Strict gut symbiont specificity in Coreoidea insects governed by interspecies competition within Caballeronia strains

Gaelle Lextrait, Srotoswini Joardar, Raynald Cossard, Yoshitomo Kikuchi, Tsubasa Ohbayashi, Peter Mergaert

Strict gut symbiont specificity in Coreoidea insects governed by interspecies competition within *Caballeronia* strains

Gaelle Lextrait[1], Srotoswini Joardar[1,2,‡], Raynald Cossard[1], Yoshitomo Kikuchi[3], Tsubasa Ohbayashi[1,4,*], Peter Mergaert[1,*]

[1]Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette, France

[2]Department of Biology, Ludwig-Maximilians-Universität München (LMU), 82152 Munich, Germany

[3]Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Hokkaido Center, 062-8517 Sapporo, Japan

[4]Institute for Agro-Environmental Sciences, National Agriculture, and Food Research Organization (NARO), 305-8604 Tsukuba, Japan

*Corresponding authors. Peter Mergaert, Institute of Integrative Biology of the Cell, avenue de la Terrasse Bât. 21, 91198 Gif-sur-Yvette, France E-mail peter.mergaert@i2bc.paris-saclay.fr. Tsubasa Ohbayashi, Institute for Agro-Environmental Sciences, National Agriculture and Food Research Organization, 3-1-3 Kannondai, 305-8604 Tsukuba, Ibaraki, Japan. E-mail: ohbayashi.tsubasa023@naro.go.jp

[‡]Present address: Helmholtz Zentrum München, Comparative Microbiome Analysis (COMI), 85764 Oberschleissheim, Germany

Running title: Insect-*Caballeronia* symbiont specificity

ABSTRACT

Host-bacteria symbioses are often specific and transgenerationaly stable. In hosts that acquire their symbionts from the environment across successive generations, selective mechanisms are required to identify and maintain beneficial partners from diverse environmental microorganisms. In Coreoidea stinkbugs, which house environmentally acquired symbionts in a specialized midgut region, bacterial competition plays a key role in shaping symbiont specificity whereby *Caballeronia* strains consistently outcompete bacteria of other genera. Here, we show that competition within the gut also occurs among *Caballeronia* strains themselves, driving specificity at a finer taxonomic scale. Specifically, the stinkbugs *Riptortus pedestris* and *Coreus marginatus*, when reared on the same soil sample, preferentially select for α- and β-subclade *Caballeronia*, respectively. Using a gnotobiotic infection system, we demonstrate that representative strains from the α-, β-, and γ-subclades can independently colonize the midgut of both insect species in monoculture. However, in pairwise co-culture infections, each host exhibits a marked selectivity for either α- or β-subclade strains, consistent with patterns observed in the soil inoculation experiment. In *R. pedestris*, we further find that interspecies competition outcomes are shaped by both priority effects and displacement mechanisms. At the molecular level, differences among symbionts in metabolic capabilities, resistance to antimicrobial peptides, and chemotactic behavior influence their competitive success in the gut. Finally, we show that in *R. pedestris*, the reproductive fitness benefits conferred by the symbiosis align with the observed strain specificity in the tested strain panel, suggesting a functional link between symbiont selection and host fitness, despite these processes occurring at distinct stages of the symbiotic relationship. Our findings highlight that the gut in Coreoidea species constitutes a multifactorial, species-specific selective environment that contributes to the colonization of the symbiotic midgut region by the best-adapted *Caballeronia* strain.

Key words: bacteria-bacteria competition, insect-bacteria interactions, gut symbiosis, *Riptortus pedestris*, *Coreus marginatus*, *Caballeronia*, chemotaxis, antimicrobial peptides, metabolic capacity

INTRODUCTION

Since their early evolutionary history, insects have developed heritable interactions with a wide range of microorganisms [1, 2]. These symbionts have expanded the repertoire of capabilities in insects and thereby contributed to their ecological success [3, 4]. To maintain beneficial partnerships, insects have evolved mechanisms that ensure reliable symbiont acquisition and partner fidelity. Most insect symbioses are stably maintained through vertical transmission, passing symbionts from parent to offspring [1, 2]. However, some insects acquire their symbionts from the environment. In such cases, each symbiont-free hatchling must identify and selectively recruit suitable symbionts but avoid harmful or non-cooperative microorganisms.

Partner choice in environmentally transmitted symbioses often involves molecular signaling between host and symbiont, as seen in well-studied systems like rhizobia–legume and *Aliivibrio fischeri*–squid interactions [5]. Insects from the infraorder Pentatomomorpha provide models to study mechanisms governing the establishment of symbiosis by environmental transmission. Many phytophagous Pentatomomorpha have a midgut with a distinctive architecture, which harbors dense populations of a single bacterial species in the crypts of a specialized, posterior gut region, called M4 [6]. Although species from the Pentatomoidea superfamily typically transmit their γ-proteobacterial M4 crypt symbionts vertically, members of the superfamilies Lygaeoidea, Coreoidea, and the family Largidae (within the Pyrrhocoroidea superfamily) harbor environmentally acquired species of the *Burkholderia s.l.* (*sensu lato*) that includes the genera *Caballeronia*, *Paraburkholderia*, and *Burkholderia*. The association of the Pentatomomorpha insects with crypt symbionts increases their fitness by enhancing developmental rate and body size, survival, reproduction, immunity, and capacity for degradation of xenobiotics [7-19].

In the bean bug *Riptortus pedestris*, belonging to the Coreoidea, *Caballeronia* dominates the gut microbiota with > 95% prevalence [7]. The gut features a constricted region (CR) between midgut sections M3 and M4 that acts as a selective barrier, permitting only motile and compatible bacteria, such as *Caballeronia insecticola*, to colonize the crypts [20]. This gate closes shortly after initial colonization, preventing late-arriving bacteria from entering [21]. Within the crypts, bacterial competition further filters symbionts, favoring strains with

robust metabolic capabilities and resistance to host-derived antimicrobial peptides (AMPs) [22-25].

Recent evidence suggests that Coreoidea species may exhibit distinct *Caballeronia* subclade-level specificity, possibly shaped by host-symbiont coevolution or alternatively, by environmental availability of *Caballeronia* subclades in soils of different geographic origin [15]. For instance, *R. pedestris* typically associates with *Caballeronia* of the α-subclade, while the western conifer seed bug *Leptoglossus occidentalis* harbors β- or δ-subclade species across global populations. In the dock bug *Coreus marginatus*, β-subclade associations dominate in Europe, while Japanese populations also carry δ-subclade symbionts (originally placed in the α-subclade, Fig. 1A) [12].

In this study, we experimentally tested whether *R. pedestris* and *C. marginatus* distinguish among *Caballeronia* subclades. First, we analyzed symbiont associations in both insects reared on the same soil. Next, we used gnotobiotic mono- and co-inoculations to assess host preferences for α-, β-, and γ-subclade strains. Finally, we identified key molecular pathways in α-subclade *Caballeronia insecticola* that promote its competitive success during gut colonization.

MATERIALS AND METHODS

Insect rearing and infection with symbionts

The *R. pedestris* TKS1 inbred line and *C. marginatus* captured in two distinct locations in Gif-sur-Yvette (48°42'17.0"N 2°07'42.2"E) and Bures-sur-Yvette (48°41'52.4"N 2°09'09.3"E), France were reared in the laboratory as described before [12, 26). Bacteria are listed in Table S1 and were cultured as described [24].

For the soil inoculation test, samples of soil from the CNRS campus (48°42'17.0"N 2°07'42.2"E), Gif-sur-Yvette, France were collected in 2019 and sieved with a 2 mm sieve to remove debris. One g of soil was wetted with 10 ml of water and applied to a cotton pad, which was used as an inoculum for 2nd instar nymphs that are most perceptive to infection by crypt symbionts [27]. Insects were reared until reaching the 3rd instar (5 days post infection (dpi) for *R. pedestris* and 7 dpi for *C. marginatus*) and their M4 midgut regions were

harvested by dissection. DNA was extracted from M4 samples and a clone library analysis of the 16S rRNA gene was performed as previously described [12]. Multiple alignment of the sequences was performed to assign them to operational taxonomical units (OTU) and to construct a molecular phylogenetic tree as described in the Supplementary Information. The nucleotide sequence data of the 16S rRNA gene obtained in the present study have been deposited in the DDBJ public database with the accession numbers LC816739-LC817101.

In the gnotobiotic infection system, a defined bacterium or set of bacteria is provided to the insects. Inoculation of insects by single fluorescently labelled strains (Table S1) or by equally abundant paired strains, each labelled with a different fluorescent protein (Table S1), was done as previously by oral administration via cotton pads, wetted with the bacterial suspension [24]. Inoculated 2nd instar nymphs of *R. pedestris* or *C. marginatus* were reared until 3rd instar as above and the symbiont colonization status was assessed in dissected guts by fluorescence microscopy or flow cytometry as described earlier [24]. In pair-wise competitions, a competition index (CI) was calculated from the flow cytometry quantification of the competing strains as the ratio of a tested GFP-strain to a mScarlet-I-strain, normalized by the ratio of the inoculum. Statistical analysis, using a Kruskal-Wallis test, Dunn post hoc test and Benjamini-Hochberg correction, was performed with R studio [28].

For fitness measurements in *R. pedestris,* insects were infected as described above with different *Caballeronia* strains and bacteria-free aposymbiotic (apo) insects were obtained by omitting bacteria in the drinking water. Survival rate and developmental time to reach adulthood were reported by daily observation. Morphological parameters were measured on acetone-dehydrated and air-dried one day old adults. Egg production in one week was counted for individual females each placed together with two males in a container as described before [29]. Statistical analysis, using a Kruskal-Wallis test, Dunn post hoc test and Benjamini-Hochberg correction, was performed with R studio [28].

Phenotypic characterization of *Caballeronia* strains

To measure the velocity of the different strains in liquid YG medium, time-lapse videos of the bacteria's movement were recorded using Leica DMI6000 B Inverted Research Microscope

(Leica Camera AG, Germany). The tracking and analysis of the motility was done using the TrackMate plugin [30] of Fiji [31].

Combined chemotaxis/motility assays were performed on swimming plates (YG medium with 0.3% agar), inoculated with the tested strains by inserting a suspension of the bacteria into the soft agar layer. Images of the plates were taken at regular time points for around 60 hours, and the diameter of the colony growth in the plate was measured on the pictures using Fiji.

For pairwise *in vitro* competition experiments, GFP- and mScarlet-tagged strains (Table S1) were mixed together in a 1:1 ratio. Strain mixtures were inoculated either in liquid MM medium with either glucose or succinate as carbon source, on standard (1.5% agar) YG plates, or on swimming plates as above. Co-cultures were grown for 24 hours at 28°C. The cells were harvested in 1X PBS buffer from across the growth for liquid cultures and solid agar plates or from the edges only of the growth zone for the swimming plates. The relative number of GFP- and mScarlet-I-labelled bacteria in the harvested cells of each sample was measured by flow cytometry and a CI was calculated as above.

Strain metabolic profiles were analyzed using four 96-well Phenotype Microarray (Biolog Inc., US) microplates, which consisted of two microplates of carbon substrates (PM1 and PM2A), one microplate of various nitrogen substrates (PM3B) and one microplate of different phosphorus and sulfur substrates (PM4A). The assays were performed following the manufacturer's instructions.

Sensitivity assays of the *Caballeronia* strains to AMPs were performed essentially as before and as detailed in the Supplementary Information [25].

Other methods and detailed procedures

Full descriptions of all above methods as well as the creation of fluorescent protein labelled strains are provided in Supplementary Information.

RESULTS

Coreus marginatus and *Riptortus pedestris* select *Caballeronia* strains of different subclades from the same soil sample

To determine whether host specificity or geographic preference of *Caballeronia* subclades is involved in the midgut colonization in Coreoidea insects, we performed a soil-inoculation experiment that mimics more closely the natural symbiont acquisition process than the standard laboratory gnotobiotic system of insect colonization with specific, selected bacterial strains. For this, 2nd instar nymphs of *R. pedestris* and *C. marginatus* were reared on the same soil suspension, prepared from a natural prairie soil sample. Bacterial diversity in the midgut crypts of the 3rd instar nymphs was investigated by generating plasmid clone libraries of 16S rRNA gene PCR products per insect specimen and sequencing of clones. In a total of 363 sequences from 22 and 17 individuals of *R. pedestris* and *C. marginatus*, respectively, 11 OTUs were obtained, which corresponded to *Caballeronia*, *Paraburkholderia*, and *Pandoraea* spp. (Table S2; Fig. S1; Fig. S2). The data suggest that the M4 of insect individuals is mostly dominated by a single phylotype (Fig. S2).

A molecular phylogenetic analysis based on the 16S rRNA gene, including the sequences of the 11 OTUs, several type strains of *Caballeronia*, *Burkholderia*, *Paraburkholderia*, and *Pandoraea*, as well as previously reported gut symbionts of different Coreoidea insects, placed the *Caballeronia* OTUs in the α-, β-, and γ-subclades (Fig. 1A). No OTU from our sampling was detected within the δ-subclade (also called "Coreoidea-clade" [15]) despite the fact that all known sequences of this subclade are derived from wild specimen of Coreoidea insects, including from Japanese specimen of *C. marginatus* [6, 12, 15, 32]. Although this clade has been detected in the midgut of Coreoidea stinkbugs, it has been rarely detected in other environments. Hence it is considered to be very low abundant in soil. Two OTUs were placed in the *Paraburkholderia* but apart from the previously defined iPBE clade, containing the gut symbionts of *Largidae* species [33-35]. Four other OTUs were in the *Pandoraea* group.

The relative proportion of the OTUs detected in the M4 midguts of *R. pedestris* and *C. marginatus* demonstrates that *Caballeronia* dominated in both insects with respectively 87.5% and 97.5% of the detected crypt strains (Fig. 1B; Table S2), in agreement with the

known specificity of these insects for *Caballeronia* spp. [7, 12]. However, strains of the α-subclade of *Caballeronia* occupied 56% in *R. pedestris*, but only 16% in *C. marginatus*. In contrast, β-strains showed the opposite tendency, with 61% detected in *C. marginatus*, but none in *R. pedestris*. The γ-strains showed 33% and 20% of occupancy in *R. pedestris* and *C. marginatus*, respectively. Thus, the dominant *Caballeronia* strains were different between *R. pedestris* and *C. marginatus* suggesting a different symbiont specificity at a fine taxonomic scale.

