



HAL
open science

Hundreds of antimicrobial peptides create a selective barrier for insect gut symbionts

Joy Lachat, Gaëlle Lextrait, Romain Jouan, Amira Boukherissa, Aya Yokota, Seonghan Jang, Kota Ishigami, Ryo Futahashi, Raynald Cossard, Delphine Naquin, et al.

► **To cite this version:**

Joy Lachat, Gaëlle Lextrait, Romain Jouan, Amira Boukherissa, Aya Yokota, et al.. Hundreds of antimicrobial peptides create a selective barrier for insect gut symbionts. Proceedings of the National Academy of Sciences of the United States of America, 2024, 121 (25), pp.e2401802121. 10.1073/pnas.2401802121 . hal-04611367

HAL Id: hal-04611367

<https://hal.science/hal-04611367>

Submitted on 13 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1

2

3 Main Manuscript for

4 Hundreds of antimicrobial peptides create a selective barrier for insect 5 gut symbionts

6 Joy Lachat^{a,1}, Gaëlle Lextrait^{a,1}, Romain Jouan^{a,1}, Amira Boukherissa^a, Aya Yokota^a, Seonghan
7 Jang^{b,c,2}, Kota Ishigami^{b,c}, Ryo Futahashi^d, Raynald Cossard^a, Delphine Naquin^a, Vlad Costache^e,
8 Luis Augusto^a, Pierre Tissières^a, Emanuele G. Biondi^a, Benoît Alunni^{a,3}, Tatiana Timchenko^a,
9 Tsubasa Ohbayashi^{a,4}, Yoshitomo Kikuchi^{b,c,*}, Peter Mergaert^{a,*}

10 ^aUniversité Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC); Gif-sur-
11 Yvette, 91198, France.

12 ^bBioproduction Research Institute, National Institute of Advanced Industrial Science and
13 Technology (AIST), Hokkaido Center; Sapporo, 062-8517, Japan.

14 ^cUnit of Applied Biological Chemistry, Graduate School of Agriculture, Hokkaido University, 060-
15 8589 Sapporo, Japan.

16 ^dNational Institute of Advanced Industrial Science and Technology (AIST); Tsukuba, 305-8566,
17 Japan.

18 ^eMIMA2 Imaging Core Facility, Microscopie et Imagerie des Microorganismes, Animaux et
19 Aliments, INRAE; Jouy-en-Josas, 78352, France.

20 ¹These authors contributed equally to this work

21 ²Present address: Infectious Disease Research Center, Korea Research Institute of Bioscience
22 and Biotechnology, Daejeon 34141, South Korea.

23 ³Present address: Institut Jean-Pierre Bourgin, INRAE, AgroParisTech, Université Paris-Saclay;
24 78000 Versailles, France.

25 ⁴Present address: Institute for Agro-Environmental Sciences, National Agriculture and Food
26 Research Organization (NARO); Tsukuba, 305-8604, Japan.

27

28 *Corresponding authors: Peter Mergaert and Yoshitomo Kikuchi

29 **Email:** peter.mergaert@i2bc.paris-saclay.fr; y-kikuchi@aist.go.jp

30 **Author Contributions:** P.M. and Y.K. designed the study, planned the experiments and
31 supervised the project. T.O., R.F, A.B. and B.A. performed transcriptome analysis. J.L., R.J., D.N.
32 performed Tn-seq. J.L., R.J. and T.T. made mutants. P.M. and E.B. performed *in vitro* peptide
33 activity assays. S.J. and K.I. provided strains. V.C. performed SEM. L.A. and P.T. performed LPS
34 characterization. J.L., G.L., A.Y., R.C and T.O. performed insect experiments. J.L., G.L., R.J, A.B.,
35 T.O., Y.K, and P.M. analyzed data. P.M. wrote the manuscript with input from Y.K. All authors
36 provided critical feedback and helped to shape the manuscript.

37 **Competing Interest Statement:** Authors declare no competing interests.

38 **Classification:** Biological Sciences, Microbiology

39 **Keywords:** Gut microbiota biogeography, antimicrobial peptides, insect, resistance

40

41 **This PDF file includes:**

42 Main Text

43 Figures 1 to 5

44

45 **Abstract (max 250 words)**

46 The spatial organization of gut microbiota is crucial for the functioning of the gut ecosystem,
47 although the mechanisms that organize gut bacterial communities in microhabitats are only partially
48 understood. The gut of the insect *Riptortus pedestris* has a characteristic microbiota biogeography
49 with a multispecies community in the anterior midgut and a mono-specific bacterial population in
50 the posterior midgut. We show that the posterior midgut region produces massively hundreds of
51 specific antimicrobial peptides (AMPs), the Crypt-specific Cysteine-Rich peptides (CCRs) that have
52 membrane-damaging antimicrobial activity against diverse bacteria but posterior midgut symbionts
53 have elevated resistance. We determined by transposon-sequencing the genetic repertoire in the
54 symbiont *Caballeronia insecticola* to manage CCR stress, identifying different independent
55 pathways, including novel AMP-resistance pathways unrelated to known membrane homeostasis
56 functions as well as cell envelope functions. Mutants in the corresponding genes have reduced
57 capacity to colonize the posterior midgut, demonstrating that CCRs create a selective barrier and
58 resistance is crucial in gut symbionts. Moreover, once established in the gut, the bacteria
59 differentiate into a CCR-sensitive state, suggesting a second function of the CCR peptide arsenal
60 in protecting the gut epithelia or mediating metabolic exchanges between the host and the gut
61 symbionts. Our study highlights the evolution of an extreme diverse AMP family that likely
62 contributes to establish and control the gut microbiota.

63 **Significance Statement (max 120 words)**

64 The microbiota is usually not homogeneously dispersed in the animal gut but spatially structured
65 in microenvironments. The microbiota in the gut of the bean bug *Riptortus pedestris* displays a
66 sharp divide between the anterior and posterior midgut with a multispecies bacterial community in
67 the anterior region and a specific, mono-species *Caballeronia* symbiont population in the posterior
68 region. We found that this insect deploys in the midgut an arsenal of several hundreds of
69 antimicrobial peptides that creates a selective environment restricting the type of bacteria from the
70 anterior midgut microbiota that have a chance to establish in the posterior midgut. This finding
71 highlights a mechanism that could contribute in the construction of an exclusive niche for beneficial
72 gut symbionts.

73

74

75 **Main Text**

76

77 **Introduction**

78

79 The animal gut is colonized by bacterial communities, which provide essential functions to the host
80 (1, 2). The phylotype richness and total abundance of this gut microbiota varies strongly among the
81 animals from low to extraordinarily high (1). Moreover, in animals ranging from humans to insects,
82 gut microbiota do not constitute a homogeneous mixture but are spatially organized and form
83 discrete bacterial communities located in specific microhabitats along the longitudinal and
84 transverse axes of the gut (3-6). How this microbial biogeography is established is only partially
85 understood but is potentially correlated with physical barriers such as mucus, peritrophic membrane
86 and crypts, gradients of chemical parameters such as pH or oxygen levels, bacteriophages and
87 nutrient availability as well as host immune effectors. Among the latter are antimicrobial peptides
88 (AMPs), which are secreted in the gut lumen and come in contact with the microbiota (7-10). AMPs
89 contribute to establish an epithelia-microbiota equilibrium along the transverse axis of the gut by
90 regulating the species composition and location of the microbiota according to the resistance and
91 sensitivity patterns of its members (10). Thus, gut commensals are expected to be resilient to AMPs
92 (11, 12) but how they adapt and how important this adaptation is for colonization of their specific
93 niche within the gut remains largely unexplored. Moreover, it is not known if AMPs exert control on
94 the spatial organization of microbiota along the longitudinal gut axis.

95 The bean bug *Riptortus pedestris* has a particular midgut organization, associated with a simple
96 microbiota displaying a characteristic biogeography. The midgut has four morphologically and
97 functionally distinct compartments, labelled M1 to M4. The anterior M1 to M3 regions are involved
98 in food digestion and have a variable and transient microbiota, which is ingested through feeding.
99 The posterior M4 region on the other hand, composed of two rows of crypts branched on a central
100 tract, does not contribute to food digestion and is associated with a stable, (nearly) mono-specific
101 and high-abundant microbiota that is also acquired from the environment and sorted out from the
102 M3 microbiota (13, 14). The M4 bacteria are very specific, belonging to the *Caballeronia* genus and
103 are mostly present as a single colonizing species, established through a multifaceted selection
104 process. A sorting organ located at the entry of the M4 region winnows out the M3 microbiota
105 allowing only a subset of species to enter the M4 (15). After a successful initial passage of bacteria
106 through the sorting organ and infection of the M4 region, secondary infections are inhibited by
107 closure of the sorting organ (16). The infecting bacteria induce in the M4 crypts developmental
108 processes, including oxygenation by tracheal formation (17) and the maturation of the crypts by
109 intestinal stem cell stimulation and apoptosis inhibition that creates the luminal space in the crypts
110 for bacterial colonization (18). Finally, microbe-microbe competition within the crypts results in the
111 elimination of the least adapted strains and the dominance of a single strain in the M4 region (19).
112 Among the bean bug colonizers, *Caballeronia insecticola* has emerged as a model species (20).
113 We took advantage of this simplified gut-microbe interaction model to explore if together with the
114 already known mechanisms, AMP challenge contributes to create the gut biogeography in *R.*
115 *pedestris* and if AMP resistance in *C. insecticola* is crucial for M4 crypt colonization.

