

SEVA 4.0: an update of the Standard European Vector Architecture database for advanced analysis and programming of bacterial phenotypes

Esteban Martínez-García¹, Sofía Fraile¹, Elena Algar¹, Tomás Aparicio¹,
Elena Velázquez¹, Belén Calles¹, Huseyin Tas¹, Blas Blázquez¹, Bruno Martín²,
Clara Prieto², Lucas Sánchez-Sampedro², Morten H.H. Nørholm³, Daniel C. Volke³,
Nicolas T. Wirth³, Pavel Dvořák⁴, Lorea Alejandre⁵, Lewis Grozinger^{5,6},
Matthew Crowther^{5,6}, Angel Goñi-Moreno⁵, Pablo I. Nikel³, Juan Nogales¹ and
Víctor de Lorenzo^{1,*}

¹Systems Biology Department, Centro Nacional de Biotecnología (CNB-CSIC), 28049 Cantoblanco-Madrid, Spain, ²Scienseed SL, 28020 Madrid, Spain, ³The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Kongens Lyngby, Denmark, ⁴Department of Experimental Biology, Faculty of Science, Masaryk University, Brno 62500 Czech Republic, ⁵Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid (INIA-CSIC), Pozuelo de Alarcón 28223, Spain and ⁶School of Computing, Newcastle University, NE4 5TG, UK

Received September 23, 2022; Revised October 22, 2022; Editorial Decision October 24, 2022; Accepted October 24, 2022

ABSTRACT

The SEVA platform (<https://seva-plasmids.com>) was launched one decade ago, both as a database (DB) and as a physical repository of plasmid vectors for genetic analysis and engineering of Gram-negative bacteria with a structure and nomenclature that follows a strict, fixed architecture of functional DNA segments. While the current update keeps the basic features of earlier versions, the platform has been upgraded not only with many more ready-to-use plasmids but also with features that expand the range of target species, harmonize DNA assembly methods and enable new applications. In particular, SEVA 4.0 includes (i) a sub-collection of plasmids for easing the composition of multiple DNA segments with MoClo/Golden Gate technology, (ii) vectors for Gram-positive bacteria and yeast and [iii] off-the-shelf constructs with built-in functionalities. A growing collection of plasmids that capture part of the standard—but not its entirety—has been compiled also into the DB and repository as a separate corpus (SEVASib) because of its value as a resource for constructing and deploying phenotypes of interest. Maintenance and curation of the DB were accompanied by dedicated diffusion and communica-

tion channels that make the SEVA platform a popular resource for genetic analyses, genome editing and bioengineering of a large number of microorganisms.

INTRODUCTION

At the onset of the recombinant DNA era, different initiatives were put forward to standardize the architecture and the nomenclature of plasmid vectors (1). In retrospect, that most—if not all—failed cannot come as a surprise, as all the benefits of standardization in terms of interoperability and reproducibility come at the cost of losing flexibility (2), an issue hardly acceptable by the time modern genetic engineering was being shaped. Yet, the development of synthetic biology (SynBio) decades later and its ambition to make biology easier to engineer has raised again the issue of standardization as one of the challenges that practitioners must seriously consider (3) and tackle for moving the field forward (4). A recent review has taken stock of where standardization lays in the contemporary landscape of SynBio-fuelled bioengineering and which are the low-hanging fruits that have been seized—or can be delivered in the near future (2). Among them stands the growing collection of molecular tools for both genetic analysis and phenotype design of bacteria beyond typical Laboratory workhorses like *Escherichia coli* and *Bacillus subtilis*. Such tools include on one hand formatted methods of DNA assembly (pioneered by the BioBrick concept; 5,6) and on the other hand, vectors

*To whom correspondence should be addressed. Tel: +34 91 585 45 36; Fax: +34 91 585 45 06; Email: vdlorenzo@cnb.csic.es

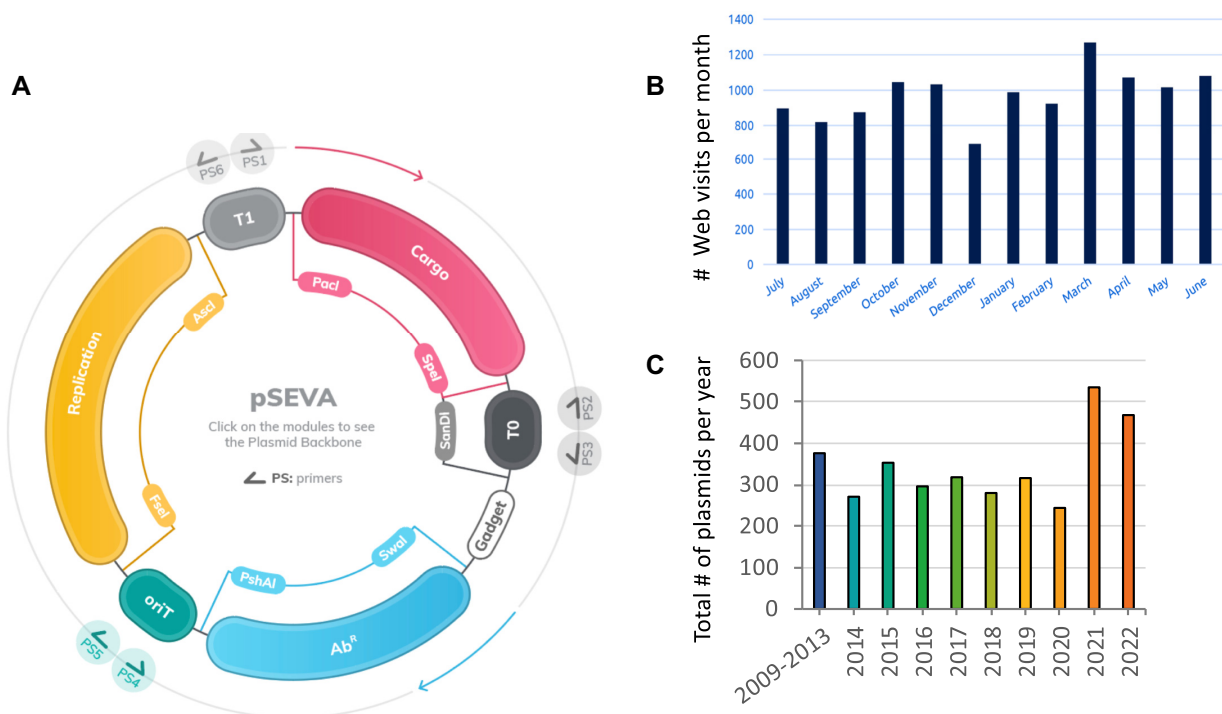


Figure 1. General organization of the SEVA plasmids and access statistics. (A) General organization of SEVA vectors as shown in the SEVA 4.0 homepage. All plasmids contain three basic modules (antibiotic marker, replication origin and cargo), the boundaries between being punctuated by unusual restriction sites. These are accompanied by a primary gadget site (additional sites are available also to this end). Clicking each of the segments enables users to visualize all choices and copy the cognate DNA sequence. The colored arrows above the cargo and Ab^R modules represent the directionality of the transcription flow of that element. (B) Visits to <https://seva-plasmids.com> during the period July 2021–June 2022. (C) Plasmids distributed since the onset of the SEVA platform.

for the deployment of given constructs in particular hosts (7–9).

