

Published in final edited form as:

BMC Methods. ; 3(1): . doi:10.1186/s44330-025-00054-3.

Automated workflow for genotyping individual transposon library variants

Lorea Alejaldre^{1,2,*}, Ana-Mariya Anhel^{2,3}, Lewis Grozinger¹, Ángel Goñi-Moreno^{1,2,*}

¹Systems Biology Department, Centro Nacional de Biotecnología (CSIC), Madrid 28049, Spain.

²Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid (UPM)- Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA/CSIC), Madrid 28223, Spain.

³Department of Biochemistry, Cell and Systems Biology, University of Liverpool, Liverpool L69 7ZX, UK.

Abstract

Background: The fields of genome engineering and synthetic biology rely on high-throughput screening to accelerate research and increase reproducibility. This often requires the use of costly equipment such as flow cytometers, specialized microfluidic devices or liquid handlers. Despite the availability of population-level screening methods, genotyping individual clones in transposon libraries remains a necessary but frequently labor-intensive task well suited for liquid-handling platforms. Although open-source automation is democratizing access to liquid handlers, it requires a minimum level of programming expertise or the availability of validated protocols to implement experimental workflows. To address these challenges, we present a step-by-step automated workflow designed for individual variant genotyping of transposon insertion libraries using an open-source OT-2 liquid handler and a custom Python package to annotate sequencing results.

Methods: The automated protocol comprises six modular steps that can be used separately or integrated into different workflows: 1) colony picking into selective media, 2) counter-selection using the backbone and integrated antibiotic resistances and preparation of glycerol stocks pre-cultures, 3) colony counter-selection and 4) control PCR assays to filter spurious integration

This work is licensed under a [BY 4.0 International license](#).

*Correspondence: Lorea Alejaldre, lorea.alejaldre@cnb.csic.es, Ángel Goñi-Moreno, angel.goni@cnb.csic.es.

Author contributions

L.A. and A.G.M. conceived the whole study. A-M.A. and L.G. carried out the computational part of the work. L.A. did experiments, supervised the workflow and contributed to the analysis of the data. All authors contributed to the discussion of the research and writing of the manuscript.

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

events, 5) arbitrary PCR assays and 6) sequencing and annotation using a custom Python package. This protocol is demonstrated using standardized vectors adhering to the Standard European Vector Architecture (SEVA) guidelines and standardized Lab Automation Protocol (LAP) scripts to ensure reproducibility and portability.

Results: The automated workflow was validated using a transposon insertion library in *Pseudomonas putida* KT2440, enabling reliable genotyping of individual variants within 4–5 days. No spurious integrations were detected, and genomic insertion sites were successfully identified for 40 out of 46 variants, demonstrating robustness, accessibility, and scalability of the open-source workflow.

Discussion: This automated workflow enables reliable genotyping while reducing hands-on time and minimizing human error. Its modular, standardized design and validated protocols promote protocol reuse and sharing while broadening access to open-source laboratory automation. While the throughput of the workflow is limited by the deck capacity of the liquid handler, it can be scaled by subsequent runs or parallel processing on multiple liquid handlers. In addition, the workflow can be further streamlined through integration of an OT-2 adapted colony picker. Overall, this open-source workflow provides a robust foundation for scalable and reproducible genotyping of individual variants in genome engineering and synthetic biology applications.

Keywords

Open-source liquid handling; Automation; Transposon; Genotyping; Workflow; Standardization

Background

Transposon insertion is a valuable tool in genome engineering and synthetic biology. Genomic libraries created by random transposon insertion have been used, for example, to identify essential genes [1]. Since the characteristics of bacterial cellular process can be spatially heterogeneous [2–4], random transposon insertions have also been used to find the best performing sites for integration of synthetic gene circuits [5]. For both of these applications, genotyping, that is, locating transposon insertions in the genome, is key to establishing correlations between position and observed phenotype [6, 7]. Several methods have been established to genotype transposon insertion library pools in a high-throughput fashion [8]. These methods often involve the use of specific transposons and high-throughput sequencing [9] or microfluidic devices [10] that are not commonly available or are in a higher price range. Other efforts based on PCR methods for genotyping insertion sites in individual colonies have also been developed [11, 12]. However, scaling these approaches to screen large numbers of mutants requires the use of laboratory automation devices. Compared to manual PCR-based genotyping methods, automated workflows markedly reduce the number of hands-on time steps traditionally required. Manual protocols often involve repetitive and time-consuming tasks such as re-streaking individual colonies on agar plates, preparing and running PCR reactions by hand, manually performing counter-selection or verifying sequencing results individually [12]. These manual steps not only increase hands-on time but also introduce variability and potential human error, particularly

when handling large mutant libraries. In contrast, automated protocols integrate these steps through precise liquid handling and programmable workflows, improving reproducibility, minimizing human error, and enabling substantially higher throughput. Fortunately, these devices are becoming more commonplace in academic research laboratories, [13, 14] including open-source liquid-handling robots, which offer reduced costs [15, 16] and software flexibility [17]. The development of high-throughput protocols tailored for these open-source devices can thus accelerate established laboratory workflows [18, 19] towards automation, and facilitate tasks such as the transposon insertion library genotyping of individual variants.

Here, we describe the use of an automated workflow for the genotyping of transposon insertion libraries created with a SEVA-Sib standardized vector and using an open-source OT-2 liquid handler (Opentrons). Version 1 of this workflow was previously used for the identification of transposon insertion library variants of the soil bacterium *Pseudomonas putida* KT2440 created with the standardized plasmid set pBLAM1-x [20]. The same experimental conditions applied in this workflow can be used with other vectors that follow Standard European Vector Architecture (SEVA) guidelines [21, 22] such as the mini-Tn5 pBAMD1-x set [23]. The use of standardized vectors bearing transposases [20, 23] unifies experimental conditions regardless of the transposase used, maximizing reproducibility with different delivery systems. Nevertheless, this workflow can be easily modified to use with transposon libraries created with other non-standardized vectors by using oligonucleotides that bind inside the integrated cargo to perform arbitrary PCRs and to non-integrative parts of the vector to perform spurious PCR controls. This workflow, which is based on previously developed PCR based methods to genotype individual variants [11, 12], increases significantly the number of variants that can be screened while increasing reproducibility and protocol sharing between laboratories, both important aspects to consider when creating an experimental workflow [24]. To this end, open-source automation protocols can use standards like the LAP format [25], which is used for this workflow, to encourage protocol sharing and creation. The LAP format constitutes a set of guidelines for creating automation scripts which are modular and user-friendly, and which also offer portability between different automation equipment, facilitating the translation of workflows to different platforms [25]. The modularity of the LAP format is twofold: it provides a standardized framework for script development and individual protocols that can be seamlessly integrated into diverse experimental workflows [26]. Each script complying with the LAP format needs an input file where variables necessary to run each protocol are specified. This minimizes human error due to script editing, enhances reproducibility and facilitates implementation by end-users. For this workflow, these input files are in excel format to facilitate implementation in a format commonly used in wet laboratories settings. The intended user profile for this workflow is a wet-lab scientist with experience in standard microbiological techniques and a basic understanding of the Opentrons OT-2 platform. No programming knowledge is required; however, users should be comfortable executing simple bash commands in a terminal environment to run the provided Python scripts and process sequencing data. This automated genotyping workflow not only streamlines the identification of transposon insertion sites, but is also the first experimental workflow integrated with LAP protocols, contributing to standardized, reproducible methodologies

