Preprint DOI: Posted on April 4, 2024
https://doi.org/10.1099/acmi.0.000755.v2

An improved genome editing system for Sphingomonadaceae

Inmaculada García-Romero[1,†], Rubén de Dios[2,†,*] and Francisca Reyes-Ramírez[1,*]

1. Departamento de Biología Molecular e Ingeniería Bioquímica, Centro Andaluz de Biología del Desarrollo, Universidad Pablo de Olavide/Consejo Superior de Investigaciones Científicas/Junta de Andalucía, 41013 Sevilla, Spain.

2. Division of Biosciences, Department of Life Sciences, Centre of Inflammation Research and Translational Medicine, College of Health, Medicine and Life Sciences, Brunel University London, Uxbridge, UK.

†: These authors contributed equally to this work.

*: Authors to whom correspondence should be addressed: R. de Dios (ruben.dediosbarranco@brunel.ac.uk) and F. Reyes-Ramírez (freyram@upo.es).

Abstract

The sphingomonads encompass a diverse group of bacteria within the Sphingomonadaceae family, with the presence of sphingolipids on their cell surface instead of lipopolysaccharide as their main common feature. They are particularly interesting for bioremediation purposes due to their capability to degrade or metabolise a variety of recalcitrant organic pollutants. However, the research and development of their full bioremediation potential has been hampered because of the limited number of tools available to investigate and modify their genome. Here, we present a markerless genome editing method for *Sphingopyxis granuli*, which can be further optimised for other sphingomonads. This procedure is based on a double recombination triggered by a DNA double strand break in the chromosome. The strength of this protocol lies in forcing the second recombination rather than favouring it by pressing a counterselection marker, thus avoiding laborious re-streaking or passaging screenings. Additionally, we introduce a modification with respect to the original protocol to increase the efficiency of the screening after the first recombination event. We show this procedure step by step and compare our modified method with respect to the original one by deleting *ecfG2*, the master regulator of the general stress response in *S. granuli*. This adds onto the genetic tool repertoire that can be applied to sphingomonads and stands as an efficient option for fast genome editing of this bacterial group.

Keywords: sphingomonads, Sphingomonadaceae, genome editing, genetic tools, mutation, bioremediation

Introduction

Sphingomonads is a bacterial group encompassing the genera *Sphingobium, Sphingopyxis, Novosphingobium, and Sphingomonas*, which classified within the Sphingomonadaceae family [1, 2]. The members of the Sphingomonadaceae family are Gram-negative alphaproteobacteria of various sizes that do not form spores. They can be motile or non-motile, and when they are, they often have a polar flagellum. The colonies they form exhibit yellow or orange tones due to the presence of carotenoids and contain sphingolipids (glycosphingolipids) in their cell envelopes instead of lipopolysaccharides (reviewed in [3]). Sphinhomonads are widely distributed in nature, as they inhabit multiple terrestrial and aquatic environments, and have been isolated from plant roots, clinical samples, and other sources [4]. Some of them have been also found to be endophytic [5, 6] and assist in phytoremediation processes (reviewed in [7]) and it has been described a facultative anaerobe, *Sphingopyxis granuli* strain TFA, which is able to growth in anaerobic conditions in the presence of nitrate [8].

Among the Sphingomonadaceae family, sphingomonads are well known for their capability to degrade recalcitrant compounds, including aromatic hydrocarbons. Examples of sphingomonads with this ability are *Sphingopyxis granuli*, which is capable of used tetralin as the sole carbon and energy source [9]; *Sphingomonas aromaticivorans* F199, which degrades biphenyl, naphthalene, m-xylene and p-cresol [10]; *Sphingomonas wittichii* RW1, which catabolises dibenzo-p-dioxin [11] or *Sphingobium chlorophenolicum* L-1, which catabolises pentachlorophenol [12]. Some members of this group are even able to metabolise pharmaceutical agents, such as *Sphingomonas sp.* MPO218, which utilises ibuprofen as carbon source [13]. Additionally, *Sphingopyxis macrogoltabida* NBRC 15033 and *Sphingopyxis sp*. PVA3 are able to degrade synthetic polymers such as polyethylene glycol (PEG) [14] and polyvinyl alcohol (PVA) [15], respectively, employed in the production of plastic items, adhesives or packaging films [16]. Due to the ability to degrade xenobiotic compounds, this group of bacteria has gained a special interest in the bioremediation field. However, as non-model microorganisms, they have been lagging behind in the development of efficient genetic tools and genome editing technologies. The further development of these tools is essential for fully understanding the biodegradation pathways and physiology of this group of bacteria and to maximise their bioremediation potential.

In this regard, basic genetic tools have been developed for mutational analyses and targeted mutagenesis in Sphongomonadaceae, including gene disruption and replacement strategies. However, these methods may cause polar effects on the genes located downstream in an operon. This hindrance led to the development of markerless gene deletion strategies based on a double recombination involving successive selection-counterselection rounds. In a first round, a plasmid harboring the upstream and downstream homologous regions of the target gene is transferred to the strain of interest and the first recombination into the chromosome is selected typically using an antibiotic resistance marker. In addition to this selection marker, plasmids used for markerless gene deletion procedures include a counterselection marker that may confer sensitivity to specific compounds, depending on the marker. By growing the cointegrate strain in the presence of this compound, only cells undergoing a second recombination and losing the plasmid backbone will be able to grow. A classic example of a counterselection marker is *sacB*, which confers sensitivity to sucrose and has been applied to different members of the Sphingomonadaceae [17-23]. The *sacB* marker confers sensitivity to sucrose, thus growing the cointegrate clones in the presence of this disaccharide after a first recombination would favour a second recombination. This second recombination event would lead either to the reconstitution of the wild-type genomic configuration or to the stable introduction of the aimed genome modification. However, the highly frequent emergence of spontaneous sucrose-resistant mutants and the need of multiple rounds of sub-culturing in the presence of sucrose make this system tedious and poorly reproducible [24]. Kaczmarczyk *et al.* (2012) [24] developed a similar strategy with a better reproducibility taking advantage of the natural streptomycin resistance of sphingomonads, which has been extensively used since its publication [25-40]. In this case, they engineered an artificial allele of *rpsL*, termed *rpsL1*, that confers sensitivity to streptomycin and can be used as a counterselection marker. Despite the unquestionable technical improvement tested for a range of sphingomonads, the streptomycin sensitivity produced by *rpsL1* varied across species, which may again need further passaging or re-streaking to achieve the second recombination.

To accelerate the double-recombination procedure, Martínez-García and de Lorenzo (2011) [41] developed a method for genome editing in *Pseudomonas putida*, subsequently modified by Wirth *et al*. (2020) [42], in which the second recombination is triggered by a DNA double-strand break. This addition would avoid the need of a counterselection marker, by-passing the re-streaking/passaging steps and hence shortening the protocol. In this method, the upstream and downstream homologous regions of the target gene are cloned in a non-replicative plasmid flanked by two SceI restriction sites termed pEMG. After the introduction of this vector, first recombination event, a second vector carrying the SceI coding gene under

an inducible promoter, termed pSW-I, is transferred into the cointegrate strain. The expression of *sceI* produces a double-strand break in the cognate restriction sites that is eventually repaired by homologous recombination, resolving the cointegrate and either producing the deletion of the target gene or a reversal to the wild-type genotype. The strength of this method lies in the need to forcefully repair the double-strand break by recombination, regardless of a counterselection, for which the strain of interest may develop secondary adaptations or resistance mutations. Thanks to its high efficiency and versatility, this genome editing strategy has been extensively applied for single gene deletions, as well as for the removal whole gene clusters and insertion of epitopes in various species [43-48]. Furthermore, it has even been adapted to other bacteria that have remained reluctant to gene manipulation, including multidrug-resistant *Acinetobacter baumannii*, in which it has been used to edit the chromosome as well as native plasmids [49].

