Experimental capture of genomic islands defines a widespread class of genetic element capable of non-autonomous transfer

Yansong Zhao[1,*], Yue Ma[2], Andrew D. Farr[1], David W. Rogers[1] & Paul B. Rainey[1,3,*]

[1] Department of Microbial Population Biology, Max Planck Institute for Evolutionary Biology, Plön, Germany

[2] Guest Group Evolutionary Medicine, Max Planck Institute for Evolutionary Biology, Plön, Germany

[3] Laboratory of Biophysics and Evolution, CBI, ESPCI Paris, Université PSL, CNRS, Paris, France

* To whom correspondence may be addressed. Emails: Y.Z. (yansongzhao@evolbio.mpg.de), P.B.R. (rainey@evolbio.mpg.de).

Abstract

Bacteria acquire new genes by horizontal gene transfer (HGT). Acquisition is typically mediated by mobile genetic elements (MGEs), however, beyond plasmids, bacteriophages and certain integrative conjugative elements (ICEs), the nature and diversity of MGEs is poorly understood. The bacterium *Pseudomonas fluorescens* SBW25 was propagated by serial transfer in the presence of filtrate obtained from garden compost communities. Genome sequencing of derived colonies revealed acquisition of three different mobile elements, each integrated immediately downstream of *tmRNA*. All are flanked by direct repeats and harbour a tyrosine integrase (*intY*), followed by a cargo of accessory genes including putative phage defence systems. Although characteristic of genomic islands, MGE-classifiers showed no matches to mobile elements. Interrogation of DNA sequence databases showed that similar elements are widespread in the genus *Pseudomonas* and beyond, with *Vibrio* Pathogenicity Island-1 (VPI-1) from *V. cholerae* being a notable example. Bioinformatic analyses demonstrate frequent transfer among diverse hosts. With focus on a single 55 kb element (I55) we show that *intY* is necessary for excision and circularisation, that the element is incapable of autonomous horizontal transfer, but is mobilizable – in the absence of direct cell-cell contact – upon addition of community filtrate. Further analyses demonstrate that I55 enhances host fitness in the presence of community filtrate, which stems in part from ability to defend against phages.

Significance Statement

The impact of horizontal gene transfer on the evolution of bacteria outpaces that driven by spontaneous mutation, but knowledge of the range of mediators, the genes mobilised, and the mechanisms of movement have largely depended on inferences stemming from bioinformatics. Here we describe a real-time evolution experiment in which a single focal strain propagated with filtrate from a complex microbial community captures genetic elements carrying a diverse cargo of genes whose mobility was previously unknown. The elements are representative of a class of mobile DNA that depend on nothing

other than a tyrosine integrase that targets a highly conserved gene. The elements encode phage resistance, require the integrase to excise from the genome, confer a significant fitness advantage, and depend on a secondary element for transfer.

Key Words

community ecology – comparative genomics – experimental evolution – lateral transduction

Introduction

There is growing awareness of the role of mobile genetic elements (MGEs), such as phages, plasmids, and integrative conjugative elements (ICEs), in shaping the structure and function of microbial communities (Arnold et al., 2022). While encoding their own replicative capacity, MGEs oftentimes capture and move genes of ecological relevance among hosts (Arnold et al., 2022; Dmitrijeva et al., 2024; Quistad et al., 2020) by horizontal gene transfer (HGT). Recently, it has even been suggested that accessory genes define a set of "community genes" that could allow selection to operate on communities by promoting the efficacy of feedback cycles that underpin differential persistence (Doolittle & Inkpen, 2018; Lenton et al., 2021; Rainey & Quistad, 2020; Roughgarden, 2020).

While classical genetics has led to discovery of various mechanisms of HGT, including conjugation (Lederberg & Tatum, 1946), transduction (Zinder & Lederberg, 1952) and transformation (Griffith, 1928), the majority of recent discoveries stem from genome analysis (Dmitrijeva et al., 2024; Groussin et al., 2021; Hackl et al., 2023) . For example, analysis of closely related strains often reveals regions with distinct evolutionary histories – often referred to as "genomic islands" (GIs) – that are then marked as having likely arisen by HGT (Koutsovoulos et al., 2022; Yuan et al., 2023). While instructive, such bioinformatic methods rarely provide insight into the causes of horizontal movement. Notable exceptions include demonstration of the mobility of ICEs (Colombi et al., 2017; J. T. Sullivan & Ronson, 1998), of plasmids (Hall et al., 2016; Lilley & Bailey, 1997b), of gene transfer agents (Lang et al., 2012; Marrs, 1974), and of phage-inducible chromosomal islands (PICIs) (Barcia-Cruz et al., 2024; Fillol-Salom et al., 2019; Lindsay et al., 1998; Novick et al., 2010). The PICIs, a class of phage satellite (Ibarra-Chávez et al., 2021) , hitchhike with, or hijack, helper bacteriophages to facilitate their own intercellular transfer.

Experimental evolution, where populations or communities are propagated under controlled conditions (Lenski, 2017; Rainey et al., 2017; Rainey & Travisano, 1998), while not typically used to study HGT, offer new possibilities to understand the range of genes disseminated by MGEs, the mediators of transfer, and functional consequences (Frazão et al., 2019; Hall et al., 2016; Lilley & Bailey, 1997a; Quistad et al., 2020; Woods et al., 2020). For example, Quistad et al. (Quistad et al., 2020), focusing on complex microbial communities founded from garden compost, employed an experimental design in which the evolutionary life of MGEs was either enhanced, or curtailed. Comparative metagenomic analyses revealed a significant flux of horizontally disseminated DNA, including putative transfer agents (Quistad et al., 2020; van Dijk et al., 2023), but in the absence of culturable hosts carrying newly acquired DNA – underpinned by further analyses of fitness and genetics – understanding is limited to inferences stemming from bioinformatic analyses.

In this study, we use experimental evolution to detect, isolate and characterise novel MGEs from environmental sources. To achieve this, we simplified the experimental strategy of Quistad *et al.* (Quistad et al., 2020) to one in which a single bacterium, *Pseudomonas fluorescens* SBW25 (Fortmann-Grote et al., 2023; Silby et al., 2009), was propagated in the presence of filtrate derived from complex compost communities. This led to capture of three GIs that define a class of MGE unclassified by current identification tools. The elements integrate immediately downstream of *tmRNA*, are diverse in terms of size and gene content, with many encoding putative phage defence systems, but also metabolic genes, and many with no matches to sequences in databases. Bioinformatic analyses based on a minimal set of core genes demonstrate the elements to be representative of a widespread class of GI whose mobility

87 has not been previously demonstrated. Detailed analysis of a single element shows that the integrase is
88 necessary for spontaneous excision from the genome and concomitant circularisation, but intracellular
89 transfer occurs only in the presence of community filtrate. The element confers a fitness benefit in the
90 presence of community filtrate, but is weakly deleterious in pure culture. Further analysis shows that
91 the fitness benefit stems in part from enhanced resistance to bacteriophages present in the community
92 filtrate.

93

Results

Pseudomonas fluorescens SBW25 propagated in the presence of compost filtrate acquired mobile genetic elements

A simple mark-recapture approach (Fig. 1a(ii)) was used to see whether SBW25 introduced into garden compost communities might acquire MGEs. To this end, SBW25 was marked with gentamycin resistance (strain MPB28806), placed in static microcosms together with wash from compost and propagated under a regime involving transfer of a 1-in-100 aliquot every 48 hours. Samples were obtained at regular intervals, however marked SBW25 was not recovered after ~5 transfers (see Fig. S1). This most likely reflected inability of SBW25 to persist in the face of competition imposed by community members. We next propagated SBW25 (marked with either GFP or RFP, strains MPB29284 and MPB29300, respectively) with filtrate derived from the community, using two slightly different protocols referred to as "live-filtrate" and "fixed-filtrate" regimes (see Fig. 1a(i) and 1a(iii), respectively). Choice of the two regimes was motivated by concerns of potential negative effects of storing samples at -80 °C, but also the benefits stemming from having archived samples. In both cases, filtrate was obtained by passing community culture through a 0.2 µm filter, thus eliminating all cells with the exception of nanobacteria (Quistad et al., 2020). For the live-filtrate regime, material was obtained from a single synchronised parallel community (compost C) that was also subject to transfer every 48 hours. On each occasion 1 mL of filtrate was obtained with 50 µL being added to each of eight independent serially-propagated replicate populations of SBW25 (Fig. 1a(i)). Populations were transferred 14 times. For the fixed-filtrate regime, a single sample of filtrate was obtained from each of three independent, seven-day-old compost communities (composts B, C and D) and stored at -80 °C. An aliquot of 50 µL from the archived sample was added at each serial transfer (Fig. 1a(iii)). Ten replicate SBW25 populations received filtrate from each of the three independent communities. Populations were transferred seven times.

At the end of each experiment, samples of derived SBW25 were plated on agar. For the live-filtrate regime, two to three colonies were selected from each of the eight replicate lines and re-sequenced. For the fixed-filtrate regime, 24 colonies were chosen at random from each replicate, grown separately overnight and then pooled, with DNA extracted from the pooled sample being re-sequenced. Mapping DNA sequence reads back to the ancestral SBW25 genome sequence revealed the presence of DNA sequences that did not match the genome (Fig. 1b). Further analysis of these reads led to identification of three GIs that had been acquired during the course of the selection experiment. The elements, designated I23, I44 and I55 (approximately 23 kb, 44 kb and 55 kb in length, respectively), were identified on five independent occasions (see Fig. 1c). From the 19 colonies sequenced from the live-filtrate regime, one contained I55 (strain MPB33357), which was acquired from compost C. Of the 720 colonies sequenced from the fixed-filtrate regime, six contained I23 (strains MPB38392, MPB38393, MPB38394, MPB38395, MPB38396, MPB38895) and two contained I44 (strains MPB38906, MPB38907). I23 was acquired independently on three occasions from compost B; I44 was detected in just a single replicate receiving filtrate from compost B.

The three GIs are unrecognised by MGE-classifiers

Mapping the ends of contigs containing each element confirmed that all three GIs inserted at the same genomic locus, immediately downstream of *tmRNA* and immediately upstream of a putative mobilisable genomic island PFI-12 (see Fig. 1d). This was confirmed by PCR analysis. Comparison of the elements shows a basic level of similarity. Upon insertion, each of the newly acquired elements is flanked by a 14 bp direct repeat (CCC CGG CTC CAC CA), which defines the extent of each element and maps perfectly to

137 the second half of the *attP* sequence in *Pseudomonas aeruginosa* PAO1 (Fiedoruk et al., 2020). One of
138 the 14 bp sequences is located in the genome of SBW25 and spans the last 11 bp of *tmRNA* and 3 bp
139 immediately downstream. The second (a duplication) is introduced by the MGE upon insertion. Each
140 element harbours a gene encoding a tyrosine integrase (IntY) near the 5'-end and a gene encoding AlpA,
141 a putative transcriptional activator of *intY* (Barcia-Cruz et al., 2024; Wang et al., 2009), at the 3'-end (see
142 Fig. 1e). It is likely that IntY catalyses site-specific recombination of the MGE into the genome, with the
143 14 bp sequence being the common core O sequence of *attB* and *attP* sites (Grindley et al., 2006; Groth
144 & Calos, 2004). The integrases each show ~20% divergence at the nucleotide level; *alpA* in I44 and I55
145 are identical, but ~20% dissimilar to *alpA* from I23. Recombination thus impacts evolution of the
146 elements.

147 Beyond the duplicated integration site and presence of an IntY plus AlpA, the three elements share a
148 ~18 kb region of sequence similarity at the 3'-end, although I44 and I55 contain gaps (see Fig. 1e for
149 alignment, and Table S4 for detailed annotations). The region encodes YqaJ (I23-*orf06*, I44-*orf23*, I55-
150 *orf28*), an HTH-containing protein (I23-*orf11*, I44-*orf28*, I55-*orf32*), inovirus Gp2 (I23-*orf19*, I44-*orf31*,
151 I55-*orf38*), AlpA (I23-*orf20*, I44-*orf32*, I55-*orf39*), along with many open reading frames predicted to
152 encode proteins with no matches to either DNA or protein sequence databases. In this region elements
153 I23 and I55 encode a putative DNA methyltransferase (I23-*orf18* and I55-*orf37*), which is predicted to be
154 part of a type II restriction-modification (RM) system (Tock & Dryden, 2005) with possible phage defence
155 activity. However, the endonuclease component of the RM system is not present, thus leaving in doubt
156 roles in phage defence.

157 Between *intY* and the region of similarity – and accounting for length differences – are various cargo
158 genes that differ among the three MGEs. In the hyper-variable region, I23 encodes a known phage
159 defence system, the Gabija system (*orf02-orf03*) (Cheng et al., 2023; Doron et al., 2018). By contrast, I55
160 encodes dynamins (*orf02-orf03*) and RecBCD (*orf23-orf25*), as well as many proteins that have no
161 matches to databases. I44 encodes two putative phage defence systems, the classic type II restriction-
162 modification system (*orf02-orf03*) (Tock & Dryden, 2005) and the DISARM system (*orf05-orf10*) (Ofir et
163 al., 2017) . In addition, I44 encodes a ~14 kb operon (*orf11-orf20*) composed of metabolic genes
164 including putative hydantoinases (Ishikawa et al., 1997), a peptidase M24, and an Asp/Glu/hydantoin
165 racemase.

166 Given that the three elements were acquired during the course of serial transfer propagation
167 experiments, they must be capable of either autonomous, or non-autonomously transfer. A search of
168 encoded genes shows no evidence of those typically associated with lateral transmission, such as
169 machinery for conjugation, or type IV pilli. Additionally, they bear no resemblance to prophages and lack
170 hallmarks of phage satellites (de Sousa et al., 2023; Ibarra-Chávez et al., 2021) including capsid-encoding
171 genes, and *cos* sites. Moreover, interrogation of databases using recognised bioinformatic pipelines,
172 including ISEScan (screening for prokaryotic insertion sequences) (Xie & Tang, 2017), geNomad
173 (screening for viruses and plasmids) (Camargo et al., 2023), SatelliteFinder (screening for phage
174 satellites) (de Sousa et al., 2023) and MacSyFinder v2 (including tools TFFscan, TXSScan, CasFinder and
175 CONJScan, screening for Type IV filament systems, bacterial secretion systems, CRISPR-Cas systems and
176 conjugative systems, respectively) (Néron et al., 2023), showed no matches to any recognised mobile
177 element.

The three GIs represent a widespread and diverse class of element in *Pseudomonas*

Given knowledge of shared features of the three elements, we turned to whole genome sequences and asked whether *Pseudomonas* genomes harbour elements with these basic features. Search criteria included the presence of *intY* downstream of *tmRNA*, as well as the presence of an *alpA*, or a 23 bp sequence (TTC AGA TTA CCC CGG CTC CAC CA, the last 14 bp being the direct repeat) that is common to the 3'-end of each element. Searches revealed 99 elements from 17,078 *Pseudomonas* genomes. Fig. 2a shows the bioinformatic workflow and Fig. 2b depicts the relationship among the elements, constructed based on previously established methods (Low et al., 2019; Nayfach et al., 2021). Briefly, 36 ORFs that are commonly present in the entire set of MGEs were selected as core genes. For each element, the core genes were concatenated in order, with gaps allowed where core genes are absent. The tree was then inferred based on the concatenated sequences. The cladogram should be interpreted in a topological sense; the edges represent evolutionary relationships only (and not distances).

Overall, the elements are highly diverse in terms of size and gene content (Fig. 2b). The largest elements are in excess of 70 kb with the smallest being less than 20 kb. No two elements are identical; however several are highly similar. For example, the element detected in *P. rhodesiae* B21-046 (NCBI RefSeq Assembly GCF_024748755.1) is 99.96% similar at the nucleotide level to I55 acquired by SBW25. As described more fully below, these two host genomes are highly divergent. Other examples of such pairs are denoted by pink boxes in Fig. 2b.

Many elements contain homologues of RadC, YqaJ, inovirus Gp2 and HTH-containing protein (Fig. 2b, 2c). These are also present in I23, I44 and I55 but were not included in the search criteria. Together, these genes, in addition to *intY* and *alpA*, likely define core element functions.

Fig. 2c lists the rank abundance of genes that occur at least 10 times across the elements, along with functional classes and corresponding COG categories (Tatusov, 2000). In terms of variable genes, those encoding functions related to DNA replication, recombination and repair and transcription are frequently represented. Furthermore, most elements contain recognisable defence systems against bacteriophages (Fig. 2b). Not surprisingly, many genes are found as singletons with 104 of the 241 functional gene classes identified via matches to protein databases being represented just once (Table S6). Additionally, most elements harbour at least one recognisable ORFs with no matches to genes in databases (Fig. 2b).

To explore the relationship among elements and host genomes, a comparison of pairwise distances between MGEs and hosts was performed. This involved construction of phylogenetic trees for both MGEs and hosts (the latter based on 16S rDNA sequences) with the results represented in Fig. 2d. The elements shown on the X- and Y- axes are ordered based on clustering shown in Fig. 2b. The region above the diagonal represents similarity among elements, with the region below showing similarity among hosts. For example, with focus on comparison between I55 (in *P. fluorescens* SBW25) and the element in assembly GCF_024748755.1 (in *P. rhodesiae* B21-046), which is marked by a pink box: the square below the diagonal is coloured dark blue indicating that the two elements are highly similar. When the same comparison is made above the diagonal, the corresponding box is light blue indicating that the host genomes carry divergent 16S rDNA sequences. Similar patterns are evident throughout; dark blue regions below the diagonal very often correspond to divergent hosts. These incongruent patterns are indicative of the transfer of similar elements across a diverse range of host genomes as

219 expected for elements capable of frequent horizontal movement (Colombi et al., 2024; Koutsovoulos et
220 al., 2022).

221 I55 transfers intercellularly only in the presence of community filtrate

222 Knowing that the three captured elements represent a widespread, frequently-transferred, but poorly-
223 studied class of MGEs, we sought understanding of mechanisms of transfer and fitness consequences
224 (the latter is addressed in the next section) with attention focussed on I55.

225 By design of the serial transfer experiment, I55 is mobile (it evidently transferred from the compost-
226 derived filtrate into the SBW25 genome). To ask whether I55 is capable of autonomous transfer from
227 one SBW25 cell to another, I55 was tagged with a kanamycin resistance (Kan^R) marker (strain
228 MPB39465/MPB39467). The gene encoding resistance (including promoter and a synthesised
229 transcription terminator) was inserted in the intergenic region between two convergently transcribed
230 genes, *orf08* and *orf09*, on I55. This position was chosen in order to minimise chances that insertion
231 disrupted element function. SBW25 carrying a chromosomally-integrated tetracycline resistance (Tet^R)
232 gene (strain MPB39500) was used as the recipient. Donor and recipient strains were co-cultured in liquid
233 medium (without shaking as in the selection experiment, see Fig. 3a) and on solid agar plates (as per a
234 standard conjugation assay). After 48 hours of co-culture, mating mixtures were vortexed and aliquots
235 added to plates containing kanamycin and tetracycline. No colonies expressing both resistance markers
236 were detected. This suggests that I55 is incapable of autonomous intercellular movement.

237 In the original serial transfer experiment, filtrate from serially-propagated compost C had been added to
238 SBW25 cultures. We thus considered it possible that one or more components contained within the
239 filtrate might be necessary to mobilise the element. To this end, the experiment above was repeated in
240 liquid medium, but this time with the addition of 50 µL of filtrate from compost C (Fig. 3a). On plating
241 mixtures to selective agar an average ~10^4 colonies (per mL) resistant to kanamycin and tetracycline
242 were detected (Fig. 3b).

243 While the result strongly implicates a component of the community filtrate as mediator in the transfer
244 of I55 between bacterial cells, competing hypotheses exist. For example, the Tet^R gene in the recipient
245 could in principle reverse-transfer to the donor (Pei et al., 2024). Alternatively, I55 might spontaneously
246 transfer to another cell at a rate below detection limit, with transfer events being amplified and
247 detected only when community filtrate imposes selection for I55. A test of these hypotheses requires
248 repetition of the experiment with an engineered form of I55 incapable of transfer. We reasoned that
249 deletion of the integrase *intY* might prevent excision of I55 from the genome (Groth & Calos, 2004) and
250 thus curtail transfer. PCR, using primers flanking the insertion junctions of I55, was performed on
251 overnight cultures of SBW25-I55 (strain MPB39024/MPB39026). The results, shown and described in Fig.
252 S2, demonstrated that a fraction of I55 is in circular form, consistent with excision of the element from
253 the genome. The integrase was then deleted (strain MPB39132/MPB39136) and the PCR analysis
254 repeated. Results show that SBW25-I55Δ*intY* is incapable of excising from the genome and generating a
255 circular form.

256 The transfer experiment was repeated using SBW25-I55Δ*intY*-Kan^R (strain MPB39565/MPB39568), with
257 addition of culture filtrate from compost C. While on average ~10^2 colonies (per mL) resistant to both
258 antibiotics were detected, the number was significantly reduced ($p < 0.0001$) compared to the case
259 where SBW25-I55-Kan^R was the donor (Fig. 3b). This not only confirms that the occurrence of double-

260 resistance cells was not due to artifacts, but also implies that the circularisation of I55 underpins
261 intercellular transfer and / or integration.

262 Next, we asked whether donor-recipient cell-cell contact is necessary for the transfer of I55. To test this,
263 the transfer process was broken into two steps (Fig. 3c). In the first step, SBW25-I55-Kan^R with filtrate
264 from compost C was cultured for 48 hrs and the filtrate collected. In the second step, recipient SBW25-
265 Tet^R was grown with addition of the filtrate obtained from the first step. The mixture was plated as
266 previously. On average, more than 10^6 colonies (per mL) resistant to both antibiotics were detected (Fig.
267 3d), indicating that transfer of I55 is independent of cell-cell contact. To test the significance of the
268 tyrosine integrase under this two-step transfer regime, SBW25-I55Δ*intY*-Kan^R was used in place of
269 SBW25-I55-Kan^R, and the assay repeated. The density of double-resistant colonies decreased to ~10^4
270 colonies (per mL) (p = 0.0152), with this 100-fold difference consistent with the fold change in the one-
271 step transfer regime (Fig. 3b, 3d).

