The Pathfinder plasmid toolkit for genetically engineering newly isolated bacteria enables the study of *Drosophila*-colonizing *Orbaceae*

Katherine M. Elston[a], Laila E. Phillips[b], Sean P. Leonard[a,b], Eleanor Young[a], Jo-anne C. Holley[b,c], Tasneem Ahsanullah[c], Braydin McReynolds[a], Nancy A. Moran[b], Jeffrey E. Barrick[a#]

[a]Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX 78712, USA

[b]Department of Integrative Biology, The University of Texas at Austin, Austin, TX 78712, USA

[c]Freshman Research Initiative, The University of Texas at Austin, Austin, TX 78712, USA

Running Head: Genetic engineering of *Drosophila*-colonizing *Orbaceae*

#Address correspondence to Jeffrey E. Barrick, jbarrick@cm.utexas.edu

Katherine M. Elston and Laila E. Phillips contributed equally to this work. Author order was determined on the basis of seniority.

ABSTRACT

Toolkits of plasmids and genetic parts streamline the process of assembling DNA constructs and engineering microbes. Many of these kits were designed with specific industrial or laboratory microbes in mind. For researchers interested in non-model microbial systems, it is often unclear which tools and techniques will function in newly isolated strains. To address this challenge, we designed the Pathfinder toolkit for quickly determining the compatibility of a bacterium with different plasmid components. Pathfinder plasmids combine three different broad-host-range origins of replication with multiple antibiotic resistance cassettes and reporters, so that sets of parts can be rapidly screened through multiplex conjugation. We first tested these plasmids in *Escherichia coli*, a strain of *Sodalis praecaptivus* that colonizes insects, and a *Rosenbergiella* isolate from leafhoppers. Then, we used the Pathfinder plasmids to engineer previously unstudied bacteria from the family *Orbaceae* that were isolated from several fly species. Engineered *Orbaceae* strains were able to colonize *Drosophila melanogaster* and could be visualized in fly guts. *Orbaceae* are common and abundant in the guts of wild-caught flies but have not been included in laboratory studies of how the *Drosophila* microbiome affects fly health. Thus, this work provides foundational genetic tools for studying new host-associated microbes, including bacteria that are a key constituent of the gut microbiome of a model insect species.

IMPORTANCE To fully understand how microbes have evolved to interact with their environments, one must be able to modify their genomes. However, it can be difficult and laborious to discover which genetic tools and approaches work for a new isolate.

Bacteria from the recently described *Orbaceae* family are common in the microbiomes of insects. We developed the Pathfinder plasmid toolkit for testing the compatibility of different genetic parts with newly cultured bacteria. We demonstrate its utility by engineering *Orbaceae* strains isolated from flies to express fluorescent proteins and characterizing how they colonize the *Drosophila melanogaster* gut. *Orbaceae* are widespread in *Drosophila* in the wild but have not been included in laboratory studies examining how the gut microbiome affects fly nutrition, health, and longevity. Our work establishes a path for genetic studies aimed at understanding and altering interactions between these and other newly isolated bacteria and their hosts.

INTRODUCTION

Researchers have isolated and sequenced many new microbes from different ecosystems and from diverse plant and animal hosts. To characterize these microbes and study how they interact with their physical environments and with other organisms, one needs genetic tools. However, most described bacterial species have never been genetically manipulated (1–3). The primary obstacle in many cases is likely that the requisite resources and know-how for microbial genetic engineering are not easily accessible to researchers who encounter non-model microbes.

Toolkits of genetic parts have been developed for molecular microbiology and synthetic biology. These kits, such as the modular cloning (MoClo) and Standard European Vector Architecture (SEVA) toolkits (4, 5), aim to be flexible and comprehensive. Their collections of interchangeable parts include promoters with a range of different expression levels and multiple reporter genes and plasmid backbones

that can be combined to assemble a genetic construct of interest (4, 6–11). While these systems facilitate complex genetic engineering tasks in well-studied laboratory and industrial bacteria, such as *Escherichia coli* and *Pseudomonas putida*, there are still gaps with respect to their applicability to all bacteria. For example, they may only have a few antibiotic resistance cassettes or rely on plasmids that replicate only in specific species. Researchers may also find assembling new plasmids according to the schemes in these kits daunting and overly cumbersome if all they want to achieve are basic tasks like expressing a single protein in a new bacterial species.

Fluorescent protein expression alone is often enough to investigate aspects of host-microbe interactions, such as bacterial localization, or to track a strain of interest within a microbiome or environmental community. However, even this rudimentary genetic modification can be challenging (1, 3, 12). A reasonable first step is to start with a broad-host-range plasmid that has been reported to replicate in diverse bacteria, but one must still empirically test whether one of these plasmids is compatible with each new species (13). Additionally, electroporation or chemical treatments to transform plasmids into cells do not work in all bacteria (14, 15). Developing these techniques through trial and error can be frustrating and time-consuming, particularly when one does not know if the plasmid being used will successfully replicate after transformation. Conjugation is often a more reliable method for delivering DNA to non-model bacteria (2, 16), and it has been incorporated into several kits that focus on engineering a wider range of bacteria, such as the Bacterial Expression Vector Archive (BEVA) (9), the bee microbiome toolkit (BTK) (6), and the Proteobacteria toolbox (8). The wide phylogenetic

distribution of natural conjugative and transmissible broad-host-range plasmids suggests that this approach should work for many bacterial species (17–19).

Although *Drosophila melanogaster* has been a model organism for genetics for over a century, research focused on its gut microbiome is a relatively new field (20–22). Laboratory studies have focused primarily on *Acetobacter* and *Lactobacillus* species (23–25), which make up only a fraction of the microbiome that is normally present in wild *Drosophila* (26). A large percentage of the natural *Drosophila* microbiome is composed of bacteria in the recently-described *Orbaceae* family (27, 28). *Orbaceae* are prevalent in a wide variety of insects (28–31) and are observed in 16S rRNA gene surveys of populations of laboratory-reared and wild flies of different species, including *D. melanogaster* (26, 29, 32–34). How *Orbaceae* colonize and interact with their hosts is relatively unexplored despite how prevalent they are in insect microbiomes.

Here we describe the Pathfinder plasmid system, a simple and robust toolkit for engineering newly cultured bacteria. First, we show how multiplex conjugation with defined subsets of Pathfinder plasmids can be used to quickly determine the compatibility of bacteria with different genetic parts. Then, we then use the Pathfinder plasmids to engineer recently cultured *Orbaceae* isolates from flies and characterize how they colonize the *D. melanogaster* gut.

RESULTS

Pathfinder plasmid toolkit design. An overview of the Pathfinder plasmid design and procedure is shown in Figure 1A and Table 1. Plasmids pSL1, pSL9, and pSL25 have 3 different origins of replication (RSF1010, pBBR1, and RP4) paired with 3

different reporters (red chromoprotein, RCP; E2-Crimson, E2C; and blue chromoprotein, BCP), respectively, along with kanamycin resistance. The reporter genes are all expressed from the broad-host-range CP25 promoter. Plasmids pSL1-7 all have an RSF1010 origin and red chromoprotein expression with one of 7 different antibiotic resistances (in order): kanamycin (Kan^R), spectinomycin ($Spec^R$), gentamicin ($Gent^R$), chloramphenicol (Cam^R), erythromycin (Ery^R), tetracycline (Tet^R), and ampicillin/carbenicillin ($Carb^R$). All Pathfinder plasmids are also BsaI dropout vectors (Type 8 parts) compatible with stage 1 of the Golden Gate assembly scheme used by the yeast and bee microbiome toolkits (6, 35). Thus, the Pathfinder plasmids can be readily reconfigured to convey and express DNA sequences other than the included reporter genes. For example, we created pSL1-GFP to express GFP instead of RCP from the same backbone as pSL1 in this way.

To confirm the functionality of the Pathfinder plasmids, we performed an initial test with *E. coli.* All plasmids except pSL5 (RSF1010, RCP, Ery^R) were transformed into the *E. coli* donor strain MFD*pir* (36), then combined equally into a mix that could be frozen down and thawed as needed for conjugation. We were able to recover *E. coli* transconjugants using this mix for every plasmid except for pSL3 (RSF1010, RCP, $Gent^R$). Colonies of *E. coli* cells containing Pathfinder plasmids expressing each of the three reporters or GFP can be identified by eye (Fig. 1B). Fluorescence can also be used to distinguish colonies with different plasmids from one another (Fig. 1B), which might be useful if the markers are expressed at lower levels in other species. To demonstrate functionality in a more distantly-related bacterium, we tested conjugation of the Pathfinder mix into *Sodalis praecaptivus* HS^T, a human wound bacterial isolate

previously shown to colonize weevils and tsetse flies (37–39). As with *E. coli*, we were able to recover transconjugants for every plasmid except for pSL3. This negative result may be due to inherent gentamicin resistance of the MFD*pir* strain interfering with conjugation. We also established that Pathfinder plasmids function in a new *Rosenbergiella* bmE01 strain we isolated from *Empoasca* leafhoppers. For bmE01, we tested only pSL1-GFP and pSL7 and successfully isolated transconjugants for both.

Pathfinder plasmids function in recently isolated fly symbionts. We then applied the Pathfinder plasmid system to a set of bacteria that we isolated from wild flies (members of order Diptera), along with an isolate, *Orbus hercynius* CN3, collected by Volkmann *et al.* (27), which likely originates from a non-*Drosophila* dipteran species breeding in boar feces. Based on phylogenies constructed using 16S rRNA genes, all of these isolates are closely related and belong to the *Orbaceae* family within the *Gammaproteobacteria*, which includes symbionts of bees and other insects (Fig. 2A).