Colonization of gnotobiotic *Riptortus pedestris* and *Coreus marginatus* by α-, β-, and γ-subclade *Caballeronia* strains

To investigate the ability of *Caballeronia* belonging to the different subclades to colonize the midguts of *R. pedestris* and *C. marginatus*, GFP-labeled strains were independently inoculated to 2nd instar nymphs of *R. pedestris* and *C. marginatus*, and the status of midgut colonization at the 3rd instar nymphs was surveyed by epifluorescence microscopy. The investigated strains included the α-strain *C. insecticola* (isolated from the gut of a *R. pedestris* specimen), three β-strains, Cm1876 (isolated from the gut of a *C. marginatus* specimen), Lo2144 (isolated from the gut of a *L. occidentalis* specimen), and *Caballeronia sordidicola* (type strain DSM17212^{T}), and the two γ- strains *Caballeronia telluris* (type strain LMG22936^{T}) and *Caballeronia glathei* (type strain JCM10563^{T}) (Fig. 2A; Table S1) [12, 15]. The gnotobiotic infection test demonstrated that the α- and β-strains displayed an almost 100% colonization ability in the midgut crypts (proportion of tested insects that were infected in the M4 crypts) of both, *R. pedestris* and *C. marginatus*, whereas the colonization rate with the γ-strains was slightly lower than the other two *Caballeronia* subclades, particularly in *C. marginatus* (Fig. 2B). Inspection of the crypt infection by fluorescence microscopy confirmed a regular infection pattern for all tested strains in the two insects except for *C. telluris* (γ-strain) in the *C. marginatus* crypts, which were only partially infected by this strain in some individuals (Figs. S2 and S3). Together, these results show that strains of the three subclades have high colonization ability in the midgut of *R. pedestris* and *C. marginatus,* though γ-strains are slightly less efficient.

We determined if selected strains from the α-, β-, and γ-subclades provide fitness advantages to the host insect. This analysis was limited to *R. pedestris* because efficient rearing conditions are available for this insect and fitness advantages derived from the crypt symbionts have been well characterized [7, 29], while the generation of offspring in *C. marginatus* in our rearing conditions was insufficient for multiple experiments. The survival rate, developmental time to reach adulthood, body mass and size of adults, and egg production were determined in *R. pedestris* apo insects or insects infected with *C. insecticola* (α-subclade strain), *Caballeronia* strain Cm1876 (β-subclade strain), or *C. telluris* (γ-subclade strain) (Fig. 2C-K). Insects infected with the three strains showed a nearly 100% survival rate, identical to apo insects, indicating that none of the bacteria had a negative effect on the host (Fig. 2C). The different measured fitness parameters showed that all three tested strains provided an enhanced fitness to the host, compared to apo insects (Fig. 2D-K). However, whereas the developmental time and body mass were similar for the insects infected with the three *Caballeronia* strains, body size and egg production were significant lower in insects infected with *Caballeronia* strain Cm1876 or *C. telluris* compared to insects carrying the natural crypt symbiont *C. insecticola* (Fig. 2E-K). Thus overall, even if all the tested *Caballeronia* strains can provide fitness services to the host, *C. insecticola* provides superior advantages.

Caballeronia interspecies competition determines the crypt occupancy

Although *Caballeronia* of the three subclades have a good colonization ability in *R. pedestris* and *C. marginatus*, the soil inoculation test demonstrated that respectively α- and β-OTUs were dominant in these two insect species (Fig. 1). To further explore a potential selection at the *Caballeronia* subclade level, pairwise bacteria-bacteria competition assays were conducted. Equal amounts of GFP- and mScarlet-I-labelled strains belonging to different *Caballeronia* subclades (Table S1) were inoculated in growth medium or fed to *R. pedestris* and *C. marginatus* nymphs and the relative amount of both strains was determined by flow cytometry after growth in *in vitro* cultures or establishment in the M4 crypts of the insects. In liquid culture, two carbon sources were used, the glycolytic carbon source glucose and the gluconeogenic carbon source succinate. The latter nutrient was used because it was demonstrated that in the crypts, the colonizing bacteria are fed with a gluconeogenic

substrate [24, 26]. Besides revealing differences in growth rate between strains, competition in liquid cultures also test for secretion of compounds that are toxic for competing bacteria. Additionally, antagonism between competitors by contact-dependent mechanisms [36] was tested on solid agar plates that favour cell-cell contact.

For *in vitro* control competitions, GFP- and mScarlet-I-labelled *C. insecticola* clones, both derivatives of the RPE75 strain (Table S1), kept each other in close balance in the two liquid media and the solid medium (Fig. 3A-C), indicating that the introduction of fluorescent markers, carried on a Tn*7* transposon inserted in the genome, has no noticeable effect on the fitness of the bacteria. In all tested strain combinations in the three *in vitro* set-ups, strain-pairs mostly kept each other in balance with only minor deviations from equal abundance after 24h of growth ($|CI| \leq 10^1$) (Fig. 3A-C). Thus, the pairwise competition assays in culture revealed no strong growth differences or antagonisms between *Caballeronia* strains.

Contrary to the cultures, pairwise combinations in the midgut crypts of the two tested insects showed mostly strong competitive interactions whereby one strain outcompeted the other ($|CI| > 10^2$). In the case of *R. pedestris*, the β-subclade strains were outcompeted by the α-strains in the midgut crypts in all tested cases (Fig. 3D). The same pairwise competitions provided in *C. marginatus* very different outcomes. Here, the β-strain Cm1876 outcompeted the α-strain *C. insecticola*, while the two other combinations resulted in a stochastic crypt colonization outcome wherein each insect was infected almost exclusively by either one of the paired strains in a random fashion (Fig. 3E). Also in the case of β- versus γ-competitions, the outcomes were strongly contrasted in the two insect species (Fig. 3D, E). The γ-strain was significantly outcompeted by the β-strains in *C. marginatus*. In contrast, the γ-strain outcompeted two of the β-strains in *R. pedestris* while the other γ–β confrontation between *C. telluris* and strain Cm1876 was the only observed case where the two strains could co-exist in the crypts of the same insects. In both insect species, the γ-strains were outcompeted by the α-strain (Fig. 3D, E). Several of the competitions in *C. marginatus* showed a higher variability in the CI between insect individuals than in *R. pedestris*, suggesting a non-negligible influence of stochastic picking up of symbionts in the former species.

Collectively, the above results demonstrate that the two stinkbug species have different selectivity for the *Caballeronia* subclade strains that they host in the M4 midgut crypts. Overall, the specificity can be summarized as $\alpha > \gamma > \beta$ for *R. pedestris* and $\beta > \alpha > \gamma$ for *C. marginatus* (Fig. 3), which is in very good correspondence to the specificity suggested by the soil inoculation experiment (Fig. 1).

Dynamics of interspecies competitions reveal different underpinning mechanisms

Out-competition of one strain by another in co-infection could be the result of priority effects whereby the timing of arrival in the crypt area will determine which strain will win the competition or alternatively, competing strains might initially co-exist in the crypts followed by displacement of one of the strains. In previous confrontation experiments between the α-strain *C. insecticola* and more distantly related *Paraburkholderia* or *Pandoraea* spp., it was shown that initially strains coexisted in the crypts and that *C. insecticola* subsequently displaced the competitor [22]. We analyzed in four of the above interactions if we could detect evidence for priority effects or displacement mechanisms. The dynamics of competition during M4 crypt colonization in *R. pedestris* was followed over time after co-infections of the insects with *C. insecticola* (α-strain) together with either strain Cm1876 (β-strain), *C. sordidicola* (β-strain), strain Lo2144 (β-strain), or *C. glathei* (γ-strain) (Fig. 4). The M4 community was analyzed at an early time point (1.5 dpi), an intermediate time point (3 dpi) and a late time point (5 dpi). Overall, three patterns of colonization dynamics could be distinguished. In the case of the *C. insecticola*/Cm1876, initially the two strains co-exist in the midgut crypts and gradually, *C. insecticola* displaces the Cm1876 strain (Fig. 4A). In contrast, in the case of the *C. insecticola*/*C. sordidicola* and *C. insecticola*/*C. glathei* interactions, *C. insecticola* dominates from the early stages, a pattern that suggests a priority effect (Fig. 4B, C). Finally, the colonization pattern of the *C. insecticola*/Lo2144 competition suggest a mixed mechanism whereby a priority effect leads to a majority of either one of the strains in the crypts followed by a displacement of the Lo2144 strain by *C. insecticola* (Fig. 4D).

Bacterial factors contributing to interspecies competition

Several molecular mechanisms underlying the efficient *R. pedestris* M4 midgut colonization capacity of *C. insecticola* have been revealed. Among them are the use of host-derived carbon sources via gluconeogenesis, chemotaxis and resistance to AMPs. Mutants of *C. insecticola* in the genes *fbp* and *pps* (gluconeogenesis)[24], *cheA* (chemotaxis; Fig. S4)[37], or *wzm* and *tpr* (AMP resistance)[25] are still capable to colonize the *R. pedestris* M4 crypts but less efficiently than the parental wild-type strain and these mutants are fully outcompeted by the wild type. To determine whether factors controlled by these five genes are also involved in interspecific bacteria-bacteria competition in the midgut of *R. pedestris,* we conducted competition assays in *R. pedestris* confronting the five mutants with the γ-strain *C. telluris*. As above, the wild-type *C. insecticola* strongly outcompeted *C. telluris* while all five mutants were severely affected to various degrees (Fig. 5A). This result indicates that the superior performance of *C. insecticola*, compared to *C. telluris*, for *R. pedestris* midgut colonization is multifactorial, depending at least on nutritional, stress-response, and chemotactic adaptations. To further elaborate this hypothesis, we compared these three parameters directly in the members of our *Caballeronia* test panel.

The swimming velocity of the six strains was determined in exponential phase bacteria growing in rich medium using videos taken by fluorescence microscopy. All strains were motile showing moderate variations in the mean velocity (Fig. S5A). Next, the chemotactic performance of the strains was compared using swimming plates in which bacteria are inoculated in soft agar and move away from the inoculation point through their combined motility and chemotactic activity. The measurement of the diameter of growing colonies in function of time highlighted differences between strains, suggesting variable chemotactic abilities (Fig. S5B). Pairwise competitions on swimming plates confirmed that the more efficient chemotactic strains outcompeted the less efficient ones at the front of the growing community (Fig. 5B). The competition indexes on the swimming plates correlated well with the corresponding competition indexes in the gut of *R. pedestris* (Fig. 3D), suggesting that chemotaxis is a strong contributor to the specificity in this species. In contrast, the pattern of competition indexes in the swimming plates (Fig. 5B) is different from the competition indexes in the *C. marginatus* gut (Fig. 3E), suggesting that in this insect, mechanisms, other than motility and chemotaxis, are more important in symbiont selection.

354 The metabolic capacities of the strains in the test panel were profiled with Biolog Phenotype
355 MicroArrays [38]. Testing the use of hundreds of potential carbon, nitrogen, phosphorous
356 and sulfur sources revealed that the six strains have distinguishable metabolite usage
357 patterns whereby the three tested β-*Caballeronia* strains are most similar to each other
358 while the α- and two γ-subclade strains are each different (Fig. S6). The γ-subclade strain *C.*
359 *telluris* seems to metabolize the largest set of compounds and only few compounds that are
360 metabolized by the other strains are not processed by it (Fig. 5C, top cluster). Nevertheless,
361 this strain is not dominating the α-subclade strain in *R. pedestris* or the β-*Caballeronia*
362 strains in *C. marginatus* (Fig. 3D, E), indicating that the superior metabolic capacity of *C.*
363 *telluris* is not sufficient for dominating in the crypts of the two stinkbug species. In contrast,
364 the α-strain *C. insecticola* metabolizes several compounds that are not processed by the β-
365 *Caballeronia* strains (Fig. 5C, middle cluster), which could contribute to the superior
366 colonization capacity of *C. insecticola* in *R. pedestris* relative to the β-strains. At the opposite,
367 the β-strains have also metabolic capacities that are absent in *C. insecticola* (Fig. 5C, bottom
368 cluster) but as these capacities are mostly shared by the tested β-strains, while they do not
369 have the same competitive advantage to *C. insecticola* in the colonization of the *C.*
370 *marginatus* gut (Fig. 3E), it is not straightforward to correlate these metabolic capacities with
371 gut colonization. Taken together, the *Caballeronia* strains have distinct metabolic activities,
372 which could contribute to the efficiency of crypt colonization although, in the absence of
373 knowledge on the specific identify of crypt metabolites, they are not easily correlated with
374 the observed competitive behavior in the *R. pedestris* and *C. marginatus* crypts.

375 We tested the resistance of the *Caballeronia* strains of the test panel against a set of AMPs,
376 including CCR AMPs produced in the posterior midgut of *R. pedestris* [25], together with the
377 innate immunity-related AMPs thanatin and riptocin [19], mammalian LL37 [39], plant
378 NCR335 [40], and bacterial polymyxin B (PMB). The three β-strains as well as the γ-clade
379 strain *C. telluris* were relatively more sensitive to the tested peptides, and in particular to the
380 CCR peptides, than the α-strain *C. insecticola* or the γ-strain *C. glathei* (Fig. 5D). Thus, AMP
381 resistance can contribute to the dominance of *C. insecticola* in the gut crypts relative to the
382 β-strains or the γ-strain *C. telluris*. In contrast, the dominance of *C. insecticola* over *C. glathei*
383 is probably independent of AMP resistance because both strains have similar resistance
384 patterns. Moreover, the β-strain Cm1876 outcompetes *C. insecticola* in the crypts of *C.*

marginatus, despite its lower resistance to AMPs. Thus, the dominance of the β-*Caballeronia* strain Cm1876 in the *C. marginatus* crypts seems to be independent of its overall metabolic performance, motility and chemotaxis, and AMP resistance, which suggests that this bacterium has a yet to be identified adaptation for the crypt colonization in this particular host.

DISCUSSION

Field-collected *R. pedestris* and *C. marginatus* are both predominantly colonized by *Caballeronia* spp. in their M4 midgut region [6, 7, 12, 41], a pattern confirmed in our study in insects of these two species reared on the same French soil sample. Our result also matches earlier observations in *R. pedestris* with soil samples from South Korea and Japan [42]. However, even if the two insect species displayed the same specificity at the genus level, we discovered a marked difference between them at a finer taxonomic level. Whereas α-*Caballeronia* dominated in *R. pedestris* infected with soil bacteria, β-*Caballeronia* prevailed in *C. marginatus*. Gnotobiotic infection experiments corroborated these species-dependent specificities for *Caballeronia* subclades. Mono-inoculation experiments in the gnotobiotic system showed strains of all *Caballeronia* subclades could colonize either host. However, co-inoculation revealed competitive hierarchies between subclade strains: the α-*Caballeronia* are more competitive than the tested β- and γ-strains, and the γ-strains more than β-strains to colonize the symbiotic organ of *R. pedestris*; in contrast, β-strains have an advantage for the colonisation of *C. marginatus* over α- and γ-strains, and α- over γ-strains. These findings confirm that species-specific subclade specificity, observed in natural specimens [12-15], are at least driven by host-bacteria compatibility and partner-choice mechanisms. We observed that less competitive strains still colonized some individuals reared on soil, likely due to ecological drift when environmental symbiont abundance is low and insects can pick up stochastically less-preferred symbiont strains [43]. Possibly, also in the gnotobiotic infection system in *C. marginatus*, stochastic symbiont selection is happening as suggested by the observed variability in the competition experiments between several of the tested strain pairs. The importance of ecological drift suggests that besides partner choice mechanisms, geographic distribution of *Caballeronia* subclades can also contribute to observed colonization patterns in wild insects.