272 To ask which step in the transfer process the compost C filtrate is necessary, we repeated the two-step
273 transfer regime, but with compost C filtrate added to the culture at the second step. No double-resistant
274 colonies were detected (Fig. 3d), indicating that transfer of I55 did not occur (or if it did then it remained
275 below the detection limit). This suggests that in order for I55 to be transferred, the compost filtrate
276 must be co-cultured with the donor. Therefore, I55 appears to require some component in the
277 community filtrate that mediates transfer to recipient cells (Fig. 3e).

278 The identity of the transfer-facilitating component of compost filtrate is unknown, but it is reasonable to
279 suspect the involvement of phages. Five phage isolates were purified from the compost C filtrate and
280 individually added to the intercellular transfer assay (Fig. 3a) in place of the community filtrate, but no
281 double-resistant cells were detected and hence no evidence of transfer of I55. Additionally, community
282 filtrate from composts B, C and D were amplified and sequenced. This was done on the off-chance that
283 we might identify I23, I44 or I55 as a co-integrate with some other element. The sequences of I23 and
284 I55 were not covered with sufficient quality by reads of any of the samples. The sequence of I44,
285 however, was covered by reads from compost B filtrate, which is the filtrate in the selection experiment
286 from which isolates carrying I44 were detected. Approximately 250 reads that partially map to the left
287 or right ends of I44 sequence were extracted and then mapped to the circular version of I44 sequence.
288 All such reads matched the circularisation junction. This suggests that most of I44 in compost B filtrate
289 was in circular form, although rolling-circle replication during amplification of DNA could cause an over-
290 estimation of the proportion of circular DNA. There was no evidence that I44 in the filtrate formed a co-
291 integrant with some other element.

292 I55 confers a fitness advantage in the presence of community filtrate, partially due to defence 293 against phages

294 While knowledge of the mechanism(s) of I55 transfer are unclear, experiments to determine fitness
295 consequence are readily achievable. The fitness of SBW25 with and without I55 were determined by
296 direct competition taking advantage of fluorescent reporters allowing the two types to be distinguished
297 by flow cytometry (Fig. 4a). However, because I55 can transfer in the presence of community filtrate (a
298 potentially confounding factor), fitness effects were determined using SBW25-I55Δ*intY*. Moreover, to
299 ensure the two competing strains have identical genetic backgrounds, ancestral SBW25 (strain
300 MPB39379/MPB39459) was re-constructed by deleting I55 from SBW25-I55Δ*intY*. Compared to the wild
301 type, I55 confers a significant fitness advantage (~6%) in the presence of compost C filtrate (p = 0.0104),

302 but is slightly costly (~2%) in the absence of the filtrate (p = 0.0033) (Fig. 4b). Addition of community
303 filtrate significantly increased the fitness benefit of I55 (p = 0.0006).

304 I55 encodes a putative phage defence system (*orf37*), and the community filtrate is rich in phages
305 (Quistad et al., 2020). We therefore hypothesised that benefit of I55 stems from defence against
306 phages. This hypothesis was tested by determining the ability of SBW25 and SBW25-I55Δ*intY* to resist
307 attack by phages. Compost C filtrate was added to both hosts, with the number of plaque forming units
308 on SBW25-I55Δ*intY* being significantly reduced compared to SBW25 (p = 0.0080) (Fig. 4d). Moreover, the
309 plaques on lawns of SBW25-I55Δ*intY* appeared visually opaque compared to those on SBW25 (see Fig.
310 4c), suggesting that infection and cell lysis is less severe when cells carry I55. Together, this result
311 indicates that I55 confers non-perfect, but significant protection against phages present in the
312 community filtrate.

313 As the community filtrate stock is limited, we purified two phages (P1 and P2) from the plaque assay of
314 community filtrate. Phage P1 was isolated by picking a random plaque from the SBW25 lawn followed
315 by iterative purification on SBW25 as the host. Phage P2 was isolated in the same way except the host
316 for the initial plaque assay was SBW25-I55Δ*intY*. The fluorescent competition assay and plaque assay
317 were repeated with the two phage isolates replacing the community filtrate, and the assay results were
318 qualitatively reproduced (see Fig. 4e, 4f). This indicates that phage defence partially explains the fitness
319 benefit of I55.

Discussion

HGT mediated by MGEs shapes the diversity and function of microbial communities, but recent studies indicate that there is much left to discover concerning both the effectors of transfer and the range of functional genes subject to horizontal movement. Motivating the work reported here, was a previous study (Quistad et al., 2020) in which frequent exposure of pooled supernatant from compost communities to fresh hosts caused a substantive fraction of community DNA to partition into the pool of horizontally disseminated material. Comparative metagenomics provided evidence of genes of ecological relevance and in some instances phages, plasmids and other known MGEs were identified, but in the absence of ability to capture unknown elements in readily culturable hosts, progress is severely curtailed.

Here we simplified our previous strategy, retaining the complexity of culture filtrate from compost communities, but exposing the diversity of elements contained within the filtrate to a single focal, culturable and genetically tractable strain of *P. fluorescens.* The focal bacterium acquired three new MGEs in a short period of time, each integrated immediately downstream of *tmRNA*. The I23, I44 and I55 family of elements (hereafter the I55 family) lack characteristics of recognised MGEs, with a range of purpose-built bioinformatic pipelines (Camargo et al., 2023; de Sousa et al., 2023; Néron et al., 2023; Xie & Tang, 2017) failing to detect similarities. While tempting to suggest that the elements are phage satellites (de Sousa et al., 2023), absent are genes encoding primases, genes for viral packaging and *cos* sites, which are regarded as defining features of this class.

Bioinformatic searches for similar elements defined by minimal set of criteria including the target site, 3'-end sequence and integrase revealed ~100 within the genus *Pseudomonas*, however, the class is more diverse. Relaxation of the strict requirement for the flanking sequence (*attP*) to match the *tmRNA* sequence from *Pseudomonas* shows that elements with essential features of I55 are widespread. In fact, the broadest manifestation of the class appears to be that of GIs, many of which insert at *tmRNA* (or *tRNA*) genes, are flanked by direct repeats, encode tyrosine recombinases, but lack genes known to be required for autonomous transfer. The so-named "phage defense elements" recently described by Hussain et al. (Hussain et al., 2021) are a case in point. Islander, a bioinformatic tool developed to search for integrases downstream of *tmRNA* or *tRNA* genes (Hudson et al., 2015) shows presence of I55-like elements (GIs) across eubacteria and archaea.

A particularly notable example is the 41 kb *Vibrio* Pathogenicity Island-1 (VPI-1) that encodes the toxin-co-regulated pilus (TCP), which is the receptor for the cholera toxin-encoding phage CTXΦ. Acquisition of VPI-1, followed by entry of CTXΦ, converts environmental *V. cholerae* strains into toxigenic strains (Faruque & Mekalanos, 2003, 2012). While it is evident from comparative genomics that VPI-1 horizontally transferred (Karaolis et al., 1999; O'Shea & Boyd, 2002), attempts to demonstrate mobility have proven controversial (Faruque et al., 2003) with the current consensus being that VPI-1 is incapable of autonomous movement. Armed with knowledge learned from acquisition of members of the I55 family by SBW25, combined with studies of I55 transfer, and the fact that VPI-1 shares all the hallmarks that define this new class, we suggest that VPI-1 is a non-autonomous MGE of the I55 family, whose movement is facilitated by a secondary factor likely common in the environment naturally occupied by environmental *V. cholerae*. Perhaps this discovery ignites attempts to define the helper element, thus paving the way for development of strategies that curtail opportunity for the agent to facilitate transfer of VPI-1.

362 Capture of I55 (and family members) by SBW25 combined with genetic tractability of the focal
363 bacterium facilitated studies of the mechanism of transfer. While much remains to be discovered,
364 unequivocal demonstration that I55 is incapable of autonomous transfer is an important negative
365 finding. Indeed, this result led us to consider the possibility that a secondary element present in
366 compost filtrate may act as a helper. That transfer did occur in the presence of compost filtrate, was
367 largely dependent on IntY (necessary for excision and circularisation of I55), and did not require cell-to-
368 cell contact, is informative. Obvious secondary factors are phages, but purified phages have thus far
369 failed to mediate transfer. Nonetheless, we suggest that I55 family elements – including the huge
370 diversity of similar GIs suspected of movement, but for which direct evidence of transfer is lacking – are
371 likely to depend on the process of lateral transduction (Chen et al., 2018; Fillol-Salom et al., 2021;
372 Humphrey et al., 2021) mediated by *pac*-type phages that package DNA using the headful mechanism.

373 A benefit to working on I55 in a genetically tractable host has been understanding of fitness
374 consequence upon acquisition. That I55 confers a significant advantage in the presence of compost
375 filtrate (but a minor cost in its absence) offers an explanation for why it was possible to detected these
376 elements in SBW25 given sequence analysis of so few derived colonies. Demonstration of fitness
377 advantage in the presence of purified phages supports the hypothesis that the advantage stems in part
378 from benefits derived from phage defence systems. Interestingly, bioinformatic predictors (Tesson et al.,
379 2022) failed to detect complete phage defence systems in I55, however orthologs of dynamins (Guo et
380 al., 2022) and RecBCD (Dillingham & Kowalczykowski, 2008) are present (see Table S4). More generally,
381 this supports the view that MGEs are a rich source of novel phage defence systems (Hussain et al., 2021;
382 Johnson et al., 2023; Rousset et al., 2022).

383 Experimental detection of I55 family elements, and with these elements being instrumental in discovery
384 of a plethora of similar MGEs across *Pseudomonas* and beyond, likely defines just a small quota of the
385 yet-to-be discovered mediators of HGT. Repetition of the experiment with filtrates from different
386 environments, or different hosts, is likely to result in further breakthroughs. This gives substance to the
387 notion that microbial communities, replete with frequently moving MGEs (and associated genetic
388 cargoes), might be shaped by a process akin to sex as understood in eukaryotic populations (Rainey &
389 Quistad, 2020). To test this hypothesis, it will be necessary to go beyond knowledge that MGEs move
390 ecologically relevant DNA, to a more complete description of the nature and diversity of mobilising
391 elements, including factors affecting capacity to capture and recombine cargo genes. Equally important
392 will be development of new approaches to directly measure the flux of horizontally disseminated DNA.
393 Finally, simplicity of the strategy outlined here stands as a useful and general approach for discovery of
394 new MGEs, and the functional genes to which these elements become linked.

395

396 # Materials and Methods

397 ## Strains and growth conditions

398 All bacterial strains used in this study are listed in Table S1. All strains were preserved in -80 °C with 30%
399 glycerol saline.

400 *Isolation of foreign elements, plasmid construction, strain construction*: Unless otherwise stated, *P.*
401 *fluorescens* SBW25 strains were cultured in Luria broth (LB) at 28 °C, while *E. coli* strains were cultured in
402 LB (with antibiotics where maintenance of plasmids was required) at 37 °C, all shaking at 220 rpm for 16-
403 24 hrs.

404 *Experimental evolution, intercellular transfer assay and fluorescent fitness assay*: Unless otherwise
405 stated, each pre-conditioning and incubation was performed in the absence of shaking at 28 °C, in 5 mL
406 freshly made M9-glucose medium in a 15 mL Falcon tube. The M9-glucose medium was M9 salt solution
407 supplemented with 2 mM $MgSO_4$, 0.1 mM $CaCl_2$ and 0.2% w/v D-glucose.

408 *Plaque assay and bacteriophage isolation*: Unless otherwise stated, bacterial host strains were cultured
409 in S-LB medium (LB medium supplemented with 10 mM $MgCl_2$ and 5 mM $CaCl_2$) at 28 °C, shaking at 220
410 rpm for 16-24 hrs until stationary phase. Exponential phase bacterial culture was obtained by 1/25
411 diluting stationary phase culture in fresh S-LB medium and then shaking incubate at 28 °C, 220 rpm for 2
412 hrs. The semi-soft agar (SSA) used consisted of S-LB medium and 0.4% agar.

413 Bacterial culture was plated on respective growth medium supplemented with 1.5% agar. Unless
414 otherwise stated, agar plates were incubated at 28 °C. Antibiotics were applied as per the following
415 concentrations: kanamycin (Kan, 50 µg/mL), nitrofurantoin (100 µg/mL), tetracycline (Tet, 12.5 µg/mL),
416 *Pseudomonas* CFC selective supplement (Millipore, as per the manual).

417 ## Compost community sampling, storage and retrieval

418 Three heaps of garden compost, named B, C and D, were sampled in Plön, Germany (54°9'44"N,
419 10°25'17"E), from May to August 2022. Upon collection, the samples were immediately transferred to
420 laboratory and processed as following. For each sample, 30 g compost was mixed with 150 mL M9 salt
421 solution, shaken vigorously for ~30 sec, and left static for ~5 min. The supernatant was distributed into 1
422 mL aliquots, to each of which 350 µL glycerol saline was added. The aliquots were frozen in -80 °C for
423 storage.

424 To retrieve the compost wash, one aliquot from the frozen stock was kept on ice for ~10 min. After
425 thawing, the mixture was pelleted by centrifuging at 15000 rcf for 1 min, resuspended in 1 mL M9 salt
426 solution, and then pelleted by centrifuging again. The washed pellet was then resuspended in 500 µL M9
427 salts solution to inoculate microcosms.

428 ## Community filtrate (fixed) generation, storage and retrieval

429 Initially, 500 µL washed suspension from one compost wash aliquot was inoculated into 100 mL M9-
430 glucose medium in a 500 mL microcosm. The microcosm was incubated without shaking at 28 °C for 7
431 days. At the end of incubation period, the whole culture was filtered with a 0.2 µm filter (Whatman).
432 The filtrate was distributed into 600 µL aliquots. Aliquots were directly frozen at -80 °C for storage.

433 To retrieve the community filtrate, one aliquot from the frozen stock was kept at room temperature for
434 ~10 min. After thawing, the aliquot was directly used to inoculate microcosms. Unless otherwise stated,
435 50 µL community filtrate was added to each microcosm.

436 As a negative control, *P. fluorescens* SBW25 was used in place of the washed suspension, and otherwise
437 followed exactly the same generation, storage and retrieval protocols to obtain the SBW25 filtrate to
438 inoculate microcosms.

439 Experimental evolution

440 *Co-culture regime*

441 Parallel lines of SBW25, cocultured with the compost community, were serially propagated. In detail,
442 SBW25-RFP-Gen^R (strain MPB28806) and 90 µL washed suspension from compost C were inoculated.
443 There were three parallel lines, with one control line where no washed suspension was added. For each
444 line, an 1/20 transfer was made every 24 hrs, for a total of five transfers. At each transfer, a dilution
445 series of the culture was spread plated on LB + CFC + Gen plates, which were then cultivated for 24 hrs.
446 The number of RFP colonies were counted and recorded.

447 *Fixed filtrate regime*

448 Parallel lines of SBW25 were serially propagated, with aliquots of filtrate from the frozen stock added at
449 each transfer. In detail, filtrate from communities B, C and D, and from SBW25, were separately used.
450 For each type of filtrate, there were five parallel lines of SBW25-GFP-Kan^R (strain MPB29284) and five
451 parallel lines of SBW25-RFP-Kan^R (strain MPB29300). For each line, an 1/100 transfer was made every 48
452 hrs, for a total of seven transfers. At each transfer, 50 µL filtrate was added to a fresh microcosm, and
453 aliquots from the old microcosm were stored at -80 °C.

454 *Live filtrate regime*

455 Here, parallel lines of SBW25 were serial propagated, with aliquots of filtrate from independently
456 propagated communities or SBW25 added at each transfer. In detail, 200 µL washed suspension from
457 compost C wash was added and transferred 1/100 every 48 hrs, yielding fresh filtrate by filtering 1mL of
458 the culture through a 0.2 µm filter (Whatman). As a negative control, *P. fluorescens* SBW25 was used in
459 place of the washed suspension, but otherwise the same protocol was followed to obtain SBW25
460 filtrate. At the same time, for each type of filtrate, there were four parallel lines of SBW25-GFP-Kan^R
461 (strain MPB29284) and four parallel lines of SBW25-RFP-Kan^R (strain MPB29300). For each line, a 1/100
462 transfer was made every 48 hrs, for a total of 14 transfers. At each transfer, 50 µL fresh filtrate was
463 added to the new microcosm, and a sample from the old microcosm were stored at -80 °C.

464 PCR and gel electrophoresis

465 All PCR primers used in this study are listed in Table S3. PCR was done with NEB Q5 polymerase, with
466 either extracted DNA or bacterial suspension as templates. Annealing temperatures and extension time
467 was determined as per NEB Tm Calculator (version 1.16.6). Unless otherwise stated, 35 cycles were
468 performed. Gel electrophoresis was performed for 5 µL PCR products with 1x DNA loading dye, in 1%
469 agarose gels.

470 Detection, isolation and annotation of foreign elements

471 For each line, frozen culture at the last transfer was serially diluted and plated out on LB plates. For the
472 live-filtrate regime, 2 to 3 colonies were picked and separately cultured overnight, from which aliquots
473 were frozen and DNA extracted with Qiagen DNeasy UltraClean Microbial kit. DNA samples were

474 sequenced by Illumina NextSeq for at least 300x coverage. For fixed-filtrate regime, 24 colonies were
475 randomly picked and cultured up overnight in 300 µL LB in 96-deep-well plates. Aliquots from each of
476 the 24 cultures were mixed in equal proportion and then DNA extractions were performed with Qiagen
477 DNeasy UltraClean Microbial kit. The leftovers in 96-deep-well plates were stored at -80 °C. DNA
478 samples were Illumina short-read sequenced (read length 150 bp) by Eurofins Scientific SE for at least
479 300x coverage.

480 Bioinformatic analysis of sequencing data was done in Geneious Prime 2022.2.1
481 (https://www.geneious.com). For each sample, raw reads were trimmed by BBDuk and mapped to *P.*
482 *fluorescens* SBW25 reference genome (Fortmann-Grote et al., 2023) under the default settings. The
483 unmapped reads were saved and then *de novo* assembled. The vast majority of contigs were short (< 1
484 kb) and apparently artefacts of sequencing technology. However, in some samples, long (> 10 kb) and
485 good quality contigs were obtained. To test if the contig inserted into the genome as a foreign element,
486 the first 50 bp and last 50 bp of the contig were blasted against the reference genome. If the two
487 regions mapped, manual comparison were used to confirm the insertion and its locus.

488 To confirm the insertion of the foreign element, two primer pairs were designed, flanking either the left
489 or the right junction of the insertion. For fixed-filtrate regime, the two primer pairs were additionally
490 used to exclude the cultured isolates (out of the 24 picked) that carried the foreign element. Such
491 isolates were then routinely whole-genome sequenced by Eurofins Scientific SE and Variations / SNPs
492 calling was subsequently performed. Table S5 shows the list of mutations detected in each isolate.

493 The three foreign elements were screened for open reading frames (ORFs) by Prokka (Seemann, 2014) in
494 KBase (Arkin et al., 2018). The ORFs were then annotated by Pfam in InterProScan (Jones et al., 2014).
495 The elements were also screened by DefenseFinder (Tesson et al., 2022) for potential phage defense
496 systems (only HMM-only hits were retained). Alignment of the 3 foreign elements was done by Easyfig
497 (M. J. Sullivan et al., 2011) based on blastn under the default settings.

498 ## Detection and analysis of similar elements in databases

499 The nucleic acid sequences of the *tmRNA*, *intY* (from the element I55) and the 23 bp sequence at the 3'-
500 end were compared to representative genomes in GTDB-tk (version R214) (Chaumeil et al., 2020).
501 Sequence comparisons revealed that 35 representative genomes contain all three sequences, and they
502 all belong to the family *Pseudomonadaceae*. This motivated a more focused and thorough search within
503 this family. Therefore, the nucleic acid sequences of the *tmRNA*, *intY* (from the element I55) and the 23
504 bp sequence at the 3'-end were compared to all available genomic sequences (17,078 in total) of strains
505 belonging to the family *Pseudomonadaceae*. Sequence comparisons revealed that 136 genomes in the
506 family *Pseudomonadaceae* contain all three sequences. Those 136 genomes were then subject to gene
507 prediction by Prodigal (v2.6.3) (Hyatt et al., 2010) and functional annotation by eggNOG-mapper (v2)
508 (Cantalapiedra et al., 2021). Based on annotations, 99 genomes fulfilled the following criteria: (1)
509 contained an annotated tmRNA; (2) contained an annotated tyrosine integrase (IntY); (3) contained at
510 least one of an annotated AlpA or the 23 bp sequence. These 99 genomes are derived from 67 species,
511 all in the genus *Pseudomonas*. In each genome, the MGEs were isolated by extracting the sequence from
512 the 3'-end of the *tmRNA* to the 3'-end of the 23 bp sequence or the 3'-end of *alpA*. All MGEs were
513 screened by DefenseFinder (Tesson et al., 2022) for potential phage defense systems.

514 To establish evolutionary relationships of the 99 detected MGEs and I23, I44 and I55, a cladogram of the
515 102 MGEs was built based on previously established methods (Low et al., 2019; Nayfach et al., 2021).
516 Briefly, 36 ORFs that are commonly present in the MGEs were selected as marker genes. For each MGE,
517 the marker genes were concatenated in order, with a gap introduced in cases where a marker gene was
518 absent. The tree of the MGEs was then inferred based on the concatenated sequences.

519 To assess host-MGE relationships, the 16S rDNA sequences of the hosts of the 102 MGEs were
520 extracted. Among them, 5 host genomes lack detectable 16S rDNA regions and hence were excluded
521 from the following analysis. For the 97 remaining genomes, phylogenetic trees were built separately
522 based on the 16S rDNA sequences by Clustal (v2.1) (Sievers et al., 2011), and based on the MGE
523 sequences by the same method as before (Low et al., 2019; Nayfach et al., 2021). The two trees were
524 converted to distance matrices in R (version 4.1.2), with rows clustered according to MGE distances.

525 Plasmid construction

526 All plasmids used in this study are listed in Table S2. Plasmid pUIsacB (plasmid MPB35522) backbone
527 was amplified by PCR. SBW25 genomic fragments flanking the region of interest, where a mutagenesis,
528 knock-out or knock-in was to be performed, were amplified by PCR. In particular, the Kan^R fragment
529 used to tag I55 was produced by cloning the Kan^R gene from plasmid pMRE-Tn*7*-152 (plasmid
530 MPB27708) followed by a synthesised transcription terminator (CCA ATT ATT GAA CAC CCT TCG GGG
531 TGT TTT TTT GTT TCT GGT CTC CC). Primers were designed with a 5' overhang to facilitate assembly.
532 Backbone and fragments were assembled (with NEBuilder HiFi DNA Assembly Master Mix) and
533 transformed into *E. coli* Top10 competent cells. Successful transformants were selected on LB + Tet
534 plates, and confirmed by PCR and Sanger sequencing across the construct.