We were able to successfully conjugate Pathfinder plasmids into each of the *Orbaceae* (Fig. 2B). In terms of origin of replication compatibility, we only observed conjugation with the pSL1 and pSL1-GFP plasmids that contain the RSF1010 origin. The pBBR1 and RP4 origin plasmids were absent from our *Orbaceae* transconjugant plates. For the antibiotic resistance panel, we achieved conjugation with all plasmids other than pSL6 (TetR) in most strains. Differences arose when strains had elevated levels of intrinsic resistance to an antibiotic (Table 2). For instance, lpD01 was highly resistant to Spec and Cam which prevented us from isolating pSL2 and pSL4 transconjugants, respectively, because untransformed cells grew on these selective plates. Rates of conjugation with the Pathfinder plasmids were similar across strains

and comparable to the rates observed for *E. coli* for many combinations of plasmids and strains (Fig. 2C). The overall average conjugation efficiency for all the *Orbaceae* strains was 1.4 ± 0.02%. BiB had the highest average conjugation efficiency (2.3 ± 0.1%), while lpD03 had the lowest (0.0026 ± 0.0002%). Despite the lower conjugation efficiencies observed for some *Orbaceae* strains, we demonstrated that each of these newly isolated bacteria can be engineered with several plasmids from the Pathfinder series.

As a basic test for functionality of the engineered constructs in the uncharacterized *Orbaceae* strains, we measured fluorescence levels from pSL1-GFP conjugants. We observed significant variation in GFP fluorescence among these strains ($F_{4,40}$ = 322.8, p < 0.0001). In particular, lpD01 had such low GFP signal that it was difficult to detect by eye that it fluoresced on a blue light transilluminator (Fig. 2D), but even this weak GFP signal was sufficient for further studies of this strain (see below).

***D. melanogaster* colonization by engineered *Orbaceae*.** We next used our engineered strains to determine if fly-associated *Orbaceae* can colonize *D. melanogaster*. We colonized conventionally reared Canton-S flies using a method we refer to as arena inoculation, in which flies were kept in a container along with an agar plate of fly diet grown with a lawn of one of our strains (see Methods for additional details). We expected that flies would ingest the bacteria while feeding. After 24 hours of inoculation, we transferred flies to fresh diet every 24 hours to eliminate bacteria that survived on the diet rather than within the flies. To ascertain whether bacteria could persist by replicating on the diet itself, we confirmed lack of growth on the yeast-glucose agar fly diet. No growth was observed on the fly diet for strains lpD01, lpD02, and lpD03. BiB and *O. hercynius* CN3 had light growth after 4-5 days, which could

potentially be a source of fly recolonization after the initial inoculation arena. Because we did not perform our inoculation with germ-free flies, we anticipated that other microbes in the fly gut might complicate our colonization experiment. However, a benefit to performing our assay with engineered fluorescent bacteria that carry an antibiotic resistance marker is that we can easily identify and select for our strain of interest within a microbiome containing other microbes.

At several time points after inoculation, we washed and crushed 5-6 flies and plated them on selective media with kanamycin. We found that each of these strains can colonize flies to some extent, in contrast to the bee-associated *Orbaceae, Gilliamella apicola* M1-2G, which was not retained at any time point (Fig. 3). Colonization of the fly-derived strains varied between time points and between individual flies in these initial tests. The most consistent findings were that lpD01 was able to robustly colonize flies at every time point, while lpD03 and BiB were lost after day 2 and 4, respectively.

Based on these preliminary results, we decided to track bacterial titer over time in flies colonized with lpD02 and lpD01. To account for variation among experimental populations of flies, we inoculated three separate arenas of flies per trial. The overall trend for lpD02 colonization follows the pattern observed in the qualitative experiment. Over time, the number of colony-forming units (CFU) per fly gradually decreases until day 11, when most flies are no longer colonized (Fig. 4A). However, for lpD01, the average CFU in each arena decreases then increases almost to the initial level seen on day 0 (Fig. 4B). Between day 0 and day 4, average CFU drops (arena 1, $p = 0.001$, arena 2, $p = 0.0045$, arena 3, $p < 0.0001$, Dunn's test with Bonferroni correction), and

most sampled flies were uncolonized on day 4 in arenas 1 and 3. The CFU per fly then increases between day 4 and day 11 for arenas 1 and 3 (p = 0.0022 and p = 0.0001, respectively, Dunn's test with Bonferroni correction). The increase between day 4 and day 11 was not significant for arena 2 (p = 1, Dunn's test with Bonferroni correction), but appears to show a similar trend with a slight temporal delay.

Visualizing lpD01 in the gut of *D. melanogaster.* Based on the results of the bacterial titer assay, we selected lpD01 to visualize colonization of the fly gut by *Orbaceae*. We inoculated 100 flies with the engineered lpD01 + pSL1-GFP strain and reared them for 11 days on fresh diet, replicating the quantification experiment. After this point we dissected the *D. melanogaster* gut and used confocal fluorescence microscopy to assess bacterial localization (Fig. 5). We observed the presence of fluorescent lpD01 in the proventriculus (cardia) of every imaged fly (Fig. 5C, D), and for 2 out of 5 flies lpD01 could also be found colonizing the crop (Fig. 5G, H). These locations are consistent with where gut-associated *Acetobacteraceae* and *Lactobacillus* strains have been observed in *D. melanogaster* (40–42). Bacterial aggregates were present in both crop and cardia (Fig. 5E), suggesting active replication. Throughout the remainder of the gut, we occasionally observed fluorescent cells (Fig. 5J, K), but lpD01 did not robustly colonize the midgut or hindgut regions.

DISCUSSION

The Pathfinder plasmid system provides a simple toolkit with a straightforward methodology for genetically modifying non-model bacteria. With this kit we successfully engineered several *Orbaceae* strains isolated from wild flies and utilized their GFP

226 expression and antibiotic resistance cassettes to facilitate colonization experiments in
227 *D. melanogaster*. To our knowledge, our study is the first to demonstrate that *D.*
228 *melanogaster* can be experimentally colonized with natural fly symbionts from the
229 *Orbaceae* family.

230 Our results illustrate a common obstacle in synthetic biology: not all plasmid
231 components work well in all bacteria (12). For example, we did not recover *Orbaceae*
232 transconjugants carrying either the RP4 or pBBR1 origins, despite reports that these
233 origins function in a wide range of bacteria in other studies (13, 43). In terms of
234 antibiotic resistance cassette compatibility, our results also matched our expectation
235 that each resistance cassette would not function in every strain. The incompatible
236 strains were mostly those with intrinsic resistance to specific antibiotics. However, the
237 basis of the incompatibility of the Tet resistance cassette with all *Orbaceae* tested is
238 unclear since most were sensitive to low levels of Tet. Possibly, the Tet resistance gene
239 on the Pathfinder plasmids does not express or function effectively in *Orbaceae*. Our
240 varied results highlight the utility of widely surveying for plasmid component functionality
241 when first working with a new strain. The information from Pathfinder informs the
242 selection of plasmids for future experiments.

243 Many other genetic toolkits have been developed with a similar goal of
244 engineering wild bacteria (9, 16, 44), and Pathfinder has fewer components than other
245 kits. We prioritized building out the complete antibiotic set with RSF1010 because of this
246 origin of replication's wide compatibility with different bacteria (45, 46). RSF1010
247 worked well in the *Orbaceae* strains, but it may not replicate in other bacteria. The host
248 ranges of many plasmid origins of replication that function in non-model bacterial

249 species have not yet been exhaustively surveyed. In the future, the kit might be
250 improved by including additional plasmid origins and pairing each of these with the
251 entire antibiotic resistance set. The current Pathfinder toolkit also relies on a single
252 promoter to drive all reporter genes. One could expand the combinations to include
253 different promoters to rapidly survey their functions in a new bacterium, for example by
254 performing multiplex conjugation and then picking the most highly fluorescent or colored
255 colonies. Combinatorial Golden Gate assembly schemes could be used to create sets
256 of plasmids with new combinations of these components.

257 The Pathfinder kit is limited to plasmid-based expression systems. Plasmid
258 transformation is a common first step in engineering a new bacterium, but it is not ideal
259 for ensuring the long-term stability of engineered constructs. Multicopy plasmids can be
260 especially burdensome when their gene products divert resources from host cell
261 replication (47, 48). Because of this, plasmid-based systems tend to be more likely to
262 rapidly lose function due to the takeover of cells with mutated plasmids that alleviate this
263 burden by inactivating engineered functions, as compared to systems engineered into
264 the chromosome (49). Another potential complication is rapid plasmid loss from a cell
265 population due to segregation in the absence of antibiotic selection. Our colonization
266 experiments in *D. melanogaster* used constant antibiotic selection to prevent plasmid
267 loss. However, administering sufficient levels of antibiotics for selection may be
268 challenging in other environments. Tools designed for chromosomal integration such as
269 transposon systems may be a better option for researchers with these concerns (44).
270 Conjugation has been used to deliver DNA to engineer bacteria *in situ* in gut and soil
271 communities (2, 16, 50). The Pathfinder plasmids are compatible with this approach and

could potentially be used to engineer bacteria that are currently unculturable outside of their hosts (1, 51).

The fluorescent *Orbaceae* strains that we built enabled us to easily screen for effective colonization of *D. melanogaster*. We observed differences among isolates in their abilities to colonize flies, as well as variable colonization levels among individual flies inoculated with the same isolate and between the preliminary trials and follow-up experiments with the lpD01 and lpD02 isolates. We colonized non-axenic flies to emulate invasion conditions in which other microbes are already present in the fly gut (40). Differences in the established gut communities of the cohorts of flies that we used could explain some of the variation in our results (41). Using germ-free flies would eliminate these effects (52). Since our flies were housed in a laboratory setting, their diets and microbiomes do not necessarily reflect wild conditions (40), which could also impact the success of *Orbaceae* relative to other gut-associated species of bacteria.