In this work, we describe the optimisation of the SceI-based genome editing method to *Sphingopyxis granuli* TFA, supporting its further applicability to the Sphingomonadaceae family. Furthermore, we implement this strategy with additions from Kaczmarczyk *et al*. (2012) to improve the detection of single-recombinant clones. The improvement of this method with respect to traditionally-used counterselection-based procedures lies in forcing a second recombination rather than just favouring it. We do this by deleting *ecfG2*, a well-known regulator of the alphaproteobacterial general stress response (GSR) and showing the effect of the deletion on the GSR activation with respect to the wild type TFA. We also provided the guidelines of this procedure in a step-by-step comprehensive protocol.

Protocol optimization: methods and results

Plasmid construction

The original plasmids to perform the SceI-based genome editing strategy, pEMG (kanamycin resistance (Km^R)) and pSW-I (ampicillin resistance (Ap^R)) were kindly provided by Prof. V. de Lorenzo (CNB, Madrid). Plasmids pMPO1409 (Km^R, carrying upstream and downstream homologous regions to delete *ecfG2* in *S. granuli* TFA) pMPO1408 (Ap^R, carrying a *ecfG2::lacZ* gene fusion) and pMPO1412 ($Km^R Str^S$ (streptomycin sensitivity)) were previously constructed as described in de Dios *et al.* (2020) [50] and González-Flores *et al*. (2019) [51], respectively. For the purpose of comparing the original protocol (using pEMG-derivative plasmids) and the subsequent improvements using pMPO1412-derivative plasmids (introducing the *rpsL1* counterselection marker in pEMG), we constructed pMPO1162, a pMPO1412-derivative carrying the above mentioned *ecfG2* homologous regions. This construction was performed by digesting pMPO1409 with SacI and XbaI (New

England Biolabs), purifying a 2 kb fragment containing the *ecfG2* homologous regions using the GFX (GE Healthcare Life Sciences) DNA purification kit, and ligating it into pMPO1412 digested with the same enzymes using T4 DNA ligase (New England Biolabs). Enzymatic reactions and purification procedures were performed as per the manufacturer´s instructions. Chemically competent *Escherichia coli* DH5α λ*pir* were transformed with ligation mixtures via heat-shock transformation.

The use of *rpsL1* counterselection improves the screening of cointegrate candidates

Traditionally, kanamycin has been used as a selection marker in sphingomonads in general and in *S. granuli* in particular. However, when selecting single-recombination events, the recombination frequency is similar to that of spontaneous kanamycin resistant mutants (Figure 1). Due to this, the distinction of cointegrates carrying pEMG derivatives with the upstream and downstream regions of the target gene required tedious screening by PCR. In our efforts to make this step more efficient, we cloned the *rpsL1* counterselection marker in pEMG, obtaining pMPO1412 in previous work [51]. In order to compare both approaches, we attempted to delete *ecfG2* using backbone plasmids with and without carrying the *rpsL1* allele in parallel. To do this, *S. granuli* TFA was electrotransformed with 200 ng of pMPO1409 or pMPO1162, or an equivalent amount of bi-distilled water as a control. All transformations were performed in biological triplicate and serial dilutions were plated on MML agar supplemented with 20 mg/L kanamycin (plain MML agar for viable cell counting). As a result, we confirmed that the spontaneous emergence of kanamycin resistant mutants in the control transformations was not significantly different from that obtained in the transformations with pMPO1409 and pMPO1162.

To identify the cointegrate clones, we performed a screening with all the resulting kanamycin resistant colonies by streaking them on MML agar plates supplemented with either 20 mg/L kanamycin alone or with 20 mg/L kanamycin and 200 mg/L streptomycin (four-fold the concentration of streptomycin we routinely use to select the wild-type TFA strain), which would negatively impact the growth of the clones carrying the *rpsL1* marker. Then, the agar plates were incubated at 30 ºC for just 16 h (all the streaks would look equally grown if incubated beyond 18-20 hours). As expected, colonies taken from control plates and those transformed with pMPO1409 grew at similar rates in media supplemented with kanamycin only or with kanamycin and streptomycin (Figure 2A,B). However, plates streaked with colonies obtained by transforming TFA with pMPO1162, harbouring *rpsL1*, showed multiple clones that grew visibly slower in the presence of kanamycin and streptomycin compared to those grown in the presence of kanamycin only (Figure 2C).

To test the efficacy of this counterselection as an improvement to the screening process, we performed an additional PCR screening using primers that annealed within the kanamycin resistance marker harboured in pMPO1409 and pMPO1162, yielding an amplicon of approximately 700 bp (KmFw: GATTGAACAAGATGGATTGC; KmRev: CGTCAAGAAGGCGATAGAAGG). To do this screening, we randomly selected 10 clones from each of the 3 pools obtained by transforming with pMPO1409, as well as 5 targeted clones transformed with pMPO1162 that grew slower in the presence of kanamycin and streptomycin. As shown in Figure 3, only 7 out of 30 clones transformed with pMPO1409 yielded a PCR product, indicative of having undergone the first recombination event. However, all 5 clones tested from the transformation with pMPO1162 yielded a PCR product. This conclusively shows how using the *rpsL1* counterselection as an indication of the first recombination event reduces the number of clones reduces the number of clones to test and yields a more targeted and efficient screening.

The transformation with pSW-I triggers the second recombination

At this point, the plasmid carrying the upstream and downstream homologous regions of the target gene (*ecfG2* in this case) flanked by the SceI restriction sites would be inserted in the TFA chromosome by a single recombination event. The next step consists in forcing the second recombination that would lead to the *ecfG2* deletion. To do this, a replicative plasmid carrying the *sceI* gene needs to be transferred into the selected cointegrate clone. To perform this step, we selected clones 3, 4 and 5 as labelled in Figure 2C, prepared electrocompetent cells of each of them and transformed them with 200 ng of pSW-I in parallel to the respective three controls with an equivalent amount of water. Selection was carried out on MML agar supplemented with ampicillin 5 mg/L. In this case, the transformation frequency with pSW-I was significantly higher than the emergence of spontaneous ampicillin resistant clones (Figure 4).

In the original *P. putida* protocol, both the pEMG derivative construct and the pSW-I plasmid are simultaneously selected prior to inducing the *sceI* expression. However, during our first attempts to use this procedure, we observed a very poor growth when selecting both genetic elements at the same time. A possible explanation would be that the presence of both the cointegrate and pSW-I simultaneously has a strong fitness cost due to the leaky expression of *sceI*. This would continuously produce double-strand breaks in the SceI target site introduced in the chromosome with the first recombination step. For this reason, rather than selecting the presence of both elements at the same time as in the original protocol [41], we selected only the presence of pSW-I. In addition to this, the original protocol requires the induction of the *sceI* expression by adding 3-methylbenzoate. This was further optimised by

Wirth *et al*. (2020) [42] by doing the pSW-I selection and *sceI* induction in a single step. However, during the optimisation of this protocol for the use on *S. granuli* TFA, we noticed that the leaky expression of *sceI* alone was enough to trigger the DNA double-strand break. This has also been observed for other bacteria, for which the addition of 3-methylbenzotae is even deleterious for the growth [49].

