

Conjugative serine recombinase engineering enables genome integration in transformation-resistant bacteria and bioproduction pathway prototyping

Michael S. Guzman^{1,2*†}, Choltipit Kiattisewee^{1,2,†}, Amanda M. Robert^{1,2+}, Jackson Comes^{1,2+}, Ryan A.L. Cardiff^{1,2}, Allan Scott³, Margaret Cook^{1,2}, Diego Alba Burbano^{1,2}, Stella Anastasakis^{1,2}, Sarah Grube^{1,2}, Kieran Heiberg¹, Brian Darst^{1,2}, Daniel Howell^{1,2}, Gara N. Dexter⁴, Robert G. Egbert³, Jesse G. Zalatan^{1,2,5}, Adam M. Guss⁴, Joshua R. Elmore^{3,*}, Alexander S. Beliaev^{6,*}, James M. Carothers^{1,2*}

1: Department of Chemical Engineering
University of Washington
Seattle, WA 98195
United States

2: UW Center for Synthetic Biology and Molecular Engineering & Sciences Institute
University of Washington
Seattle, WA 98195
United States

3: Biological Sciences Division, Pacific Northwest National Laboratory
Richland, WA 99354
United States

4: Biosciences Division, Oak Ridge National Laboratory
Oak Ridge, TN 37831
United States

5: Department of Chemistry
University of Washington
Seattle, WA 98195
United States

6: Environmental Molecular Sciences Division,
Pacific Northwest National Laboratory
Richland, WA 99354
United States

†, + These authors contributed equally.

* Correspondence to:

Michael S. Guzman: msguzman@uw.edu

Joshua Elmore: joshua.elmore@pnnl.gov

Alex Beliaev: alex.beliaev@pnnl.gov

James M. Carothers: jcaroth@uw.edu

Abstract

Non-model bacteria offer unique metabolic capabilities for sustainable bioproduction, yet their limited genetic accessibility hinders systematic strain development. Here we present conjugation-based serine recombinase-assisted genome engineering (cSAGE), a portable platform for sequential, site-specific genomic integration in transformation-resistant bacteria. Using *Rhodobacter sphaeroides* as a testbed, we show that genome integration fails by electroporation and is not rescued by disruption of three candidate restriction endonucleases. Conjugative delivery of the same integration vectors recovers genomic integrants at frequencies of approximately 10^{-4} per recipient cell, an improvement of at least five orders of magnitude over electroporation. We demonstrate site-specific genomic integration across eight bacterial hosts spanning four classes and two phyla, as well as three sequential payload integrations using a standardized, non-replicating plasmid toolkit. We then use cSAGE to install the same bioproduction pathway in two *Rhodopseudomonas palustris* strains: the reference strain CGA009 and the environmental isolate P4, for which genetic engineering has not previously been reported. Under anoxygenic photosynthetic growth conditions, engineered P4 produces *p*-vinylphenol from *p*-coumarate, whereas CGA009 engineered with the same construct, does not. By enabling standardized genome engineering and cross-host evaluation of bioproduction pathways, cSAGE provides a general framework for non-model strain prototyping and biotransformation discovery.

Introduction

Efficient microbial conversion of underutilized waste-derived feedstocks into valuable products is central to a sustainable bioeconomy¹. Across agricultural, chemical, and wastewater management sectors, large reservoirs of carbon exist in the form of volatile fatty acids², lignocellulosic residues³, and industrial off-gases⁴. Harnessing these resources relies on leveraging the native metabolic capabilities of non-model bacteria and environmental isolates, together with engineered functions for efficient chemical and fuel production. Among these, purple nonsulfur bacteria (PNSB) represent a particularly versatile and underexploited group of microorganisms for bioproduction⁵. PNSB couple CO₂ fixation, photosynthesis, and carbon-efficient aromatic compound degradation within a single organism, allowing them to use low-energy infrared light to catabolize waste carbon feedstocks, including lignin-derived aromatics and syngas^{6–10}. Despite their metabolic flexibility and the availability of curated genome-scale metabolic models predicting production capabilities^{11,12}, the use of PNSB for sustainable biomanufacturing remains limited by the lack of robust, stable, and portable genetic tools^{6,7,13}. Like other non-model organisms, many PNSB strains remain recalcitrant to transformation, creating a major bottleneck for strain engineering¹⁴.

Over the past decade, many advanced tools have been developed to address barriers in microbial genetics and engineering. However, these tools require tradeoffs among DNA delivery, host compatibility, and genome-integration throughput. For example, approaches such as conventional allelic exchange¹⁵, CRISPR-Cas systems¹⁶, transposon-based methods¹⁷, and recombineering¹⁸ are often limited by transformation efficiency, host-range compatibility, or the need for host-specific genetic components¹⁶. These approaches can require host-compatible delivery vectors, replication origins, expression systems, selectable markers, or counterselection strategies. Methods that rely on replicating plasmids also require compatible replication origins and may necessitate additional plasmid-curing steps^{16–20}. Moreover, replicating plasmids typically require prior screening of functional genetic parts for each new organism²¹. As a result, genome integration strategies that perform well in some species often fail to generalize across bacteria²².

Beyond transformation efficiency and host range, genome integration throughput presents a further limitation for non-model strain construction¹⁵. Widely used Tn7 transposon-based strategies typically support a single integration event^{19,23} at the conserved *attTn7* site, after which target immunity limits reuse of that site. CRISPR-Cas systems require cloning of strain-specific homology-directed repair plasmids for each successive modification²⁰. Although both approaches can be made compatible with conjugative DNA delivery^{19,23} successive modifications require either additional genomic target sites or newly designed repair constructs. Recombinase-based landing-pad systems such as CRAGE similarly support conjugative delivery and broad cross-phylum host range, but current implementations support at most two sequential integration sites^{24,25}. To our knowledge, no current platform combines conjugative delivery, multiple orthogonal integration sites, marker recycling, and a standardized non-replicating plasmid toolkit for sequential payload integration across diverse bacterial hosts. This gap limits scalable strain construction in non-model organisms.

Serine recombinase-based integration systems have recently emerged as versatile tools for genome engineering in non-model bacteria²⁶. Serine-recombinase-assisted genome engineering (SAGE) enables site-specific, unidirectional integration of DNA into the host chromosome by catalyzing recombination between a plasmid-encoded, phage-derived attachment site (*attP*) and a corresponding bacterial chromosomal attachment site (*attB*)²⁷. Once a poly-*attB* landing pad composed of multiple orthogonal *attB* sites is installed in the genome, SAGE supports efficient and stable integration of large or multi-gene payloads by using distinct recombinases to direct successive *attP*-flanked cargos to cognate *attB* sites. This architecture enables sequential payload integration without constructing new host-specific homology arms for each insertion. However, broader use of SAGE has been limited by its reliance on electroporation for DNA delivery. Although electroporation is effective in many bacteria, its performance is strongly host dependent, requiring optimization of electroporation conditions and competent-cell preparation for each new species¹⁵. Even after extensive optimization, some non-model bacteria remain refractory to electroporation²⁸.

Conjugation provides several advantages that make it an attractive alternative to electroporation for DNA delivery. First, conjugation is a highly conserved mechanism of horizontal gene transfer across the bacterial kingdom that functions in hosts ranging from laboratory *E. coli* strains to non-model microbes^{29–31}. Second, conjugative DNA is initially transferred as single-stranded DNA and then converted to double-stranded DNA in the recipient cell, which can reduce its exposure to some host restriction barriers relative to delivery of double-stranded DNA^{32–34}. Accordingly, several recent broad-host-range genetic toolkits rely on conjugation as a robust and portable DNA-delivery strategy^{19,23,25}. Conjugation has become an established method for introducing DNA into bacteria that are otherwise refractory to transformation, including PNSB, gut microbiota, marine and soil bacteria, and extremophiles²⁹. We therefore reasoned that coupling conjugative delivery to SAGE could reduce its dependence on host-specific electroporation while preserving sequential, site-specific payload integration.

Here we present conjugation-based SAGE (cSAGE), a portable genome engineering toolkit that combines conjugative delivery with sequential, site-specific payload integration. In cSAGE, all plasmids are designed in the Standard European Vector Architecture (SEVA) format^{35,36} with an RK2/RP4 origin of transfer (*oriT*) for conjugative delivery and a *pir*-dependent R6K origin of replication that supports plasmid maintenance in specialized donor strains but prevents replication in recipient hosts lacking the π replication protein (Pir)³⁷. We paired this with a conjugation-compatible Tn7 strategy²³ for landing-pad integration. Once a compatible poly-*attB* landing pad is installed, the same recombinase and payload plasmids can be deployed across diverse bacteria without host-specific redesign.

We first use *Rhodobacter sphaeroides* to test whether conjugative delivery can overcome a defined barrier to SAGE. SAGE fails by electroporation in this host, a barrier not rescued by restriction endonuclease knockout (Fig. 1), but is resolved by cSAGE, which increases recovery of genomic integrants by at least five orders of magnitude relative to delivery of the same vectors by electroporation. We next show that cSAGE supports site-specific genomic integration across a diverse panel of bacteria, including several non-model PNSB strains, two *Pseudomonas* species, and *Rhodococcus jostii*. The toolkit also supports sequential multi-payload integration and marker recycling. Finally, we apply cSAGE to compare

strain-specific differences in photosynthetic conversion of the lignin-derived aromatic *p*-coumarate to the polymer precursor *p*-vinylphenol across *Rhodopseudomonas palustris* strains, including the wastewater-derived environmental isolate P4. The same integrated pathway produced *p*-vinylphenol in P4 but not in the reference strain CGA009. Together, these results establish cSAGE as a portable platform for sequential genome engineering and comparative strain prototyping in genetically inaccessible bacteria.

Results

SAGE fails by electroporation in Rhodobacter sphaeroides, a barrier not rescued by restriction-modification knockout

Serine recombinase-assisted genome engineering (SAGE) enables single-copy, site-specific integration of genetic cargo through recombination between a chromosomally integrated poly-*attB* landing pad and a plasmid-borne *attP* site²⁶. This system was previously validated in *Rhodopseudomonas palustris* CGA009 and several other species²⁶, where two non-replicating plasmids, an integrase expression plasmid and an *attP*/cargo plasmid, are delivered by electroporation (Fig. 1A-B). *Rhodobacter sphaeroides* have a long history of use in bacterial photosynthesis research, yet remain resistant to standard transformation methods^{7,38}. To extend SAGE to this host, we introduced a poly-*attB* landing pad cassette into the *R. sphaeroides* 2.4.1 genome by allelic exchange. Restriction-modification systems are a common barrier to foreign DNA uptake in transformation-resistant bacteria, and knockout of the *R. sphaeroides* type II restriction endonuclease *RshI* was previously reported to improve transformation efficiency^{38,39}. We therefore integrated the landing pad cassette directly at the *RshI* locus, simultaneously installing the poly-*attB* array and disrupting *RshI*, generating a knockout strain ($\Delta rshI$) (Fig. 1A).

We then tested whether genetic payloads could be delivered with SAGE to this landing pad strain. We co-electroporated the integrase plasmid (pGW31) and *attP*/cargo plasmid (pGW60) into *R. palustris* CGA009 as a positive control and into the *R. sphaeroides* landing pad strain, and quantified transformation efficiency by selecting for the cargo plasmid's resistance marker (Fig. 1C). Electroporation yielded high transformation efficiency in *R. palustris* CGA009, consistent with prior reports²⁶. In *R. sphaeroides*, however, transformants were recovered at or near the limit of detection, despite the landing pad strain already lacking *RshI*. Given that the most extensive published optimization of *R. sphaeroides* electroporation achieved only a ~10-fold improvement (reaching ~7,000 transformants/ μ g DNA for a single replicating plasmid)³⁹, a magnitude far short of the barrier observed here for two non-replicating SAGE vectors, we did not pursue further optimization of electroporation parameters for *R. sphaeroides*. We instead investigated restriction-modification as a testable hypothesis for the observed barrier.

To ask whether additional, uncharacterized restriction-modification systems account for this barrier, we identified two further candidate type II restriction endonucleases encoded in the *R. sphaeroides* genome, RSP_RS17455 and RSP_RS21745. We generated a series of strains with sequential in-frame deletions (Supplementary Table 1) and compared their transformation

efficiency to wild-type *R. sphaeroides* by electroporation of pBBR1, a broad-host-range replicating plasmid⁴⁰. We used pBBR1 rather than the non-replicating SAGE vectors for this comparison because, as a single autonomously replicating plasmid, it provides a higher baseline transformation efficiency and isolates the restriction-endonuclease variable from the added complexity of two-plasmid, non-replicating SAGE delivery (Fig. 1D).

Transformation efficiency did not improve with successive restriction endonuclease knockouts. Knockout strains showed transformation efficiency comparable to wild-type *R. sphaeroides* ($P = 0.86$, one-way ANOVA; Fig. 1D). These results indicate that the three restriction endonucleases examined here are not sufficient to explain the electroporation barrier in *R. sphaeroides*, and that additional, poorly defined mechanisms likely limit DNA uptake.

Conjugation-based SAGE resolves the engineering bottleneck in Rhodobacter sphaeroides

Because the tested restriction-modification knockouts did not improve electroporation efficiency in *R. sphaeroides*, we evaluated conjugation as an alternative delivery route. Conjugation delivers DNA as a single-stranded intermediate directly through a mating pore^{32–34}. This bypasses several steps required for electroporation, including membrane permeabilization and cytoplasmic entry of double-stranded DNA, that may limit uptake in this host²⁸. Conjugation is also broadly compatible with Gram-negative bacteria and does not require host-specific electrocompetent cell preparation, making it an attractive strategy for extending SAGE to transformation-resistant hosts^{29–31}. Conjugative delivery has previously enabled genome engineering in other systems facing similar barriers, including Tn7-based transposition^{19,23}.

To enable conjugative delivery, we redesigned the SAGE plasmids by incorporating an RK2/RP4 *oriT*, generating the integrase plasmid pCK1205 and the *attP*/cargo plasmid pAR054 (Fig. 2A). We further restructured both plasmids into the Standard European Vector Architecture (SEVA) format, a modular design widely adopted for non-model bacterial hosts^{35,36} (Fig. 2A, Supplementary Fig. 1). We also replaced the replicon with R6K, a low-copy origin that requires the *pir* gene for replication³⁷. The *pir*-dependent R6K origin lowers plasmid copy number and metabolic burden during cloning in *pir*⁺ *E. coli*⁴¹ and ensures the plasmids do not replicate in *R. sphaeroides* or other recipient hosts following conjugation³⁷, so that only cells with a stable, integrated cargo cassette retain antibiotic resistance.