We observed fluorescent *Orbaceae* cells in the crop and proventriculus regions of the foregut in a majority of the flies we colonized with lpD01. The *Drosophila* foregut tends to be more hospitable for bacterial colonization than the midgut, which has a lower pH and undergoes peristalsis along with continual turnover of the peritrophic membrane (21, 41). Accordingly, we observed very few lpD01 cells in the midgut, and other stable colonizers of the *D. melanogaster* gut like *Lactobacillus plantarum* and *Acetobacter thailandicus* also principally localize to the crop and proventriculus (40, 41). Our current results do not reveal whether lpD01 is attaching to and forming biofilms in the gut as has been observed for *L. plantarum* (41).

294 *Orbaceae* are widespread insect symbionts, but their roles in host biology and the
295 reasons for their host specificity are largely unexplored. *D. melanogaster* offers
296 sophisticated genetic resources for understanding the host side of these microbiome
297 interactions. Our results show that it is also possible to genetically modify *Orbaceae* to
298 begin to dissect these relationships. Because the Pathfinder plasmids are compatible
299 with established Golden Gate assembly schemes and parts libraries (6, 35), they can be
300 used to build new constructs, including systems for knocking out or inserting genes into
301 the bacterial chromosome (6, 44, 53). Such tools would facilitate future studies of
302 insect-*Orbaceae* interactions. It may also be possible to use these genetic tools to
303 control insects that are agricultural pests (1, 54), by isolating and engineering *Orbaceae*
304 native to the tephritid fruit fly *Bactrocera dorsalis* (33), for example.
305
306 **MATERIALS AND METHODS**
307 **Isolation of bmE01 from leafhoppers.** *Empoasca* sp. leafhoppers were
308 collected by sweep netting *Salvia* sp. plants on the University of Texas campus in
309 Austin, TX (30.289160, -97.738927). Individual leafhoppers were washed by soaking in
310 70% ethanol for one minute, followed by another minute in 10% bleach. Next, each
311 leafhopper was rinsed 3 times with sterile water and then crushed in 100 µL of sterile
312 saline. Then, 50 µL of 5 different leafhopper samples were plated onto separate Brain
313 Heart Infusion (BHI) agar containing cycloheximide at 100 µg/mL. Colonies were picked
314 and identified based on PCR and Sanger sequencing of the 16S rRNA gene with
315 primers 16SA1F (5'-AGAGTTTGATCMTGGCTCAG) and 16SB1R (5'-
316 TACGGYTACCTTGTTACGACTT) (55). The bme01 isolate studied here was predicted

317 with high confidence to be a *Rosenbergiella* species by both the Ribosomal Database
318 Project classifier tool (56), and BLAST searches of the NCBI 16S ribosomal RNA
319 sequence database (57).

320 **Isolation of BiB, lpD01, lpD02, and lpD03 from flies.** Wild flies were collected
321 at the Brackenridge Field Laboratory in Austin, TX (30.284326, -97.778522). Traps were
322 prepared by adding fermented banana, yeast, and twigs to punctured plastic bottles.
323 Traps were hung from trees for 7 days, and flies were collected each day. Flies were
324 placed on ice or at 4°C immediately after collection and processed within 24 hours.

325 Flies were immobilized on ice and photographed for morphological identification.
326 They were then washed with 10% bleach to remove surface microbes for 1 minute
327 followed by rinsing in sterile water for 1 minute to remove residual bleach. Legs and
328 wings were removed and placed in 95% ethanol to preserve host DNA. Next, each fly
329 was placed in 50 µl of Insectagro DS2 insect growth medium (IGM) (Corning, VA, USA)
330 and homogenized using a sterile plastic pestle. Dilutions of homogenate were plated on
331 heart infusion agar (HIA) with 5% sheep's blood. Agar plates were initially incubated at
332 37°C and 5% CO_2 (*e.g.*, in the case of BiB), but 30°C was later used due to superior
333 growth (*e.g.*, in the case of lpD01). Clear or off-white and slower-growing colonies were
334 passaged onto fresh plates multiple times to obtain pure cultures.

335 To identify the *Orbaceae* bacteria, we amplified and Sanger sequenced the 16S
336 rRNA gene using 27F (5'-AGAGTTTGATCMTGGCTCAG) and *Orbaceae*-specific
337 primer Orb742R (5'-ATCTCAGCGTCAGTATCTGTCCAGAA). Host insects were
338 identified both by morphology and by sequencing PCR amplicons of the barcode region
339 of the COI gene with primers LCO1490F (5'-GGTCAACAAATCATAAAGATATTGG) and

HCO2198R (5'-TAAACTTCAGGGTGACCAAAAAATCA) (58). The phylogenetic tree was assembled from 16S ribosomal RNA gene sequences. 16S rRNA gene sequences were aligned in Geneious using MUSCLE (59), and all sites containing ≥50% gaps were stripped. This masked alignment used to infer a phylogenetic tree using IQ-TREE with default options and nonparametric bootstrapping (60). The tree was visualized using iTOL (v5) (61).

Growth and maintenance of bacterial strains. *Escherichia coli* DH5α, *E. coli* MFD*pir*, and *Sodalis praecaptivus* HS^T were grown in LB broth and on LB agar at 37°C. Media was supplemented with 0.3 mM diaminopimelic acid (DAP) for MFD*pir* growth. Following isolation, *Rosenbergiella* was grown at 30°C on BHI broth or agar. *Orbus hercynius* CN3 was acquired from the German Collection of Microorganisms and Cell Cultures (DSMZ) (DSM 22228). Following isolation, *Orbaceae* strains were determined to be culturable in either IGM or BHI broth, and on BHI agar or HIA with or without 5% defibrinated sheep's blood. For robust growth, BHI and HIA + 5% sheep's blood were preferred, but media were selected based on the needs of the assay. All fly *Orbaceae* were grown at 30°C with 5% CO_2. *Gilliamella apis* M1-2G was grown on HIA + 5% sheep's blood at 35°C with 5% CO_2. Antibiotic concentrations used in this study are shown in Table 2.

MIC tests to determine antibiotic susceptibility. We performed MIC assays to determine the appropriate selective antibiotic concentration for each bacterial strain. To do so, we prepared 96-well plates with 2-fold dilutions of each antibiotic, ranging from 400 – 6.25 μg/mL in 100 μL media. One microliter of each strain was inoculated in triplicate for each condition. After allowing the strains to grow for 1-3 days, the plates

were inspected visually to determine the minimum inhibitory concentration for each antibiotic. These MIC values were used to guide how much antibiotic was used for selection during the Pathfinder conjugation process (Table 2).

Assembly of the Pathfinder plasmids. All cloning procedures were carried out in *E. coli* strain NEB5alpha (#C2987H, New England Biolabs) cultured overnight aerobically at 37°C in LB broth or solid LB agar. Antibiotics were supplemented when necessary for plasmid selection or maintenance at the following concentrations: Kanamycin (Kan) (50 μg/mL), Spectinomycin (Spec) (60 μg/mL), Gentamicin (Gent) (25 μg/mL), Chloramphenicol (Cam) (20 μg/mL), Erythromycin (Ery) (250 μg/mL), Tetracycline (Tet) (10 μg/mL), and Carbenicillin (Carb) (100 μg/mL).

We designed the Pathfinder plasmid series to have a variety of broad-host-range origins, antibiotic resistance genes, and highly expressed visible reporters suitable for rapid identification and testing in newly isolated bacteria. We started with pBTK402, a broad-host-range plasmid with a RSF1010 origin that we previously engineered to remove any BsaI and BsmBI cut sites and make it suitable for Golden Gate Assembly (6). The pBTK402 plasmid was designed to function as a Type 8 dropout vector for the BTK Golden Gate assembly scheme, and all of the main Pathfinder plasmids retain this attribute. We replaced the weakly expressed *rfp* on pBTK402 with a visible red chromoprotein (RCP) (Bba_E1010) expressed from the strong CP25 promoter and associated ribosome binding site (RBS) from plasmid pBTK569. This promoter-RBS combination appears to lead to robust protein expression in Proteobacteria. This plasmid was re-designated “pSL1” and all subsequent pSL plasmids are derived from pSL1. Plasmid pSL1-GFP replaces the RCP reporter with GFP. Plasmids pSL2 – pSL7

were constructed by replacing the Kanamycin resistance allele (*aphA-1*, Kan^R) present in pSL1 with an alternate antibiotic resistance allele and associated upstream promoter. pSL9 and pSL25 replace the RSF1010/RCP origin and reporter with pBBR1/E2-Crimson (E2C) and RP4/blue chromoprotein (BCP), respectively (see Table 1).

To construct plasmids, we first designed assembly primers using Benchling (http://www.benchling.com) and added Golden Gate-compatible BsmBI cut sites. We ordered DNA primers from Integrated DNA Technologies and then amplified PCR products using either Q5 Hot-Start Master Mix (#M0494S, New England Biolabs) or KOD XL (#71087-3, Millipore Sigma) according to manufacturer's instructions. We purified PCR products with a QIAquick PCR Purification Kit (#28104, QIAGEN), assembled them using a NEBridge Golden Gate Assembly Kit (BsmBI-v2) (#E1602S, New England Biolabs), and then electroporated 1 µL of the reaction into electrocompetent NEB5alpha. Cells recovered for 1 hour and were then plated on appropriate selective media. Plasmids were initially verified by Sanger sequencing of the assembly junctions and later by whole-plasmid Illumina sequencing on an iSeq 100. Three plasmids (pSL2, pSL4, and pSL5) contain a duplication of a BsaI restriction site and the adjacent CP25 promoter and ribosome binding site for RCP. Whether this duplication affects RCP expression is unknown. These plasmids still function as Type 8 dropout vectors for Golden Gate assembly.