535 Strain construction

536 *Two-step allelic replacement (2SAR)*

537 The method is adapted from Hmelo et al. (Hmelo et al., 2015) and extensively described by Farr et al.
538 (Farr et al., 2024). Briefly, recipient (*P. fluorescens* SBW25), donor (*E. coli* carrying the plasmid used for
539 allelic exchange) and helper (*E. coli* carrying plasmid pRK2013, plasmid MPB35623) strains were cultured
540 overnight until stationary phase. Then, 1 mL recipient culture was heat-shocked at 42 °C for 20 min, and
541 subsequently mixed with 500 µL donor and 500 µL helper culture. The mixture was pelleted and
542 resuspended in 100 µL LB, which was then spotted onto an LB plate for incubation overnight. The
543 cultivated cell layer was then plated out on an LB + nitrofurantoin + Tet plate and incubated overnight to
544 select for trans-conjugates. Single trans-conjugates were then picked and streaked onto TYS10 plates,
545 which were subsequently incubated overnight. Finally, PCR and Sanger sequencing were used to screen
546 for the desired mutants.

547 *Fluorescent tagging*

548 The method is adapted from Schlechter et al. (Schlechter et al., 2018). Briefly, recipient (*P. fluorescens*
549 SBW25) and donor (*E. coli* carrying the plasmid pMRE-Tn*7*-152 or pMRE-Tn*7*-155, MPB27708 or
550 MPB27709, respectively) strains were cultured overnight until stationary phase. Then, 1 mL recipient
551 culture was heat-shocked at 42 °C for 20 min, and subsequently mixed with 1 mL donor culture. The
552 mixture was pelleted and resuspended in 100 µL LB, which was then spotted onto an LB plate for
553 incubation overnight. The cultivated cell layer was then plated out on an LB + nitrofurantoin + Kan +
554 0.1% w/v arabinose plate and incubated overnight to select for trans-conjugates. Single trans-conjugates
555 were picked, and integration of the fluorescent tag was then confirmed by PCR.

556 ## Preparation of SBW25-I55 and SBW25 (reconstructed) for assays

557 DNA sequencing of the evolved isolate carrying I55 (strain MPB33357) revealed, in addition to presence
558 of the mobile element, two SNPs elsewhere in the genome (see Table S5) that had arisen during the
559 course of serial propagation. Additionally, this isolate also harboured the Tn*7*-152 cassette. While there
560 was no reason to suspect any connection to I55, these SNPs were reverted to the ancestral SBW25
561 sequence, and Tn*7*-152 removed, by site-directed mutagenesis using the 2SAR strategy described above.
562 Prior to these mutagenesis steps, two independent colonies were picked from the evolved isolate and
563 treated as biological replicates in all subsequent genetic manipulations. The resulting genotype, SBW25-
564 I55 (strain MPB39024/MPB39026) is in principle isogenic with SBW25 (strain MPB29212). However, to
565 be more certain of this fact, the ancestral SBW25 genotype used in the fitness assays and plaque assays
566 was re-generated from SBW25-I55 by removal of the element, giving reconstructed SBW25 (strain
567 MPB39379/MPB39459).

568 ## Intercellular transfer assay

569 *One-step regime*: The donor, a SBW25 strain with a Kan^R-tagged I55 (with or without *intY*, strains
570 MPB39465/MPB39467 or MPB39565/MPB39568, respectively), and the recipient, a SBW25 strain with a
571 Tet^R marker (strain MPB39500), were cultivated separately for 48 hrs for pre-conditioning. Then 25 µL of
572 each culture and 50 µL of community filtrate from frozen stocks (autoclaved or not) were inoculated
573 into the same microcosm and then cultivated for a further 48 hrs. A dilution series of the culture was
574 then spread plated on LB + Kan + Tet plates, which were then cultivated for 24 hrs. The number of
575 colonies were counted and recorded.

576 *Two-step regime*: For the first step, the donor, a SBW25 strain with a Kan^R-tagged I55 (with or without
577 *intY*, strains MPB39465/MPB39467 or MPB39565/MPB39568, respectively), were cultivated for 48 hrs
578 for pre-conditioning. Then 50 µL of the donor culture, with or without 50 µL of community filtrate from
579 frozen stocks, was added to microcosms and then cultivated for a further 48 hrs. A stock of donor-
580 culture filtrate was obtained by filtering the culture through a 0.2 µm filter (Whatman). The filtrate was
581 distributed into 100 µL aliquots. Aliquots were directly frozen in -80 °C for storage. To retrieve the
582 donor-culture filtrate, one aliquot from the frozen stock was kept at room temperature for ~10 min.

583 For the second step, the recipient, a SBW25 strain with a Tet^R marker (strain MPB39500), was cultivated
584 for 48 hrs for pre-conditioning. Then 50 µL of the recipient culture and 50 µL of the donor-culture filtrate
585 from frozen stocks (obtained in the first step), with or without 50 µL of community filtrate from frozen
586 stocks, were added into the same microcosm and then cultivated for a further 48 hrs. A dilution series of
587 the culture was then spread plated on LB + Kan + Tet plates, which were then cultivated for 24 hrs. The
588 number of colonies were counted and recorded.

589 *Conjugation regime*: The donor, a SBW25 strain with Kan^R-tagged I55 (strain MPB39465/MPB39467),
590 recipient, a SBW25 strain with a Tet^R marker (strain MPB39500), and helper, *E. coli* carrying plasmid
591 pRK2013 (plasmid MPB35623), were cultured overnight until stationary phase. Then, 1 mL recipient
592 culture was heat-shocked at 42 °C for 20 min, and subsequently mixed with 500 µL donor and 500 µL
593 helper culture. The mixture was pelleted and resuspended in 100 µL LB, which was then spotted onto an
594 LB plate for incubation overnight. The cultivated cell layer was then plated out on an LB + Kan + Tet plate
595 and incubated overnight to select for trans-conjugates.

596 ## Community filtrate sequencing and analysis

597 To sequence each community filtrate, a 600 µL aliquot of filtrate was extracted with Qiagen DNeasy
598 PowerSoil Pro kit. The concentrations of eluted DNA were in general lower than ~1 ng/µL and
599 insufficient for subsequent sequencing. To amplify the DNA, 10 µL of eluted DNA was amplified with
600 Cytiva GenomiPhi V3 Ready-To-Go DNA Amplification kit. The amplified DNA was extracted again with
601 Qiagen DNeasy PowerSoil Pro kit. The DNA samples Illumina short-read sequenced (read length 150 bp)
602 by Eurofins Scientific SE for ~100M paired-end reads.

603 Bioinformatic analysis of sequencing data was done in Geneious Prime 2022.2.1
604 (https://www.geneious.com). For each sample, raw reads were trimmed by BBDuk and separately
605 mapped to the sequences of I23, I44 and I55.

606 ## Fluorescent fitness assay

607 The protocol was adapted from Barnett et al. (Barnett et al., 2024). Two competitor strains, one with a
608 GFP tag and the other with an RFP tag, were cultivated separately for 48 hrs for pre-conditioning. Then
609 25 µL of each culture, and 50 µL of community filtrate or 1 µL of phage lysate (see following sections)
610 from frozen stocks were inoculated into the same microcosm, which was then cultivated for a further 48
611 hrs. At the beginning and end of the second incubation period, the ratios of GFP to RFP cells were
612 determined by flow cytometry (MACSQuant VYB), denoted by r_0 and r_1, and total densities of cells in
613 the microcosm were determined by spread plating on LB plates, denoted by n_0 and n_1. The relative
614 fitness of GFP to RFP cells is the ratio of Malthusian parameters (Lenski et al., 1991), given by

615 $$1 + \frac{\log\left(\frac{r_1}{r_0}\right)}{\log\left(\frac{n_1(1+r_0)}{n_0(1+r_1)}\right)}.$$

616 ## Plaque assay

617 To determine the phage titre, a dilution series of community filtrate or phage lysate (see following
618 sections) was made. Then, 50 µL diluted material and 400 µL bacterial host overnight culture were
619 mixed into 4 mL SSA (kept at 55 °C). The mixture was immediately poured onto a LB plate. The plaque
620 counts were determined after ~40 hrs incubation. Pictures of the plates were taken with Bio-Rad
621 ChemiDoc Imager, under the colorimetric setting with exposure time 0.1 sec.

622 ## Bacteriophage isolation and storage

623 To purify phages from the community filtrate, a plaque assay (Clokie & Kropinski, 2009) using the filtrate
624 was performed. A single isolated plaque was then picked, on which a plaque assay was performed. This
625 was repeated at least three times. Finally, a single isolated plaque from the last assay was picked and
626 amplified in exponential-phase bacterial host culture for ~3 hrs. The obtained phage lysate was
627 chloroform extracted and stored in -80 °C with 50% glycerol saline. For the lysate extraction control, an
628 exponential-phase bacterial host culture without added phages was extracted and stored in the same
629 way.

630 ## Statistical analysis

631 All experiments were done for at least 4 independent biological replicates. Statistical analyses were
632 performed with GraphPad Prism (v. 10.1.2). Welch's *t*-test and Mann-Whitney test were performed for

parametric and non-parametric pairwise comparisons, respectively. In each plot, p values are indicated above the bars: $p > 0.05$ not significant (ns), $p \leq 0.05$ (∗), $p \leq 0.01$ (∗∗), $p \leq 0.001$ (∗∗∗), $p \leq 0.0001$ (∗∗∗∗).

636 ## Miscellaneous

637 ### Data availability

638 Supporting information, including sequencing raw reads, bioinformatic codes, gel images and plaque
639 assay images, will be deposited in Edmond (https://edmond.mpg.de/).

640 ### Author contributions

641 Y.Z., A.D.F., D.W.R. and P.B.R. designed research; Y.Z. and Y.M. performed research; all authors analysed
642 data; and Y.Z. and P.B.R. wrote the paper; all authors approved the final version.

643 ### Acknowledgements

644 We acknowledge the Max Planck Society for generous core funding. We thank members of Department
645 of Microbial Population Biology (past and present) for support, technical assistance and discussion.

646 ### Competing interests

647 The authors declare no competing interest.

648

649 ## References

650 Arkin, A. P., Cottingham, R. W., Henry, C. S., Harris, N. L., Stevens, R. L., Maslov, S., Dehal, P., Ware, D.,
651 Perez, F., Canon, S., Sneddon, M. W., Henderson, M. L., Riehl, W. J., Murphy-Olson, D., Chan, S.
652 Y., Kamimura, R. T., Kumari, S., Drake, M. M., Brettin, T. S., … Yu, D. (2018). KBase: The United
653 States Department of Energy Systems Biology Knowledgebase. *Nature Biotechnology*, *36*(7),
654 566–569. https://doi.org/10.1038/nbt.4163

655 Arnold, B. J., Huang, I.-T., & Hanage, W. P. (2022). Horizontal gene transfer and adaptive evolution in
656 bacteria. *Nature Reviews Microbiology*, *20*(4), 206–218. https://doi.org/10.1038/s41579-021-
657 00650-4

658 Barcia-Cruz, R., Goudenège, D., Moura de Sousa, J. A., Piel, D., Marbouty, M., Rocha, E. P. C., & Le Roux,
659 F. (2024). Phage-inducible chromosomal minimalist islands (PICMIs), a novel family of small
660 marine satellites of virulent phages. *Nature Communications*, *15*(1), 664.
661 https://doi.org/10.1038/s41467-024-44965-1

662 Barnett, M., Zeller, L., & Rainey, P. B. (2024). Experimental evolution of evolvability. *BioRxiv*.
663 https://doi.org/10.1101/2024.05.01.592015

664 Camargo, A. P., Roux, S., Schulz, F., Babinski, M., Xu, Y., Hu, B., Chain, P. S. G., Nayfach, S., & Kyrpides, N.
665 C. (2023). Identification of mobile genetic elements with geNomad. *Nature Biotechnology*.
666 https://doi.org/10.1038/s41587-023-01953-y

667 Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., & Huerta-Cepas, J. (2021). eggNOG-
668 mapper v2: Functional annotation, orthology assignments, and domain prediction at the
669 metagenomic scale. *Molecular Biology and Evolution*, *38*(12), 5825–5829.
670 https://doi.org/10.1093/molbev/msab293

671 Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., & Parks, D. H. (2020). GTDB-Tk: A toolkit to classify
672 genomes with the Genome Taxonomy Database. *Bioinformatics*, *36*(6), 1925–1927.
673 https://doi.org/10.1093/bioinformatics/btz848
674 Chen, J., Quiles-Puchalt, N., Chiang, Y. N., Bacigalupe, R., Fillol-Salom, A., Chee, M. S. J., Fitzgerald, J. R.,
675 & Penadés, J. R. (2018). Genome hypermobility by lateral transduction. *Science*, *362*(6411), 207–
676 212. https://doi.org/10.1126/science.aat5867
677 Cheng, R., Huang, F., Lu, X., Yan, Y., Yu, B., Wang, X., & Zhu, B. (2023). Prokaryotic Gabija complex senses
678 and executes nucleotide depletion and DNA cleavage for antiviral defense. *Cell Host & Microbe*,
679 *31*(8), 1331-1344.e5. https://doi.org/10.1016/j.chom.2023.06.014
680 Clokie, M. R. J., & Kropinski, A. M. (Eds.). (2009). *Bacteriophages* (Vol. 501). Humana Press.
681 https://doi.org/10.1007/978-1-60327-164-6
682 Colombi, E., Bertels, F., Doulcier, G., McConnell, E., Pichugina, T., Sohn, K. H., Straub, C., McCann, H. C.,
683 & Rainey, P. B. (2024). Rapid dissemination of host metabolism–manipulating genes via
684 integrative and conjugative elements. *Proceedings of the National Academy of Sciences*, *121*(11),
685 e2309263121. https://doi.org/10.1073/pnas.2309263121
686 Colombi, E., Straub, C., Künzel, S., Templeton, M. D., McCann, H. C., & Rainey, P. B. (2017). Evolution of
687 copper resistance in the kiwifruit pathogen *Pseudomonas syringae* pv. *Actinidiae* through
688 acquisition of integrative conjugative elements and plasmids. *Environmental Microbiology*,
689 *19*(2), 819–832. https://doi.org/10.1111/1462-2920.13662
690 de Sousa, J. A. M., Fillol-Salom, A., Penadés, J. R., & Rocha, E. P. C. (2023). Identification and
691 characterization of thousands of bacteriophage satellites across bacteria. *Nucleic Acids*
692 *Research*, *51*(6), 2759–2777. https://doi.org/10.1093/nar/gkad123

693 Dillingham, M. S., & Kowalczykowski, S. C. (2008). RecBCD enzyme and the repair of double-stranded
694 DNA breaks. *Microbiology and Molecular Biology Reviews*, *72*(4), 642–671.
695 https://doi.org/10.1128/MMBR.00020-08
696 Dmitrijeva, M., Tackmann, J., Matias Rodrigues, J. F., Huerta-Cepas, J., Coelho, L. P., & von Mering, C.
697 (2024). A global survey of prokaryotic genomes reveals the eco-evolutionary pressures driving
698 horizontal gene transfer. *Nature Ecology & Evolution*, *8*(5), 986–998.
699 https://doi.org/10.1038/s41559-024-02357-0
700 Doolittle, W. F., & Inkpen, S. A. (2018). Processes and patterns of interaction as units of selection: An
701 introduction to ITSNTS thinking. *Proceedings of the National Academy of Sciences*, *115*(16),
702 4006–4014. https://doi.org/10.1073/pnas.1722232115
703 Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., Amitai, G., & Sorek, R. (2018).
704 Systematic discovery of antiphage defense systems in the microbial pangenome. *Science*,
705 *359*(6379), eaar4120. https://doi.org/10.1126/science.aar4120
706 Farr, A. D., Vasileiou, C., Lind, P. A., & Rainey, P. B. (2024). An extreme mutational hotspot in *nlpD*
707 depends on transcriptional induction of *rpoS*. *BioRxiv*.
708 https://doi.org/10.1101/2024.07.11.603030
709 Faruque, S. M., & Mekalanos, J. J. (2003). Pathogenicity islands and phages in *Vibrio cholerae* evolution.
710 *Trends in Microbiology*, *11*(11), Article 11. https://doi.org/10.1016/j.tim.2003.09.003
711 Faruque, S. M., & Mekalanos, J. J. (2012). Phage-bacterial interactions in the evolution of toxigenic *Vibrio*
712 *cholerae*. *Virulence*, *3*(7), Article 7. https://doi.org/10.4161/viru.22351
713 Faruque, S. M., Zhu, J., Asadulghani, Kamruzzaman, M., & Mekalanos, J. J. (2003). Examination of diverse
714 toxin-coregulated pilus-positive *Vibrio cholerae* strains fails to demonstrate evidence for *Vibrio*
715 pathogenicity island phage. *Infection and Immunity*, *71*(6), 2993–2999.
716 https://doi.org/10.1128/IAI.71.6.2993-2999.2003

717 Fiedoruk, K., Zakrzewska, M., Daniluk, T., Piktel, E., Chmielewska, S., & Bucki, R. (2020). Two lineages of
718 *Pseudomonas aeruginosa* filamentous phages: Structural uniformity over integration
719 preferences. *Genome Biology and Evolution*, *12*(10), 1765–1781.
720 https://doi.org/10.1093/gbe/evaa146
721 Fillol-Salom, A., Bacarizo, J., Alqasmi, M., Ciges-Tomas, J. R., Martínez-Rubio, R., Roszak, A. W., Cogdell,
722 R. J., Chen, J., Marina, A., & Penadés, J. R. (2019). Hijacking the hijackers: *Escherichia coli*
723 pathogenicity islands redirect helper phage packaging for their own benefit. *Molecular Cell*,
724 *75*(5), 1020-1030.e4. https://doi.org/10.1016/j.molcel.2019.06.017
725 Fillol-Salom, A., Bacigalupe, R., Humphrey, S., Chiang, Y. N., Chen, J., & Penadés, J. R. (2021). Lateral
726 transduction is inherent to the life cycle of the archetypical *Salmonella* phage P22. *Nature*
727 *Communications*, *12*(1), 6510. https://doi.org/10.1038/s41467-021-26520-4
728 Fortmann-Grote, C., Hugoson, E., Summers, J., Theodosiou, L., & Rainey, P. B. (2023). Genome update
729 for *Pseudomonas fluorescens* isolate SBW25. *Microbiology Resource Announcements*, *12*(2),
730 e00637-22. https://doi.org/10.1128/mra.00637-22
731 Frazão, N., Sousa, A., Lässig, M., & Gordo, I. (2019). Horizontal gene transfer overrides mutation in
732 *Escherichia coli* colonizing the mammalian gut. *Proceedings of the National Academy of Sciences*,
733 *116*(36), 17906–17915. https://doi.org/10.1073/pnas.1906958116
734 Griffith, F. (1928). The significance of pneumococcal types. *Journal of Hygiene*, *27*(2), 113–159.
735 https://doi.org/10.1017/S0022172400031879
736 Grindley, N. D. F., Whiteson, K. L., & Rice, P. A. (2006). Mechanisms of site-specific recombination.
737 *Annual Review of Biochemistry*, *75*(1), 567–605.
738 https://doi.org/10.1146/annurev.biochem.73.011303.073908
739 Groth, A. C., & Calos, M. P. (2004). Phage integrases: Biology and applications. *Journal of Molecular*
740 *Biology*, *335*(3), 667–678. https://doi.org/10.1016/j.jmb.2003.09.082

741 Groussin, M., Poyet, M., Sistiaga, A., Kearney, S. M., Moniz, K., Noel, M., Hooker, J., Gibbons, S. M.,
742 Segurel, L., Froment, A., Mohamed, R. S., Fezeu, A., Juimo, V. A., Lafosse, S., Tabe, F. E., Girard,
743 C., Iqaluk, D., Nguyen, L. T. T., Shapiro, B. J., … Alm, E. J. (2021). Elevated rates of horizontal gene
744 transfer in the industrialized human microbiome. *Cell*, *184*(8), 2053-2067.e18.
745 https://doi.org/10.1016/j.cell.2021.02.052
746 Guo, L., Sattler, L., Shafqat, S., Graumann, P. L., & Bramkamp, M. (2022). A bacterial dynamin-like
747 protein confers a novel phage resistance strategy on the population level in *Bacillus subtilis*.
748 *MBio*, *13*(1). https://journals.asm.org/doi/10.1128/mbio.03753-21
749 Hackl, T., Laurenceau, R., Ankenbrand, M. J., Bliem, C., Cariani, Z., Thomas, E., Dooley, K. D., Arellano, A.
750 A., Hogle, S. L., Berube, P., Leventhal, G. E., Luo, E., Eppley, J. M., Zayed, A. A., Beaulaurier, J.,
751 Stepanauskas, R., Sullivan, M. B., DeLong, E. F., Biller, S. J., & Chisholm, S. W. (2023). Novel
752 integrative elements and genomic plasticity in ocean ecosystems. *Cell*, *186*(1), 47-62.e16.
753 https://doi.org/10.1016/j.cell.2022.12.006
754 Hall, J. P. J., Wood, A. J., Harrison, E., & Brockhurst, M. A. (2016). Source–sink plasmid transfer dynamics
755 maintain gene mobility in soil bacterial communities. *Proceedings of the National Academy of*
756 *Sciences*, *113*(29), 8260–8265. https://doi.org/10.1073/pnas.1600974113
757 Hmelo, L. R., Borlee, B. R., Almblad, H., Love, M. E., Randall, T. E., Tseng, B. S., Lin, C., Irie, Y., Storek, K.
758 M., Yang, J. J., Siehnel, R. J., Howell, P. L., Singh, P. K., Tolker-Nielsen, T., Parsek, M. R.,
759 Schweizer, H. P., & Harrison, J. J. (2015). Precision-engineering the *Pseudomonas aeruginosa*
760 genome with two-step allelic exchange. *Nature Protocols*, *10*(11), 1820–1841.
761 https://doi.org/10.1038/nprot.2015.115
762 Hudson, C. M., Lau, B. Y., & Williams, K. P. (2015). Islander: A database of precisely mapped genomic
763 islands in tRNA and tmRNA genes. *Nucleic Acids Research*, *43*(D1), D48–D53.
764 https://doi.org/10.1093/nar/gku1072