in genome engineering. For a single plate, the OT-2 executes the workflow more slowly than a human operator (Step 2 - counter-selection and glycerol stocks pre-cultures 9 min 37 s versus 4 min 25 s manually, Step 3-colony counter-selection 12 min 6 s versus 9 min 55 s, and Steps 4 and 5 - control and arbitrary PCRs 55 min 14 s versus 14 min 14 s). However, automation substantially reduces hands-on time, allows researchers to remain off the bench, and ensures reproducibility through detailed run logs, plate maps, and minimized human error. Step 3, in particular, benefits from automation, as it is prone to manual errors when selecting samples from two separate plates. When scaling to multiple plates or libraries, these timing differences become less critical: the OT-2 maintains consistent performance, whereas human operators experience fatigue and reduced speed. Overall, this automated workflow enhances robustness, reliability, and reproducibility, particularly for large or complex mutant libraries.

Materials and procedure

This workflow is divided into the following steps: 1) colony picking in selective media 2) counter-selection and glycerol stocks pre-cultures 3) Colony counter-selection 4) control PCRs 5) arbitrary PCRs and 6) Sequencing and annotation. An outline of the workflow is shown in Fig. 1. All steps follow a 96-well plate format for high-throughput analysis. Except for step 1), which requires a colony picker, all steps have been automated. Recently, the modification of the OT-2 liquid handler to perform as a colony picker has been implemented [27]. This could be integrated in future versions to make the workflow completely automated.

The workflow described in this peer-reviewed article has been published online on protocols.io (<http://dx.doi.org/10.17504/protocols.io.kqdg394jzg25/v2>) and is attached as Supplementary S1 File. Individual OT-2 protocols that form the workflow are also described in protocols.io: counter-selection (<http://dx.doi.org/10.17504/protocols.io0.5qpvor5xdv4o/v4>), Media dispensing and inoculation V3(<http://dx.doi.org/10.17504/protocols.io.q26g7yb3kgwz/v3>), PCR sample preparation (<http://dx.doi.org/10.17504/protocols.io.n92ldpyzn15b/v3>) and Bacterial Genome annotation using BLASTN V4(<http://dx.doi.org/10.17504/protocols.io.dm6gpjrb1gzp/v4>). Input files with variables used for running this workflow are available as S2–S6 supplementary files. Individual OT-2 protocols described in protocols.io are run with scripts that follow LAP format and that can be accessed through the LAP repository (<http://www.laprepo.com>) where more extensive documentation on the LAP format and experimental validation information for each script is available: counter-selection (<https://laprepo.com/protocol/2-criteria-counter-selection-v-2-0-0/>), Media dispensing and inoculation (<https://laprepo.com/protocol/cell-inoculation-in-different-media-v2-0-0/>), PCR sample preparation (<https://www.laprepo.com/protocol/pcr-mix-preparation-and-temperature-profile-v-2-0-0/>). All the scripts used in this workflow have been developed in Python 3.7.1, Opentrons Protocol API version 2.14 and Opentrons App software Version 7.0.2. Scripts generated with software versions are publicly available in the LAP repository(<https://www.laprepo.com>).

The Bacterial Genome and annotation using BLASTN protocol described in protocols.io does not strictly follow the LAP format but is also available in the LAP repository as

a LAPutility (<https://laprepo.com/protocol/bacterial-insertions-in-genome-annotation-using-blastn-v3-0-0/>). The LAP repository describes individual scripts but not specific step-by-step protocols or workflows, which to date have been described in protocols.io. When available, LAP repository entries provide links to protocols.io entries describing a step-by-step protocol to implement the script. Note that a LAP script could be found linked to different experimental workflows in protocols.io. To facilitate script development and download, the latest version of all of our LAP format scripts can also be directly cloned from the Biocomputation Lab's Github repository <https://github.com/BiocomputationLab/LAPrepository>. The genome annotation python package can also be directly cloned from the Biocomputation Lab's Github repository <https://github.com/BiocomputationLab/tnatlas>. Installation of the TnAtlas package can be done from the source code using pip install in the appropriate directory or directly from the python package repository Pypi using python3 -m pip install tnatlas. For TnAtlas to work, python ≥ 3.8 and blastn ≥ 2.12 should be installed. Installation of fastqc and sickle are required for sequencing quality control reports and for trimming based on sequencing quality. Supplementary files S8–S13 are labware descriptions used to run OT-2 automation protocols to run this workflow in our lab. In addition, a step-by-step guide on how to run scripts in the OT-2 is described in a separate protocols.io entry (<http://dx.doi.org/10.17504/protocols.io/0.3byl49jxjgo5/v1>).

Molecular biology reagents and oligonucleotides

Control PCRs and arbitrary PCR 1 are performed with Phire Green Hot Start II PCR Master mix (Thermo Fisher Scientific), arbitrary PCR 2 uses Phire Hot start II PCR Master Mix (Thermo Fisher Scientific). Oligonucleotides are described in Table 1.

Bacterial growth media, antibiotics and conjugation

Bacteria were grown in LB media (Invitrogen) or M9-citrate (1X Minimal Medium salts (MP Biomedicals), 2 mM Magnesium sulfate, 20 mM trisodium citrate). Antibiotic concentrations were 500 $\mu\text{g ml}^{-1}$ ampicillin (Fisher Scientific) and 50 $\mu\text{g ml}^{-1}$ kanamycin (Fisher Scientific). Bacterial conjugation was done using plasmid pBLAM1-2-TS2m as previously described [20].

General notes and troubleshooting

The workflow has been developed for an OT-2 liquid handler (opentrons) but its standardization using LAP format facilitates its translation to other automation devices. The scripts included in this workflow have been validated in the software versions described in the LAP repository (<http://www.laprepo.com>).

- Step 1 - Colony picking in selective media It should be noted that Step 1 has not been automated in this version of the workflow but could be automated in future versions through the use of a colony picker. This function could be implemented in an open-source OT-2 liquid handler [27], as the one used in this study, without significantly increasing the cost of the equipment necessary to operate this workflow.
- Optional step-Master 96-well plate There is an additional optional step in the workflow when genotyping multiple libraries at the same time after step 3)

Colony counter-selection, the creation of a master 96-well plate prior to step 4) control PCRs as shown in Fig. 1.

