurs in the opposite direction to the cargo. Per default, the cargo vector uses the high copy number ori pRO1600/ColE1. For the kanamycin resistance marker, this vector was unstable, so the medium copy number ori pBR322/ROP was used. All cargo vectors are offered with the default SEVA MCS (pUC18-related polylinker without *lacZ $\alpha$* ,<sup>7</sup>) or the *lacZ $\alpha$* -pUC18 MCS (Table 1). The cargo of choice can be inserted into the cargo vector or into the final vector, depending on enzyme compatibility and cloning strategy.

**Up and down flanking homology regions.** The choice of *up* and *down* flanking homology regions depends on and needs to be adjusted to the *Bacillus* strain and experiment. The regions need specific overhangs for the subsequent assembly. This can be achieved by either of the following two strategies: (i) PCR-amplified fragments of choice (without overhangs) are cloned blunt end into the EcoRV-linearized pSEVA243X (*up*) or pSEVA243Y (*down*), respectively. As a result, the insert receives the matching MCS-IIS encoded on these vectors which can be used for Golden Gate assembly. (ii) Incorporation of one restriction site to the primers will directly allow the use of the PCR product for Golden Gate assembly.

**pBSd: replication part in *E. coli*.** The *destination vector* carries the *replication part* of the *final vector*. Its resistance (ampicillin marker) differs from all other *entry vectors* to allow selection for the correct backbone after vector assembly. A high copy number (pBSd141R) and medium copy number version (pBSd191R) are provided. Both carry an mRFP cassette in their default MCS, thereby allowing red/white screening: After Golden Gate assembly, colonies with the original *destination vector* will appear red and therefore can be discarded. If different *E. coli* features are required for the *final vector*, they can be exchanged in the *destination vector* or the *final vector*, depending on needs and enzyme compatibilities.

**Assembly, transformation and integration of pBS vectors.** The reaction for the vector assembly contains the *entry parts*, one of the type IIS restriction enzymes, ligase and buffer. The specific protocols depend on the enzyme used and can be found in the Materials and Method section. Competent *E. coli* cells are subsequently transformed with the reaction mix and plated on selective media containing ampicillin (and X-Gal in case of blue/white screening). Verification of the *final vector* can be achieved in a two-step procedure using test digests followed by sequencing. For the latter, primers TM3782 and TM3783 (if pBSd141R was used as *destination vector*) or TM3783 and TM5128 (pBSd191R) are recommended. The sequencing results should cover the *up* and *down* region as well as the adjacent assembly scars to ensure correct assembly of all parts. We tested all *entry parts* of the toolbox (Table 1) for their assembly efficiencies (Tables S4–S6) and found that in all cases it was more than sufficient to test six colonies of the correct color to obtain a correct *final vector*. If the high copy number *destination vector* is used, or selection for the correct cargo is possible, e.g. via blue/white screening, assembly efficiencies are even higher.