We next transformed all pSL plasmids into the DAP auxotrophic conjugation donor strain *E. coli* MFD*pir* (36). We could not successfully transform pSL5, however, due to intrinsic erythromycin resistance in MFD*pir*. These strains of MFD*pir* were used individually or combined for subsequent multiplex conjugation assays. A preliminary

study examined various aspects of how the RSF1010-based Pathfinder plasmids function in *E. coli*, including how different antibiotic markers and concentrations affect plasmid copy number and reporter output and how stably they are maintained over many serial transfers in laboratory cultures with or without antibiotic selection (62).

Conjugation of the Pathfinder plasmids into insect associated bacteria. We created the Pathfinder conjugation mix by first growing up each of the donor strains separately to saturation. At this point we measured the optical density at 600 nm (OD600) of each strain and resuspended it at an OD600 of 1. Equal volumes of each strain were combined into a single tube along with 16% glycerol then distributed into PCR tubes and frozen at −80°C. For conjugations, this mix can be thawed and added straight to the first conjugation step.

To perform the conjugation itself, we started with 1 mL of an overnight culture of the target bacteria. The culture was pelleted (1000 × *g* for 6 minutes) and washed once with 145 mM NaCl saline to remove any residual media, and then resuspended in saline to OD600 = 1. Twenty-five µL of the sample was combined with 25 µL of the thawed Pathfinder plasmid mix and spot-plated on media compatible for the growth of both *E. coli* and *Orbaceae* spp. (BHI for all strains except lpD02, which was grown on BHI + 5% sheep blood) plus DAP. After 1-2 days of growth, we scraped up all the growth from the conjugation spot and suspended it in 1 mL of saline. This was washed and resuspended twice with 1 mL sterile saline to remove residual DAP. The resuspended sample was divided into 5 equal portions and serially diluted 10-fold to a final dilution of 1×10^{-5}. To plate these dilutions, 5 µL of each replicate and dilution were spotted onto antibiotic plates at each bacterium's MIC, along with a zero-antibiotic control. Plates were left to

dry and then placed in an incubator at the optimal temperature for each bacterium. If one of the antibiotic concentrations proved to be too low or too high following conjugation, it was adjusted accordingly, and the procedure was repeated (Table 2). To confirm successful conjugation, colonies were first visually examined for expression of fluorescent reporter genes. Then, one colony from each antibiotic condition was regrown in a liquid medium, and a dilution of this culture was used as template for whole-cell PCR and 16S rRNA gene sequencing with primers 16SA1F and 16SB1R, as described above. Colonies were counted and the efficiency of conjugation was determined relative to growth on the zero-antibiotic control plate.

Imaging bacterial colonies. To ensure accurate categorization of each of the fluorescent reporters in strains where the fluorescence was not as bright, we visualized plates on a Typhoon 9500 imager (GE Healthcare Bio-Sciences, Uppsala, Sweden). To visualize RCP fluorescence, we imaged with 532 nm excitation and 575 nm emission. To distinguish E2C fluorescence from RCP, we imaged the plate with 635 nm excitation and 665 nm emission. Images were processed and counted in Fiji (version 1.53q) (63).

Measuring GFP expression in different bacteria. To measure the level of GFP expression, we assessed three separate colonies picked from the Pathfinder conjugation plate. Each colony was first grown to saturation, then re-grown with a starting OD600 of 0.05 in 5 mL Insectagro DS2 in a test tube. After re-growth, 1 mL of each culture was pelleted and resuspended in 600 µL saline. OD600 and GFP expression (485 nm excitation, 535 nm emission) were measured for 200 µL of each resuspension in triplicate using a Tecan Infinite 200 Pro plate reader. GFP expression was normalized to the OD600 measurement for each sample.

Growth and maintenance of *Drosophila* stocks. *D. melanogaster* Canton-S were acquired from the Bloomington Drosophila Stock Center (Bloomington, IN). Fly stocks were reared on Formula 4-24 Instant Drosophila Medium (Carolina Biological Supply Company, Burlington, NC). For experiments, stocks were swapped to a yeast-glucose agar (YGA) diet containing brewer's yeast, D-glucose, agar, and water (52). Nystatin (10 µg/mL) and kanamycin (10 µg/mL) were added to the diet where specified to prevent the growth of fungal contaminants and ensure the maintenance of engineered strains, respectively. Stocks were maintained at 25°C during experiments, with a 12L:12D photoperiod in Percival I-36LLVL incubators (Perry, IA, USA).

Inoculation of *Drosophila* with *Orbaceae* strains. *D. melanogaster* was inoculated with the engineered *Orbaceae* using a method we refer to as arena inoculation. For these experiments, adult female *D. melanogaster* were first removed from the commercial diet and transferred to YGA diet for 48 hours to clear their gut from the preservatives in the commercial diet. In the meantime, HIA + 5% sheep blood + Kan plates were inoculated with 100 µL of OD600 = 1 *Orbaceae* + pSL1-GFP and allowed to grow for 1-2 days as needed. The bacterial plate was then taped to the bottom of a lidded 12oz SelecTE plastic container "arena" (Berry Plastics, Evansville, IN) modified with a mesh vent on its lid and a small port for administering CO2. YGA flies were added to an arena and allowed to feed for 24 hours. After this inoculation step, flies were transferred daily to fresh YGA + nystatin + Kan diet tubes.

Quantification of bacterial colonization. Each bacterial strain was first delivered to female adult flies 3-5 days post-eclosion using the methodology described above. For this set of experiments, 3 different arenas were used to inoculate 3 separate

478 populations of female *D. melanogaster.* Equal numbers of flies were placed into each
479 arena, ranging from 100-150 in each arena for each round of this experiment. After
480 inoculation, each separate population of flies was transferred to their own tube of YGA +
481 nystatin + Kan diet. Flies were transferred to fresh diet daily. At day(s) 0, 2, 4, 7, and 11
482 post-inoculation, 10 flies from each independent population (30 total) were crushed and
483 plated to determine their quantity of bacterial colonization.

484 To prepare samples for crushing, we first washed them in 10% bleach to ensure
485 that any bacteria plated would not be from the outside of the fly body. Flies were
486 immobilized for this procedure by first placing them at −20°C for 1 minute, and then
487 keeping the tubes on ice during the washes. They were soaked in 500 µL 10% bleach
488 for 1 minute, washed with 500 µL 1× phosphate buffered saline (PBS), then
489 resuspended in 200ul PBS and homogenized with a sterile plastic micropestle. A set of
490 1:10 serial dilutions in PBS were then carried out for each sample. Five microliters of
491 each dilution were spotted onto agar plates (HIA + 5% sheep blood, Kan 50 µg/mL,
492 nystatin 10 µg/mL) in triplicate. The remaining homogenized fly mixture was then
493 centrifuged at 1000 ×*g* for 6 minutes to pellet the remaining bacteria. The pellet was
494 resuspended in 15 µL saline and plated to ensure the detection of bacteria that were
495 present in low abundance.

496 **Imaging *Drosophila* guts with confocal microscopy.** Fifty female *D.*
497 *melanogaster* were colonized with lpD01 + pSL1-GFP in two separate arenas and
498 maintained for 11 days along with an uncolonized control population. To ensure that
499 flies were colonized with the bacteria, 3 flies from each population were crushed and
500 plated on day 9 as described in the prior methods section. On day 11, all flies were

transferred to a tube without diet for 12-18 hours to empty the gut of any diet-associated particles that may complicate imaging. The entire gut of a selection of the flies was dissected and mounted in PBS on a glass slide. The dissected guts were imaged using a Leica MZ16 Fluorescent Stereoscope in the GFP channel to visualize bacterial colonization. Images were linearly adjusted to highlight the bacterial localization using Fiji software (version 1.53q) (63).

Data availability. The data that support the findings of this study and newly isolated bacteria reported here are available from the authors upon reasonable request. The Pathfinder plasmids have been deposited in Addgene (Table 1).

ACKNOWLEDGEMENTS

We thank Bibiana Toro and other members of the 2018 UT Austin iGEM team for assistance with Pathfinder toolkit development and testing. We thank John McCutcheon for the gift of *Sodalis praecaptivus*. We thank Thilini Wijesekera for providing stocks of *D. melanogaster* Canton-S. This research made use of equipment at the Microscopy and Imaging Facility of the Center for Biomedical Research Support at The University of Texas at Austin (RRID# SCR_021756). Funding was provided by the U.S. Army Research Office (W911NF-20-1-0195) to J.E.B. and N.A.M. and by the U.S. National Institutes of Health (R35GM131738) to N.A.M. K.M.E. acknowledges a UT Austin University Graduate Continuing Fellowship, T.A. and B.M. acknowledge UT Austin TIDES Summer Research Fellowships, and B.M. acknowledges an Army Educational Outreach Program (AEOP) Undergraduate Research Apprenticeship.

REFERENCES

1. Elston KM, Leonard SP, Geng P, Bialik SB, Robinson E, Barrick JE. 2022. Engineering insects from the endosymbiont out. Trends Microbiol 30:79–96.
2. Brophy JAN, Triassi AJ, Adams BL, Renberg RL, Stratis-Cullum DN, Grossman AD, Voigt CA. 2018. Engineered integrative and conjugative elements for efficient and inducible DNA transfer to undomesticated bacteria. Nat Microbiol 3:1043–1053.
3. Waller MC, Bober JR, Nair NU, Beisel CL. 2017. Toward a genetic tool development pipeline for host-associated bacteria. Curr Opin Microbiol 38:156–164.
4. Durante-Rodríguez G, De Lorenzo V, Martínez-García E. 2014. The standard European vector architecture (SEVA) plasmid toolkit. Methods Mol Biol 1149:469–478.
5. Iverson S V., Haddock TL, Beal J, Densmore DM. 2016. CIDAR MoClo: Improved MoClo assembly standard and new *E. coli* part library enable rapid combinatorial design for synthetic and traditional biology. ACS Synth Biol 5:99–103.
6. Leonard SP, Perutka J, Powell JE, Geng P, Richhart DD, Byrom M, Kar S, Davies BW, Ellington AD, Moran NA, Barrick JE. 2018. Genetic engineering of bee gut microbiome bacteria with a toolkit for modular assembly of broad-host-range plasmids. ACS Synth Biol 7:1279–1290.
7. Martínez-Garćía E, Aparicio T, Goñi-Moreno A, Fraile S, De Lorenzo V. 2015. SEVA 2.0: An update of the Standard European Vector Architecture for de-/re-construction of bacterial functionalities. Nucleic Acids Res 43:D1183–D1189.