765 Humphrey, S., Fillol-Salom, A., Quiles-Puchalt, N., Ibarra-Chávez, R., Haag, A. F., Chen, J., & Penadés, J. R.
766 (2021). Bacterial chromosomal mobility via lateral transduction exceeds that of classical mobile
767 genetic elements. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-
768 021-26004-5
769 Hussain, F. A., Dubert, J., Elsherbini, J., Murphy, M., VanInsberghe, D., Arevalo, P., Kauffman, K., Rodino-
770 Janeiro, B. K., Gavin, H., Gomez, A., Lopatina, A., Le Roux, F., & Polz, M. F. (2021). Rapid
771 evolutionary turnover of mobile genetic elements drives bacterial resistance to phages. *Science*,
772 *374*(6566), 488–492. https://doi.org/10.1126/science.abb1083
773 Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal:
774 Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*,
775 *11*(1), 119. https://doi.org/10.1186/1471-2105-11-119
776 Ibarra-Chávez, R., Hansen, M. F., Pinilla-Redondo, R., Seed, K. D., & Trivedi, U. (2021). Phage satellites
777 and their emerging applications in biotechnology. *FEMS Microbiology Reviews*, *45*(6), fuab031.
778 https://doi.org/10.1093/femsre/fuab031
779 Ishikawa, T., Watabe, K., Mukohara, Y., & Nakamura, H. (1997). Mechanism of stereospecific conversion
780 of DL-5-substituted hydantoins to the corresponding L-amino acids by *Pseudomonas* sp. Strain
781 NS671. *Bioscience, Biotechnology, and Biochemistry*, *61*(1), 185–187.
782 https://doi.org/10.1271/bbb.61.185
783 Johnson, M. C., Laderman, E., Huiting, E., Zhang, C., Davidson, A., & Bondy-Denomy, J. (2023). Core
784 defense hotspots within *Pseudomonas aeruginosa* are a consistent and rich source of anti-phage
785 defense systems. *Nucleic Acids Research*, *51*(10), 4995–5005.
786 https://doi.org/10.1093/nar/gkad317
787 Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A.,
788 Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R.,

789 & Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification.
790 *Bioinformatics*, *30*(9), 1236–1240. https://doi.org/10.1093/bioinformatics/btu031
791 Karaolis, D. K. R., Somara, S., Maneval, D. R., Johnson, J. A., & Kaper, J. B. (1999). A bacteriophage
792 encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria.
793 *Nature*, *399*(6734), Article 6734. https://doi.org/10.1038/20715
794 Koutsovoulos, G. D., Granjeon Noriot, S., Bailly-Bechet, M., Danchin, E. G. J., & Rancurel, C. (2022). AvP:
795 A software package for automatic phylogenetic detection of candidate horizontal gene
796 transfers. *PLOS Computational Biology*, *18*(11), e1010686.
797 https://doi.org/10.1371/journal.pcbi.1010686
798 Lang, A. S., Zhaxybayeva, O., & Beatty, J. T. (2012). Gene transfer agents: Phage-like elements of genetic
799 exchange. *Nature Reviews Microbiology*, *10*(7), 472–482. https://doi.org/10.1038/nrmicro2802
800 Lederberg, J., & Tatum, E. L. (1946). Gene recombination in *Escherichia coli*. *Nature*, *158*(4016), 558–
801 558. https://doi.org/10.1038/158558b0
802 Lenski, R. E. (2017). Experimental evolution and the dynamics of adaptation and genome evolution in
803 microbial populations. *The ISME Journal*, *11*(10), 2181–2194.
804 https://doi.org/10.1038/ismej.2017.69
805 Lenski, R. E., Rose, M. R., Simpson, S. C., & Tadler, S. C. (1991). Long-term experimental evolution in
806 *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *The American*
807 *Naturalist*, *138*(6), 1315–1341. https://doi.org/10.1086/285289
808 Lenton, T. M., Kohler, T. A., Marquet, P. A., Boyle, R. A., Crucifix, M., Wilkinson, D. M., & Scheffer, M.
809 (2021). Survival of the systems. *Trends in Ecology & Evolution*, *36*(4), 333–344.
810 https://doi.org/10.1016/j.tree.2020.12.003
811 Lilley, A. K., & Bailey, M. J. (1997a). Impact of plasmid pQBR103 acquisition and carriage on the
812 phytosphere fitness of *Pseudomonas fluorescens* SBW25: Burden and benefit. *Applied and*

Environmental Microbiology, *63*(4), 1584–1587. https://doi.org/10.1128/aem.63.4.1584-1587.1997

Lilley, A. K., & Bailey, M. J. (1997b). The acquisition of indigenous plasmids by a genetically marked pseudomonad population colonizing the sugar beet phytosphere is related to local environmental conditions. *Applied and Environmental Microbiology*, *63*(4), 1577–1583. https://doi.org/10.1128/aem.63.4.1577-1583.1997

Lindsay, J. A., Ruzin, A., Ross, H. F., Kurepina, N., & Novick, R. P. (1998). The gene for toxic shock toxin is carried by a family of mobile pathogenicity islands in *Staphylococcus aureus*. *Molecular Microbiology*, *29*(2), 527–543. https://doi.org/10.1046/j.1365-2958.1998.00947.x

Low, S. J., Džunková, M., Chaumeil, P.-A., Parks, D. H., & Hugenholtz, P. (2019). Evaluation of a concatenated protein phylogeny for classification of tailed double-stranded DNA viruses belonging to the order Caudovirales. *Nature Microbiology*, *4*(8), 1306–1315. https://doi.org/10.1038/s41564-019-0448-z

Marrs, B. (1974). Genetic recombination in *Rhodopseudomonas capsulata*. *Proceedings of the National Academy of Sciences*, *71*(3), 971–973. https://doi.org/10.1073/pnas.71.3.971

Nayfach, S., Páez-Espino, D., Call, L., Low, S. J., Sberro, H., Ivanova, N. N., Proal, A. D., Fischbach, M. A., Bhatt, A. S., Hugenholtz, P., & Kyrpides, N. C. (2021). Metagenomic compendium of 189,680 DNA viruses from the human gut microbiome. *Nature Microbiology*, *6*(7), 960–970. https://doi.org/10.1038/s41564-021-00928-6

Néron, B., Denise, R., Coluzzi, C., Touchon, M., Rocha, E. P. C., & Abby, S. S. (2023). MacSyFinder v2: Improved modelling and search engine to identify molecular systems in genomes. *Peer Community Journal*, *3*, e28. https://doi.org/10.24072/pcjournal.250

835 Novick, R. P., Christie, G. E., & Penadés, J. R. (2010). The phage-related chromosomal islands of Gram-
836 positive bacteria. *Nature Reviews Microbiology*, *8*(8), 541–551.
837 https://doi.org/10.1038/nrmicro2393
838 Ofir, G., Melamed, S., Sberro, H., Mukamel, Z., Silverman, S., Yaakov, G., Doron, S., & Sorek, R. (2017).
839 DISARM is a widespread bacterial defence system with broad anti-phage activities. *Nature*
840 *Microbiology*, *3*(1), 90–98. https://doi.org/10.1038/s41564-017-0051-0
841 O'Shea, Y. A., & Boyd, E. F. (2002). Mobilization of the *Vibrio* pathogenicity island between *Vibrio*
842 *cholerae* isolates mediated by CP-T1 generalized transduction. *FEMS Microbiology Letters*,
843 *214*(2), 153–157. https://doi.org/10.1111/j.1574-6968.2002.tb11339.x
844 Pei, T.-T., Luo, H., Wang, Y., Li, H., Wang, X.-Y., Zhang, Y.-Q., An, Y., Wu, L.-L., Ma, J., Liang, X., Yan, A.,
845 Yang, L., Chen, C., & Dong, T. (2024). Filamentous prophage Pf4 promotes genetic exchange in
846 *Pseudomonas aeruginosa*. *The ISME Journal*, *18*(1), wrad025.
847 https://doi.org/10.1093/ismejo/wrad025
848 Quistad, S. D., Doulcier, G., & Rainey, P. B. (2020). Experimental manipulation of selfish genetic elements
849 links genes to microbial community function. *Philosophical Transactions of the Royal Society B:*
850 *Biological Sciences*, *375*, 20190681. https://doi.org/10.1098/rstb.2019.0681
851 Rainey, P. B., & Quistad, S. D. (2020). Toward a dynamical understanding of microbial communities.
852 *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*, 20190248.
853 https://doi.org/10.1098/rstb.2019.0248
854 Rainey, P. B., Remigi, P., Farr, A. D., & Lind, P. A. (2017). Darwin was right: Where now for experimental
855 evolution? *Current Opinion in Genetics & Development*, *47*, 102–109.
856 https://doi.org/10.1016/j.gde.2017.09.003
857 Rainey, P. B., & Travisano, M. (1998). Adaptive radiation in a heterogeneous environment. *Nature*,
858 *394*(6688), 69–72. https://doi.org/10.1038/27900

Roughgarden, J. (2020). Holobiont evolution: Mathematical model with vertical vs. horizontal microbiome transmission. *Philosophy, Theory, and Practice in Biology*, *12*(20220112). https://doi.org/10.3998/ptpbio.16039257.0012.002

Rousset, F., Depardieu, F., Miele, S., Dowding, J., Laval, A.-L., Lieberman, E., Garry, D., Rocha, E. P. C., Bernheim, A., & Bikard, D. (2022). Phages and their satellites encode hotspots of antiviral systems. *Cell Host & Microbe*, *30*(5), 740-753.e5. https://doi.org/10.1016/j.chom.2022.02.018

Schlechter, R. O., Jun, H., Bernach, M., Oso, S., Boyd, E., Muñoz-Lintz, D. A., Dobson, R. C. J., Remus, D. M., & Remus-Emsermann, M. N. P. (2018). Chromatic bacteria – A broad host-range plasmid and chromosomal insertion toolbox for fluorescent protein expression in bacteria. *Frontiers in Microbiology*, *9*, 3052. https://doi.org/10.3389/fmicb.2018.03052

Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, *7*(539), 1–6. https://doi.org/10.1038/msb.2011.75

Silby, M. W., Cerdeño-Tárraga, A. M., Vernikos, G. S., Giddens, S. R., Jackson, R. W., Preston, G. M., Zhang, X.-X., Moon, C. D., Gehrig, S. M., Godfrey, S. A., Knight, C. G., Malone, J. G., Robinson, Z., Spiers, A. J., Harris, S., Challis, G. L., Yaxley, A. M., Harris, D., Seeger, K., … Thomson, N. R. (2009). Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens*. *Genome Biology*, *10*(5), R51. https://doi.org/10.1186/gb-2009-10-5-r51

Sullivan, J. T., & Ronson, C. W. (1998). Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. *Proceedings of the National Academy of Sciences*, *95*(9), 5145–5149. https://doi.org/10.1073/pnas.95.9.5145

883 Sullivan, M. J., Petty, N. K., & Beatson, S. A. (2011). Easyfig: A genome comparison visualizer.
884 *Bioinformatics*, *27*(7), 1009–1010. https://doi.org/10.1093/bioinformatics/btr039
885 Tatusov, R. L. (2000). The COG database: A tool for genome-scale analysis of protein functions and
886 evolution. *Nucleic Acids Research*, *28*(1), 33–36. https://doi.org/10.1093/nar/28.1.33
887 Tesson, F., Hervé, A., Mordret, E., Touchon, M., d'Humières, C., Cury, J., & Bernheim, A. (2022).
888 Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nature*
889 *Communications*, *13*(1), 2561. https://doi.org/10.1038/s41467-022-30269-9
890 Tock, M. R., & Dryden, D. T. (2005). The biology of restriction and anti-restriction. *Current Opinion in*
891 *Microbiology*, *8*(4), 466–472. https://doi.org/10.1016/j.mib.2005.06.003
892 van Dijk, B., Buffard, P., Farr, A. D., Giersdorf, F., Meijer, J., Dutilh, B. E., & Rainey, P. B. (2023).
893 Identifying and tracking mobile elements in evolving compost communities yields insights into
894 the nanobiome. *ISME Communications*, *3*(90). https://doi.org/10.1101/2023.02.02.526783
895 Wang, X., Kim, Y., & Wood, T. K. (2009). Control and benefits of CP4-57 prophage excision in *Escherichia*
896 *coli* biofilms. *The ISME Journal*, *3*(10), 1164–1179. https://doi.org/10.1038/ismej.2009.59
897 Woods, L. C., Gorrell, R. J., Taylor, F., Connallon, T., Kwok, T., & McDonald, M. J. (2020). Horizontal gene
898 transfer potentiates adaptation by reducing selective constraints on the spread of genetic
899 variation. *Proceedings of the National Academy of Sciences*, *117*(43), 26868–26875.
900 https://doi.org/10.1073/pnas.2005331117
901 Xie, Z., & Tang, H. (2017). ISEScan: Automated identification of insertion sequence elements in
902 prokaryotic genomes. *Bioinformatics*, *33*(21), 3340–3347.
903 https://doi.org/10.1093/bioinformatics/btx433
904 Yuan, L., Lu, H., Li, F., Nielsen, J., & Kerkhoven, E. J. (2023). HGTphyloDetect: Facilitating the
905 identification and phylogenetic analysis of horizontal gene transfer. *Briefings in Bioinformatics*,
906 *24*(2), bbad035. https://doi.org/10.1093/bib/bbad035

907 Zinder, N. D., & Lederberg, J. (1952). Genetic exchange in *Salmonella*. *Journal of Bacteriology*, *64*(5),

908 679–699. https://doi.org/10.1128/jb.64.5.679-699.1952

909

910

Figure Captions

Fig. 1. Experimental evolution and detection of foreign elements

a. Three serial-propagation regimes to evolve SBW25 populations. (i) live filtrate regime; (ii) co-culture regime; (iii) fixed filtrate regime. Components within each box are specific to each regime, and components outside of boxes are common to all regimes.

b. The pipeline for screening and confirming SBW25 strains carrying foreign elements inserted into the genome.

c. Basic information for the 3 foreign elements detected. The table lists the number of base pairs of each foreign element, the name of the compost community from which the community filtrate was generated, the serial-propagation regime, the number of parallel lines from which isolates with the foreign element were detected, and the number of isolates containing the foreign element in each such line.

d. Insertion site of the foreign elements in the SBW25 genome. Upon insertion, two 14 bp direct repeats flank the foreign element. The direct repeats at the 5'-end and 3'-end of the element were treated as from SBW25 genome and from the foreign element, respectively, despite the exact origin of each repeat is unknown.

e. The aligned genetic maps of the 3 foreign elements, I23, I44 and I55. The alignment was created by Easyfig (M. J. Sullivan et al., 2011). Parallelograms between MGEs indicate similar genomic regions based on blastn. The ORFs were numbered serially from 5-end to 3'-end for each MGE, and coloured according to annotations. Detailed annotations of each ORF are listed in Table S4.

932 Fig. 2. Bioinformatic analysis of similar elements

933 **a**. Bioinformatic pipeline for detection of similar elements to I55. Bullet points below 'sequence
934 comparison' and 'sequence selection' represent the criteria for comparison and searching, respectively.

935 **b**. Cladogram and genomic representations of I23, I44, I55 and similar MGEs (identified by NCBI RefSeq
936 assembly indices). Basis for construction of the tree is summarised in the text. Shape of the tree
937 represents inferred evolutionary relationship but not evolutionary distances.

938 **c**. Rank abundance of annotated genes of different functional classes in panel b. Genes occurring at least
939 10 times are illustrated. A full rank abundance table is in Table S6. Colours represent COG functional
940 categories.

941 **d**. Pairwise distances between the MGEs (shown in lower-left), and pairwise distances between their
942 corresponding hosts (shown in upper-right). MGEs and hosts are identified by NCBI RefSeq assembly
943 indices. Basis for construction and interpretation of the heatmap is summarised in the text. Pink boxes
944 highlight pairs of strains with almost identical MGE, but remarkably different 16S sequence. Annotations
945 of those MGEs are boxed in pink in panel b.

946

947 Fig. 3. Intercellular transfer of I55

948 **a**. Schematic of the one-step intercellular transfer assay.

949 **b**. The densities of CFUs on LB + Kan + Tet plates in the intercellular transfer assay. Data are mean ± s.d.
950 The labels A/C and D.L. represent autoclaved filtrate and the detection limit of the assay, respectively.
951 For each treatment, 8 to 12 biological replicates were used. Mann-Whitney test was applied to test the
952 effect of IntA ($p < 0.0001$) when the community filtrate was added.

953 **c**. Schematic of the two-step intercellular transfer assay.

954 **d**. The densities of CFUs on LB + Kan + Tet plates in the two-stage intercellular transfer assay. Data are
955 mean ± s.d. The label D.L. represents the detection limit of the assay. For each treatment, 8 biological
956 replicates were used. Mann-Whitney test was applied to test the effect of IntA ($p = 0.0152$) when the
957 community filtrate was added at the first step.

958 **e**. Proposed model for the transfer of I55 via an unknown component in the community filtrate.

959

960 Fig. 4. Fitness advantage and phage defence activity of I55

961 **a**. Schematic of fluorescent fitness assay. The diagram indicates the case where SBW25-I55Δ*intY* was
962 tagged with GFP and the wild type was tagged with RFP. The reverse tagging strategy, where SBW25-
963 I55Δ*intY* was tagged with RFP and the wild type was tagged with GFP, is also conducted to eliminate any
964 fitness difference due to the choice of fluorescent protein markers.

965 **b**. Fluorescent fitness assay results for SBW25-I55Δ*intY* relative to wild type, in presence or absence of
966 community filtrate. Data are mean ± 95% CI. In the Filtrate + and - treatments, stored filtrate aliquots
967 from compost C and from SBW25 were added, respectively. The colour of each data point corresponds
968 to the colour of the fluorescent marker added to SBW25-I55Δ*intY*. For each treatment, 8 biological
969 replicates (4 for either tagging strategy) were used. One sample *t*-tests were applied to test the relative
970 fitness values against 1 when the community filtrate was added (p = 0.0104) and not added (p = 0.0033).
971 Paired Welch's *t*-test was applied to test the effect of community filtrate (p = 0.0006).

972 **c**. Typical photos of the plaque assay of the community filtrate against SBW25-I55Δ*intY* or wild type as
973 hosts. Picture contrast was altered to facilitate visual realisation of the plaques. Photos of three
974 additional biological replicates can be found in supporting information.

975 **d**. Plaque forming unit (PFU) results of the plaque assay in panel c. Data are mean ± s.d. For each strain,
976 four biological replicates were used. Paired Welch's *t*-test was applied to test the effect of I55Δ*intY* (p =
977 0.0080).

978 **e**. Fluorescent fitness assay results for SBW25-I55Δ*intY* relative to wild type, in presence or absence of
979 the two phage isolates (P1 and P2). Data are mean ± 95% CI. In the P1, P2 and Ctrl treatments, stored
980 phage lysate P1, phage lysate P2 and lysate extraction control (see Materials and Methods) were added,
981 respectively. The colour of each data point corresponds to the colour of the fluorescent marker added to
982 SBW25-I55Δ*intY*. For each treatment, 4 biological replicates (2 for either tagging strategy) were used.

983 **f**. PFU results of plaque assay of P1 and P2 (in (i) and (ii) respectively) against SBW25-I55Δ*intY* or wild
984 type as hosts. Data are mean ± s.d. For each strain, 4 biological replicates were used. Paired Welch's *t*-
985 tests were applied to test the effect of I55Δ*intY* (p = 0.0011 in the case of P1, and p = 0.0082 in the case
986 of P2).

987

988 ## Supplementary Figure Captions and Tables

989 **Fig. S1**. SBW25 density in the co-culture regime (see Fig. 1a(ii) and Materials and Methods)
990 SBW25 co-cultured with compost C community decreased in abundance over transfers. By the 5th
991 transfer the density was only 10^3 ~ 10^4 cells (per mL) across the three replicate lines. By contrast, mono-
992 culture of SBW25 maintained a density of 10^8 ~ 10^9 cells (per mL).

993

994 Fig. S2. Excision and circularisation of I55

995 **a**. The optical density of overnight LB shaking cultures of SBW25-I55, SBW25-I55Δ*intY* and wild type
996 (WT). Data are mean ± s.d. Brown-Forsythe and Welch ANOVA test was applied to test the difference
997 between the three strains (p = 0.3551).

998 **b**. PCR of overnight cultures of SBW25-I55, SBW25-I55Δ*intY* and WT, with combination of primers
999 ①②③④. Photos of 3 more biological replicates can be found in supporting information.

1000 **c**. A proposed explanation for the presence of bands for SBW25-I55 with primer pairs ①④ and ②③.
1001 Primers ① and ② flank the left insertion junction, and primers ③ and ④ flanks the right. Primers ①
1002 and ④ match the genome, and primers ② and ③ match I55. The presence of bands on SBW25-I55
1003 with pairs ①+② and ③+④ confirms the integration of I55 in the genome. The presence of band on
1004 SBW25-I55 with pair ①+④, at a length that is the same as WT, suggests that in the overnight culture
1005 of SBW25-I55 there are some cells with I55 excised from the genome. The presence of band on SBW25-
1006 I55 with pair ②+③ suggests that I55 somehow circularises so that primers ② and ③ flank the
1007 circularisation junction. Sanger sequencing of the PCR products of ②+③ confirms that the circular I55
1008 contains exactly one 14 bp sequence at the junction, and of ①+④ confirms that I55 excises in a scar-
1009 free manner. We therefore proposed that there are two forms of I55 in the population of overnight
1010 culture of SBW25-I55, the integrated form (which gives bands with pairs ①+② and ③+④) and the
1011 circular form (which gives bands with pairs ①+④ and ②+③). To test whether the tyrosine integrase
1012 (IntY) is necessary for excision, we deleted *intY* and did the same PCR. The strain SBW25-I55Δ*intY*
1013 showed bands with pairs ①+② and ③+④, but no band with pairs ①+④ and ②+③. This
1014 confirms that IntY is necessary for excision and circularisation of I55.