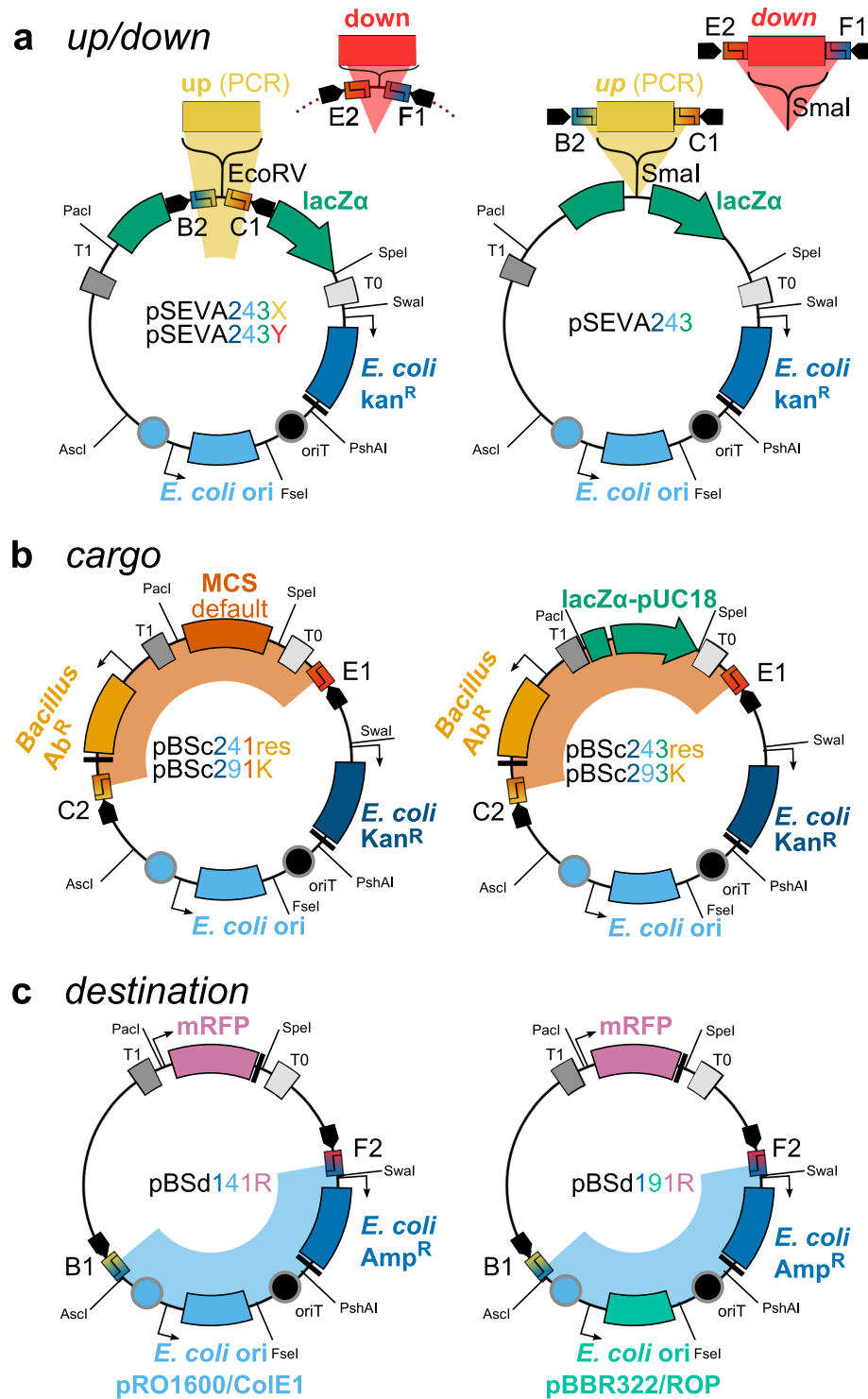
Prior to transformation of *B. subtilis* W168, all vectors have to be linearized, e.g. using ApaI, to ensure chromosomal integration by double homologous recombination. All transformations were successful – except for vectors carrying the tetracycline resistance marker where only for one of two loci could be targeted (Tables S4, S5). The transformation efficiencies depended on the selection marker (Tables 2, 3 and S4–S6). Correct integration was verified by colony PCR or physiological tests; e.g. a starch test for integration into the *amyE* locus.

As a proof of concept, we analyzed the expression of the reporter gene *mkate2* under control of the xylose-inducible promoter  $P_{xylA}$  at five loci spread across the chromosome (Table 4): *amyE*, *ykoS*, *ypqP* (prophage SP $\beta$ ), *nkd*, and *thrC* at positions 28, 111, 183–195, 203, and 283° on the circular chromosome, respectively. *amyE* and *thrC* are early-discovered and frequently used integration loci close to the ori, encoding a starch-hydrolyzing alpha-amylase and the threonine synthase, respectively. The 130 kb-large prophage SP $\beta$ , which itself is inserted at the *ypqP* locus, is known to be not essential and was targeted to demonstrate the possibility of deleting large genomic regions. The two remaining genes, *nkd* and *ykoS*, encode a nucleoside diphosphate kinase and a gene of unknown function, respectively. Those non-essential genes were chosen based on their chromosomal location. Since the same reporter was used in all vector constructs, the reporter-cargo was added at the *entry vector* level to pBSc241M, resulting in pBSc241M- $P_{xylA}$ -*mkate2*. pBSd141R (ori pRO/ColE1, high copy number) was chosen as the *destination vector* and primers were designed for PCR-product assembly of *up* and *down* fragments via BsaI. Assembly and verification of four *final vectors* was achieved within five working days. The vector carrying homology regions for integration into *nkd* was unstable (small colonies), necessitating a change to pBSd191R (ori pBR322/ROP, medium copy number) as a *destination vector*. After transformation of *B. subtilis* W168, the resulting integrants were verified by starch or threonine auxotrophy tests (*amyE* and *thrC*, respectively), or colony PCR for the remaining loci (Table 4). The xylose-inducible promoter  $P_{xylA}$  was fully induced with xylose in exponentially growing cultures. The fluorescence intensity of *B. subtilis* cells was quantified in triplicates in a microtiter plate reader as a measure for mKate2 production (Fig. 5). The results highlight a dependence of the expression levels on the chromosomal location: As demonstrated previously, genes close to the ori (0°/360°) tend to be expressed at higher levels than those close to the termination region (180°). One exception was the *ykoS*-locus, located at the replication termination site, which had a higher reporter activity than the constructs inserted at the ori-proximal *ypqP* or *nkd* sites. Nevertheless, these results are in good agreement with a recent study<sup>31</sup>, thereby demonstrating that our vector toolbox can for example be used to study the effect of chromosome location on expression levels in a simple and straightforward manner.