116

117

118 Results

119

120 The *Riptortus pedestris* midgut expresses hundreds of AMP-like genes

121 A preliminary transcriptome analysis of the M4 midgut region has identified a novel class of
122 secretory peptides, which we call the Crypt-specific Cysteine-Rich peptides or CCRs (21, 22). In
123 order to define the expression pattern of *CCR* genes, the transcriptome was determined by RNA-
124 seq in midgut regions of insects that were reared for different times in the presence or absence of
125 the *C. insecticola* gut symbiont (Fig. 1A). The pooled sequencing reads were assembled in a set of
126 unique transcripts and encoded proteins. Hidden Markov Models based on the previously identified
127 *CCR* sequences were used to identify in the newly generated transcriptome the complete set of
128 *CCR* sequences. This analysis revealed 310 *CCR* transcripts (SI Appendix, Data S1). Together,
129 these transcripts encode 217 distinct *CCR* peptides derived from 126 putative genes. Closely
130 related transcripts and peptide variants could arise from recently duplicated genes, alternative
131 splicing of gene transcripts or allelic variation present in the rearing population, although the latter
132 is expected to be low since it is an inbred line derived from a single pair. The *CCR* peptides do not
133 show high similarity apart from a pattern of conserved cysteine residues (Fig. 1B). Despite their
134 sequence divergence, AlphaFold2 predicted similar folds for tested *CCR* peptides, consisting of
135 three pairs of β -sheets that are probably connected by cystine bridges (Fig. 1C). Differential
136 expression analysis revealed that the majority of the *CCR* genes are most strongly expressed in
137 the midguts of symbiotic insects (Fig. 1D and SI Appendix, Data S1). Subsets of genes were
138 specific for the M3, M4B and the majority for the M4 region carrying the *C. insecticola* bacteria,
139 suggesting that the encoded peptides target the symbionts. Moreover, the *CCRs* are among the
140 most strongly expressed transcripts in the overall transcriptome (Fig. 1E), suggesting a primordial
141 role of the peptides in the midgut. The *CCR* genes did not exhibit similarity to known sequences of
142 other organisms. However the taxonomically restricted nature of the genes as well as the structure
143 of the *CCRs*, being small, secreted and characterized by conserved cysteine residues, are features
144 shared with AMP gene families (10) and AMP prediction tools confirmed this presumption (Fig. 1C
145 and SI Appendix, Data S1). Whole mount *in situ* hybridization with the infected-M4-specific gene
146 *CCR0043* showed that the gene is expressed uniformly by the epithelial cells in all M4 crypts (Fig.

147 1F and SI Appendix, Fig. S1). This pattern contrasts with the mammalian small intestine where
148 specialized cell types at the base of crypts express AMP genes (23).

149

150 **CCRs have antibacterial activity but gut colonizers are resistant**

151 We selected CCRs for chemical synthesis on the basis of a consistent prediction of AMP activity,
152 an independently confirmed transcript sequence by cloned cDNA sequencing (21), taking into
153 account the diversity of expression patterns, including peptides expressed in apo and/or sym
154 insects and in the M4B and/or M4, with high or medium expression levels, and favouring smaller
155 peptides to increase feasibility of successful peptide synthesis (synthesis attempts for several
156 initially selected peptides failed) (Fig. 1C and SI Appendix, Table S1). A total of seven CCRs,
157 together with thanatin and riptocin, two known innate immunity-related AMPs of *R. pedestris* (24),
158 LL37 and NCR335, from mammal and plant origin respectively (25, 26) and bacterial polymyxin B
159 (PMB), were tested for growth inhibiting activity against a panel of taxonomically diverse bacterial
160 species consisting in *Bacillus subtilis*, *Sinorhizobium meliloti*, *Paraburkholderia fungorum* and
161 *C. insecticola*. The first two species are typical, well-studied soil bacteria known to be unable to
162 colonize the *R. pedestris* M4 crypts while the latter two can efficiently proliferate in the crypts (19).
163 In agreement with the bioinformatics predictions, the CCRs had growth inhibiting activity against
164 *B. subtilis* and *S. meliloti* although with variable strengths (Fig. 2A, B). On the other hand, the two
165 species, *P. fungorum* and *C. insecticola*, that are able to colonize the gut crypts, are not or only
166 weakly affected by the tested CCRs (Fig. 2A, B). This pattern of sensitivity/resistance to CCRs
167 matches with the response of these species to PMB and in part to the other tested peptides.

168

169 **CCRs have membrane-damaging bactericidal activity**

170 CFU counting showed the bacterial reduction from 10^7 CFU to no colonies after treatment of
171 sensitive *S. meliloti* with the CCR1659 peptide for a few hours, indicating that the growth inhibition
172 results from a bactericidal activity, similarly as for PMB (Fig. 3A). The bactericidal activity of
173 CCR1659 was abolished by prior Proteinase K treatment of the peptide and inhibited by the
174 presence of the divalent cations Ca^{2+} and Mg^{2+} , which interfere with the electrostatic interaction of
175 AMPs with negatively charged membrane lipids and diminish the activity of membrane-targeting
176 AMPs (27) (Fig. 3A; SI Appendix, Fig. S2). To acquire insight in the killing mode of CCR1659, we
177 tested the hypothesis that the peptide disrupt bacterial membranes, like PMB and the other tested
178 AMPs do (28-30). Outer and inner membrane integrities in *S. meliloti* were consecutively damaged
179 by both CCR1659 and PMB treatment, as measured respectively by 1-N-PhenylNaphthylamine
180 (NPN) and Propidium Iodide (PI) uptake leading to enhanced fluorescence (Fig. 3B). In agreement
181 with the membrane disruption, fluorescence microscopy showed that FITC-modified CCR1659
182 labelled the envelope of *S. meliloti* cells in a similar way as polylysine-FITC, which is a polycation
183 known to interact with negatively charged membranes of bacteria (31, 32) (Fig. 3C). Binding of
184 CCR1659 to the envelope suggests that its killing efficiency depends on the strength of envelope
185 binding. To test this assumption, we measured with flow cytometry the binding level of CCR1659-
186 FITC and polylysine-FITC to the above panel of species. Strikingly, CCR1659-sensitive *S. meliloti*
187 and *B. subtilis* were strongly labeled with these two molecules while resistant *C. insecticola* and
188 *P. fungorum* only weakly (Fig. 3D). Thus, the level of binding to cells is correlated with the
189 susceptibility/resistance pattern. Scanning electron microscopy (SEM) of CCR1659-treated
190 *S. meliloti* cells further confirmed the membrane-perturbing activity of the peptide that provoked the
191 formation of fibrous materials from damaged cells similarly as PMB (Fig. 3E; SI Appendix, Fig. S3)
192 and similarly as reported for bacteria and yeasts treated with other types of membrane-disrupting
193 AMPs (33-38). Together, this data reveal that the M4 symbiotic region of the gut produces a
194 remarkably large arsenal of CCR peptides with membrane-damaging AMP activity.

195

196 **The *Caballeronia insecticola* genetic repertoire determining AMP resistance**

197 Species that colonize the midgut display a high level of resistance to CCRs and other AMPs
198 suggesting that resistance is a prerequisite for efficient gut colonization. To test this hypothesis, we
199 aimed to identify the resistance determinants in *C. insecticola* and assess if they control gut
200 colonization. A transposon mutant library (39) was used to perform a Tn-seq screen with PMB,
201 since PMB has a similar membrane action as CCR peptides and is commercially accessible in
202 sufficient quantities for Tn-seq experiments. The screen, performed with three sub-lethal PMB
203 concentrations, resulted in 54 genes whose mutation provoked a fitness defect with the highest
204 concentration. With the lower PMB concentrations, subsets of these genes were identified
205 suggesting a multifactorial resistance with some mechanisms contributing more strongly than
206 others (Fig. 4A, B; SI Appendix, Fig. S4 and Data S2). In agreement with the membrane-targeting
207 mode of action of PMB, the majority of fitness genes are involved in the generation of bacterial
208 envelope components, including LPS, peptidoglycan, phospholipids, hopanoids and membrane
209 protein machineries. In order to validate the Tn-seq results, we constructed insertion and deletion
210 mutants in 11 genes selected among the 54 PMB fitness genes. These genes are predicted to be
211 involved in the biosynthesis of the LPS core (*dedA*, *waaC* and *waaF*) (40-42), LPS O-antigen (*wbiF*,
212 *wbiG*, *wbil*, *wzm* and *rfbA*) (43), peptidoglycan (*dedA*), membrane protein machineries (*tolB* and
213 *tolQ*) (44, 45), in addition to a gene (*tpr*) encoding a tetratricopeptide repeat protein of unknown
214 function. Complementing strains were constructed for some of the mutants. Sensitivity assays with
215 PMB and colistin (COL), another polymyxin-family AMP, confirmed that each mutant had an 8-
216 32-fold increased sensitivity compared to the WT (Fig. 4C) while the complemented mutants were
217 restored to WT-levels (SI Appendix, Fig. S5). Thus, the Tn-seq analysis correctly identified genetic
218 determinants for PMB resistance in *C. insecticola*.

219 In line with the sensitivity of the mutants to PMB and the membrane-attacking properties of
220 CCRs, we found that all mutants were more sensitive than WT for at least one of the tested CCR
221 peptides and the other available AMPs (Fig. 4C, D). The *tolB* and *tolQ* mutants were the least
222 sensitive and displayed only a slight difference compared to WT for all tested peptides. The *dedA*
223 and *tpr* mutants were strongly affected by the CCR1659 peptide (Fig. 4D) and moderately by the
224 other tested peptides. The mutants *wzm*, *wbiF*, *wbiG*, *wbil* and *rfbA* were sensitive to several of the
225 tested CCRs although in many cases, enhanced sensitivity was not resulting in a complete growth
226 inhibition but in a retarded and lesser growth compared to untreated control and the WT grown with
227 the same peptide concentration. The *waaC* and *waaF* mutants were the most strongly affected,
228 being more sensitive than WT to all tested peptides and at higher peptide concentrations, their
229 growth was completely blocked (Fig. 4D). Taken together, the *C. insecticola* genes that were
230 revealed by the PMB Tn-seq screen, contribute also to resistance towards other membrane-
231 attacking AMPs, including the CCRs.