To deliver these plasmids, we used a tri-parental mating strategy (Fig. 2B). Two *E. coli* WM6026 donor strains, each *pir*⁺ and $\Delta dapA$ and therefore auxotrophic for diaminopimelic acid (DAP)⁴², were used: one carrying the integrase plasmid and the other carrying the *attP*/cargo plasmid. Both donor strains were co-cultured with the SAGE-enabled *R. sphaeroides* recipient, allowing conjugative transfer of both plasmids into the same recipient cell via the mating pilus. Mating mixtures were then plated on medium lacking DAP and containing the antibiotic corresponding to the cargo plasmid's resistance marker. This selects against both donor strains, which cannot grow without DAP, and for recipient cells with stable, site-specific integration of the cargo cassette, since the non-replicating SAGE plasmids are otherwise lost after cell division.

Using this conjugation-based approach (cSAGE), we compared electroporation and conjugation efficiency (both of which were normalized to total recipient CFUs) directly in *R. palustris* CGA009 and *R. sphaeroides*, using the same R6K-based cSAGE vectors for both

delivery methods to isolate delivery method as the sole variable (Fig. 2C). Conjugation and electroporation gave comparable recovery frequencies in *R. palustris* CGA009 ($P = 0.71$, Welch's t -test; Fig. 2C) but conjugation yielded at least five orders of magnitude more recovered integrants than electroporation in *R. sphaeroides* (Fig. 2C). In *R. sphaeroides*, cSAGE efficiently and reproducibly achieved transconjugant recovery at frequencies of $\sim 10^{-4}$ per recipient cell (1 in 10,000), whereas electroporation efficiency fell below our limit of detection ($\leq 10^{-9}$). Under the matched tri-parental mating conditions used here, the tested mini-Tn7 configuration yielded no recoverable transconjugants in *R. sphaeroides* (Supplementary Fig. 2). Published Tn7 workflows using an equivalent tri-parental mating scheme have similarly failed to recover transconjugants¹⁹. In the published study, reliable recovery required either an additional conjugal helper strain or a recipient engineered to overexpress the Tn7 transposase, with mean frequencies of 10^{-4} to 10^{-3} across the tested configurations¹⁹. cSAGE achieves comparable transconjugant recovery ($\sim 10^{-4}$ per recipient cell) using a standardized tri-parental mating scheme and non-replicating vectors.

Conjugation-based SAGE enables efficient genomic integration across diverse bacteria

To evaluate the portability of cSAGE beyond *R. sphaeroides*, we assembled a diverse panel of PNSB from the Alphaproteobacteria and Betaproteobacteria (Table 1)^{43,44}. For *R. palustris* P4, *Rhodospirillum rubrum*, and *Rubrivivax gelatinosus*, we integrated the poly-*attB* landing pad using a transposon-based strategy²³ (Fig. 3A). Tn7 systems have been widely adopted for single-copy gene integration in non-model organisms, but transposition immunity restricts integration to a single event per *attTn7* site²³, precluding iterative use of Tn7 itself for additional integrations⁴⁵. By using Tn7 only once to install a poly-*attB* landing pad containing ten orthogonal *attB* sites, we can enable iterative engineering at this locus. We introduced a single standardized delivery plasmid (pRC414) to SAGE-enable multiple hosts without redesigning integration vectors for each species. pRC414 itself carries an RK2/RP4 *oriT* (Fig. 3A) and was delivered to each recipient by conjugation, consistent with our conjugation-based delivery strategy (Fig. 2B).

Using this panel of five PNSB strains, we screened eight orthogonal serine recombinases: Bxb1, R4, TG1, ϕ BT1, RV, MR11, ϕ C1, and BL3. Each recombinase was delivered by tri-parental conjugation as in Fig. 2B. cSAGE produced detectable transconjugants in all five PNSB strains, and for nearly all recombinase-host combinations (Fig. 3C). To enable this comparison across recombinases and hosts within a consistent genetic context, we employed a poly-*attP* reporter plasmid series, each carrying one of eight *attP* sites paired with a constitutively expressed mRFP reporter (Fig. 3B)²⁶. When paired with the cognate recombinase expression plasmid, integration is directed specifically to the matching *attB* site, resulting in chromosomal insertion of the mRFP cassette, confirmed by colony PCR. As a benchmark, we also measured conjugation efficiency using a pBBR1 control plasmid.

Both integration efficiency and accuracy influence the practical utility of a recombinase. We quantified conjugation efficiency as the frequency of transconjugant colonies per viable recipient cell and integration accuracy as the fraction of screened colonies with on-target integration. Bxb1, R4, and TG1 consistently yielded high conjugation efficiency across all

strains, with some values comparable to pBBR1 control (Fig. 3C). The remaining ϕ C1, ϕ BT1, RV, and BL3 recombinases showed variable conjugation efficiency across species, with ϕ C1 yielding detectable transconjugants only in *R. palustris* CGA009. Bxb1, TG1, R4, and RV were also among the most accurate recombinases tested across all five species. The remaining recombinases showed more variable integration accuracy compared to Bxb1, TG1, R4, and RV.

No transconjugants were recovered in *R. gelatinosus* CBS using the replicating pBBR1 control, whereas cSAGE integrations were readily obtained in this host (Fig. 3C). To assess whether this portability extends beyond PNSB, we evaluated cSAGE in phylogenetically distinct bacteria, including *Pseudomonas putida* KT2440^{46,47}, *Pseudomonas sorgoleonovorans* SO81⁴⁸, and *Rhodococcus jostii* RHA1⁴⁹, using the three highest-performing recombinases identified in PNSB (Bxb1, R4, and TG1) (Supplementary Fig. 3). Electroporation of *P. sorgoleonovorans* was previously found to be highly inefficient, requiring conjugation for gene deletion and DNA integration⁴⁸. We observed detectable integration across all three hosts although transconjugant recovery frequency and integration accuracy varied among recombinase-host combinations. While this dataset is not intended as a comprehensive host-range analysis, cSAGE supported genomic integration across four bacterial classes spanning two phyla.

Iterative genome engineering workflows enabled by cSAGE

Iterative strain engineering in non-model bacteria involves repeated rounds of genomic integration to incrementally assemble complex or burdensome genetic programs, such as multi-gene metabolic pathways that cannot be introduced in a single step¹⁵. In practice, such workflows are limited by the repertoire of effective selectable markers^{21,50} due to intrinsic resistance to antibiotics¹³, cross-resistance conferred by some antimicrobial resistance genes⁵¹, and inconsistent performance of counterselection systems across species⁵⁰. As a result, many existing approaches combine sequential rounds of integration with intermediate marker or backbone removal to enable marker reuse, adding experimental steps and imposing additional host-specific requirements²². These limitations are particularly prevalent in non-model organisms such as PNSB, where counterselection systems are often unavailable, unreliable, or require extensive characterization alongside selection markers⁵². Engineering workflows that enable rapid construction of multi-payload strains without mandatory marker recycling, while retaining optional marker excision, remain scarce. To address this need, we developed two conjugation-compatible workflows: a multi-payload integration toolkit that enables sequential construction of marked strains without counterselection, and a ϕ C31-based marker-recycling strategy that supports removal of antibiotic resistance markers where compatible counterselection systems exist.

We tested whether multiple payloads could be integrated sequentially into cognate sites within the poly-*attB* landing pad without cross-recombination or interference. We designed payload plasmids containing distinct *attP* sites matched to orthogonal *attB* loci to enable independent recombinase-mediated integration events. We constructed three orthogonal serine recombinase systems—Bxb1, R4, and TG1—based on their strong performance and paired each with a unique antibiotic-resistance marker (Km^R , Sm^R , or Gm^R) for independent selection (Fig. 4A). As a proof of concept, we introduced orthogonal fluorescent reporters (mRFP, sfGFP, and

mTagBFP2) as cargos into each of the three attachment sites. We performed sequential rounds of integration in both *R. palustris* CGA009 and *R. sphaeroides*, selecting only for the newly introduced antibiotic marker at each step and without intermediate excision of previously integrated backbones. Despite the introduction of repetitive R6K backbone sequences from prior integrations, fluorescence from all three reporters was retained through the successive rounds of integration (Fig. 4B-C; Supplementary Fig. 5).

While the multi-payload integration strategy enables strain construction without backbone excision, the generation of antibiotic-resistance-free strains remains important for many downstream applications. In the SAGE system, ϕ C31 integrase catalyzes excision of plasmid backbones flanked by *attP* and *attB* sites following payload integration, leaving behind only the desired cargo and a small recombination scar (*attL*) with minimal expected impact on host physiology²⁶. However, delivery of ϕ C31 integrase in the original implementation relied on either a replicative mSF^{ts1} plasmid (a *Pseudomonas*-specific origin⁵³) or a non-conjugative pUC-derived plasmid²⁶. These delivery systems have limited applicability in PNSB where mSF^{ts1} does not replicate and electroporation is inefficient. To evaluate marker recycling and backbone excision in a fully conjugation-based workflow, we reimplemented the ϕ C31 integrase helper plasmid in the conjugative R6K vector. We constructed a poly-*attP* target plasmid containing constitutively expressed *mRFP* and a *sacB* gene for counterselection (Fig. 5A). We integrated this *sacB*-containing cargo into *R. palustris* CGA009 and *R. sphaeroides* using a Bxb1-based cSAGE protocol. Strains with the chromosomally integrated cargo then underwent diparental conjugation to deliver the ϕ C31-encoding R6K plasmid. Following conjugation and sucrose counterselection, we observed marker curing in both species (Fig. 5B-C). In *R. palustris* CGA009, 16 of 16 colonies yielded the expected ~1.5 kb PCR product, consistent with precise excision of the plasmid backbone (Fig. 5B). Similarly, in *R. sphaeroides*, 14 of 16 colonies yielded the expected PCR product; no diagnostic amplicon was obtained from the remaining two colonies, leaving their curing status unresolved.

The multi-payload integration toolkit supports reliable, counterselection-free construction of strains carrying multiple integrated payloads, making it particularly well suited for workflows where rapid pathway assembly, testing, and comparison are prioritized. In parallel, ϕ C31-mediated marker recycling provides an efficient route to generate antibiotic-resistance-free strains when compatible counterselection systems are available. These capabilities establish a flexible engineering framework for strain construction that supports both rapid prototyping and marker-free strain construction.

Light-powered aromatic bioproduction in purple nonsulfur bacteria

PNSB represent one of the most metabolically versatile groups of microorganisms known, capable of thriving under anaerobic conditions while using infrared light energy and a broad range of inorganic and organic compounds as energy and carbon sources⁵. *R. palustris*, in particular, can oxidize diverse substrates, including CO₂, H₂, and organic acids, making it a promising chassis for circular bioeconomy applications⁶. The ability of *R. palustris* to catabolize lignin-derived aromatics under anoxygenic conditions^{54–56} further distinguishes this species as a candidate for valorizing industrial waste streams such as lignocellulosic hydrolysates^{10,56} and

wastewater effluents⁸. However, this metabolic versatility is underpinned by complex regulatory networks that respond to light availability, O₂ tension, and nutrient status to control distinct carbon assimilation and redox balancing pathways⁵⁷. Effective bioproduction in PNSB, therefore, requires not only understanding their metabolic potential but also genome engineering tools that enable predictable control of heterologous pathway expression across species.

For well-studied PNSB such as *R. palustris* CGA009, genome-scale metabolic models (GEMs) provide a useful framework for predicting metabolic fluxes and evaluating production potential⁵⁸. However, GEM construction is labor-intensive, requires extensive curation and experimental validation, and models are often not readily transferrable even between closely related strains⁵⁹. This limitation becomes particularly apparent when working with environmental isolates. For example, we recently sequenced and annotated the genome of the wastewater isolate *R. palustris* P4 (GenBank accession number CM151374.1). P4 has a previously characterized ability to couple carbon monoxide oxidation to hydrogen production through the water-gas shift reaction^{43,60,61}, but its capacity to degrade aromatic compounds had not been previously characterized. To our knowledge, P4 had also not previously been genetically engineered, making this study the first demonstration of genome engineering in this strain. Comparative genomic analysis revealed numerous strain-level differences relative to CGA009, complicating *a priori* prediction of how native metabolism would interact with an engineered pathway (Supplementary Table 6 and 7). In this context, cSAGE provides an empirical approach to evaluate metabolically distinct but poorly characterized environmental isolates for bioproduction. Rather than optimizing a pathway within a single host, our goal was to deploy the same genetic constructs across multiple hosts to identify productive strain backgrounds.

To illustrate this capability, we focused on *R. palustris* strains as hosts for aromatic bioconversion. *R. palustris* possesses well-characterized pathways for uptake and catabolism of *p*-coumarate (*p*-CA) and related aromatics, including a conserved coumarate degradation operon (the *CouAB* genes and associated transport systems^{54,62}). Because substrate uptake and native catabolic capacity are prerequisites for conversion, *R. palustris* provides a relevant chassis for evaluating aromatic bioproduction. We therefore integrated a codon-optimized *BaPAD* expression cassette encoding a phenolic acid decarboxylase from *Bacillus atrophaeus*. *BaPAD* converts *p*-CA to *p*-vinylphenol (*p*-VP)^{63,64}, a precursor for polymers⁶⁵, adhesives⁶⁶, and photolithography materials⁶⁷ (Fig. 6A). Because *R. palustris* strains differ in their genomic content (Supplementary Tables 6 and 7), we hypothesized that host-specific physiology could strongly influence bioproduction outcomes even when the heterologous pathway is introduced.

Before integrating *BaPAD*, we used cSAGE to benchmark promoter activity across the two *R. palustris* strains using a set of synthetic constitutive promoters characterized in *E. coli* and other Gram-negative bacteria⁶⁸ (Fig. 6B; Supplementary Fig. 6). Promoter characterization is important for controlling heterologous gene expression. Although the rank correlation of these synthetic promoters is generally conserved in previously characterized species⁶⁸, their activity can vary among hosts and is best determined experimentally. Using mRFP as a reporter, we identified BBa_J23110 as a medium-strength promoter within each hosts' promoter panel, though its absolute expression level differed between hosts. We then used cSAGE to integrate the BBa_J23110::*BaPAD* expression cassette. The engineered strains were cultivated photoheterotrophically with *p*-CA as the sole carbon source (Fig. 6C, Supplementary Fig. 7).