8. Schuster LA, Reisch CR. 2021. A plasmid toolbox for controlled gene expression across the Proteobacteria. Nucleic Acids Res 49:7189–7202.

9. Geddes BA, Mendoza-Suárez MA, Poole PS. 2019. A Bacterial Expression Vector Archive (BEVA) for flexible modular assembly of golden gate-compatible vectors. Front Microbiol 10:3345.

10. Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. 2011. A modular cloning system for standardized assembly of multigene constructs. PLoS One 6:e16765.

11. Young R, Haines M, Storch M, Freemont PS. 2021. Combinatorial metabolic pathway assembly approaches and toolkits for modular assembly. Metab Eng 63:81–101.

12. Freed E, Fenster J, Smolinski SL, Walker J, Henard CA, Gill R, Eckert CA. 2018. Building a genome engineering toolbox in nonmodel prokaryotic microbes. Biotechnol Bioeng 115:2120–2138.

13. Jain A, Srivastava P. 2013. Broad host range plasmids. FEMS Microbiol Lett 348:87–96.

14. Aune TEV, Aachmann FL. 2010. Methodologies to increase the transformation efficiencies and the range of bacteria that can be transformed. Appl Microbiol Biotechnol 85:1301–1313.

15. Mazodier P, Davies J. 1991. Gene transfer between distantly related bacteria. Annu Rev Genet 25:147–171.

16. Ronda C, Chen SP, Cabral V, Yaung SJ, Wang HH. 2019. Metagenomic engineering of the mammalian gut microbiome in situ. Nat Methods 16:167–170.

17. Von Wintersdorff CJH, Penders J, Van Niekerk JM, Mills ND, Majumder S, Van Alphen LB, Savelkoul PHM, Wolffs PFG. 2016. Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. Front Microbiol 7:173.

18. Akiyama T, Asfahl KL, Savin MC. 2010. Broad-host-range plasmids in treated wastewater effluent and receiving streams. J Environ Qual 39:2211–2215.

19. Brooks LE, Kaze M, Sistrom M. 2019. Where the plasmids roam: Large-scale sequence analysis reveals plasmids with large host ranges. Microb Genomics 5:e000244.

20. Douglas AE. 2018. The *Drosophila* model for microbiome research. Lab Anim (NY). Nature Publishing Group.

21. Ludington WB, Ja WW. 2020. *Drosophila* as a model for the gut microbiome. PLoS Pathog 16:e1008398.

22. Broderick NA, Lemaitre B. 2012. Gut-associated microbes of *Drosophila melanogaster*. Gut Microbes. Taylor & Francis.

23. Newell PD, Chaston JM, Wang Y, Winans NJ, Sannino DR, Wong ACN, Dobson AJ, Kagle J, Douglas AE. 2014. In vivo function and comparative genomic analyses of the *Drosophila* gut microbiota identify candidate symbiosis factors. Front Microbiol 5:576.

24. Shin SC, Kim S-H, You H, Kim B, Kim AC, Lee K-A, Yoon J-H, Ryu J-H, Lee W-J. 2011. *Drosophila* microbiome modulates host developmental and metabolic homeostasis via insulin signaling. Science 334:670–674.

593 25. Martino ME, Joncour P, Leenay R, Gervais H, Shah M, Hughes S, Gillet B, Beisel
594 C, Leulier F. 2018. Bacterial adaptation to the host's diet is a key evolutionary
595 force shaping *Drosophila-Lactobacillus* symbiosis. Cell Host Microbe 24:109-
596 119.e6.
597 26. Chandler JA, Morgan Lang J, Bhatnagar S, Eisen JA, Kopp A. 2011. Bacterial
598 communities of diverse *Drosophila* species: Ecological context of a host–microbe
599 model system. PLoS Genet 7:e1002272.
600 27. Volkmann M, Skiebe E, Kerrinnes T, Faber F, Lepka D, Pfeifer Y, Holland G,
601 Bannert N, Wilharm G. 2010. *Orbus hercynius* gen. nov., sp. nov., isolated from
602 faeces of wild boar, is most closely related to members of the orders
603 "*Enterobacteriales*" and *Pasteurellales*. Int J Syst Evol Microbiol 60:2601–2605.
604 28. Kwong WK, Moran NA. 2013. Cultivation and characterization of the gut
605 symbionts of honey bees and bumble bees: description of *Snodgrassella alvi* gen.
606 nov., sp. nov., a member of the family *Neisseriaceae* of the *Betaproteobacteria*,
607 and *Gilliamella apicola*. Int J Syst Evol Microbiol 63:2008–2018.
608 29. Martinson VG, Carpinteyro-Ponce J, Moran NA, Markow TA. 2017. A distinctive
609 and host-restricted gut microbiota in populations of a cactophilic *Drosophila*
610 species. Appl Environ Microbiol 83:e01551-17.
611 30. Hammer TJ, McMillan WO, Fierer N. 2014. Metamorphosis of a butterfly-
612 associated bacterial community. PLoS One 9:86995.

613 31. Kim JY, Lee J, Shin NR, Yun JH, Whon TW, Kim MS, Jung MJ, Roh SW, Hyun
614 DW, Bae JW. 2013. *Orbus sasakiae* sp. nov., a bacterium isolated from the gut of
615 the butterfly *Sasakia charonda*, and emended description of the genus *Orbus*. Int
616 J Syst Evol Microbiol 63:1766–1770.
617 32. Martinson VG, Douglas AE, Jaenike J. 2017. Community structure of the gut
618 microbiota in sympatric species of wild *Drosophila*. Ecol Lett 20:629–639.
619 33. Liu S-H, Chen Y, Li W, Tang G-H, Yang Y, Jiang H-B, Dou W, Wang J-J. 2018.
620 Diversity of bacterial communities in the intestinal tracts of two geographically
621 distant populations of *Bactrocera dorsalis* (Diptera: Tephritidae). J Econ Entomol
622 111:2861–2868.
623 34. Gupta AK, Nayduch D, Verma P, Shah B, Ghate H V., Patole MS, Shouche YS.
624 2012. Phylogenetic characterization of bacteria in the gut of house flies (*Musca*
625 *domestica* L.). FEMS Microbiol Ecol 79:581–593.
626 35. Lee ME, DeLoache WC, Cervantes B, Dueber JE. 2015. A highly characterized
627 yeast toolkit for modular, multipart assembly. ACS Synth Biol 4:975–986.
628 36. Ferrières L, Hémery G, Nham T, Guérout AM, Mazel D, Beloin C, Ghigo JM.
629 2010. Silent mischief: Bacteriophage Mu insertions contaminate products of
630 *Escherichia coli* random mutagenesis performed using suicidal transposon
631 delivery plasmids mobilized by broad-host-range RP4 conjugative machinery. J
632 Bacteriol 192:6418–6427.

633 37. Chari A, Oakeson KF, Enomoto S, Grant Jackson D, Fisher MA, Dale C. 2015.
634 Phenotypic characterization of *Sodalis praecaptivus* sp. nov., a close non-insect-
635 associated member of the *Sodalis*-allied lineage of insect endosymbionts. Int J
636 Syst Evol Microbiol 65:1400–1405.
637 38. Medina Munoz M, Spencer N, Enomoto S, Dale C, Rio RVM. 2020. Quorum
638 sensing sets the stage for the establishment and vertical transmission of *Sodalis*
639 *praecaptivus* in tsetse flies. PLOS Genet 16:e1008992.
640 39. Enomoto S, Chari A, Clayton AL, Dale C. 2017. Quorum Sensing Attenuates
641 Virulence in *Sodalis praecaptivus*. Cell Host Microbe 21:629–636.
642 40. Pais IS, Valente RS, Sporniak M, Teixeira L. 2018. *Drosophila melanogaster*
643 establishes a species-specific mutualistic interaction with stable gut-colonizing
644 bacteria. PLoS Biol 16:e2005710.
645 41. Obadia B, Güvener ZT, Zhang V, Ceja-Navarro JA, Brodie EL, Ja WW, Ludington
646 WB. 2017. Probabilistic invasion underlies natural gut microbiome stability. Curr
647 Biol 27:1999-2006.e8.
648 42. Dodge R, Jones EW, Zhu H, Obadia B, Martinez DJ, Wang C, Aranda-Díaz A,
649 Aumiller K, Liu Z, Voltolini M, Brodie EL, Huang KC, Carlson JM, Sivak DA,
650 Spradling AC, Ludington WB. 2021. A gut commensal niche regulates stable
651 association of a multispecies microbiota. bioRxiv 2021.09.30.462663.
652 43. Kovach ME, Elzer PH, Steven Hill D, Robertson GT, Farris MA, Roop RM,
653 Peterson KM. 1995. Four new derivatives of the broad-host-range cloning vector
654 pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene 166:175–
655 176.