## Discussion

*B. subtilis* is a versatile heterologous host with powerful genetics allowing precise genomic manipulation. But so far, the available integrative vectors could not easily be adapted to different strains or, more importantly, other *bacilli*, since no modular toolbox was available to allow free combination of resistance markers and integration sites. To fill this gap, we present a rigorously evaluated toolbox for the construction of integrative vectors with customizable flanking homology regions for *Bacillus* sp. The *final* pBS *vector* can be assembled within one week from four *entry parts* via efficient Golden Gate cloning. Our vector suite offers the choice of four type IIS restriction endonucleases (AarI, BbsI, BsaI and BsmBI) for Golden Gate assembly, seven *Bacillus* antibiotic resistance markers, and two different MCSs, available on either a high or medium copy number *E. coli* backbone. The toolbox is adjusted to and widely compatible with the *E. coli* SEVA standard to allow reusability of its cargos or vector

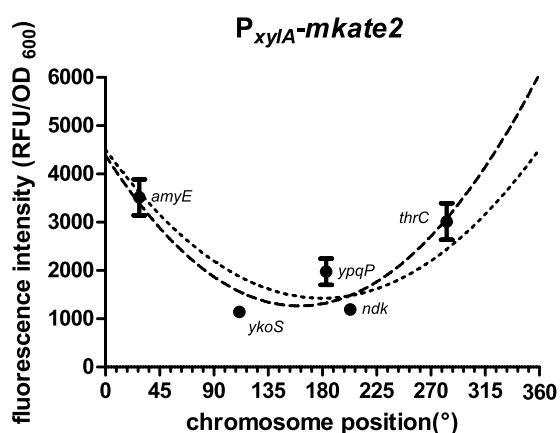




**Figure 4.** Vector suite for the generation of *Bacillus* SEVA siblings. Schematic representation of the vector architectures, details are listed in Table 1. (a) Vectors for flanking homology regions. *Up* fragments (PCR product) can be stored in pSEVA243X and *down* in pSEVA243Y, each linearized with *EcoRV*. The respective MCS-IIS are encoded on the vectors. If required restriction site are already encoded on the primer overhangs, fragments can be stored in pSEVA243 or used directly for Golden Gate assembly. (b) *Cargo* vectors carry one of the following *Bacillus* antibiotic markers: Ble, Cat, Kan, MLS, Spc, Tet, Zeo and either the default MCS (pBSc241res) or the *lacZα*\*-pUC18 MCS for blue/white-screening. Vectors carrying the *Bacillus* kanamycin resistance marker utilize the medium copy number pBR322/ROP all, others the high copy number pRO1600/ColE1. Backbones are also available without *Bacillus* marker to allow insertion of a new or customized marker. (c) *Destination* vectors carry an mRFP1-cassette as cargo for red/white screening and an ampicillin resistance marker for selection in *E. coli*. They are available with high (pRO1600/ColE1) or medium (pBR322/ROP) copy number origins of replication.

Locus	% Light red <sup>§</sup>	% White	% Red	Total colonies	Correct test digest	Correct Sequencing	<i>B. subtilis</i> transf. <sup>#</sup>	
<i>amyE</i>	69.8	19.8	10.3	1160	6/6	1/1	+	4/4
<i>ypqP</i>	92.2	2.9	4.9	1236	6/6	1/1	+	4/4
<i>ykoS</i>	91.6	3.2	5.3	1900	6/6	1/1	++	4/4
<i>ndk</i> <sup>&amp;</sup>		75	25	496	12/12	1/1	++	4/4
<i>thrC</i>	92.1	5.2	2.6	4580	6/6	1/1	+	4/4