Despite complete consumption of *p*-CA substrate, the CGA009_BaPAD strain produced no detectable extracellular *p*-VP (Fig. 6C, Supplementary Fig. 7). This result is consistent with native *p*-CA metabolism limiting net product accumulation. Flux simulations using the iAN1128 GEM adapted for CGA009_BaPAD⁵⁸ predicted no net flux toward *p*-VP production and instead routed carbon through native aromatic degradation pathways and into central metabolic intermediates such as acetyl-CoA. In contrast, *R. palustris* P4_BaPAD produced measurable amounts of *p*-VP, converting approximately 6% of the supplied *p*-CA to product (Fig. 6C). The *p*-CA catabolic operon has an identical nucleotide sequence in CGA009 and P4 (Supplementary Fig. 8). Comparative analysis of KEGG orthologies identified additional functions in P4 associated with phenylpropanoid metabolism, redox processes, and transport. Differences in redox balance, pathway regulation, product turnover or export could therefore contribute to the observed divergence in *p*-vinylphenol accumulation between strains (Supplementary Tables 6 and 7).

Together, these results demonstrate that host physiology can strongly influence pathway performance, even among closely related strains. More broadly, this experiment illustrates how cSAGE enables direct evaluation of bioproduction pathways across hosts using empirical performance data. By enabling stable genome integration across hosts, cSAGE reduces confounding associated with plasmid copy number and genetic instability¹³. This capability provides a practical route to identify productive strain backgrounds early in the strain development pipeline, particularly for non-model organisms for which predictive models are unavailable or insufficient.

Discussion

Poor transformation efficiency remains one of the most common barriers to genome engineering in microbial synthetic biology¹⁴. Conventional domestication strategies scale poorly—each new host typically requires constructing bespoke allelic exchange vectors, compatible origins of replication, and genetic parts validated from scratch^{15,22}. This challenge is not limited to poorly characterized organisms: in *R. sphaeroides*, restriction endonuclease knockout, a targeted intervention, failed to rescue electroporation efficiency (Fig. 1), demonstrating that host-specific troubleshooting does not always resolve transformation barriers even when a plausible cause is identified. Conjugative delivery of the same vectors instead increased recovery of genomic integrants by at least five orders of magnitude (Fig. 2), indicating that delivery, not integration, was limiting. By coupling conjugative DNA delivery with serine recombinase-mediated genomic integration and a narrow-range R6K vector backbone, the same pre-built donor system applies to new organisms with minimal development, including previously uncharacterized environmental isolates such as *R. palustris* P4, for which no genetic tools existed prior to this work. cSAGE relies on Tn7 transposition or allelic exchange for the one-time installation of its landing pad, complementary methods that remain the most effective tools for that specific step but require host- and locus-specific optimization (e.g., homology arm design) for each new target (Table 3). To facilitate community adoption of cSAGE, we have deposited a curated set of plasmids to Addgene, summarized in Table 2.

The variation we observe across recombinases (Fig. 3), even when expressed from identical promoters and delivered by conjugation, suggests that performance is in part determined by intrinsic sequence preferences and host genomic context^{69,70}. RV achieved 100% on-target integration across all five PNSB despite low overall conjugation efficiency (Fig. 3), suggesting that accuracy and efficiency reflect distinct mechanistic properties that could be independently optimized. Applying motif-scanning or pseudo-*attB* site prediction⁷¹ approaches to bacterial genomes could reveal the mechanistic basis for these differences and reduce the need for empirical screening when porting SAGE to new hosts.

Out of the PNSB strains tested in this work, *R. palustris* CGA009 and P4 were the most closely related, despite exhibiting divergent *p*-VP production phenotypes (Fig. 6C). While both CGA009 and P4 possess the *p*-CA catabolic operon (Supplementary Fig. 8), P4 produced measurable *p*-VP while CGA009 did not. Reporter output from J23110 constructs integrated at the respective landing pads was higher in P4 than in CGA009 ($P < 0.0001$, one-way ANOVA; Fig. 6B), raising the possibility that promoter and enzyme dosage contribute to differences in *p*-VP titer. However, a ~5-fold difference in expression may not completely explain the binary outcome we observed: CGA009 produced no *p*-VP despite complete consumption of *p*-CA (Fig. 6C). This is consistent with the GEM simulations, which predict zero net flux toward *p*-VP in CGA009 regardless of BaPAD expression level, because competing native catabolic pathways are predicted to consume the available flux upstream of the heterologous step (Fig. 6A). The underlying basis for this divergence remains to be resolved, but comparative genomic analysis reveals that P4 encodes additional functions related to phenylpropanoid metabolism, redox processes, and transport relative to CGA009 (Supplementary Tables 6 and 7), any of which could contribute to the observed difference.

Several non-mutually exclusive mechanisms could also account for the distinct production phenotype. Differences in *p*-CA import or CoA ligation kinetics could alter the balance between BaPAD-mediated decarboxylation and native catabolic flux. CGA009 encodes two redundant, high-affinity *p*-CA transport systems⁷², and comparative genomic analysis indicates P4 differs in transport-related loci (Supplementary Tables 6 and 7), potentially allowing CGA009 to import and co-ligate *p*-CA (via CouB) more efficiently and limit the pool of free *p*-CA available to BaPAD⁶². Differences in regulatory control of the coumarate degradation (*cou*) operon, which in CGA009 is controlled by the MarR-family repressor CouR and derepressed specifically by *p*-coumaroyl-CoA⁶², could similarly affect how quickly catabolic flux activates. Strain differences in *p*-VP export capacity may also contribute, since vinylphenol derivatives are more toxic to Gram-negative bacteria than their parent phenolics, and are not reliable inducers for known aromatic acid efflux systems⁷³. Distinguishing among these possibilities will require targeted experiments, including flux measurements, transporter knockouts, and *in vitro* enzyme assays⁶⁴. Directly deleting competing catabolic genes could in principle redirect flux toward *p*-VP, but carries a risk of unintended loss of function. The naturally genome-reduced *R. palustris* strain DSM127, which lacks these catabolic genes, has lost the capacity for anaerobic aromatic degradation entirely⁷⁴. Targeted, additive engineering approaches such as cSAGE, which leave native metabolism intact, may therefore be preferable for metabolically versatile organisms such as PNSB, where the functional consequences of pathway deletion are not always predictable.

By combining broad-host-range genome engineering with photosynthetic chassis, cSAGE extends the reach of the circular bioeconomy into organisms capable of valorizing underutilized carbon streams. The ability of PNSB to integrate CO₂ fixation, aromatic catabolism, and phototrophic growth positions them as uniquely powerful platforms for carbon-neutral bioproduction from lignin-derived aromatics, organic acids, and syngas^{6,7}. The *p*-VP route demonstrated here illustrates this potential: BaPAD-catalyzed decarboxylation removes only the carboxyl group from *p*-CA, retaining the aromatic ring rather than requiring the extensive ring-cleavage and reassembly typical of central catabolic degradation^{56,63}. The multi-payload integration and promoter characterization enabled here lay the groundwork for predictive strain design and for coupling genome engineering with metabolic and systems-level models^{11,58}. Beyond PNSB, cSAGE also extends to taxonomically diverse hosts including Gram-positive *R. jostii* RHA1 and Gram-negative *P. putida* KT2440 and *P. sorgoleonovorans* SO81 (Supplementary Fig. 3), relevant to bioremediation⁷⁵, bioproduction⁴⁷, and agriculture⁴⁸, confirming that the platform is not confined to any single phylogenetic group. Broadly, cSAGE serves as a platform for expanding genetic accessibility and identifying productive strain backgrounds, with future extensions to CRISPR-based regulation^{76,77}, synthetic pathway construction⁷⁸, and genetic code expansion⁷⁹.

Methods

Plasmid and strain construction

Plasmids used in this study are listed in Supplementary Table 2. Key plasmids used in this study, including Addgene IDs, are listed in Table 2. All cSAGE plasmids were assembled in the Standard European Vector Architecture (SEVA) format³⁶. Plasmids were constructed either using In-Fusion Snap Assembly (Takara Bio), unless otherwise noted, and verified by whole-plasmid Nanopore sequencing. A complete list of plasmids, assembly primers, and notes can be found in Supplementary Table 5. Base strains carrying poly-*attB* landing pads were generated for each PNSB host to enable site-specific integration. The *R. sphaeroides* landing pad and restriction endonuclease gene deletion strains were generated by two-step allelic exchange using pK18mobSacB-derived vectors. *R. palustris* P4, *R. rubrum*, and *R. gelatinosus* landing pads were delivered by a Tn7 transposon-based system we adapted for this study, where a Tn7 delivery vector carrying the poly-*attB* cassette and flanking FRT sites allows for subsequent excision of the antibiotic resistance marker. The Tn7 expression plasmid was co-transformed as either a non-replicating (pJMP1039) or replicating (pJMP8045) helper plasmid. pJMP1039 was a gift from Carol Gross & Jason Peters & Oren Rosenberg (Addgene plasmid #119239; <http://n2t.net/addgene:119239>; RRID:Addgene_119239). pJMP8045 was a gift from Jason Peters (Addgene plasmid #220900; <http://n2t.net/addgene:220900>; RRID:Addgene_220900). We used PCR to verify correct cassette insertion into each organism. Reporter and pathway integration strains were generated by tri-parental conjugation using *E. coli* WM6026⁴² donors carrying the integrase expression plasmid and the cognate *attP* payload plasmid. Single colonies were screened for correct integration by colony PCR and reporter fluorescence, and confirmed integrants were archived for subsequent engineering steps. For sequential integration experiments, orthogonal serine integrases (Bxb1, TG1, and R4) were paired with distinct antibiotic resistance markers and reporters (sfGFP, mRFP, and mTagBFP2) to enable iterative selection without marker curing, as described in the section “Sequential integration of multiple payloads”. Marker recycling was performed using ϕ C31 integrase-mediated excision and *sacB* counterselection, as described in the section “Marker recycling via ϕ C31 integrase and *sacB* counterselection”.

Electrocompetent cell preparation. Electrocompetent cells were prepared as previously described for *R. palustris*²⁶, with the same protocol applied to *R. sphaeroides* which is comparable to optimized approaches developed for this species³⁹. Briefly, cells were grown aerobically to mid-to-late log phase (OD₆₀₀ 0.5-0.9), harvested, and washed three times in ice-cold 10% glycerol before resuspension at 1:400 the original culture volume. Electroporation was performed in 0.1-cm cuvettes with ~100 ng plasmid DNA for both the *attP*/cargo plasmid and serine recombinase plasmid at 1.75 kV, 25 μ F, and 200 Ω . Cells were recovered for 6 h at 30°C with shaking prior to plating on selective medium.

Serine recombinase screening assays

To assess integrase performance across hosts, we used a standard tri-parental conjugation protocol to deliver both the non-replicating poly-*attP* target plasmid and the corresponding

non-replicating integrase expression plasmid into recipient strains (Supplementary Table 2). All conjugations were carried out using the *E. coli* donor strain WM6026, which carries the *pir* genes necessary for R6K replication and is auxotrophic for diaminopimelic acid (DAP) to allow for counterselection post-mating. Donor WM6026 cells were grown overnight in LB with 0.1 mM DAP and appropriate antibiotics. Separately, PNSB recipients were cultured in rich medium (e.g., 3xYPS-MOPS for *R. palustris*, LB for *R. sphaeroides*) for 2-5 days until mid- to late-log phase. For tri-parental conjugations, WM6026 carrying the integrase plasmid and WM6026 carrying the poly-*attP* target plasmid were each washed and mixed with the recipient strain at equal optical densities ($OD_{600} = 1.0$) in a 1:1:1 ratio. The cell mixtures were centrifuged, resuspended in ~50 μ L of fresh media, and spotted onto non-selective agar plates supplemented with 0.1 mM DAP and incubated overnight (~16 h) to allow for conjugative DNA transfer and recombination. After overnight incubation at 30°C, conjugation mixtures were harvested, washed three times in fresh medium to remove residual DAP, and serially diluted for plating onto selective agar (lacking DAP) containing the appropriate antibiotic for selection of a single-*attP* or poly-*attP* target plasmid. Control recipient-only plates were used to quantify conjugation efficiency as transconjugants per viable recipient. Colonies were counted after 3–5 days of incubation at 30°C. Integration accuracy was assessed by colony PCR using a primer flanking the intended *attB* site and a single primer internal to the integrated target plasmid (Supplementary Table 5, Supplementary Fig. 10). In *R. palustris* CGA009 and P4, *R. rubrum*, and *R. gelatinosus*, kanamycin selection (poly-*attP* plasmid pSA29) was used to quantify conjugation efficiency following delivery of the payload and recombinase plasmids. However, in *R. sphaeroides*, kanamycin selection proved unreliable, likely due to antibiotic marker incompatibility, so we substituted spectinomycin (poly-*attP* plasmid pAR038) as the selectable marker unless otherwise indicated (Supplementary Fig. 4). Conjugation efficiencies were determined as the number of colonies on selective media (transconjugants) per number of viable recipient cells.

Tn7 benchmarking study. To benchmark cSAGE against a standard Tn7-based delivery approach under matched conditions, we performed tri-parental conjugation using the mini-Tn7 helper plasmid pJMP1039 (R6K-based Tn7 transposase) and the *attTn7* cargo plasmid pCK383 (J23111::mRFP, spectinomycin resistance marker) (Supplementary Table 2), following the tri-parental mating protocol described above for recombinase screening, with *R. sphaeroides* as the recipient. Mating mixtures were plated on selective medium (spectinomycin) across a dilution series (10^{-1} to 10^{-7}), and viable recipient cell density was determined in parallel by plating the same OD-normalized recipient culture on non-selective medium. No transconjugant colonies were recovered at any dilution across two independent conjugations, yielding a calculated detection limit of $\leq 10^{-10}$ transconjugants per recipient CFU (Supplementary Fig. 2).

Reporter fluorescence quantification

Single colonies were inoculated in a 2 ml 96-DWP containing rich broth media with antibiotic selection (500 μ L media). Plates were incubated at 30°C on a plate shaker for 24-48 h to enable sufficient outgrowth and fluorescent reporter expression. Optical density (OD_{660}) and fluorescence were measured on a BioTek Synergy H1 plate reader and compared to WT

controls to ascertain RFU/OD₆₆₀. The following plate reader settings were used, unless otherwise noted: sfGFP: 485nm/528nm, mTagBFP2: 400nm/455nm, and mRFP 540nm/600nm.