It is in this context that the first SEVA database was launched in 2013 (10) with the purpose of bringing conceptual and material SynBio assets, in particular plasmid vectors, for analysis and construction of complex phenotypes to a much wider variety of bacteria. The proposed standard, (which has remained identical over subsequent versions) involves a few simple assembly rules of the functional DNA segments of the vectors i.e. replication origin, antibiotic marker and cargo. These segments are combined with a default *oriT* sequence for optional conjugal mobilization from the assembly host (typically *E. coli*) to the ultimate recipient (Figure 1A). The vectors resulting from such an arrangement are then assigned a code that acts as a unique identifier and since version SEVA 2.0 (11), they are also linked to a complete description in Synthetic Biology Open Language (SBOL) format (12). The community of users has grown ever since (Figure 1B) and the top 10 requested vectors have been distributed to hundreds of users (total plasmids distributed thus far exceeds 3400; Figure 1C) and the DB cited in the scientific literature >700 times at this time.

Since the latest update (13), the Open Access policy of the SEVA platform and its bottom-up, community-driven development have generated a substantial and useful feedback that is at the basis of the database update presented below. SEVA 4.0 incorporates a significant number of changes,

modifications and improvements that are briefly described in the following paragraphs and fully reflected in the current webpage. Special efforts have been made to render SEVA vectors compatible with DNA assembly methods based on Type-II restriction enzymes as well as easing the interplay of users with more friendly interfaces for guiding vector selection and plasmid construction. On a different side, the growing availability of *in silico* analysis tools of DNA sequences and plasmid display software has advised to encourage users outsourcing such assets rather than having them embedded in the platform. Finally, we present below how the diffusion through social media channels attached to the DB has created a conspicuous, growing community of global users. Ultimately, the ambition of SEVA is to serve the Synthetic Biology community by offering an open and user-friendly collection of genetic tools that meets a whole range of needs from simple students' projects all the way to large-scale biotechnological applications.

DATABASE DESCRIPTION

Database organization

As was the case with its predecessor, the updated SEVA database (SEVA-DB 4.0, <http://seva-plasmids.com/>) serves primarily as an annotated and information-rich index of functional DNA sequences and constructs that are available to the community. The platform maintains a simple

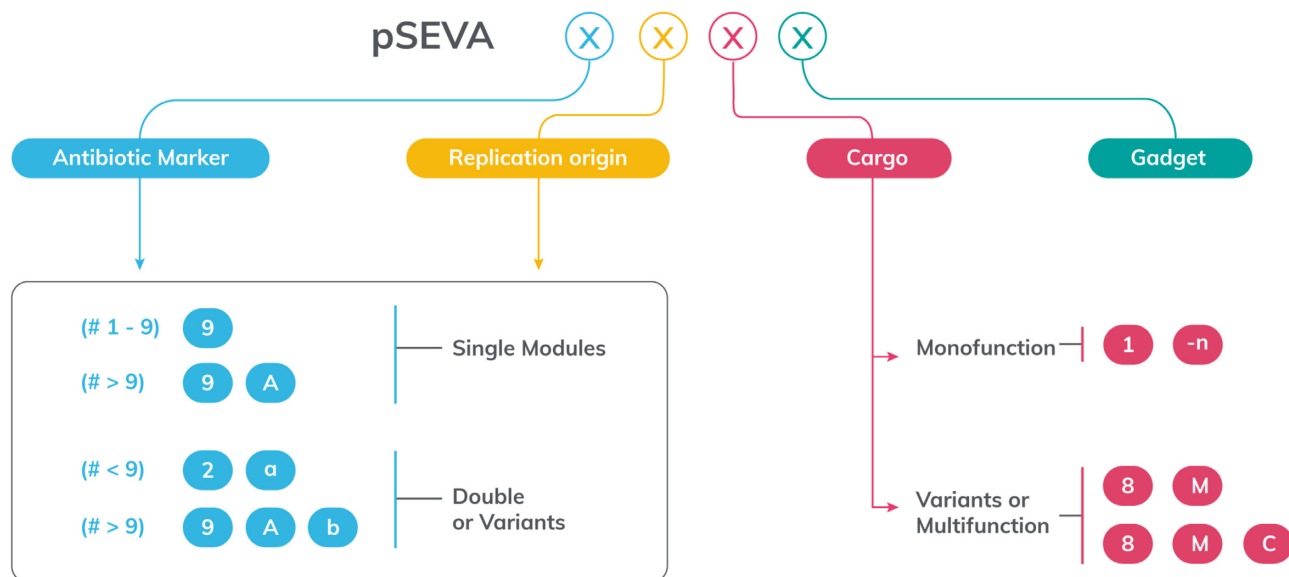


Figure 2. Updated nomenclature of SEVA vectors. The figure sketches the expanded rules for naming each of the four positions available for composing a complete SEVA code. The first position identifies the antibiotic resistance. The second position is the origin of replication (a sole numeric code 1 to 9 and then a capital letter is added). In case of either variants or addition(s) of either the antibiotic marker or a second replication origin, a lower-case letter is then inserted next. The third position is the cargo, which can be mono-function (named 1 to n) or variants/multi-function thereof (number followed by a capital letter). Finally, the fourth position is for the gadgets, which are designated by lowercase Greek letters (α to ω). For instance, pSEVA237M means a Km^R plasmid with a pBBR1 origin of replication and a promoterless *msf*GFP reporter (mono-function). pSEVA2313R means Km^R, pBBR1 origin and a P_{EM7}→mCherry cargo (bi-function). pSEVA2a2b8R α means Km^R plasmid selectable in Gram-positive hosts with a double RK2-SCP2* replication origin, a bi-functional cargo *xyf*/S-P_m→ mCherry and a *hok-sok* gadget (see text for explanation)

architecture, which consists of a relational database as the records storage layer, a series of modules and utilities that are hosted by an application server and a web-based presentation layer (13). The last is endowed with an explicit set of standards and keywords that apply to all constructs. The outputs of the search for each vector include links to GenBank (or alternatively to interim *gbk* or *gb* file of the plasmid), and to the sequences in Synthetic Biology Open language (SBOL) format (see below). However, built-in visualization of physical maps of each of the plasmids (which was available in earlier DB versions) has been removed and users are encouraged instead to export the DNA sequences of interest to exterior platforms of their choice, e.g. DNASTAR Lasergene (<https://www.dnastar.com/>) or Benchling (<https://www.benchling.com/>).

The graphic identity of the new website and its main sections have been largely kept as in the previous version (13), although it incorporates some important subsections and utilities that were not available before. The *Structure* section provides users with a basic description of the SEVA format as well as information about the steps needed for researchers to contribute to the standard—these have not changed significantly. The rules that govern the structure of the canonical SEVA plasmid collection are the very same proposed in 2013 (10) and we make a point of keeping them permanently as such. In this way, earlier vectors do not become obsolete and, instead, improvements can be entered by building on existing constructs. Such a canon involves DNA segments encoding specific plasmid-related functions which have been either synthesized or edited to remove restriction enzymes and endowed with rare restriction sites for assembly with an *oriT* sequence and transcriptional termi-

nators in a fixed arrangement. Some of such rare sites can later be used for insertion of what we call *gadgets*, which endow some extra functionality to the vector (see below).

