This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 870294.

MIXed plastics biodegradation and UPcycling using microbial communities
MIX-UP

Deliverable D7.3

Update SEVA collection

Dissemination				
Lead	**International partners involved**	**Level**	**Type**	**Delivery Month**
CSIC	☐ NTU ☐ BUCT ☐ IPE ☐ TU	☒ PU (Public) ☐ CO (Confidential) ☐ Cl (Classified)	☒ R (Document, report) ☐ DEM (Demonstrator, pilot, prototype) ☐ DEC (Websites, patent fillings, videos etc.) ☐ ORDP (Open Research Data Pilot) ☐ ETHICS ☐ OTHER	M36

Grant Agreement:	870294
Project acronym:	MIX-UP
Project title:	MIXed plastics biodegradation and UPcycling using microbial communities
Start of the project:	01.01.2020 (48 months)
Work package:	WP7
Deliverable number:	D7.3
Title of deliverable:	Update SEVA collection
Deliverable due date:	31/12/2022
Date of submission:	29/12/2022
First Author:	Victor de Lorenzo (CNB-CSIC)
Contributors:	Juan Nogales (CNB-CSIC)
Version:	1.1

VERSION CONTROL

Version	Date	Author (Name, Institution)	Comments
1.0	16/12/2022	Victor de Lorenzo (CNB-CSIC)	First draft
1.1	22/12/2022	Hendrik Ballerstedt (RWTH)	Minor textual changes

Short description of the content of this deliverable: Update SEVA collection with at least three new platforms for genome editing and site-specific sequence diversification of chromosomal segments (CSIC).

1 Update of the SEVA collection.

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

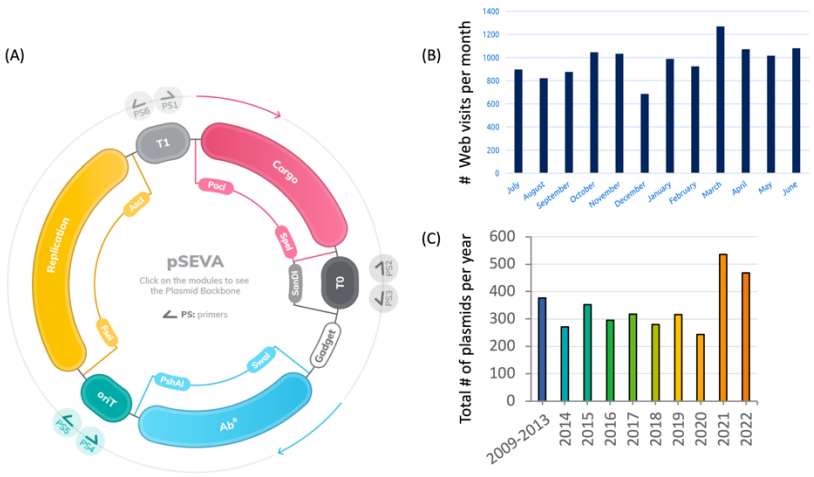

Figure 1. General organization of the SEVA plasmids and access statistics. (A) General organization of SEVA vectors as shown in the SEVA 4.0 homepage. All plasmids contain three basic modules (replication origin, antibiotic marker, and cargo), the boundaries between being punctuated by unusual restriction sites. These are accompanied by a primary gadget site (additional sites are also available to this end). Clicking each segment 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.

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 [1] (Fig. 1A)., This can be unequivocally cyphered according to a code (Fig. 2) which in this case needs to be updated as the plasmid collection

increases in number and utilities. Apart from 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.