For these reasons, we directly screened the ampicillin resistant, kanamycin sensitive clones obtained after the transformation with pSW-I (50 clones per transformation) by streaking them on MML agar plates supplemented with either 20 mg/L kanamycin, 5 mg/L ampicillin or 50 mg/L streptomycin (Supplementary Figure S1). After this, we obtained that, for each of the three independent pSW-I transformations, 48%, 62% and 30% of the clones were kanamycin sensitive, indicating that they had undergone a second recombination event during the selection process.

To assess if the second recombination had led to the deletion of *ecfG2* or if the clone had reverted to the wild type genotype, we screened 10 clones of each transformation by PCR. We used primers Seq_ecfG2_Fw2 (ACCGATTTTGCCCATGGCTTC) and Seq_ecfG2_Rv (CGAACGGAAACAGAGGTGATC), which would yield a product of approximately 1 kb in the case of the wild type configuration or approximately 0.5 kb in the case of the *ecfG2* deletion. As a result, 21 out of 30 total clones had suffered the deletion of *ecfG2* (Figure 5).

As a final step, pSW-I have to be cured from the deletion mutant. To do this, we inoculated one positive clone from each pSW-I transformation in MML broth in the absence of ampicillin. After two passages, cells were serially diluted and plated on MML agar supplemented with 50 mg/L streptomycin. To conclusively assess the loss of both pMPO1162 and pSW-I, 50 colonies obtained from plating each individual clone were streaked on MML agar supplemented with either 20 mg/L kanamycin, 5 mg/L ampicillin or 50 mg/L streptomycin. As a result, none of the clones grew in the presence of kanamycin or ampicillin, but all grew in the presence of streptomycin (Supplementary Figure S2), which indicated the curation of all genetic devices used in this deletion strategy. A scheme summarising the recombinations and genetic rearrangements undergone during this procedure is shown in Figure 6.

The deletion of *ecfG2* abolishes the expression of *nepR2*

EcfG2 is an extracytoplasmic function sigma factor that acts as the main regulator of the GSR in *S. granuli* TFA [49, 50]. Although *ecfG2* is essential under stressing conditions, a mutant in this gene shows a similar fitness compared to the wild type TFA in the absence of stress. Among its target genes, *nepR2* encodes an anti-sigma factor that exerts a negative feedback loop on the activation of the GSR [49]. To show the phenotype of the new

constructed Δ*ecfG2* mutant, we transformed the Δ*ecfG2* mutant along with the wild type TFA with pMPO1408, an integrative vector carrying a *nepR2::lacZ* gene fusion. Thus, EcfG2-mediated activation of the GSR, and thus *nepR2*, would lead to the production of the β-galactosidase enzyme that breaks down X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) into a blue precipitate. To visualize this, we streaked both the wild type TFA and the Δ*ecfG2* deletion mutant, both carrying the *nepR2::lacZ* fusion, on a MML agar plate supplemented with 25 mg/L X-gal and incubated it at 30 ºC for 5 days. As a result, we observed that the wild type TFA yielded a blue colour, whereas the Δ*ecfG2* deletion mutant would not produce this precipitate (Figure 7). This is coherent with previous studies, showing that a mutant in *ecfG2* has a null ability to activate the GSR, including the expression of *nepR2* [50].

Discussion

In this work, we describe an efficient genome editing procedure with potential applicability in sphingomonads, using *S. granuli* as a model. Traditionally, mutational studies involving this group of bacteria have been developed via marked mutation or counterselection-mediated double recombination. The *sacB* gene has been the most frequently used counterselection marker, although alternative counterselection markers have been described, including *pheS*, which provides sensitivity to p-chlorophenylalanine [52]. However, there are representatives of this group that are naturally sensitive to this compound [24], making it not generally applicable to these procedures. Even in the case of *sacB*, previous reports have shown that its presence in the absence of selection can generate a certain toxicity that leads to accumulating mutations that inactivate it [53], thus rendering it ineffective for counterselection. This has also been observed when applying this strategy in members of the Sphingomonadaceae [24]. The optimised protocol we present here takes advantage of the natural streptomycin resistance of sphingomonads to indicate a successful first recombination, combined with a DNA double-strand break induction that efficiently triggers a second recombination. This strategy offers a fast-track procedure compared to the previously mentioned double recombination-based genome editing protocols.

Although we present an optimisation of the procedure to *S. granuli*, the protocol is further amenable to different modifications that may tailor it to the specificities of other sphingomonads. For example, we use electrotransformation as a means to introduce plasmids in *S. granuli*. However, both pMPO1412 (and the parental pEMG) and pSW-I are mobilisable via biparental or triparental mating, which may be an alternative for sphingomonads in which electrotransformation is not applicable. Furthermore, the selection

markers in the pMPO1412 (and pEMG) and pSW-I can be exchanged according to the resistance profile of other sphingomonads. In this regard, several versions of pEMG and pSW-I derivatives with different selection markers are available through the Standard European Vector Architecture platform [54]. Another aspect that may be attuned to the requirements of other sphingomonads is the Sce-I expression induction step. In the case of *S. granuli*, the leaky expression of Sce-I, which is inducible by 3-methylbenzoate under the XylS-*Pm* system [41], was enough to trigger the second recombination. However, this expression system may behave differently in other species, and the addition of the inducer in the selection plates after introducing pSW-I may be required [42].

All in all, we describe a powerful tool for genome editing in *S. granuli* that can be further tailored to the requirements of other sphingomonad models.

Summary step-by-step protocol

Here, we describe a step-wise protocol to perform this genome editing strategy on *S. granuli*, which can be used as a base for optimisation to other sphingomonad species and laboratory methods (e.g. culture media, incubation times, selection/screening procedures). The protocol is described once the pMPO1412-derivative vector has been constructed and purified.

<u>First recombination event</u>

Day 1: inoculation

- Inoculate 3 ml of MML broth with wild type *S. granuli* TFA. Incubate at 30 ºC, 180 rpm to saturation (typically 24 h). MML broth recipe is provided as supplementary material.

Day 2: electrotransformation with the pMPO1412-derivative plasmid

- Prepare *S. granuli* electrocompetent cells (our in-house protocol has been provided as supplementary material). Alternatively, the pMPO1412 derivative can be introduced in the target strain by triparental mating.
- Electrotransform with 200 ng of the purified pMPO1412-derivative plasmid using an electroporator.
- Reconstitute the electroporated cell mixture with 1 ml of ice-chilled MML broth supplemented with 0.5 M sorbitol or 10% glycerol. Incubate at 30 ºC, 180 rpm for 1.5 h.

- Perform serial dilution and plating on MML agar plates supplemented with kanamycin 20 mg/L for selection. Incubate the plates at 30 ºC for 4-5 days.

Day 3: first recombination screening

- Perform dual streaking of 50-100 clones on two MML agar plates: one supplemented with kanamycin 20 mg/L and other one supplemented with kanamycin 20 mg/L and streptomycin 200 mg/L. Incubate the plates at 30 ºC for approximately 16 h.