**Table 4.** Assembly efficiencies\* and reporter activity depending on integration sites. \*Data is shown for one Golden Gate assembly using BsaI and the following *entry* parts: PCR products *up* and *down* with BsaI-overhangs, pBSd141R and pBSc141M\_P<sub>*xylA*</sub>-*mkate2*. <sup>§</sup>Colonies appear light red if carrying the correct *final vector* and red if the *destination vector* is unchanged. Only light red colonies were used for test digest and correct test digests were used for sequencing. The correct chromosomal integration was verified as described in Material and Methods. <sup>#</sup>Number of colonies per 100  $\mu$ l of transformation mixture: ++, >1000; +, >100 and number of colonies with verified chromosomal integration from number of tested colonies. <sup>&</sup>pBSd191R was used as *destination vector*. No colonies of “light red” color could be identified due to lower copy number of the *final vector*, in comparison to the other constructs. As a consequence, white colonies were used for further verification.



**Figure 5.** Maximal promoter activity of P<sub>*xylA*</sub> depends on chromosomal location. The reporter construct P<sub>*xylA*</sub>-*mkate2* was integrated into the *B. subtilis* genome at five different chromosomal loci, as indicated. mKate expression was maximally induced with the addition of 0.2% xylose and fluorescence intensity was measured as an indicator for mKate abundance. The fluorescence intensity is given as a function of the chromosome position. The error bars show the standard deviation of three independent biological replicates. The data was fitted to two second order polynomial functions: dashed line: no constrains,  $R^2 = 0.85$ ; dotted line: minimum was set to  $X = 180$ ,  $R^2 = 0.72$ .

parts. All components provided in this toolbox were successfully tested for assembly efficiency, functionality, and usability in *B. subtilis*. The supplementary assembly guide (Text S1) provides necessary information to facilitate a fast and efficient cloning process.

In the course of developing our vectors, a few combinations were discovered which were difficult to handle on high copy number vectors (ori pRO1600/ColE1): (i) The *Bacillus* kanamycin resistance marker was prone to mutations and is therefore offered with the medium copy number ori pBR322/ROP. (ii) The *lacZ* $\alpha$ \*-pUC18 MCS differs from the original *lacZ* $\alpha$ -pUC18 MCS in a single-nucleotide polymorphism causing a premature stop codon in *lacZ* $\alpha$  that does not impair blue/white screening. There was a strong selection pressure against the original *lacZ* $\alpha$ -pUC18 MCS (very small colonies), which caused transposon integrations and single nucleotide polymorphisms in *lacZ* $\alpha$ . Consequently, we used the mutated but fully functional *lacZ* $\alpha$ \*-pUC18 MCS for our constructs. (iii) The *final ndk*-vectors only contained the *ndk down* fragment when using the high copy number pBSd141R as *destination vector*. However, assembly with the medium copy number vector pBSd191R was successful in the first attempt.

Based on our experience, the use of a medium copy number vector is highly recommended as the first trouble shooting strategy in case of cloning issues, especially if slowly growing colonies occur. Also, low copy number oris (e.g. p15A and pSC101) or different directionality of the inserts could be used in case toxicity issues occur, which result in genetic instabilities even on medium copy number vectors.

Furthermore, the tetracycline resistance marker could not confer resistance in *B. subtilis* W168 when the *amyE* locus was targeted. However, when using a different integration locus (*lacA*), the transformation was successful. Due to the dependence of expression levels on the chromosomal region, we suspect this to be causing the transformation issues.

The *Bacillus* SEVA siblings toolbox was developed to provide a versatile starting point for the efficient construction of personalized, yet standardized vectors. Both, the *entry* parts as well as *final vectors* can be customized to meet personal needs. The MCS or cargo and ori can easily be exchanged according to the SEVA standard and new or modified resistance markers can be inserted in markerless *cargo vectors* via AscI and MluI, e.g. markers flanked with target sites for recombinases. These systems would allow directed, recombinase-mediated removal of the marker after chromosomal integration (for reviews on recombinase-mediated cassette exchange, e.g. by Flp and Cre/*loxP*, see<sup>32,33</sup>).

It should also be pointed out that the free choice of chromosomal integration loci provided by our vector toolbox allows for replacing even larger (non-essential) chromosomal areas directly in the process of plasmid integration. We have demonstrated this possibility by integrating an *mkate2*-expressing plasmid into the *ypqP* locus of the *B. subtilis* chromosome, thereby deleting the 130 kb prophage SP $\beta$  (Table 4). This combined integration/deletion at any desired chromosomal position will greatly enhance the possibilities in genetically manipulating *B. subtilis*.