Subclade-level specificity is also seen in other Coreoidea bugs. For example, *Leptoglossus occidentalis* typically hosts β- or δ-*Caballeronia*, even when co-infected with α-strains [15]. However, specificity varies within the Pentatomomorpha. Largidae species for example are selectively colonized by *Paraburkholderia* of the iPBE subclade rather than *Caballeronia* [33-35], whereas Blissidae and Bertidae, are more permissive, accepting species of multiple genera of the *Burkholderia s.l.* [6, 13, 44-47].

Symbiont specificity within the Pentatomomorpha echoes other environmentally transmitted symbioses, like rhizobia-legume or squid-*Aliivibrio* associations, where host-symbiont compatibility is mediated by molecular signaling and variations on the conserved theme of signals account for symbiont specificity [48, 49]. Likewise, we can expect that signal production and recognition are fundamental in the selection of the midgut crypt symbionts by the stinkbugs. Though the specific chemical cues in stinkbug-*Caballeronia* interactions remain unidentified and their identification is a crucial future challenge, our findings suggest that tested *Caballeronia* strains produce conserved signals while subtle differences in signal production and receptor recognition between subclades may guide symbiont selection. The here inferred chemotactic cues produced in the gut and their recognition by chemotaxis receptors of the gut bacteria may constitute one example of a signaling system controlling the stinkbug-*Caballeronia* symbiosis.

Beyond signaling, we show here that microbial competition in the gut strongly shapes the strain composition in the crypts. Crypt colonization begins with a bottleneck at the constricted region (CR), limiting passage to a few thousand bacteria into the crypt region [20, 21]. Rapid expansion of this founding population to ~10^7 cells follows, offering two key windows for competition to operate: early priority effects during CR passage and establishment of the founding population and later displacement during outgrowth of the founding population.

Microorganisms generally compete in two ways: indirectly, through adaptation to the environment, or directly, by harming each other via chemical warfare [50]. Direct competition involves secreting antimicrobials like bacteriocins or injecting toxins using systems such as Type VI secretion (T6SS) or contact-dependent inhibition [36]. Although culture experiments showed no clear antagonism between tested strains, antimicrobial or toxin-mediated competition still may occur in the midgut crypts. For instance, *C. insecticola*

possesses a T6SS that could target competing strains in the crypts, though evidence is still lacking: T6SS genes were downregulated in single-strain infections [26], and T6SS mutants showed no competitive disadvantage against *Pandoraea* or *Paraburkholderia* strains [22].

In contrast to direct competition, indirect competition likely plays a dominant role in midgut crypt colonization and here we identified three mechanisms. One of them is chemotaxis, guiding bacteria toward the M4 midgut region. Flagellar motility is essential for entry into the crypt region in *R. pedestris* through the CR [20; 51, 52]. We found that a *cheA* mutant of *C. insecticola*, though able to colonize alone, is outcompeted by its wild type and by *C. telluris*—which itself is usually outcompeted by the wild-type *C. insecticola*. This suggests that chemotactic efficiency influences competitive success. Supporting this, *Caballeronia* strains of the test panel showed significant differences in motility and chemotaxis, which correlated with competition outcomes *in vitro* that resembled strongly competition outcomes in the crypts of *R. pedestris* (Fig. 5B vs. Fig. 3D). Priority effects observed during infection may result from faster strains reaching the crypts earlier and establishing dominance, especially as the CR closes after the passage of the initial colonizers [21], preventing late arrivers from entering.

Indirect competition may also depend on the ability of strains to adapt to the crypt environment, affecting their growth within the crypt community. Because crypt bacteria feed on metabolites secreted by the crypt epithelium rather than ingested food, due to the crypt's disconnection from the anterior digestive midgut region [20, 26], nutrient availability in crypts is probably limited in both quantity and diversity. Thus, bacterial growth competition, particularly in the case of displacement of one strain by another, likely hinges on their metabolic gene repertoire for harvesting crypt nutrients. Metabolic phenotyping of the tested *Caballeronia* strains indeed revealed their distinct capabilities. For example, we previously showed that gluconeogenesis and taurine and inositol harvesting contribute to *C. insecticola*'s gut fitness, with mutants in these pathways able to colonize but outcompeted by wild-type strains [24]. Here, we demonstrate that gluconeogenesis mutants also lose the ability to outcompete an otherwise less efficient *Caballeronia* s, supporting the idea that available nutrient resources in the crypts and matching metabolic capacities in the crypt colonizers determine the outcome of competitions. Differences in crypt nutrients between *R. pedestris* and *C. marginatus*, which feed on soybean seeds [53] and *Rumex* seeds [54]

478 respectively, likely contribute to their specificity for different *Caballeronia* subclades, as their
479 distinct diets are expected to influence crypt metabolite composition.

480 Stress conditions in the crypt environment can limit bacterial growth and thus that
481 potentially regulates interspecies competitions. The *R. pedestris* crypts produce massively a
482 high variety of antimicrobial peptides, known as crypt-specific cysteine rich peptides (CCRs)
483 that can kill or inhibit growth of bacteria [25]. Typical crypt colonizing bacteria are resistant
484 to these peptides but CCR-sensitive mutants of *C. insecticola*, such as the here tested *wzm*
485 and *tpr* mutants, have a strongly reduced fitness for crypt colonization and are outcompeted
486 by wild-type *C. insecticola* or by a competing species [25; this study]. The tested *Caballeronia*
487 strains have distinct resistance profiles to CCRs. Overall, this suggest that the adaptation of
488 candidate gut symbionts to grow in the presence of the CCRs contributes to the selection
489 process during competitive crypt colonization. Moreover, *R. pedestris* and *C. marginatus*
490 possibly produce different panels of CCRs in order to select their matching microbial
491 partners.

492 The majority of our tested competitions results in a "winner takes all" outcome, with one of
493 the competitors dominating and the other one nearly eliminated ($|CI| > 10^2$). This
494 observation reflects well the colonization pattern seen in insects reared on soil or in wild-
495 captured insects of different bug species, which have midgut crypts that are mostly
496 dominated by a single OTU [this work; 12, 13, 33, 29, 42, 45, 55]. Because the bacterial
497 population in the crypts is generated by a small founding population that multiplies within
498 the crypts [21], the nearly mono-strain composition of the final crypt population needs to be
499 explained by either strong differences in the number of bacteria of competing strains
500 composing the crypt founding community, despite an equal composition in the inoculum in
501 our experimental competitions (e.g., due to different efficiency of strains in passage through
502 the CR) or by strong differences in the growth rate of the strains within the crypts (e.g., due
503 to differences in the adaptations to the nutritional or stress conditions in the crypts).

504 Although all tested strains provide an enhanced fitness to the host compared to apo insects,
505 the most competitive symbiont of *R. pedestris* provides higher reproductive benefits to the
506 insect than the less favoured strains, despite the fact that the selection of the symbiont is
507 happening in the early life of the host, well before the host can assess the benefits received
508 from the symbiont during adulthood and reproduction. Other studies on related bug-

Caballeronia associations reported a nearly the same degree of host benefit provided by more or less distantly related symbionts [8, 22, 56]. However, these studies measured fitness parameters like survival, developmental time, and body size, while the strongest difference that we have detected between strains is egg production, which was not compared between symbiont strains in the other studies. Possibly, determining the reproduction fitness in other strain comparisons, or other parameters like immune priming [19], could reveal differences in strain performances also in other cases. Admittedly, our observation of larger fitness provided by the most competitive strain is based on comparing a small number of strains comprising one competitive and two less-competitive ones. The hypothesis of a connection between symbiont selection and fitness will require further testing involving a larger panel of strains.

The possible selection of optimal symbionts by *R. pedestris* contrasts starkly with the selection of rhizobium symbionts by legumes, which are unable to pick out from the soil efficient nitrogen fixing symbionts from inefficient strains because the nitrogen fixation *per se* is not active during the selection and infection process but only at the late stage of the interaction when the nodule is formed and occupied with the selected rhizobium strain [57, 58]. However, at the level of the whole plant that carries many nodules, the host applies sanctions to the least efficient nodules. Because sanctioned nodules contain fewer viable bacteria than non-sanctioned ones, efficient nitrogen-fixing symbionts nevertheless progressively outcompete less efficient strains [59, 60]. The proposed alignment of the symbiont-host specificity with the host benefits obtained from the symbiont in *R. pedestris* suggests a different scenario. Contrary to the rhizobia-legume symbiosis, the insects can potentially monitor already in the infecting bacteria the capacity to produce the services that will contribute ultimately to the host fitness at a later stage. These services are at present not clearly defined at the molecular level and their identification will be helpful to better understand the details of symbiont selection in these insects. In addition, the crypt symbionts could not only deliver specific metabolites to the host but also need to provide bacterial biomass to replace crypt bacteria that are consumed in the M4B region, an anterior sub-region of M4 whose function is to digest symbiont cells and absorb the derived nutrients that support insect development [26]. In that case, a *Caballeronia* strain with higher proliferation efficiency in the crypts could dominate the bacteria-bacteria competition

within the crypts in the early stage of the interaction and supply in a late stage substantial more bacterial biomass, resulting in more nutritional benefits to the host.

ACKNOWLEDGEMENTS

We thank Minhyung Jung from the Kikuchi lab for helpful comments on the manuscript. This work benefited from financial support by Saclay Plant Sciences-SPS to P.M., by the ANR grant ANR-19-CE20-0007 to P.M., and by a JSPS-CNRS Bilateral Open Partnership Joint Research Project (18KK0211) and a CNRS International Research Project to Y.K. and P.M. The present work has benefited from Imagerie-Gif core facility supported by l'Agence Nationale de la Recherche (ANR-11-EQPX-0029/Morphoscope, ANR-10-INBS-04/FranceBioImaging; ANR-11-IDEX-0003-02/ Saclay Plant Sciences). G.L. was supported by a Ph.D. fellowship from the French Ministry of Higher Education, Research, and Innovation and T.O by JSPS Research Fellowships for Young Scientist (14J03996, 20170267 and 19J01106).

AUTHOR CONTRIBUTIONS

PM, TO, YK, and GL conceived and designed the study. GL , SJ, RC, YK, TO, and PM performed the experiments and analysed the results. PM wrote the paper and all authors amended and commented on the various versions of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

DATA AVAILABILITY

The nucleotide sequence data of the 16S rRNA gene obtained in the present study have been deposited in the DDBJ public database with the accession numbers LC816739-LC817101. All other data generated or analyzed during this study are included in this published article and its supplementary information files.

References

1. Fisher RM, Henry LM, Cornwallis CK *et al.* The evolution of host-symbiont dependence. *Nat Commun* 2017;**8**:15973. https://doi.org/10.1038/ncomms15973
2. Cornwallis CK, van 't Padje A, Ellers J *et al.* Symbioses shape feeding niches and diversification across insects. *Nat Ecol Evol* 2023;**7**:1022-44. https://doi.org/10.1038/s41559-023-02058-0
3. Chouvenc T, Šobotník J, Engel MS *et al*. Termite evolution: mutualistic associations, key innovations, and the rise of Termitidae. *Cell Mol Life Sci* 2021;**78**:2749-69. https://doi.org/10.1007/s00018-020-03728-z
4. García-Lozano M, Henzler C, Porras MÁG *et al.* Paleocene origin of a streamlined digestive symbiosis in leaf beetles. *Curr Biol* 2024;**34**:1621-34.e9. https://doi.org/10.1016/j.cub.2024.01.070
5. Stubbendieck RM, Li H, Currie CR. Convergent evolution of signal-structure interfaces for maintaining symbioses. *Curr Opin Microbiol* 2019;**50**:71-8. https://doi.org/10.1016/j.mib.2019.10.001
6. Kikuchi Y, Hosokawa T, Fukatsu T. An ancient but promiscuous host-symbiont association between *Burkholderia* gut symbionts and their heteropteran hosts. *ISME J* 2011;**5**:446-60. https://doi.org/10.1038/ismej.2010.150
7. Kikuchi Y, Hosokawa T, Fukatsu T. Insect-microbe mutualism without vertical transmission: a stinkbug acquires a beneficial gut symbiont from the environment every generation. *Appl Environ Microbiol* 2007;**73**:4308-16. https://doi.org/10.1128/AEM.00067-07
8. Hunter MS, Umanzor EF, Kelly SE *et al.* Development of common leaf-footed bug pests depends on the presence and identity of their environmentally acquired symbionts. *Appl Environ Microbiol* 2022;**88**:e0177821. https://doi.org/10.1128/AEM.01778-21
9. Koga R, Moriyama M, Onodera-Tanifuji N *et al.* Single mutation makes *Escherichia coli* an insect mutualist. *Nat Microbiol* 2022;**7**:1141-50. https://doi.org/10.1038/s41564-022-01179-9
10. Hosokawa T, Kikuchi Y, Nikoh N *et al.* Strict host-symbiont cospeciation and reductive genome evolution in insect gut bacteria. *PLoS Biol* 2006;**4**:e337. https://doi.org/10.1371/journal.pbio.0040337
11. Hosokawa T, Ishii Y, Nikoh N *et al.* Obligate bacterial mutualists evolving from environmental bacteria in natural insect populations. *Nat Microbiol* 2016;**1**:15011. https://doi.org/10.1038/nmicrobiol.2015.11
12. Ohbayashi T, Itoh H, Lachat J *et al. Burkholderia* Gut symbionts associated with European and Japanese populations of the dock bug *Coreus marginatus* (Coreoidea: Coreidae). *Microbes Environ* 2019;**34**:219-22. https://doi.org/10.1264/jsme2.ME19011
13. Ravenscraft A, Thairu MW, Hansen AK *et al*. Continent-scale sampling reveals fine-scale turnover in a beneficial bug symbiont. *Front Microbiol* 2020;**11**:1276. https://doi.org/10.3389/fmicb.2020.01276
14. Acevedo TS, Fricker GP, Garcia JR *et al.* The importance of environmentally acquired bacterial symbionts for the squash bug (*Anasa tristis*), a significant agricultural pest. *Front Microbiol* 2021;**12**:719112. https://doi.org/10.3389/fmicb.2021.719112
15. Ohbayashi T, Cossard R, Lextrait G *et al.* Intercontinental diversity of *Caballeronia* gut symbionts in the conifer pest bug *Leptoglossus occidentalis*. *Microbes Environ* 2022;**37**:ME22042. https://doi.org/10.1264/jsme2.ME22042