232

233 **Different pathways contribute to AMP resistance in *Caballeronia insecticola***

234 Because the tested AMPs interfere with bacterial membrane function, we characterized the cell
235 envelope of the mutants. Since some of the mutated genes are known or suspected to be involved
236 in LPS biosynthesis, we analyzed the LPS structure of all mutants by PAGE profiling and by mass
237 spectrometry analysis of their lipid A moiety, which is proposed to be a direct target of PMB (28, 29)
238 (Fig. 5A, B). The *tpr*, *dedA* and *tolQ* mutants had a PAGE LPS profile that was indistinguishable
239 from the WT. The *wzm*, *rfbA*, *wbiF*, *wbiG* and *wbil* mutants produced a similar LPS that lacked the
240 O-antigen but had a lipid A/core oligosaccharide moiety that was indistinguishable from WT while
241 the *waaC* and *waaF* mutants had an altered lipid A/core moiety, in agreement with the predicted
242 heptosyl-transferase activity of the encoded enzymes that perform the first steps of the core
243 oligosaccharide synthesis. Mass spectrometry analysis of the lipid A moieties suggested that none

244 of the mutants had an altered lipid A structure and notably, that all mutants produced lipid A carrying
245 the 4-amino-4-deoxy-L arabinose (Ara4N) modification that is known to confer PMB resistance in
246 related bacterial species (41, 42, 46) (Fig. 5B; SI Appendix, Fig. S6).

247 We assessed the steady-state outer membrane integrity of the mutants by NPN labeling and
248 sensitivity to detergents (SI Appendix, Fig. S7). The *waaC* and *waaF* mutants had a higher NPN-
249 derived fluorescence and slightly higher sensitivity to the non-ionic detergent Triton X100 and the
250 cationic detergent CTAB than the WT, while the other mutants were similar to WT. The *tolB* and
251 *tolQ* mutants on the other hand were more sensitive to the anionic detergent SDS than to other
252 tested strains. Overall, this indicates that although the outer membrane in some mutants has a
253 reduced robustness, the AMP sensitivity of the mutants is not a direct consequence of a generic
254 membrane instability but of the deficiency of specific resistance mechanisms.

255 The capacity of the bacterial envelope to bind membrane-disrupting AMPs is a parameter
256 influencing AMP sensitivity. The *waaC* and *wbiF* LPS mutants showed indeed a strong labeling of
257 their envelope with CCR1659-FITC, contrary to the WT that did not show any labelling (Fig. 5C).
258 However, the *tpr* mutant was also not labeled. Therefore, we quantified the relative capacity of the
259 envelope of all the mutants to bind membrane-disrupting AMPs by labeling the cells with the
260 fluorescent polylysine-FITC peptide or CCR1659-FITC, followed by flow cytometry analysis
261 (Fig. 5D). All the mutants with altered LPS (*waaC*, *waaF*, *wzm*, *rfaA*, *wbiF*, *wbiG* and *wbiI*) had a
262 strongly enhanced labeling with both peptides indicating a more accessible cell surface for AMP
263 binding. However, the *dedA* and *tpr* mutants displayed a peptide labeling that was identical to the
264 WT while the *tolB* and *tolQ* mutants were even labelled less intensively. Thus, the LPS mutants
265 might be more sensitive to the AMPs because of the higher accessibility of their membranes for
266 interactions with AMPs but the sensitivity of the *dedA*, *tpr* and *tolBQ* mutants has to be explained
267 by a different mechanism. Interestingly, crypt-colonizing *C. insecticola* bacteria have lost their O-
268 antigen after establishing in the crypts (24) and thus have an LPS that is similar to the LPS of the
269 *wzm*, *rfaA*, *wbiF*, *wbiG* and *wbiI* mutants. In agreement, bacteria isolated from the crypts are
270 hypersensitive to PMB and the CCR1659 peptide (Fig. 4C, D) and they strongly bind polylysine-
271 FITC and CCR1659-FITC (Fig. 5D).

272 To confirm that the set of mutants are affected in different pathways for AMP resistance, we
273 created the *waaC/tpr*, *waaC/dedA* and *waaC/wbiF* double mutants. We reasoned that if genes are
274 part of the same pathway, double mutants should not show an additive phenotype compared to the
275 single mutants, while in case genes are in separate pathways, double mutants might display a more
276 severe phenotype than single mutants. We found that the three double mutants were more
277 sensitive than the corresponding single mutants to PMB and CCR1659 and bound more CCR1659-
278 FITC (SI Appendix, Fig. S8), suggesting that indeed “*waaC* and *tpr*” or “*waaC* and *dedA*” or “*waaC*
279 and *wbiF*” define different pathways to PMB resistance. The synthetic phenotype of the *waaC/wbiF*
280 mutant further suggest that the LPS core and the O-antigen constitute two distinct barriers for AMPs
281 to reach the membrane.

282 SEM of untreated WT and *tpr*, *dedA*, *tolB*, *waaC* and *wzm* mutants showed that the mutants
283 affect the bacterial envelope in various ways (SI Appendix, Fig. S9). SEM of CCR1659-treated cells
284 reveals that the response to the peptide in the *waaC* and *tpr* mutant is markedly different. In the
285 *waaC* mutant, very strong membrane distortions are visible and frequent cell lysis, indicated by the
286 cellular material released from cells. The *tpr* mutant on the other hand shows only minor
287 modifications on the cell surface, similar to WT, although infrequent release of large amounts of
288 cellular material was also observed (Fig. 5E). Collectively, the properties of the single and double
289 mutants suggest that in *C. insecticola* different mechanisms contribute to AMP resistance.

290

291 **AMP resistance in *Caballeronia insecticola* is crucial for midgut colonization**

292 Since the midgut crypts are the site of intensive AMP production, we next analyzed the capacity of
293 the AMP sensitivity mutants to colonize the M4 midgut region of the *R. pedestris* midgut. As a
294 preliminary test and to exclude that gut colonization phenotypes can be attributed to trivial reasons,
295 we confirmed that each mutant has similar growth patterns as WT (SI Appendix, Fig. S10A) and is
296 motile (SI Appendix, Fig. S10B) since motility is crucial for colonization of the M4 crypts (14).
297 Analysis at 5 days post infection (dpi) of second instar nymphs showed that the 11 mutants had
298 the capacity to colonize the crypts although they were to various extends less efficient than the WT.
299 The WT had a 100% efficiency (n=10) and the number of bacteria per gut was consistently high
300 ($>10^7$ genome copies per gut). In contrast, the mutants displayed a large variability in colonization
301 level between insect individuals, ranging from a wild-type colonization level for some individuals to
302 a failure to establish in the crypts in other individuals (Fig. 6A). The *waaC* and *tpr* mutants were
303 particularly affected in agreement with their strong AMP sensitivity. This intriguing probabilistic
304 colonization of the gut by the mutants is reminiscent to stochastic colonization of the *Drosophila*
305 gut by underperforming *Lactobacillus plantarum* strains while a strong colonizer strain had a 100%
306 efficiency (47).

307 Next, we evaluated the fitness of the mutants in M4 colonization when they were in competition
308 with WT. Insects were infected with fifty-fifty mixtures of RFP-marked WT and one of the mutants
309 (or WT as a control) that were marked with GFP. The outcome of the competitions was analyzed
310 at 5 dpi by fluorescence microscopy of dissected M4 midguts and flow cytometry quantification of
311 their bacterial content (Fig. 6B). In the control competition, RFP- and GFP-marked WT were kept
312 in balance in the M4 crypts. However, in competitions with the mutants, the WT nearly completely
313 outcompeted each of them, confirming their reduced colonization capacity.

314 Finally, we also tested if the mutants have maintained or lost the capacity to outcompete a less
315 efficient crypt colonizing species. We previously showed that *P. fungorum* can efficiently colonize
316 the M4 crypts in the absence of competing strains but that it is outcompeted by *C. insecticola* when
317 both strains are co-infecting the *R. pedestris* midgut (19). Here, the outcompetition in the M4 crypts
318 of *P. fungorum* by *C. insecticola* WT in co-infection experiments was confirmed while *wzm*, *waaC*,
319 *tolB*, *tpr* and *dedA* mutants were significantly less efficient in outcompeting *P. fungorum* (Fig. 6C).
320 Thus, high AMP resistance in *C. insecticola* is an important factor contributing to the efficiency of
321 this strain in occupying the *R. pedestris* gut.

322

323

324 Discussion

325

326 The microbiota biogeography in the *R. pedestris* midgut shows a sharp divide between the anterior
327 midgut, which has a highly variable, diverse and relatively low abundant microbiota, and the
328 posterior midgut region that carries in striking contrast a dense mono-specific bacterial population
329 that is strictly a *Caballeronia* species. The two principal findings from this work are that this posterior
330 midgut region and the immediately adjacent anterior region is a highly challenging environment for
331 bacteria because of the abundant presence of symbiosis-specific, membrane-damaging
332 antimicrobial CCRs (Figs. 1-3) and that resistance to these AMPs is crucial for bacteria to colonize
333 the crypts in the posterior midgut (Figs. 4-6). Thus, we propose that the CCRs are new players,
334 acting together with previously identified sorting mechanisms (14, 16-19), in the creation of the
335 biogeography by eliminating sensitive bacteria. A prediction that follows from this hypothesis is that
336 the prevention of CCR production in the gut (for example by the knock-down of a global regulator
337 of CCR gene expression) would alleviate the strict selectivity for *Caballeronia* in the M4 crypts.