The current resistance markers located on *cargo vectors* were tested for their functionality only in *B. subtilis* W168. Since they originate from broad host range vectors, they are expected to be functionally expressed in many low G + C Gram-positive bacteria (phylum *Firmicutes*). The *Bacillus* SEVA siblings might therefore be suitable for e.g. other bacilli, increasing the number of species where customized vectors can be used for genetic manipulation. Indeed, preliminary results from an ongoing study using our vectors reported similar or even better vector assembly efficiencies as those described in Table 2. Choramphenicol- or erythromycin-resistant mutants were readily obtained in *Paenibacillus polymyxa* (order: *Bacillales*, family: *Paenibacillaceae*) (Christoph Engl, personal communication).

If transformation efficiencies are too low, conjugation can be performed using the *oriT* already included in the SEVA standard, which provides efficient transfer into various Gram-negative and -positive hosts<sup>7,34,35</sup>. It can be exchanged with a conjugation marker of choice in the *destination* or *final vector* to meet the needs of the target organism. Recombination efficiency of homologous sequences into the chromosome varies more than 30-fold along the *B. subtilis* chromosome<sup>36</sup>, and even more between species. In general, integration rates can be improved by using longer stretches of DNA or a vector with a temperature-sensitive replicon (e.g. based on pMAD<sup>3,37</sup>). Vector replication not only increases the number of vectors per cell and number of cells carrying a vector (by passing it on to the next generation), but also supports the second recombination event during which the vector backbone is excised from the chromosome<sup>3</sup>. A temperature-sensitive origin of replication can therefore be added to the *destination* or *final vector* to improve recombination with the chromosome.

Moreover, the construction logic described for the *Bacillus* SEVA siblings – that is, the restriction enzymes used for the assembly of the final integrative plasmids – can of course also be applied for developing other SEVA-compatible integration vectors for completely unrelated microorganisms in a similar fashion. But this would require adjusting the corresponding resistance cassettes applicable to these bacteria.

Here, we present the first fully modular, yet standardized vector toolkit for integrative vectors in *B. subtilis* and beyond. We hope that the *Bacillus* SEVA siblings vector toolbox will prove to be useful for projects throughout the *Bacillus* world. In case personalized *entry vectors* (*pBSc*, *pBSd*) are created, we would like to encourage sharing them with the *Bacillus* community, e.g. via the SEVA or BGSC collections, to further improve the tools available for genetically manipulating these powerful organisms.

## Material and Methods

**Bacterial strains and growth conditions.** All strains used in this study are listed in Table S1. *B. subtilis* and *E. coli* were routinely grown in lysogeny broth (LB) medium (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl) at 37°C with agitation (220 rpm). Solid media additionally contained 1.5% (w/v) agar. Selective media contained appropriate antibiotics, as provided in Table 3.

**Transformation.** *E. coli* (XL1 blue, Agilent Technologies, Santa Clara, CA, USA) competent cells were prepared and transformed according to the rubidium chloride method<sup>38</sup>, achieving  $\sim 5 \times 10^6$  colony forming units (CFU) per  $\mu\text{g}$  pUC18 DNA. Transformations of *B. subtilis* were carried out as described previously<sup>17,22</sup>. The integration of plasmids into the *B. subtilis* genome was checked on starch plates (*amyE*), with minimal medium lacking threonine (*thrC*) or colony PCR (*lacA*, *ypqP*, *ykoS*, *ndk*). Detailed protocols were published previously<sup>17,39</sup>.

**DNA manipulation.** Vectors and plasmids used in this study are listed in Table 1. General cloning procedure, such as endonuclease restriction digest, ligation and PCR, was performed with enzymes and buffers from New England Biolabs® (NEB; Ipswich, MA, USA) or Thermo Scientific™ (Waltham, MA, USA) according to the respective protocols. Phusion® polymerase was used for PCRs if the resulting fragment was further used, otherwise OneTaq® was the polymerase of choice. PCR-purification was performed with the *HiYield PCR Gel Extraction/PCR Clean-up Kit* (Süd-Laborbedarf GmbH (SLG), Gauting, Germany). Plasmids were prepared using alkaline lysis and subsequent DNA precipitation. All plasmids created during this study are listed in Table 1, their construction is described in supplemental Table S2 and all primer sequences are given in Table S3.