16. Lee JB, Park KE, Lee SA *et al.* Gut symbiotic bacteria stimulate insect growth and egg production by modulating hexamerin and vitellogenin gene expression. *Dev Comp Immunol* 2017;**69**:12-22. https://doi.org/10.1016/j.dci.2016.11.019
17. Kikuchi Y, Fukatsu T. Live imaging of symbiosis: spatiotemporal infection dynamics of a GFP-labelled *Burkholderia* symbiont in the bean bug *Riptortus pedestris*. *Mol Ecol* 2014;**23**:1445-56. https://doi.org/10.1111/mec.12479
18. Kikuchi Y, Hayatsu M, Hosokawa T *et al.* Symbiont-mediated insecticide resistance. *Proc Natl Acad Sci U S A* 2012;**109**:8618-22. https://doi.org/10.1073/pnas.1200231109
19. Kim JK, Lee JB, Huh YR *et al. Burkholderia* gut symbionts enhance the innate immunity of host *Riptortus pedestris*. *Dev Comp Immunol* 2015;**53**:265-9. https://doi.org/10.1016/j.dci.2015.07.006
20. Ohbayashi T, Takeshita K, Kitagawa W *et al.* Insect's intestinal organ for symbiont sorting. *Proc Natl Acad Sci U S A* 2015;**112**:E5179-88. https://doi.org/10.1073/pnas.1511454112
21. Kikuchi Y, Ohbayashi T, Jang S *et al*. *Burkholderia insecticola* triggers midgut closure in the bean bug *Riptortus pedestris* to prevent secondary bacterial infections of midgut crypts. *ISME J* 2020;**14**:1627-38. https://doi.org/10.1038/s41396-020-0633-3
22. Itoh H, Jang S, Takeshita K *et al.* Host-symbiont specificity determined by microbe-microbe competition in an insect gut. *Proc Natl Acad Sci U S A* 2019;**116**:22673-82. https://doi.org/10.1073/pnas.1912397116
23. Kim JK, Jang HA, Won YJ *et al.* Purine biosynthesis-deficient *Burkholderia* mutants are incapable of symbiotic accommodation in the stinkbug. *ISME J* 2014;**8**:552-63. https://doi.org/10.1038/ismej.2013.168
24. Jouan R, Lextrait G, Lachat J *et al.* Transposon sequencing reveals the essential gene set and genes enabling gut symbiosis in the insect symbiont *Caballeronia insecticola*. *ISME Commun* 2024;**4**:ycad001. https://doi.org/10.1093/ismeco/ycad001
25. Lachat J, Lextrait G, Jouan R *et al.* Hundreds of antimicrobial peptides create a selective barrier for insect gut symbionts. *Proc Natl Acad Sci U S A* 2024;**121**:e2401802121. https://doi.org/10.1073/pnas.2401802121
26. Ohbayashi T, Futahashi R, Terashima M *et al.* Comparative cytology, physiology and transcriptomics of *Burkholderia insecticola* in symbiosis with the bean bug *Riptortus pedestris* and in culture. *ISME J* 2019;**13**:1469-83. https://doi.org/10.1038/s41396-019-0361-8
27. Kikuchi Y, Hosokawa T, Fukatsu T. Specific developmental window for establishment of an insect-microbe gut symbiosis. *Appl Environ Microbiol* 2011;**77**:4075-81. https://doi.org/10.1128/AEM.00358-11
28. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2021. https://www.R-project.org/
29. Jang S, Ishigami K, Mergaert P *et al*. Ingested soil bacteria breach gut epithelia and prime systemic immunity in an insect. *Proc Natl Acad Sci U S A* 2024;**121**:e2315540121. https://doi.org/10.1073/pnas.2315540121
30. Ershov D, Phan MS, Pylvänäinen JW *et al.* TrackMate 7: integrating state-of-the-art segmentation algorithms into tracking pipelines. *Nat Methods* 2022;**19**:829-32. https://doi.org/10.1038/s41592-022-01507-1
31. Schindelin J, Arganda-Carreras I, Frise E *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* 2012;**9**:676-82. https://doi.org/10.1038/nmeth.2019

32. Sullivan LT, Kelly SE, Ravenscraft A *et al*. Acquisition of an obligate environmental symbiont may be limited in the arboreal environment. *FEMS Microbiol Ecol* 2025;**101**:fiaf045. https://doi.org/10.1093/femsec/fiaf045
33. Takeshita K, Matsuura Y, Itoh H *et al. Burkholderia* of plant-beneficial group are symbiotically associated with bordered plant bugs (Heteroptera: Pyrrhocoroidea: Largidae). *Microbes Environ* 2015;**30**:321-9. https://doi.org/10.1264/jsme2.ME15153
34. Gordon ER, McFrederick Q, Weirauch C. Phylogenetic evidence for ancient and persistent environmental symbiont reacquisition in Largidae (Hemiptera: Heteroptera). *Appl Environ Microbiol* 2016;**82**:7123-33. https://doi.org/10.1128/AEM.02114-16
35. Yashima R, Terata Y, Sakamoto K *et al. Paraburkholderia largidicola* sp. nov., a gut symbiont of the bordered plant bug *Physopelta gutta*. *Int J Syst Evol Microbiol* 2024;**74**:006411. https://doi.org/10.1099/ijsem.0.006411
36. García-Bayona L, Comstock LE. Bacterial antagonism in host-associated microbial communities. *Science* 2018;**361**:eaat2456. https://doi.org/10.1126/science.aat2456
37. Ishigami K, Lirette A-O, Kikuchi Y. Evolution in spatiotemporal infection patterns of *Burkholderia sensu lato* lineages in the gut of *Riptortus pedestris*. submitted to *Appl Environ Microbiol*
38. Mackie AM, Hassan KA, Paulsen IT *et al*. Biolog Phenotype Microarrays for phenotypic characterization of microbial cells. *Methods Mol Biol* 2014;**1096**:123-30. https://doi.org/10.1007/978-1-62703-712-9_10
39. Burton MF, Steel PG. The chemistry and biology of LL-37. *Nat Prod Rep* 2009;**26**:1572-84. https://doi.org/10.1039/b912533g
40. Farkas A, Maróti G, Kereszt A *et al*. Comparative analysis of the bacterial membrane disruption effect of two natural plant antimicrobial peptides. *Front Microbiol* 2017;**8**:51. https://doi.org/10.3389/fmicb.2017.00051
41. Kikuchi Y, Meng XY, Fukatsu T. Gut symbiotic bacteria of the genus *Burkholderia* in the broad-headed bugs *Riptortus clavatus* and *Leptocorisa chinensis* (Heteroptera: Alydidae). *Appl Environ Microbiol* 2005;**71**:4035-43. https://doi.org/10.1128/AEM.71.7.4035-4043.2005
42. Gook DH, Jung M, Kim S *et al*. Species diversity of environmentally-transmitted bacteria colonizing *Riptortus pedestris* (Hemiptera: Alydidae) and symbiotic effects of the most dominant bacteria. *Sci Rep* 2023;**13**:15166. https://doi.org/10.1038/s41598-023-42419-0
43. Chen JZ, Kwong Z, Gerardo NM *et al*. Ecological drift during colonization drives within-host and between-host heterogeneity in an animal-associated symbiont. *PLoS Biol* 2024;**22**:e3002304. https://doi.org/10.1371/journal.pbio.3002304
44. Boucias DG, Garcia-Maruniak A, Cherry R *et al.* Detection and characterization of bacterial symbionts in the Heteropteran, *Blissus insularis*. *FEMS Microbiol Ecol* 2012;**82**:629-41. https://doi.org/10.1111/j.1574-6941.2012.01433.x
45. Itoh H, Aita M, Nagayama A *et al.* Evidence of environmental and vertical transmission of *Burkholderia* symbionts in the oriental chinch bug, *Cavelerius saccharivorus* (Heteroptera: Blissidae). *Appl Environ Microbiol* 2014;**80**:5974-83. https://doi.org/10.1128/AEM.01087-14
46. Xu Y, Buss EA, Boucias DG. Culturing and characterization of gut symbiont *Burkholderia* spp. from the southern chinch bug, *Blissus insularis* (Hemiptera: Blissidae). *Appl Environ Microbiol* 2016;**82**:3319-30. https://doi.org/10.1128/AEM.00367-16

47. Xu Y, Buss EA, Boucias DG. Environmental transmission of the gut symbiont *Burkholderia* to phloem-feeding *Blissus insularis*. *PLoS One* 2016;**11**:e0161699. https://doi.org/10.1371/journal.pone.0161699
48. Poole P, Ramachandran V, Terpolilli J. Rhizobia: from saprophytes to endosymbionts. *Nat Rev Microbiol* 2018;**16**:291-303. https://doi.org/10.1038/nrmicro.2017.171
49. Visick KL, Stabb EV, Ruby EG. A lasting symbiosis: how *Vibrio fischeri* finds a squid partner and persists within its natural host. *Nat Rev Microbiol* 2021;**19**:654-65. https://doi.org/10.1038/s41579-021-00557-0
50. Ghoul M, Mitri S. The ecology and evolution of microbial competition. *Trends Microbiol* 2016;**24**:833-45. https://doi.org/10.1016/j.tim.2016.06.011
51. Kinosita Y, Kikuchi Y, Mikami N *et al.* Unforeseen swimming and gliding mode of an insect gut symbiont, *Burkholderia* sp. RPE64, with wrapping of the flagella around its cell body. *ISME J* 2018;**12**:838-48. https://doi.org/10.1038/s41396-017-0010-z
52. Yoshioka A, Shimada YY, Omori T *et al.* Bacteria break through one-micrometer-square passages by flagellar wrapping. *bioRxiv* 2025;2025.02.12.637796. https://doi.org/10.1101/2025.02.12.637796
53. Mainali BP, Kim HJ, Yoon YN *et al.* Evaluation of different leguminous seeds as food sources for the bean bug *Riptortus pedestris*. *J Asia-Pac Entomol*. 2014;**17**:115–7. https://doi.org/10.1016/j.aspen.2013.11.007
54. Hrušková M, Honěk A, Pekár S. *Coreus marginatus* (Heteroptera: Coreidae) as a natural enemy of *Rumex obtusifolius* (Polygonaceae). *Acta Oecologica* 2005;**28**:281–7. https://doi.org/10.1016/j.actao.2005.05.004
55. Itoh H, Hori T, Sato Y *et al.* Infection dynamics of insecticide-degrading symbionts from soil to insects in response to insecticide spraying. ISME J. 2018;**12**:909-20. https://doi.org/10.1038/s41396-017-0021-9
56. Stoy KS, Chavez J, De Las Casas V *et al.* Evaluating coevolution in a horizontally transmitted mutualism. *Evolution* 2023;**77**:166-85. https://doi.org/10.1093/evolut/qpac009
57. Bourion V, Heulin-Gotty K, Aubert V *et al.* Co-inoculation of a pea core-collection with diverse rhizobial strains shows competitiveness for nodulation and efficiency of nitrogen fixation are distinct traits in the interaction. *Front Plant Sci* 2018;**8**:2249. https://doi.org/10.3389/fpls.2017.02249
58. Westhoek A, Field E, Rehling F *et al.* Policing the legume-*Rhizobium* symbiosis: a critical test of partner choice. *Sci Rep* 2017;**7**:1419. https://doi.org/10.1038/s41598-017-01634-2
59. Daubech B, Remigi P, Doin de Moura G *et al.* Spatio-temporal control of mutualism in legumes helps spread symbiotic nitrogen fixation. *Elife* 2017;**6**:e28683. https://doi.org/10.7554/eLife.28683
60. Westhoek A, Clark LJ, Culbert M *et al.* Conditional sanctioning in a legume-*Rhizobium* mutualism. *Proc Natl Acad Sci U S A* 2021;**118**:e2025760118. https://doi.org/10.1073/pnas.2025760118

Figure legends

Figure 1. Molecular phylogenetic analysis of gut symbionts of *Riptortus pedestris* and *Coreus marginatus*. A. Maximum-likelihood tree of 1,285 aligned nucleotide sites of the 16S rRNA gene. Numbers at the tree nodes indicate bootstrap values (%) with 1,000 replicates and bootstrap values of more than 50 are shown. Accession numbers in the DNA database (DDBJ/EMBL/GenBank) are shown in square brackets. Gut symbionts of *R. pedestris* (*Rp*) and *C. marginatus* (*Cm*) identified in the soil inoculation test are shown in blue. *Caballeronia* strains used in insect inoculation tests are shown in red. The δ-subclade is also known as "Coreoidea-clade" [15]. GS: Gut symbiont. Scale bar refers to a phylogenetic distance of 0.01 nucleotide substitutions per site. **B.** Relative abundance of α-, β-, and γ-*Caballeronia*, *Paraburkholderia*, and *Pandoraea* among gut symbionts of *R. pedestris* and *C. marginatus* in the soil inoculation test. The number of investigated insects is shown on top of the graphs, and the corresponding OTUs and their counts are provided in Fig. S2 and Table S2.

Figure 2. The crypt colonization ability of *Riptortus pedestris* and *Coreus marginatus* and fitness advantages in *Riptortus pedestris* provided by *Caballeronia* strains. A. Phylogenetic tree, based on the 16S rRNA gene, of *Caballeronia* strains used in bacterial inoculation tests in *R. pedestris* and *C. marginatus*. Abbreviations of bacterial strain are provided in brackets. **B.** Second instar nymphs were inoculated with single strains as indicated. The colonization rate in the midgut crypts of *R. pedestris* and *C. marginatus* is indicated. The 3rd instar nymphs of *R. pedestris* (n = 20-59) and *C. marginatus* ($n \geq 5$) were investigated at 5 dpi and 7 dpi, respectively. The bacterial names are provided in abbreviated form as indicated in A. **C-K.** Fitness advantages in *R. pedestris* provided by *Caballeronia* strains. *R. pedestris* individuals were colonize by the indicated *Caballeronia* strains. The following fitness parameters were determined: Insect survival rate (C); Developmental time until reaching adulthood (D); number of eggs produced during the first week after commencement of egg-laying by one female reared in the presence of two males (E); Body weight of males (F); Body length of males (G); Thorax width of males (H); Body weight of females (I); Body length of females (J); Thorax width of females (K). Standard deviations are indicated by vertical bars. The numbers in the histograms indicate the number of analyzed samples n. Different letters indicate significant differences ($P < 0.05$) determined by a Kruskal-Wallis test with the Dunn post hoc test and P-value adjustment with the Benjamini-Hochberg method.