1015

1016 **Table S1**. List of bacterial strains

1017

Number	Name	Genotype	Construction notes
MPB29212	SBW25_Wild_Type	SBW25	Lab stock
MPB28806	SBW25 Tn7-PntpII-mScarlet-I-Gm	SBW25 *glmS*::Tn*7*-145(RFP, Gen^R)	Lab stock
MPB29284	Tn7_152_CFC_LB_SBW25	SBW25 *glmS*::Tn*7*-152(GFP, Kan^R)	Conjugation of MPB29212 with MPB27708
MPB29300	Tn7_155_CFC_LB_SBW25	SBW25 *glmS*::Tn*7*-155(RFP, Kan^R)	Conjugation of MPB29212 with MPB27709
MPB38392	SFF_b_GFP_4_iso24_23kb	SBW25 *tmRNA*::I23 *glmS*::Tn*7*-152(GFP, Kan^R) isolate1*	Isolated from fixed-filtrate experiment, compost b, GFP, line 4
MPB38393	SFF_b_GFP_5_iso02_23kb	SBW25 *tmRNA*::I23 *glmS*::Tn*7*-152(GFP, Kan^R) isolate2*	Isolated from fixed-filtrate experiment, compost b, GFP, line 5
MPB38394	SFF_b_GFP_5_iso04_23kb	SBW25 *tmRNA*::I23 *glmS*::Tn*7*-152(GFP, Kan^R) isolate3*	Isolated from fixed-filtrate experiment, compost b, GFP, line 5
MPB38395	SFF_b_GFP_5_iso10_23kb	SBW25 *tmRNA*::I23 *glmS*::Tn*7*-152(GFP, Kan^R) isolate4*	Isolated from fixed-filtrate experiment, compost b, GFP, line 5
MPB38396	SFF_b_GFP_5_iso13_23kb	SBW25 *tmRNA*::I23 *glmS*::Tn*7*-152(GFP, Kan^R) isolate5*	Isolated from fixed-filtrate experiment, compost b, GFP, line 5
MPB38895	SFF_b_RFP_3_iso13_23kb	SBW25 *tmRNA*::I23 *glmS*::Tn*7*-155(RFP, Kan^R) isolate6*	Isolated from fixed-filtrate experiment, compost b, RFP, line 3
MPB38906	SFF_b_RFP_4_iso06_44kb	SBW25 *tmRNA*::I44 *glmS*::Tn*7*-155(RFP, Kan^R) isolate1*	Isolated from fixed-filtrate experiment, compost b, RFP, line 4
MPB38907	SFF_b_RFP_4_iso23_44kb	SBW25 *tmRNA*::I44 *glmS*::Tn*7*-155(RFP, Kan^R) isolate2*	Isolated from fixed-filtrate experiment, compost b, RFP, line 4
MPB33357	SF_G_C_2_T14_iso04	SBW25 *tmRNA*::I55 *glmS*::Tn*7*-152(GFP, Kan^R) isolate1*	Isolated from live-filtrate experiment, compost c, GFP, line 2
MPB38103/MPB38109	Iso04_rev_PFLU4414	SBW25 *tmRNA*::I55 *glmS*::Tn*7*-152(GFP, Kan^R) *pflu3805*SNP	2SAR of two independent colonies of MPB33357 with MPB38033
MPB38572/MPB38574	Iso04_rev_PFLU4414_3805	SBW25 *tmRNA*::I55 *glmS*::Tn*7*-152(GFP, Kan^R)	2SAR of MPB38103/MPB38109 with MPB38391
MPB39024/MPB39026	Iso04_rev_PFLU4414_3805_del_Tn7	SBW25 *tmRNA*::I55	2SAR of MPB38572/MPB38574 with MPB38943
MPB39132/MPB39136	SBW25_I55_Δint	SBW25 *tmRNA*::(I55 Δ*intY*)	2SAR of MPB39024/MPB39026 with MPB38939
MPB39504/MPB39507	SBW25_I55_Δint_Tn7_GFP	SBW25 *tmRNA*::(I55 Δ*intY*) *glmS*::Tn*7*-152(GFP, Kan^R)	Conjugation of MPB39132/MPB39136 with MPB27708
MPB39514/MPB39516	SBW25_I55_Δint_Tn7_RFP	SBW25 *tmRNA*::(I55 Δ*intY*) *glmS*::Tn*7*-152(GFP, Kan^R)	Conjugation of MPB39132/MPB39136 with MPB27709
MPB39379/MPB39459	SBW25_reconstructed	SBW25	2SAR of MPB39132/MPB39136 with MPB39223
MPB39508/MPB39510	SBW25_reconstructed_Tn7_GFP	SBW25 *glmS*::Tn*7*-152(GFP, Kan^R)	Conjugation of MPB39379/MPB39459 with MPB27708
MPB39518/MPB39520	SBW25_reconstructed_Tn7_RFP	SBW25 *glmS*::Tn*7*-155(RFP, Kan^R)	Conjugation of MPB39379/MPB39459 with MPB27709
MPB39465/MPB39467	SBW25_I55_KanR_InsertS1	SBW25 *tmRNA*::(I55 (*orf08-orf09*)::Kan^R)	2SAR of MPB39024/MPB39026 with MPB39269
MPB39565/MPB39568	SBW25_I55_Δint_ KanR_InsertS1	SBW25 *tmRNA*::(I55 Δ*intY* (*orf08-orf09*)::Kan^R)	2SAR of MPB39132/MPB39136 with MPB39269
MPB39500	SBW25-TetR	SBW25 *glmS*::Tn7(AGACTGTTGCGCGCTGGCGGTGGCG, Tet^R)	Lab stock, see https://doi.org/10.1007/s00239-023-10103-6

1018

1019 * For variations/mutations, see Table S5.

1020

1021 Table S2. List of *E. coli* plasmids

1022

Number	Name	Description	Source
MPB27708	pMRE-Tn7-152	Plasmid to tag *Pseudomonas* with GFP and Kan^R	https://www.addgene.org/118566/
MPB27709	pMRE-Tn7-155	Plasmid to tag *Pseudomonas* with RFP and Kan^R	https://www.addgene.org/118569/
MPB35522	pUIsacB	Empty plasmid for 2-step allelic replacement	Lab stock, see https://www.biorxiv.org/content/10.1101/2024.07.11.603030v1
MPB35623	pRK2013	Helper plasmid for 2-step allelic replacement	Lab stock, see https://www.biorxiv.org/content/10.1101/2024.07.11.603030v1
MPB38033	rev_PFLU4414	Plasmid to revert the SNP in *pflu4414*	this study
MPB38391	rev_PFLU3805	Plasmid to revert the SNP in *pflu3805*	this study
MPB38943	del_Tn7	Plasmid to delete the Tn*7* cassette	this study
MPB38939	del_55kbInt	Plasmid to delete the gene *intY* of I55	this study
MPB39223	del_tmIns	Plasmid to delete any inserts downstream of *tmRNA*	this study
MPB39269	KanR_InsertS1	Plasmid to insert a KanR cassette between *orf08* and *orf09* of I55	this study

1023

1024

1025 Table S3. List of primers

1026

Number	Name	Purpose	Sequence
YZ001	YZ_SBW_5797955F	To detect potential insertions downstream of tmRNA	GTTACCGCTGCGATGTTGAC
YZ002	YZ_SBW_5799313R	To detect potential insertions downstream of tmRNA	GCGTGCTGATGCTGCTAATC
YZ003	YZ_In_1016R	To confirm the presence of I55	GTGGCCAGATCAGGGAAGAC
YZ004	YZ_In_54359F	To confirm the presence of I55	ACACGACACACGAAGAGACC
YZ005	pUIsacB_fwd	To amplify the pUIsacB plasmid backbone	AGCACTACATCAACTGACTA
YZ006	pUIsacB_rev	To amplify the pUIsacB plasmid backbone	CTGAACCAAGATAGCTGTAC
YZ007	2SAR_4880278_Insert_fwd	To construct the plasmid for reverting the SNP in *pflu4414*	GTACAGCTATCTTGGTTCAGGAAATCAACCCATCGCCTG
YZ008	2SAR_4880278_Insert_rev	To construct the plasmid for reverting the SNP in *pflu4414*	TAGTCAGTTGATGTAGTGCTAACGCCTGTCCTATGTGG
YZ009	pUIsacB_seq_fwd	To Sanger sequence the assembled regions of the pUIsacB plasmid	ATTTGTCCTACTCAGGAGAGCGTTC
YZ010	4880278_Insert_4962R	To Sanger sequence the plasmid for reverting the SNP in *pflu4414*	GCTGAGTTCGAATCCCTGCT
YZ011	4880278_Insert_4804F	To Sanger sequence the plasmid for reverting the SNP in *pflu4414*	CTTGCCATGCAGTTGGTCG
YZ012	4880278_Insert_5486R	To Sanger sequence the plasmid for reverting the SNP in *pflu4414*	GGCCGGCGAAATTTTAGAGC
YZ013	4880278_Insert_5275F	To Sanger sequence the plasmid for reverting the SNP in *pflu4414*	CGAATCAACGTGACGCTCAC
YZ014	4880278_Insert_5956R	To Sanger sequence the plasmid for reverting the SNP in *pflu4414*	CGACTGGGGCCGTTTCAT
YZ015	4880278_Insert_5777F	To Sanger sequence the plasmid for reverting the SNP in *pflu4414*	ACACGCTTGATCACCTGACC
YZ016	pUIsacB_seq_rev	To Sanger sequence the assembled regions of the pUIsacB plasmid	AGCGTTCTGAACAAATCCAGATG
YZ084	SBW_23kb_insert_1	To confirm the presence of I23	CCGTTTCTCTTTGGGGTCGA
YZ085	SBW_23kb_insert_2	To confirm the presence of I23	ACTCAAGAGCCTCGCCAAAG
YZ086	2SAR_rev_3805_1	To construct the plasmid for reverting the SNP in *pflu3805*	GTACAGCTATCTTGGTTCAGCCAGCAACGGGTTGTTCTTG
YZ087	2SAR_rev_3805_2	To construct the plasmid for reverting the SNP in *pflu3805*	TAGTCAGTTGATGTAGTGCTCCACTGATCTTGACGCCAGA
YZ088	rev_PFLU3805_4997F	To Sanger sequence the plasmid for reverting the SNP in *pflu3805*	CCTGTTCGCTGAAAATGGCC
YZ089	rev_PFLU3805_5687R	To Sanger sequence the plasmid for reverting the SNP in *pflu3805*	ACACGATGACGACGGTTCTG
YZ091	PFLU4414_Final_Seq_2	To confirm the genotype of mutants after allelic replacement in *pflu4414*	GAACGCATCCACGGCTAATG
YZ092	PFLU3805_Final_Seq_1	To confirm the genotype of mutants after allelic replacement in *pflu3805*	GACACGTCGACCTTCTGGAAG
YZ093	PFLU3805_Final_Seq_2	To confirm the genotype of mutants after allelic replacement in *pflu3805*	CGAGTTGATTGGGGTGGCC
YZ098	PFLU4414_Final_Seq_1_1	To confirm the genotype of mutants after allelic replacement in *pflu4414*	CGCCGACCATGTTCATGATG
YZ102	SBW_44kb_insert_1	To confirm the presence of I44	GCATAGGCTGTGGTCACTGA
YZ103	SBW_44kb_insert_2	To confirm the presence of I44	AGACCAAGGCCGTCAAGAAG
YZ105	2SAR_del_55kbInt_1	To construct the plasmid for deleting the tyrosine integrase (*orf01*) of I55	GTACAGCTATCTTGGTTCAGGTTTAGGATTCGACGCCG
YZ106	2SAR_del_55kbInt_2	To construct the plasmid for deleting the tyrosine integrase (*orf01*) of I55	GAAGTCTCTCGGGTGTATTCCAGTGCTGTTTATG
YZ107	2SAR_del_55kbInt_3	To construct the plasmid for deleting the tyrosine integrase (*orf01*) of I55	TGGAATACACCCGAGAGACTTCAACCAGTGA

YZ108	2SAR_del_55kbInt_4	To construct the plasmid for deleting the tyrosine integrase (*orf01*) of I55	TAGTCAGTTGATGTAGTGCTTCTTCGTTCACCACATTCCG
YZ109	del_55kbInt_4941F	To Sanger sequence the plasmid for deleting the tyrosine integrase (*orf01*) of I55	TGGACGAAGCGGCTAAAACT
YZ110	del_55kbInt_5514R	To Sanger sequence the plasmid for deleting the tyrosine integrase (*orf01*) of I55	TCGCCATTCAGCTGTCCAAT
YZ111	del_55kbInt_Final_Seq_1	To confirm the genotype of mutants after allelic replacement in the tyrosine integrase (*orf01*) of I55	ACGTTAGCTCAGCCGAGGT
YZ112	2SAR_del_Tn7_1	To construct the plasmid for deleting the Tn*7* cassette	GTACAGCTATCTTGGTTCAGTGAGCTATCAGTTGCGCCTG
YZ113	2SAR_del_Tn7_2	To construct the plasmid for deleting the Tn*7* cassette	TAGTCAGTTGATGTAGTGCTAATCCGACCTGACCTTGCTG
YZ114	del_Tn7_4864F	To Sanger sequence the plasmid for deleting the Tn*7* cassette	TTGCGGGTTACGGATGACC
YZ115	del_Tn7_5273F	To Sanger sequence the plasmid for deleting the Tn*7* cassette	GACGTTGGTGGATCAGGGAG
YZ116	del_Tn7_5670F	To Sanger sequence the plasmid for deleting the Tn*7* cassette	CAATGGGCCGTGTTTGAGC
YZ117	del_Tn7_5084R	To Sanger sequence the plasmid for deleting the Tn*7* cassette	TTGGAAGGTGAGCTGCTGTC
YZ118	del_Tn7_5480R	To Sanger sequence the plasmid for deleting the Tn*7* cassette	CTGTCGCCGATCCTCTACAC
YZ119	del_Tn7_5845R	To Sanger sequence the plasmid for deleting the Tn*7* cassette	CAATGGACAGCACCGTGGAA
YZ120	del_Tn7_Final_Seq_1	To confirm the genotype of mutants after allelic replacement in the Tn*7* cassette	ACACCGCCCAACGTGATG
YZ121	del_Tn7_Final_Seq_2	To confirm the genotype of mutants after allelic replacement in the Tn*7* cassette	ACACCCTGTTCGTGACCATC
YZ124	2SAR_KanR_InsertS1_1	To construct the plasmid for inserting an Kan^R cassette between *orf08* and *orf09* of I55	GTACAGCTATCTTGGTTCAGGTCCGCAGAACTGTTGGAG
YZ125	2SAR_KanR_InsertS1_2	To construct the plasmid for inserting an Kan^R cassette between *orf08* and *orf09* of I55	GATTAGCTGTGTCGCTACTG
YZ126	2SAR_KanR_InsertS1_3	To construct the plasmid for inserting an Kan^R cassette between *orf08* and *orf09* of I55	CCGCTCACCGAAGCTGATTA
YZ127	2SAR_KanR_InsertS1_4	To construct the plasmid for inserting an Kan^R cassette between *orf08* and *orf09* of I55	TAGTCAGTTGATGTAGTGCTGGCGACTTCGGTATTCCAG
YZ128	2SAR_KanR_InsertS1_5	To construct the plasmid for inserting an Kan^R cassette between *orf08* and *orf09* of I55	CAGTAGCGACACAGCTAATCGGGAGACCAGAAACAAAAA AACACCCCGAAGGGTGTTCAATAATTGGTCCGCTTCCTTTA GCAGCC
YZ129	2SAR_KanR_InsertS1_6	To construct the plasmid for inserting an Kan^R cassette between *orf08* and *orf09* of I55	TAATCAGCTTCGGTGAGCGGTCCTAGGTGCTCGAGTGG
YZ130	KanR_InsertS1_5008F	To Sanger sequence the plasmid for inserting an Kan^R cassette between *orf08* and *orf09* of I55	TGGCACACGATACCGAAGAC
YZ131	KanR_InsertS1_5508F	To Sanger sequence the plasmid for inserting an Kan^R cassette between *orf08* and *orf09* of I55	AGCGTGCGAGCTCTAAAACA
YZ132	KanR_InsertS1_6011F	To Sanger sequence the plasmid for inserting an Kan^R cassette between *orf08* and *orf09* of I55	ACTCTTCCGAGCAAAGGACG
YZ133	KanR_InsertS1_6662F	To Sanger sequence the plasmid for inserting an Kan^R cassette between *orf08* and *orf09* of I55	GCAGGAAATGGGTCATCGGT
YZ134	KanR_InsertS1_5189R	To Sanger sequence the plasmid for inserting an Kan^R cassette between *orf08* and *orf09* of I55	TTGCAGCAGTGACAAGGACA
YZ135	KanR_InsertS1_5695R	To Sanger sequence the plasmid for inserting an Kan^R cassette between *orf08* and *orf09* of I55	TTGATCTTGGGAGAAGCGGC
YZ136	KanR_InsertS1_6177R	To Sanger sequence the plasmid for inserting an Kan^R cassette between *orf08* and *orf09* of I55	GGACAGCCGGTATAAAGGGA
YZ137	KanR_InsertS1_6853R	To Sanger sequence the plasmid for inserting an Kan^R cassette between *orf08* and *orf09* of I55	TCTCACACGTTCGTTCGACC
YZ144	2SAR_del_tmIns_1	To construct the plasmid for deleting MGEs inserted downstream of *tmRNA*	GTACAGCTATCTTGGTTCAGGAGATCGACACCGCCGTCTC
YZ145	2SAR_del_tmIns_2	To construct the plasmid for deleting MGEs inserted downstream of *tmRNA*	TAGTCAGTTGATGTAGTGCTGCGTGCTGATGCTGCTAATC
YZ146	del_tmIns_5225F	To Sanger sequence the plasmid for deleting MGEs inserted downstream of *tmRNA*	GGGTCGCTGAGGGTTAACTT
YZ147	del_tmIns_5699F	To Sanger sequence the plasmid for deleting MGEs inserted downstream of *tmRNA*	ATGACGCTCGTGTCCAACTT
YZ148	del_tmIns_4886R	To Sanger sequence the plasmid for deleting MGEs inserted downstream of *tmRNA*	CTGTGCAGTGTACGCCAAAC

YZ149	del_tmIns_5575R	To Sanger sequence the plasmid for deleting MGEs inserted downstream of *tmRNA*	GGTACAGACCCTGTTCGTCG
YZ150	del_tmIns_Final_Seq_1	To confirm the genotype of mutants after allelic replacement downstream of *tmRNA*	TTACCGCTGCGATGTTGACT
YZ151	del_tmIns_Final_Seq_2	To confirm the genotype of mutants after allelic replacement downstream of *tmRNA*	GGAGCGTCCGTGAGAGATTC
YZ161	Tn7R-109	To confirm the insertion of Tn7 cassette in SBW25 genome	CAGCATAACTGGACTGATTTCAG
YZ162	SBW25-glmS	To confirm the insertion of Tn7 cassette in SBW25 genome	CACCAAAGCTTTCACCACCCAA

1027

1028

1029 Table S4. Annotations of MGEs I23, I44 and I55

1030

Element	ORF name	Consensus annotation	Minimum coordinate	Maximum coordinate	Length	Direction of transcription	Prokka product	Pfam domains	DefenseFinder prediction
I23	*orf01 (intY)*	tyrosine integrase	186	1442	1257	forward	Prophage integrase IntA	PF13356 (Arm-DNA-bind_3; aa 8-94); PF22022 (Phage_int_M; aa 107-200); PF00589 (Phage_integrase; aa 214-378).	NA
I23	*orf02*	GajA	1631	3436	1806	reverse	ATP-dependent DNA helicase Rep	PF00580 (UvrD-helicase; aa 5-125, 133-239).	Gabija__GajB_1.
I23	*orf03*	GajB	3462	5312	1851	reverse	hypothetical protein	PF13175 (AAA_15; aa 1-72, 310-381); PF20469 (OLD-like_TOPRIM; aa 429-500).	Gabija__GajA.
I23	*orf04*	unannotated	5637	5966	330	forward	hypothetical protein	NA	NA
I23	*orf05*	DUF932	6062	7030	969	forward	hypothetical protein	PF06067 (DUF932; aa 77-302).	NA
I23	*orf06*	YqaJ	7106	8110	1005	forward	hypothetical protein	PF09588 (YqaJ; aa 29-172).	NA
I23	*orf07*	Gp3-like	8201	9151	951	forward	hypothetical protein	PF18897 (Gp3-like; aa 12-206).	NA
I23	*orf08*	RadC	9123	9620	498	forward	hypothetical protein	PF04002 (RadC; aa 44-163).	NA
I23	*orf09*	unannotated	9634	9843	210	forward	hypothetical protein	NA	NA
I23	*orf10*	unannotated	9903	10208	306	forward	hypothetical protein	NA	NA
I23	*orf11*	HTH-containing protein	10195	10542	348	reverse	hypothetical protein	PF01381 (HTH_3; aa 17-71).	NA
I23	*orf12*	unannotated	10681	11034	354	forward	hypothetical protein	NA	NA
I23	*orf13*	unannotated	11155	11973	819	forward	hypothetical protein	NA	NA
I23	*orf14*	DNA topoisomerase	12150	12920	771	reverse	hypothetical protein	PF08378 (NERD; aa 33-145); PF01396 (zf-C4_Topoisom; aa 217-245).	NA
I23	*orf15*	T5orf172 helicase	13022	14206	1185	reverse	hypothetical protein	PF10544 (T5orf172; aa 263-374).	NA
I23	*orf16*	DEAD-like helicase	14199	16265	2067	reverse	hypothetical protein	PF00270 (DEAD; aa 28-193).	NA
I23	*orf17*	unannotated	16268	16372	105	reverse	hypothetical protein	NA	NA
I23	*orf18*	DNA methyltransferase	16369	19131	2763	reverse	hypothetical protein	PF20464 (MmeI_N; aa 1-159); PF20465 (MmeI_hel; aa 170-250); PF20473 (MmeI_Mtase; aa 325-606); PF20466 (MmeI_TRD; aa 629-832); PF20467 (MmeI_C; aa 834-911).	RM_Type_IIG__Type_IIG.
I23	*orf19*	inovirus Gp2	19299	20000	702	reverse	hypothetical protein	PF11726 (Inovirus_Gp2; aa 50-231).	NA
I23	*orf20 (alpA)*	AlpA	20507	20740	234	reverse	hypothetical protein	PF05930 (Phage_AlpA; aa 1-51).	NA
I23	*orf21*	unannotated	20810	22114	1305	reverse	hypothetical protein	NA	NA
I23	*orf22*	unannotated	22346	23020	675	forward	hypothetical protein	NA	NA