- **Step 4 - Control PCRs** This protocol has been validated using *Pseudomonas putida* KT2440 but it can be used without modification in other bacterial systems where transposon libraries are created using SEVA-Sib vectors. Implementation of this workflow in laboratories using non-SEVA-Sib vectors will only need to include their specific oligonucleotides and PCR conditions to implement this workflow. Oligonucleotides should anneal in non-integrative parts of the plasmid such as PS3/PS4 which amplify the antibiotic cassette and transposase and PS5/PS6 that amplify the origin or replication. Colony PCRs yield non-specific bands, therefore it is crucial to include a negative control, in this case the empty strain *Pseudomonas putida* KT2440. A positive control using the integrative plasmid also helps identifying spurious events, a band of 2964 bp for PS3/PS4 and a band at 665 bp for PS5/PS6 indicate that non-integrative parts of the plasmid are present. Note that even with non-specific bands due to colony PCR, when colonies are positive for spurious events, these become the predominant bands in the PCR. If no PCR product or very faint bands are obtained at this step, excessive colony material may have inhibited the reaction; repeating the PCR with a smaller amount of colony material is recommended. Rates of spurious events are related to the efficiency of integration, sequence to integrate and transposase used. If there are high-rates of spurious events detected, we recommend repeating the conjugation protocol or changing to a different transposase.
- **Step 5 - Arbitrary PCRs** Note that amplification with arbitrary PCR is seen as a smear even in successfully sequenced variants. At this step, the absence of a band in randomly sampled PCR products indicates that the reaction has failed. While this protocol has been validated using *Pseudomonas putida* KT2440, the oligonucleotides used in this step were originally designed for *Staphylococcus* [11] and have been used in other Gram-positive and Gram-negative bacteria.
- **Step 6 - Sequencing and annotation** Absence of sequencing results for individual clones can result from poor sequencing quality. We recommend repeating the PCR if insufficient sample diversity is available with alternative arbitrary primers [11] or purifying the product from the first arbitrary PCR amplification prior to running the second arbitrary PCR. Alternatively, multiple possible insertion sites may be detected when running TnAtlas. This typically occurs when the transposon integrates into repetitive genomic regions, such as ribosomal DNA. Because additional PCRs are unlikely to resolve such insertions, we recommend whole-genome sequencing of the variant if further characterization is required.

Expected results

This workflow constitutes a step-by-step guide to automatize transposon insertion genotyping by incorporating commonly used protocols [11, 23] into an automation format. The use of open-source automation and standardized plasmids [20, 22, 23] intends to increase reproducibility and provide higher accessibility to the development of the

workflow. This section is intended to provide the necessary information and files for experimental and automation scientists to validate the workflow in their own laboratory. The development of the workflow includes two overnight growths (excluding initial plating of colonies) and is generally completed in 4–5 days depending on the number of colonies to be analysed and liquid handlers available. The workflow herein described has been validated with a previously generated transposon insertion library through conjugation of plasmid pBLAM1-2-TS2m into the soil bacterium *Pseudomonas putida* KT2440 as acceptor strain. Individual colonies were picked into a 96-well plate containing M9-citrate selective media and the corresponding antibiotic marker kanamycin (Step 1). Subsequent steps have been automated and are operated in an OT-2 liquid handler by running the corresponding script in LAP format and variables file necessary for each step. After overnight growth at 30°C with shaking of the plate from Step 1, it was replicated in the OT-2 into a counter-selection plate containing the backbone antibiotic ampicillin in M9-citrate selective media and a plate containing Luria-Bertani media plus kanamycin using VariablesPlateIncubation.xlsx variables file (Supplementary S2 File) to run LAP-CellMediaInoculation-OT2-2.0.0 script included in OT-2 Media dispensing and culture inoculation protocol V3 (Step 2). Note that while the protocol is demonstrated using a single input plate for clarity, it can be run sequentially or in parallel across multiple OT-2 liquid handlers. After overnight growth at 30°C with shaking of plates obtained in Step 2, the OD_{600 nm} of cultures grown in selective media with the cargo antibiotic kanamycin and the backbone antibiotic ampicillin was measured and values were entered into VariablesCounterSelection.xlsx (Supplementary S3 File) required to run LAP-ColonyCounterSelection-OT2-2.0.0 script included in OT-2 Counter-Selection V4 protocol (Step 3). The number of wells selected by the liquid handler after running the script revealed the number of colonies that had a spurious event and grew in the presence of the backbone antibiotic ampicillin (1/96 colonies). A low number of wells growing with the backbone antibiotic is optimal, however this depends on the efficiency of the integration. Selected wells that presented growth in the cargo antibiotic and no growth in the backbone antibiotic were inoculated into a 96-well plate containing 30% glycerol and a 96-well plate containing molecular grade milliQ water for subsequent PCR steps (Steps 4 and 5). After running the script, the layout of the selected colonies is obtained as output file map_CounterSelectionTS2.xlsx shown in Table 2. This layout serves as the map of selected colonies for the final glycerol stock and 96-well plate that will be used for subsequent PCRs in steps 4 and 5 of the workflow. Note that the name of the transposon insertion library to be analysed (i.e TS2m) can be included in VariablesCounter-Selection.xlsx (Supplementary S3 File). Step 3 can be performed with up to two initial 96-well plates in the same run.

The workflow was subsequently run for the first 46 samples on the 96-well plate shown in Table 2 to allow the following steps to be completed within a single plate for both primer pairs including positive and negative controls, though the approach can be expanded to accommodate up to 4 complete additional 96-well plates or 384 PCR reactions. Although this setup is not demonstrated here, the workflow supports parallel processing of multiple libraries up to step 3, which can then be combined in a Master 96-well plate using running LAP-NplateMerging-OT2-2.0.0.py before initiating step 4. As spurious events need to be controlled for when creating transposon insertion libraries [23], control PCRs using

oligonucleotides PS3/PS4 amplifying the antibiotic cassette and transposase and PS5/PS6 that amplify the origin or replication, non-integrative parts of the SEVA-Sib plasmids [21], were performed in the OT-2 (Step 4). The number of wells to test, identity of each well in source plate, number of primer pairs and name of the file containing the final layout of the plate was specified in variablesPCR.xlsx (Supplementary S4 File) needed to run LAP-PCR-OT2-2.0.0 script included in OT-2 PCR sample preparation protocol V3. It should be highlighted that the map shown in Table 2 was used as input in sheet “mapPlate1” in Supplementary S4 File. After running the protocol, an output file with the final layout of the spurious PCR controls with primer pairs PS3/PS4 and PS5/PS6 is obtained (Table 3). Note that wells E12 and G12 were inoculated with the wild-type KT2440 and wells F12 and H12 with plasmid pBLAM1-2 to use as negative and positive controls respectively for PS3/PS4 and PS5/PS6 reactions. These controls are crucial to determine the existence of spurious controls as colony PCR amplifications yield non-specific bands. Agarose gels of the PCR products revealed no detectable spurious events in the integration of the library Fig. 2. In agreement with the low number of colonies that grow with the backbone antibiotic, this result indicates efficient integration of our construct with a negligible number of spurious integration events.