Figure 3. Competition assays of *Caballeronia* strains in culture and in insect midgut crypts. A, **B.** Competition index of *Caballeronia* strains in culture in Minimum Medium with glucose (A) or succinate (B) as carbon source (n = 4). **C.** Competition index of *Caballeronia* strains in growth in contact on solid agar plates with YG medium. **D, E.** Competition index of *Caballeronia* strains in the midgut crypts of *R. pedestris* (n = 10) (D), or *C. marginatus* (n = 5-14) (E). Midguts of *R. pedestris* and *C. marginatus* were investigated at 5 dpi and 7 dpi, respectively. The used strains are abbreviated according to the key at the bottom right. The competition index indicates the ratio of a tested β- or γ-strain to the *C. insecticola* (α) or C.

telluris (γ) strain, normalized by the ratio of the inoculum (inocula were prepared with close to equal amounts of the two competing strains). Dots and black lines show the competition index in individual insects and the mean of the replicates, respectively. Letters indicate significant differences ($P < 0.05$) using a Kruskal-Wallis test with the Dunn post hoc test and *P*-value adjustment with the Benjamini-Hochberg method.

Figure 4. Dynamics of competition of *Caballeronia* strains in midgut crypts of *Riptortus pedestris*. A. Competition index of Cm versus In. **B.** Competition index of So versus In. **C.** Competition index of Gl versus In. **D.** Competition index of Lo versus In. Competitions were determined in the midgut crypts of *R. pedestris* ($n = 10$) at 1.5, 3 and 5 dpi. The competition index indicates the ratio of the two tested strains, normalized by one of the inoculum. Dots and black lines show the competition index in individual insects and the mean of the replicates, respectively. The used strains are abbreviated according to the key at the bottom right. Different letters indicate statistically significant differences ($P < 0.05$). Statistical significance was analyzed by Kruskal–Wallis test, Dunn post hoc test and *P*-value adjustment with the Benjamini-Hochberg method.

Figure 5. Bacterial functions contributing to interspecies competition. A. Competition assay of *C. telluris* (γ) and *C. insecticola* (α) wild type (WT) and mutants in midgut crypts of *R. pedestris*. An equal mixture of mScarlet-I-labeled *C. telluris* (γ) and GFP-labeled *C. insecticola* mutants was fed to *R. pedestris* and the relative abundance of mScarlet-I- and GFP-labeled bacteria colonizing the midgut crypts at 5 dpi ($n = 10$) was determined. The competition index indicates the ratio of the *C. telluris* strain to the tested *C. insecticola* WT or mutant strains, normalized by the ratio of the inoculum. The *tpr* and *wzm* mutants have reduced AMP resistance. The *cheA* mutant has lost chemotaxis and the *fbp* and *pps* mutants cannot perform gluconeogenesis. Dots and black lines show the competition index in individual insects and the mean of the replicates, respectively. Different letters indicate statistically significant differences ($P < 0.05$). Statistical significance was analyzed by Kruskal–Wallis test, Dunn post hoc test and *P*-value adjustment with the Benjamini-Hochberg method. **B.** Combined motility and chemotaxis competition assays on 0.3% YG agar swimming plates. GFP and mScarlet-I-labeled strains were inoculated in a 1:1 ratio in the center of the plates and cells were harvested after 24h growth at the edge of the growth zone for measurement of the relative abundance of the GFP and mScarlet-I-labeled strains. **C.** Heatmaps and hierarchical cluster analysis of metabolite usage by *Caballeronia* strains using Biolog Phenotype MicroArrays. Metabolites are in rows and strains in columns. The top cluster shows the metabolic capacity of strains for all metabolites used by *C. telluris*. The middle cluster contains metabolites used by *C. insecticola* but not by one or more β-*Caballeronia* strains. The lower cluster contains the metabolites used by the β-*Caballeronia* strains but not by *C. insecticola*. The full dataset of 384 conditions is shown in Fig. S6. The compound order in the three clusters is available in the Supplementary Data Set 1. **D.** Resistance of *Caballeronia* strains to AMPs. The heat map shows the minimal concentrations of growth inhibition of the indicated wild-type and mutant strains by the listed AMPs. Minimal

concentrations are indicated in µg/mL. The color key of the heat maps in panels C and D are indicated at their right. The used strains in all panels are abbreviated according to the key at the top right.

SUPPLEMENTARY INFORMATION

Strict gut symbiont specificity in Coreoidea insects governed by interspecies competition within *Caballeronia*

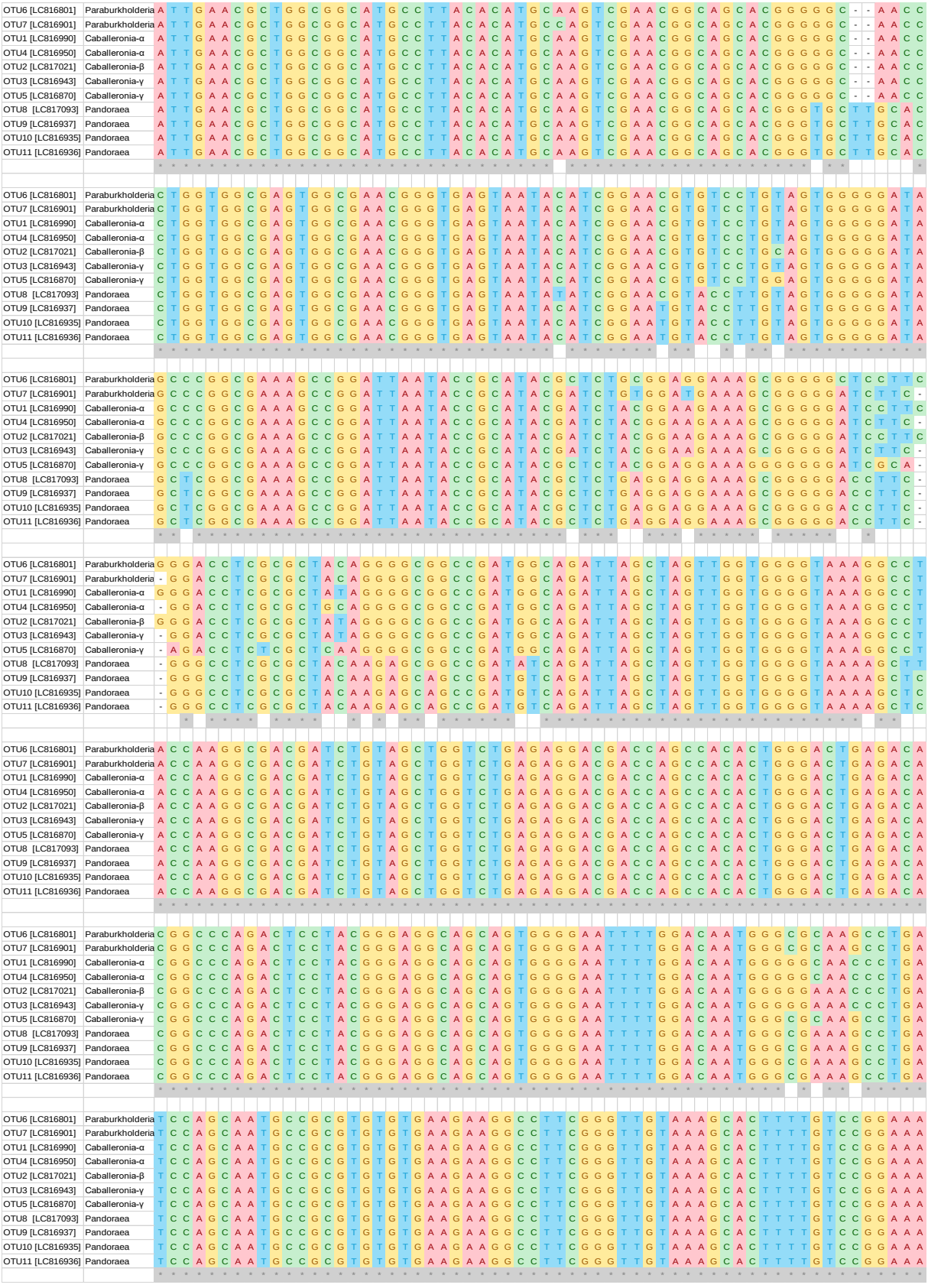
Gaelle Lextrait[1], Srotoswini Joardar[1,2], Raynald Cossard[1], Yoshitomo Kikuchi[3], Tsubasa Ohbayashi[1,4], Peter Mergaert[1]

[1]Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette, France
[2]Department of Biology, Ludwig-Maximilians-Universität München (LMU), 82152 Munich, Germany
[3]Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Hokkaido Center, 062-8517 Sapporo, Japan
[4]Institute for Agro-Environmental Sciences, National Agriculture and Food Research Organization (NARO), 305-8604 Tsukuba, Japan

Correspondence to PM peter.mergaert@i2bc.paris-saclay.fr or TO ohbayashi.tsubasa023@naro.go.jp

Content

SUPPLEMENTARY FIGURES




OTU6 [LC816801] Paraburkholderia
OTU7 [LC816901] Paraburkholderia
OTU1 [LC816990] Caballeronia-α
OTU4 [LC816950] Caballeronia-α
OTU2 [LC817021] Caballeronia-β
OTU3 [LC816943] Caballeronia-γ
OTU5 [LC816870] Caballeronia-γ
OTU8 [LC817093] Pandoraea
OTU9 [LC816937] Pandoraea
OTU10 [LC816935] Pandoraea
OTU11 [LC816936] Pandoraea

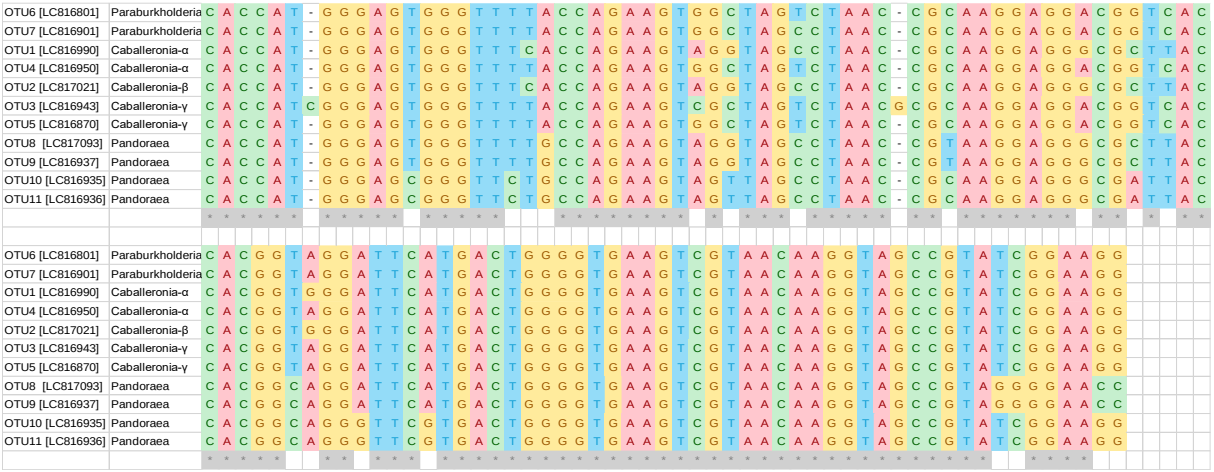


Figure S1. Alignment of OTUs identified in the midgut crypts of *Riptortus pedestris* and *Coreus marginatus*. The sequences of 16S rRNA gene PCR products were aligned with the MAFFT program. An asterisk indicates a correspondence of sequences for all OTUs.

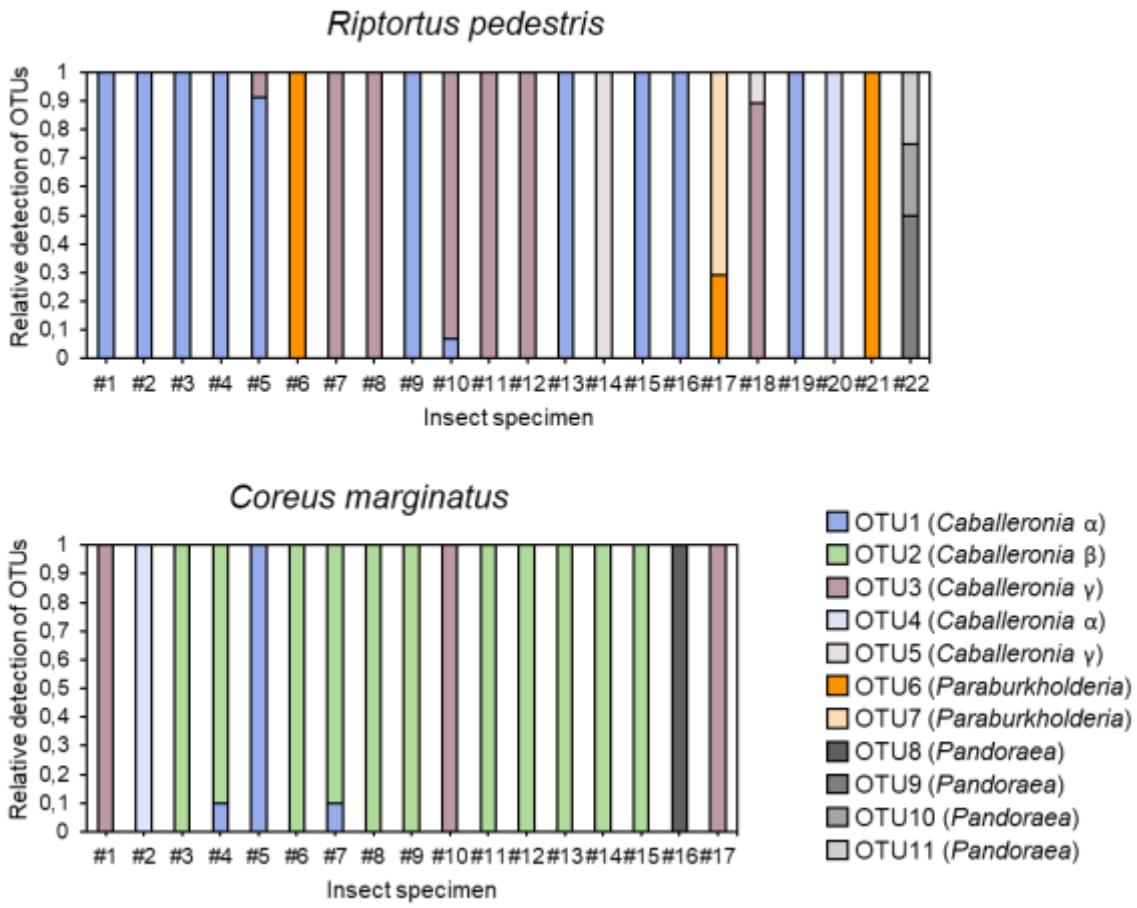


Figure S2. Distribution of OTUs of the 16S rRNA gene in individual *Riptortus pedestris* and *Coreus marginatus* M4 midgut samples. Bar plots representing the relative proportion of OTUs detected in each M4 midgut sample from *R. pedestris* and *C. marginatus* individuals reared on the same soil sample (see Table S2 for data).