The *Backbone modules* section constitutes the core of the platform, as the user has access to an intuitive graphic tour through the different functional segments of the plasmids and where visitors can see and copy all relevant sequences of the SEVA frame (Figure 1A). This section also contains a graphic representation of the rules for naming the plasmids that follow the SEVA standard and the procedure to assemble its code. The members of the canonical collection are named with a non-ambiguous code which due to the growing number of entries, was revised in the current version as indicated in Figure 2 to describe e.g. variants of existing functional segments as well as a DNA fragment with two or more functions (see below).

The next section (*Find your plasmid*; Figure 3) is the one which has been more significantly improved and offers more choices for users. The key subsection is named *Find your canonical plasmid*, where visitors can now separately select among many more antibiotic markers, replication origins and cargoes (instead of the earlier obligation to proceed stepwise). Optional gadgets include stabilization parts for securing plasmid retention in the absence of antibiotic selection and a yeast origin of replication for inter-kingdom transfer. Finally, users can survey a limited number of preset off-the-shelf functionalities for some specific applications (see below). This canonical collection (which can be perused also as a separate Table) includes the catalogue of SEVA vectors immediately available which fulfil every rule of the standard. While most plasmids of interest use to be ready in such a normative repository, combinations of func-

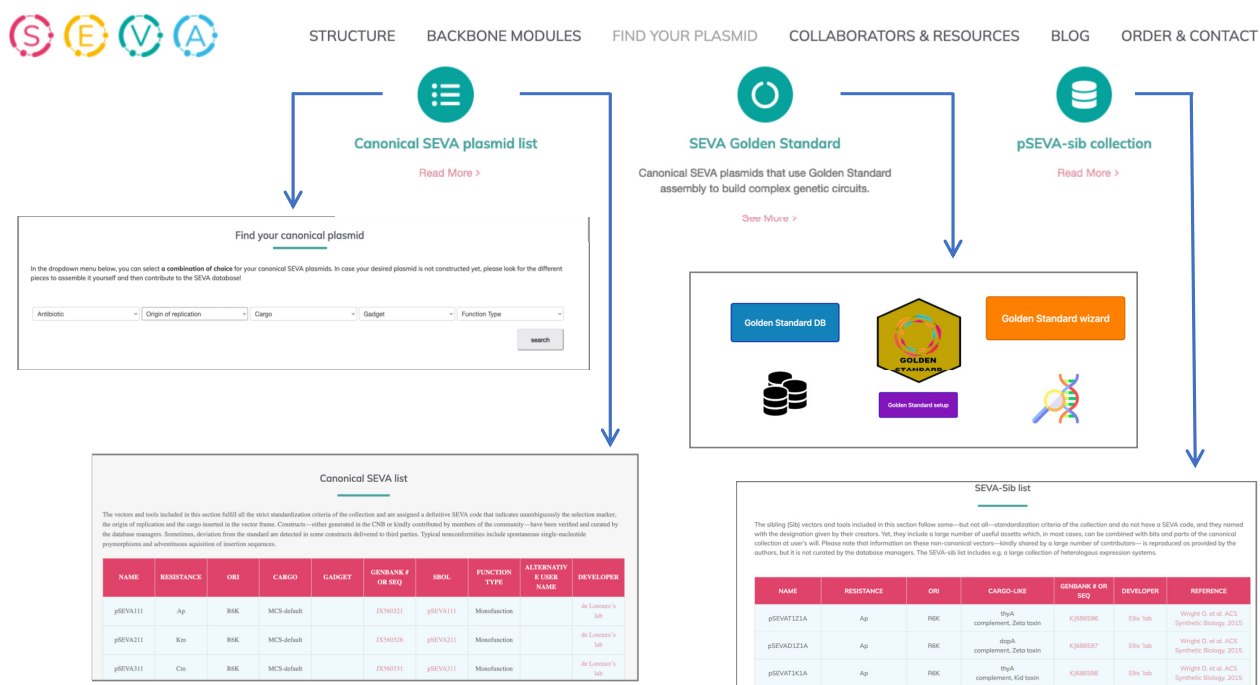


Figure 3. Details and utilities of the *Find your plasmid* section. The core of the database embodies the three subsections listed on top. Clicking each of them leads users to a different utility, the most useful of them being the Find your canonical plasmid, which can be composed according to necessities and then searched through the plasmid collection. The same section also leads to the resource named SEVA Golden Standard, which guides users through a complete roadmap for assembling complex constructs with Type IIS restriction technology. Separate Tables with the lists of canonical and non-canonical (SEVA-sib) vectors can be accessed as well.

tional parts that may not be physically assembled yet is facilitated to users-to-be by providing precursors following the directions spelled out in this same section. Availability of sequences in Genbank and SBOL formats enable their export for multiple visualisation options as indicated above.

Highlights of the updated canonical SEVA collection

The most conspicuous addition to the ready-to-use, canonical vector list is the whole collection of plasmids grouped under the heading *SEVA Golden standard* (cargo #19 and derivatives). These plasmids enable a Type IIS assembly method for combinatorial multi-part composition of standardized genetic elements to create complex genetic constructs with many transcription units. The detailed description of such platform is described in (14) and this section of the SEVA webpage is linked to a distinct repository of composable parts and vectors along with tools for guiding the design of DNA segments with both stored and user-specific sequences.

Apart from the *SEVA Golden standard*, the updated DB and repository has been value-added with new and noteworthy cloning and expression vectors as well as pre-made constructs. These can immediately be used for particular purposes of ample interest regarding analyses and construction of bacterial phenotypes. First comes new SEVA-formatted cargoes e.g. (i) mTurquoise2 (in plasmid pSEVA2313T) where the gene encoding this marker was placed under the control of the constitutive P_{EM7} promoter (15), (ii) an intracellular pH sensor (assembled in

plasmid pSEVA2513P), designed for assessing this parameter in Gram-negative bacteria. The cognate bi-functional cargo encodes a pH-sensitive reporter expressed through P_{EM7} (16), [iii] expanded tools for genome editing e.g. pSEVA1213S and pSEVA6213S, which supply the I-SceI meganuclease used for gene deletion in Gram-negative bacteria (17) endowed with the P_{EM7} promoter and bearing either an Ap^R or Gm^R resistance gene for specific applications. [iv] heterologous expression systems, e.g. responsive to either phosphate starvation (cargo #17 and its derivatives; 18) or rhamnose (cargo #18) as well as variants of existing ones thereof e.g. pSEVA6311 (a Gm^R vector encoding the cyclohexanone-inducible $ChnR/P_{chnB}$ system; 19,20). New expression systems also include a suite of calibrated constitutive promoters (cargoes #20-24) as well as an orthogonal T7 promoter (cargo #25) for high-processivity transcription of heterologous genes in Gram-negative hosts (21) that conditionally express the T7 RNA polymerase gene (for instance, integrated in the chromosome; 22). In some plasmids (e.g. pSEVA2325M or pSEVA2425M) the T7 promoter drives expression of an *msfGFP* marker (23) for calibration of transcriptional activity in given host strains.

Still within the realm of fully-standardized plasmids it is worth to highlight pSEVA2a2d1, which bear a double origin of replication (24) that enables constructs to become established in Gram-positive hosts (25). This adds to the already existing branch of SEVA vectors that replicate both in Gram-negative bacteria and actinomycetes (26). Also worthy of note is the yeast/bacteria shuttle vector

pSEVA222SB, endowed with a CEN6-ARS209-URA origin of replication-selection marker insert at the gadget site of the plasmid backbone. This vector has been instrumental for assembling very large constructs using *S. cerevisiae* as the host for recombination of multiple DNA fragments (27).