I44	*orf01 (intY)*	tyrosine integrase	185	1441	1257	forward	Prophage integrase IntA	PF13356 (Arm-DNA-bind_3; aa 8-94); PF22022 (Phage_int_M; aa 107-200); PF00589 (Phage_integrase; aa 214-377).	NA
I44	*orf02*	type II REase	1609	2364	756	forward	hypothetical protein	PF01844 (HNH; aa 180-235).	RM_Type_II__Type_II_REases
I44	*orf03*	type II Mtase	2467	3531	1065	forward	putative BsuMI modification methylase subunit YdiO	PF00145 (DNA_methylase; aa 4-336).	RM_Type_II__Type_II_MTases
I44	*orf04*	VSR endonuclease	3482	3898	417	reverse	Very short patch repair protein	PF03852 (Vsr; aa 2-64).	NA
I44	*orf05*	DrmC	3916	4554	639	reverse	Cardiolipin synthase B	PF13091 (PLDc_2; aa 67-200).	DISARM__drmC
I44	*orf06*	DrmB	4682	6538	1857	reverse	hypothetical protein	PF09369 (MZB; aa 483-583).	DISARM__drmB
I44	*orf07*	DrmA	6541	10575	4035	reverse	hypothetical protein	NA	DISARM__drmA
I44	*orf08*	unannotated	10579	10908	330	reverse	hypothetical protein	NA	NA
I44	*orf09*	DrmMI	10905	14981	4077	reverse	hypothetical protein	PF07669 (Eco57I; aa 845-931).	DISARM_1__drmMI
I44	*orf10*	DrmD	14978	18103	3126	reverse	RNA polymerase-associated protein RapA	PF00176 (SNF2-rel_dom; aa 134-317); PF00271 (Helicase_C; aa 514-620).	DISARM_1__drmD
I44	*orf11*	MFS transporter	18695	19495	801	reverse	Glycerol-3-phosphate transporter	PF07690 (MFS_1; aa 29-257).	NA
I44	*orf12*	alcohol dehydrogenase	19609	20700	1092	reverse	NAD-dependent methanol dehydrogenase	PF00465 (Fe-ADH; aa 8-351).	NA
I44	*orf13*	avidin	20847	21623	777	reverse	hypothetical protein	PF01382 (Avidin; aa 3-85).	NA
I44	*orf14*	MFS transporter	21691	23055	1365	reverse	Putative niacin/nicotinamide transporter NaiP	PF07690 (MFS_1; aa 46-378).	NA
I44	*orf15*	Asp/Glu/hydantoin racemase	23118	23939	822	reverse	Hydantoin racemase	PF01177 (Asp_Glu_race; aa 20-225).	NA
I44	*orf16*	unannotated	23960	24850	891	reverse	hypothetical protein	NA	NA
I44	*orf17*	metallopeptidase	24907	26292	1386	reverse	Xaa-Pro aminopeptidase	PF05195 (AMP_N; aa 7-123); PF00557 (Peptidase_M24; aa 173-438).	NA
I44	*orf18*	hydantoinase B	26320	27978	1659	reverse	Acetophenone carboxylase delta subunit	PF02538 (Hydantoinase_B; aa 5-522).	NA
I44	*orf19*	hydantoinase A	27992	30076	2085	reverse	Acetophenone carboxylase gamma subunit	PF05378 (Hydant_A_N; aa 9-185); PF01968 (Hydantoinase_A; aa 208-500); PF19278 (Hydant_A_C; aa 514-689).	NA

I44	*orf20*	HTH-type transcriptional regulator	30253	32856	2604	reverse	HTH-type transcriptional regulator MalT	PF00196 (GerE; aa 802-856).	NA
I44	*orf21*	unannotated	33248	33571	324	forward	hypothetical protein	NA	NA
I44	*orf22*	DUF932	33664	34632	969	forward	hypothetical protein	PF06067 (DUF932; aa 77-302).	NA
I44	*orf23*	YqaJ	34707	35711	1005	forward	hypothetical protein	PF09588 (YqaJ; aa 29-172).	NA
I44	*orf24*	Gp3-like	35790	36740	951	forward	hypothetical protein	PF18897 (Gp3-like; aa 12-206).	NA
I44	*orf25*	RadC	36712	37209	498	forward	hypothetical protein	PF04002 (RadC; aa 44-163).	NA
I44	*orf26*	unannotated	37223	37432	210	forward	hypothetical protein	NA	NA
I44	*orf27*	unannotated	37534	37779	246	forward	hypothetical protein	NA	NA
I44	*orf28*	HTH-containing protein	37787	38134	348	reverse	hypothetical protein	PF01381 (HTH_3; aa 17-71).	NA
I44	*orf29*	unannotated	38272	38658	387	forward	hypothetical protein	NA	NA
I44	*orf30*	unannotated	38675	39493	819	forward	hypothetical protein	NA	NA
I44	*orf31*	inovirus Gp2	39957	40916	960	reverse	hypothetical protein	PF11726 (Inovirus_Gp2; aa 149-232).	NA
I44	*orf32 (alpA)*	AlpA	41252	41482	231	reverse	hypothetical protein	PF05930 (Phage_AlpA; aa 1-51).	NA
I44	*orf33*	unannotated	41531	42988	1458	reverse	hypothetical protein	NA	NA
I44	*orf34*	unannotated	43158	43829	672	forward	hypothetical protein	NA	NA
I55	*orf01 (intY)*	tyrosine integrase	185	1441	1257	forward	Prophage integrase IntA	PF13356 (Arm-DNA-bind_3; aa 8-94); PF22022 (Phage_int_M; aa 107-200); PF00589 (Phage_integrase; aa 214-378).	NA
I55	*orf02*	dynamin	2594	4903	2310	forward	hypothetical protein	PF00350 (Dynamin_N; aa 59-226).	NA
I55	*orf03*	dynamin	4914	7166	2253	forward	hypothetical protein	PF00350 (Dynamin_N; aa 40-177).	NA
I55	*orf04*	unannotated	7235	7738	504	forward	hypothetical protein	NA	NA
I55	*orf05*	unannotated	7735	8307	573	forward	hypothetical protein	NA	NA
I55	*orf06*	unannotated	8349	10991	2643	forward	hypothetical protein	NA	NA
I55	*orf07*	unannotated	11420	12151	732	forward	hypothetical protein	NA	NA
I55	*orf08*	unannotated	12148	13686	1539	forward	hypothetical protein	NA	NA
I55	*orf09*	helicase	14030	16783	2754	reverse	RNA polymerase-associated protein RapA	PF04851 (ResIII; aa 36-245); PF00271 (Helicase_C; aa 755-814).	NA
I55	*orf10*	unannotated	16770	18677	1908	reverse	hypothetical protein	NA	NA
I55	*orf11*	unannotated	18674	20056	1383	reverse	hypothetical protein	NA	NA
I55	*orf12*	unannotated	20681	21103	423	forward	hypothetical protein	NA	NA
I55	*orf13*	DUF1232	21186	21524	339	forward	hypothetical protein	PF06803 (DUF1232; aa 50-84).	NA
I55	*orf14*	unannotated	21652	22302	651	forward	hypothetical protein	NA	NA
I55	*orf15*	unannotated	22316	22447	132	reverse	hypothetical protein	NA	NA

I55	*orf16*	AAA	22605	24695	2091	forward	ATP-dependent zinc metalloprotease FtsH	PF00004 (AAA; aa 255-377, 491-614).	NA
I55	*orf17*	WYL	24791	25813	1023	forward	hypothetical protein	PF13280 (WYL; aa 156-330).	NA
I55	*orf18*	unannotated	25859	26395	537	forward	hypothetical protein	NA	NA
I55	*orf19*	Sir2	26521	27174	654	reverse	NAD-dependent protein deacylase	PF02146 (SIR2; aa 20-196).	NA
I55	*orf20*	unannotated	27411	27668	258	reverse	hypothetical protein	NA	NA
I55	*orf21*	Sel1	27658	28299	642	reverse	hypothetical protein	PF08238 (Sel1; aa 96-127, 133-166, 171-200).	NA
I55	*orf22*	unannotated	28421	28585	165	forward	hypothetical protein	NA	NA
I55	*orf23*	RecD	28573	30666	2094	reverse	RecBCD enzyme subunit RecD	PF21185 (RecD_N; aa 40-153); PF13245 (AAA_19; aa 205-357); PF13538 (UvrD_C_2; aa 615-662).	NA
I55	*orf24*	RecB	30663	34337	3675	reverse	RecBCD enzyme subunit RecB	PF00580 (UvrD-helicase; aa 13-452); PF13361 (UvrD_C; aa 459-660); PF13361 (UvrD_C; aa 710-848); PF12705 (PDDEXK_1; aa 1072-1214).	NA
I55	*orf25*	RecC	34334	37804	3471	reverse	RecBCD enzyme subunit RecC	PF04257 (Exonuc_V_gamma; aa 11-360); PF17946 (RecC_C; aa 835-1081).	NA
I55	*orf26*	unannotated	38042	38371	330	forward	hypothetical protein	NA	NA
I55	*orf27*	DUF932	38462	39430	969	forward	hypothetical protein	PF06067 (DUF932; aa 77-302).	NA
I55	*orf28*	YqaJ	39513	40514	1002	forward	hypothetical protein	PF09588 (YqaJ; aa 28-171).	NA
I55	*orf29*	Gp3-like	40605	41498	894	forward	hypothetical protein	PF18897 (Gp3-like; aa 12-206).	NA
I55	*orf30*	unannotated	41554	41742	189	forward	hypothetical protein	NA	NA
I55	*orf31*	unannotated	41803	42126	324	forward	hypothetical protein	NA	NA
I55	*orf32*	HTH-containing protein	42136	42483	348	reverse	hypothetical protein	PF01381 (HTH_3; aa 17-71).	NA
I55	*orf33*	unannotated	42622	42975	354	forward	hypothetical protein	NA	NA
I55	*orf34*	unannotated	43189	43812	624	forward	hypothetical protein	NA	NA
I55	*orf35*	MUG113	43980	45188	1209	reverse	hypothetical protein	PF13455 (MUG113; aa 285-373).	NA
I55	*orf36*	DEAD-like helicase	45181	47247	2067	reverse	hypothetical protein	PF00270 (DEAD; aa 28-193).	NA
I55	*orf37*	DNA methyltransferase	47349	50105	2757	reverse	hypothetical protein	PF20464 (MmeI_N; aa 1-159); PF20465 (MmeI_hel; aa 170-250); PF20473 (MmeI_Mtase; aa 325-605); PF20466 (MmeI_TRD; aa 627-826); PF20467 (MmeI_C; aa 827-909).	RM_Type_IIG__Type_IIG
I55	*orf38*	inovirus Gp2	50350	51321	972	reverse	hypothetical protein	NA	NA
I55	*orf39 (alpA)*	AlpA	51661	51888	228	reverse	hypothetical protein	PF05930 (Phage_AlpA; aa 1-51).	NA
I55	*orf40*	unannotated	51950	53260	1311	reverse	hypothetical protein	NA	NA

I55	*orf41*	unannotated	53417	54289	873	forward	hypothetical protein	NA	NA

- **ORF name**: This is the name given to a gene on an MGE. On each MGE, genes are named serially and consecutively from 5'-end (near tyrosine integrase) to 3'-end in the format *orf#*, where # is a two-digit number. The tyrosine integrase always has #=01, the genes following it have #=02, 03, 04, etc. For some well-recognisable genes, common gene names are indicated inside round brackets following *orf#*.
- **Consensus annotation**: This is the annotation given to a gene after manual checks on Prokka and Pfam predictions. Genes are called "unannotated" when neither predictors find results.
- **Minimum coordinate**: This is the position of the first nucleotide of a gene. For each MGE, the nucleotides are numbered from the 5'-end, with the first nucleotide immediately downstream of the 14 bp repeat (near the 3'-end of *tmRNA*) labelled as 1.
- **Maximum coordinate**: This is the position of the last nucleotide of a gene.
- **Length**: This is the number of nucleotides in a gene, and is equal to Maximum coordinate - Minimum coordinate + 1.
- **Direction of transcription**: This is the direction of transcription of a gene. Genes labelled as "forward" are transcribed from the 5'-end to the 3'-end along the annotated strand (i.e., the strand where the tyrosine integrase is at the 5'-end). Genes labelled as "forward" are transcribed from the 5'-end to the 3'-end along the other stand.
- **Prokka product**: This is the protein product predicted by Prokka in KBase. Proteins without predictions are labelled as "hypothetical protein".
- **Pfam domains**: This is the protein domain(s) predicted by Pfam in InterProScan. For each predicted domain, the Pfam Id is listed, with the corresponding Pfam name and mapped region (given by the interval of amino acids along the amino acid sequence) given in round brackets. For proteins with multiple predicted domains, all domains are listed and separated by semicolons. Proteins with no predicted domains are labelled as "NA".
- **DefenseFinder prediction**: This is the gene name of putative phage defence genes predicted by DenfenseFinder. Genes not predicted to be involved in phage defence activities are labelled as "NA".

1049 Table S5. Mutations detected in isolates with MGEs.

1050

Strain number	Genotype	Gene locus	Genome coordinate	Type of mutation	Change	CDS position	Effect on protein
MPB38392	SBW25 *tmRNA*::I23 *glmS*::Tn*7*-152(GFP, Kan^R) isolate1	*pflu4414*	4878944-4878945	insertion (tandem repeat)	T(2) -> T(3)	1489-1490	frameshift
MPB38393	SBW25 *tmRNA*::I23 *glmS*::Tn*7*-152(GFP, Kan^R) isolate2	*pflu4414 - pflu4415*	4880381-4880570	deletion	del 190 bp	*pflu4414*: 1-53; *pflu4415*: 664-789	*pflu4415*: truncation; *pflu4414*: no translation
MPB38394	SBW25 *tmRNA*::I23 *glmS*::Tn*7*-152(GFP, Kan^R) isolate3	*pflu4414*	4880259	SNP (transition)	G -> A	175	truncation
MPB38395	SBW25 *tmRNA*::I23 *glmS*::Tn*7*-152(GFP, Kan^R) isolate4	*pflu3022*	3295100-3295889	deletion	del 790 bp	387-1176	truncation
		pflu4414	4879083-4879084	insertion (tandem repeat)	T(5) -> T(6)	1350-1351	frameshift
MPB38396	SBW25 *tmRNA*::I23 *glmS*::Tn*7*-152(GFP, Kan^R) isolate5	*pflu4414 - pflu4415*	4880381-4880570	deletion	del 190 bp	*pflu4414*: 1-53; *pflu4415*: 664-789	*pflu4415*: truncation; *pflu4414*: no translation
MPB38895	SBW25 *tmRNA*::I23 *glmS*::Tn*7*-155(RFP, Kan^R) isolate6	None	-	-	-	-	-
MPB38906	SBW25 *tmRNA*::I44 *glmS*::Tn*7*-155(RFP, Kan^R) isolate1	*pflu4744*	5217912	SNP (transition)	G -> A	116	substitution (S39N)
MPB38907	SBW25 *tmRNA*::I44 *glmS*::Tn*7*-155(RFP, Kan^R) isolate2	*pflu4744*	5217912	SNP (transition)	G -> A	116	substitution (S39N)
MPB33357	SBW25 *tmRNA*::I55 *glmS*::Tn*7*-152(GFP, Kan^R) isolate1	*pflu3805*	4199776	SNP (transversion)	A -> T	33	substitution (N11K)
		pflu4414	4880277-4880278	insertion (tandem repeat)	C(5) -> C(6)	157-158	frameshift

1051

1052

1053 Table S6. Rank abundance of genes carried by similar MGEs

1054

Domains	COG category	Description	Counts
Phage_integrase	L	Belongs to the 'phage' integrase family	111
AlpA	K	transcriptional regulator	109
Inovirus_Gp2	S	Inovirus Gp2	101
WYL	K	regulation of single-species biofilm formation	87
DUF932	S	Domain of unknown function (DUF932)	80
YqaJ	L	COG5377 Phage-related protein	78
RadC	L	RadC-like JAB domain	66
DUF1524,DUF262	S	Protein of unknown function DUF262	42
HsdM_N,N6_Mtase	V	Type I restriction-modification system methyltransferase subunit	40
Methylase_S	V	PFAM Type I restriction modification DNA specificity domain	31
N6_Mtase	V	Type II restriction enzyme, methylase	24
DUF1788	S	Domain of unknown function (DUF1788)	23
AAA_5	V	GTPase subunit of restriction endonuclease	20
Dynamin_N	T	ATPase. Has a role at an early stage in the morphogenesis of the spore coat	20
HTH_3,HTH_31	K	sequence-specific DNA binding	20
MMR_HSR1	P	Dynamin family	17
ADP_ribosyl_GH	O	ADP-ribosylglycohydrolase	16
DNA_methylase	H	Belongs to the class I-like SAM-binding methyltransferase superfamily. C5-methyltransferase family	16
DUF262	S	Protein of unknown function DUF262	16
Helicase_C,SNF2_N	KL	Transcription regulator that activates transcription by stimulating RNA polymerase (RNAP) recycling in case of stress conditions such as supercoiled DNA or high salt concentrations. Probably acts by releasing the RNAP, when it is trapped or immobilized on tightly supercoiled DNA. Does not activate transcription on linear DNA. Probably not involved in DNA repair	15
Vsr	L	May nick specific sequences that contain T G mispairs resulting from m5C-deamination	15
EcoR124_C,HSDR_N,ResIII	L	Subunit R is required for both nuclease and ATPase activities, but not for modification	14
Lon_2,Lon_C	O	Putative ATP-dependent Lon protease	14
PDDEXK_1,UvrD-helicase,UvrD_C	L	A helicase nuclease that prepares dsDNA breaks (DSB) for recombinational DNA repair. Binds to DSBs and unwinds DNA via a highly rapid and processive ATP-dependent bidirectional helicase activity. Unwinds dsDNA until it encounters a Chi (crossover hotspot instigator) sequence from the 3' direction. Cuts ssDNA a few nucleotides 3' to the Chi site. The properties and activities of the enzyme are changed at Chi. The Chi-altered holoenzyme produces a long 3'-ssDNA overhang and facilitates RecA-binding to the ssDNA for homologous DNA recombination and repair. Holoenzyme degrades any linearized DNA that is unable to undergo homologous recombination. In the holoenzyme this subunit contributes ATPase, 3'-5' helicase, exonuclease activity and loads RecA onto ssDNA	14
Helicase_C	KL	Helicase conserved C-terminal domain	13
HTH_1,LysR_substrate	K	transcriptional regulator	13

McrBC	V	McrBC 5-methylcytosine restriction system component	13
HSDR_N,ResIII	L	Subunit R is required for both nuclease and ATPase activities, but not for modification	12
HTH_17	K	TIGRFAM DNA binding domain, excisionase family	12
DUF1998	S	Domain of unknown function (DUF1998)	11
PglZ	H	PglZ domain	11
Ank_2,Ank_4,Ank_5	G	response to abiotic stimulus	10
N6_N4_Mtase	H	DNA methylase N-4 N-6	10
ResIII	KL	Type III restriction enzyme, res subunit	10
DEAD,DUF1998,Helicase_C	L	COG1205 Distinct helicase family with a unique C-terminal domain including a metal-binding cysteine cluster	9
DUF1819	S	Putative inner membrane protein (DUF1819)	9
DUF45	S	Metal-dependent hydrolase	9
HTH_Tnp_1	L	Transposase and inactivated derivatives	9
Lon_2,Lon_C,MIT_C	O	ATP-dependent Lon-type protease	9
PLDc_2	I	phospholipase d transphosphatidylase	9
AAA_15	L	AAA ATPase domain	8
BPD_transp_1	P	Binding-protein-dependent transport system inner membrane component	8
DNA_processg_A	LU	DNA recombination-mediator protein A	8
DUF697,MMR_HSR1	D	50S ribosome-binding GTPase	8
Transp_cyt_pur	FH	Permease for cytosine/purines, uracil, thiamine, allantoin	8
AAA_15,AAA_21,AAA_23	T	AAA domain, putative AbiEii toxin, Type IV TA system	7
DEAD,Helicase_C,RecQ_Zn_bind	L	DNA helicase	7
Eco57I,N6_Mtase	V	COG1002 Type II restriction enzyme, methylase subunits	7
AAA_11,AAA_12,AAA_30,PDDEXK_1,RNase_H_2	L	RNase_H superfamily	6
AAA_19,AAA_30,UvrD_C_2	L	A helicase nuclease that prepares dsDNA breaks (DSB) for recombinational DNA repair. Binds to DSBs and unwinds DNA via a highly rapid and processive ATP-dependent bidirectional helicase activity. Unwinds dsDNA until it encounters a Chi (crossover hotspot instigator) sequence from the 3' direction. Cuts ssDNA a few nucleotides 3' to the Chi site. The properties and activities of the enzyme are changed at Chi. The Chi-altered holoenzyme produces a long 3'-ssDNA overhang and facilitates RecA-binding to the ssDNA for homologous DNA recombination and repair. Holoenzyme degrades any linearized DNA that is unable to undergo homologous recombination. In the holoenzyme this subunit has ssDNA-dependent ATPase and 5'-3' helicase activity. When added to pre-assembled RecBC greatly stimulates nuclease activity and augments holoenzyme processivity. Negatively regulates the RecA-loading ability of RecBCD	6
AAA_23	L	ATPase involved in DNA repair	6
ABC_tran,DUF726,Dynamin_N	S	ATPase. Has a role at an early stage in the morphogenesis of the spore coat	6
CxxCxxCC	S	PFAM Uncharacterised protein family (UPF0153)	6
DUF2357,PDDEXK_7	S	PD-(D/E)XK nuclease superfamily	6