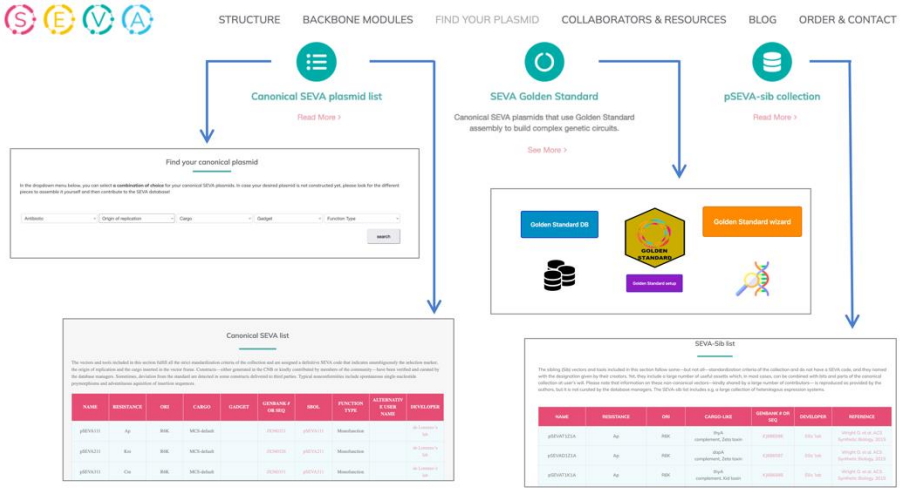

Figure 2. Details and utilities of the *Find you plasmid* section. The core of the database embodies the three subsections listed on top. Clicking each of them leads users to a different utility, the most useful of them being the Find your canonical plasmid, which can be composed according to necessities 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 lists of canonical and non-canonical (SEVAsib) vectors can be accessed as well.

Despite the number of excellent plasmid repositories available to molecular and synthetic biologists, 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 (replication origin, antibiotic markers, 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 the publication of a number of articles that further develop the concept in different directions explicitly related to the original platform [2, 3]. In other cases, authors take on the overall SEVA view for their genetic tool development progress but depart from the standard to different degrees and purposes [4, 5, 6, 7, 8, 9, 10]. 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 [11, 12, 13, 14, 15]. 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 [16]. 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 [17] asks also for novel, dedicated tools that are still in their infancy and deserve standardization efforts as well.

2 Site-specific sequence diversification of chromosomal segments.

The ability of T7 RNA polymerase ($RNAP^{T7}$) fusions to cytosine deaminases (CdA) for entering C→T changes in any DNA segment downstream of a T7 promoter was exploited for hyperdiversification of defined genomic portions of *Pseudomonas putida* KT2440. To this end, test strains were constructed in which the chromosomally encoded *pyrF* gene (the prokaryotic homologue of yeast URA3) was flanked by T7 transcription initiation and termination signals and also carried plasmids expressing constitutively either high-activity (lamprey's) or low-activity (rat's) CdA-$RNAP^{T7}$ fusions. The DNA segment-specific mutagenic action of these fusions was then tested in strains lacking or not uracil-DNA glycosylase (UDG), i.e., Δ*ung*/*ung*+ variants. The resulting diversification was measured by counting single nucleotide changes in clones resistant to fluoroorotic acid (5FOA), which otherwise is transformed by wild-type PyrF into a toxic compound. Although the absence of UDG dramatically increased mutagenic rates with both CdA-$RNAP^{T7}$ fusions, the most active variant—pmCDA1—caused the extensive appearance of 5FOA-resistant colonies in the wild-type strain not limited to C→T but including also a range of other changes. Furthermore, the presence/absence of UDG activity swapped cytosine deamination preference between DNA strands. These qualities provided the basis of a robust system for the continuous evolution of preset genomic portions of *P. putida* and beyond (Fig. 3).