Registered SEVA siblings

Visitors of the SEVA webpage can also find a link to the list of SEVA sibling (SEVAsib) constructs in the section *Find your plasmid*, next to the *SEVA Golden standard* sector (Figure 3). The vectors and tools included in this divide follow some—but not all—standardization criteria of the collection, they do not have a SEVA code, and they are named with the designation provided by their creators. Yet, they include a large number of useful assets which, in most cases, can be combined with bits and parts of the canonical collection at the user's will. Please note that information on these non-canonical vectors—kindly shared by a large number of contributors—is reproduced as provided by the authors, but it is not curated by the database managers. Yet, the SEVAsib materials are available through the SEVA platform under the same terms of the canonical collection. However, specific details on their utilization may have to be requested to the original authors. Under these criteria, a suite of vectors and constructs that minimally diverge in respect to the standard were also incorporated to the update of SEVAsib section whenever rigorous maintenance of the format was not considered essential for the main purpose underlying their creation. For instance, some Ap^R vectors of the SEVAsib list bear a variant of the *bla* gene that has not been edited to fulfil the canon, but it delivers a much better selection with lower antibiotic concentrations (28). Also, the options for Type IIS-restriction mediated DNA assembly include not only the above-mentioned SEVA Golden vector collection, but also assets for what has been described as the SEVA 3.1 platform (29). This comprises a library of vectors tailored for Golden Gate-based assembly that work in conjunction with a core set of standard primers. This enables interoperability of the SEVA and BioBrick formats with minimal cloning and design efforts. That the deviation from the canon involved some swapping of the order of restriction sites in the cognate vectors (the so-called pSEVAb series)—while keeping the rest of the SEVA architecture—was an acceptable payoff for the interoperability of the plasmid vector series with the BioBrick assembly.

Another remarkable set of constructs of the SEVAsib collection comprises a whole inverter package for automated genetic circuit design in Gram-negative bacteria. This was aimed at extending the host range of the CELLO platform initially designed only for *E. coli* (30) towards a variety of non-canonical laboratory hosts (31–33). Also, a separate vector pSEVA6511-GIII was tailored for recombination-independent insertion of short DNA sequences with different purposes. One of its applications involves introduction of unique identifiers (i.e. genetic barcodes; 34,35) in permissive genomic sites of bacteria of industrial or environmental interest for the sake of strain safety and traceability (36).

Besides, the updated webpage includes plasmids which act as suicide deliverers of transposon vectors (37). These endow users with the possibility of easy cloning and subsequent random insertion of functional cargoes with different antibiotic-resistance markers. A set of such vectors are based on Tn5 (38) while another set is built on the Mariner transposase system—which, instead of a bias towards GC-rich regions, targets TA sites (39). In this last case, the plasmids bore the hyperactive form of the mariner Himar1 transposase (MarC9; 40) as well as cognate inverted repeats with a site for the Mm1 type IIs enzyme (41,42) that cuts ~20 bp away from its recognition site. This design allows inference of the genomic context of the inserted cargo and simplifies creation and analysis of next-generation sequencing libraries. While none of the transposon-delivery vectors—whether Tn5-based or Mariner-based—follows the SEVA rules altogether, the compatibility and boundaries of cloning sites makes transfer of cargoes among platforms straightforward.

Finally, the homepage of the platform incorporates (i) updated links to the addresses of depositors along with a connection to the autonomous *Plasmid collection of Ghent University* (within the Belgian Co-ordinated Collection of Microorganisms), a useful resource for Synthetic Biology-based biotechnology, (ii) access to the collection of blog posts in social media (see below) and (iii) contact forms for users to request the plasmids free of charge.

Updating the nomenclature of SEVA modules

As the physical number of modules within SEVA plasmids continues to expand, the nomenclature of the collection has to evolve as well. While SEVA 3.0 included a detailed description of how-to code variant and double origins of replication nothing was stated about the antibiotic module (13). To clarify this, the update (Figure 2) reflects the nomenclature of different module variants encoding the same antibiotic resistance. While new antibiotics beyond #9 would be represented by adding a capital letter (9A, 9B and so on), different DNA variants for the same antibiotic resistance would be coded by adding a lower-case letter (similarly to the ORI module). For example, the new antibiotic codes #2a (*aadD* gene from pUB110-*S. aureus*) and #3a (*cat* gene from pC194-*S. aureus*) correspond to resistance cassettes derived from Gram-positive bacteria, that allows cognate bacterial cells to grow in kanamycin or chloramphenicol respectively (43–45). Moreover, the update includes SEVA plasmids adapted to modular cloning (Golden Standard; 14) which, as indicated above, leverages Type IIS assembly. The specific nomenclature of these plasmids is explained in (14).

Another new feature included in the 4.0 update is the grouping of SEVA plasmids into two types, mono-function or multi-function, depending on the contents of the cargo module. To meet that change, a new column has been added to the canonical SEVA plasmid list in the webpage. The mono-function group includes cargoes endowed with a unique and clear application (e.g. MCS). The four types of cargoes applications originally described in SEVA 1.0 (10) belong to the mono-function class. These applications include the most common genetic operations: cloning,

genome editing, promoter-probe and heterologous DNA expression systems. However, in some cases two functional applications can be joined together within a cargo module (the whole DNA sequence being flanked by *PacI* and *SpeI*) to deliver a new function. This multi-function category is not limited to just two functional applications, but as many as necessary. However, all such DNA sequences within this multi-functional cargo must fulfil the SEVA standard requirements (10). Multi-function plasmids include, e.g. constitutive/inducible expression of a reporter gene, or inducible expression of the *I-SceI* enzyme. This new criterion has allowed us to transfer plasmids previously listed as siblings into the canonical collection and thus rename them to meet the updated SEVA code. In order to ease searches, an additional column with the earlier plasmid name has also been included in the canonical Table where necessary. The presence of multi-functional cargoes in given vectors is thus noted by first stating the leading part (e.g. P_{EM7} would be #13) followed by one or two letters corresponding to the downstream component (e.g. cargo $P_{EM7} \rightarrow$ mCherry would be #13R; see legend to Figure 2 for some examples)

SEVA sequences in SBOL format: an update

SEVA 2.0 (11) enabled translation of vectors into the Synthetic Biology Open Language (SBOL) format, a data standard for genetic circuit design, the advantages of which have been widely argued before (12). In the current update, SEVAhub (<http://sevahub.es>), the design repository providing access to SEVA vectors converted to the Synthetic Biology Open Language (SBOL) (12) has been updated to the latest version of SynBioHub (46). This version adds new features (e.g. validation steps or improved SBOL-visual representation; 47) and bug fixes as outlined at <https://github.com/SynBioHub/synbiohub/releases>. Moreover, the new version lists the federated SynBioHub instances. The SEVA-to-SBOL conversion process has been improved as well: Previous SBOL files were generated manually by developing a software package to combine part sequences. In this update, we provide an automatic conversion tool that allows individual generation of SBOL encoded vectors and provides feedback on potential issues that may not initially be derived from the corresponding GenBank files (e.g. double annotations on the same sequence). Users can access this conversion tool at <https://github.com/Biocomputation-CBGP/GBK2SBOL>. Resulting SBOL files can then be used for the in-depth analysis of sequences and constructs (48) leading to advanced data integration strategies (49).