Since all variants were negative for spurious events, we proceeded to run arbitrary PCRs [23] (Step 5), starting with arbitrary PCR1 using primer pairs ARB6/ME-O-Km-Ext-F. The file variablesPCR.xlsx was modified to perform arbitrary PCR1 and used to run LAP-PCR-OT2-2.0.0 script included in OT-2 PCR sample preparation protocol V2. The updated variablesPCR.xlsx used for arbitrary PCR1 is available in Supplementary S5 File. To run LAP-PCR-OT2-2.0.0 to perform arbitrary PCR2 the file variablesPCR.xlsx was modified to include the 96-well skirted plate from arbitrary PCR1 as source plate (Supplementary S6 File). Primer pair ARB2/ME-O-Km-Int-F is used for arbitrary PCR 2. In each step a.xlsx file containing the layout of the variants was obtained, the name of this file can be modified in variablesPCR.xlsx file to include relevant keywords for a given experiment. For arbitrary PCR1 and PCR2 this layout was identical, corresponding to that in Table 4. An agarose gel with the PCR products of a few variants verified the correct amplification Fig. 3.

Last in the workflow is sequencing and annotation (Step 6), 10 μ L of each PCR product obtained after arbitrary PCR 2 was sent to sequencing with 10 μ L of ME-O-Km-Int-F oligonucleotide to Macrogen, Spain. Sequencing results were used as input for the tfind tool of the Python package TnAtlas, as described in the protocol Bacterial genome annotation using BLASTN. Other files used as input were the *Pseudomonas putida* KT2440 TIGR reference genome in GENBANK format and GENBANK sequences used to identify transposon integrations. The tfind tool produced a summary output in excel format with a row for each probable alignment to the reference genome and columns as described in Table 5.

The results confirmed the insertion into the genome of 40 out of 46 variants (Supplementary S7 File). Low sequencing quality and not the automation protocol prevented determination of the genomic location of the remaining variants either due to low amplification efficiency from arbitrary primers or colony PCR impurities. This can be solved by purifying the product from the first arbitrary PCR amplification prior to proceeding or by using alternative

arbitrary primers [11]. However, enough variants distributed across the genome were obtained for posterior phenotypic analysis. Note that when a sequence aligns multiple times to the reference genome and its location cannot be uniquely determined this is indicated as the “multiple candidates” column being True. Other parameters related to alignment performance are shown in the raw output file for the annotation script in Supplementary S7 File. Overall, the description of this workflow provides evidence of its validity to identify genomic insertions in an accessible pipeline that uses open-source automation. The scale of our experimental demonstration effectively validates the workflow; however, it can be further scaled by increasing the number of variants per library or processing multiple libraries in parallel prior to scaling up. The use of excel format to input script variables allows an easier implementation by experimental scientists with little background in automation. Finally, the use of LAP format facilitates script development for future versions in laboratory settings with automation expertise.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Not applicable.

Funding

Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This work was supported by the the Contract No. 101,044,360 of the EU (ERC-2021-COG ECCO), Grant No. Y2020/TCS-6555 (BioSinT-CM) funded by the Comunidad de Madrid, and Grants No. PID2020-117205 GA-I00 (MULTI-SYBIO), No. CNS2022-135951 (BIOELECTRIC), and No. PID2023-152470NB-I00 (MULTI-SYNBIO) funded by MICIU/AEI/10.13039/501100011033 and FEDER (UE).

Data availability

All information and data needed to reproduce the study is available in the protocols.io repository, <http://dx.doi.org/10.17504/protocols.io.kqdg394jzg25/v2>, and in the Supplementary Information online

References

1. van Opijnen T, Levin HL. Transposon insertion sequencing, a global measure of gene function. *Annu Rev Genet.* 2020; 54 (1) 337–65. DOI: 10.1146/annurev-genet-112618-043838 [PubMed: 32886545]
2. Montero Llopis P, Jackson AF, Sliusarenko O, Surovtsev I, Heinritz J, Emonet T, et al. Spatial organization of the flow of genetic information in bacteria. *Nature.* 2010; 466 (7302) 77–81. DOI: 10.1038/nature09152 [PubMed: 20562858]
3. Kim J, Goñi-Moreno A, de Lorenzo V. Subcellular Architecture of the xyl gene expression flow of the TOL catabolic plasmid of *Pseudomonas putida* mt-2. *Mbio.* 2021; 12 (1) 10–1128. DOI: 10.1128/mBio.03685-20 [PubMed: 33622725]
4. Kim J, Goñi-Moreno A, Calles B, de Lorenzo V. Spatial organization of the gene expression hardware in *Pseudomonas putida*. *Environ Microbiol.* 2019; 21 (5) 1645–58. [PubMed: 30689295]
5. Hueso-Gil A, Calles B, de Lorenzo V. In vivo sampling of intracellular heterogeneity of *Pseudomonas putida* enables multiobjective optimization of genetic devices. *ACS Synth Biol.* 2023; doi: 10.1021/acssynbio.3c00009 [PubMed: 37196337]

6. Goñi-Moreno Á, Benedetti I, Kim J, de Lorenzo V. Deconvolution of gene expression noise into spatial dynamics of transcription factor–promoter interplay. *ACS Synth Biol.* 2017; 6 (7) 1359–69. DOI: 10.1021/acssynbio.6b00397 [PubMed: 28355056]
7. Stoof R, Wood A, Goñi-Moreno A. A model for the spatiotemporal design of gene regulatory circuits. *ACS Synth Biol.* 2019; 8 (9) 2007–16. [PubMed: 31429541]
8. Cain AK, Barquist L, Goodman AL, Paulsen IT, Parkhill J, van Opijnen T. A decade of advances in transposon-insertion sequencing. *Nat Rev Genet.* 2020; 21 (9) 526–40. DOI: 10.1038/s41576-020-0244-x [PubMed: 32533119]
9. van Opijnen T, Bodi KL, Camilli A. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat Methods.* 2009; 6 (10) 767–72. DOI: 10.1038/nmeth.1377 [PubMed: 19767758]
10. Thibault D, Jensen PA, Wood S, Qabar C, Clark S, Shainheit MG, et al. Droplet tn-Seq combines microfluidics with tn-seq for identifying complex single-cell phenotypes. *Nat Commun.* 2019; 10 (1) 5729 doi: 10.1038/s41467-019-13719-9 [PubMed: 31844066]
11. Das S, Noe JC, Paik S, Kitten T. An improved arbitrary primed PCR method for rapid characterization of transposon insertion sites. *J Microbiol Methods.* 2005; 63 (1) 89–94. [PubMed: 16157212]
12. Martínez-García, E, Aparicio, T, de Lorenzo, V, Nikel, PI. In vitro mutagenesis Methods in Molecular Biology. Reeves, A, editor. Vol. 1498: Humana Press; New York (NY): 2017. 273–293.
13. Kitney R, Adeogun M, Fujishima Y, Goñi-Moreno Á, Johnson R, Maxon M, et al. Enabling the advanced bioeconomy through public policy supporting bio-foundries and engineering biology. *Trends In Biotechnol.* 2019; 37 (9) 917–20. [PubMed: 31036350]
14. Tellechea-Luzardo J, Otero-Muras I, Goñi-Moreno A, Carbonell P. Fast biofoundries: coping with the challenges of biomanufacturing. *Trends In Biotechnol.* 2022. [PubMed: 35012773]
15. Tegally H, San JE, Giandhari J, de Oliveira T. Unlocking the efficiency of genomics laboratories with robotic liquid-handling. *BMC Genomics.* 2020; 21 (1) 729. doi: 10.1186/s12864-020-07137-1 [PubMed: 33081689]
16. Torres-Acosta MA, Lye GJ, Dikicioglu D. Automated liquid-handling operations for robust, resilient, and efficient bio-based laboratory practices. *Biochemical Eng J.* 2022; 188 108713 doi: 10.1016/j.bej.2022.108713
17. Chory EJ, Gretton DW, DeBenedictis EA, Esvelt KM. Enabling high-throughput biology with flexible open-source automation. *Mol Syst Biol.* 2021; 17 (3) e9942 doi: 10.15252/msb.20209942 [PubMed: 33764680]
18. Beal J, Weiss R, Densmore D, Adler A, Appleton E, Babb J, et al. An end-to-end workflow for engineering of biological networks from high-level specifications. *ACS Synth Biol.* 2012; 1 (8) 317–31. [PubMed: 23651286]
19. Goñi-Moreno A, Carajona M, Kim J, Martinez-Garcia E, Amos M, de Lorenzo V. An implementation-focused bio/algorithmic workflow for synthetic biology. *ACS Synth Biol.* 2016; 5 (10) 1127–35. [PubMed: 27454551]
20. Alejaldre L, Anhel AM, Goñi-Moreno. pBLAM1-x: standardized transposon tools for high-throughput screening. *Synth Biol.* 2023; 8 (1) ysad012 doi: 10.1093/synbio/ysad012 [PubMed: 37388964]
21. Martínez-García 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.* 2015; 43 (D1) D1183–89. DOI: 10.1093/nar/gku1114 [PubMed: 25392407]
22. Martínez-García E, Fraile S, Algar E, Aparicio T, Velázquez E, Calles B, et al. SEVA 4.0: an update of the Standard European vector Architecture database for advanced analysis and programming of bacterial phenotypes. *Nucleic Acids Res.* 2023; 51 (D1) D1558–67. DOI: 10.1093/nar/gkac1059 [PubMed: 36420904]
23. Martínez-García E, Aparicio T, de Lorenzo V, Nikel PI. New transposon tools tailored for metabolic engineering of Gram-negative microbial cell factories. *Front Bioeng and Biotechnol.* 2014; 2 doi: 10.3389/fbioe.2014.00046 [PubMed: 25389526]
24. Jessop-Fabre MM, Sonnenschein N. Improving reproducibility in synthetic biology. *Front Bioeng and Biotechnol.* 2019; 7 doi: 10.3389/fbioe.2019.00018 [PubMed: 30805337]