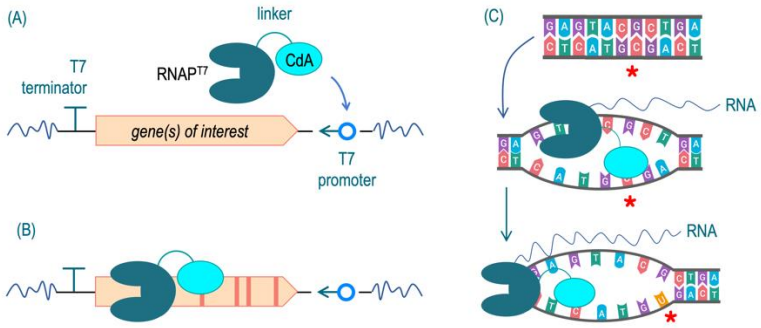

Figure 3. Schematic representation of the segment-specific mutagenic regime caused by CdA-$RNAP^{T7}$ fusions *in vivo*. (A) The basic arrangement involves one or more gene of interest—or any DNA sequence thereof— flanked by a T7 promoter and a T7 terminator. In the sketch, the direction of transcription from P_{T7} is opposite to native gene reading, but can be otherwise if desired. The CdA-$RNAP^{T7}$ fusion is then expressed in trans, bound to P_{T7} and transcription initiated by the polymerase activity of the fusion. (B) As CdA-$RNAP^{T7}$ proceeds downstream of the P_{T7}, the cytosine deamination moiety may find C residues and change them to U counterparts. (C) Mutation occurs as the ssDNA of the transcription bubble exposes C bases, which become substrates of CdA. If the resulting U is not corrected by the repair machinery of the cell, it will base-pair with A during the replication process, generating a change from the original C:G pair to a mutated T:A pair. Note that changes can occur, albeit with different frequencies in both DNA strands.

Recent strategies for all-*in vivo* continuous evolution of defined segments of bacterial genomes have been based the discovery that fusions of cytosine deaminases (CdA) to the bacteriophage T7 RNAP ($RNAP^{T7}$) caused high mutagenic rates C→T (and accordingly G→A) in any DNA sequence located downstream of a T7 promoter [18, 19, 20, 21, 22]. On this basis, a number of genetic platforms exploiting different base editors fused to $RNAP^{T7}$ have been developed in which the bases of the ssDNA that

become exposed along transcription become the substrate of the deamination reaction. This occurrence has been exploited to set off the evolution of delimited DNA segments *in vivo* without affecting the rest of the genome. Given the poor efficacy of T7 phage transcription terminators ($T7_T$) to make $RNAP^{T7}$ come off the engaged DNA [18, 23], one useful strategy to define the downstream limit of the mutagenized region involves the expression of gRNA-dCas9 complexes for blocking the progression of the polymerase [19]. While the available wealth of results accredits the mutagenic activity of CdA-$RNAP^{T7}$ fusions in *E. coli*, whether the same applies to other bacterial species and strains of biotechnological interest remains unknown. In this deliverable, we document the ability of CdA-$RNAP^{T7}$ fusions to boost DNA variability in a preset genomic segment of a *P. putida* KT2440 derivative delimited by a T7 promoter and a T7 terminator, the efficacy and types of mutations caused under various conditions (Fig. 4), and the confinement of the diversification regime to the desired DNA portion of the cell's genetic complement. The results not only verify the performance of the approach in this bacterium, but they also provide a complete genetic platform for the continuous evolution of specified sectors of the *P. putida* KT2440 chromosome.