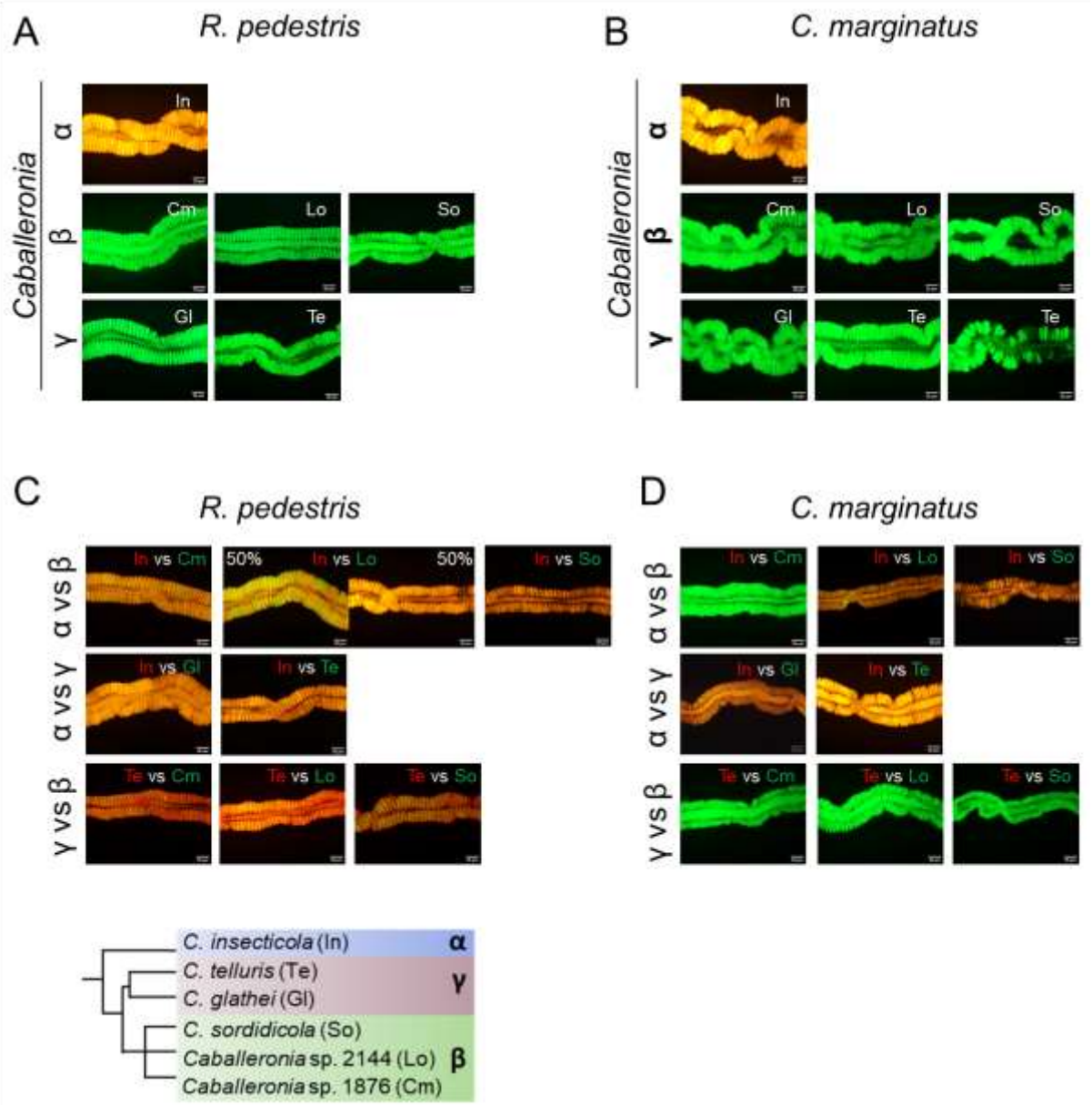


Figure S3. The colonization ability of *Caballeronia* species in the midgut of *Riptortus pedestris* and *Coreus marginatus*. A. Representative micrographs of midguts of *R. pedestris*, infected with a single *Caballeronia* strain and dissected at 5 dpi. Scale bars are 30 µm. **B.** Representative micrographs of midgut crypts of *C. marginatus*, infected with a single *Caballeronia* strain and dissected at 7 dpi. The crypt region of individuals infected with *C. telluris* (Te) is either fully colonized (as in left Te picture) or in some cases only partially colonized (as in right Te picture). **C.** Representative micrographs of midguts of *R. pedestris*, co-infected with two *Caballeronia* strains and dissected at 5 dpi. **D.** Representative micrographs of midgut crypts of *C. marginatus*, co-infected with two *Caballeronia* strains and dissected at 7 dpi. Scale bars are 40 µm.

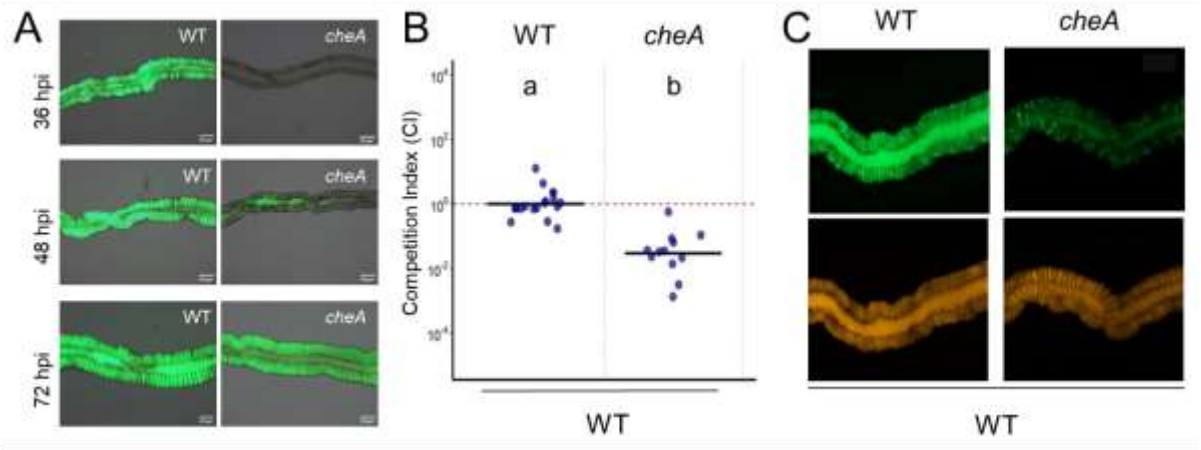


Figure S4. The colonization ability of the *Caballeronia cheA* mutant in the *Riptortus pedestris* midgut. A. Microscopy image of the M4 colonization at 36, 48 and 72 hpi of R. *pedestris* with the wild-type strain (WT) and the *cheA* mutant (n = 10), showing the delayed colonization dynamics by the *cheA* mutant compared to the wild-type strain. **B**. Competition index of *Caballeronia* mutant *cheA* with the wild-type strain in *R. pedestris* (n > 10). Letters indicate significant differences (P < 0.05) using the Kruskal-Wallis test with the post-hoc pairwise Wilcoxon rank with the Benjamini-Hochberg *P*-value adjustment method. **C.** Microscopy images of pairwise competitions between GFP- and mScarlet-I-labelled *C. insecticola* strains.

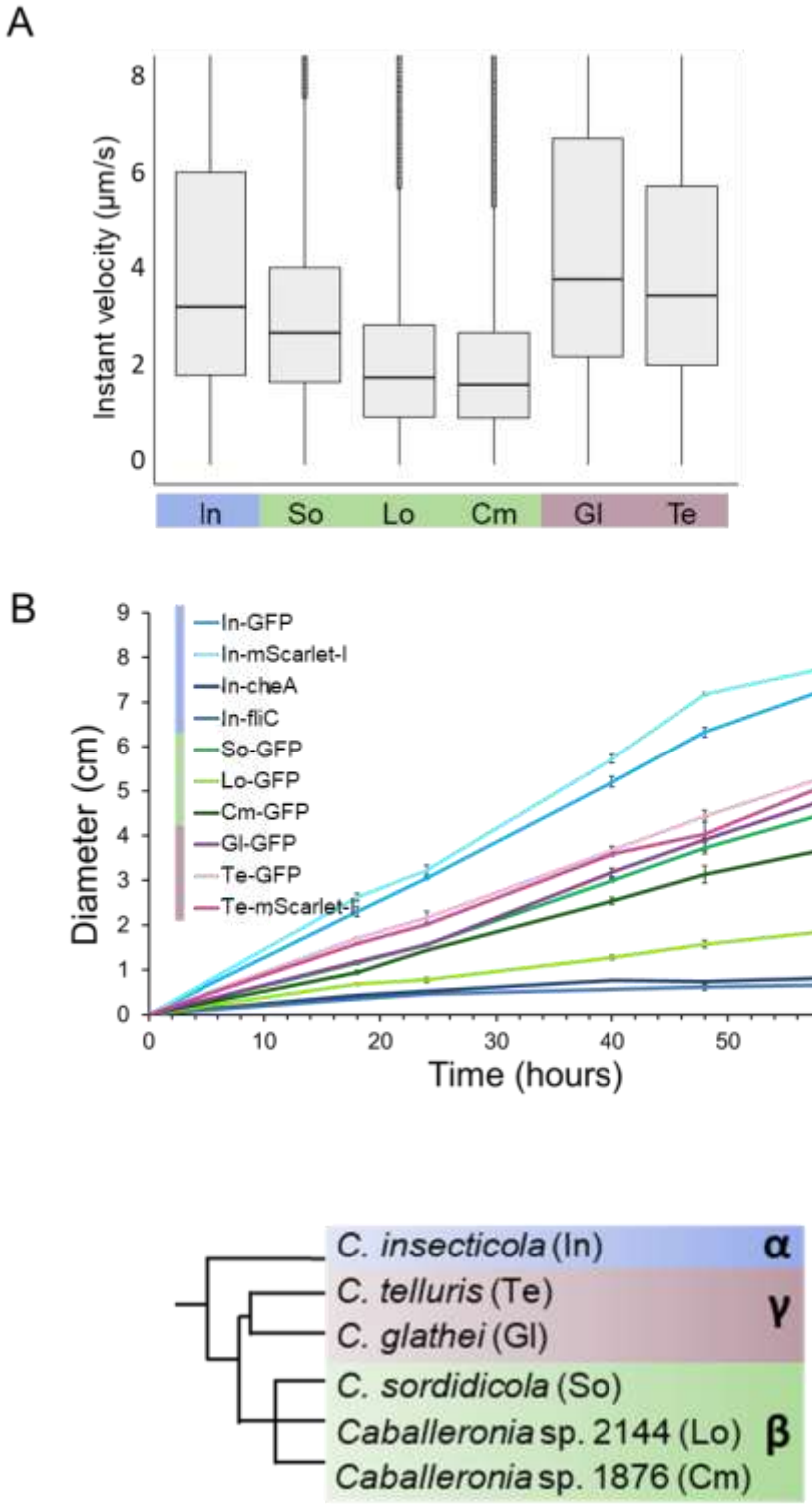


Figure S5. Motility and chemotaxis of *Caballeronia* species. A. Box plots representing the velocity of strains in liquid YG medium. Box limits indicate the range of the central 50% of the data, the black line within the box indicates the median of the data set, and the thick black lines running above the whiskers represent the outlier values. **B.** The combined motility and chemotaxis of strains measured on 0.3% YG agar swimming plates. The graphs shows the diameter of growth on the plates in function of time. The used strains are abbreviated according to the key at the bottom.

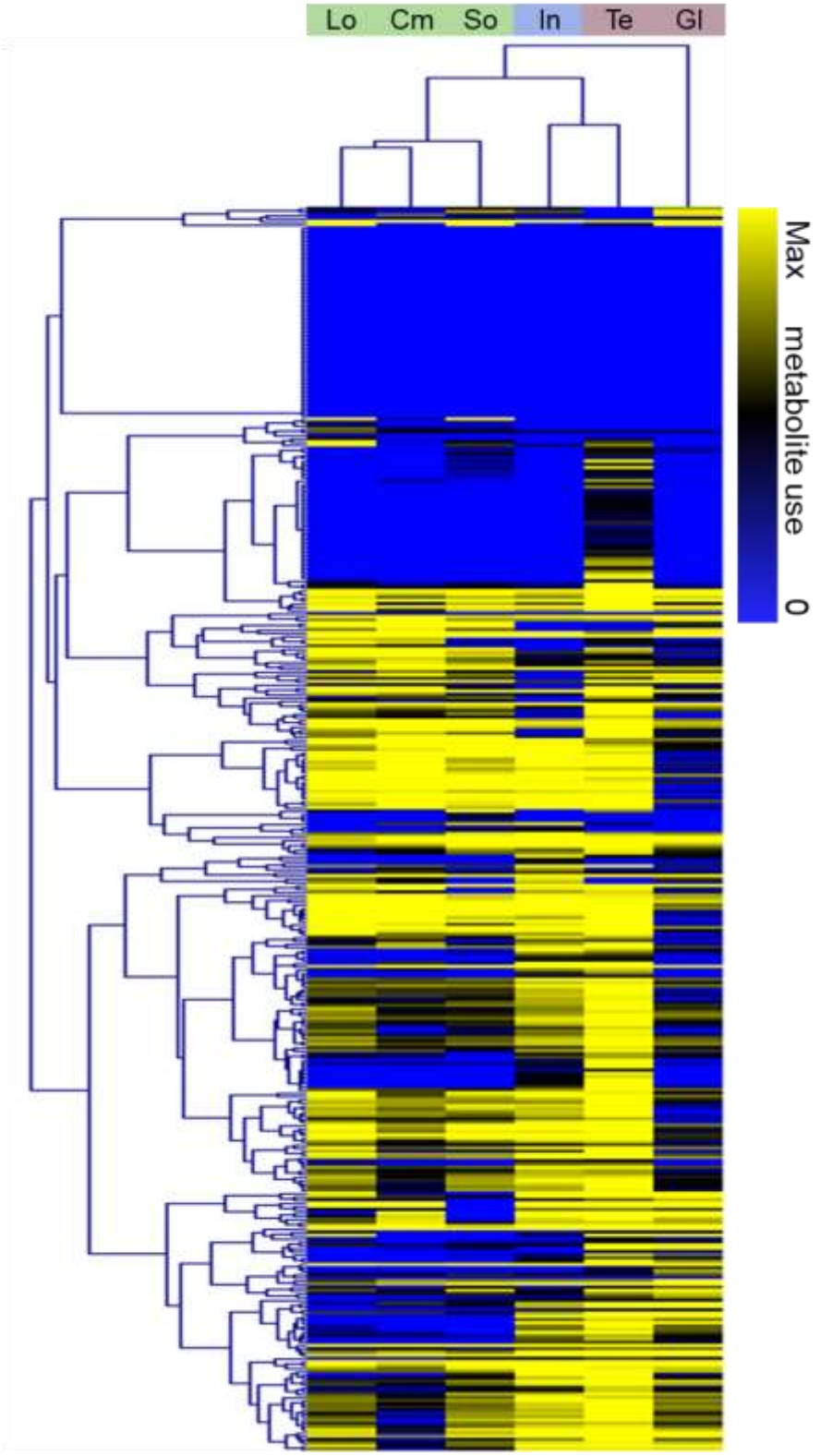


Figure S6. Biolog phenotyping of *Caballeronia* strains. The heat map shows metabolite usage by *Caballeronia* strains for the combined 486 tested conditions by Biolog plates PM1, PM2A, PM3B and PM4A. Data represents areas under the absorbance curves in function of time, normalized to 1 relative to the maximal value in each microplate. Metabolites and strains were ordered with Pearson correlation using hierarchical cluster analysis. The compound order in is available in the Supplementary Data Set 1.