25. Anhel AM, Alejaldre L, Goñi-Moreno A. The laboratory automation protocol (lap) format and repository: a platform for enhancing workflow efficiency in synthetic biology. *ACS Synth Biol.* 2023; doi: 10.1021/acssynbio.3c00397 [PubMed: 37982688]
26. Fananás-Pueyo I, Anhel AM, Goñi-Moreno A, Onate-Sánchez L, Carrera-Castano G. Workflow to select functional promoter DNA baits and screen arrayed gene libraries in yeast. *Current protocols.* 2024; 4 e70059 doi: 10.1002/cpz1.70059 [PubMed: 39570200]
27. Del Olmo Lianes I, Yubero P, Gómez-Luengo N, Espeso JDR. Technical upgrade of an open-source liquid handler to support bacterial colony screening. *Front Bioeng and Biotechnol.* 2023; 11 1202836 doi: 10.3389/fbioe.2023.1202836 [PubMed: 37404684]

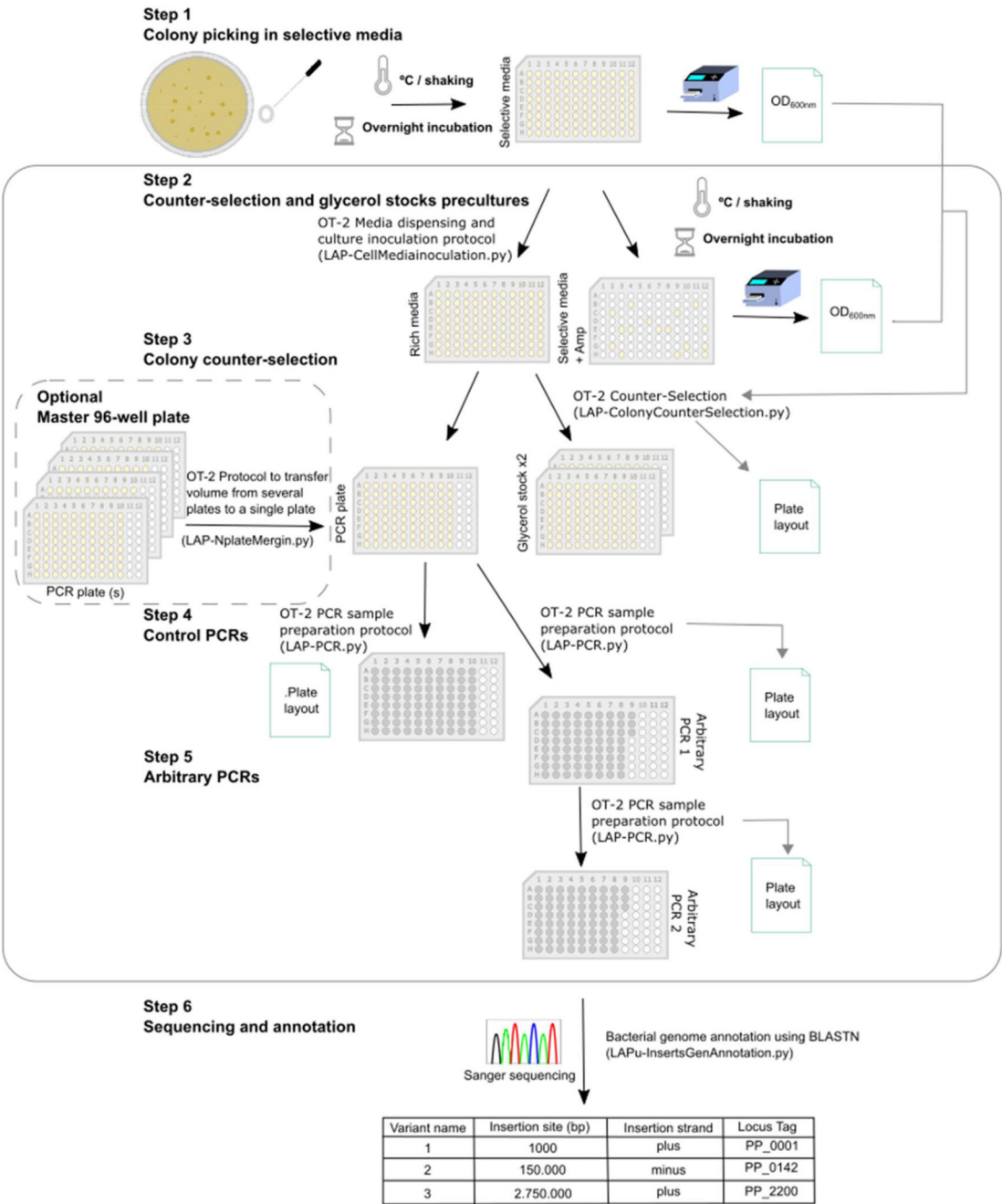


Fig. 1. Schematic representation of workflow outline. Steps 2 to 5 are performed in an OT-2 by running standardized lap scripts. An optional Master 96-well plate step resulting from previous or parallel runs of the workflow is also indicated