Day 4: first recombination screening results

- By direct visualization, select the clones that grew well on MML agar with kanamycin only but grew slower on MML agar with kanamycin and streptomycin.
- To validate these recombinant clones, perform a colony PCR using specific primers to the plasmid introduced (we typically use primers annealing in the kanamycin resistance marker).
- Inoculate one of the validated cointegrate clones in 3 ml of MML broth to continue on to the second recombination event.

Day 5: electrotransformation with pSW-I

- Prepare electrocompetent cells of the cointegrate clone.
- Electrotransform the cointegrate clone with 200 ng of purified pSW-I.
- Reconstitute the electroporated cell mixture with 1 ml of ice-chilled MML broth supplemented with 0.5 M sorbitol or 10% grycerol. Incubate at 30 ºC, 180 rpm for 1.5 h.
- Perform serial dilution and plating on MML agar plates supplemented with ampicillin 5 mg/L for selection. Incubate the plates at 30 ºC for 4-5 days.

Day 6: second recombination screening

- Perform multiple streaking of 50 clones on three MML agar plates: one supplemented with kanamycin 20 mg/L, other one supplemented with ampicillin 5 mg/L and one last plate supplemented with streptomycin 50 mg/L. Incubate the plates at 30 ºC for 24 h.

Day 7: second recombination screening results and mutation screening

- By direct visualization, select the clones that grew on MML agar with ampicillin only but did not grow on MML agar with kanamycin and ampicillin.

- To validate these recombinant clones and identify those that underwent the genetic modification, perform a colony PCR using specific primers flanking the gene targeted for deletion in order to distinguish PCR products of different sizes between the wild type and the mutant strain. Other modifications may require alternative validation approaches (e.g. other primer combinations for DNA insertions, Sanger sequencing for point mutation)

- Inoculate one of the validated mutant clones from those streaked on plain MML agar in 3 ml of MML broth start the curation on pSW-I Incubate at 30 ºC, 180 rpm for 24 h.

Day 8: pSW-I curation

- Perform a passage of the mutant strain by doing a 1/500 dilution in fresh MML broth. Incubate at 30 ºC, 180 rpm for 24 h.

Day 9: pSW-I curation

- Repeat passaging as in Day 8.

Day 10: pSW-I curation

- Streak on MML agar to obtain isolated colonies. Incubate at 30 ºC for 24 h.

Day 11: validation of pSW-I-cured mutant clones

- Perform dual streaking of 50 clones on two MML agar plates: one supplemented with ampicillin 5 mg/L and a plain MML agar. Incubate the plates at 30 ºC for 24 h.

Day 12: selection of pSW-I-cured clones

- By direct visualization, select the clones that grew on plain MML agar but did not grow on MML agar supplemented with ampicillin and neither with kanamycin. Re-streak on plain MML agar to obtain isolated colonies. Use these isolated colonies for cryoconservation of the newly generated mutant strain.

Data summary

All data and protocols used or generated through this work have been provided within this article or in the associates supplementary files.

Acknowledgements

We would like to thank Prof Eduardo Santero, Prof Fernando Govantes and Dr Elena Rivas-Marín, as well as the rest of the member of the Division of Microbiology of the University Pablo de Olavide (Seville, Spain), for their advice and suggestions during the optimisation of this method, and Dr Guadalupe Martín-Cabello for the technical support.

Funding

This work has been supported by the grant PID2021-125491NB-I00 funded by MICIU/AEI/10.13039/501100011033 (Ministerio de Ciencia, Innovación y Universidades / Agencia Estatal de Investigación) and FEDER (UE) to FR-R. IGR is supported by a postdoctoral contract (PAIDI 2020, POSTDOC_21_00064) funded by the Andalusian Government (Junta de Andalucía). RD is supported by a Biotechnology and Biological Sciences Research Council New Investigator Award (BB/V007823/1) awarded to Dr Ronan McCarthy.

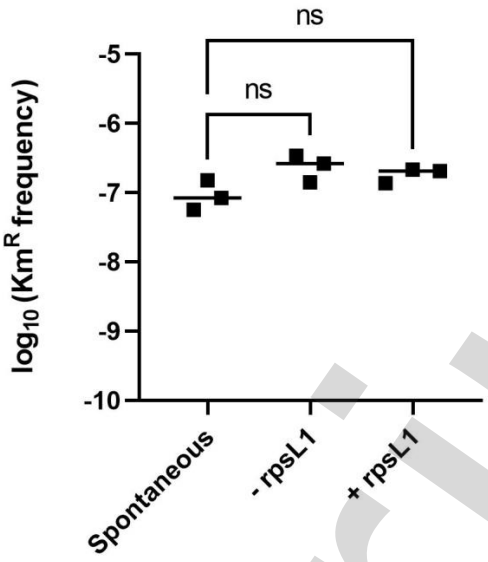
Author contribution

IGR designed and performed experimental work and drafted and edited the manuscript. RD originally conceived this work, designed and performed experimental work and drafted and edited the manuscript. FRR designed and supervised the work and edited the manuscript.

Conflict of interest

RD is an Editor for Access Microbiology.