**Golden Gate assemblies of *final vectors*.** Golden Gate assemblies were performed using T4 DNA ligase (30 WU) from Thermo Scientific™ with the accompanied buffer. BsaI, BbsI, BsmBI and BtgZI were purchased from NEB, AarI from Thermo Scientific™. *Entry* parts were diluted to 20 or 40 nM stock concentrations and 2 or 1  $\mu\text{l}$  were used per 15  $\mu\text{l}$  reaction, respectively. For all enzymes, 0.5  $\mu\text{l}$  were used per reaction, except for AarI where 1.5  $\mu\text{l}$  were necessary because of its lower activity. BsaI and AarI are active in ligase buffer, but for AarI the

accompanying oligonucleotides were supplemented for optimal efficiency. For BbsI half of the ligase buffer was replaced by NEBuffer 2.1 and for BsmBI half was replaced by NEBuffer 3.1.

The general assembly protocol was 37 °C, 30 min; 16 °C, 30 min; (37 °C, 3 min; 16 °C, 5 min) × 15; 37 °C, 10 min; 50 °C, 10 min; 80 °C, 10 min. The exception was BsmBI, where 55 °C were required for the first incubation step and ligase was only added afterwards. 7.5 µl of the final reaction were used for *E. coli* transformation.

**Measurement of P<sub>xyIA</sub> activity.** Fluorescence intensity of strains carrying a transcriptional fusion of the xylose-inducible promoter P<sub>xyIA</sub> and the fluorescence reporter *mkate2* were assayed using a Synergy<sup>TM</sup> NEOALPHAB multi-mode microplate reader from BioTek® (Winooski, VT, USA). The reader was controlled using the software Gen5<sup>TM</sup> (version 2.06). Cells were inoculated 1:1000 from fresh overnight cultures and grown to OD<sub>600</sub> ~0.2, treated with 0.2% xylose and grown for two hours. Cells were harvested by centrifugation and resuspended in phosphate buffered saline (137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4). 200 µl per well in 96-well plates (black walls, clear bottom; Greiner Bio-One, Frickenhausen, Germany) were measured for their OD<sub>600</sub> and mKate-fluorescence using the monochromator with following parameters: endpoint measurement, gain: 100, excitation wavelength: 588 nm, emission wavelength: 633 nm.

**Data availability.** The vectors generated in this study are available from the Bacillus Genetic Stock Center (BGSC, <http://www.bgsc.org/>, accession numbers ECE701-26) and the SEVA collection ([http://wwwuser.cnb.csic.es/~seva/?page\\_id=19](http://wwwuser.cnb.csic.es/~seva/?page_id=19)). For sequence information, the following accession numbers apply: GenBank, KY995178 to KY995203; ACS Synthetic Biology Registry (<https://acs-registry.jbei.org/>), JPUB\_008862 to JPUB\_008887).