Ordering SEVA plasmids

One of the trademarks of the SEVA platform is making standardized vectors freely available to all types of users, whether academic or industrial. All users have to do for ordering SEVA plasmids and Golden Standard vectors is filling the request form of the corresponding section of the SEVA webpage. Petitioners find there an easy procedure for asking (a reasonable number of) specific vectors. These are then shipped free of charge to individuals who contact through institutional email addresses of research centres,

universities and accredited companies. For bulk orders of plasmids (10 or more), please contact the Webmaster as some handling fees could be applicable. Please note that we do not send materials requested through commercial email platforms or private addresses. Note also that we disclaim carrier liability. While most of the functional parts entered in the SEVA vectors enjoy freedom of operation, we cannot rule out that some given sequences are subject to some intellectual property (IP) rights. Should utilization of SEVA plasmids eventually led to a commercial application, users are encouraged to inspect the IP status of the corresponding DNA segments and deal with the issue with the original IP holder. In this way, IP and potential licensing issues are handed over to end-user scenarios rather than raising them upfront.

Diffusion and communication: building a community of users

The setup of the first SEVA webpage in 2013 (10) was followed later (13) by the opening of a linked Twitter account (<https://twitter.com/SEVAplasmids>) along with an interactive blog with comments, novelties and posts related to vector development. More recently, these have been added with a free newsletter sent every three months to subscribers. External access to these platforms has enabled us to map and follow the evolution of the interest in the SEVA resource, especially in the period of time since the last update in 2020. As seen in Figure 1B, the number of website visits have remained basically stable during the last period (over 1000 per month), and the number of plasmids distributed (Figure 1C) kept over an average of 300 per year. Furthermore, adopting the number of followers in Twitter as a proxy of the interest and the professional allocation of users, allowed us to analyse and understand the communities created around the project as well as determining whether the platform reaches relevant audiences. The results (Figure 4) break down the field in seven distinct communities, the most differentiated being those self-described as Synthetic Biologists and Biotechnology Service providers. Beyond the number of citations of the SEVA platform since its first edition (>700 at the time of writing this article), these figures suggest an increased interest in the sort of genetic tools available through the site and they foreshow its further growth and demand. The same can be said for visits to the SEVA blog (<https://seva-plasmids.com/blog/>) and the quarterly Newsletter to which one can subscribe through the webpage as well. Worth to mention also is the considerable visibility and appeal bestowed by the explanatory animation video (<https://www.youtube.com/watch?v=I8gZjXDRQ2I>) which appears in the SEVA homepage and provides a quick, user-friendly glimpse of what the platform is all about.

OUTLOOK

What could be called the SEVA space includes a collection of general-purpose, broad host range cloning vectors, heterologous expression systems, reporters and specific purpose-constructs. The entries of the canonical vector collection, which is the core of the corpus, follows rigorously the arrangement of functional DNA segments established in the earliest publication (Figure 1A; 10) and which

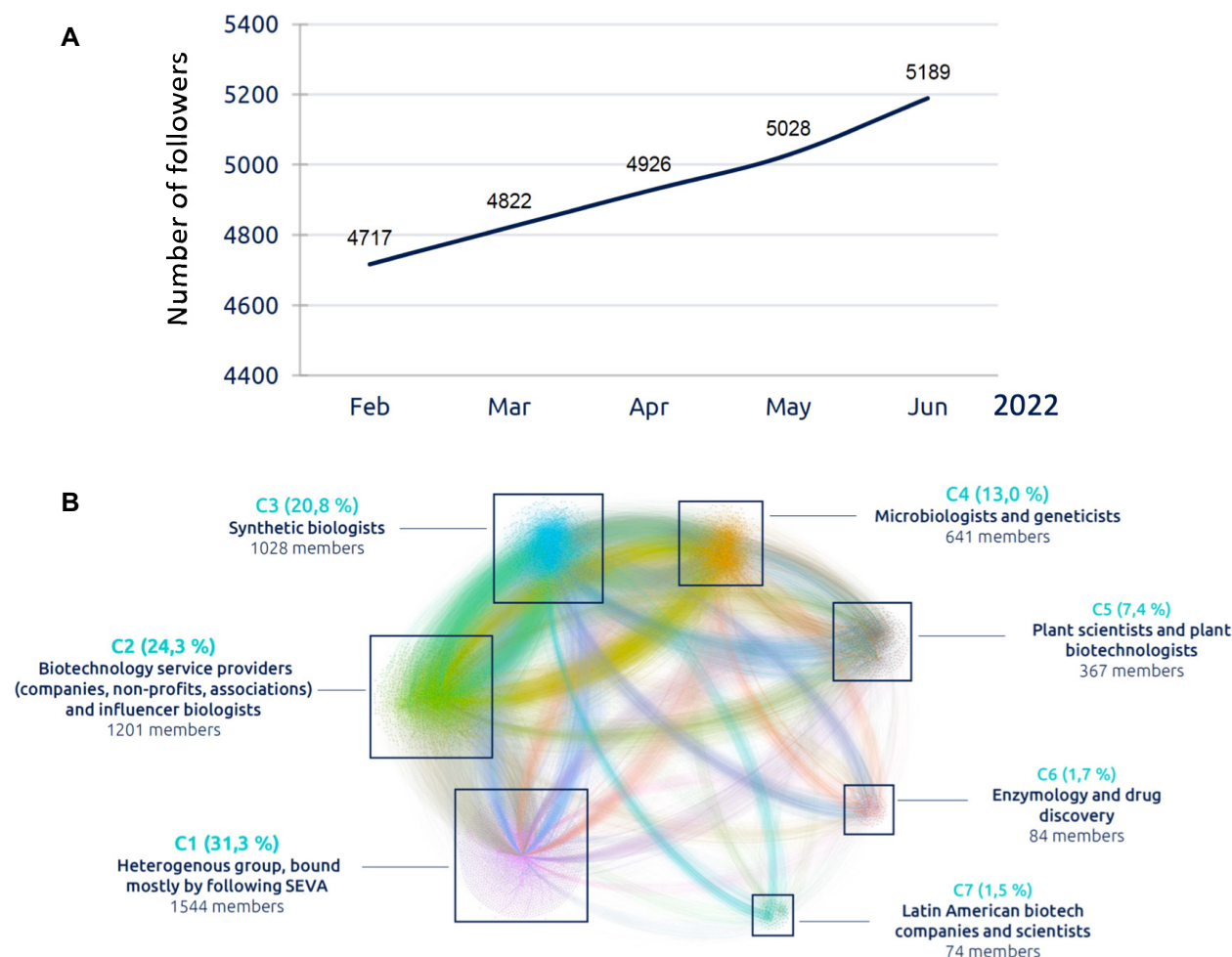


Figure 4. Mapping the community of SEVA users. The upper panel shows the steady increase of followers of the SEVA Twitter account <https://twitter.com/SEVAplasmids> in the first part of 2022. The profiles of all followers were downloaded through the Twitter API (application programming interface) and the resulting network of users visualized using Gephi software (<https://gephi.org/>) for complex node analysis. Each square represents a community of followers (nodes) with shared interests, language or geographic location (C1–C7), labelled with the percentage share of followers in that group. Denser clusters in the visualisation (such as C3 or C4) indicate closely-knit communities, wherein many users follow each other. Note predominance of synthetic biologists and biotechnology companies.

can be unequivocally cyphered according to a code (Figure 2) which in this case needs to be updated as the plasmid collection increases in number and utilities. Apart of the canonical set, the platform also includes a considerable number of constructs (SEVAsib) that capture part of the standard—but not all—without loss of functionality for specific purposes. These are also included in the repository and can be retrieved through the same channels and terms of the canonical collection.