338 The expression of the majority of the CCR genes is correlated with crypt colonization because
339 they are specifically expressed in the M4 crypt region of the midgut and they are frequently induced
340 by bacterial colonization of the crypts (Fig. 1). A few of them are also expressed in the upstream
341 M3 midgut region, where they still may have a function related to crypt colonization, for example
342 by preselecting bacterial species. Overall, their gene expression pattern, combined with their

343 secretory nature suggesting that they are released in the lumen of the midgut, is consistent with a
344 function of the CCRs in interacting with the bacterial community during midgut infection as well as
345 during M4 colonization.

346 Our analyses demonstrated that CCR peptides act through membrane interaction and damage
347 (Fig. 3), similarly to most AMPs produced by eukaryotic organisms (30, 48). AMPs damage
348 membranes by first interacting with negative charges exposed on the membrane. In many Gram -
349 negative bacteria, the negative charges carried by phosphate groups on the lipid A moiety of LPS
350 are particularly important for this electrostatic interaction (29, 30). However, in *C. insecticola*,
351 including in the AMP-sensitive mutants, these lipid A charges are converted into positive charges
352 through the Ara4N modifications and therefore, the lipid A is likely not the target of AMPs in *C.*
353 *insecticola*. We propose that the O-antigen and core oligosaccharide of LPS form a safeguard
354 around the cell that limits the access of AMPs to their targets in the membrane. This hypothesis is
355 consistent with the enhanced sensitivity and peptide-binding of mutants without O-antigen or LPS
356 core (Fig. 5). The direct targets of the AMPs are presently unknown but could be revealed by the
357 analysis of the other genetic determinants of AMP resistance in *C. insecticola* identified here,
358 including *tpr*, *dedA*, *tol-pal*, *tamAB*, *rpoE*, and hopanoid and phospholipid biosynthesis genes.

359 We conclude from our infection experiments that the reduced resilience to CCRs of *C.*
360 *insecticola* mutants in different resistance mechanisms makes them less apt to colonize the midgut
361 crypts (Fig. 6). This correlates with the inability of strongly sensitive bacterial species to colonize
362 the midgut (Fig. 2) (19). Presumably, AMP-resistance is critical during the initial infection stages,
363 when a few hundred cells enter into the crypt region and this founder population subsequently
364 multiplies rapidly, in two to three days, to a crypt-space-filling population of about 10^7 - 10^8 bacteria
365 (14, 16). The surprisingly large diversity of CCR peptides, several of them already expressed in the
366 M3 and M4 before the microbiota establishment, could be an adaptation to create a selective
367 environment that restricts the type of bacteria from the anterior midgut microbiota that have a
368 chance to establish in the M4 crypts and that favors optimal beneficial *Caballeronia* strains. Such
369 a molecular filter of bacteria could arise from additive, synergistic or specific antimicrobial activities
370 of different CCR peptides towards distinct bacteria. Indeed, the tested CCRs have variable
371 antimicrobial efficiency against different bacterial species and *C. insecticola* mutants. Recent
372 insights from *Drosophila* and other models have changed the previous view on AMPs as generic,
373 non-specific antimicrobials by the demonstration that they can display a degree of specificity and
374 synergism. Accordingly, AMP repertoires in organisms dynamically evolve according to the
375 diversity of microbes encountered in the natural environment (30, 48, 49). The hundreds of diverse
376 CCR peptides might be an extreme example of such an evolutionary process.

377 On the other hand, once established in the M4 crypts, *C. insecticola* loses its O-antigen by an
378 unknown mechanism (24), which renders them sensitive to the CCR peptides (Fig. 4C, D). This
379 suggests a second function of the CCR peptide arsenal - in particular for those peptides encoded
380 by the late-expressed genes - that could be related to the protection of the crypt epithelia and
381 prevention of the bacteria breaching these epithelia. Indeed, in *R. pedestris* the crypt epithelium
382 lacks mucus or peritrophic protective layers and is therefore in direct contact with the microbiota
383 (17). Additionally, the membrane fragilization of the crypt-colonizing bacteria by the CCRs could
384 facilitate the retrieval of nutrients from the bacteria (50), suggesting that the insect tames the gut
385 symbionts with the CCRs. Although these additional functions should be confirmed in future
386 studies, they may also be related to the CCR diversification in *R. pedestris*.

387

388

389 **Materials and Methods**

390

391 Detailed protocols for all used procedures are available in SI Appendix. Bacterial strains and
392 materials used are listed in SI Appendix, Table S1. Tn-seq screening with PMB was done with an
393 available *Himar1* transposon mutant Tn-seq library (39). Plasmid constructs for mutagenesis or

394 mutant complementation were obtained by gene synthesis or Gibson cloning. Plasmids were
395 introduced in *C. insecticola* by mating and deletion mutants were obtained by double homologous
396 recombination. Strains were tagged with fluorescent proteins using mutagenesis with modified Tn7
397 transposons. Antimicrobial peptide activity and detergent sensitivity assays were performed in 96-
398 well plates and determining growth curves in the presence of test peptide dilution series or by cfu
399 counting after peptide exposure during variable times. Membrane interactions of CCR peptides
400 were determined by NPN and PI uptake assays, fluorescence microscopy, flow cytometry and
401 SEM. LPS was obtained by phenol extraction. Total LPS was analyzed by SDS-PAGE and the
402 lipid A fraction, obtained by triethylamine-citrate treatment, was analyzed by matrix-assisted laser
403 desorption ionization–time of flight mass spectrometer.

404 The rearing conditions for *R. pedestris* and the procedures for colonization assays were done
405 as before (13, 39) and are detailed in SI Appendix. Colonization efficiency by tested bacterial
406 strains were determined by light and fluorescence microscopy, qPCR and flow cytometry.
407 Transcriptome analysis of the midgut of colonized and aposymbiotic insects was done by RNA-seq
408 and Illumina sequencing. Whole-mount *in situ* hybridization on the *R. pedestris* midgut was
409 performed with digoxigenin (DIG)-labeled *CCR043* cRNA probe. Bioinformatic procedures for data
410 analysis are available in SI Appendix.

411

412

413 **Data availability**

414

415 RNA-seq sequencing data are available in the Sequence Read Archive (SRA), BioProject
416 accession no. PRJNA1006624. The *de novo* assembled transcriptome was deposited in the
417 Transcriptome Shotgun Assembly (TSA), BioProject accession no. PRJNA1006624. Tn-seq
418 sequencing data were deposited in SRA, BioProject accession no. PRJNA890438.

419

420

421 **Acknowledgments**

422

423 We are grateful to Olga Soutourina (University Paris-Saclay, France) and Yu Matsuura (University
424 of the Ryukyus, Japan) for critical reading of the manuscript and constructive comments. This work
425 benefited from financial support by Saclay Plant Sciences-SPS, by the ANR grant ANR-19-CE20-
426 0007, and by a JSPS-CNRS Bilateral Open Partnership Joint Research Project (18KK0211) and a
427 CNRS International Research Project to Y.K. and P.M. Y.K. was supported by the MEXT KAKENHI
428 (21K18241, 22H05068, 22B303). J.L., G.L. and R.J. were supported by Ph.D. fellowships from the
429 French Ministry of Higher Education, Research, and Innovation and A.B. benefited from a PhD
430 contract in the frame of the CNRS 80|PRIME – 2021 program. This work was supported by JSPS
431 Research Fellowships for Young Scientist to S.J. (21F21090), K.I. (22KJ0057) and T.O. (14J03996,
432 20170267 and 19J01106). Tn-seq sequencing and data treatment were performed by the I2BC
433 high-throughput sequencing facility, supported by France Génomique (funded by the French
434 National Program Investissement d’Avenir ANR-10-INBS-09). This work has benefited from the
435 facilities and expertise of MIMA2 (Université Paris-Saclay, INRAE, AgroParisTech, 78350, Jouy-
436 en-Josas, France).