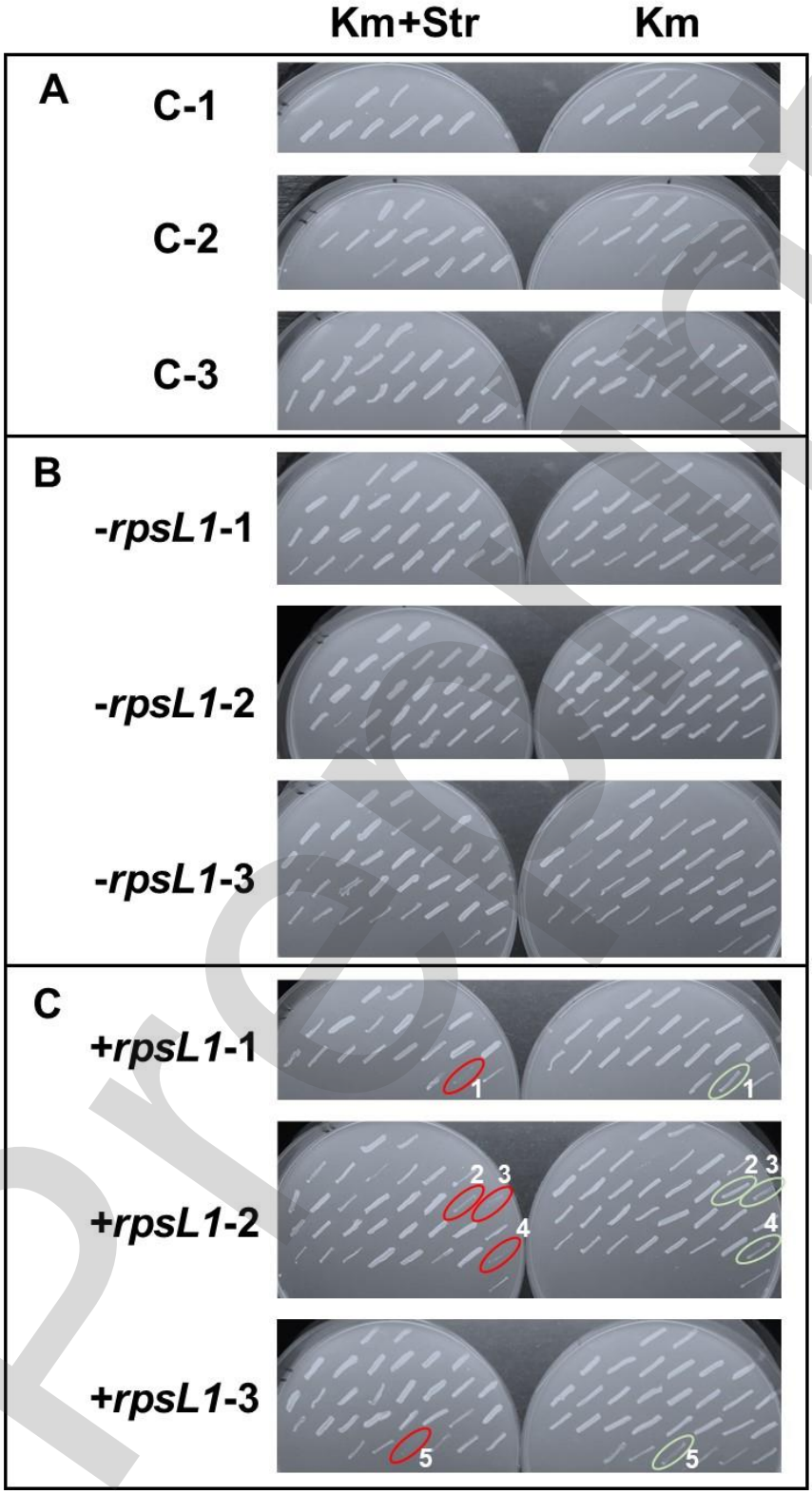
Figures and figure legends



387

388 **Figure 1**. Logarithmic representation of *S. granuli* transformation frequency when
389 transforming with pMPO1409 (*-rpsL1*) and pMPO1162 (*+rpsL1*) compared to a control
390 (spontaneous). 200 ng of the respective plasmid (an equivalent volume of water in the case
391 of the control) were used in each case. Each transformation was repeated with three
392 independently prepared aliquot of electrocompetent *S. granuli* TFA cells. Individualised
393 colony counts for each repeat and frequency calculations are shown in Supplementary Table
394 S1. One-way ANOVA analysis was performed between the different samples; n.s.: non-
395 significant.

396



397 **Figure 2.** Streaks of clones obtained from the control electroporation (A), or
398 electrotransformations with pMPO1409 (-*rpsL1*) (B) or pMPO1162 (+*rpsL1*) (C). The red-

labelled streaks do not grow or grow slowly with Km 20 mg/L plus Str 200 mg/L when compared to the same streaks growing with only Km 20 mg/L (labelled in green). Results for the previously mentioned three independent electrotransformation replicates are shown, as well as controls in which the electroporation was performed with an equivalent volume of water.

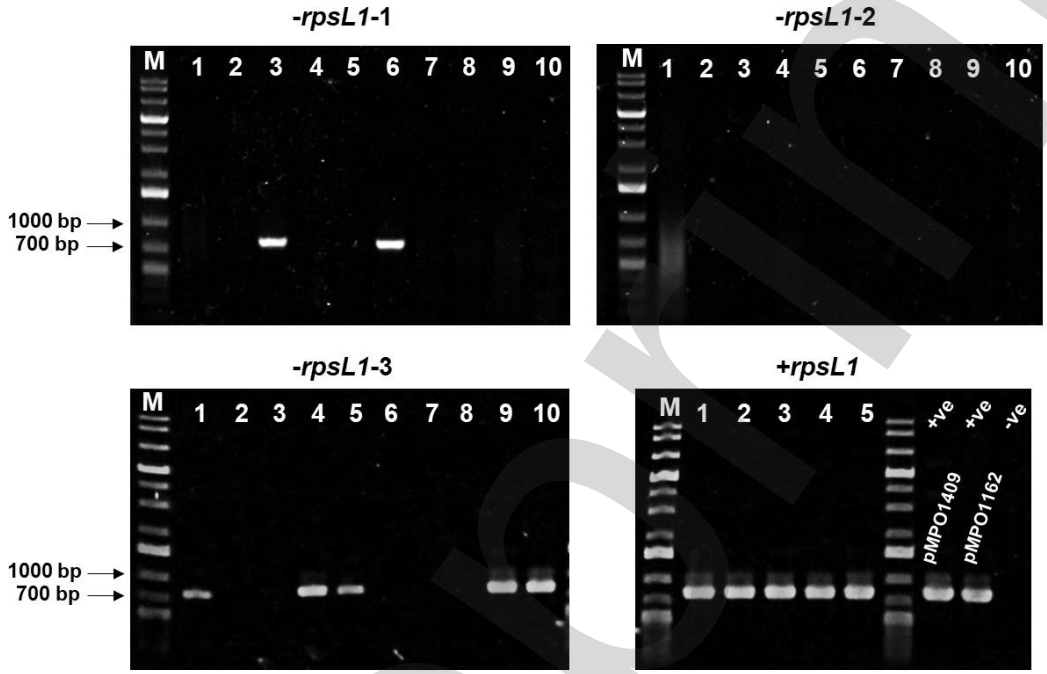


Figure 3. Agarose gel electrophoresis (0.8%) of PCR products from different clones after electrotransforming with pMPO1409 (-*rpsL1*) or pMPO1162 (+*rpsL1*), using the primers KmFw (GATTGAACAAGATGGATTGC) and KmRev (CGTCAAGAAGGCGATAGAAGG). The plasmids pMPO1409 and pMPO1162 were used as positive (+ve) controls, respectively, and no DNA was added in negative (-ve) control.

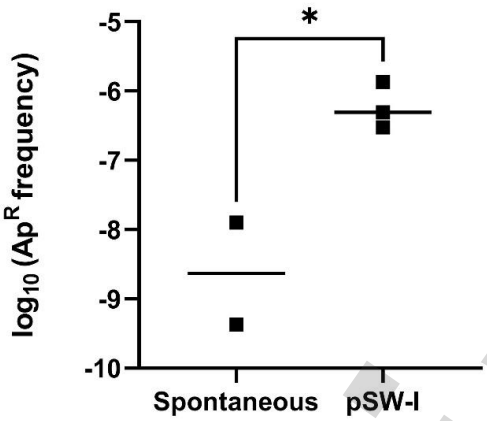


Figure 4. Logarithmic representation of transformation frequency when introducing pSW-I into each independent cointegrate clone compared to the respective negative controls. 200 ng of plasmid were used in each case, and an equivalent amount of water was used for the negative controls. Individualised colony counts for each repeat and frequency calculations are shown in Supplementary Table S2. T-Student analysis was performed between the different samples. * = $p<0.05$.