Despite the number of excellent plasmid repositories available to molecular biologists and synthetic biologist, the distinguishing marks of the SEVA platform include (i) the emphasis on compositional standards for the sake of interoperability and reproducibility, (ii) the adoption of broad host range functionalities (antibiotic markers, replication origin, cargoes) and (iii) open and free upfront access to the materials at stake. While other repositories of biological materials may also take all or some of these features, we strongly advocate standardization as the way to go for moving the field of synthetic biology forward.

Since the last update of the SEVA DB we have noted publication of a number of articles that further develop the concept in different directions explicitly related to the original platform (50,51). In other cases, authors take on the overall SEVA view for the progress of their own developments in genetic tools but depart from the standard to different degrees and purposes (52–58). Finally, some conceptually independent proposals go also along the path of standardization of genetic assets without any explicit connection to the SEVA concept or format (59–63). It would be ideal that all these eventually converge in a single frame that eases genetic analyses and engineering of a large variety of microorganisms with compatible and comparable tools.

In the meantime, some additional expansion of the SEVA platform could be expected for improving e.g. the host range of the activities encoded in the constructs. This involves not just the replication origins but also expression signals for enabling genes to be expressed in a large variety of hosts (64). Also, there is considerable room for standardizing mini-

transposon delivery vectors (and other integrative tools) beyond the prototypes currently listed in the SEVAsib Table. Finally, the growing trend to move from strain analysis and engineering to the same applied to microbiomes (65) asks also for novel, dedicated tools that are still in their infancy and deserve standardization efforts as well. These challenges will hopefully be met in subsequent updates of the platform.

DATA AVAILABILITY

All relevant data is available through the SEVA platform (<https://seva-plasmids.com>). No new data were generated or analysed in support of this research.

ACKNOWLEDGEMENTS

Authors are indebted to all members of the Environmental Synthetic Laboratory in Madrid as well as to different international teams who supplied plasmids (whether canonical or siblings) to the SEVA collection. In particular, we thank members of the SEM group (DTU Biosustain, Denmark) for their help in constructing some of the functional parts reported in this work. The contributions of Ryan Kniewel (Prieto Laboratory, CIB-CSIC Madrid, Spain), Jesús Torres Bacete (Nogales Laboratory, CNB-CSIC Madrid, Spain), Daniel Daley (Stockholm, Sweden), Christos Batianis and Stamatios Damalas (Martins dos Santos Laboratory, Wageningen) and Luis López Maury (University of Sevilla) are also gratefully acknowledged.

FUNDING

The SEVA repository has been developed and maintained with funds of the SYCOLIM [ERA-COBIOTECH 2018-PCI2019-111859-2] Project of the Spanish Ministry of Science and Innovation, SYN BIO4FLAV [H2020-NMBP-TR-IND/H2020-NMBP-BIO-2018-814650]; MIX-UP [MIX-UP H2020-BIO-CN-2019-870294] Contracts of the European Union; BIOSINT-CM [Y2020/TCS-6555] Project of the Comunidad de Madrid-European Structural and Investment Funds (FSE, FECER); P.I.N. acknowledges financial support by the Novo Nordisk Foundation [NNF20CC0035580, TARGET (NNF21OC0067996)]; European Union's Horizon 2020 Research and Innovation Programme [814418 (SinFonia)]; M.H.H.N. acknowledges funding by the Novo Nordisk Foundation [NNF20CC0035580]; P.D. was funded by Czech Science Foundation Project 22-12505S; A.G.M. was supported by the Grants BioSinT-CM [Y2020/TCS-6555]; CONTEXT (Atracción de Talento Program) [2019-T1/BIO-14053] Projects of the Comunidad de Madrid, MULTI-SYSBIO [PID2020-117205GA-I00]; Severo Ochoa Program for Centres of Excellence in R&D [CEX2020-000999-S] funded by MCIN/AEI/10.13039/501100011033 and the ECCO (ERC-2021-COG-101044360) Contract of the EU. Funding for open access charge: European Commission Grant SYN BIO4FLAV [H2020-NMBP-TR-IND/H2020-NMBP-BIO-2018-814650].

Conflict of interest statement. None declared.