437

438

439 **References**

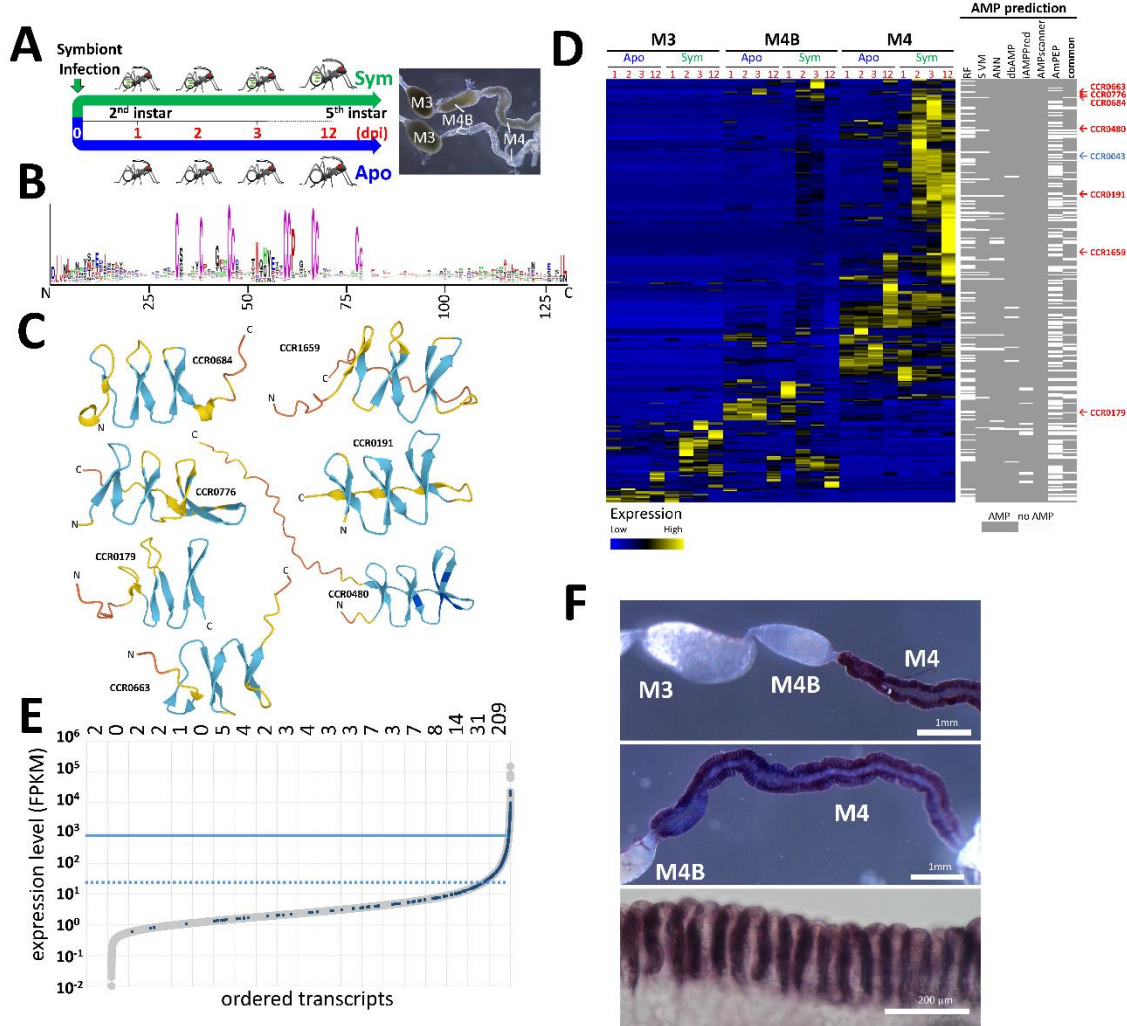
440

- 441 1. N. A. Moran, H. Ochman, T. J. Hammer, Evolutionary and ecological consequences of gut
442 microbial communities. *Annual Review of Ecology, Evolution, and Systematics* 50, 451-475
443 (2019). doi: 10.1146/annurev-ecolsys-110617-062453.

- 444 2. J. C. Clemente, L. K. Ursell, L. W. Parfrey, R. Knight, The impact of the gut microbiota on
445 human health: an integrative view. *Cell* 148, 1258-1270 (2012). doi:
446 10.1016/j.cell.2012.01.035.
- 447 3. G. McCallum, C. Tropini, The gut microbiota and its biogeography. *Nature Reviews*
448 *Microbiology* (2023). doi: 10.1038/s41579-023-00969-0.
- 449 4. G. P. Donaldson, S. M. Lee, S. K. Mazmanian, Gut biogeography of the bacterial microbiota.
450 *Nature Reviews Microbiology* 14, 20-32 (2016). doi: 10.1038/nrmicro3552.
- 451 5. Z. Yao, Z. Cai, Q. Ma, S. Bai, Y. Wang, P. Zhang, Q. Guo, J. Gu, B. Lemaitre, H. Zhang,
452 Compartmentalized PGRP expression along the dipteran *Bactrocera dorsalis* gut forms a zone
453 of protection for symbiotic bacteria. *Cell Reports* 41, 111523 (2022). doi:
454 10.1016/j.celrep.2022.111523.
- 455 6. R. Dodge, E.W. Jones, H. Zhu, B. Obadia, D.J. Martinez, C. Wang, A. Aranda-Díaz, K.
456 Aumiller, Z. Liu, M. Voltolini, E. L. Brodie, K. C. Huang, J. M. Carlson, D. A. Sivak, A. C.
457 Spradling, W. B. Ludington, A symbiotic physical niche in *Drosophila melanogaster* regulates
458 stable association of a multi-species gut microbiota. *Nature Communications* 14, 1557 (2023).
459 doi: 10.1038/s41467-023-36942-x.
- 460 7. Z. Hu, C. Zhang, L. Sifuentes-Dominguez, C. M. Zarek, D. C. Propheter, Z. Kuang, Y. Wang,
461 M. Pendse, K. A. Ruhn, B. Hassell, C. L. Behrendt, B. Zhang, P. Raj, T. A. Harris-Tryon, T. A.
462 Reese, L. V. Hooper, Small proline-rich protein 2A is a gut bactericidal protein deployed during
463 helminth infection. *Science* 374, eabe6723 (2021). doi: 10.1126/science.abe6723.
- 464 8. D. Sun, R. Bai, W. Zhou, Z. Yao, Y. Liu, S. Tang, X. Ge, L. Luo, C. Luo, G. F. Hu, J. Sheng,
465 Z. Xu, Angiogenin maintains gut microbe homeostasis by balancing α -Proteobacteria and
466 Lachnospiraceae. *Gut* 70, 666-676 (2021). doi: 10.1136/gutjnl-2019-320135.
- 467 9. R. L. Gallo, L. V. Hooper, Epithelial antimicrobial defence of the skin and intestine. *Nature*
468 *Reviews Immunology* 12, 503-516 (2012). doi: 10.1038/nri3228.
- 469 10. P. Mergaert, Role of antimicrobial peptides in controlling symbiotic bacterial populations.
470 *Natural Product Reports* 35, 336-356 (2018). doi: 10.1039/c7np00056a.
- 471 11. T. W. Cullen, W. B. Schofield, N. A. Barry, E. E. Putnam, E. A. Rundell, M. S. Trent, P. H.
472 Degnan, C. J. Booth, H. Yu, A. L. Goodman, Antimicrobial peptide resistance mediates
473 resilience of prominent gut commensals during inflammation. *Science* 347, 170-175 (2015).
474 doi: 10.1126/science.1260580.
- 475 12. A. Arias-Rojas, D. Frahm, R. Hurwitz, V. Brinmann, I. Iatsenko, Resistance to host
476 antimicrobial peptides mediates resilience of gut commensals during infection and aging in
477 *Drosophila*. *Proceedings of the National Academy of Sciences USA* 120, e2305649120
478 (2023). doi: 10.1073/pnas.2305649120.
- 479 13. Y. Kikuchi, T. Hosokawa, T. Fukatsu, Insect-microbe mutualism without vertical transmission:
480 a stinkbug acquires a beneficial gut symbiont from the environment every generation. *Applied*
481 *and Environmental Microbiology* 73, 4308-4316 (2007). doi: 10.1128/AEM.00067-07.
- 482 14. T. Ohbayashi, K. Takeshita, W. Kitagawa, N. Nikoh, R. Koga, X. Y. Meng, K. Tago, T. Hori,
483 M. Hayatsu, K. Asano, Y. Kamagata, B. L. Lee, T. Fukatsu, Y. Kikuchi, Insect's intestinal organ
484 for symbiont sorting. *Proceedings of the National Academy of Sciences USA* 112, E5179-5188
485 (2015). doi: 10.1073/pnas.1511454112.
- 486 15. S. Jang, K. Ishigami, P. Mergaert, Y. Kikuchi, Ingested soil bacteria breach gut epithelia and
487 prime systemic immunity in an insect. *Proceedings of the National Academy of Sciences USA*
488 121, e2315540121 (2024). doi: 10.1073/pnas.2315540121.
- 489 16. Y. Kikuchi, T. Ohbayashi, S. Jang, P. Mergaert, *Burkholderia insecticola* triggers midgut
490 closure in the bean bug *Riptortus pedestris* to prevent secondary bacterial infections of midgut
491 crypts. *The ISME Journal* 14, 1627-1638 (2020). doi: 10.1038/s41396-020-0633-3.
- 492 17. S. Jang, P. Mergaert, T. Ohbayashi, K. Ishigami, S. Shigenobu, H. Itoh, Y. Kikuchi, Dual
493 oxidase enables insect gut symbiosis by mediating respiratory network formation. *Proc. Natl.*
494 *Acad. Sci. USA* 118, e2020922118 (2021). doi: 10.1073/pnas.2020922118.
- 495 18. S. Jang, Y. Matsuura, K. Ishigami, P. Mergaert, Y. Kikuchi, Symbiont coordinates stem cell
496 proliferation, apoptosis, and morphogenesis of gut symbiotic organ in the stinkbug-