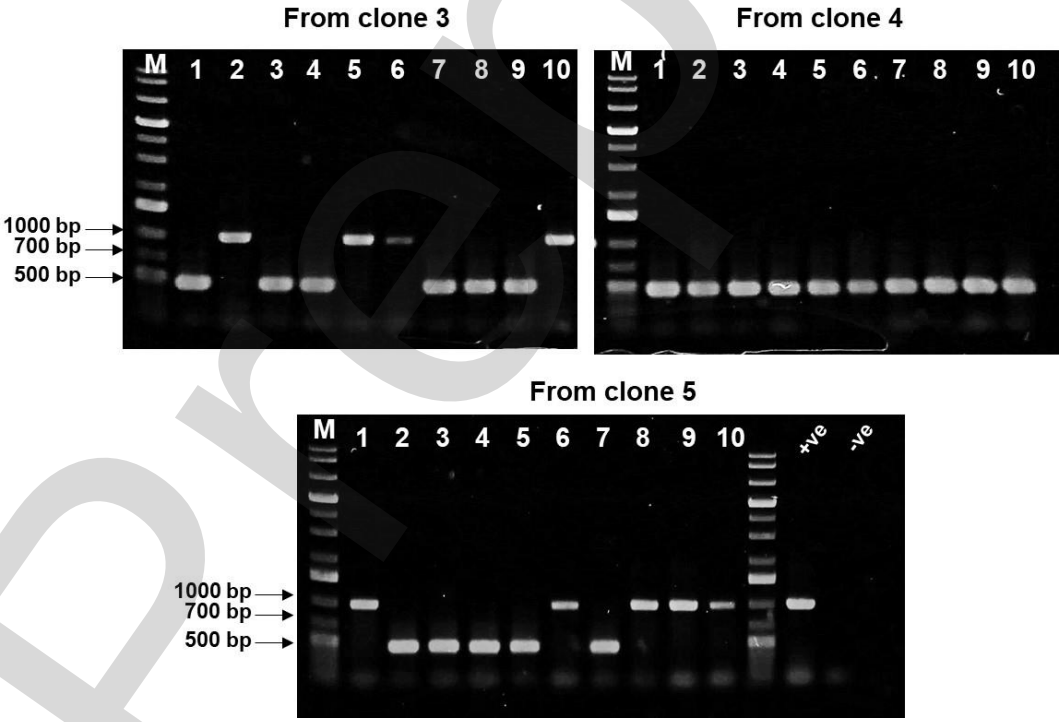
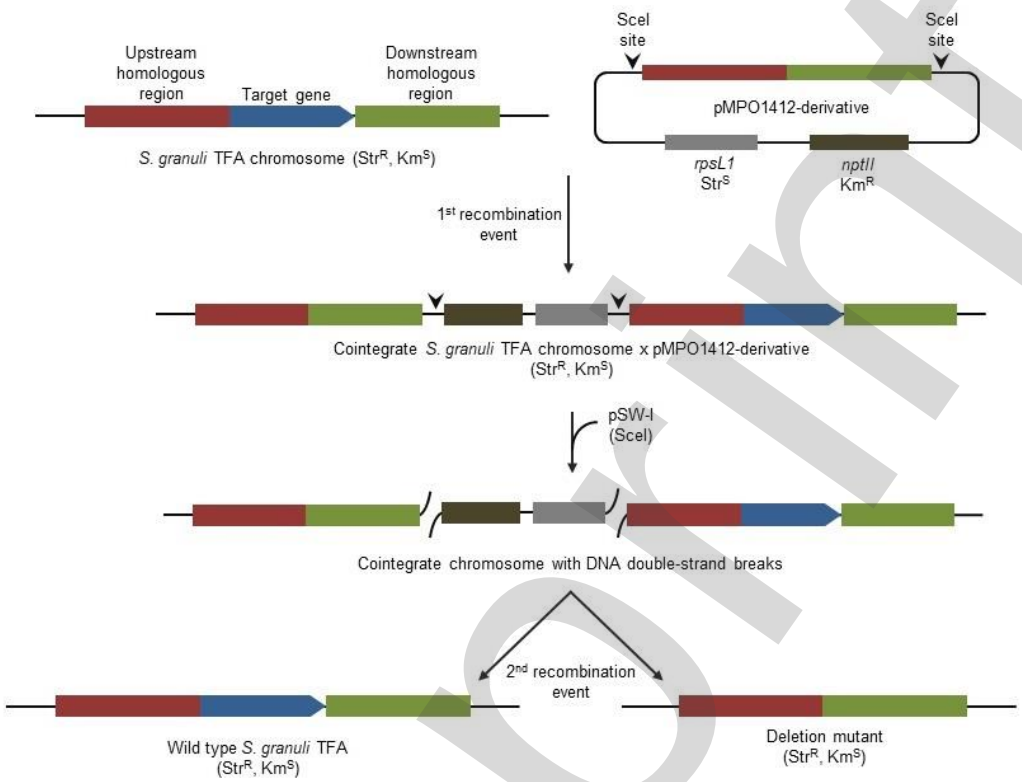


Figure 5. Agarose gel electrophoresis (0.8%) of PCR product from different clones of second recombination event, using the oligos Seq_ecfG2_Fw2 (ACCGATTTTGCCCATGGCTTC) and Seq_ecfG2_Rv (CGAACGGAAACAGAGGTGATC). The plasmid used in the electroporation were pSW-I. Genomic DNA from wild type TFA

strain was used as positive (+ve) control and no DNA was added in negative (-ve) control.



Wild type fragment is ~1000 bp and the deletion of *ecfG2* is shown as a ~500 bp fragment.

Figure 6. Schematic representation of the genetic rearrangements and recombinations occurring during this genome editing procedure for a generic target gene (*ecfG2* in the example developed in this work), as well as the two possible outcomes: a reconstitution of the wild-type genotype or the aimed genome modification.

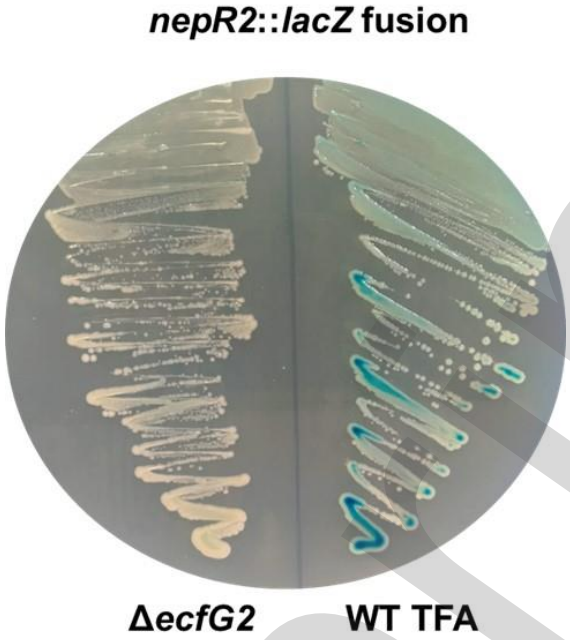


Figure 7. Expression of *nepR2::lacZ* fusion in the mutant Δ*ecfG2* compared to the wild type (WT) TFA strain. EcfG2 is essential for activating the expression of *nepR2* (de Dios *et al*., 2020). Thus, the *nepR2::lacZ* fusion (born in pMPO1408) in the Δ*ecfG2* background does not yield a blue colour in the presence of X-gal. Both strains were streaked on MML agar plates supplemented with X-gal 25 mg/L and were incubated at 30 ºC for 5 days.