REFERENCES

- Novick, R.P., Clowes, R.C., Cohen, S.N., Curtiss, R. 3rd, Datta, N. and Falkow, S. (1976) Uniform nomenclature for bacterial plasmids: a proposal. *Bacteriol. Rev.*, **40**, 168–189.
- Beal, J., Goñi-Moreno, A., Myers, C., Hecht, A., de Vicente, M.D.C., Parco, M., Schmidt, M., Timmis, K., Baldwin, G., Friedrichs, S. *et al.* (2020) The long journey towards standards for engineering biosystems: are the molecular biology and the biotech communities ready to standardise? *EMBO Rep.*, **21**, e50521.
- Tas, H., Amara, A., Cueva, M.E., Bongaerts, N., Calvo-Villamañán, A., Hamadache, S. and Vavitsas, K. (2020) Are synthetic biology standards applicable in everyday research practice? *Microb. Biotechnol.*, **13**, 1304–1308.
- Porcar, M., Danchin, A. and de Lorenzo, V. (2015) Confidence, tolerance, and allowance in biological engineering: the nuts and bolts of living things. *Bioessays*, **37**, 95–102.
- Shetty, R.P., Endy, D. and Knight, T.F. (2008) Engineering biobrick vectors from biobrick parts. *J. Biol. Eng.*, **2**, 5.
- Røkke, G., Korvald, E., Pahr, J., Oyås, O. and Lale, R. (2014) BioBrick assembly standards and techniques and associated software tools. *Methods Mol. Biol.*, **1116**, 1–24.
- Dong, H. and Zhang, D. (2014) Current development in genetic engineering strategies of *Bacillus* species. *Microb. Cell Fact.*, **13**, 63.
- Lee, N., Hwang, S., Lee, Y., Cho, S., Palsson, B. and Cho, B.K. (2019) Synthetic biology tools for novel secondary metabolite discovery in streptomyces. *J. Microbiol. Biotechnol.*, **29**, 667–686.
- Tas, H., Nguyen, C.T., Patel, R., Kim, N.H. and Kuhlman, T.E. (2015) An integrated system for precise genome modification in *Escherichia coli*. *PLoS One*, **10**, e0136963.
- Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., de Las Heras, A., Páez-Espino, A.D., Durante-Rodríguez, G., Kim, J., Nikel, P.I. *et al.* (2013) The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.*, **41**, D666–D675.
- Martínez-García, E., Aparicio, T., Goni-Moreno, A., Fraile, S. and de Lorenzo, V. (2015) SEVA 2.0: an update of the standard european vector architecture for de/-re-construction of bacterial functionalities. *Nucleic Acids Res.*, **43**, D1183–D1189.
- Madsen, C., Goñi Moreno, A. and P, U.P. U., Palchick, Z., Roehner, N., Atallah, C., Bartley, B., Choi, K., Cox, R.S., Gorochoowski, T. *et al.* (2019) Synthetic biology open language (SBOL) version 2.3. *J. Integr. Bioinform.*, **16**, 20190025.
- Martínez-García, E., Goñi-Moreno, A., Bartley, B., McLaughlin, J., Sánchez-Sampedro, L., Pascual Del Pozo, H., Prieto Hernández, C., Marletta, A.S., De Lucrezia, D., Sánchez-Fernández, G. *et al.* (2020) SEVA 3.0: an update of the standard european vector architecture for enabling portability of genetic constructs among diverse bacterial hosts. *Nucleic Acids Res.*, **48**, D1164–D1170.
- Blázquez, B., Torres-Bacete, J., San Leon, D., Kniewel, R., Martínez, I., Sordon, S., Wilczak, A., Salgado, S., Huszcza, E., Popłoński, J. *et al.* (2022) Golden standard: a complete standard, portable, and interoperable moclo tool for model and non-model bacterial hosts. bioRxiv doi: <https://doi.org/10.1101/2022.09.20.508659>, 20 September 2022, preprint: not peer reviewed.
- Wirth, N.T., Kozaeva, E. and Nikel, P.I. (2020) Accelerated genome engineering of *Pseudomonas putida* by I-SceI mediated recombination and CRISPR-Cas9 counterselection. *Microb. Biotech.*, **13**, 233–249.
- Arce-Rodríguez, A., Volke, D.C., Bense, S., Häussler, S. and Nikel, P.I. (2019) Non-invasive, ratiometric determination of intracellular pH in *Pseudomonas* species using a novel genetically encoded indicator. *Microb. Biotech.*, **12**, 799–813.
- Martínez-García, E. and de Lorenzo, V. (2011) Engineering multiple genomic deletions in Gram-negative bacteria: analysis of the multi-resistant antibiotic profile of *Pseudomonas putida* KT2440. *Environ. Microbiol.*, **13**, 2702–2716.
- Torres-Bacete, J., Luis García, J. and Nogales, J. (2021) A portable library of phosphate-depletion based synthetic promoters for customizable and automata control of gene expression in bacteria. *Microb. Biotech.*, **14**, 2643–2658.
- Benedetti, I., de Lorenzo, V. and Nikel, P.I. (2016) Genetic programming of catalytic *Pseudomonas putida* biofilms for boosting biodegradation of haloalkanes. *Metab. Eng.*, **33**, 109–118.