- 497 *Caballeronia* symbiosis. *Frontiers in Physiology* 13, 1071987 (2023). doi:
498 10.3389/fphys.2022.1071987.
- 499 19. H. Itoh, S. Jang, K. Takeshita, T. Ohbayashi, N. Ohnishi, X. Y. Meng, Y. Mitani, Y. Kikuchi,
500 Host-symbiont specificity determined by microbe-microbe competition in an insect gut.
501 *Proceedings of the National Academy of Sciences USA* 116, 22673-22682 (2019). doi:
502 10.1073/pnas.1912397116.
- 503 20. K. Takeshita, H. Tamaki, T. Ohbayashi, X. Y. Meng, T. Sone, Y. Mitani, C. Peeters, Y. Kikuchi,
504 P. Vandamme, *Burkholderia insecticola* sp. nov., a gut symbiotic bacterium of the bean bug
505 *Riptortus pedestris*. *International Journal of Systematic and Evolutionary Microbiology* 68,
506 2370-2374 (2018). doi: 10.1099/ijsem.0.002848.
- 507 21. R. Futahashi, K. Tanaka, M. Tanahashi, N. Nikoh, Y. Kikuchi, B. L. Lee, T. Fukatsu, Gene
508 expression in gut symbiotic organ of stinkbug affected by extracellular bacterial symbiont.
509 *PLoS One* 8, e64557 (2013). doi: 10.1371/journal.pone.0064557.
- 510 22. T. Ohbayashi, R. Futahashi, M. Terashima, Q. Barrière, F. Lamouche, K. Takeshita, X. Y.
511 Meng, Y. Mitani, T. Sone, S. Shigenobu, T. Fukatsu, P. Mergaert, Y. Kikuchi, Comparative
512 cytology, physiology and transcriptomics of *Burkholderia insecticola* in symbiosis with the bean
513 bug *Riptortus pedestris* and in culture. *The ISME Journal* 13, 1469-1483 (2019). doi:
514 10.1038/s41396-019-0361-8.
- 515 23. H. C. Clevers, C. L. Bevins, Paneth cells: maestros of the small intestinal crypts. *Annual*
516 *Review of Physiology* 75, 289-311 (2013). doi: 10.1146/annurev-physiol-030212-183744.
- 517 24. J. K. Kim, D. W. Son, C. H. Kim, J. H. Cho, R. Marchetti, A. Silipo, L. Sturiale, H. Y. Park, Y.
518 R. Huh, H. Nakayama, T. Fukatsu, A. Molinaro, B.L. Lee, Insect gut symbiont susceptibility to
519 host antimicrobial peptides caused by alteration of the bacterial cell envelope. *Journal of*
520 *Biological Chemistry* 290, 21042-21053 (2015). doi: 10.1074/jbc.M115.651158.
- 521 25. M. F. Burton, P. G. Steel, The chemistry and biology of LL-37. *Natural Product Reports* 26,
522 1572-1584 (2009). doi: 10.1039/b912533g.
- 523 26. A. Farkas, G. Maróti, A. Kereszt, É. Kondorosi, Comparative Analysis of the Bacterial
524 Membrane Disruption Effect of Two Natural Plant Antimicrobial Peptides. *Frontiers in*
525 *Microbiology* 8, 51 (2017). doi: 10.3389/fmicb.2017.00051.
- 526 27. S. J. Ko, E. Park, A. Asandei, J. Y. Choi, S. C. Lee, C. H. Seo, T. Luchian, Y. Park, Bee venom -
527 derived antimicrobial peptide melectin has broad-spectrum potency, cell selectivity, and salt-
528 resistant properties. *Scientific Reports* 10, 10145 (2020). doi: 10.1038/s41598-020-66995-7.
- 529 28. L. Poirel, A. Jayol, P. Nordmann, Polymyxins: Antibacterial activity, susceptibility testing, and
530 resistance mechanisms encoded by plasmids or chromosomes. *Clinical Microbiology Reviews*
531 30, 557-596 (2017). doi: 10.1128/CMR.00064-16.
- 532 29. C. A. Moubareck, Polymyxins and bacterial membranes: A review of antibacterial activity and
533 mechanisms of resistance. *Membranes* 10, 181 (2020). doi: 10.3390/membranes10080181.
- 534 30. B. P. Lazzaro, M. Zasloff, J. Rolf, Antimicrobial peptides: Application informed by evolution.
535 *Science* 368, eaau5480 (2020). doi: 10.1126/science.aau5480.
- 536 31. R. Spohn, L. Daruka, V. Lázár, A. Martins, F. Vidovics, G. Grézal, O. Méhi, B. Kintsés, M.
537 Számel, P. K. Jangir, B. Csörgő, Á. Györkei, Z. Bódi, A. Faragó, L. Bodai, I. Földesi, D. Kata,
538 G. Maróti, B. Pap, R. Wirth, B. Papp, C. Pál, Integrated evolutionary analysis reveals
539 antimicrobial peptides with limited resistance. *Nature Communications* 10, 4538 (2019). doi:
540 10.1038/s41467-019-12364-6.
- 541 32. F. F. Rossetti, I. Reviakine, G. Csúcs, F. Assi, J. Vörös, M. Textor, Interaction of poly(L-lysine)-
542 g-poly(ethylene glycol) with supported phospholipid bilayers. *Biophysical Journal* 87, 1711-
543 1721 (2004). doi: 10.1529/biophysj.104.041780.
- 544 33. B. O. Schroeder, D. Ehmann, J. C. Precht, P. A. Castillo, R. Küchler, J. Berger, M. Schaller,
545 E. F. Stange, J. Wehkamp, Paneth cell α -defensin 6 (HD-6) is an antimicrobial peptide.
546 *Mucosal Immunology* 8, 661-671 (2015). doi: 10.1038/mi.2014.100.
- 547 34. N. Shagahi, M. Bhave, E. A. Palombo, A. H. A. Clayton, Revealing the sequence of
548 interactions of PuroA peptide with *Candida albicans* cells by live-cell imaging. *Scientific*
549 *Reports* 7, 43542 (2017). doi: 10.1038/srep43542.

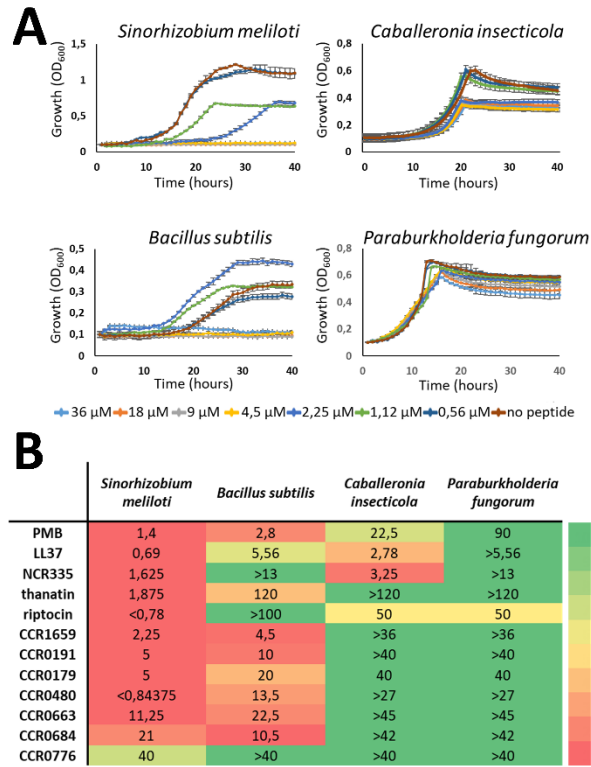
- 550 35. J. Montiel, J. A. Downie, A. Farkas, P. Bihari, R. Herczeg, B. Bálint, P. Mergaert, A. Kereszt, É.
551 Kondorosi, Morphotype of bacteroids in different legumes correlates with the number and
552 type of symbiotic NCR peptides. *Proceedings of the National Academy of Sciences USA* 114,
553 5041-5046 (2017). doi: 10.1073/pnas.1704217114.
- 554 36. A. Farkas, B. Pap, É. Kondorosi, G. Maróti, Antimicrobial activity of NCR plant peptides
555 strongly depends on the test assays. *Frontiers in Microbiology* (2018). doi:
556 10.3389/fmicb.2018.02600.
- 557 37. H. Zhuo, X. Zhang, M. Li, Q. Zhang, Y. Wang, Antibacterial and anti-inflammatory properties
558 of a novel antimicrobial peptide derived from LL-37. *Antibiotics* 11, 754 (2022). doi:
559 10.3390/antibiotics11060754.
- 560 38. J. Shi, C. Chen, D. Wang, Z. Wang, Y. Liu, The antimicrobial peptide LI14 combats multidrug-
561 resistant bacterial infections. *Communications biology* 5, 926 (2022). doi: 10.1038/s42003-
562 022-03899-4.
- 563 39. R. Jouan, G. Lextrait, J. Lachat, A. Yokota, R. Cossard, D. Naquin, T. Timtchenko, Y. Kikuchi,
564 T. Ohbayashi, P. Mergaert, Transposon sequencing reveals the essential gene set and genes
565 enabling gut symbiosis in the insect symbiont *Caballeronia insecticola*. *ISME Communications*
566 (2024). doi: 10.1093/ismeco/ycad001.
- 567 40. J. K. Kim, H. A. Jang, M. S. Kim, J. H. Cho, J. Lee, F. Di Lorenzo, L. Sturiale, A. Silipo, A.
568 Molinaro, B. L. Lee, The lipopolysaccharide core oligosaccharide of *Burkholderia* plays a
569 critical role in maintaining a proper gut symbiosis with the bean bug *Riptortus pedestris*.
570 *Journal of Biological Chemistry* 292, 19226-19237 (2017). doi: 10.1074/jbc.M117.813832.
- 571 41. P. R. Panta, S. Kumar, C. F. Stafford, C. E. Billiot, M. V. Douglass, C. M. Herrera, M. S. Trent,
572 W. T. Doerrler, A DedA family membrane protein is required for *Burkholderia thailandensis*
573 colistin resistance. *Frontiers in Microbiology* 10, 2532 (2019). doi: 10.3389/fmicb.2019.02532.
- 574 42. P. R. Panta, W. T. Doerrler, A link between pH homeostasis and colistin resistance in bacteria.
575 *Scientific Reports* 11, 13230 (2021). doi: 10.1038/s41598-021-92718-7.
- 576 43. J. K. Kim, H. Y. Park, B. L. Lee, The symbiotic role of O-antigen of *Burkholderia* symbiont in
577 association with host *Riptortus pedestris*. *Developmental & Comparative Immunology* 60, 202-
578 208 (2016). doi: 10.1016/j.dci.2016.02.009.
- 579 44. B. Sit, V. Srisuknimit, E. Bueno, F. G. Zingl, K. Hullahalli, F. Cava, M. K. Waldor, Undecaprenyl
580 phosphate translocases confer conditional microbial fitness. *Nature* 613, 721-728 (2023). doi:
581 10.1038/s41586-022-05569-1.
- 582 45. J. Szczepaniak, C. Press, C. Kleanthous, The multifarious roles of Tol-Pal in Gram-negative
583 bacteria. *FEMS Microbiology Reviews* 44, 490-506 (2020). doi: 10.1093/femsre/fuaa018.
- 584 46. M. A. Valvano, Remodelling of the Gram-negative bacterial Kdo₂-lipid A and its functional
585 implications. *Microbiology* 168, 001159 (2022). doi: 10.1099/mic.0.001159.
- 586 47. B. Obadia, Z. T. Güvener, V. Zhang, J. A. Ceja-Navarro, E. L. Brodie, W. W. Ja, W. B.
587 Ludington, Probabilistic invasion underlies natural gut microbiome stability. *Current Biology*
588 27, 1999-2006 (2017). doi: 10.1016/j.cub.2017.05.034.
- 589 48. M. Zasloff, Antimicrobial peptides of multicellular organisms. *Nature* 415, 389-395 (2002). doi:
590 10.1038/415389a.
- 591 49. M. A. Hanson, L. Grollmus, B. Lemaitre, Ecology-relevant bacteria drive the evolution of host
592 antimicrobial peptides in *Drosophila*. *Science* 381, eadg5725 (2023). doi:
593 10.1126/science.adg5725.
- 594 50. P. Mergaert, Y. Kikuchi, S. Shigenobu, E. C. M. Nowack, Metabolic integration of bacterial
595 endosymbionts through antimicrobial peptides. *Trends Microbiology* 25, 703-712 (2017). doi:
596 10.1016/j.tim.2017.04.007.
597