## References

- Wright, O., Delmans, M., Stan, G. B. & Ellis, T. GeneGuard: A modular plasmid system designed for biosafety. *ACS Synth Biol* **4**, 307–316, <https://doi.org/10.1021/sb500234s> (2015).
- Shimada, T. *et al.* Classification and strength measurement of stationary-phase promoters by use of a newly developed promoter cloning vector. *J Bacteriol* **186**, 7112–7122, <https://doi.org/10.1128/JB.186.21.7112-7122.2004> (2004).
- Arnaud, M., Chastanet, A. & Debarbouille, M. New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl Environ Microbiol* **70**, 6887–6891, <https://doi.org/10.1128/AEM.70.11.6887-6891.2004> (2004).
- Sinah, N., Williams, C. A., Piper, R. C. & Shields, S. B. A set of dual promoter vectors for high throughput cloning, screening, and protein expression in eukaryotic and prokaryotic systems from a single plasmid. *BMC Biotechnol* **12**, 54, <https://doi.org/10.1186/1472-6750-12-54> (2012).
- Andrianantoandro, E., Basu, S., Karig, D. K. & Weiss, R. Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol* **2**(2006), 0028, <https://doi.org/10.1038/msb4100073> (2006).
- Martinez-Garcia, E., Aparicio, T., Goni-Moreno, A., Fraile, S. & de Lorenzo, V. SEVA 2.0: an update of the Standard European Vector Architecture for de-/re-construction of bacterial functionalities. *Nucleic Acids Res* **43**, D1183–1189, <https://doi.org/10.1093/nar/gku1114> (2015).
- Silva-Rocha, R. *et al.* The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res* **41**, D666–675, <https://doi.org/10.1093/nar/gks1119> (2013).
- Standard European Vector Architecture, [http://wwwuser.cnb.csic.es/~seva/?page\\_id=17](http://wwwuser.cnb.csic.es/~seva/?page_id=17) and [http://wwwuser.cnb.csic.es/~seva/?page\\_id=19](http://wwwuser.cnb.csic.es/~seva/?page_id=19) (accessed March 2017)
- Kim, S. H., Cavaleiro, A. M., Rennig, M. & Norholm, M. H. SEVA Linkers: A versatile and automatable DNA backbone exchange standard for synthetic biology. *ACS Synth Biol* **5**, 1177–1181, <https://doi.org/10.1021/acssynbio.5b00257> (2016).
- van Dijk, J. M. & Hecker, M. *Bacillus subtilis*: from soil bacterium to super-secreting cell factory. *Microb Cell Fact* **12**, 3, <https://doi.org/10.1186/1475-2859-12-3> (2013).
- Hohmann, H.-P., van Dijk, J. M., Krishnappa, L. & Prágai, Z. In *Industrial Biotechnology*, 221–297 (Wiley-VCH Verlag GmbH & Co. KGaA, 2017).
- Liu, Y., Li, J., Du, G., Chen, J. & Liu, L. Metabolic engineering of *Bacillus subtilis* fueled by systems biology: Recent advances and future directions. *Biotechnol Adv* **35**, 20–30, <https://doi.org/10.1016/j.biotechadv.2016.11.003> (2017).
- Mell, J. C. & Redfield, R. J. Natural competence and the evolution of DNA uptake specificity. *J Bacteriol* **196**, 1471–1483, <https://doi.org/10.1128/JB.01293-13> (2014).
- Zucca, S., Pasotti, L., Mazzini, G., De Angelis, M. G. & Magni, P. Characterization of an inducible promoter in different DNA copy number conditions. *BMC Bioinformatics* **13**(Suppl 4), S11, <https://doi.org/10.1186/1471-2105-13-S4-S11> (2012).
- Adams, C. W. & Hatfield, G. W. Effects of Promoter Strengths and Growth Conditions on Copy Number of Transcription-Fusion Vectors. *J Biol Chem* **259**, 7399–7403 (1984).
- Guerout-Fleury, A. M., Frandsen, N. & Stragier, P. Plasmids for ectopic integration in *Bacillus subtilis*. *Gene* **180**, 57–61 (1996).
- Radeck, J. *et al.* The *Bacillus* BioBrick Box: generation and evaluation of essential genetic building blocks for standardized work with *Bacillus subtilis*. *J Biol Eng* **7**, 29, <https://doi.org/10.1186/1754-1611-7-29> (2013).
- Hartl, B., Wehr, W., Wiegert, T., Homuth, G. & Schumann, W. Development of a new integration site within the *Bacillus subtilis* chromosome and construction of compatible expression cassettes. *J Bacteriol* **183**, 2696–2699, <https://doi.org/10.1128/JB.183.8.2696-2699.2001> (2001).
- Schmalisch, M. *et al.* Small genes under sporulation control in the *Bacillus subtilis* genome. *J Bacteriol* **192**, 5402–5412, <https://doi.org/10.1128/JB.00534-10> (2010).
- Stülke, J. *et al.* Induction of the *Bacillus subtilis* *ptsGHI* operon by glucose is controlled by a novel antiterminator, GlcT. *Mol Microbiol* **25**, 65–78 (1997).
- Wach, A. PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in *S. cerevisiae*. *Yeast* **12**, 259–265 (1996).
- Harwood, C. R. & Cutting, S. M. *Molecular Biological Methods for Bacillus* (John Wiley & Sons, 1990).
- Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, e3647, <https://doi.org/10.1371/journal.pone.0003647> (2008).
- Engler, C., Gruetzner, R., Kandzia, R. & Marillonnet, S. Golden gate shuffling: a one-pot DNA shuffling method based on type II restriction enzymes. *PLoS One* **4**, e5553, <https://doi.org/10.1371/journal.pone.0005553> (2009).
- Khasanov, F. K., Zvingila, D. J., Zainullin, A. A., Prozorov, A. A. & Bashkurov, V. I. Homologous recombination between plasmid and chromosomal DNA in *Bacillus subtilis* requires approximately 70 bp of homology. *Mol Gen Genet* **234**, 494–497, <https://doi.org/10.1007/bf00538711> (1992).
- Gibson, D. G. Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides. *Nucleic Acids Res* **37**, 6984–6990, <https://doi.org/10.1093/nar/gkp687> (2009).