DUF445	S	membrane	6
EH_Signature	S	EH_Signature domain	6
Exonuc_V_gamma	L	A helicase nuclease that prepares dsDNA breaks (DSB) for recombinational DNA repair. Binds to DSBs and unwinds DNA via a highly rapid and processive ATP-dependent bidirectional helicase activity. Unwinds dsDNA until it encounters a Chi (crossover hotspot instigator) sequence from the 3' direction. Cuts ssDNA a few nucleotides 3' to the Chi site. The properties and activities of the enzyme are changed at Chi. The Chi-altered holoenzyme produces a long 3'-ssDNA overhang and facilitates RecA-binding to the ssDNA for homologous DNA recombination and repair. Holoenzyme degrades any linearized DNA that is unable to undergo homologous recombination. In the holoenzyme this subunit recognizes the wild- type Chi sequence, and when added to isolated RecB increases its ATP-dependent helicase processivity	6
Helicase_C,Mrr_cat,SNF2_N, SWIM	KL	COG0553 Superfamily II DNA RNA helicases, SNF2 family	6
PDDEXK_1	L	PD-(D/E)XK nuclease superfamily	6
Amidohydro_3	F	cytosine deaminase	5
DDE_Tnp_1,DDE_Tnp_1_6,DUF772	L	Transposase DDE domain	5
DUF1232	S	Protein of unknown function (DUF1232)	5
DUF3987	T	Protein of unknown function (DUF3987)	5
Helicase_C,ResIII,SNF2_N	L	PFAM Helicase conserved C-terminal domain	5
HNH_2	V	restriction endonuclease	5
Hydant_A_N,Hydantoinase_A	EQ	N-methylhydantoinase A acetone carboxylase, beta subunit	5
Hydantoinase_B	EQ	Hydantoinase B/oxoprolinase	5
Isochorismatase	Q	Isochorismatase family	5
MFS_1	EGP	Major Facilitator Superfamily	5
MFS_1,Sugar_tr	EGP	Major facilitator superfamily	5
MotB_plug,OmpA	N	Membrane MotB of proton-channel complex MotA/MotB	5
PDDEXK_7	S	PD-(D/E)XK nuclease superfamily	5
UvrD-helicase,UvrD_C,UvrD_C_2	L	Uvrd rep helicase	5
Virulence_RhuM	S	COG3943 Virulence protein	5
AAA	O	COG0464 ATPases of the AAA class	4
AAA_13	L	AAA domain	4
AAA_5,DUF3578	V	Domain of unknown function (DUF3578)	4
ABC_tran,TOBE_2	P	Belongs to the ABC transporter superfamily	4
DAO	E	FAD dependent oxidoreductase	4
DUF2950	N	Protein of unknown function (DUF2950)	4
DUF3077	S	Protein of unknown function (DUF3077)	4
DUF4357	S	Domain of unknown function (DUF4357)	4

Fer2_4	C	2Fe-2S iron-sulfur cluster binding domain	4
Fer2_BFD,Pyr_redox_2	C	FAD-dependent pyridine nucleotide-disulphide oxidoreductase	4
Helicase_C,ResIII	L	Type III restriction protein res subunit	4
His_Phos_1,NUDIX	L	nUDIX hydrolase	4
HTH_30,PucR	QT	transcriptional regulator	4
Ketoacyl-synt_C,ketoacyl-synt	IQ	catalyzes a condensation reaction in fatty acid biosynthesis addition of an acyl acceptor of two carbons from malonyl-ACP	4
Pkinase	KLT	Protein tyrosine kinase	4
Ribonuc_L-PSP	J	Endoribonuclease L-PSP	4
SBP_bac_6,SBP_bac_8	E	Bacterial extracellular solute-binding protein	4
AAA_11,AAA_12,NERD,Pkinase	L	AAA domain	3
Arm-DNA-bind_3,Phage_integrase	L	Belongs to the 'phage' integrase family	3
DDE_Tnp_IS66,DDE_Tnp_IS66_C,LZ_Tnp_IS66,zf-IS66	S	zinc-finger binding domain of transposase IS66	3
DEAD,Helicase_C	L	helicase superfamily c-terminal domain	3
DJ-1_Pfpl,HTH_18	K	COG4977 Transcriptional regulator containing an amidase domain and an AraC-type DNA-binding HTH domain	3
DUF3387,HSDR_N,ResIII	F	Subunit R is required for both nuclease and ATPase activities, but not for modification	3
Helicase_C,PLDc_2,SNF2_N	L	SNF2 family N-terminal domain	3
HTH_21,rve	L	Transposase and inactivated derivatives	3
Mrr_cat	V	Restriction endonuclease	3
NA37	S	37-kD nucleoid-associated bacterial protein	3
RelA_SpoT	S	Region found in RelA / SpoT proteins	3
RHH_1	S	Ribbon-helix-helix protein, copG family	3
Sel1	S	Sel1-like repeats.	3
TnpB_IS66	L	Transposase	3
AAA_15,AAA_21,ABC_tran	L	(ABC) transporter	2
AAA_19,NERD,UvrD_C_2	L	COG0210 Superfamily I DNA and RNA helicases	2
AAA_19,UvrD-helicase	L	UvrD/REP helicase N-terminal domain	2
AAA_19,UvrD-helicase,UvrD_C,UvrD_C_2	L	UvrD/REP helicase N-terminal domain	2
Acyl-CoA_dh_1,Acyl-CoA_dh_M,Acyl-CoA_dh_N	I	Acyl-CoA dehydrogenase, middle domain	2
adh_short_C2	IQ	PFAM Short-chain dehydrogenase reductase SDR	2
AIPR	S	AIPR protein	2
Aldedh	C	Aldehyde dehydrogenase family	2

CoA_transf_3	C	CoA-transferase family III	2
DUF1837	S	Domain of unknown function (DUF1837)	2
DUF3223	S	Protein of unknown function (DUF3223)	2
DUF3883	O	Molecular chaperone. Has ATPase activity	2
DUF4868	S	Domain of unknown function (DUF4868)	2
DUF726	S	Protein of unknown function (DUF726)	2
EcoEI_R_C,HSDR_N,Helicase_C,ResIII	L	EcoEI R protein C-terminal	2
GAF,HATPase_c,HisKA,PAS_2,PHY	T	PAS fold	2
HATPase_c_3	L	Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase	2
HNH	L	HNH nucleases	2
HpcH_HpaI	G	HpcH/HpaI aldolase/citrate lyase family	2
HTH_29,rve,rve_2,rve_3	L	COG2801 Transposase and inactivated derivatives	2
HTH_31	K	Helix-turn-helix XRE-family like proteins	2
HxlR	K	transcriptional regulator	2
MaoC_dehydrat_N,MaoC_dehydratas	I	N-terminal half of MaoC dehydratase	2
MUG113,T5orf172	P	T5orf172	2
NmrA	GM	NmrA family	2
PDDEXK_1,Peptidase_M78,UvrD-helicase,UvrD_C	L	Unwinds DNA duplexes with 3' to 5' polarity with respect to the bound strand and initiates unwinding most effectively when a single-stranded region is present	2
Plug,TonB_dep_Rec	P	TonB dependent receptor	2
PRiA4_ORF3	L	Plasmid pRiA4b ORF-3-like protein	2
RES	S	RES	2
Response_reg	T	response regulator	2
SIR2_2	F	SIR2-like domain	2
UvrD_C_2	L	UvrD-like helicase C-terminal domain	2
3HCDH,3HCDH_N	I	3-hydroxyacyl-coa dehydrogenase	1
AAA_11,AAA_12,DUF3320,DUF4011,DUF559	L	AAA domain	1
AAA_16,GAF,GAF_2,HATPase_c,HisKA,PAS_3,Pkinase	T	Histidine kinase	1
AAA_18,AAA_33,CPT	S	AAA domain	1
AAA_19,UvrD-helicase,UvrD_C	L	DNA helicase	1
AAA_21	S	AAA domain, putative AbiEii toxin, Type IV TA system	1

AAA_30,UvrD_C_2	L	PIF1-like helicase	1
Abhydrolase_1,Abhydrolase_6	I	Alpha beta hydrolase	1
Abhydrolase_1,Hydrolase_4	I	hydrolases or acyltransferases (alpha beta hydrolase superfamily)	1
ABM	S	Antibiotic biosynthesis monooxygenase	1
AcetylCoA_hyd_C,AcetylCoA_hydro	C	acetyl-CoA hydrolase	1
Acetyltransf_1,Acetyltransf_10,tRNA_m1G_MT	K	acetyltransferase	1
ACR_tran	V	TIGRFAM transporter, hydrophobe amphiphile efflux-1 (HAE1) family	1
Acyl-CoA_dh_1,Acyl-CoA_dh_N	I	acyl-CoA dehydrogenase	1
Adenylsucc_synt	F	Adenylosuccinate synthetase	1
ADH_N,ADH_zinc_N	C	Zinc-binding dehydrogenase	1
ADH_N,ADH_zinc_N,ADH_zinc_N_2	C	alcohol dehydrogenase	1
ADH_N_2,ADH_zinc_N	S	N-terminal domain of oxidoreductase	1
adh_short	S	Belongs to the short-chain dehydrogenases reductases (SDR) family	1
adh_short,adh_short_C2	IQ	KR domain	1
Amidohydro_2	S	Amidohydrolase	1
AMP_N,Peptidase_M24	E	Aminopeptidase P, N-terminal domain	1
ApbA,ApbA_C	H	Ketopantoate reductase PanE/ApbA C terminal	1
AraC_binding,HTH_18	K	Transcriptional regulator	1
Arginase	E	Belongs to the arginase family	1
Biotin_lipoyl_2,HlyD_3,HlyD_D23	M	Belongs to the membrane fusion protein (MFP) (TC 8.A.1) family	1
CitMHS	C	TIGRFAM citrate H symporter, CitMHS family	1
Cytochrome_CBB3	C	photosynthesis	1
DDE_3,HTH_23,HTH_32,HTH_33	L	COG3335 Transposase and inactivated derivatives	1
DDE_Tnp_1	L	PFAM Protein	1
DDE_Tnp_1,DUF772	L	COG3039 Transposase and inactivated derivatives, IS5 family	1
DDE_Tnp_1,Methylase_S	V	type I restriction modification DNA specificity domain	1
DEAD,Helicase_C,Pribosyltran,RecQ_Zn_bind,UvrD-helicase,UvrD_C	L	helicase superfamily c-terminal domain	1
DJ-1_PfpI	S	DJ-1/PfpI family	1

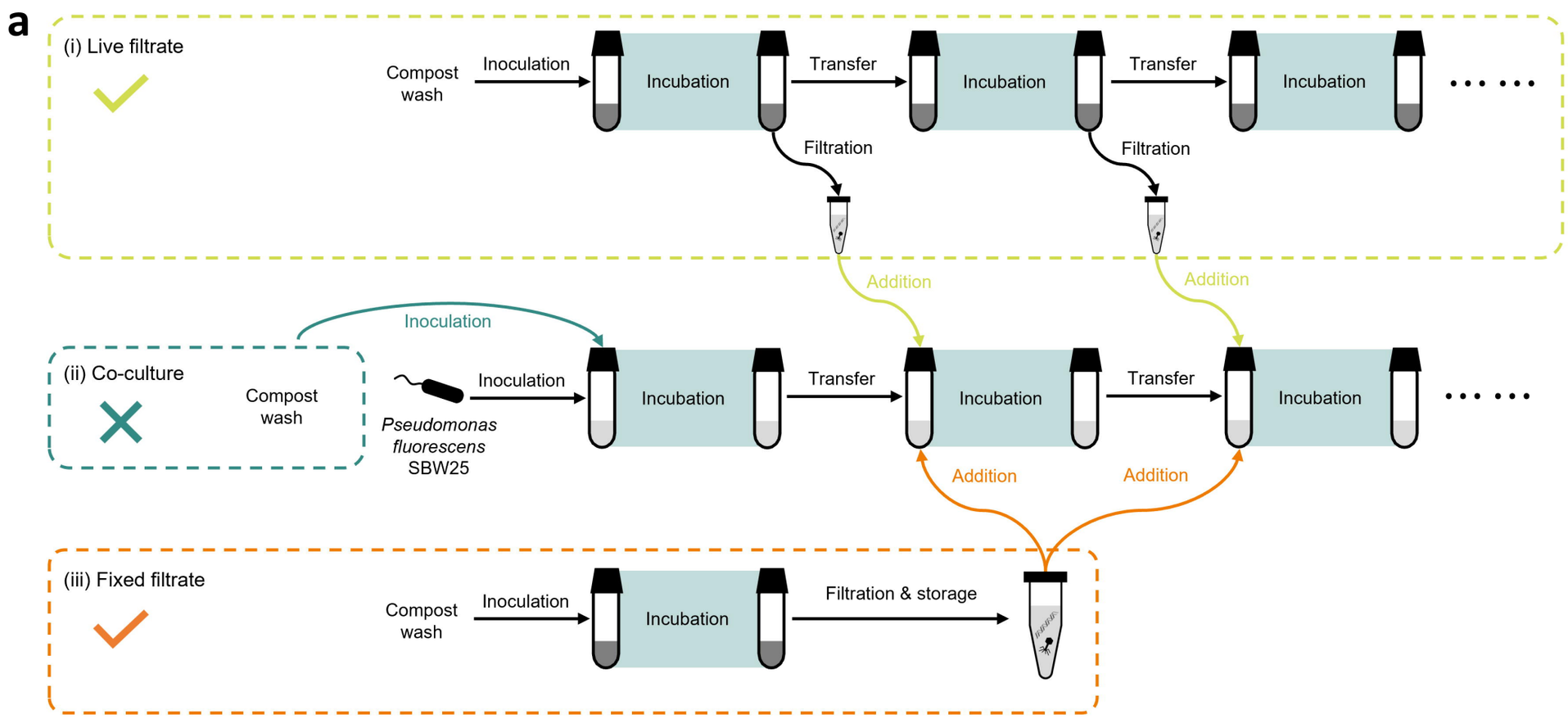
DoxX_2	S	DoxX-like family	1
DUF1302	M	Protein of unknown function (DUF1302)	1
DUF2236	S	Uncharacterized protein conserved in bacteria (DUF2236)	1
DUF2867	S	Protein of unknown function (DUF2867)	1
DUF2938	S	Protein of unknown function (DUF2938)	1
DUF2946	S	Protein of unknown function (DUF2946)	1
DUF4113,IMS,IMS_C,IMS_HHH	L	DNA polymerase	1
DUF4420	S	Putative PD-(D/E)XK family member, (DUF4420)	1
DUF928,SIR2_2	F	5-carbamoylmethyl uridine residue modification	1
Dynamin_N,MMR_HSR1,Patatin	M	Esterase of the alpha-beta hydrolase superfamily	1
EcoEI_R_C,Helicase_C,ResIII	L	EcoEI R protein C-terminal	1
EcoRII-C	L	EcoRII C terminal	1
ETF	C	Electron transfer flavoprotein	1
ETF,ETF_alpha	C	Electron transfer flavoprotein	1
ETF_QO,FAD_binding_2,NAD_binding_8,Thi4	C	Electron transfer flavoprotein-ubiquinone oxidoreductase	1
Fe-ADH	C	Iron-containing alcohol dehydrogenase	1
Fer4_12,Fer4_14,Radical_SAM	S	Radical SAM superfamily	1
Fer4_12,Radical_SAM	S	Radical SAM superfamily	1
Flg_new,LRR_5,LRR_6	S	regulation of response to stimulus	1
GAF,GAF_2,GGDEF,PAS_3	T	GAF PAS GGDEF domain-containing protein	1
GerE	K	helix_turn_helix, Lux Regulon	1
GerE,Response_reg,Trans_reg_C	K	response regulator	1
GntR,UTRA	K	UTRA	1
Histone_HNS	-	-	1
HTH_11,WYL	K	HTH domain	1
HTH_18	K	AraC Family	1
HTH_20	K	transcriptional regulator	1
HTH_21,rve,rve_3	L	PFAM Integrase catalytic region	1
HTH_23,HTH_24,rve	L	Transposase and inactivated derivatives	1
HTH_26	K	transcriptional regulator	1
HTH_3,Peptidase_M78	K	IrrE N-terminal-like domain	1

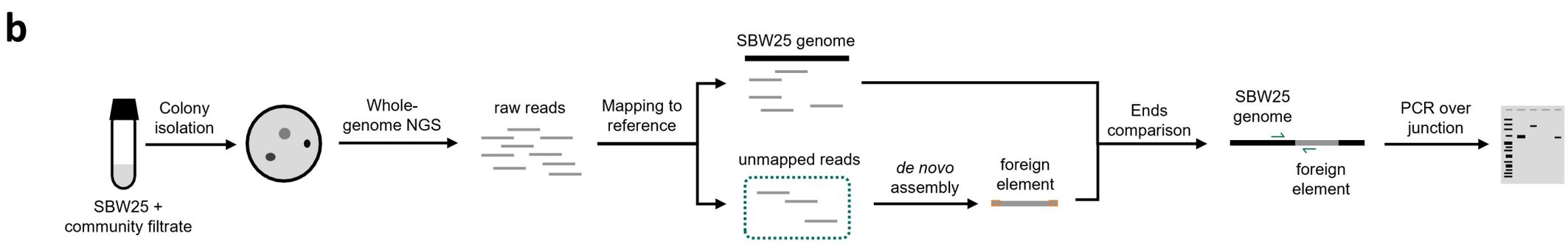
HTH_IclR,IclR	K	transcriptional regulator	1
Integrase_1,Phage_int_SAM_2	L	Belongs to the 'phage' integrase family	1
Lactamase_B	S	Metallo-beta-lactamase superfamily	1
Lumazine_bd_2	S	Putative lumazine-binding	1
LysR_substrate	K	LysR substrate binding domain	1
M20_dimer,Peptidase_M20	S	Peptidase dimerisation domain	1
MaoC_dehydratas	I	MaoC like domain	1
MS_channel,cNMP_binding	MT	mechanosensitive ion channel	1
N6_N4_Mtase,TypeIII_RM_meth	L	DNA methylase	1
NAD_binding_10,NmrA	GM	NmrA-like family	1
NERD,zf-C4_Topoisom	L	COG0551 Zn-finger domain associated with topoisomerase type I	1
NUDIX	L	NUDIX domain	1
OEP	M	RND efflux system, outer membrane lipoprotein	1
OmpA	N	Flagellar Motor Protein	1
OsmC	O	response to oxidative stress	1
Oxidored_FMN	C	PFAM NADH flavin oxidoreductase NADH oxidase	1
PCuAC	S	Copper chaperone PCu(A)C	1
PDDEXK_4	S	PD-(D/E)XK nuclease superfamily	1
PKD,Tannase	C	PFAM blue (type 1) copper domain protein	1
Prok-JAB	S	Prokaryotic homologs of the JAB domain	1
PrpF	P	protein conserved in bacteria	1
Radical_SAM	C	Elongator protein 3, MiaB family, Radical SAM	1
Radical_SAM,TRAM,UPF0004	J	Catalyzes the methylthiolation of an aspartic acid residue of ribosomal protein S12	1
RloB	S	RloB-like protein	1
RraA-like	H	Catalyzes the aldol cleavage of 4-hydroxy-4-methyl-2- oxoglutarate (HMG) into 2 molecules of pyruvate. Also contains a secondary oxaloacetate (OAA) decarboxylase activity due to the common pyruvate enolate transition state formed following C-C bond cleavage in the retro-aldol and decarboxylation reactions	1
SCO1-SenC	S	SCO1/SenC	1
SIR2	K	NAD-dependent lysine deacetylase and desuccinylase that specifically removes acetyl and succinyl groups on target proteins. Modulates the activities of several proteins which are inactive in their acylated form	1
TauD	Q	Taurine catabolism dioxygenase TauD, TfdA family	1
Tautomerase_2	S	tautomerase	1
TetR_C_13,TetR_N	K	transcriptional regulator	1
TetR_N	K	Bacterial regulatory proteins, tetR family	1

ThiF	H	ThiF family	1
Thiolase_C,Thiolase_N	I	Acetyl-CoA acetyltransferase	1
Thioredoxin	O	OST3 / OST6 family, transporter family	1
Thymidylate_kin	F	Phosphorylation of dTMP to form dTDP in both de novo and salvage pathways of dTTP synthesis	1
Transposase_mut	L	Transposase	1
tRNA_anti-like	S	tRNA_anti-like	1
UvrD-helicase,UvrD_C	L	UvrD-like helicase C-terminal domain	1
Z1	L	Z1 domain	1