599

600 **Figure 1. CCRs are symbiosis-specific AMP-like peptides.** (A) Experimental setup for
 601 transcriptome analysis. First day second instars were divided in two groups. To one of them, *C.*
 602 *insecticola* symbionts were administered (green, Sym) and the other group remained free of
 603 symbionts (blue, Apo). Insects were dissected in the second (1, 2, and 3 days post inoculation [dpi])
 604 or fifth instar (12 dpi) and the M3, M4B and M4 regions were harvested for transcriptome analysis.
 605 The pictures at the right show representative guts of a Sym insect at 3 dpi (top) and a same age
 606 Apo insect (bottom). (B) Logo profile of the mature CCR peptides identified in the transcriptome,
 607 highlighting the sequence diversity of the peptides and the ten conserved cysteine residues. (C)
 608 AlphaFold2 structural predictions of examples of CCR peptides showing antiparallel β -sheets
 609 carrying the cysteine residues. (D) Blue-black-yellow heat map of the relative expression profile of
 610 the identified CCR genes and white-grey heat map of AMP predictions. Sample identity in the
 611 expression heat map is indicated at the top and is according to panel A. AMP prediction tools are
 612 Random Forest (RF), Support Vector Machine (SVM), Artificial Neural Network (ANN),
 613 AMPpredictor (dbAMP); iAMPpred; Antimicrobial Peptide Scanner (AMPscanner); (AmPEP_v1
 614 and AmPEP_v2). A consensus prediction (6 out of 7 positive predictions) is indicated in the last
 615 column. The peptides used for functional characterization are indicated at the right of the heat
 616 maps. (E) Whole-mount *in situ* hybridization with a CCR0043 antisense probe on the dissected

617 midgut of a 3 dpi symbiotic insect. Positive signal appears with a blue-brownish color. CCR0043 is
618 specifically expressed in the M4 and uniformly in all crypts. Control *in situ* hybridizations on the gut
619 of aposymbiotic insects and with a sense probe on symbiotic insects are shown in Supplemental
620 Figure S2. (F) CCR transcript expression levels. Transcripts are ordered according to their
621 expression level in the x-axis and their expression levels (FPKM, Fragments Per Kilobase of
622 transcript per Million mapped reads) are plotted in the y-axis. All transcripts are indicated with grey
623 dots and the CCR transcripts are indicated with blue crosses. The dotted and plain blue horizontal
624 lines correspond to the mean expression level of all transcripts and CCR transcripts, respectively.
625 The numbers above the plot indicate the number of CCR transcripts present in 5-percentile bins of
626 transcripts. 77 % of the CCR transcripts are among the 10 % highest expressed transcripts in the
627 midgut.
628



629

630

Figure 2. CCR peptides are AMPs. (A) Growth inhibition of the indicated bacterial species by

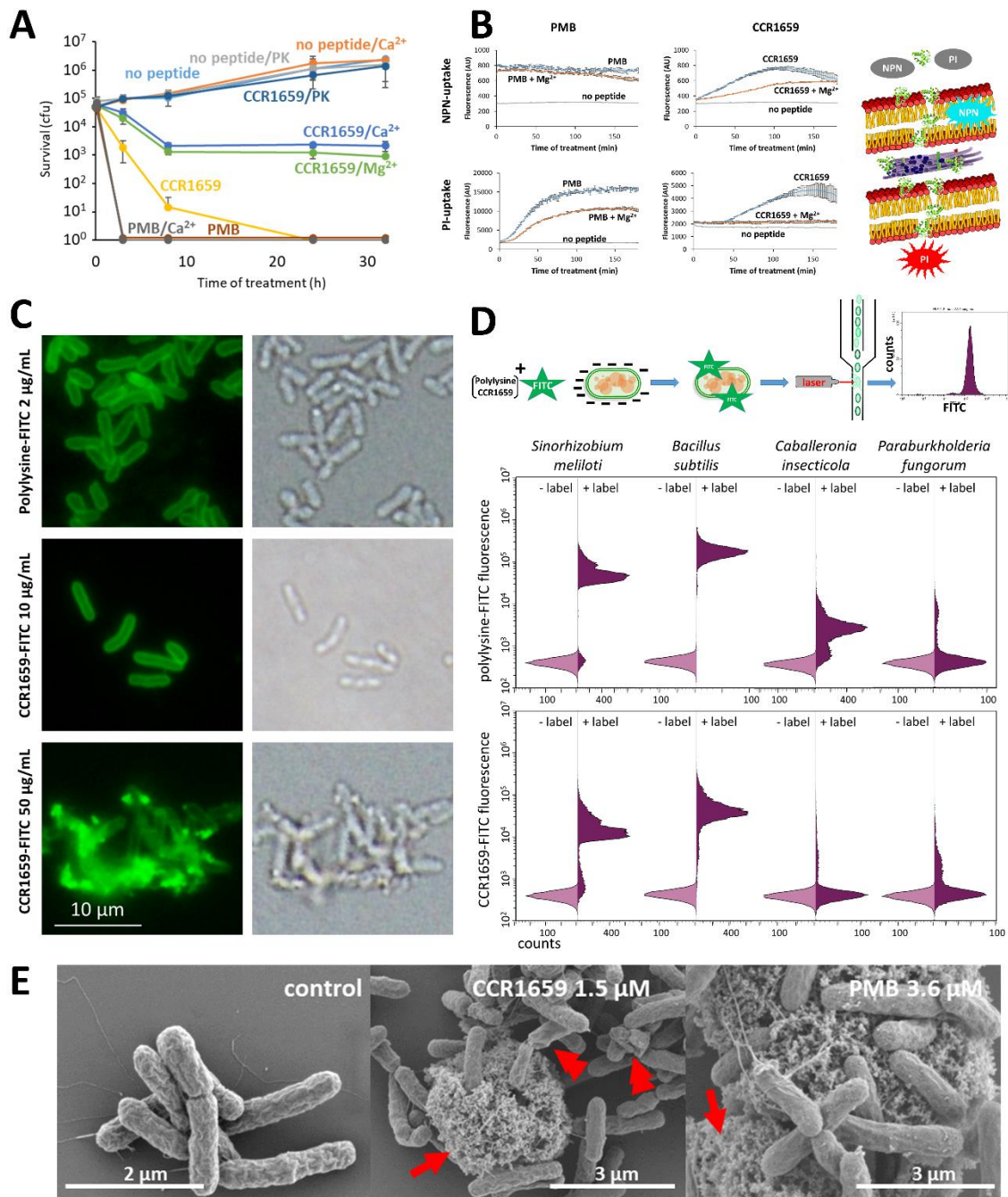
631

different concentrations of CCR1659. Error bars are standard deviations (n=3). (B) Minimal

632

concentrations (in μM) of growth inhibition of the indicated strains by various peptides.