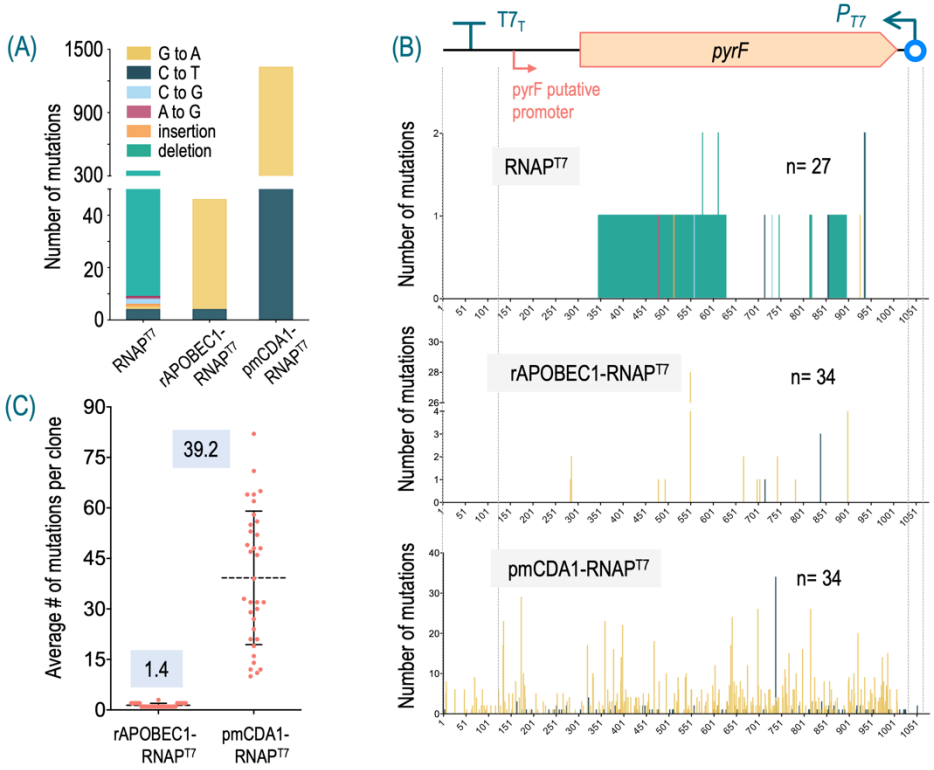

Figure 4. Characterization of *pyrF* mutations found in $5FOA^R$ colonies expressing CdAs. (A) Total number and types of mutations associated to each construct borne by *P. putida* PYRC Δ*ung*, indicating the base substitutions found. (B) Distribution and number of mutations throughput the *pyrF* segment of *P. putida* PYRC Δ*ung* (sketched on top) in $5FOA^R$ clones carrying the construct indicated and the number of clones analyzed in each case. The boundaries of the *pyrF* DNA sequence, the T7 promoter (P_{T7}) and T7 terminator ($T7_T$) are indicated along the location of the putative *pyrF* promoter. Note very different scales of the Y axis. The base changes are tagged into the coding sequence of *pyrF*. Mutations types are indicated with the same color codes. Extension of deletions are indicated with the boundaries of the nucleotides erased from the cognate DNA. (C) Average number of mutations per clone found in the $5FOA^R$ colonies analyzed for each CdA-$RNAP^{T7}$ fusions. Single values are represented with red dots, means and standard deviations are with black lines. Final figures are highlighted in each case.

3 Conclusions

The work presented in this deliverable provides the genetic tools and the variables to consider for setting continuous diversification of any DNA segment of the genome of *P. putida*. Since all constructs are borne by broad host range plasmids, it is plausible that the very same platform can be effectively reused in other Gram-negative bacteria. Specifically, by using a *pyrF*-based reporter system, the efficacy of each CdA-RNAP^{T7} fusion as a mutagenic agent has been determined, the role of UDG in the process settled, and the efficacy of a T7 terminator to restrain DNA diversification beyond a prefixed site documented. Given the degeneracy of the genetic code and that not every base change necessarily originates an inactive *pyrF* mutant, it is likely that the actual DNA diversification figures are at least 3-fold higher than the ones shown throughout this work. As indicated above, the desired mutagenesis levels can be modulated at will by picking one of the fusions available and playing with having either *ung*$^-$ or *ung*$^+$ strains as hosts of the process—or engineering a transient of UDG-minus phenotype upon the addition of a compatible plasmid expressing the inhibitory peptide UGI. Although not tested directly in this work, the length of the diversified genomic segment could be limited by engineering a series of terminators or a dCas9-based device [19] to inhibit the advancing CdA-RNAP^{T7}. In other cases, such *rowing base editors* can be let to proceed through longer DNA segments in the genome if desired. In sum, we believe that the hereby described devices may become a phenomenal addition to the already rich toolset available for many types of bioengineering endeavors in *P. putida* [24, 25].

Literature cited

1. 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 (2013).
2. Kim SH, Cavaleiro AM, Rennig M, Nørholm MHH. SEVA Linkers: A Versatile and Automatable DNA Backbone Exchange Standard for Synthetic Biology. *ACS Synthetic Biology* **5**, 1177-1181 (2016).
3. Valenzuela-Ortega M, French C. Joint universal modular plasmids (JUMP): a flexible vector platform for synthetic biology. *Synthetic Biology* **6**, (2021).
4. Lammens E-M, Boon M, Grimon D, Briers Y, Lavigne R. SEVAtile: a standardised DNA assembly method optimised for *Pseudomonas*. *Microbial Biotechnology* **15**, 370-386 (2022).
5. Schuster LA, Reisch CR. A plasmid toolbox for controlled gene expression across the Proteobacteria. *Nucleic Acids Research* **49**, 7189-7202 (2021).
6. Geddes BA, Mendoza-Suárez MA, Poole PS. A Bacterial Expression Vector Archive (BEVA) for Flexible Modular Assembly of Golden Gate-Compatible Vectors. *Front Microbiol* **9**, 3345 (2018).
7. Falkenberg KB, *et al.* The ProUSER2.0 Toolbox: Genetic Parts and Highly Customizable Plasmids for Synthetic Biology in *Bacillus subtilis*. *ACS Synthetic Biology* **10**, 3278-3289 (2021).
8. Opel F, *et al.* Generation of Synthetic Shuttle Vectors Enabling Modular Genetic Engineering of Cyanobacteria. *ACS Synthetic Biology* **11**, 1758-1771 (2022).
9. Silva-Rocha R, Pontelli MC, Furtado GP, Zaramela LS, Koide T. Development of New Modular Genetic Tools for Engineering the Halophilic Archaeon *Halobacterium salinarum*. *PLOS ONE* **10**, e0129215 (2015).