1055

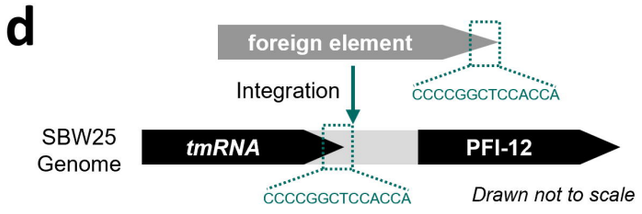
1056

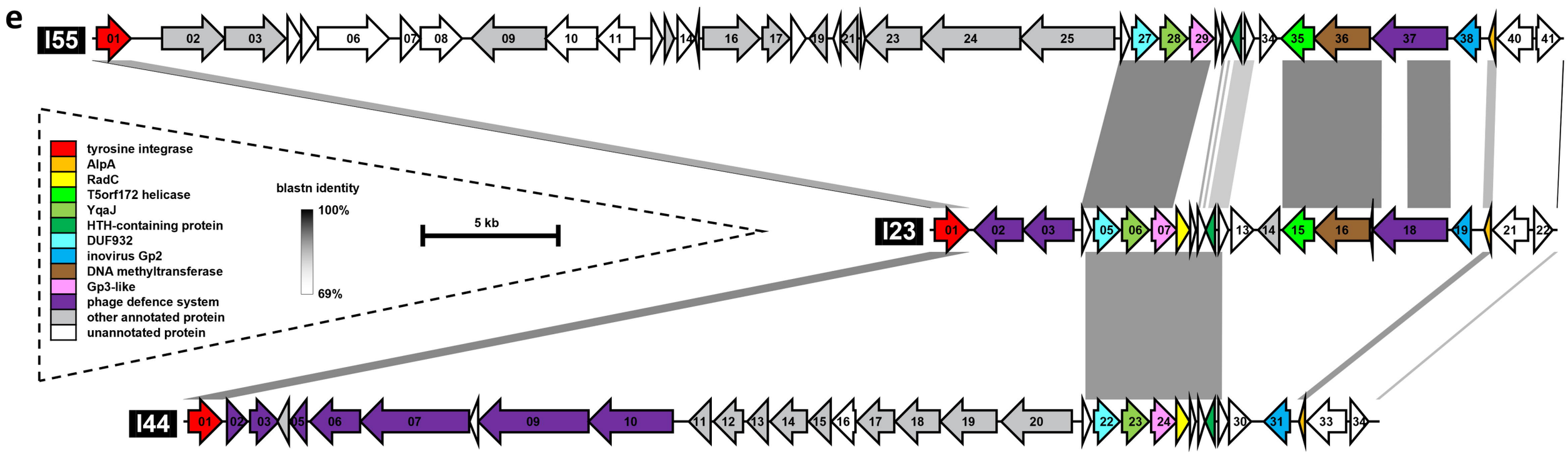


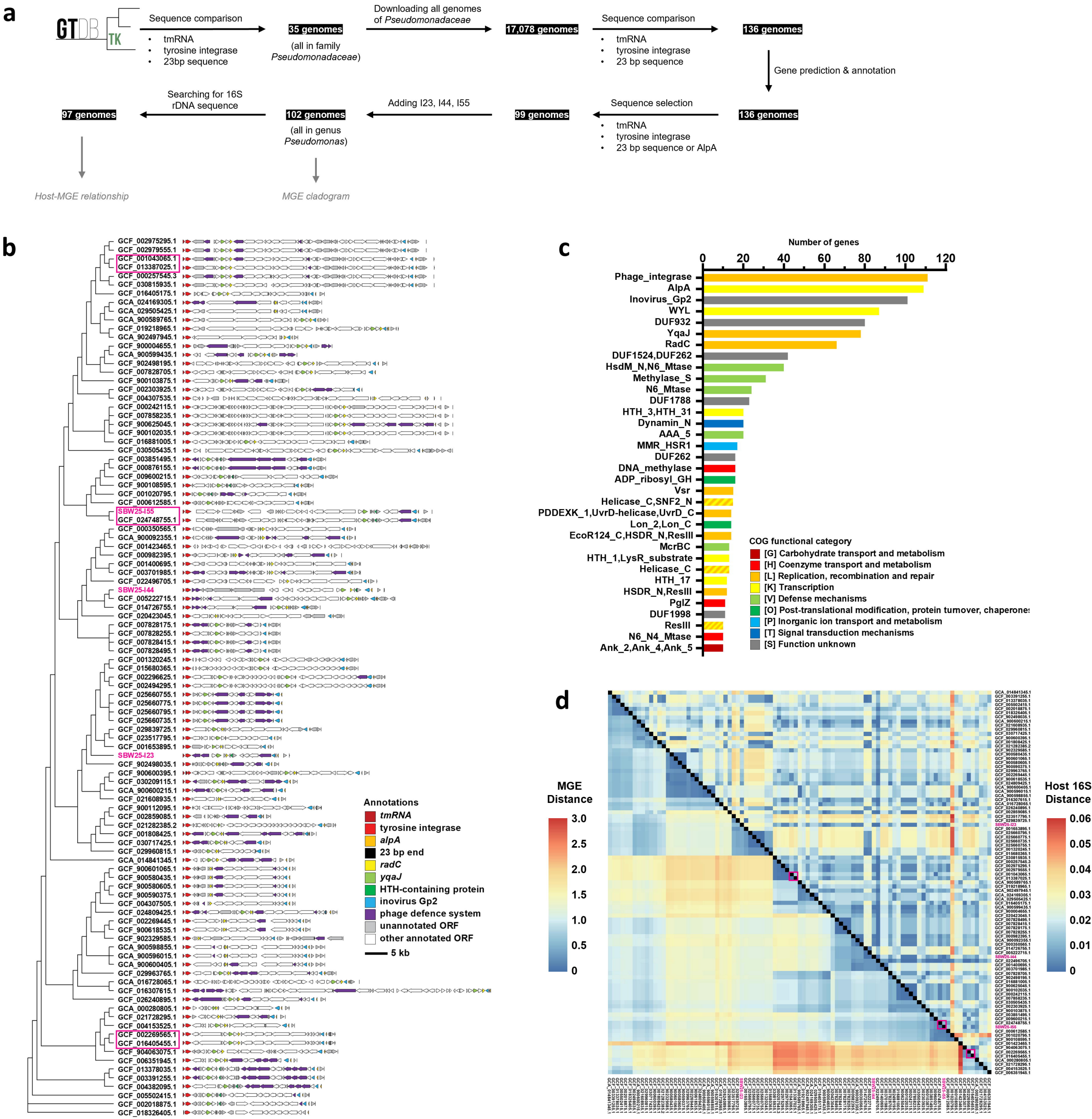


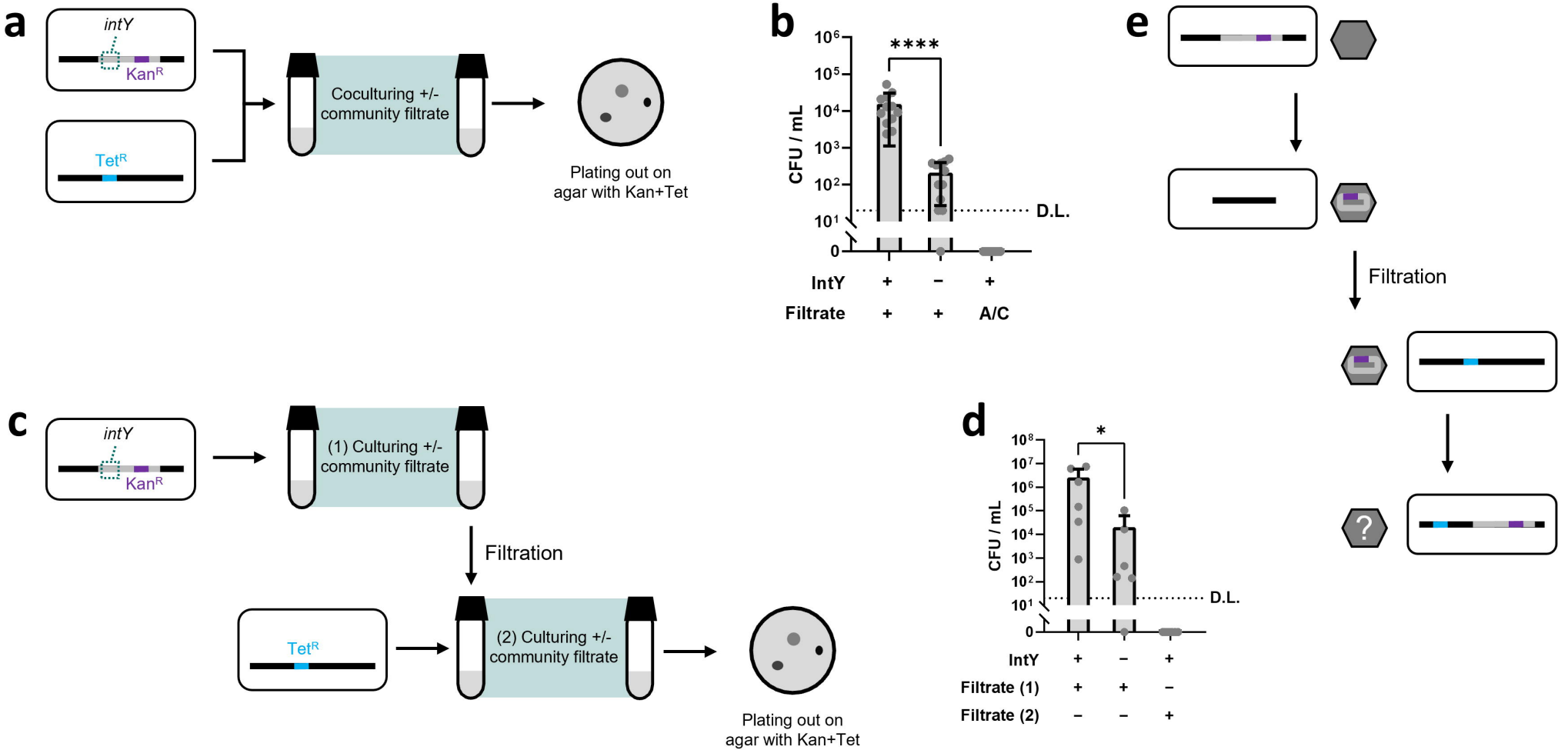
c

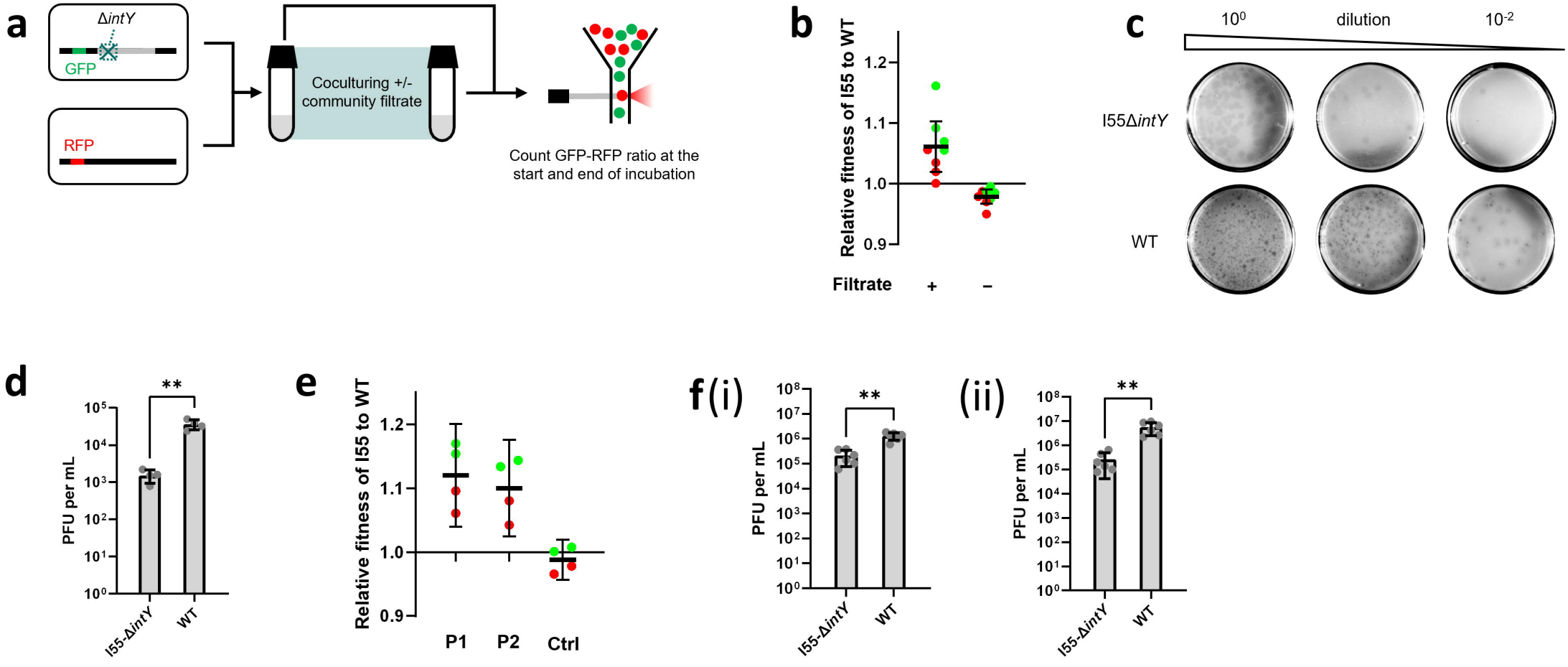
Foreign element	Length / bp	Community filtrate	Regime	Frequency
I23	23,199	Compost B	Fixed filtrate	3 lines in 10 lines (1, 1 and 4 out of 24 isolates)
I44	44,236	Compost B	Fixed filtrate	1 line in 10 lines (2 out of 24 isolates)
I55	54,419	Compost C	Live filtrate	1 line in 8 lines (1 out of 3 isolates)

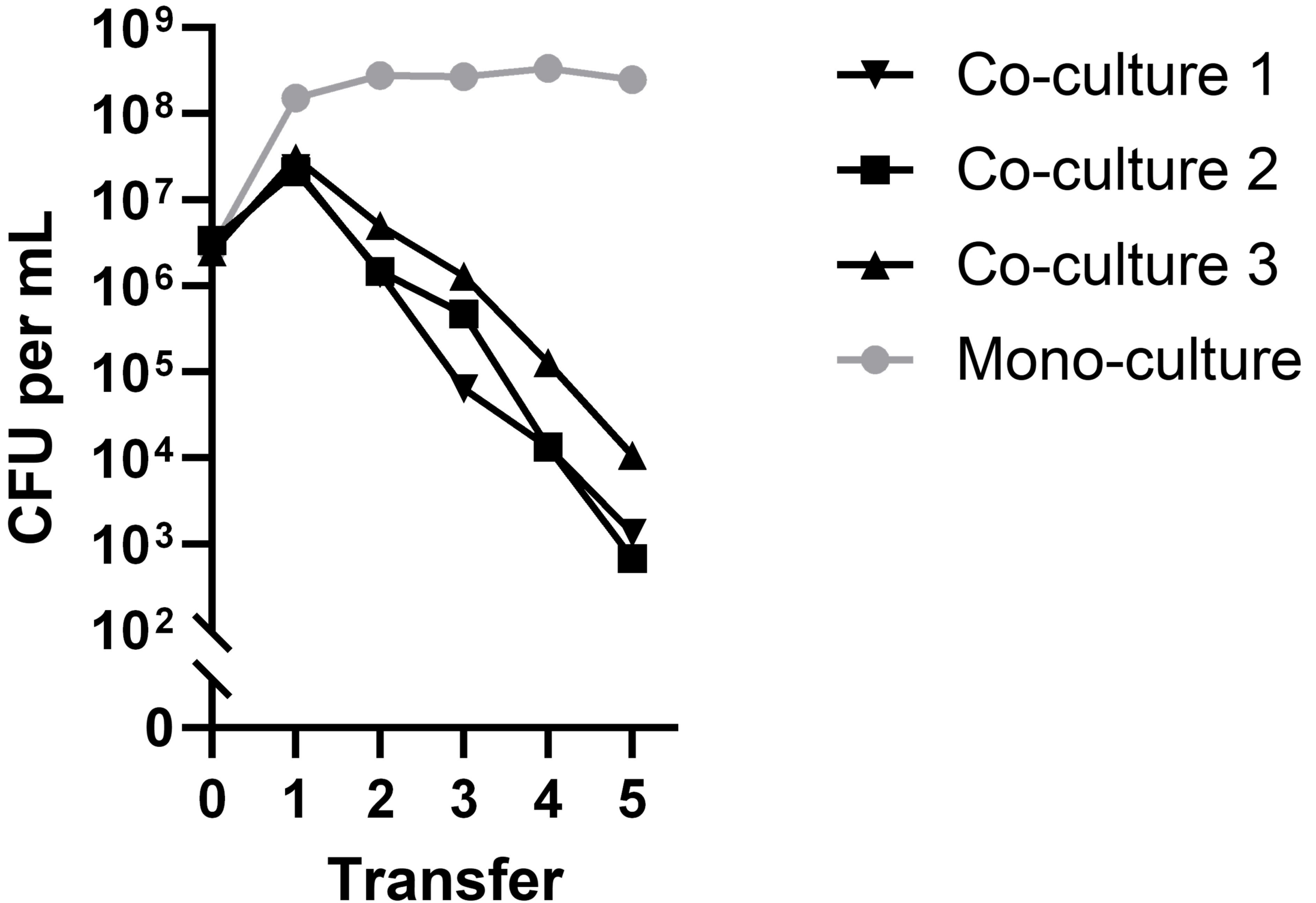


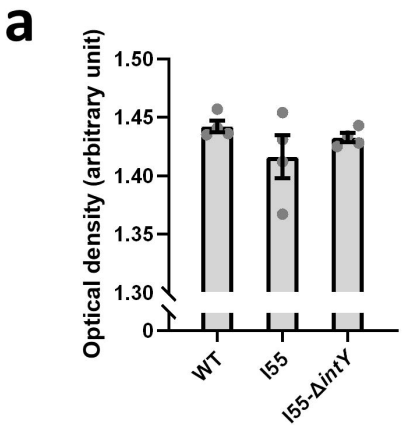


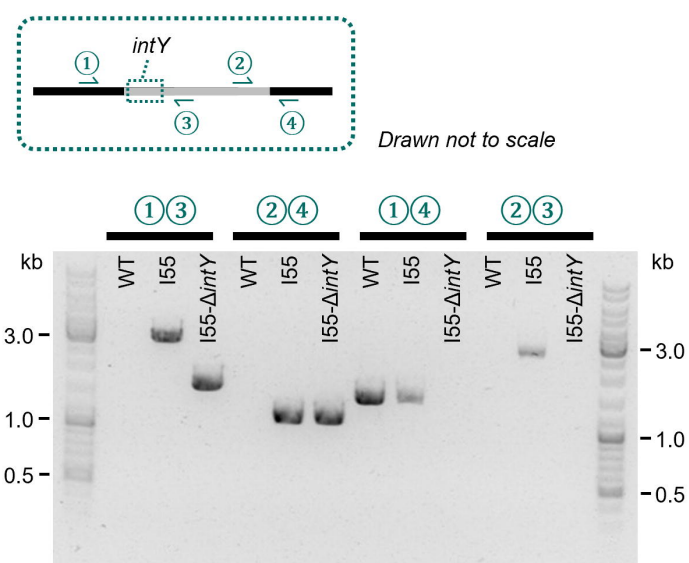

a
GTDB TK
Sequence comparison
tmRNA
tyrosine integrase
23bp sequence
35 genomes
(all in family Pseudomonadaceae)
Downloading all genomes of Pseudomonadaceae
17,078 genomes
Sequence comparison
tmRNA
tyrosine integrase
23 bp sequence
136 genomes
Gene prediction & annotation
136 genomes
Sequence selection
tmRNA
tyrosine integrase
23 bp sequence or AlpA
99 genomes
Adding I23, I44, I55
102 genomes
(all in genus Pseudomonas)
Searching for 16S rDNA sequence
97 genomes
Host-MGE relationship
MGE cladogram
b
SBW25-I55
SBW25-I44
SBW25-I23
Annotations
tmRNA
tyrosine integrase
alpA
23 bp end
radC
yqaJ
HTH-containing protein
inovirus Gp2
phage defence system
unannotated ORF
other annotated ORF
5 kb
c
Number of genes
0
20
40
60
80
100
120
Phage_integrase
AlpA
Inovirus_Gp2
WYL
DUF932
YqaJ
RadC
DUF1524,DUF262
HsdM_N,N6_Mtase
Methylase_S
N6_Mtase
DUF1788
HTH_3,HTH_31
Dynamin_N
AAA_5
MMR_HSR1
DUF262
DNA_methylase
ADP_ribosyl_GH
Vsr
Helicase_C,SNF2_N
PDDEXK_1,UvrD-helicase,UvrD_C
Lon_2,Lon_C
EcoR124_C,HSDR_N,ResIII
McrBC
HTH_1,LysR_substrate
Helicase_C
HTH_17
HSDR_N,ResIII
PglZ
DUF1998
ResIII
N6_N4_Mtase
Ank_2,Ank_4,Ank_5
COG functional category
[G] Carbohydrate transport and metabolism
[H] Coenzyme transport and metabolism
[L] Replication, recombination and repair
[K] Transcription
[V] Defense mechanisms
[O] Post-translational modification, protein turnover, chaperones
[P] Inorganic ion transport and metabolism
[T] Signal transduction mechanisms
[S] Function unknown
d
MGE Distance
3.0
2.5
2.0
1.5
1.0
0.5
0
Host 16S Distance
0.06
0.05
0.04
0.03
0.02
0.01
0

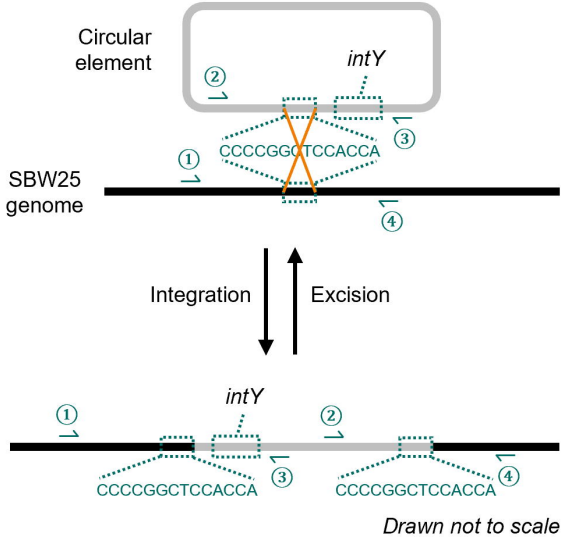

a
intY
KanR
TetR
Coculturing +/-
community filtrate
Plating out on
agar with Kan+Tet
b
****
CFU / mL
D.L.
IntY + − +
Filtrate + + A/C
c
intY
KanR
(1) Culturing +/-
community filtrate
Filtration
TetR
(2) Culturing +/-
community filtrate
Plating out on
agar with Kan+Tet
d
*
CFU / mL
D.L.
IntY + − +
Filtrate (1) + + −
Filtrate (2) − − +
e
Filtration
?


a
ΔintY
GFP
RFP
Coculturing +/-
community filtrate
Count GFP-RFP ratio at the
start and end of incubation
b
Relative fitness of I55 to WT
1.2
1.1
1.0
0.9
Filtrate
+
–
c
10⁰
dilution
10⁻²
I55ΔintY
WT
d
**
PFU per mL
I55-ΔintY
WT
e
Relative fitness of I55 to WT
P1
P2
Ctrl
f (i)
(ii)
**
PFU per mL
I55-ΔintY
WT


CFU per mL
10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
0
0
1
2
3
4
5
Transfer
Co-culture 1
Co-culture 2
Co-culture 3
Mono-culture


a
Optical density (arbitrary unit)
1.50
1.45
1.40
1.35
1.30
0
WT
I55
I55-ΔintY


b
intY
①
②
③
④
Drawn not to scale
①③
②④
①④
②③
kb
WT
I55
I55-ΔintY
3.0
1.0
0.5


c
Circular element
intY
CCCCGGCTCCACCA
SBW25 genome
Integration
Excision
Drawn not to scale