Promoter activity assays

To benchmark promoter activity in *R. palustris* CGA009 and *R. sphaeroides*, Anderson series (<https://parts.igem.org/>) promoters were cloned upstream of an mRFP reporter and integrated into the genome at the SAGE landing pad (Supplementary Tables 2 and 3). Strains were cultivated under routine aerobic cultivation conditions at 30°C, and fluorescence was measured (Ex/Em: 540nm/600nm) and calculated as described above.

Marker recycling via ϕ C31 integrase and *sacB* counterselection

To enable antibiotic marker recycling following integration, we adapted a marker curing strategy using ϕ C31 integrase-mediated excision and *sacB*-based counterselection (Elmore et al. 2023). PNSB strains containing a genomically integrated poly-*attP* target plasmid (pMG97 or pJC009) were used as recipients for the ϕ C31 plasmid (pSG010). WM6026 donor strains harboring the non-replicating ϕ C31 integrase expression plasmid were prepared as described above and conjugated into these strains using di-parental mating. After mating, cells were plated onto agar containing sucrose (concentrations used for each strain are noted in Supplementary Table 4), excluding DAP, to select for marker loss and counterselect *E. coli*. Purple colonies appearing after 3–5 days of incubation at 30°C were screened by colony PCR using primers flanking the Bxb1 *attB* site (Supplementary Table 5). Successful marker curing was indicated by the loss of the antibiotic marker and plasmid backbone, while retention of the payload was verified by *mRFP* fluorescence.

Sequential integration of multiple payloads

To demonstrate multi-payload genome engineering using cSAGE, we constructed a three-plasmid toolkit in which each plasmid encodes a different fluorescent reporter (sfGFP, mTagBFP2, or mRFP), a unique integrase attachment site (Bxb1, R4, or TG1), and an orthogonal antibiotic resistance marker (Km^R, Sm^R, or Gm^R). Integration was performed sequentially in *R. palustris* CGA009 and *R. sphaeroides* using tri-parental conjugation as described above, starting first with mTagBFP2 (pJC001), then mRFP (pMG083 for CGA009 or pJC008 for *R. sphaeroides*), and lastly, sfGFP (pBD0132). After each integration round, transconjugants were screened for reporter expression and validated by colony PCR to confirm integration at each intended *attB* site (Supplementary Table 5). Confirmed single-copy integrants were subcultured and used as recipients for the subsequent integration. Each round of integration was performed using only the antibiotic marker relevant to the newly introduced plasmid, with no selection for previously integrated markers. Fluorescence was quantified for each strain to evaluate reporter expression and confirm stable maintenance of all three integrated payloads.

Genomic integration of BaPAD

p-VP producing strains were generated by tri-parental conjugation of pRC442 (J23110::BaPAD) and pCK1205 (Bxb1 helper plasmid) with either *R. palustris* CGA009 (JE4632), *R. palustris* P4

(sMG127) or *R. sphaeroides* 2.4.1 (AVS4) (Supplementary Table 1). pRC442 contains a codon-optimized phenolic acid decarboxylase gene from *Bacillus atrophaeus* (BaPAD), the Bxb1 *attP* site, and a kanamycin selection marker. Successful integrants were screened as described above.

Bioproduction of *p*-vinylphenol

Metabolites were quantified by preparing a standard curve in Siström's media and fitting it to a linear equation. Media containing *p*-coumarate (*p*-CA) was prepared as previously described⁸⁰. As a result of *p*-CA insolubility, sonication followed by filtration was used to make 10 mM *p*-CA in Siström's minimal media. *R. palustris* strains were fed exclusively *p*-CA. Subsequent bioproduction experiments were inoculated from a starting anaerobic culture. Cultures were incubated at 30°C with infrared light until reaching stationary phase (7 days), at which point they were sampled periodically. For quantitative analysis, 1 mL of cell cultures were extracted using a 1 mL syringe and centrifuged for 5 min at 20,000×g. Supernatants were loaded into 0.2 µm, 10 kDa PES polypropylene filter 96-well plates (Analytical, cat. no. 964014) and centrifuged for 20 min at 1,500×g. *p*-vinylphenol (*p*-VP) and *p*-CA were quantified using high-performance liquid chromatography (HPLC). Filtered supernatants were collected in HPLC vials with pierceable caps. Samples were analyzed by injecting 50 µL into an Agilent HPLC system equipped with a diode array detector set at 254 nm and 310 nm and separated using a ZORBAX RX-C18 reverse-phase column (4.6 mm × 150 mm, 5 µm particle size; Agilent Technologies). The mobile phases were (A) acetonitrile plus 0.1% trifluoroacetic acid (TFA), and (B) water plus 0.1% TFA. The flow was held at 1.000 mL/min. The mobile phase gradient follows: hold at 10% solvent A from 0.0 to 2.0 min, 10% to 50% solvent A from 2.0 to 12.0 min, hold at 50% solvent A from 12.0 to 14.0 min, 50% to 10% solvent A from 14.0 to 14.5 min, hold at 10% solvent A from 14.5 to 20.0 min. The retention times for the detected compounds were: 11.5 min (*p*-VP) and 7.1 min (*p*-CA). *p*-CA and *p*-VP concentrations were determined according to a standard curve of known concentrations (Supplementary Fig. 9).

Optical density measurements

Optical density (OD₆₆₀) of bacterial cultures in Hungate tubes was measured using a custom-built, non-invasive photometer adapted from the TubeOD design⁸¹. The device consists of a 3D-printed tube holder with an LED (596 nm) and a photosensor positioned on opposite sides of the culture tube to measure transmitted light (Supplementary Fig. 11). Light intensity was recorded as an analog voltage inversely proportional to cell density and logged using a LabJack U3-HV. Species-specific calibration curves were generated relative to a benchtop spectrophotometer and used to convert voltage to OD₆₆₀ values. Python scripts are located on the GitHub repository <https://github.com/carothersresearch/K-BOT>.

Genome-scale metabolic modeling

Flux values for the metabolic map of *R. palustris* CGA009 were calculated using COBRApy with a genome-scale model based on iAN1128, with additional reactions to represent the heterologous PAD enzyme and transport reactions for the product of the engineered pathway, *p*-vinylphenol. Simulations were performed under photoheterotrophic conditions with *p*-CA as

the sole carbon source. The code used to generate these flux values is available at <https://github.com/carothersresearch/cSAGE>.

Comparative genomic analysis

Alignments of the coumarate catabolic pathway genes between *R. palustris* CGA009 and P4 were performed using Mauve plugin in Geneious Prime v2026.0. We sequenced and assembled the P4 genome using Oxford Nanopore Sequencing serviced by Plasmidsaurus Inc. (with 100x coverage depth) and deposited it to GenBank under accession number CM151374.1. KEGG orthology (KO) analysis was performed using BLASTKoala for KO assignment. Restriction endonucleases were identified in the *R. sphaeroides* 2.4.1 genome assembly (GenBank accession number GCF_000012905) by scanning for the Gene Ontology (GO) term for endonuclease activity (GO:0004519). *RshI* was previously characterized³⁸.

Statistics and Reproducibility

Replicate numbers and experimental units are stated in each figure legend. Experiments used independent biological cultures or independent conjugation reactions as specified. For fluorescence and growth assays, values represent the mean of independent cultures with error bars indicating one standard deviation. For HPLC-based metabolite quantification, standard curves were prepared fresh for each experiment, and quantification was performed using at least three biological replicates. Integration efficiencies are reported as the mean proportion of colonies with correct integration, as determined by colony PCR.

Data availability

All data in this study are available from the corresponding authors upon request. The source data underlying Figs. 1c-d, 2c, 3, 4b-c, 6b-c and Supplementary Figs. 3-5, 7, 9, and 11 are provided as a Source Data file. Key plasmids from the cSAGE toolkit have been deposited to Addgene (accession numbers listed in Table 2). Additional plasmids generated in this study are available from the corresponding authors upon request.

Author contributions

†M.S.G. and C.K. contributed equally to this work. *J.C. and A.R. contributed equally to this work. M.S.G., C.K., R.L.C., D.A.B. and J.E. conceived and designed the study. M.S.G., C.K., J.C., A.R., A.S., R.L.C., D.A.B., M.C., S.G., S.A., B.D., G.N.D, K.H. and D.H. performed experiments and analyzed the data. M.S.G., J.M.C., J.Z., J.E., R.G.E., A.M.G., and A.B. supervised the project and guided data interpretation. A.B. and J.M.C. acquired funding. M.S.G., C.K., A.R., J.C., R.L.C., M.C., J.E., and J.M.C. wrote the manuscript with input from all authors.

Conflicts of interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: JGZ and JMC are advisors to Wayfinder Biosciences.

Acknowledgements

We thank: Carrie Harwood and Yasuhiro Oda at the University of Washington for providing essential strains, protocols, and reagents that enabled this work; Huimin Zhao at the University of Illinois Urbana-Champaign for sharing *E. coli* WM6026; Eric Schaedig at The National Renewable Energy Laboratory and Carrie Eckhert at Oak Ridge National Laboratory who provided *Rubrivivax gelatinosus* CBS, as well as guidance and protocols for working with this strain. This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0023091. This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Figures and Tables

Table 1. Bacteria used in this study.

Organism	Classification	Origin	Primary bioproduction use cases
<i>Rhodopseudomonas palustris</i> CGA009	Alphaproteobacteria	Soil	Aromatic compound valorization ^{54–56,82}
<i>Rhodopseudomonas palustris</i> P4	Alphaproteobacteria	Anaerobic digester	Fermentative CO oxidation and hydrogen production ^{43,60,61}
<i>Rhodobacter sphaeroides</i> 2.4.1	Alphaproteobacteria	Freshwater	Isoprenoid biosynthesis and biohydrogen production ^{6,7,13}
<i>Rhodospirillum rubrum</i> ATCC 11170	Alphaproteobacteria	Freshwater	Anaerobic syngas fermentation ⁹
<i>Rubrivivax gelatinosus</i> CBS	Betaproteobacteria	Soil	Photosynthetic fermentation of syngas and organic acids ⁸³
<i>Rhodococcus jostii</i> RHA1	Actinomycetes	Soil	Lignin depolymerization and conversion of aromatic compounds ^{49,75}
<i>Pseudomonas sorgoleonovorans</i> SO81	Gammaproteobacteria	Rhizosphere	Plant-growth promotion and utilization of plant-derived carbon sources ⁴⁸
<i>Pseudomonas putida</i> KT2440	Gammaproteobacteria	Soil	Metabolic engineering platform for aromatic compound valorization ^{46,47}

Table 2. cSAGE plasmid toolkit.

Plasmid	Addgene ID	Intended use and key features
pRC414	251503	Tn7-compatible landing pad integration; integrates poly- <i>attB</i> cassette at <i>attTn7</i>
pCK1205	251504	Integrase expression, Bxb1
pCK1206	251505	Integrase expression, R4
pCK1207	251506	Integrase expression, TG1
pSG010	251507	Integrase expression, ΦC31 (marker curing / backbone excision)
pMG065	251508	Payload integration, Bxb1 <i>attP</i>
pJC001	251509	Payload integration, R4 <i>attP</i>
pBD0132	251510	Payload integration, TG1 <i>attP</i>
pJC009	251511	Payload integration with counterselection, poly- <i>attP</i>
pRC453	253620	FLP recombinase, <i>sacB</i>
pAR038	257828	Payload integration, poly- <i>attP</i>
pAR079	257829	Payload integration, poly- <i>attP</i>

Table 3. Comparison of cSAGE to alternative genome engineering approaches.

Feature	Tn7-based systems	CRAGE	CRISPR/HDR	SAGE	cSAGE
<i>Conjugative delivery</i>	✓ ^{19,23}	✓ ²⁵	✓ ²⁰	✗ ²⁶	✓ (Fig. 2, 4)
<i>Iterative integrations</i>	✗ (single site) ^{19,23}	✗ capped at 2 sites (CRAGE-DUET) ²⁴	✓ unlimited (new construct required per edit) ²⁰	✓ (poly- <i>attB</i> landing pad) ²⁶	✓ (poly- <i>attB</i> landing pad, Fig. 4)
<i>Marker recycling</i>	✓ Flp-FRT ^{19,23}	✓ Cre- <i>lox</i> cassette exchange ^{24,25}	✓ ²⁰	✓ (φC31 excision) ²⁶	✓ (φC31 excision, conjugation-compatible, Fig. 5)
<i>Standardized toolkit across hosts</i>	✗ (<i>attTn7</i> only conserved in Gram-negatives) ^{19,23}	✓ (validated across multiple phyla) ²⁵	✗ (strain-specific HDR construct required) ²⁰	✓ (validated across multiple phyla) ²⁶	✓ (identical plasmids across two phyla; Table 1, Figs. 2-3)