References

1. Takeuchi, M., K. Hamana, and A. Hiraishi, *Proposal of the genus Sphingomonas sensu stricto and three new genera, Sphingobium, Novosphingobium and Sphingopyxis, on the basis of phylogenetic and chemotaxonomic analyses.* Int J Syst Evol Microbiol, 2001. **51**(Pt 4): p. 1405-17.
2. Kosako, Y., et al., *Proposal of Sphingomonadaceae fam. nov., consisting of Sphingomonas Yabuuchi et al. 1990, Erythrobacter Shiba and Shimidu 1982, Erythromicrobium Yurkov et al. 1994, Porphyrobacter Fuerst et al. 1993, Zymomonas Kluyver and van Niel 1936, and Sandaracinobacter Yurkov et al. 1997, with the type genus Sphingomonas Yabuuchi et al. 1990.* Microbiol Immunol, 2000. **44**(7): p. 563-75.
3. Glaeser, S.P. and P. Kämpfer, *The Family Sphingomonadaceae.*, in *The Prokaryotes*, E. Rosenberg, et al., Editors. 2014, Springer: Berlin, Heidelberg. p. 641-707.
4. Balkwill, D.L., J.K. Fredrickson, and M.F. Romine, *Sphingomonas and Related Genera*, in *The Prokaryotes*, M. Dworkin, et al., Editors. 2006, Springer: New York. p. 605-629.

5. Battu, L., et al., *Assembly of genomic reads of elite indica rice cultivar onto 2101 reference bacterial genomes for identification of co-sequenced endophytic bacteria.* Data Brief, 2017. **12**: p. 305-312.
6. Battu, L., et al., *Genome inside genome: NGS based identification and assembly of endophytic Sphingopyxis granuli and Pseudomonas aeruginosa genomes from rice genomic reads.* Genomics, 2017. **109**(3-4): p. 141-146.
7. Gatheru Waigi, M., K. Sun, and Y. Gao, *Sphingomonads in Microbe-Assisted Phytoremediation: Tackling Soil Pollution.* Trends Biotechnol, 2017. **35**(9): p. 883-899.
8. García-Romero, I., et al., *Genomic analysis of the nitrate-respiring Sphingopyxis granuli (formerly Sphingomonas macrogoltabida) strain TFA.* BMC Genomics, 2016. **17**: p. 93.
9. Floriano, B., E. Santero, and F. Reyes-Ramírez, *Biodegradation of Tetralin: Genomics, Gene Function and Regulation.* Genes (Basel), 2019. **10**(5).
10. Romine, M.F., et al., *Complete sequence of a 184-kilobase catabolic plasmid from Sphingomonas aromaticivorans F199.* J Bacteriol, 1999. **181**(5): p. 1585-602.
11. Wittich, R.M., et al., *Metabolism of dibenzo-p-dioxin by Sphingomonas sp. strain RW1.* Appl Environ Microbiol, 1992. **58**(3): p. 1005-10.
12. Copley, S.D., et al., *The whole genome sequence of Sphingobium chlorophenolicum L-1: insights into the evolution of the pentachlorophenol degradation pathway.* Genome Biol Evol, 2012. **4**(2): p. 184-98.
13. Aulestia, M., et al., *Isolation and genomic characterization of the ibuprofen-degrading bacterium Sphingomonas strain MPO218.* Environ Microbiol, 2021. **23**(1): p. 267-280.
14. Ohtsubo, Y., et al., *Complete Genome Sequence of Sphingopyxis macrogoltabida Type Strain NBRC 15033, Originally Isolated as a Polyethylene Glycol Degrader.* Genome Announc, 2015. **3**(6).
15. Yamatsu, A., et al., *Isolation and characterization of a novel poly(vinyl alcohol)-degrading bacterium, Sphingopyxis sp. PVA3.* Appl Microbiol Biotechnol, 2006. **72**(4): p. 804-11.
16. Kawai, F., *Sphingomonads involved in the biodegradation of xenobiotic polymers.* J Ind Microbiol Biotechnol, 1999. **23**(4-5): p. 400-407.
17. Masai, E., et al., *Genetic and biochemical characterization of a 2-pyrone-4, 6-dicarboxylic acid hydrolase involved in the protocatechuate 4, 5-cleavage pathway of Sphingomonas paucimobilis SYK-6.* J Bacteriol, 1999. **181**(1): p. 55-62.
18. García-Romero, I., et al., *SuhB, a small non-coding RNA involved in catabolite repression of tetralin degradation genes in Sphingopyxis granuli strain TFA.* Environ Microbiol, 2018. **20**(10): p. 3671-3683.
19. Zhu, L., et al., *Cloning and characterization of genes involved in nostoxanthin biosynthesis of Sphingomonas elodea ATCC 31461.* PLoS One, 2012. **7**(4): p. e35099.
20. Li, A., et al., *A Carotenoid- and Poly-β-Hydroxybutyrate-Free Mutant Strain of.* mSphere, 2019. **4**(5).
21. Kontur, W.S., et al., *Novosphingobium aromaticivorans uses a Nu-class glutathione S-transferase as a glutathione lyase in breaking the β-aryl ether bond of lignin.* J Biol Chem, 2018. **293**(14): p. 4955-4968.
22. Cai, C., et al., *Developing Rhodococcus opacus and Sphingobium sp. coculture systems for valorization of lignin-derived dimers.* Biotechnol Bioeng, 2022. **119**(11): p. 3162-3177.
23. Basta, T., et al., *Detection and characterization of conjugative degradative plasmids in xenobiotic-degrading Sphingomonas strains.* J Bacteriol, 2004. **186**(12): p. 3862-72.
24. Kaczmarczyk, A., J.A. Vorholt, and A. Francez-Charlot, *Markerless gene deletion system for sphingomonads.* Appl Environ Microbiol, 2012. **78**(10): p. 3774-7.
25. Song, D., et al., *The variations of native plasmids greatly affect the cell surface hydrophobicity of sphingomonads.* mSystems, 2023: p. e0086223.