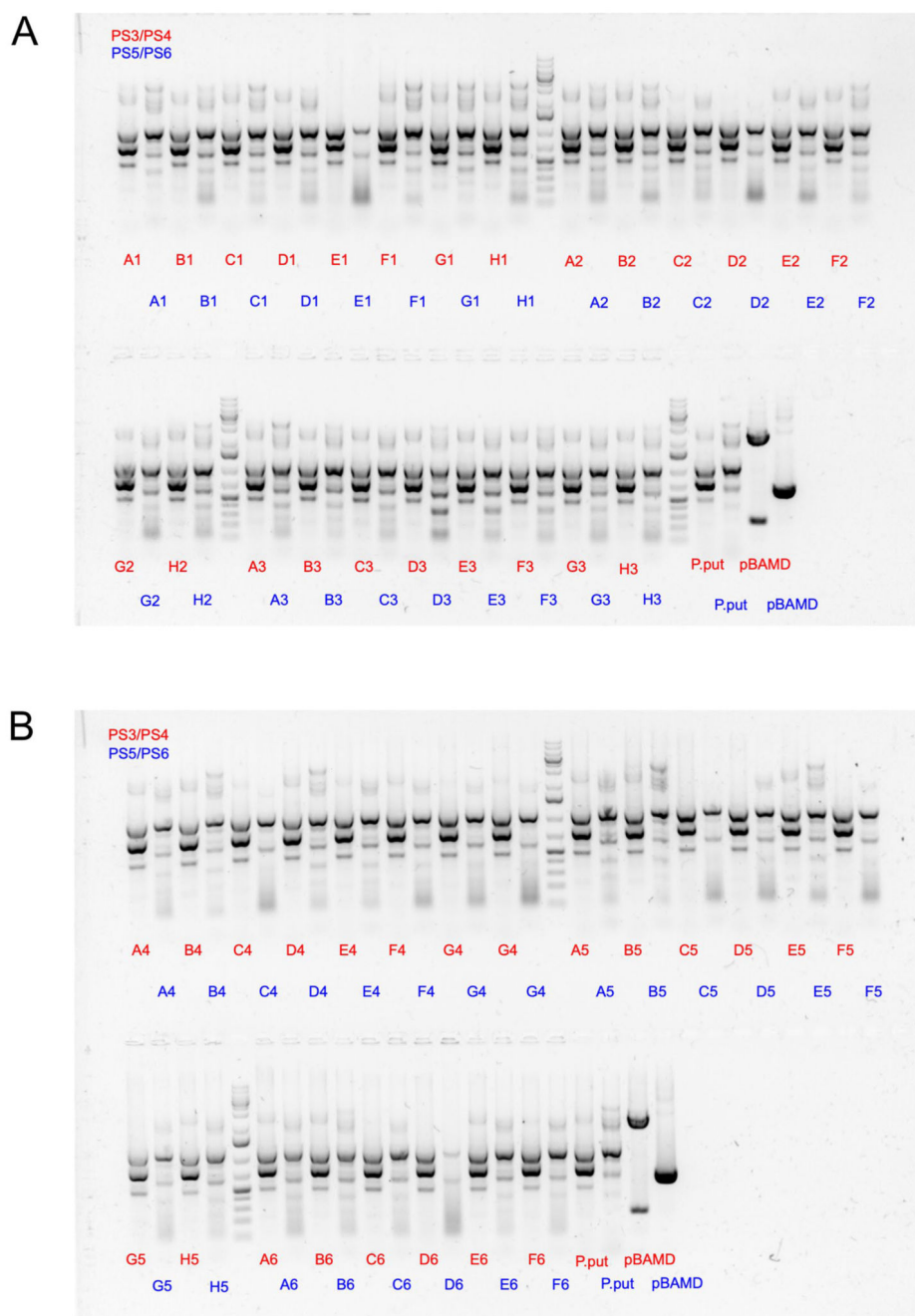


Fig. 2. Spurious control PCR results. Spurious control PCRs with primer pairs PS3/PS4 (red) and PS5/PS6 (blue). Expected band sizes for PS3/PS4 positive control is 2494 bp and 665 bp for PS5/PS6. Note that numbering corresponds to that of the 96-well plate and not to that of the variant names. PCR reactions for PS3/PS4 and PS5/PS6 from wells A1 to H3 (**A**) and from wells A4 to F6 (**B**). GeneRuler 1kb plus DNA ladder (Fisher Scientific) is used in all gels as reference