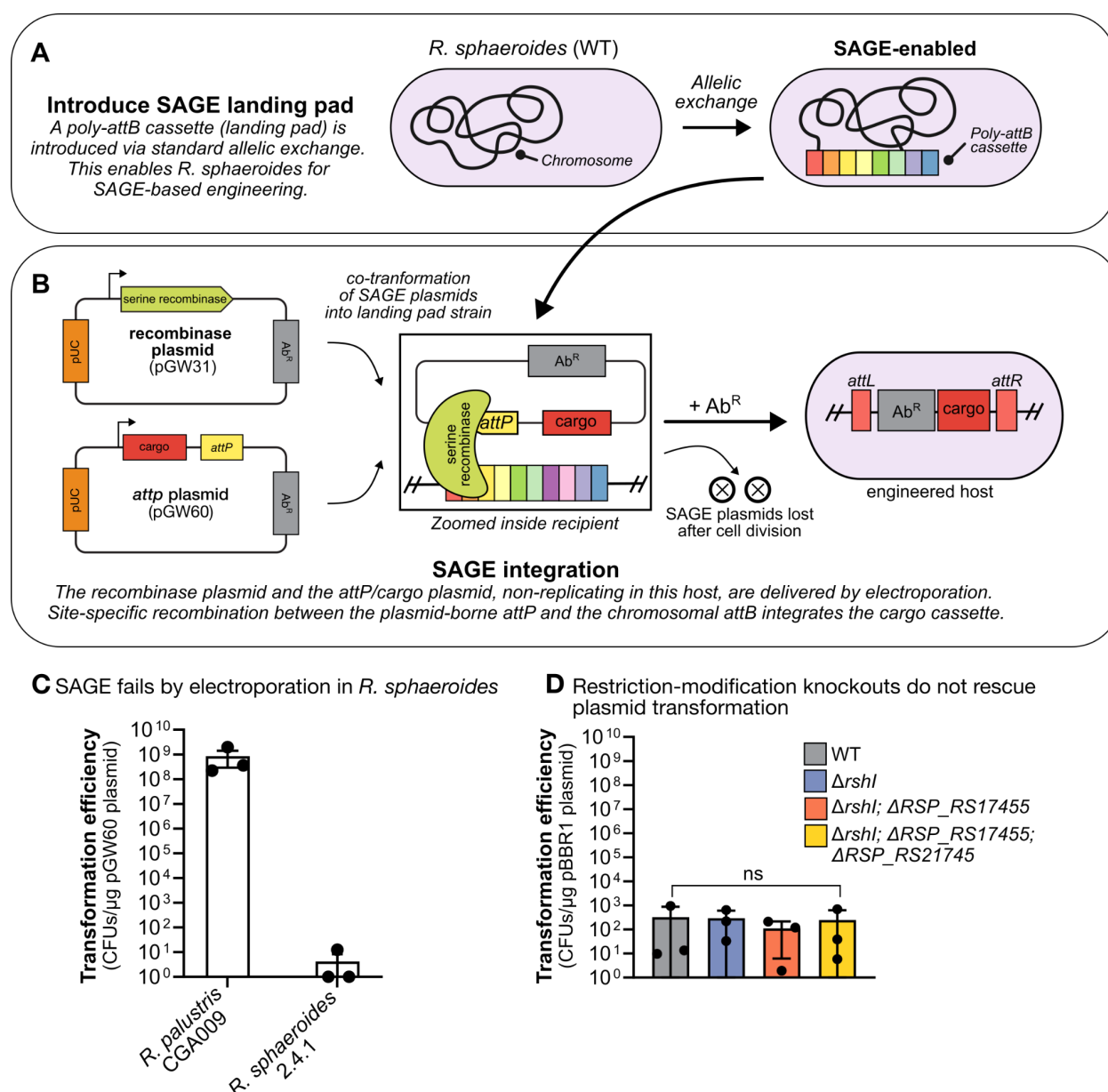
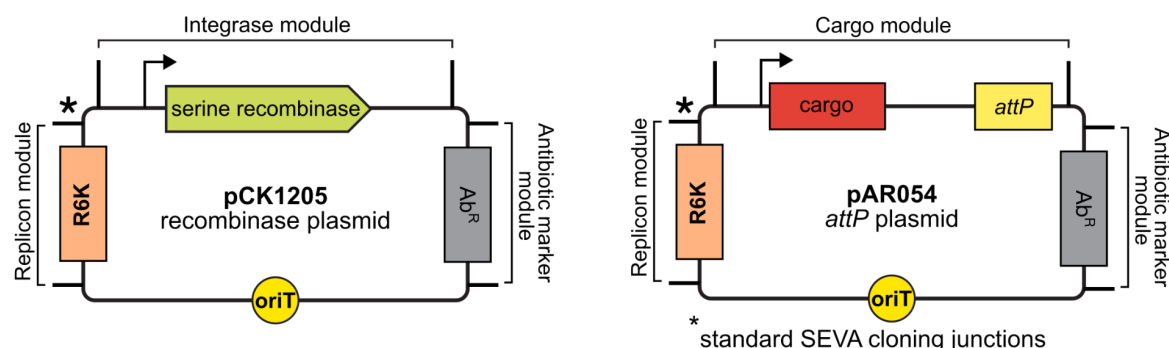


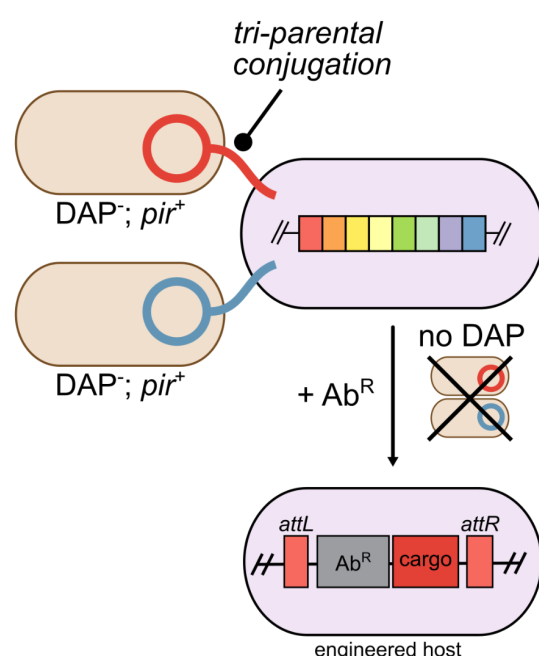
Figure 1. SAGE fails by electroporation in *Rhodobacter sphaeroides*. (A) Overview of landing pad installation. A poly-attB cassette containing multiple orthogonal attB sites is introduced into the *R. sphaeroides* chromosome at the *RshI* site by standard allelic exchange, generating a SAGE-enabled landing pad strain. (B) Overview of SAGE integration. The integrase plasmid (pGW31) carries a Bxb1 recombinase and pGW60 carries the cognate Bxb1 attP site. Both plasmids, non-replicating in this host, are co-transformed into the landing pad strain. The serine recombinase catalyzes site-specific recombination between the plasmid-borne attP and the chromosomal attB, integrating the cargo cassette and forming attL and attR junctions. The non-replicating plasmids are lost after cell division, leaving the engineered host with a single stably integrated cassette. (C) SAGE fails by electroporation in *R. sphaeroides*. Transformation efficiency (CFUs/ μ g pGW60 plasmid) following electroporation of *R. palustris* CGA009 (positive control) and *R. sphaeroides* 2.4.1 ($n = 3$; bars = mean $\pm$ s.d.). (D)

Restriction-modification knockouts do not rescue electroporation efficiency. Transformation efficiency (CFUs/ μ g pBBR1 plasmid) in *R. sphaeroides* WT and restriction endonuclease knockout strains tested using the replicating plasmid pBBR1 (pCK524) ($n = 3$; bars = mean $\pm$ s.d.). *RSP_RS17455* and *RSP_RS21745* encode putative restriction endonucleases. The *P* values were determined by one-way ANOVA followed by a pairwise test with Bonferroni adjustment (ns, not significant).

A Vector modifications enabling conjugation-based SAGE



B Conjugation-based SAGE (cSAGE) workflow



C cSAGE enables efficient integration in both electroporation-resistant and -competent hosts

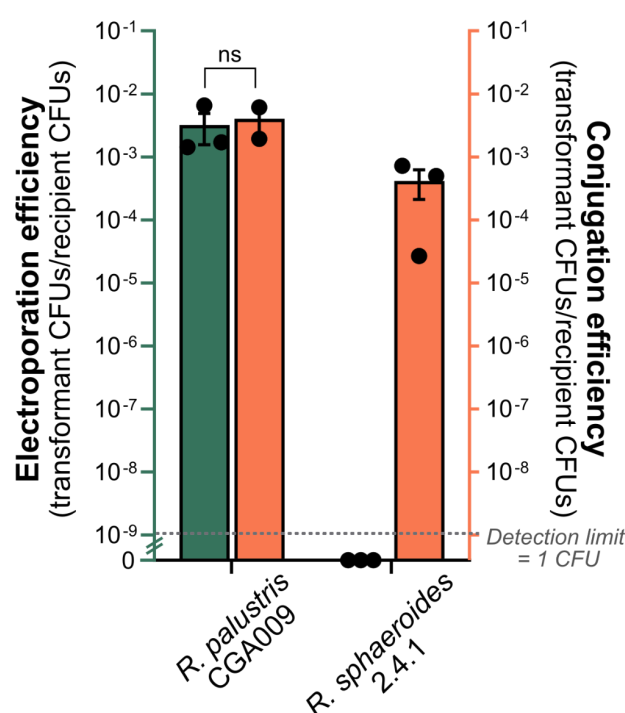
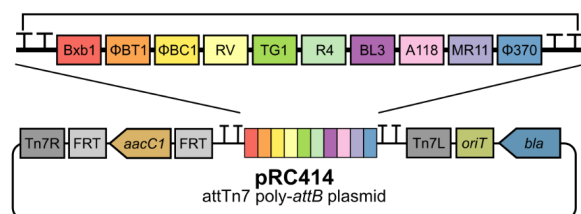


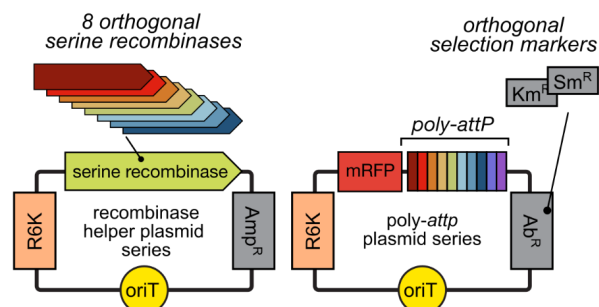
Figure 2. Conjugation-based SAGE (cSAGE) resolves the engineering bottleneck in *Rhodobacter sphaeroides*. (A) Vector modifications enabling cSAGE. SAGE plasmids were redesigned with a modular SEVA architecture, with each module flanked by standard SEVA cloning junctions (*). The integrase plasmid (left, pCK1205) and attP/cargo plasmid (right, pAR054) were made conjugation-compatible by addition of an RK2/RP4 origin of transfer (oriT). The replicon was changed to R6K, a low-copy origin that requires *pir* in trans for replication. (B) cSAGE workflow. Two DAP-auxotrophic *E. coli* donor strains, one carrying the recombinate plasmid and the other the attP/cargo plasmid, are mated with a SAGE-enabled recipient strain by tri-parental conjugation, transferring each plasmid via the mating pilus. Mating mixtures are then plated on medium lacking DAP and containing the antibiotic corresponding to the cargo plasmid's resistance marker. This selects against donor growth, since the donors are DAP-auxotrophic, and for recipient cells with a stably integrated cargo cassette, since the non-replicating cSAGE plasmids are lost after division. (C) cSAGE enables efficient integration

in both electroporation-resistant and -competent hosts. Electroporation efficiency (left, y-axis) and conjugation efficiency (right, y-axis) are shown for R6K-based cSAGE vectors in *R. palustris* CGA009 and *R. sphaeroides* 2.4.1. Conjugation efficiency and electroporation efficiencies are expressed as transconjugant (transformant) CFUs per total recipient CFUs ($n \geq 2$; bars = mean $\pm$ s.d.). Electroporation efficiencies are normalized to input DNA concentration. The *P* values were determined by Welch's *t*-test (ns, not significant). Source data is provided as a Source Data File.

A Poly-*attB* (landing pad) integration Tn7 plasmid



B Maps of cSAGE plasmids used for efficiency experiments



C Integration efficiency and accuracy across species

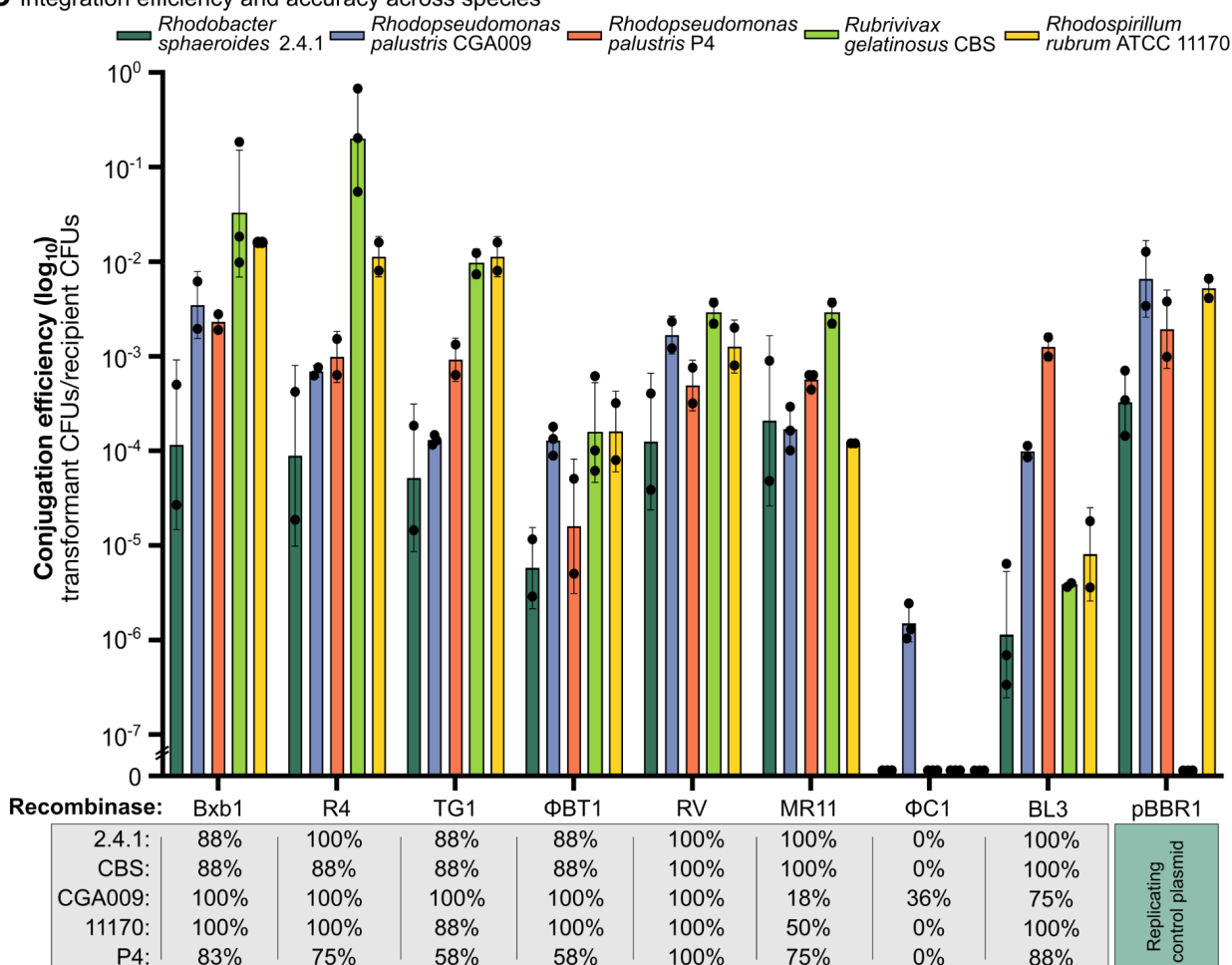
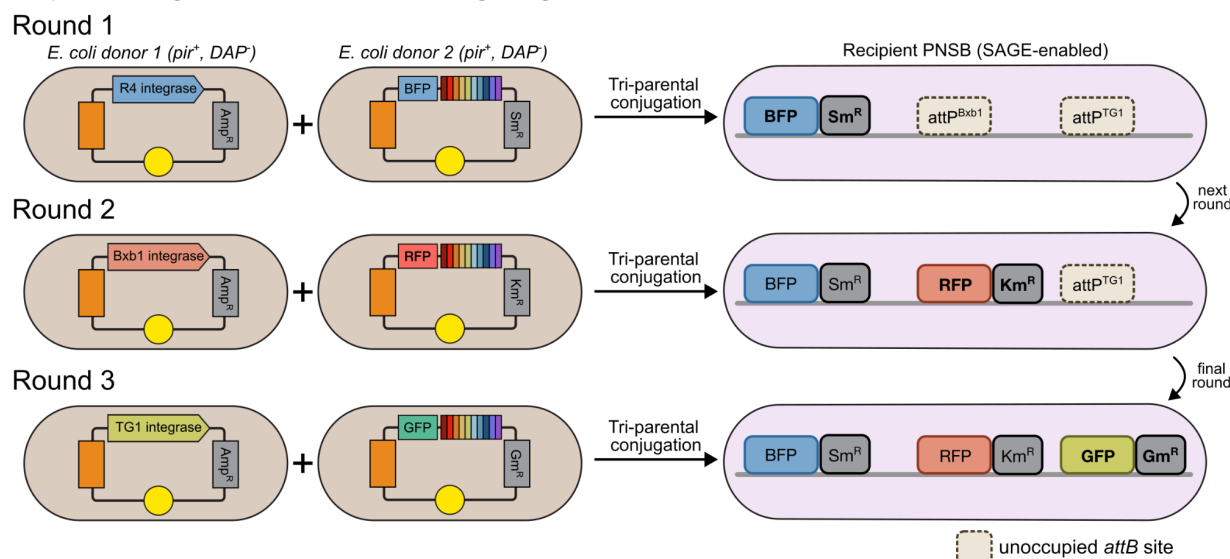