656 44. Wiles TJ, Wall ES, Schlomann BH, Hay EA, Parthasarathy R, Guillemin K. 2018.
657 Modernized tools for streamlined genetic manipulation and comparative study of
658 wild and diverse proteobacterial lineages. MBio 9:e01877-18.
659 45. Meyer R. 2009. Replication and conjugative mobilization of broad host-range IncQ
660 plasmids. Plasmid 62:57–70.
661 46. Bishé B, Taton A, Golden JW. 2019. Modification of RSF1010-based broad-host-
662 range plasmids for improved conjugation and cyanobacterial bioprospecting.
663 iScience 20:216–228.
664 47. Ceroni F, Algar R, Stan G-BB, Ellis T. 2015. Quantifying cellular capacity
665 identifies gene expression designs with reduced burden. Nat Methods 12:415–
666 418.
667 48. Renda BA, Hammerling MJ, Barrick JE. 2014. Engineering reduced evolutionary
668 potential for synthetic biology. Mol Biosyst 10:1668–1678.
669 49. Englaender JA, Jones JA, Cress BF, Kuhlman TE, Linhardt RJ, Koffas MAG.
670 2017. Effect of genomic integration location on heterologous protein expression
671 and metabolic engineering in *E. coli*. ACS Synth Biol 6:710–720.
672 50. Rubin BE, Diamond S, Cress BF, Crits-Christoph A, Lou YC, Borges AL, Shivram
673 H, He C, Xu M, Zhou Z, Smith SJ, Rovinsky R, Smock DCJ, Tang K, Owens TK,
674 Krishnappa N, Sachdeva R, Barrangou R, Deutschbauer AM, Banfield JF,
675 Doudna JA. 2021. Species- and site-specific genome editing in complex bacterial
676 communities. Nat Microbiol 2021 71 7:34–47.

677 51. Lewis WH, Tahon G, Geesink P, Sousa DZ, Ettema TJG. 2020. Innovations to
678 culturing the uncultured microbial majority. Nat Rev Microbiol 2020 194 19:225–
679 240.
680 52. Koyle ML, Veloz M, Judd AM, Wong ACN, Newell PD, Douglas AE, Chaston JM.
681 2016. Rearing the fruit fly *Drosophila melanogaster* under axenic and gnotobiotic
682 conditions. J Vis Exp 54219.
683 53. Tucker AT, Nowicki EM, Boll JM, Knauf GA, Burdis NC, Trent MS, Davies BW.
684 2014. Defining gene-phenotype relationships in *Acinetobacter baumannii* through
685 one-step chromosomal gene inactivation. MBio 5:e01313-14.
686 54. Berasategui A, Shukla S, Salem H, Kaltenpoth M. 2016. Potential applications of
687 insect symbionts in biotechnology. Appl Microbiol Biotechnol 100:1567–1577.
688 55. Fukatsu T, Nikoh N. 1998. Two intracellular symbiotic bacteria from the mulberry
689 psyllid *Anomoneura mori* (Insecta: Homoptera). Appl Environ Microbiol 64:3599–
690 3606.
691 56. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naïve Bayesian classifier for
692 rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl
693 Environ Microbiol 73:5261–5267.
694 57. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Connor R,
695 Funk K, Kelly C, Kim S, Madej T, Marchler-Bauer A, Lanczycki C, Lathrop S, Lu
696 Z, Thibaud-Nissen F, Murphy T, Phan L, Skripchenko Y, Tse T, Wang J, Williams
697 R, Trawick BW, Pruitt KD, Sherry ST. 2022. Database resources of the national
698 center for biotechnology information. Nucleic Acids Res 50:D20–D26.

699 58. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for
700 amplification of mitochondrial cytochrome c oxidase subunit I from diverse
701 metazoan invertebrates. Mol Mar Biol Biotechnol 3:294–299.
702 59. Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, Madhusoodanan
703 N, Kolesnikov A, Lopez R. 2022. Search and sequence analysis tools services
704 from EMBL-EBI in 2022. Nucleic Acids Res 50:W276–W279.
705 60. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler
706 A, Lanfear R, Teeling E. 2020. IQ-TREE 2: New models and efficient methods for
707 phylogenetic inference in the genomic era. Mol Biol Evol 37:1530–1534.
708 61. Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for
709 phylogenetic tree display and annotation. Nucleic Acids Res 49:W293–W296.
710 62. Young E. 2021. Honors thesis. The University of Texas at Austin, Austin, TX.
711 Antibiotic markers: An overlooked design choice in synthetic biology? DOI:
712 10.26153/tsw/43814
713 63. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T,
714 Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein
715 V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: An open-source platform for
716 biological-image analysis. Nat Methods 9:676–682.
717 64. Campbell RE, Tour O, Palmer AE, Steinbach PA, Baird GS, Zacharias DA, Tsien
718 RY. 2002. A monomeric red fluorescent protein. Proc Natl Acad Sci U S A
719 99:7877–7882.

65. Alexeyev MF. 1999. The pKNOCK series of broad-host-range mobilizable suicide vectors for gene knockout and targeted DNA insertion into the chromosome of gram-negative bacteria. Biotechniques 26:824–6, 828.

66. Bryan EM, Bae T, Kleerebezem M, Dunny GM. 2000. Improved vectors for nisin-controlled expression in gram-positive bacteria. Plasmid 44:183–190.

67. Prior JE, Lynch MD, Gill RT. 2010. Broad-host-range vectors for protein expression across gram negative hosts. Biotechnol Bioeng 106:326–332.

68. Dammeyer T, Timmis KN, Tinnefeld P. 2013. Broad host range vectors for expression of proteins with (Twin-) Strep-tag, His-tag and engineered, export optimized yellow fluorescent protein. Microb Cell Fact 12:49.

69. Liljeruhm J, Funk SK, Tietscher S, Edlund AD, Jamal S, Wistrand-Yuen P, Dyrhage K, Gynnå A, Ivermark K, Lövgren J, Törnblom V, Virtanen A, Lundin ER, Wistrand-Yuen E, Forster AC. 2018. Engineering a palette of eukaryotic chromoproteins for bacterial synthetic biology. J Biol Eng 12:8.

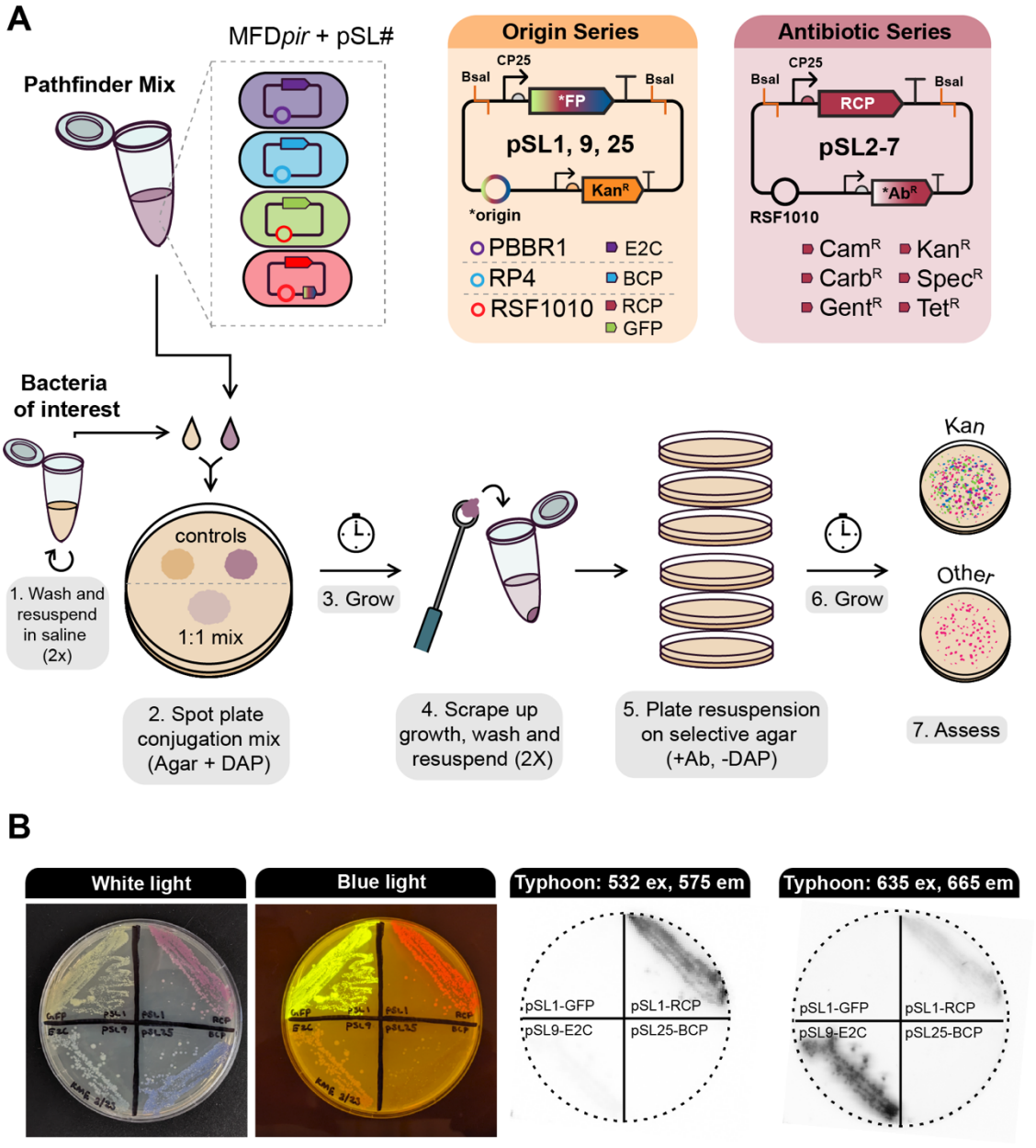


FIG 1 The Pathfinder plasmid system. (A) Plasmid maps and workflow for multiplex conjugation into a bacterium of interest. (B) Visualizing Pathfinder reporters. The same agar plate containing streaks of *E. coli* MFD*pir* donor strains, each with a plasmid expressing a different fluorescent protein or chromoprotein is shown in all images. The leftmost panel shows the plate under white light, the second shows the plate on a blue light transilluminator, and the last two panels show the plate imaged using a Typhoon 9500 FLA system with two different excitation and emission settings.