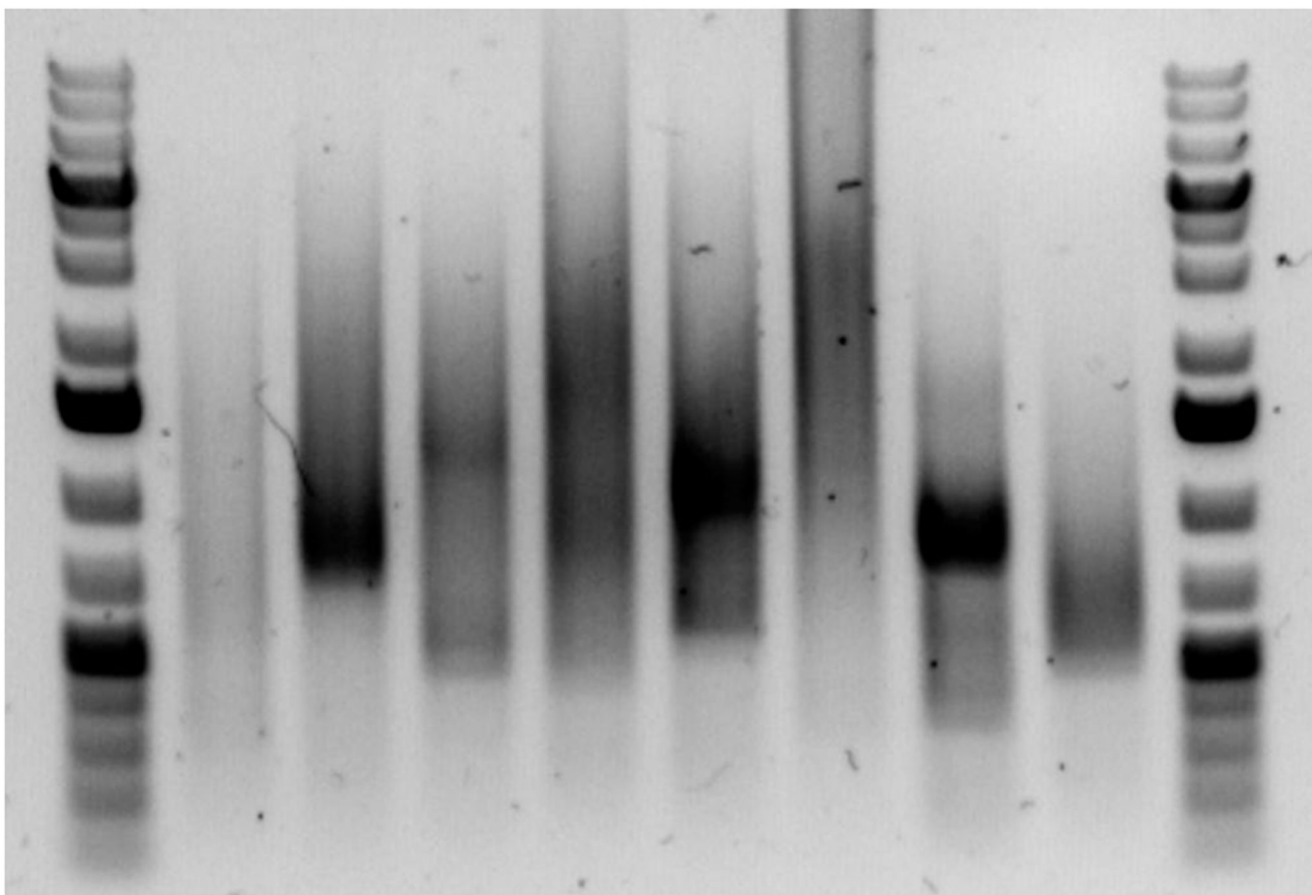


Fig. 3.
Arbitrary PCR2 of randomly selected variants. GeneRuler 1kb plus DNA ladder (Fisher Scientific) is used as reference

Table 1 Oligonucleotides used for control PCRs and arbitrary PCRs

Name	Sequence	Workflow step	Reference
PS3	GAACGCTCGGTTGCCGC	Step 4 - Control PCRs	[23]
PS4	CCAGCCTCGCAGAGCAGG	Step 4 - Control PCRs	[23]
PS5	CCCTGCTTCGGGGTCATT	Step 4 - Control PCRs	[23]
PS6	GGACAAATCCGCCGCCCT	Step 4 - Control PCRs	[23]
ME-O-Km-Ext-F	CGTCTGTTTCAGAAATATGGCAT	Step 5 - Arbitrary PCR1	[23]
ARB6	GGCACGCGTCGACTAGTACNNNNNNNNNACGCC	Step 5 - Arbitrary PCR1	[11,23]
ME-O-Km-Int-F	ATCTGATGCTGGATGAATTTTC	Step 5 - Arbitrary PCR2	[23]
ARB2	GGCACGCGTCGACTAGTAC	Step 5 - Arbitrary PCR2	[11,23]

Table 2 Layout of selected colonies after counter-selection

Row	Column											
	1	2	3	4	5	6	7	8	9	10	11	12
A	A1 TS2m	A2 TS2m	A3 TS2m	A4 TS2m	A5 TS2m	A6 TS2m	A7 TS2m	A8 TS2m	A9 TS2m	B10 TS2m	B11 TS2m	B12 TS2m
B	B1 TS2m	B2 TS2m	B3 TS2m	B4 TS2m	B5 TS2m	B6 TS2m	B7 TS2m	B8 TS2m	B9 TS2m	C10 TS2m	C11 TS2m	C12 TS2m
C	C1 TS2m	C2 TS2m	C3 TS2m	C4 TS2m	C5 TS2m	C6 TS2m	C7 TS2m	C8 TS2m	C9 TS2m	D10 TS2m	D11 TS2m	D12 TS2m
D	D1 TS2m	D2 TS2m	D3 TS2m	D4 TS2m	D5 TS2m	D6 TS2m	D7 TS2m	D8 TS2m	D9 TS2m	E10 TS2m	E11 TS2m	E12 TS2m
E	E1 TS2m	E2 TS2m	E3 TS2m	E4 TS2m	E5 TS2m	E6 TS2m	E7 TS2m	E8 TS2m	E9 TS2m	F10 TS2m	F11 TS2m	F12 TS2m
F	F1 TS2m	F2 TS2m	F3 TS2m	F4 TS2m	F5 TS2m	F6 TS2m	F7 TS2m	F8 TS2m	F9 TS2m	G10 TS2m	G11 TS2m	G12 TS2m
G	G1 TS2m	G2 TS2m	G3 TS2m	G4 TS2m	G5 TS2m	G6 TS2m	G7 TS2m	G8 TS2m	H9 TS2m	H10 TS2m	H11 TS2m	H12 TS2m
H	H1 TS2m	H2 TS2m	H3 TS2m	H4 TS2m	H5 TS2m	H6 TS2m	H7 TS2m	H8 TS2m	A10 TS2m	A11 TS2m	A12 TS2m	

Table 3 Layout of selected colonies for spurious control PCR with primer pairs PS3/PS4 and PS5/PS6

Row	Column											
	1	2	3	4	5	6	7	8	9	10	11	12
A	A1 TS2m Slot 1 with Set 1	A2 TS2m Slot 1 with Set 1	A3 TS2m Slot 1 with Set 1	A4 TS2m Slot 1 with Set 1	A5 TS2m Slot 1 with Set 1	A6 TS2m Slot 1 with Set 1	C1 TS2m Slot 1 with Set 2	C2 TS2m Slot 1 with Set 2	C3 TS2m Slot 1 with Set 2	C4 TS2m Slot 1 with Set 2	C5 TS2m Slot 1 with Set 2	C6 TS2m Slot 1 with Set 2
B	B1 TS2m Slot 1 with Set 1	B2 TS2m Slot 1 with Set 1	B3 TS2m Slot 1 with Set 1	B4 TS2m Slot 1 with Set 1	B5 TS2m Slot 1 with Set 1	B6 TS2m Slot 1 with Set 1	D1 TS2m Slot 1 with Set 2	D2 TS2m Slot 1 with Set 2	D3 TS2m Slot 1 with Set 2	D4 TS2m Slot 1 with Set 2	D5 TS2m Slot 1 with Set 2	D6 TS2m Slot 1 with Set 2
C	C1 TS2m Slot 1 with Set 1	C2 TS2m Slot 1 with Set 1	C3 TS2m Slot 1 with Set 1	C4 TS2m Slot 1 with Set 1	C5 TS2m Slot 1 with Set 1	C6 TS2m Slot 1 with Set 1	E1 TS2m Slot 1 with Set 2	E2 TS2m Slot 1 with Set 2	E3 TS2m Slot 1 with Set 2	E4 TS2m Slot 1 with Set 2	E5 TS2m Slot 1 with Set 2	E6 TS2m Slot 1 with Set 2
D	D1 TS2m Slot 1 with Set 1	D2 TS2m Slot 1 with Set 1	D3 TS2m Slot 1 with Set 1	D4 TS2m Slot 1 with Set 1	D5 TS2m Slot 1 with Set 1	D6 TS2m Slot 1 with Set 1	F1 TS2m Slot 1 with Set 2	F2 TS2m Slot 1 with Set 2	F3 TS2m Slot 1 with Set 2	F4 TS2m Slot 1 with Set 2	F5 TS2m Slot 1 with Set 2	F6 TS2m Slot 1 with Set 2
E	E1 TS2m Slot 1 with Set 1	E2 TS2m Slot 1 with Set 1	E3 TS2m Slot 1 with Set 1	E4 TS2m Slot 1 with Set 1	E5 TS2m Slot 1 with Set 1	E6 TS2m Slot 1 with Set 1	G1 TS2m Slot 1 with Set 2	G2 TS2m Slot 1 with Set 2	G3 TS2m Slot 1 with Set 2	G4 TS2m Slot 1 with Set 2	G5 TS2m Slot 1 with Set 2	G6 TS2m Slot 1 with Set 2
F	F1 TS2m Slot 1 with Set 1	F2 TS2m Slot 1 with Set 1	F3 TS2m Slot 1 with Set 1	F4 TS2m Slot 1 with Set 1	F5 TS2m Slot 1 with Set 1	F6 TS2m Slot 1 with Set 1	H1 TS2m Slot 1 with Set 2	H2 TS2m Slot 1 with Set 2	H3 TS2m Slot 1 with Set 2	H4 TS2m Slot 1 with Set 2	H5 TS2m Slot 1 with Set 2	H6 TS2m Slot 1 with Set 2
G	G1 TS2m Slot 1 with Set 1	G2 TS2m Slot 1 with Set 1	G3 TS2m Slot 1 with Set 1	G4 TS2m Slot 1 with Set 1	G5 TS2m Slot 1 with Set 1	G6 TS2m Slot 1 with Set 1	A1 TS2m Slot 1 with Set 2	A2 TS2m Slot 1 with Set 2	A3 TS2m Slot 1 with Set 2	A4 TS2m Slot 1 with Set 2	A5 TS2m Slot 1 with Set 2	A6 TS2m Slot 1 with Set 2
H	H1 TS2m Slot 1 with Set 1	H2 TS2m Slot 1 with Set 1	H3 TS2m Slot 1 with Set 1	H4 TS2m Slot 1 with Set 1	H5 TS2m Slot 1 with Set 1	B1 TS2m Slot 1 with Set 1	B2 TS2m Slot 1 with Set 2	B3 TS2m Slot 1 with Set 2	B4 TS2m Slot 1 with Set 2	B5 TS2m Slot 1 with Set 2	B6 TS2m Slot 1 with Set 2	B7 TS2m Slot 1 with Set 2