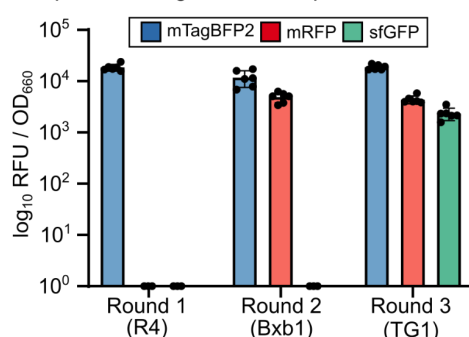
Figure 3. cSAGE recombinase benchmarking across diverse hosts. (A) Design of the Tn7-based poly-*attB* landing pad delivery plasmid. The genomic landing pad consists of ten orthogonal serine recombinase attachment sites (*attB*), flanked by transcriptional terminators to prevent transcriptional readthrough and interference with neighboring loci. A Tn7 delivery vector (pRC414) was constructed to integrate the poly-*attB* cassette into Gram-negative bacterial genomes. The *aacC1* gentamicin resistance marker is flanked by FRT sites for optional removal via Fip recombinase. (B) Maps of recombinase and poly-*attP* plasmids used for recombinase screening. Integration was tested by tri-parental conjugation, delivering two non-replicating plasmids: (i) a poly-*attP* cargo plasmid carrying a constitutively expressed mRFP reporter and one of eight orthogonal *attP* sites, and (ii) an integrase expression plasmid encoding the

cognate serine recombinase (Bxb1, ϕ BT1, ϕ C1, RV, TG1, R4, BL3, or MR11). Two additional landing pad sites, A118 and ϕ 370, are included for future expansion but were not tested in this screen. **(C)** Recombinase activity and fidelity across purple nonsulfur bacteria (PNSB). Conjugation efficiencies (number of transconjugant colonies per recipient CFU; $\log_{10}$ scale) are shown for eight recombinases tested in *Rhodobacter sphaeroides* 2.4.1, *Rhodopseudomonas palustris* CGA009, *R. palustris* P4, *Rubrivivax gelatinosus* CBS, and *Rhodospirillum rubrum* ATCC 11170. A replicating control plasmid (pBBR1) was used to benchmark conjugation efficiency ($n \geq 2$; bars = mean $\pm$ s.d.). The table below summarizes integration accuracy, defined as the proportion of colonies with correct on-target integration confirmed by colony PCR ($n \geq 8$; bars = mean $\pm$ s.d.). Source data is provided as a Source Data File.

A Sequential integrations without marker curing using cSAGE



B Sequential integrations: *R. sphaeroides* 2.4.1



C Sequential integrations: *R. palustris* CGA009

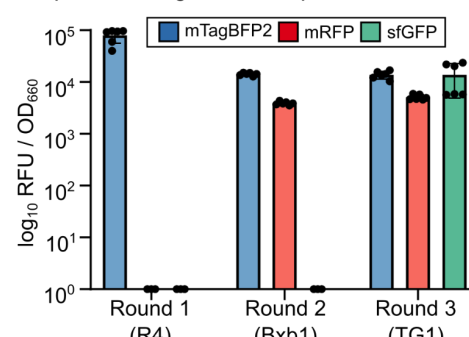
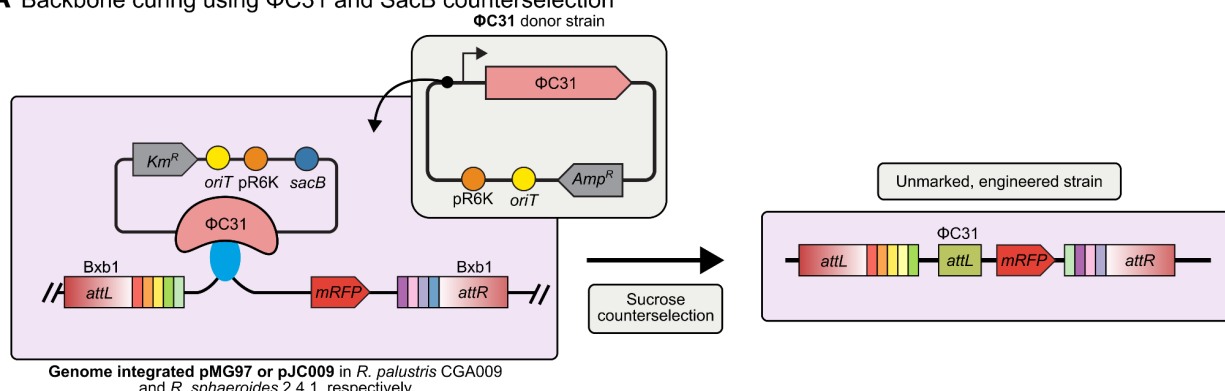
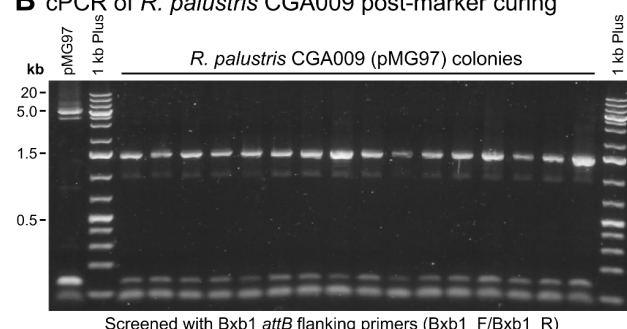


Figure 4. Sequential multi-payload integration via cSAGE without marker curing. (A) Three rounds of tri-parental conjugation deliver orthogonal fluorescent reporter cassettes sequentially into a SAGE-enabled PNSB recipient. In each round, two *E. coli* donor strains (*pir*⁺, DAP⁻) carry a non-replicating integrase expression plasmid and a cognate non-replicating *attP*/cargo plasmid, respectively. Round 1 delivers R4 integrase and a BFP cargo with a spectinomycin resistance marker (Sm^R); Round 2 delivers Bxb1 integrase and an RFP cargo with kanamycin resistance (Km^R); Round 3 delivers TG1 integrase and a GFP cargo with gentamicin resistance (Gm^R). Each round selects only for the newly introduced antibiotic marker, with no selection for or excision of previously integrated cassettes. Unoccupied *attB* sites remaining in the landing pad after each round are shown as dashed boxes. (B-C) Fluorescence quantification after each integration round in *Rhodobacter sphaeroides* 2.4.1 (B) and *Rhodospseudomonas palustris* CGA009 (C). Fluorescence from all reporters present in the strain (newly integrated and previously integrated) was quantified at each round to assess expression and confirm absence of cassette dropout (*n* = 6; bars = mean ± s.d.).

A Backbone curing using Φ C31 and SacB counterselection



B cPCR of *R. palustris* CGA009 post-marker curing



C cPCR of *R. sphaeroides* 2.4.1 post-marker curing

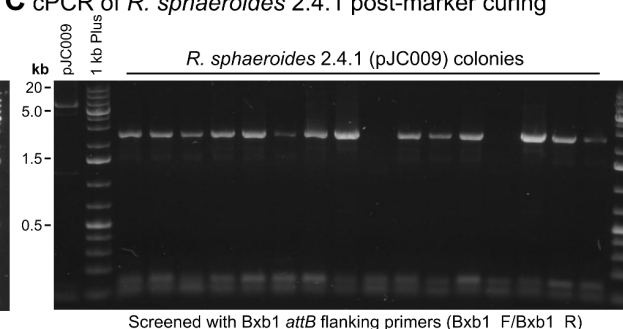
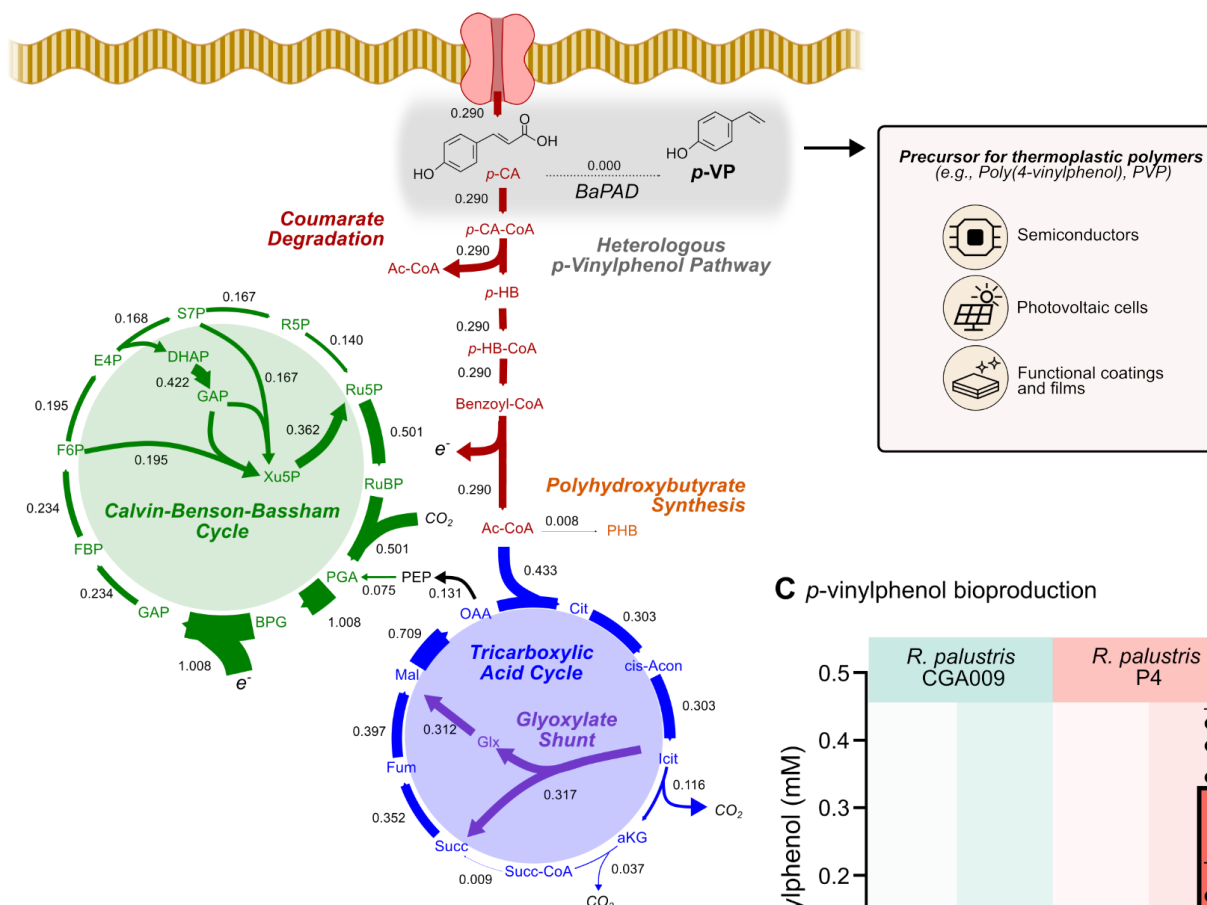
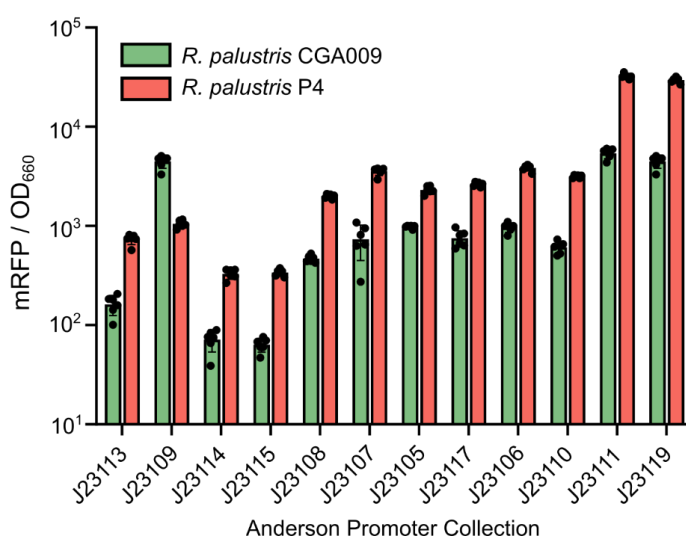


Figure 5. Marker recycling in cSAGE using Φ C31 integrase-mediated excision and sucrose counterselection. (A) Φ C31-based marker curing strategy. Genome-integrated poly-attP plasmids (pMG097 or pJC009) containing an *mRFP* reporter and a *sacB* counterselection cassette were flanked by Φ C31 att sites and delivered to *Rhodopseudomonas palustris* CGA009 or *Rhodobacter sphaeroides* 2.4.1. A non-replicating Φ C31 integrase expression plasmid (pSG010) was subsequently introduced by conjugation to catalyze site-specific recombination between the Φ C31 att sites, excising the plasmid backbone and antibiotic resistance marker. The resulting markerless strain retains only the integrated cargo and minimal attL/attR scars. Following Φ C31-mediated recombination, sucrose counterselection was applied to remove residual cells retaining the *sacB* cassette/backbone. (B-C) Verification of marker excision. Colony PCR using primers flanking the Bxb1 attB integration site were used to distinguish between uncured (larger amplicon) and cured (smaller amplicon) strains. Expected band sizes: pMG097, 5125 bp (uncured) and 1461 bp (cured); pJC009, 6400 bp (uncured) and 2313 bp (cured). Thermo Scientific GeneRuler 1KB Plus DNA Ladder was used as a reference.