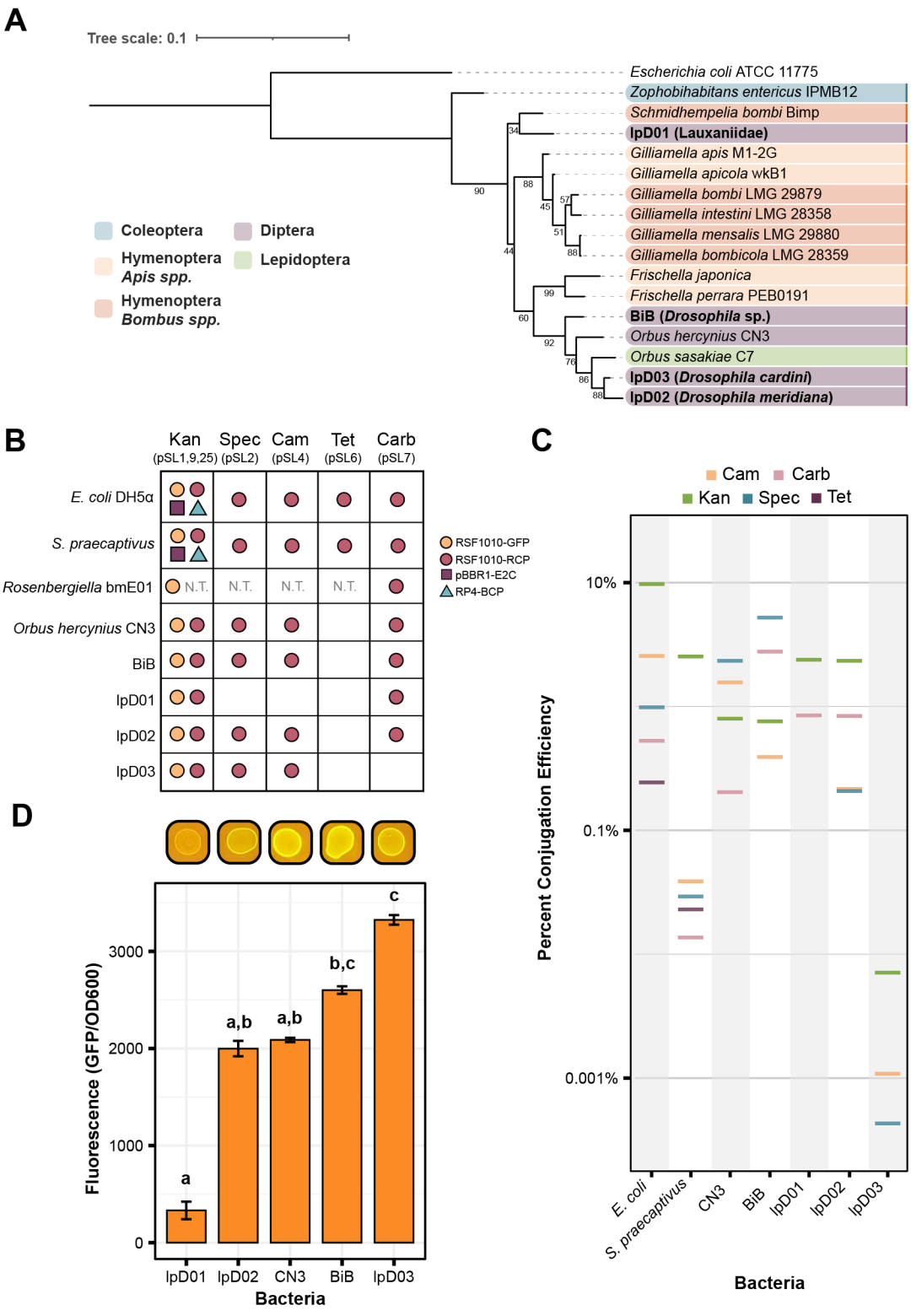
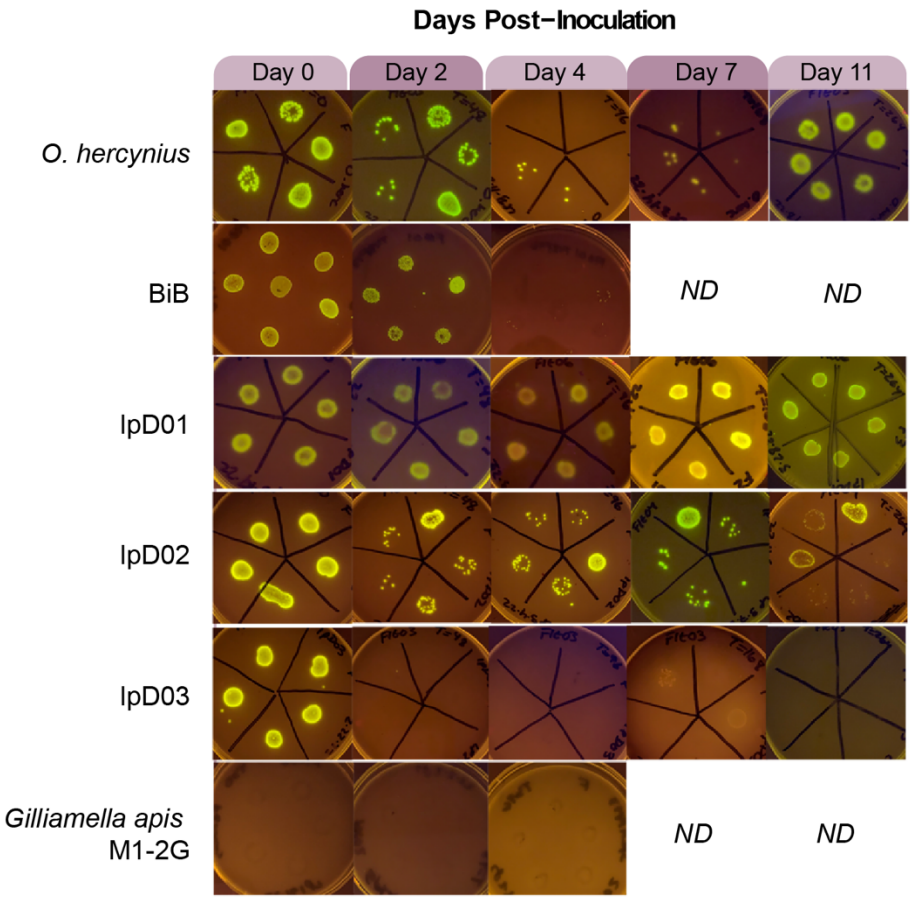


FIG 2 Recently isolated *Orbaceae* strains can be engineered with the Pathfinder plasmid system. (A) 16S rRNA gene sequence phylogeny showing the relationship between the isolated species in this paper (in bold) and other members of the family

Orbaceae. Strains are color-coded based on the taxonomic order of their original insect host. For each *Orbaceae* strain first reported in this study the closest taxonomic identifier that could be established for the host is shown in parentheses. Bootstrap values are depicted next to their respective branches. (B) Table depicting the compatibility of the bacteria in this study with each of the Pathfinder plasmids. Compatibility is defined as isolation and verification of at least a single transconjugant colony for each plasmid. Dots are used to represent compatibility and are color-coded according to the origin of replication and reporter gene on the plasmid. (C) The efficiency of conjugation under each of the antibiotic conditions in the study is plotted as a percentage of transconjugants relative to growth on nonselective media. Data are plotted on a log scale. Except in the Kan condition, where multiple plasmids may be present as shown in (B), all plasmids are RSF1010-RCP. (D) Bar chart showing the level of GFP expression for each strain engineered with pSL1-GFP (RSF1010-GFP) normalized to an OD600 reading for that same strain. Images above the chart show the appearance of each strain on a blue light transilluminator. Letters above each bar designate groups that are significantly different at the $p < 0.01$ level calculated using Dunn's test with Bonferroni correction.

764



765

766 **FIG 3** Engineered *Orbaceae* strains can colonize *D. melanogaster.* Images were taken
767 of bacterial growth on selective media containing kanamycin following the growth of
768 bacteria from crushed flies in the days following inoculation. Ability to colonize the host
769 is determined by the presence of many GFP-expressing colonies. Days on which we
770 were unable to collect data or stopped the experiment are indicated by “ND”.