20. Nikel, P.I., Benedetti, I., Wirth, N.T., de Lorenzo, V. and Calles, B. (2022) Standardization of regulatory nodes for engineering heterologous gene expression: a feasibility study. *Microb. Biotechnol.*, **15**, 2250–2265.
21. Rong, M., He, B., McAllister, W.T. and Durbin, R.K. (1998) Promoter specificity determinants of T7 RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 515–519.
22. Calero, P., Volke, D.C., Lowe, P.T., Gotfredsen, C.H., O'Hagan, D. and Nikel, P.I. (2020) A fluoride-responsive genetic circuit enables in vivo biofluorination in engineered *Pseudomonas putida*. *Nat. Commun.*, **11**, 5045.
23. Volke, D.C., Turlin, J., Mol, V. and Nikel, P.I. (2020) Physical decoupling of xyls/Pm regulatory elements and conditional proteolysis enable precise control of gene expression in *Pseudomonas putida*. *Microb. Biotechnol.*, **13**, 222–232.
24. Aparicio, T., Silbert, J., Cepeda, S. and de Lorenzo, V. (2022) Propagation of recombinant genes through complex microbiomes with synthetic mini-RP4 plasmid vectors. *Bio Design Res.*, **2022**, 9850305.
25. Ceglowski, P., Boitsov, A., Chai, S. and Alonso, J.C. (1993) Analysis of the stabilization system of pSM19035-derived plasmid pBT233 in *Bacillus subtilis*. *Gene*, **136**, 1–12.
26. García-Gutiérrez, C., Aparicio, T., Torres-Sánchez, L., Martínez-García, E., de Lorenzo, V., Villar, C.J. and Lombó, F. (2020) Multifunctional SEVA shuttle vectors for actinomycetes and Gram-negative bacteria. *Microbiol. Open*, **9**, 1135–1149.
27. Silbert, J., Lorenzo, V.d. and Aparicio, T. (2021) Refactoring the conjugation machinery of promiscuous plasmid RP4 into a device for conversion of gram-negative isolates to *hfr* strains. *ACS Synth. Biol.*, **10**, 690–697.
28. Cumming, A.J., Khananisho, D., Harris, R., Bayer, C.N., Nørholm, M.H.H., Jamshidi, S., Ilag, L.L. and Daley, D.O. (2022) Antibiotic-efficient genetic cassette for the TEM-1 β -lactamase that improves plasmid performance. *ACS Synth. Biol.*, **11**, 241–253.
29. Damalas, S.G., Batianis, C., Martin-Pascual, M., de Lorenzo, V. and Martins Dos Santos, V.A.P. (2020) SEVA 3.1: enabling interoperability of DNA assembly among the SEVA, biobricks and type IIS restriction enzyme standards. *Microb. Biotechnol.*, **13**, 1793–1806.
30. Nielsen, A.A., Der, B.S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E.A., Ross, D., Densmore, D. and Voigt, C.A. (2016) Genetic circuit design automation. *Science*, **352**, aac7341.
31. Tas, H., Goñi-Moreno, Á. and Lorenzo, V. (2021) A standardized inverter package borne by broad host range plasmids for genetic circuit design in Gram-negative bacteria. *ACS Synth. Biol.*, **10**, 213–217.
32. Tas, H., Grozinger, L., Goñi-Moreno, A. and de Lorenzo, V. (2021) Automated design and implementation of a NOR gate in *Pseudomonas putida*. *Synth. Biol.*, **6**, ysab024.
33. Tas, H., Grozinger, L., Stoof, R., de Lorenzo, V. and Goñi-Moreno, Á. (2021) Contextual dependencies expand the re-usability of genetic inverters. *Nat. Commun.*, **12**, 355.
34. Tellechea-Luzardo, J., Winterhalter, C., Wiedera, P., Kozyra, J., de Lorenzo, V. and Krasnogor, N. (2020) Linking engineered cells to their digital twins: a version control system for strain engineering. *ACS Synth. Biol.*, **9**, 536–545.
35. Tellechea-Luzardo, J., Hobbs, L., Velázquez, E., Pelechova, L., Woods, S., de Lorenzo, V. and Krasnogor, N. (2022) Versioning biological cells for trustworthy cell engineering. *Nat. Commun.*, **13**, 765.
36. Velázquez, E., Al-Ramahi, Y., Tellechea-Luzardo, J., Krasnogor, N. and de Lorenzo, V. (2021) Targetron-assisted delivery of exogenous DNA sequences into *Pseudomonas putida* through CRISPR-Aided counterselection. *ACS Synth. Biol.*, **10**, 2552–2565.
37. Martínez-García, E., Aparicio, T., Lorenzo, V.d. and Nikel, P.I. (2017) Engineering Gram-negative microbial cell factories using transposon vectors. *Methods Mol. Biol.*, **1498**, 273–293.
38. Martínez-García, E., Aparicio, T., de Lorenzo, V. and Nikel, P.I. (2014) New transposon tools tailored for metabolic engineering of gram-negative microbial cell factories. *Front. Bioeng Biotechnol.*, **2**, 46.
39. Choudhery, S., Brown, A.J., Akusobi, C., Rubin, E.J., Sasseti, C.M. and Ioerger, T.R. (2021) Modeling site-specific nucleotide biases affecting *himar1* transposon insertion frequencies in tseq data sets. *Msystems*, **6**, e0087621.
40. Lampe, D.J., Akerley, B.J., Rubin, E.J., Mekalanos, J.J. and Robertson, H.M. (1999) Hyperactive transposase mutants of the *himar1* mariner transposon. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 11428–11433.
41. van Opijnen, T., Bodi, K.L. and Camilli, A. (2009) Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat. Methods*, **6**, 767–772.
42. Morgan, R.D., Bhatia, T.K., Lovasco, L. and Davis, T.B. (2008) MmeI: a minimal type II restriction-modification system that only modifies one DNA strand for host protection. *Nucleic Acids Res.*, **36**, 6558–6570.
43. McKenzie, T., Hoshino, T., Tanaka, T. and Sueoka, N. (1986) The nucleotide sequence of pUB110: some salient features in relation to replication and its regulation. *Plasmid*, **15**, 93–103.
44. Horinouchi, S. and Weisblum, B. (1982) Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. *J. Bacteriol.*, **150**, 815–825.
45. Biswas, I., Jha, J.K. and Fromm, N. (2008) Shuttle expression plasmids for genetic studies in *Streptococcus mutans*. *Microbiology*, **154**, 2275–2282.
46. McLaughlin, J.A., Myers, C.J., Zundel, Z., Mısırlı, G., Zhang, M., Ofiteru, I.D., Goñi-Moreno, A. and Wipat, A. (2018) SynBioHub: a standards-enabled design repository for synthetic biology. *ACS Synth. Biol.*, **7**, 682–688.
47. Beal, J., Nguyen, T., Goroehowski, T.E., Goñi-Moreno, A., Scott-Brown, J., McLaughlin, J.A., Madsen, C., Aleritsch, B., Bartley, B., Bhakta, S. *et al.* (2019) Communicating structure and function in synthetic biology diagrams. *ACS Synth. Biol.*, **8**, 1818–1825.
48. Crowther, M., Wipat, A. and Goñi-Moreno, Á. (2022) A network approach to genetic circuit designs. *ACS Synth. Biol.*, **11**, 3058–3066.
49. Tellechea-Luzardo, J., Otero-Muras, I., Goñi-Moreno, A. and Carbonell, P. (2022) Fast biofoundries: coping with the challenges of biomanufacturing. *Trends Biotechnol.*, **40**, 831–842.
50. Kim, S.H., Cavaleiro, A.M., Rennig, M. and Nørholm, M.H.H. (2016) SEVA linkers: a versatile and automatable DNA backbone exchange standard for synthetic biology. *ACS Synth. Biol.*, **5**, 1177–1181.
51. Valenzuela-Ortega, M. and French, C. (2021) Joint universal modular plasmids (JUMP): a flexible vector platform for synthetic biology. *Synth. Biol.*, **6**, ysab003
52. Lammens, E.-M., Boon, M., Grimon, D., Briers, Y. and Lavigne, R. (2022) SEVAtile: a standardised DNA assembly method optimised for *Pseudomonas*. *Microb. Biotech.*, **15**, 370–386.
53. Schuster, L.A. and Reisch, C.R. (2021) A plasmid toolbox for controlled gene expression across the proteobacteria. *Nucleic Acids Res.*, **49**, 7189–7202.
54. Geddes, B.A., Mendoza-Suárez, M.A. and Poole, P.S. (2018) A bacterial expression vector archive (BEVA) for flexible modular assembly of golden gate-compatible vectors. *Front. Microbiol.*, **9**, 3345.
55. Falkenberg, K.B., Mol, V., de la Maza Larrea, A.S., Pogrebnyakov, I., Nørholm, M.H.H., Nielsen, A.T. and Jensen, S.I. (2021) The prouser2.0 toolbox: genetic parts and highly customizable plasmids for synthetic biology in *bacillus subtilis*. *ACS Synth. Biol.*, **10**, 3278–3289.
56. Opel, F., Siebert, N.A., Klatt, S., Tüllinghoff, A., Hantke, J.G., Toepel, J., Bühler, B., Nürnberg, D.J. and Klähn, S. (2022) Generation of synthetic shuttle vectors enabling modular genetic engineering of cyanobacteria. *ACS Synth. Biol.*, **11**, 1758–1771.
57. Silva-Rocha, R., Pontelli, M.C., Furtado, G.P., Zaramela, L.S. and Koide, T. (2015) Development of new modular genetic tools for engineering the halophilic archaeon *Halobacterium salinarum*. *PLoS One*, **10**, e0129215.
58. Bayer, C.N., Rennig, M., Ehrmann, A.K. and Nørholm, M.H.H. (2021) A standardized genome architecture for bacterial synthetic biology (SEGA). *Nat. Comm.*, **12**, 5876.
59. Elmore, J.R., Furches, A., Wolff, G.N., Gorday, K. and Guss, A.M. (2017) Development of a high efficiency integration system and promoter library for rapid modification of *Pseudomonas putida* KT2440. *Metab. Eng. Comm.*, **5**, 1–8.
60. Nora, L.C., Gonçalves, R.A., Martins-Santana, L., Ferreira, B.H., Rodrigues, F. and Silva-Rocha, R. (2019) Synthetic and minimalist vectors for *Agrobacterium tumefaciens*-mediated transformation of fungi. *Genet. Mol. Biol.*, **42**, 395–398.

61. Verdú,C., Sanchez,E., Ortega,C., Hidalgo,A., Berenguer,J. and Mencia,M. (2019) A modular vector toolkit with a tailored set of thermosensors to regulate gene expression in *Thermus thermophilus*. *ACS Omega*, **4**, 14626–14632.
62. Ma,X., Liang,H., Cui,X., Liu,Y., Lu,H., Ning,W., Poon,N.Y., Ho,B. and Zhou,K. (2019) A standard for near-scarless plasmid construction using reusable DNA parts. *Nat. Comm.*, **10**, 3294.
63. Enghiad,B., Xue,P., Singh,N., Boob,A.G., Shi,C., Petrov,V.A., Liu,R., Peri,S.S., Lane,S.T., Gaither,E.D. *et al.* (2022) PlasmidMaker is a versatile, automated, and high throughput end-to-end platform for plasmid construction. *Nat. Commun.*, **13**, 2697.
64. De Wannemaeker,L., Bervoets,I. and De Mey,M. (2022) Unlocking the bacterial domain for industrial biotechnology applications using universal parts and tools. *Biotechnol. Adv.*, **60**, 108028.
65. Leggieri,P.A., Liu,Y., Hayes,M., Connors,B., Seppälä,S., O'Malley,M.A. and Venturelli,O.S. (2021) Integrating systems and synthetic biology to understand and engineer microbiomes. *Ann. Rev. Biomed. Eng.*, **23**, 169–201.