A Genome-scale model for *R. palustris* CGA009



B Promoter activity screening



C p-vinylphenol bioproduction

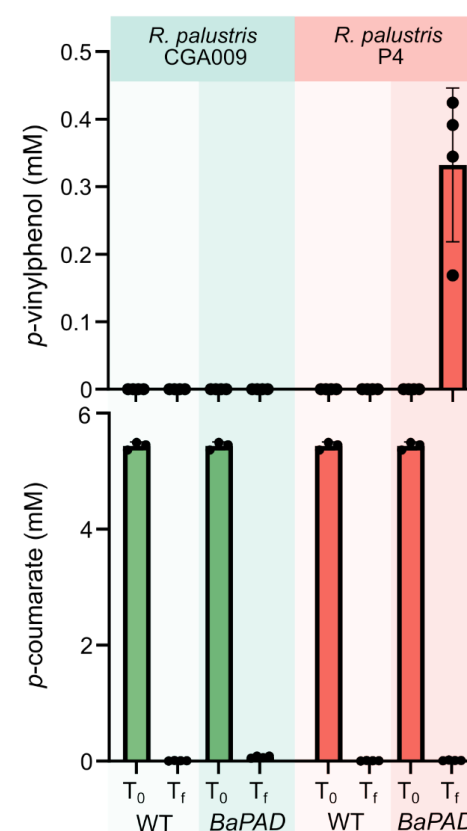


Figure 6. Photosynthetic bioproduction of p-vinylphenol in engineered PNSB. (A) Simulations of an iAN1128-based genome-scale model of engineered phenolic acid decarboxylase (PAD)-expressing *R. palustris* CGA009 grown photoheterotrophically on p-coumarate (p-CA). **(B)** Promoter activity across hosts. Fluorescence of integrated promoter

constructs in *R. palustris* P4 and CGA009. Data show relative fluorescence units (RFU) normalized to optical density (OD₆₆₀) for replicate colonies from transformation plates ($n = 6$; bars = mean $\pm$ s.d.). **(C)** *p*-vinylphenol production and *p*-CA consumption in *R. palustris* P4 and CGA009. Metabolite concentrations (*p*-VP and *p*-CA) were measured at inoculation (T_0) and after cultivation (T_f) for both wild-type (WT) and engineered (BaPAD) strains. *p*-VP is normalized to biomass at the time of sampling ($n \geq 3$; bars = mean $\pm$ s.d.). Source data is provided as a Source Data File.

References

1. Garza Elizondo, A. M., del Valle Kessra, I., Teixeira Prates, E., Komp, E., Phillips, E. K., Ashok, N., Jacobson, D. A., Webb, E. G., Bomble, Y. J., Alexander, W. G., Tannous, J., Tsai, C.-J., Parrott, W. A., Yang, X., Urbanowicz, B. R., Bartley, L. E., Maranas, C. D., Tuskan, G. A., Guss, A. M. & Eckert, C. A. Building an expanded bio-based economy through synthetic biology. *Biotechnol. Adv.* **87**, 108775 (2026).
2. Luo, K., Pang, Y., Yang, Q., Wang, D., Li, X., Lei, M. & Huang, Q. A critical review of volatile fatty acids produced from waste activated sludge: enhanced strategies and its applications. *Environ. Sci. Pollut. Res.* **26**, 13984–13998 (2019).
3. Wang, F., Ouyang, D., Zhou, Z., Page, S. J., Liu, D. & Zhao, X. Lignocellulosic biomass as sustainable feedstock and materials for power generation and energy storage. *J. Energy Chem.* **57**, 247–280 (2021).
4. Teixeira, L. V., Moutinho, L. F. & Romão-Dumaresq, A. S. Gas fermentation of C1 feedstocks: commercialization status and future prospects. *Biofuels Bioprod. Biorefining* **12**, 1103–1117 (2018).
5. Morrison, H. M. & Bose, A. Purple non-sulfur bacteria for biotechnological applications. *J. Ind. Microbiol. Biotechnol.* **52**, kuae052 (2025).
6. Brown, B., Wilkins, M. & Saha, R. *Rhodopseudomonas palustris*: A biotechnology chassis. *Biotechnol. Adv.* **60**, 108001 (2022).
7. Orsi, E., Beekwilder, J., Eggink, G., Kengen, S. W. M. & Weusthuis, R. A. The transition of *Rhodobacter sphaeroides* into a microbial cell factory. *Biotechnol. Bioeng.* **118**, 531–541 (2021).
8. Lu, H., Zhang, G., He, S., Zhao, R. & Zhu, D. Purple non-sulfur bacteria technology: a promising and potential approach for wastewater treatment and bioresources recovery. *World J. Microbiol. Biotechnol.* **37**, 161 (2021).
9. Revelles, O., Tarazona, N., García, J. L. & Prieto, M. A. Carbon roadmap from syngas to polyhydroxyalkanoates in *Rhodospirillum rubrum*. *Environ. Microbiol.* **18**, 708–720 (2016).
10. Kathol, M., Chowdhury, N. B., Immethun, C., Alsiyabi, A., Morris, D., Naldrett, M. J. & Saha, R. High enzyme promiscuity in lignin degradation mechanisms in *Rhodopseudomonas palustris* CGA009. *Appl. Environ. Microbiol.* **0**, e00573-25 (2025).
11. Imam, S., Yilmaz, S., Sohmen, U., Gorzalski, A. S., Reed, J. L., Noguera, D. R. & Donohue, T. J. IRsp1095: A genome-scale reconstruction of the *Rhodobacter sphaeroides* metabolic network. *BMC Syst. Biol.* **5**, (2011).
12. Cannon, W. R., King, E., Huening, K. A. & North, J. A. Redox poise in *R. rubrum* phototrophic growth drives large-scale changes in macromolecular pathways. *PLOS Comput. Biol.* **21**, e1013015 (2025).
13. Immethun, C. M., Kathol, M., Changa, T. & Saha, R. Synthetic Biology Tool Development Advances Predictable Gene Expression in the Metabolically Versatile Soil Bacterium *Rhodopseudomonas palustris*. *Front. Bioeng. Biotechnol.* **10**, (2022).
14. Freed, E., Fenster, J., Smolinski, S. L., Walker, J., Henard, C. A., Gill, R. & Eckert, C. A. Building a genome engineering toolbox in nonmodel prokaryotic microbes. *Biotechnol. Bioeng.* **115**, 2120–2138 (2018).
15. Riley, L. A. & Guss, A. M. Approaches to genetic tool development for rapid domestication of non-model microorganisms. *Biotechnol. Biofuels* **14**, 30 (2021).
16. Yan, Q. & Fong, S. S. Challenges and Advances for Genetic Engineering of Non-model Bacteria and Uses in Consolidated Bioprocessing. *Front. Microbiol.* **8**, 2060 (2017).
17. Choi, K.-H., Gaynor, J. B., White, K. G., Lopez, C., Bosio, C. M., Karkhoff-Schweizer, R. R. & Schweizer, H. P. A Tn7-based broad-range bacterial cloning and expression system. *Nat. Methods* **2**, 443–448 (2005).

18. Corts, A., Thomason, L. C., Costantino, N. & Court, D. L. Recombineering in Non-Model Bacteria. *Curr. Protoc.* **2**, e605 (2022).
19. Hall, A. N., Hall, B. W., Kinney, K. J., Olsen, G. G., Banta, A. B., Noguera, D. R., Donohue, T. J. & Peters, J. M. Tools for genetic engineering and gene expression control in *Novosphingobium aromaticivorans* and *Rhodobacter sphaeroides*. *Appl. Environ. Microbiol.* **0**, e00348-24 (2024).
20. Mougiakos, I., Orsi, E., Ghiffary, M. R., Post, W., de Maria, A., Adiego-Perez, B., Kengen, S. W. M., Weusthuis, R. A. & van der Oost, J. Efficient Cas9-based genome editing of *Rhodobacter sphaeroides* for metabolic engineering. *Microb. Cell Factories* **18**, (2019).
21. Gilbert, C., Brumwell, S. L., Crits-Christoph, A., Kang, S. C., Martin-Moldes, Z., Alsharif, W., Esmurria, A., Nguyen, M.-A., Lee, H. H. & Ostrov, N. A scalable framework for high-throughput identification of functional origins of replication in non-model bacteria. 2023.05.19.541510 Preprint at <https://doi.org/10.1101/2023.05.19.541510> (2023).
22. Rohmat, R. K. W., Irvine, T. C. T., Joshi, S. H.-N., Bailey, A. M., Jenkins, C., Ulaeto, D., Buscaill, P. & Gorochowski, T. E. Emerging trends in genome integration tools for precision engineering of diverse bacterial species. *Synth. Biol.* **10**, ysaf019 (2025).
23. Peters, J. M., Koo, B.-M., Patino, R., Heussler, G. E., Hearne, C. C., Qu, J., Inclan, Y. F., Hawkins, J. S., Lu, C. H. S., Silvis, M. R., Harden, M. M., Osadnik, H., Peters, J. E., Engel, J. N., Dutton, R. J., Grossman, A. D., Gross, C. A. & Rosenberg, O. S. Enabling genetic analysis of diverse bacteria with Mobile-CRISPRi. *Nat. Microbiol.* **4**, 244–250 (2019).
24. Wang, B., Zhao, Z., Jabusch, L. K., Chiniquy, D. M., Ono, K., Conway, J. M., Zhang, Z., Wang, G., Robinson, D., Cheng, J.-F., Dangl, J. L., Northen, T. R. & Yoshikuni, Y. CRAGE-Duet Facilitates Modular Assembly of Biological Systems for Studying Plant–Microbe Interactions. *ACS Synth. Biol.* **9**, 2610–2615 (2020).
25. Wang, G., Zhao, Z., Ke, J., Engel, Y., Shi, Y.-M., Robinson, D., Bingol, K., Zhang, Z., Bowen, B., Louie, K., Wang, B., Evans, R., Miyamoto, Y., Cheng, K., Kosina, S., De Raad, M., Silva, L., Luhrs, A., Lubbe, A., Hoyt, D. W., Francavilla, C., Otani, H., Deutsch, S., Washton, N. M., Rubin, E. M., Mouncey, N. J., Visel, A., Northen, T., Cheng, J.-F. & Yoshikuni, Y. CRAGE enables rapid activation of biosynthetic gene clusters in undomesticated bacteria. *Nat. Microbiol.* **4**, 2498–2510 (2019).
26. Elmore, J. R., Dexter, G. N., Baldino, H., Huenemann, J. D., Francis, R., Peabody, G. L. 5th, Martinez-Baird, J., Riley, L. A., Simmons, T., Coleman-Derr, D., Guss, A. M. & Egbert, R. G. High-Throughput Genetic Engineering of Nonmodel and Undomesticated Bacteria via Iterative Site-Specific Genome Integration. <https://www.science.org> (2023).
27. Merrick, C. A., Zhao, J. & Rosser, S. J. Serine Integrases: Advancing Synthetic Biology. *ACS Synth. Biol.* **7**, 299–310 (2018).
28. Brumwell, S. L., Lord, A., Corts, A. D., Harken, S., Crits-Christoph, A., Nguyen, M.-A., Leung, J., Mandler, K., Esmurria, A., Gilbert, C., Ling, L., Lee, H. H. & Ostrov, N. Active learning guides automated discovery of DNA delivery via electroporation for non-model microbes. 2025.11.18.689155 Preprint at <https://doi.org/10.1101/2025.11.18.689155> (2025).
29. Tominaga, Y., Ohshiro, T. & Suzuki, H. Conjugative plasmid transfer from *Escherichia coli* is a versatile approach for genetic transformation of thermophilic *Bacillus* and *Geobacillus* species. *Extremophiles* **20**, 375–381 (2016).
30. Grohmann, E., Muth, G. & Espinosa, M. Conjugative Plasmid Transfer in Gram-Positive Bacteria. *Microbiol. Mol. Biol. Rev.* **67**, 277–301 (2003).
31. Davison, J. Genetic Exchange between Bacteria in the Environment. *Plasmid* **42**, 73–91 (1999).