26. Lu, H. and Y. Huang, *Transcriptome Analysis of Novosphingobium pentaromativorans US6-1 Reveals the Rsh Regulon and Potential Molecular Mechanisms of N -acyl-l-homoserine Lactone Accumulation.* Int J Mol Sci, 2018. **19**(9).
27. Cecil, J.H., et al., *Rapid, Parallel Identification of Catabolism Pathways of Lignin-Derived Aromatic Compounds in Novosphingobium aromaticivorans.* Appl Environ Microbiol, 2018. **84**(22).
28. Gottschlich, L., et al., *Complex general stress response regulation in Sphingomonas melonis Fr1 revealed by transcriptional analyses.* Sci Rep, 2019. **9**(1): p. 9404.
29. Fujita, M., et al., *Iron acquisition system of Sphingobium sp. strain SYK-6, a degrader of lignin-derived aromatic compounds.* Sci Rep, 2020. **10**(1): p. 12177.
30. Yu, Z., et al., *The LuxI/LuxR-Type Quorum Sensing System Regulates Degradation of Polycyclic Aromatic Hydrocarbons via Two Mechanisms.* Int J Mol Sci, 2020. **21**(15).
31. Higuchi, Y., et al., *Roles of two glutathione S-transferases in the final step of the β-aryl ether cleavage pathway in Sphingobium sp. strain SYK-6.* Sci Rep, 2020. **10**(1): p. 20614.
32. Liang, J., et al., *Benzo[a]pyrene might be transported by a TonB-dependent transporter in Novosphingobium pentaromativorans US6-1.* J Hazard Mater, 2021. **404**(Pt A): p. 124037.
33. Presley, G.N., et al., *Pathway discovery and engineering for cleavage of a β-1 lignin-derived biaryl compound.* Metab Eng, 2021. **65**: p. 1-10.
34. Huang, J., et al., *Coinducible Catabolism of 1-Naphthol via Synergistic Regulation of the Initial Hydroxylase Genes in Sphingobium sp. Strain B2.* Appl Environ Microbiol, 2021. **87**(11).
35. Gao, C., et al., *Bisphenol A biodegradation by Sphingonomas sp. YK5 is regulated by acyl-homoserine lactone signaling molecules.* Sci Total Environ, 2022. **802**: p. 149898.
36. Fujita, M., et al., *Functional roles of multiple Ton complex genes in a Sphingobium degrader of lignin-derived aromatic compounds.* Sci Rep, 2021. **11**(1): p. 22444.
37. Linz, A.M., et al., *Aromatic Dimer Dehydrogenases from Novosphingobium aromaticivorans Reduce Monoaromatic Diketones.* Appl Environ Microbiol, 2021. **87**(24): p. e0174221.
38. Higuchi, Y., et al., *The Catabolic System of Acetovanillone and Acetosyringone in Sphingobium sp. Strain SYK-6 Useful for Upgrading Aromatic Compounds Obtained through Chemical Lignin Depolymerization.* Appl Environ Microbiol, 2022. **88**(16): p. e0072422.
39. Wang, Z., et al., *The quorum sensing system of Novosphingobium sp. ERN07 regulates aggregate formation that promotes cyanobacterial growth.* Sci Total Environ, 2022. **851**(Pt 2): p. 158354.
40. Jiang, W., et al., *Comparative Genomic Analysis of Carbofuran-Degrading Sphingomonads Reveals the Carbofuran Catabolism Mechanism in Sphingobium sp. Strain CFD-1.* Appl Environ Microbiol, 2022. **88**(22): p. e0102422.
41. Martínez-García, E. and V. de Lorenzo, *Engineering multiple genomic deletions in Gram-negative bacteria: analysis of the multi-resistant antibiotic profile of Pseudomonas putida KT2440.* Environ Microbiol, 2011. **13**(10): p. 2702-16.
42. Wirth, N.T., E. Kozaeva, and P.I. Nikel, *Accelerated genome engineering of Pseudomonas putida by I-SceI-mediated recombination and CRISPR-Cas9 counterselection.* Microb Biotechnol, 2020. **13**(1): p. 233-249.
43. Leal-Morales, A., et al., *Transcriptional organization and regulation of the Pseudomonas putida flagellar system.* Environ Microbiol, 2022. **24**(1): p. 137-157.
44. de Dios, R., E. Santero, and F. Reyes-Ramírez, *The functional differences between paralogous regulators define the control of the general stress response in Sphingopyxis granuli TFA.* Environ Microbiol, 2022. **24**(4): p. 1918-1931.
45. Kohlstedt, M., et al., *Biobased PET from lignin using an engineered cis, cis-muconate-producing Pseudomonas putida strain with superior robustness, energy and redox properties.* Metab Eng, 2022. **72**: p. 337-352.
46. Owen, S.V., et al., *Prophages encode phage-defense systems with cognate self-immunity.* Cell Host Microbe, 2021. **29**(11): p. 1620-1633.e8.

47. Nies, S.C., et al., *High titer methyl ketone production with tailored Pseudomonas taiwanensis VLB120.* Metab Eng, 2020. **62**: p. 84-94.
48. Hobmeier, K., et al., *Anaplerotic Pathways in Halomonas elongata: The Role of the Sodium Gradient.* Front Microbiol, 2020. **11**: p. 561800.
49. de Dios, R., K. Gadar, and R.R. McCarthy, *A high-efficiency scar-free genome-editing toolkit for Acinetobacter baumannii.* J Antimicrob Chemother, 2022. **77**(12): p. 3390-3398.
50. de Dios, R., et al., *Two paralogous EcfG σ factors hierarchically orchestrate the activation of the General Stress Response in Sphingopyxis granuli TFA.* Sci Rep, 2020. **10**(1): p. 5177.
51. González-Flores, Y.E., et al., *The response of Sphingopyxis granuli strain TFA to the hostile anoxic condition.* Scientific Reports, 2019. **9**(1): p. 6297.
52. Kast, P., *pKSS--a second-generation general purpose cloning vector for efficient positive selection of recombinant clones.* Gene, 1994. **138**(1-2): p. 109-14.
53. Trebosc, V., et al., *A Novel Genome-Editing Platform for Drug-Resistant Acinetobacter baumannii Reveals an AdeR-Unrelated Tigecycline Resistance Mechanism.* Antimicrob Agents Chemother, 2016. **60**(12): p. 7263-7271.
54. Martínez-García, E., 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): p. D1558-D1567.

We would like to thank the reviewers for their constructive feedback, we believe it has helped to improve the quality of this manuscript. Please find below a point-by-point response to the comments.

Reviewer 1

The manuscript entitled "An improved genome editing system for Sphingomonadaceae" describes a detailed protocol for the construction of deletion mutants. The protocol has been optimized for Sphingopysis granuli and validated by constructing a deletion mutant in the ecfG2 gene.

In general, the manuscript is well organized, and demonstrates that the improved protocol is superior to previous and similar methodologies. More important, it can used in other bacteria recalcitrant to genetic manipulation, if they are naturally resistant to streptomycin.

Nevertheless, I have some suggestion that could improve the manuscript.

1. This is a methodological work, and the results is the improved protocol. Therefore, instead of calling "Methods" the section found in line 124, I suggest calling it "Results".

We appreciate this suggestion, and we agree that, as a Methods manuscript, the result is the method itself. However, we also think that the methodological explanation offered does not qualify as just results. Considering this suggestion, we have corrected “Methods” to “Protocol optimisation: methods and results” in an attempt to merge both concepts.

2. There is not "Discussion" section. It would be advisable to introduce a short discussion comparing the improved protocol to previous ones, giving some conclusions, and highlighting the possible applicability of the methodology to other bacteria.

Thank you very much for this suggestion. We have now included a brief Discussion section explaining the advantages of this strategy with respect to other previous approaches and introducing specific parts of the procedure that are amenable to adaptation if using sphingomonad species different to our *S. granuli* model.

Reviewer 2

This report describes a modification of the SceI-based genome editing method that is currently used in Pseudomonas, to be used in Sphingomonadaceae. The clever approach of the authors, includes the forcing of a second recombination in a different mating process from that what incorporates in the target strain the plasmid containing the SceI sites. This smart technique has been possible due to the vast knowledge that the authors have accumulated about Sphingomonas granuli strain TFA as can be deduced from their publications, and kindly applied here.

However, some minor comments will improve the final quality of the paper.

A short phrase would be included to emphasize that the deletion of ecfG2 is not deleterious, and that the mutant bacteria behaves similarly to wt strain.

Thank you for this comment, we have now included a sentence explaining this.

L237-238: “Although *ecfG2* is essential under stressing conditions, a mutant in this gene shows a similar fitness compared to the wild type TFA in the absence of stress”.

Line 67. In this regard, delete s

Thank you for this comment, the text has been corrected accordingly.

Line 182. Rephrase… drastically reduces the number of clones reduces the number of

Thank you for this comment, the text has been corrected accordingly.

Line 349. Rephrase …clones after obtained…

Thank you for this comment, the text has been corrected accordingly.

Throughout the text: To be smarter avoid colloquial use of techniques. Plasmids are not transformed into strains. Strains are transformed with plasmids. Examples: Lines 344, 363 and others. Please revise the whole text.

Thank you for this suggestion, we have corrected this throughout the text and in the supplementary material.