10. Bayer CN, Rennig M, Ehrmann AK, Nørholm MHH. A standardized genome architecture for bacterial synthetic biology (SEGA). *Nature Communications* **12**, 5876 (2021).

11. Elmore JR, Furches A, Wolff GN, Gorday K, Guss AM. Development of a high efficiency integration system and promoter library for rapid modification of *Pseudomonas putida* KT2440. *Metabolic Engineering Communications* **5**, 1-8 (2017).

12. Nora LC, Gonçales RA, Martins-Santana L, Ferreira BH, Rodrigues F, Silva-Rocha R. Synthetic and minimalist vectors for *Agrobacterium tumefaciens*-mediated transformation of fungi. *Genet Mol Biol* **42**, 395-398 (2019).

13. Verdú C, Sanchez E, Ortega C, Hidalgo A, Berenguer J, Mencía M. A Modular Vector Toolkit with a Tailored Set of Thermosensors To Regulate Gene Expression in *Thermus thermophilus*. *ACS Omega* **4**, 14626-14632 (2019).

14. Ma X, *et al.* A standard for near-scarless plasmid construction using reusable DNA parts. *Nature Communications* **10**, 3294 (2019).

15. Enghiad B, *et al.* PlasmidMaker is a versatile, automated, and high throughput end-to-end platform for plasmid construction. *Nature Communications* **13**, 2697 (2022).

16. De Wannemaeker L, Bervoets I, De Mey M. Unlocking the bacterial domain for industrial biotechnology applications using universal parts and tools. *Biotechnol Adv* **60**, 108028 (2022).

17. Leggieri PA, *et al.* Integrating systems and synthetic biology to understand and engineer microbiomes. *Annual Review of Biomedical Engineering* **23**, 169-201 (2021).

18. Moore CL, Papa LJ, 3rd, Shoulders MD. A Processive Protein Chimera Introduces Mutations across Defined DNA Regions In Vivo. *J Am Chem Soc* **140**, 11560-11564 (2018).

19. Alvarez B, Mencia M, de Lorenzo V, Fernandez LA. In vivo diversification of target genomic sites using processive base deaminase fusions blocked by dCas9. *Nat Commun* **11**, 6436 (2020).

20. Cravens A, Jamil OK, Kong D, Sockolosky JT, Smolke CD. Polymerase-guided base editing enables in vivo mutagenesis and rapid protein engineering. *Nat Commun* **12**, 1579 (2021).

21. Chen H, *et al.* Efficient, continuous mutagenesis in human cells using a pseudo-random DNA editor. *Nat Biotechnol* **38**, 165-168 (2020).

22. Park H, Kim S. Gene-specific mutagenesis enables rapid continuous evolution of enzymes in vivo. *Nucleic Acids Res* **49**, e32 (2021).

23. Tarnowski MJ, Gorochowski TE. Massively parallel characterization of engineered transcript isoforms using direct RNA sequencing. *Nat Commun* **13**, 434 (2022).

24. Martínez-García E, de Lorenzo V. Molecular tools and emerging strategies for deep genetic/genomic refactoring of Pseudomonas. *Curr Opin Biotechnol* **47**, 120-132 (2017).

25. Martin-Pascual M, *et al.* A navigation guide of synthetic biology tools for Pseudomonas putida. *Biotechnol Adv* **49**, 107732 (2021).

==========