Table 4 Layout of selected colonies after arbitrary PCR1 and PCR2

Row	Column	1	2	3	4	5	6	7	8	9	10	11	12
A	A1 TS2m Slot 1 with Set 1	A2 TS2m Slot 1 with Set 1	A3 TS2m Slot 1 with Set 1	A4 TS2m Slot 1 with Set 1	A5 TS2m Slot 1 with Set 1	A6 TS2m Slot 1 with Set 1							
B	B1 TS2m Slot 1 with Set 1	B2 TS2m Slot 1 with Set 1	B3 TS2m Slot 1 with Set 1	B4 TS2m Slot 1 with Set 1	B5 TS2m Slot 1 with Set 1	B6 TS2m Slot 1 with Set 1							
C	C1 TS2m Slot 1 with Set 1	C2 TS2m Slot 1 with Set 1	C3 TS2m Slot 1 with Set 1	C4 TS2m Slot 1 with Set 1	C5 TS2m Slot 1 with Set 1	C6 TS2m Slot 1 with Set 1							
D	D1 TS2m Slot 1 with Set 1	D2 TS2m Slot 1 with Set 1	D3 TS2m Slot 1 with Set 1	D4 TS2m Slot 1 with Set 1	D5 TS2m Slot 1 with Set 1	D6 TS2m Slot 1 with Set 1							
E	E1 TS2m Slot 1 with Set 1	E2 TS2m Slot 1 with Set 1	E3 TS2m Slot 1 with Set 1	E4 TS2m Slot 1 with Set 1	E5 TS2m Slot 1 with Set 1	E6 TS2m Slot 1 with Set 1							
F	F1 TS2m Slot 1 with Set 1	F2 TS2m Slot 1 with Set 1	F3 TS2m Slot 1 with Set 1	F4 TS2m Slot 1 with Set 1	F5 TS2m Slot 1 with Set 1	F6 TS2m Slot 1 with Set 1							
G	G1 TS2m Slot 1 with Set 1	G2 TS2m Slot 1 with Set 1	G3 TS2m Slot 1 with Set 1	G4 TS2m Slot 1 with Set 1	G5 TS2m Slot 1 with Set 1								
H	H1 TS2m Slot 1 with Set 1	H2 TS2m Slot 1 with Set 1	H3 TS2m Slot 1 with Set 1	H4 TS2m Slot 1 with Set 1	H5 TS2m Slot 1 with Set 1								

Table 5 The fields of the summary output excel file produced by the tfind tool, with a short description and example values for each

Field	Description	Example
name	Name of the sequenced sample	23CACAA000-1_premix
read length	Length of the sequence (after trimming)	398
multiple candidates	True if there are multiple alignments to the genome with equal scores, false otherwise	True
transposon	Name of the transposon sequence that is present in the sample	pBLAM1-2_IR
transposon eval	BLASTN expected value for the alignment of transposon to sample (lower is better)	4,16713E-35
transposon bit score	BLASTN bit score for the alignment of transposon to sample (higher is better)	130,386
transposon identity	Proportion of exact matches in alignment of transposon to sample (higher is better)	0,986486486
transposon length	Length of the alignment between transposon and sample	74
transposon start	Position in sample where alignment with transposon begins	0
transposon end	Position in sample where alignment with transposon ends	73
genome offset	Number of base pairs between end of transposon alignment and beginning of genome alignment	0
gap sequence	Sequence of sample between end of transposon alignment and beginning of genome alignment	ATCG
insertion locus	Locus in the genome at which the genome alignment begins, if any	['PP_0168']
insertion gene	If insertion locus is a gene, the gene name, otherwise 'None'	[None]
sequence type	GENBANK sequence type of the insertion locus	['CDS']
insertion strand	Strand orientation of the insertion locus	['plus']
insertion product	Product(s) of the insertion locus, if any	['putative surface adhesion protein']
genome	Accession of the genome to which this sample was aligned	AE015451.2
genome alignment eval	BLASTN expected value for the alignment of genome to sample (lower is better)	1,1589E-115
genome alignment bit score	BLASTN bit score for the alignment of genome to sample (higher is better)	412,923
genome alignment identity	Proportion of exact matches in alignment of transposon to sample (higher is better)	0,896341463
genome alignment length	Length of the alignment between genome and sample	328
genome alignment mismatch	Number of mismatches in the alignment between genome and sample	34
genome alignment gaps	Number gaps in the alignment between genome and sample	6
genome start	Position in genome where alignment with sample begins	195745
genome end	Position in genome where alignment with sample ends	197870
genome strand	Strand orientation in the genome of the alignment between genome and sample	minus
genome loci	All genomic loci involved in the alignment	['PP_0168', 'PP_0169']
genome prefix	First <i>N</i> base pairs of the genomic alignment (default <i>N</i> = 9)	TACCCGCTT
genome alignment	Details of alignment of genome prefix to genome. indicates exact match.. indicates mismatch, – indicates a gap	–