32. Samuel, B., Mittelman, K., Croitoru, S. Y., Ben Haim, M. & Burstein, D. Diverse anti-defence systems are encoded in the leading region of plasmids. *Nature* **635**, 186–192 (2024).
33. Dimitriu, T., Szczelkun, M. D. & Westra, E. R. Various plasmid strategies limit the effect of bacterial restriction–modification systems against conjugation. *Nucleic Acids Res.* **52**, 12976–12986 (2024).
34. Purdy, D., O'Keeffe, T. A. T., Elmore, M., Herbert, M., McLeod, A., Bokori-Brown, M., Ostrowski, A. & Minton, N. P. Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier. *Mol. Microbiol.* **46**, 439–452 (2002).
35. Martínez-García, E., Fraile, S., Algar, E., Aparicio, T., Velázquez, E., Calles, B., Tas, H., Blázquez, B., Martín, B., Prieto, C., Sánchez-Sampedro, L., Nørholm, M. H. H., Volke, D. C., Wirth, N. T., Dvořák, P., Alejandre, L., Grozinger, L., Crowther, M., Goñi-Moreno, A., Nikel, P. I., Nogales, J. & de Lorenzo, V. SEVA 4.0: an update of the Standard European Vector Architecture database for advanced analysis and programming of bacterial phenotypes. *Nucleic Acids Res.* **51**, D1558–D1567 (2023).
36. Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., de las Heras, A., Páez-Espino, A. D., Durante-Rodríguez, G., Kim, J., Nikel, P. I., Platero, R. & de Lorenzo, V. The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* **41**, D666–D675 (2013).
37. Rakowski, S. A. & Filutowicz, M. Plasmid R6K replication control. *Plasmid* **69**, 231–242 (2013).
38. Jun, D., Saer, R. G., Madden, J. D. & Beatty, J. T. Use of new strains of *Rhodobacter sphaeroides* and a modified simple culture medium to increase yield and facilitate purification of the reaction centre. *Photosynth. Res.* **120**, 197–205 (2014).
39. Lee, Y. R., Lee, J., Hong, S., Lee, S. Y., Lee, W.-H., Koh, M., Chang, I. S. & Lee, S. Optimization of Electroporation Conditions for Introducing Heterologous DNA into *Rhodobacter sphaeroides*. *J. Microbiol. Biotechnol.* **34**, 1–6 (2024).
40. Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M. 2nd & Peterson, K. M. Four New Derivatives of the Broad-Host-Range Cloning Vector pBBR1MCS, Carrying Different Antibiotic-Resistance Cassettes. *Gene* vol. **166** 175–176 (1995).
41. Metcalf, W. W., Jiang, W. & Wanner, B. L. Use of the rep technique for allele replacement to construct new *Escherichia coli* hosts for maintenance of R6K λ origin plasmids at different copy numbers. *Gene* **138**, 1–7 (1994).
42. Blodgett, J. A. V., Thomas, P. M., Li, G., Velasquez, J. E., van der Donk, W. A., Kelleher, N. L. & Metcalf, W. W. Unusual transformations in the biosynthesis of the antibiotic phosphinothricin tripeptide. *Nat. Chem. Biol.* **3**, 480–485 (2007).
43. Jung, G. Y., Jung, H. O., Kim, J. R., Ahn, Y. & Park, S. Isolation and characterization of *Rhodospseudomonas palustris* P4 which utilizes CO with the production of H₂. *Biotechnol. Lett.* **21**, 525–529 (1999).
44. Hu, P., Lang, J., Wawrousek, K., Yu, J., Maness, P.-C. & Chen, J. Draft Genome Sequence of *Rubrivivax gelatinosus* CBS. *J. Bacteriol.* **194**, 3262 (2012).
45. Stellwagen, A. E. & Craig, N. L. Avoiding self: two Tn7-encoded proteins mediate target immunity in Tn7 transposition. *EMBO J.* **16**, 6823–6834 (1997).
46. Loeschcke, A. & Thies, S. *Pseudomonas putida*—a versatile host for the production of natural products. *Appl. Microbiol. Biotechnol.* **99**, 6197–6214 (2015).

47. Nikel, P. I., Chavarría, M., Danchin, A. & de Lorenzo, V. From dirt to industrial applications: *Pseudomonas putida* as a Synthetic Biology chassis for hosting harsh biochemical reactions. *Curr. Opin. Chem. Biol.* **34**, 20–29 (2016).
48. Sorgoleone degradation by sorghum-associated bacteria; an opportunity for enforcing plant growth promotion | bioRxiv.
<https://www.biorxiv.org/content/10.1101/2023.05.26.542311v1.abstract>.
49. Cappelletti, M., Presentato, A., Piacenza, E., Firrincieli, A., Turner, R. J. & Zannoni, D. Biotechnology of *Rhodococcus* for the production of valuable compounds. *Appl. Microbiol. Biotechnol.* **104**, 8567–8594 (2020).
50. Lammens, E.-M., Nikel, P. I. & Lavigne, R. Exploring the synthetic biology potential of bacteriophages for engineering non-model bacteria. *Nat. Commun.* **11**, 5294 (2020).
51. Kiattisewee, C. How many plasmids can bacteria carry? A synthetic biology perspective. *Open Biol.* **15**, 240378 (2025).
52. Ledermann, R., Strebel, S., Kampik, C. & Fischer, H.-M. Versatile Vectors for Efficient Mutagenesis of *Bradyrhizobium diazoefficiens* and Other Alphaproteobacteria. *Appl. Environ. Microbiol.* **82**, 2791–2799 (2016).
53. Silo-Suh, L. A., Elmore, B., Ohman, D. E. & Suh, S.-J. Isolation, characterization, and utilization of a temperature-sensitive allele of a *Pseudomonas* replicon. *J. Microbiol. Methods* **78**, 319–324 (2009).
54. Pan, C., Oda, Y., Lankford, P. K., Zhang, B., Samatova, N. F., Pelletier, D. A., Harwood, C. S. & Hettich, R. L. Characterization of Anaerobic Catabolism of p-Coumarate in *Rhodopseudomonas palustris* by Integrating Transcriptomics and Quantitative Proteomics*. *Mol. Cell. Proteomics* **7**, 938–948 (2008).
55. Austin, S., Kontur, W. S., Ulbrich, A., Oshlag, J. Z., Zhang, W., Higbee, A., Zhang, Y., Coon, J. J., Hodge, D. B., Donohue, T. J. & Noguera, D. R. Metabolism of Multiple Aromatic Compounds in Corn Stover Hydrolysate by *Rhodopseudomonas palustris*. *Environ. Sci. Technol.* **49**, 8914–8922 (2015).
56. Harwood, C. S. & Gibson, J. Anaerobic and aerobic metabolism of diverse aromatic compounds by the photosynthetic bacterium *Rhodopseudomonas palustris*. *Appl. Environ. Microbiol.* **54**, 712–717 (1988).
57. Hädicke, O., Grammel, H. & Klamt, S. Metabolic network modeling of redox balancing and biohydrogen production in purple nonsulfur bacteria. *BMC Syst. Biol.* **5**, (2011).
58. Navid, A., Jiao, Y., Wong, S. E. & Pett-Ridge, J. System-level analysis of metabolic trade-offs during anaerobic photoheterotrophic growth in *Rhodopseudomonas palustris*. *BMC Bioinformatics* **20**, 233 (2019).
59. Fang, X., Lloyd, C. J. & Palsson, B. O. Reconstructing organisms in silico: genome-scale models and their emerging applications. *Nat. Rev. Microbiol.* **18**, 731–743 (2020).
60. Oh, Y. Fermentative hydrogen production by a new chemoheterotrophic bacterium *Rhodopseudomonas palustris* P4. *Int. J. Hydrog. Energy* **27**, 1373–1379 (2002).
61. Oh, Y.-K., Seol, E.-H., Kim, M.-S. & Park, S. Photoproduction of hydrogen from acetate by a chemoheterotrophic bacterium *Rhodopseudomonas palustris* P4. *Int. J. Hydrog. Energy* **29**, 1115–1121 (2004).
62. Hirakawa, H., Schaefer, A. L., Greenberg, E. P. & Harwood, C. S. Anaerobic p-Coumarate Degradation by *Rhodopseudomonas palustris* and Identification of CouR, a MarR Repressor Protein That Binds p-Coumaroyl Coenzyme A. *J. Bacteriol.* **194**, 1960–1967 (2012).
63. Liu, R.-Y., Liu, Z.-H., Li, B.-Z. & Yuan, Y.-J. Biotransformation of lignin into 4-vinylphenol derivatives toward lignin valorization. *Green Chem.* **26**, 1770–1789 (2024).

64. Rodriguez, A., Meadows, J. A., Sun, N., Simmons, B. A. & Gladden, J. M. Evaluation of bacterial hosts for conversion of lignin-derived p-coumaric acid to 4-vinylphenol. *Microb. Cell Factories* **20**, 181 (2021).
65. Uppalapati, S., Kong, N., Norberg, O., Ramström, O. & Yan, M. Ionization of covalent immobilized poly(4-vinylphenol) monolayers measured by ellipsometry, QCM and SPR. *Appl. Surf. Sci.* **343**, 166–171 (2015).
66. Peshkova, S. & Li, K. Investigation of Poly(4-Vinylphenol) As a Wood Adhesive. *Wood Fiber Sci.* 41–48 (2003).
67. Zhang, G., Kim, C. & Choi, W. Poly(4-vinylphenol) as a new stable and metal-free sensitizer of titania for visible light photocatalysis through ligand-to-metal charge transfer process. *Catal. Today* **281**, 109–116 (2017).
68. Kiattisewee, C., Dong, C., Fontana, J., Sugianto, W., Peralta-Yahya, P., Carothers, J. M. & Zalatan, J. G. Portable bacterial CRISPR transcriptional activation enables metabolic engineering in *Pseudomonas putida*. *Metab. Eng.* **66**, 283–295 (2021).
69. Singh, S., Ghosh, P. & Hatfull, G. F. Attachment site selection and identity in Bxb1 serine integrase-mediated site-specific recombination. *PLoS Genet.* **9**, e1003490 (2013).
70. Durrant, M. G., Fanton, A., Tycko, J., Hinks, M., Chandrasekaran, S. S., Perry, N. T., Schaepe, J., Du, P. P., Lotfy, P., Bassik, M. C., Bintu, L., Bhatt, A. S. & Hsu, P. D. Systematic discovery of recombinases for efficient integration of large DNA sequences into the human genome. *Nat. Biotechnol.* **41**, 488–499 (2023).
71. McClain, H. M., Lowrey, L. C., Quinto, L. B., Torrance, E. L., Gagliano, T. R., Isaacs, F. J., Schoeniger, J. S., Williams, K. P. & Mageeney, C. M. Integrase-On-Demand: bioprospecting integrases for targeted genomic insertion of genetic cargo. *Nucleic Acids Res.* **54**, gkag106 (2026).
72. Salmon, R. C., Cliff, M. J., Rafferty, J. B. & Kelly, D. J. The CouPSTU and TarPQM Transporters in *Rhodopseudomonas palustris*: Redundant, Promiscuous Uptake Systems for Lignin-Derived Aromatic Substrates. *PLOS ONE* **8**, e59844 (2013).
73. Licandro-Seraut, H., Roussel, C., Perpetuini, G., Gervais, P. & Cavin, J.-F. Sensitivity to vinyl phenol derivatives produced by phenolic acid decarboxylase activity in *Escherichia coli* and several food-borne Gram-negative species. *Appl. Microbiol. Biotechnol.* **97**, 7853–7864 (2013).
74. Oda, Y., Nelson, W. C., Alexander, W. G., Nguyen, S., Egbert, R. G. & Harwood, C. S. A *Rhodopseudomonas* strain with a substantially smaller genome retains the core metabolic versatility of its genus. *Appl. Environ. Microbiol.* **91**, e02056-24 (2025).
75. Larkin, M. J., Kulakov, L. A. & Allen, C. C. Biodegradation and *Rhodococcus* – masters of catabolic versatility. *Curr. Opin. Biotechnol.* **16**, 282–290 (2005).
76. Burbano, D. A., Kiattisewee, C., Karanjia, A. V., Cardiff, R. A. L., Faulkner, I. D., Sugianto, W. & Carothers, J. M. CRISPR Tools for Engineering Prokaryotic Systems: Recent Advances and New Applications. *Annu. Rev. Chem. Biomol. Eng.* **15**, 389–430 (2024).
77. Elmore, J. R., Shrestha, R., Wilson, A., Van Fossen, E., Frank, A., Francis, R. M., Baldino, H., Stephenson, M., Gupta, B., Rivera, J. & Egbert, R. G. Reinforced CRISPR interference enables reliable multiplex gene repression in phylogenetically distant bacteria. 2025.04.14.648646 Preprint at <https://doi.org/10.1101/2025.04.14.648646> (2025).
78. Huenemann, J. D., Carroll, A. L., Elmore, J. R., Dexter, G. N., Mand, T. D., Klingeman, D. M., Alexander, W. G. & Guss, A. M. Combinatorial pathway engineering using multiplex Serine recombinase-Assisted Genome Engineering (mSAGE). 2025.07.11.664204 Preprint at <https://doi.org/10.1101/2025.07.11.664204> (2025).
79. Van Fossen, E. M., Stephenson, M., Frank, A., Akella, S., Wooldridge, R., Wilson, A., You, Y., Nakayasu, E., Egbert, R. & Elmore, J. Efficient genetic code expansion tools enable in

- vivo study of lysine acetylation in non-model bacteria. 2025.04.14.648644 Preprint at <https://doi.org/10.1101/2025.04.14.648644> (2025).
80. Fenster, J. A., Werner, A. Z., Tay, J. W., Gillen, M., Schirokauer, L., Hill, N. C., Watson, A., Ramirez, K. J., Johnson, C. W., Beckham, G. T., Cameron, J. C. & Eckert, C. A. Dynamic and single cell characterization of a CRISPR-interference toolset in *Pseudomonas putida* KT2440 for β -ketoadipate production from p-coumarate. *Metab. Eng. Commun.* **15**, e00204 (2022).
 81. Deutzmann, J. S., Callander, G., Gu, W., Müller, A. L., McCully, A. L., Ahn, J. K., Kracke, F. & Spormann, A. M. Low-Cost Clamp-On Photometers (ClampOD) and Tube Photometers (TubeOD) for Online Cell Density Determination. *Front. Microbiol.* **12**, (2022).
 82. Harwood, C. S. & Gibson, J. Uptake of benzoate by *Rhodopseudomonas palustris* grown anaerobically in light. *J. Bacteriol.* **165**, 504–509 (1986).
 83. Eckert, C. A., Freed, E., Wawrousek, K., Smolinski, S., Yu, J. & Maness, P.-C. Inactivation of the uptake hydrogenase in the purple non-sulfur photosynthetic bacterium *Rubrivivax gelatinosus* CBS enables a biological water–gas shift platform for H₂ production. *J. Ind. Microbiol. Biotechnol.* **46**, 993–1002 (2019).