27. Gibson, D. G., Smith, H. O., Hutchison, C. A. 3rd, Venter, J. C. & Merryman, C. Chemical synthesis of the mouse mitochondrial genome. *Nat Methods* **7**, 901–903, <https://doi.org/10.1038/nmeth.1515> (2010).
28. Weber, E., Engler, C., Gruetzner, R., Werner, S. & Marillonnet, S. A modular cloning system for standardized assembly of multigene constructs. *PLoS ONE* **6**, e16765, <https://doi.org/10.1371/journal.pone.0016765> (2011).
29. Sarrion-Perdigones, A. *et al.* GoldenBraid: An iterative cloning system for standardized assembly of reusable genetic modules. *PLoS ONE* **6**, e21622, <https://doi.org/10.1371/journal.pone.0021622> (2011).
30. Sarrion-Perdigones, A. *et al.* GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. *Plant Physiol* **162**, 1618–1631, <https://doi.org/10.1104/pp.113.217661> (2013).
31. Sauer, C. *et al.* Effect of genome position on heterologous gene expression in *Bacillus subtilis*: An unbiased analysis. *ACS Synthetic Biology* **5**, 942–947, <https://doi.org/10.1021/acssynbio.6b00065> (2016).
32. Turan, S., Zehe, C., Kuehle, J., Qiao, J. & Bode, J. Recombinase-mediated cassette exchange (RMCE) - a rapidly-expanding toolbox for targeted genomic modifications. *Gene* **515**, 1–27, <https://doi.org/10.1016/j.gene.2012.11.016> (2013).
33. Dong, H. & Zhang, D. Current development in genetic engineering strategies of *Bacillus* species. *Microb Cell Fact* **13**, 63, <https://doi.org/10.1186/1475-2859-13-63> (2014).
34. Martinez-Garcia, E., Calles, B., Arevalo-Rodriguez, M. & de Lorenzo, V. pBAM1: an all-synthetic genetic tool for analysis and construction of complex bacterial phenotypes. *BMC Microbiol* **11**, 38, <https://doi.org/10.1186/1471-2180-11-38> (2011).
35. Trieu-Cuot, P., Carlier, C., Martin, P. & Courvalin, P. Plasmid transfer by conjugation from *Escherichia coli* to Gram-positive bacteria. *FEMS Microbiol Lett* **48**, 289–294, <https://doi.org/10.1111/j.1574-6968.1987.tb02558.x> (1987).
36. Vagner, V. & Ehrlich, S. D. Efficiency of homologous DNA recombination varies along the *Bacillus subtilis* chromosome. *J Bacteriol* **170**, 3978–3982 (1988).
37. Rachinger, M. *et al.* Size unlimited markerless deletions by a transconjugative plasmid-system in *Bacillus licheniformis*. *J Biotechnol* **167**, 365–369, <https://doi.org/10.1016/j.jbiotec.2013.07.026> (2013).
38. Mülhardt, C. *Der Experimentator. Molekularbiologie/Genomics*. 6 edn, (Spektrum Akademischer Verlag, 2009).
39. Sambrook, J. & Russell, D. W. *Molecular Cloning - a laboratory manual* (Cold Spring Harbor Laboratory Press, 2001).
40. Joseph, P., Fantino, J. R., Herbaud, M. L. & Denizot, F. Rapid orientated cloning in a shuttle vector allowing modulated gene expression in *Bacillus subtilis*. *FEMS Microbiol Lett* **205**, 91–97 (2001).
41. Guerout-Fleury, A. M., Shazand, K., Frandsen, N. & Stragier, P. Antibiotic-resistance cassettes for *Bacillus subtilis*. *Gene* **167**, 335–336 (1995).
42. Balbás, P. & Bolívar, F. In *Recombinant Gene Expression: Reviews and Protocols* (eds Paulina Balbás & Argelia Lorence) 77–90 (Humana Press, 2004).

## Acknowledgements

We thank Wilfried Meijer for fruitful initial design discussions and Beate Lüdke for assistance with *B. subtilis* transformations. This work was supported by the Deutsche Forschungsgemeinschaft (DFG) priority program SPP1617 ‘Phenotypic Heterogeneity and Sociobiology of Bacterial Populations’ [grant number MA 2837/3-2 to TM]. Funding for open access charge: dito.

## Author Contributions

J.R. and T.M. conceptualized *Bacillus* SEVA siblings. J.R. detailed and planned the experiments, analyzed the data and created the figures. J.R., D.M. and N.L. cloned the vectors and J.R. performed Golden Gate assemblies. J.R. and T.M. wrote the manuscript. All authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-017-14329-5>.

**Competing Interests:** The authors declare that they have no competing interests.

**Change History:** A correction to this article has been published and is linked from the HTML version of this paper. The error has been fixed in the paper.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017