SUPPLEMENTARY TABLES

Strain	Description	Reference
Caballeronia insecticola RPE75 (In)	rifampicin-resistant derivative of wild type strain RPE64, isolated from *Riptortus pedestris* , japan	8
Caballeronia insecticola RPE225 (In-GFP)	Tn*7*-GFP derivative of RPE75	2
Caballeronia insecticola Scarlet (In-mScarlet-I)	Tn*7*-mScarlet-I derivative of RPE75	1
Cm1876 (Cm)	strain isolate of *Coreus marginatus* , France	5
Cm1876 (Cm-GFP)	Tn*7*-GFP derivative of the strain Cm1876	This work
Caballeronia sordidicola (So)	Type strain DSM17212^{T}; isolated from a culture of the fungus Phanerochaete sordida, South Korea	Type strain
Caballeronia sordidicola (So-GFP)	Tn*7*-GFP derivative of *C. sordidicola*	This work
Lo2144 (Lo)	Strain isolate of *Leptoglossus occidentalis* , France	6
Lo2144 (Lo-GFP)	Tn*7*-GFP derivative of strain Lo2144	This work
Caballeronia glathei (Gl)	Type strain JCM10563^{T}, isolated from soil, Germany	Type strain
Caballeronia glathei (Gl-GFP)	Tn*7*-GFP derivative of *C. glathei*	This work
Caballeronia telluris (Te)	Type strain LMG22936^{T}, isolated from soil, Netherlands	Type strain
Caballeronia telluris (Te-GFP)	Tn*7*-GFP derivative of *C. telluris*	This work
Caballeronia telluris (Te-mScarlet-I)	Tn*7*-mScarlet-I derivative of *C. telluris*	This work
Caballeronia insecticola RPE75.Δ*wzm*	*wzm* (BRPE64_RS10560) deletion mutant	3
Caballeronia insecticola RPE75.Δ*tpr*	*tpr (BRPE64_RS10075)* deletion mutant	3
Caballeronia insecticola RPE75.*cheA::* Tn*5*	*cheA(BRPE64_RS19435)* Tn*5* transposon insertion mutant	4
Caballeronia insecticola RPE75.Δ*fbp*	*fbp (BRPE64_RS03750)* deletion mutant	1
Caballeronia insecticola RPE75.Δ*pps*	*pps (BRPE64_RS05810)* deletion mutant	1
WM3064.pURR25	*E. coli* carring Tn*7*-GFP on plasmid	9
WM3064.pUX-BF13	*E. coli* helper carring transposase	9
S17-1λpir.pMRE-Tn*7*-135	*E. coli* carring Tn*7*-mScarlett-I on plasmid	7

Table S1. Bacterial strains used in this study.

1. Jouan R, Lextrait G, Lachat J *et al.* Transposon sequencing reveals the essential gene set and genes enabling gut symbiosis in the insect symbiont *Caballeronia insecticola*. *ISME Commun* 2024;**4**:ycad001. https://doi.org/10.1093/ismeco/ycad001
2. Kikuchi Y, Fukatsu T. Live imaging of symbiosis: spatiotemporal infection dynamics of a GFP-labelled *Burkholderia* symbiont in the bean bug *Riptortus pedestris*. *Mol Ecol* 2014;**23**:1445-56. https://doi.org/10.1111/mec.12479
3. Lachat J, Lextrait G, Jouan R *et al.* Hundreds of antimicrobial peptides create a selective barrier for insect gut symbionts. *Proc Natl Acad Sci U S A* 2024;**121**:e2401802121. https://doi.org/10.1073/pnas.2401802121
4. Ohbayashi T, Takeshita K, Kitagawa W *et al.* Insect's intestinal organ for symbiont sorting. *Proc Natl Acad Sci U S A* 2015;**112**:E5179-88. https://doi.org/10.1073/pnas.1511454112
5. Ohbayashi T, Itoh H, Lachat J *et al. Burkholderia* gut symbionts associated with European and Japanese populations of the dock bug *Coreus marginatus* (Coreoidea: Coreidae). *Microbes Environ* 2019;**34**:219-22. https://doi.org/10.1264/jsme2.ME19011
6. Ohbayashi T, Cossard R, Lextrait G *et al.* Intercontinental diversity of *Caballeronia* gut symbionts in the conifer pest bug *Leptoglossus occidentalis*. *Microbes Environ* 2022;**37**:ME22042. https://doi.org/10.1264/jsme2.ME22042
7. Schlechter RO, Jun H, Bernach M *et al.* Chromatic bacteria - a broad host-range plasmid and chromosomal insertion toolbox for fluorescent protein expression in bacteria. *Front Microbiol* 2018;**9**:3052. https://doi.org/10.3389/fmicb.2018.03052
8. Takeshita K, Tamaki H, Ohbayashi T *et al. Burkholderia insecticola* sp. nov., a gut symbiotic bacterium of the bean bug *Riptortus pedestris*. *Int J Syst Evol Microbiol* 2018;**68**:2370-4. https://doi.org/10.1099/ijsem.0.002848
9. Teal TK, Lies DP, Wold BJ, Newman DK. Spatiometabolic stratification of *Shewanella oneidensis* biofilms. *Appl Environ Microbiol* 2006;**72**:7324-30. https://doi.org/10.1128/AEM.01163-06

Sample	Individual	Caballeronia α	Caballeronia β	Caballeronia γ	Caballeronia α	Caballeronia γ	Paraburkholderia	Paraburkholderia	Pandoraea	Pandoraea	Pandoraea	Pandoraea	
name	**No.**	**OTU1**	**OTU2**	**OTU3**	**OTU4**	**OTU5**	**OTU6**	**OTU7**	**OTU8**	**OTU9**	**OTU10**	**OTU11**	**Total**
R. pedestris													
R7	#1	1	0	0	0	0	0	0	0	0	0	0	14
R11	#2	1	0	0	0	0	0	0	0	0	0	0	13
R15	#3	1	0	0	0	0	0	0	0	0	0	0	13
R16	#4	1	0	0	0	0	0	0	0	0	0	0	11
R17	#5	0,91	0	0,09	0	0	0	0	0	0	0	0	11
R18	#6	0	0	0	0	0	1	0	0	0	0	0	4
R20	#7	0	0	1	0	0	0	0	0	0	0	0	4
R25	#8	0	0	1	0	0	0	0	0	0	0	0	6
R26	#9	1	0	0	0	0	0	0	0	0	0	0	3
R27	#10	0,07	0	0,93	0	0	0	0	0	0	0	0	15
R28	#11	0	0	1	0	0	0	0	0	0	0	0	6
R29	#12	0	0	1	0	0	0	0	0	0	0	0	14
R31	#13	1	0	0	0	0	0	0	0	0	0	0	10
R32	#14	0	0	0	0	1	0	0	0	0	0	0	11
R33	#15	1	0	0	0	0	0	0	0	0	0	0	13
R37	#16	1	0	0	0	0	0	0	0	0	0	0	13
R38	#17	0	0	0	0	0	0,29	0,71	0	0	0	0	14
R39	#18	0	0	0,89	0	0,11	0	0	0	0	0	0	9
R40	#19	1	0	0	0	0	0	0	0	0	0	0	4
R41	#20	0	0	0	1	0	0	0	0	0	0	0	5
R42	#21	0	0	0	0	0	1	0	0	0	0	0	3
R43	#22	0	0	0	0	0	0	0	0	0,5	0,25	0,25	4
	Rel. Total	0,53	0,00	0,27	0,03	0,06	0,06	0,05	0,00	0,01	0,01	0,01	
	Abs. Total	105	0	53	5	12	11	10	0	2	1	1	200
C. marginatus													
C3	#1	0	0	1	0	0	0	0	0	0	0	0	11
C4	#2	0	0	0	1	0	0	0	0	0	0	0	12
C6	#3	0	1	0	0	0	0	0	0	0	0	0	10
C7	#4	0,1	0,9	0	0	0	0	0	0	0	0	0	10
C8	#5	1	0	0	0	0	0	0	0	0	0	0	13
C13	#6	0	1	0	0	0	0	0	0	0	0	0	9
C30	#7	0,1	0,9	0	0	0	0	0	0	0	0	0	10
C31	#8	0	1	0	0	0	0	0	0	0	0	0	12
C32	#9	0	1	0	0	0	0	0	0	0	0	0	12
C33	#10	0	0	1	0	0	0	0	0	0	0	0	16
C34	#11	0	1	0	0	0	0	0	0	0	0	0	5
C35	#12	0	1	0	0	0	0	0	0	0	0	0	18
C36	#13	0	1	0	0	0	0	0	0	0	0	0	6
C37	#14	0	1	0	0	0	0	0	0	0	0	0	5
C38	#15	0	1	0	0	0	0	0	0	0	0	0	5
C39	#16	0	0	0	0	0	0	0	1	0	0	0	4
C40	#17	0	0	1	0	0	0	0	0	0	0	0	5
	Rel. Total	0,09	0,61	0,20	0,07	0,00	0,00	0,00	0,02	0,00	0,00	0,00	
	Abs. Total	15	100	32	12	0	0	0	4	0	0	0	163

Table S2. Distribution of OTUs of the 16S rRNA gene in *Riptortus pedestris* and *Coreus marginatus* M4 midgut samples. The numbers in the OTU columns indicate the proportion of clones belonging to each OTU per insect sample. The figures in the Rel. Total row indicate the proportion of clones belonging to each OTU per insect, while the figures in the Abs. Total row indicate the total number of sequences for each OTU per insect. The figures in the "Total" column indicate the number of sequences obtained per insect and in total for all the insects per species.

SUPPLEMENTARY MATERIALS AND METHODS

Insects and bacterial strains

The *R. pedestris* TKS1 inbred line used in this study is derived from a single male and female pair that were collected in 2007, from a soybean field in Tsukuba, Ibaraki, Japan. *C. marginatus* specimens were collected in 2017 from Rumex plants growing on the CNRS campus, Gif-sur-Yvette, France (48°42'17.0"N 2°07'42.2"E) and in 2022 in the Bures – Gif-sur-Yvette water reservoir basin of Bures-sur-Yvette, France (48°41'52.4"N 2°09'09.3"E).

R. pedestris was reared in Petri dishes (90 mm in diameter and 20 mm height at 25°C under a long-day regimen (16 h light and 8 h dark) and fed with soybean seeds and cotton pads with distilled water containing 0.05% ascorbic acid. *C. marginatus* was reared in the same way but fed with a mixture of roasted peanut and pistachio seeds.

Bacterial strains used in this study are listed in Table S1. *Caballeronia* species were cultured in Yeast Glucose (YG) medium (0.5% yeast extract, 0.4% glucose, 0.1% NaCl) or MM medium [1], supplemented with 30 µg/mL of kanamycin (Km) or 30 µg/mL of rifampicin (Rif), if necessary, at 28°C on a gyratory shaker at 180 rpm. *Escherichia coli* strains, used for the generation of fluorescent protein tagged bacteria, were grown at 37°C in Luria-Bertani (LB) medium (5 g/L yeast extract, 10 g/L tryptone, 5 g/L NaCl), supplemented with 300 µg /mL of diaminopimelic acid (DAP), 50 µg/mL of Km, 100 µg/mL of ampicillin (Amp), or 100 µg/mL of chloramphenicol (Cm).

Generation of fluorescent protein tagged *Caballeronia* species

GFP-labelled *Caballeronia* species were generated by inserting a mini-Tn*7* transposon containing p*lac*::*GFP* into the *attTn7* site downstream of the *glmS* gene in the chromosome of *Caballeronia* species via triparental conjugation as described previously [2]. Briefly, 200 µL overnight cultures of *Caballeronia* species were diluted into 10 mL fresh YG and incubated at 28°C at 180 rpm to an exponential phase (OD_{600} = 0.5-1). These cultures were washed twice in fresh YG medium without antibiotics by centrifugation at 4,000 rpm for 10 minutes. Bacterial pellets were resuspended in fresh YG medium to obtain a final OD_{600} from 5 to 10. The transposon donor strain WM3064.pURR25, the helper WM3064.pUX-BF13 were prepared in a similar way and the *Caballeronia* species, donor and helper were mixed at a 1:1:1 ratio.

The conjugation mix was spotted on YG agar plates containing DAP. After overnight incubation at 28°C, bacterial spots were resuspended in 1 mL of YG medium, diluted in ten-fold series and plated out on YG with Km 30 µg/mL. DAP was not added to counter-select the *E. coli* strains. mScarlet-I-labelled *Caballeronia* strains were generated by inserting a Tn*7* transposon encoding mScarlet-I. The plasmid pMRE-Tn7-135 [3], carrying the Tn*7* transposon with the mScarlet-I gene and the chloramphenicol resistance marker and the Tn*7* transposase genes under the control of an arabinose-inducible promoter, was introduced into the target *Caballeronia* strains by electroporation. The electroporated cells were inoculated on YG medium with 0.1% L-arabinose, which induces the expression of the plasmid-encoded transposase and incubated for 1 h at 30°C at 180 rpm. Then, the cultures were plated on YG medium supplemented with Cm and incubated at 28°C for 3 days. For GFP- and mScarlet-I-labelled *Caballeronia* strains, the brightest colonies were selected by blue LED light illumination, and purified on YG agar supplemented with antibiotics.