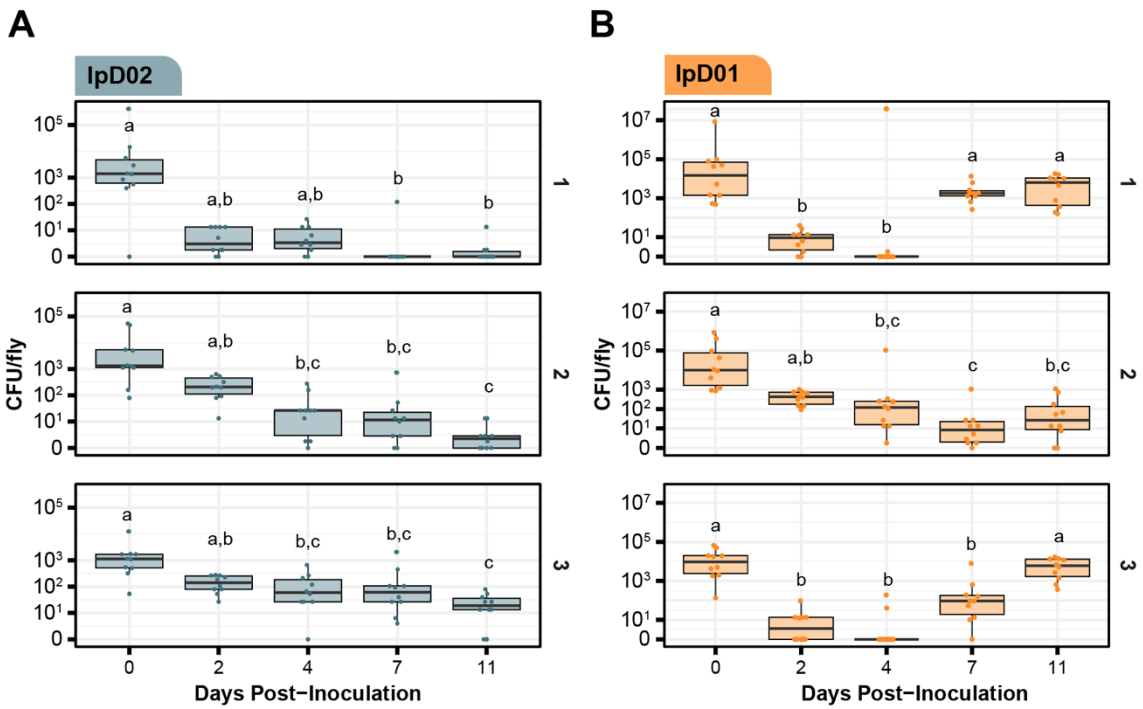
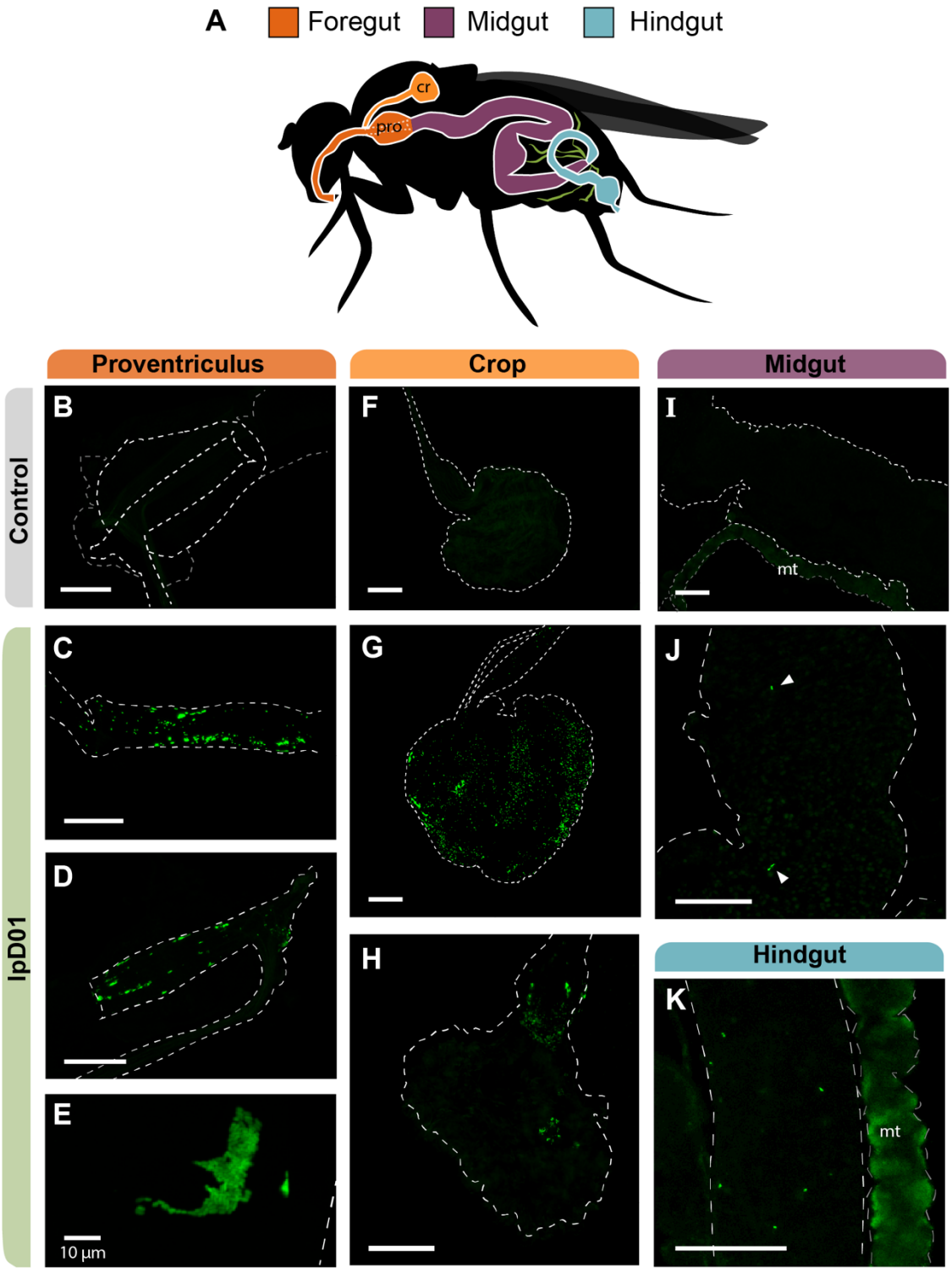


FIG 4 Dynamics of *D. melanogaster* colonization by two newly isolated *Orbaceae* strains. CFU/fly was measured following inoculation with either (A) lpD02 or (B) lpD01. CFUs in each of ten flies per arena were measured at each time point. Data are plotted on a pseudo-log scale so the full range of colonization levels can be shown. Results from three independent arenas are shown in subpanels labeled 1, 2, and 3. Letters above each boxplot represent groupings that are significantly different from one another ($p < 0.05$, Dunn's test with Bonferroni correction).



779

780 **FIG 5** Confocal microscopy of the *D. melanogaster* gut colonized with fluorescent
781 lpD01. (A) Schematic of the *Drosophila* gut with colors designating three regions,
782 foregut, midgut, and hindgut. The crop (cr) and proventriculus (pro) are labelled due to
783 their relevance for the localization of the strain. Malpighian tubules (mt) are depicted in
784 green. (B-K) Confocal images of the dissected gut of one uncolonized (B, F, I) and three

785 colonized flies. Images (C, G, J and K) are from the same fly, (D, H) are from a second
786 fly, and (E) is a from a third. Images (B-D, F-K) were captured at 10× magnification, and
787 image (E) was captured at 40×. Outlines for each of the relevant gut regions are
788 represented by white dashed lines and arrows were added to (J) to point out individual
789 bacterial cells. The Malpighian tubules (mt) exhibited autofluorescence in all flies. GFP
790 intensities were linearly adjusted in each image to highlight bacterial localization. The
791 scale bar for each image represents 100 μm except where indicated.

792 **Table 1** Pathfinder plasmids
793

Plasmid	Addgene accession	Origin	Origin source	Reporter	Reporter source	Antibiotic resistance gene	Resistance source
pSL1	180422	RSF1010	pBTK402 (6)	RCP (mRFP1)	BBa_E1010 (64)	Kan (*aphA-1*)	pBTK402 (6)
pSL1-GFP[a]	180420	RSF1010	pBTK402 (6)	GFP	pBTK520 (6)	Kan (*aphA-1*)	pBTK402 (6)
pSL2[b]	190998	RSF1010	pBTK402 (6)	RCP (mRFP1)	Bba_E1010 (64)	Spec (*aadA*)	pBTK403 (6)
pSL3	190999	RSF1010	pBTK402 (6)	RCP (mRFP1)	Bba_E1010 (64)	Gent (*aacC1*)	pKNOCK-Gm (65)
pSL4[b]	191000	RSF1010	pBTK402 (6)	RCP (mRFP1)	Bba_E1010 (64)	Cam (*cat1*)	pYTK001 (35)
pSL5[b]	191001	RSF1010	pBTK402 (6)	RCP (mRFP1)	Bba_E1010 (64)	Ery (*ermB*)	pMSP3535 (66)
pSL6	191002	RSF1010	pBTK402 (6)	RCP (mRFP1)	Bba_E1010 (64)	Tet (*tetC*)	pBMTBX-4 (67)
pSL7	191003	RSF1010	pBTK402 (6)	RCP (mRFP1)	Bba_E1010 (64)	Carb (*tem116*)	pBTK401 (6)
pSL9	191004	pBBR1	pBBR1MCS-2 (43)	E2C	pBTK570 (6)	Kan (*aphA-1*)	pBTK402 (6)
pSL25	191005	RP4	pTD-C_sYFPTwinStrep (68)	BCP (amilCP)	amilCP chromoprotein (69)	Kan (*aphA-1*)	pBTK402 (6)

794 [a]Plasmid is not compatible with Golden Gate assembly.

795 [b]Plasmid contains a duplication of one BsaI site and the adjacent CP25 promoter and
796 ribosome binding site for RCP. It is still compatible with Golden Gate assembly.

Table 2 Antibiotic concentrations used in this study[a]

		CARB	CAM	KAN	SPEC	TET
DH5α	**MIC**	100	20	50	60	10
	Pathfinder	100	20	50	60	10
Sodalis praecaptivus	**MIC**	200	6.25	6.25	50	12.5
	Pathfinder	200	6.25	6.25	50	12.5
bmE01	**MIC**	400	25	200	400	6.25
	Pathfinder	***800**	***50**	***400**	***800**	***12.5**
***Orbus hercynius* CN3**	**MIC**	800	6.25	25	25	6.25
	Pathfinder	800	6.25	25	***100**	***3**
BiB	**MIC**	25	6.25	400	100	6.25
	Pathfinder	25	6.25	400	***200**	***3**
lpD01	**MIC**	6.25	400	12.5	400	200
	Pathfinder	***25**	400	***50**	***800**	200
lpD02	**MIC**	200	6.25	50	100	6.25
	Pathfinder	***400**	6.25	***100**	100	***3**
lpD03	**MIC**	800	6.25	25	100	6.25
	Pathfinder	***6400**	6.25	***100**	100	***3**

[a]All concentrations are provided in μg/mL. Starred (*) and bolded values indicate when concentrations different from the MIC were used for Pathfinder conjugation assays.