Soil inoculation test in *Riptortus pedestris* and *Coreus marginatus*

Samples of soil from the CNRS campus, Gif-sur-Yvette, France (48°42'17.0"N 2°07'42.2"E) were collected in 2019 and sieved with a 2 mm sieve to remove debris. In preparation of the inoculation with the soil sample, the cotton pads with distilled water containing 0.05% ascorbic acid were removed from the petri dishes with 2nd instar nymphs of *R. pedestris* or *C. marginatus* and the nymphs were maintained overnight without water to make them thirsty, which facilitates the subsequent ingestion of bacteria. The 2nd instar is the most susceptible developmental stage for acquisition of the crypt symbionts [4]. The following day, cotton pads wetted with 1 g of soil suspension in 10 mL of water were reintroduced in the petri dishes. Throughout the experiment, the two species were housed in separate petri dishes with 10 insects per dish. *R. pedestris* and *C. marginatus* nymphs in their third instar, at five days and seven days after inoculation, respectively, were digested and their M4 crypt region was collected in 100 µL PBS (137 mM NaCl, 8.1 mM Na_2HPO_4, 2.7 mM KCl, and 1.5 mM KH_2PO_4, pH 7.5).

Identification of gut symbionts by 16S rRNA sequencing and molecular phylogenetic analysis

DNA was extracted from the dissected midgut crypts of individual third instar nymphs using a phenol-chloroform extraction method and a clone library analysis of the 16S rRNA gene was performed as previously reported [5]. The obtained sequences were assembled by the ATSQ software (version 5.2; Software Development, Japan), followed by manual corrections. Assembled 1500bp long sequences were associated to the most homologous bacterial species/strains by BLAST comparison. Sequences of over 99% identity were assigned to the same operational taxonomical unit (OTU).

Multiple alignment of the 16S rRNA gene was performed with MAFFT on the EMBL-EBI server [6]. A molecular phylogenetic tree was generated by the maximum likelihood (ML) method with removal of gap-including and ambiguous sites, and with bootstrap analysis (1,000 replicates) in the MEGA software (version 10.1.8) [7]. We selected the Tamura-Nei model of nucleotide substitutions with gamma distributed and invariant sites (G+I).

Single strain infection test in *Riptortus pedestris* and *Coreus marginatus*

Two hundred µL of overnight cultures of GFP- or mScarlet-I-labelled *Caballeronia* strains were diluted into 10 mL YG medium and incubated at 28°C at 180 rpm until exponential phase (OD_{600} = 0.5-1). These cultures were diluted to a final concentration of 10^7 cfu/mL in distilled water with 0.05% ascorbic acid. The second instar nymphs of *R. pedestris* and *C. marginatus* were prepared as above, by removing drinking water, and infected by oral administration via new cotton pads wetted with the bacterial suspension. Aposymbiotic insects were obtained providing only 0.05% ascorbic acid water without bacteria. These insects were reared until reaching the third instar (5 dpi for *R. pedestris* and 7 dpi for *C. marginatus*) before dissection in PBS, under a binocular microscope and analysis of the colonization state of the midgut crypts by fluorescence microscopy observation using a Nikon, Eclipse 80i microscope (Nikon Corporation, Japan).

For fitness measurements in *R. pedestris,* insects were infected as described above with different *Caballeronia* species. Until the adult stage, dead insects were counted every day, and the day of moulting from the last larval stage into the adult stage was reported to determine development time. For morphological characterization, one day old adults were immersed in acetone. Acetone was refreshed two times and the insects were dried in free air. Each

individual was weighted and thorax width and insect total length were reported using measuring pliers. In order to calculate the effect of the symbionts on reproduction, one female and two males were placed together in a container and egg production was counted as described before [8]. Mated females begin oviposition 6 to 7 days after molting to adulthood. Upon commencement of egg-laying, the total egg count was recorded after 1 week. Statistical analysis, using a Kruskal-Wallis test, Dunn post hoc test and Benjamini-Hochberg correction, was performed with R studio (version 4.5.0) [9].

Two strain competition assays in *Riptortus pedestris* and *Coreus marginatus*

In each pairwise competition assay, we combined GFP- and mScarlet-I-labelled *Caballeronia* strains (Table S1). Exponential phase cultures of each strain were diluted to a final concentration of 10^7 cfu/mL in distilled water with 0.05% ascorbic acid and combined in equal amounts. The composition of the mixture of the two strains was verified by flow cytometry analysis as described below and corrections in the mixture were introduced until reaching a 1:1 mixture. Inoculation of *R. pedestris* and *C. marginatus* insects with the mixed inoculum was performed as described above. These insects were maintained until the third instar nymphs.

Symbiotic organs were dissected under a binocular microscope in sterilized PBS. In part of the samples, a qualitative estimation of the colonization state of the midgut crypts by the GFP- and mScarlet-I-labelled strains was determined by fluorescence microscopy using a Nikon, Eclipse 80i microscope (Nikon Corporation, Japan). In the remaining samples, each M4 was individually grounded by a sterile pestle in a 1.5 mL tube containing 100 µL of sterilized PBS. The Pestle was washed with 400 µL of PBS and these samples were stored at 4°C for a maximum of 5 days until flow cytometry analysis. The relative number of GFP- and mScarlet-I-tagged bacterial strains in each sample was analysed by flow cytometry using a CytoFlex S machine (Beckman Coulter, US). GFP signal was excited by a 488 nm laser and a 525/40 nm band pass filter was used for detection; mScarlet-I signal was excited by a 561 nm laser and a 610/20 nm band pass filter was used for detection. In data acquisition of flow cytometry analysis using the Cytexpert software (version 2.4), a first gating for bacterial cells was made on the forward-scatter (FSC)-side scatter (SSC) dot plot to focus on bacteria, and then doublets were discarded using the SSC_Area and SSC_Height dot plot. Data acquisition for a total of

50,000-100,000 bacteria was recorded for each condition. Thresholds for considering positive GFP and mScarlet-I bacteria were determined using cultures of non-fluorescent bacteria as a negative control and cultures of the corresponding tagged strains as positive controls. A competition index (CI) was calculated as the ratio of a tested GFP-strain to a mScarlet-I-strain, normalized by the ratio of the inoculum, which was always close to 1. Statistical analysis, using a Kruskal-Wallis test, Dunn post hoc test and Benjamini-Hochberg correction, was performed with R studio (version 4.5.0) [9].

Motility and chemotaxis measurements

To measure the velocity of the different strains in liquid YG medium, time-lapse videos of the bacteria's movement were made using Leica DMI6000 B Inverted Research Microscope (Leica Camera AG, Germany). Cultures of the GFP-labelled strains in exponential phase (OD_{600}<1), grown in YG medium, were diluted to OD_{600} = 0.05, and loaded on a glass slide inside a Frame-Seal (Bio-Rad Laboratories, Inc, US) barricade, then covered with a piece of Polydimethylsiloxane (PDMS) to allow the entry of oxygen during microscopy. For the imaging, the YFP fluorescence filter and 10X objective was used. Videos of 10 seconds at a frequency of 10 frames per second were recorded. The tracking and analysis of the motility was done using the TrackMate plugin [10] of the Fiji software (version 2.9.0) [11], which identifies the movement of a fluorescent particle (bacteria) from frame to frame, forms tracks of individual particles across all the frames and extract the coordinates of the particles along the tracks. From the coordinates and time difference between frames, instantaneous velocities are calculated. For each strain, from 500-1,500 tracks about 50,000 to 150,000 instant velocities were calculated, which were represented in box plots using R Studio (R version 4.5.0) [9].

To perform the chemotaxis measurements, swimming plates (YG medium with 0.3% agar) were inoculated by inserting into the soft agar layer 2 µL of a suspension of the tested strains, grown in exponential phase (OD_{600} < 1) and concentrated to OD_{600} = 2. The plates were incubated at 28°C in a damp environment to make sure the plates did not dry out. Images of the plates were taken at regular time points for around 60 hours, and the diameter of growth in the plate was measured on the pictures using Fiji.

Two strain competition assays in liquid culture and on agar plates

For the pairwise competition experiments, a GFP-tagged and an mScarlet-tagged strain were mixed together in a 1:1 ratio, verified by flow cytometry, and adjusted as above. The strain mixtures were then inoculated either in liquid cultures with MM medium and either glucose or succinate as carbon source (inoculum at OD_{600} = 0.05), on standard (1.5% agar) YG plates (spot inoculum of 50 µL at OD_{600} = 0.05 or OD_{600} = 2; the liquid of the spots was air dried to create inter-strain cell-cell contact on the agar plates during colony growth), or on swimming plates as above (YG medium with 0.3% agar; inoculum at OD_{600} = 2). Co-cultures were grown for 24 hours at 28°C. The cells were harvested in 1X PBS buffer from across the growth for liquid cultures and solid agar plates or from the edges only of the growth zone for the swimming plates. The relative number of GFP- and mScarlet-I-labelled bacteria in the harvested cells of each sample was measured by flow cytometry as above.

Biolog phenotype microarrays

The metabolic profile of strains was analysed using four 96-well Phenotype Microarray (Biolog Inc., US] microplates, which consisted of two microplates of carbon substrates (PM1 and PM2A), one microplate of various nitrogen substrates (PM3B) and one microplate of different phosphorus and sulfur substrates (PM4A). The principle of the assay is based on the measurement of purple color formation from the irreversible redox reaction of the Biolog tetrazolium dye, in response to the production of reduced nicotinamide adenine dinucleotide (NADH) by bacterial respiration, which is active when the bacteria are able to uptake and metabolize the test compound present in the well. The assay was performed according to the manufacturer's protocol for Gram-negative bacteria. For the analysis, single colonies of the strains were harvested from a freshly prepared YG agar plate and suspended in 1X IF-0 solution (Biolog Inc., US). That suspension was further diluted using IF-0+Dye Mix A solution (Biolog Inc., US) to attain $OD_{600nm} \approx 0.07$. The prepared bacterial suspension was then added to each of the wells of the PM1 and PM2A plates. For the plates lacking carbon sources (PM3B and PM4A), the pre-prepared 100X carbon stock solution (2 M glucose, 200 µM ferric citrate) was added additionally in the IF-0+Dye Mix A solution for the preparation of the sample. The plates were incubated at 28°C in a humidified box. Kinetic absorbance data at OD_{490nm} and OD_{750nm} for each well were obtained at regular time intervals during 62h using a Spark

Multimode Microplate Reader (Tecan, Switzerland). A corrected signal was calculated by subtracting the OD_{750nm} measurement from the OD_{490nm} measurement and the area under the kinetic curve was calculated in Excel (Microsoft Corporation, US). Values were normalized per plate and the full data set for the 4 arrays and all strains was analyzed and visualized in a heat map by MEV software (version 4.8.0) [12] using the hierarchical clustering function with Pearson Correlation distance metric and average linkage settings.

Sensitivity assays to antimicrobial peptides

Precultures of tested strains were grown in MM medium with glucose as carbon source. Overnight grown cultures were diluted to an OD_{600nm} = 0.3 in fresh medium and grown until they reached OD_{600nm} ≈ 1. The cells were pelleted by centrifugation, resuspended in fresh medium and diluted to OD_{600nm} = 0.05. These cell suspensions were dispatched in a 96-well plate at 146.25 µL per well and 3.75 µL AMPs [13], dissolved in water at a 40-fold concentration, were added and mixed with the bacterial suspension by pipetting up and down. The AMPs were tested at the following concentrations (all in µg/mL): Polymyxin B (PMB) at 50, 25, 12.5, 6.25, 3.125, 1.5625, 0.78; riptocin, thanatin, CCR0179, CCR0191, CCR0776, CCR1659 at 300, 250, 125, 50; LL37 at 40, 25, 12.5, 5; NCR335 at 100, 50, 25, 12.5. The 96-well plates were incubated in a SPECTROstar Nano plate incubator (BMG Labtech GmbH, Germany). The growth of the cultures in the wells was monitored by measuring the OD_{600nm} and data points were collected every hour for 48 h. Plates were incubated at 28°C with double orbital shaking at 200 rpm. Data and growth curves were analyzed using Microsoft Excel (Microsoft Corporation, US). The minimal concentration was determined at which growth was diminished compared to the untreated control. The assays were performed in biological triplicates for all peptides.

Nucleotide sequence accession numbers

The nucleotide sequence data of the 16S rRNA gene obtained in the present study have been deposited in the DDBJ public database with the accession numbers LC816739-LC817101.

References

1. Jouan R, Lextrait G, Lachat J *et al.* Transposon sequencing reveals the essential gene set and genes enabling gut symbiosis in the insect symbiont *Caballeronia insecticola*. *ISME Commun* 2024;**4**:ycad001. https://doi.org/10.1093/ismeco/ycad001
2. Kikuchi Y, Fukatsu T. Live imaging of symbiosis: spatiotemporal infection dynamics of a GFP-labelled *Burkholderia* symbiont in the bean bug *Riptortus pedestris*. *Mol Ecol* 2014;**23**:1445-56. https://doi.org/10.1111/mec.12479
3. Schlechter RO, Jun H, Bernach M *et al.* Chromatic bacteria - A broad host-range plasmid and chromosomal insertion toolbox for fluorescent protein expression in bacteria. *Front Microbiol* 2018;**9**:3052. https://doi.org/10.3389/fmicb.2018.03052
4. Kikuchi Y, Hosokawa T, Fukatsu T. Specific developmental window for establishment of an insect-microbe gut symbiosis. *Appl Environ Microbiol* 2011;**77**:4075-81. https://doi.org/10.1128/AEM.00358-11
5. Ohbayashi T, Itoh H, Lachat J *et al. Burkholderia* gut symbionts associated with European and Japanese populations of the dock bug *Coreus marginatus* (Coreoidea: Coreidae). *Microbes Environ* 2019;**34**:219-22. https://doi.org/10.1264/jsme2.ME19011
6. Li W, Cowley A, Uludag M *et al.* The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res* 2015;**43(W1)**:W580-4. https://doi.org/10.1093/nar/gkv279
7. Kumar S, Stecher G, Li M *et al.* MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* 2018;**35**:1547-9. https://doi.org/10.1093/molbev/msy096
8. Jang S, Ishigami K, Mergaert P, Kikuchi Y. Ingested soil bacteria breach gut epithelia and prime systemic immunity in an insect. *Proc Natl Acad Sci U S A* 2024;**121**:e2315540121. https://doi.org/10.1073/pnas.2315540121
9. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2021. https://www.R-project.org/
10. Ershov D, Phan MS, Pylvänäinen JW *et al.* TrackMate 7: integrating state-of-the-art segmentation algorithms into tracking pipelines. *Nat Methods* 2022;**19**:829-32. https://doi.org/10.1038/s41592-022-01507-1
11. Schindelin J, Arganda-Carreras I, Frise E *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* 2012;**9**:676-82. https://doi.org/10.1038/nmeth.2019
12. Saeed AI, Sharov V, White J *et al.* TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* 2003;**34**:374-8. https://doi.org/10.2144/03342mt01
13. Lachat J, Lextrait G, Jouan R *et al.* Hundreds of antimicrobial peptides create a selective barrier for insect gut symbionts. *Proc Natl Acad Sci U S A* 2024;**121**:e2401802121. https://doi.org/10.1073/pnas.2401802121