633

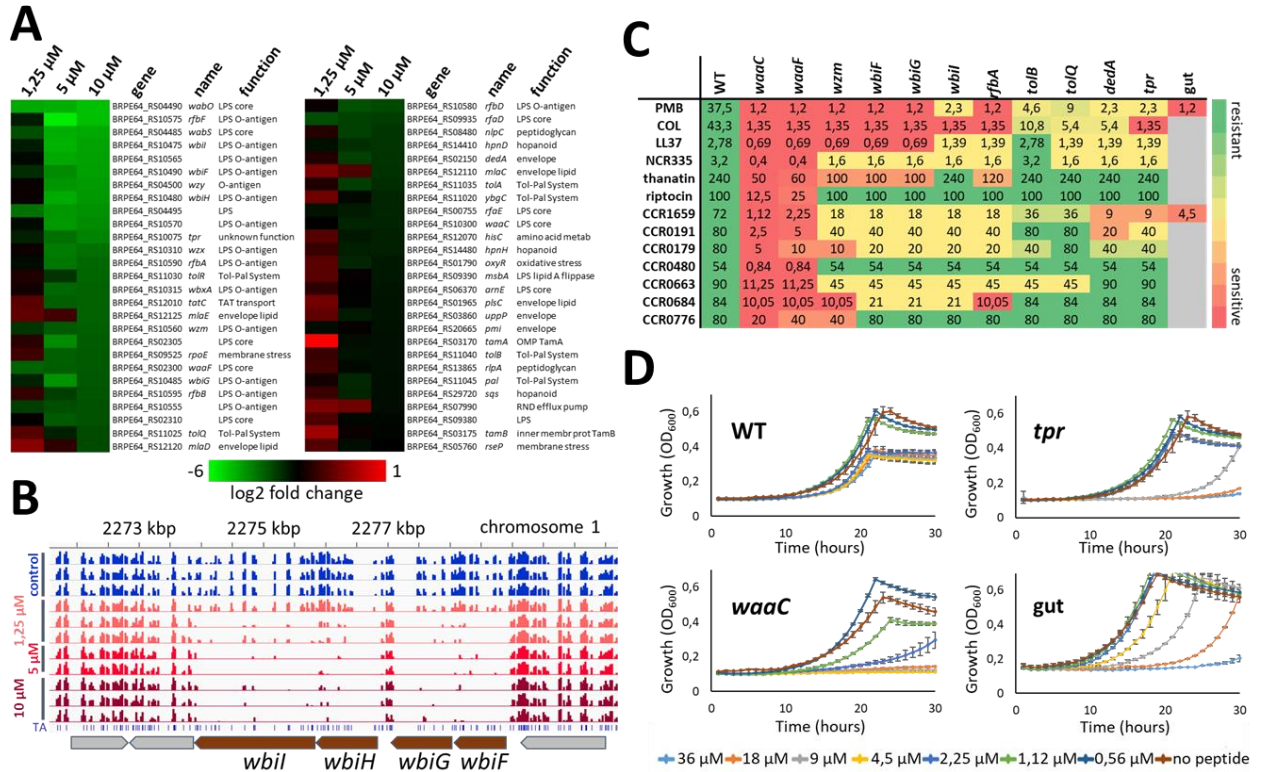


634

635 **Figure 3. CCR peptides target membranes.** (A) Bactericidal activity of 25 μM CCR1659 and 25
 636 μM PMB. PK: proteinase K; Ca²⁺: activity in the presence of 5 mM CaCl₂; Mg²⁺ activity in the
 637 presence of 5 mM MgCl₂. Error bars are standard deviations (n=3). (B) NPN and PI uptake by *S.*
 638 *melliloti* cells in response to treatment with 10 μM CCR1659 or 10 μM PMB in the presence or
 639 absence of 5 mM MgCl₂. NPN is a lipophilic dye that fluoresces in hydrophobic environments such
 640 as bacterial phospholipids exposed by outer membrane damage; PI is a membrane impermeant
 641 DNA-intercalating dye that fluoresces upon DNA binding in the cytoplasm, indicative of
 642 permeabilisation of both the outer and inner membrane. Error bars are standard deviations (n=3).

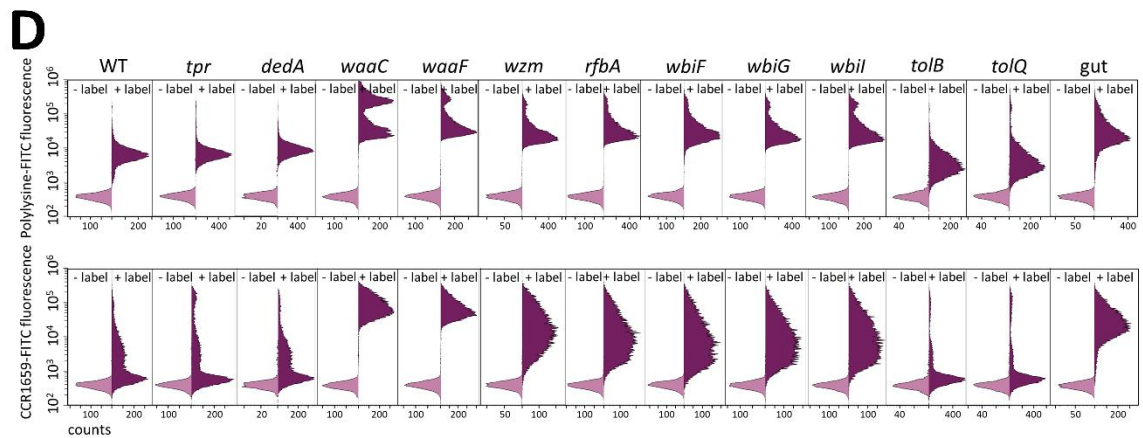
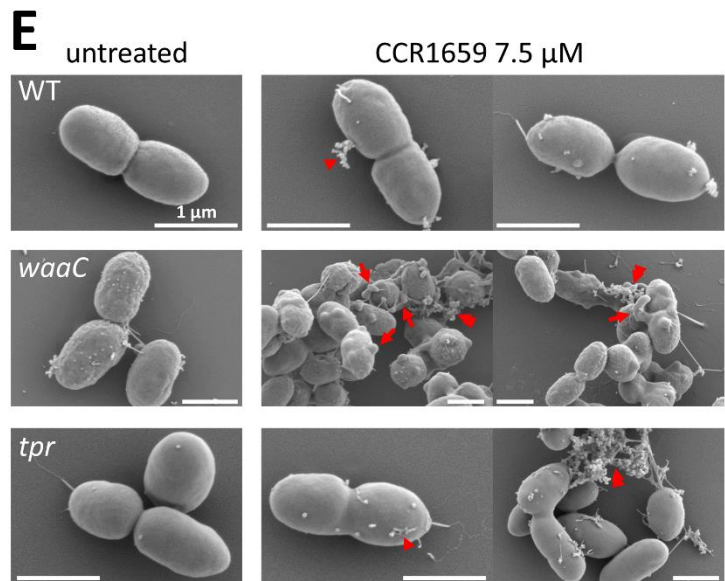
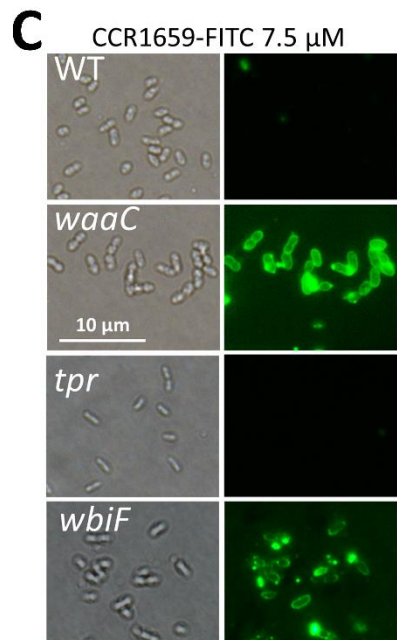
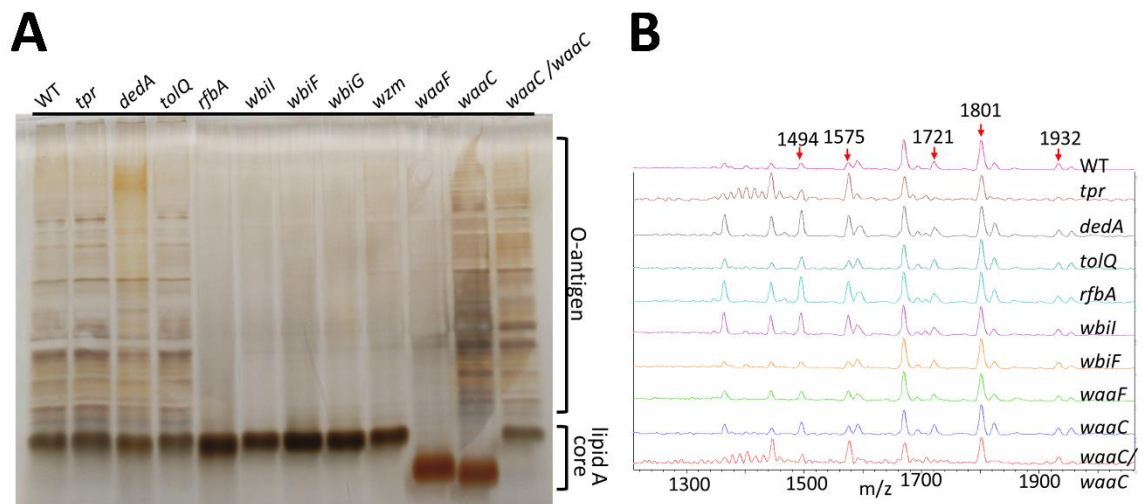
17

643 (C) Fluorescence microscopy (left) of *S. meliloti* cells treated with Polylysine-FITC or CCR1659-
644 FITC at the indicated concentrations. Corresponding bright field images are shown at the right. (D)
645 Flow cytometry analysis of Polylysine-FITC (top) or CCR1659-FITC binding by the indicated
646 strains. Light purple histograms are control measurements without fluorescent label (-label); the
647 dark purple histograms are in the presence of the fluorescent label (+label). (E) SEM micrographs
648 of untreated *S. meliloti* cells (left) or treated with 1.5 μ M CCR1659 (middle) or with 3.6 μ M PMB
649 (right). The arrows indicate cellular material released from cells. The double arrowheads indicate
650 cells with lost turgor.
651



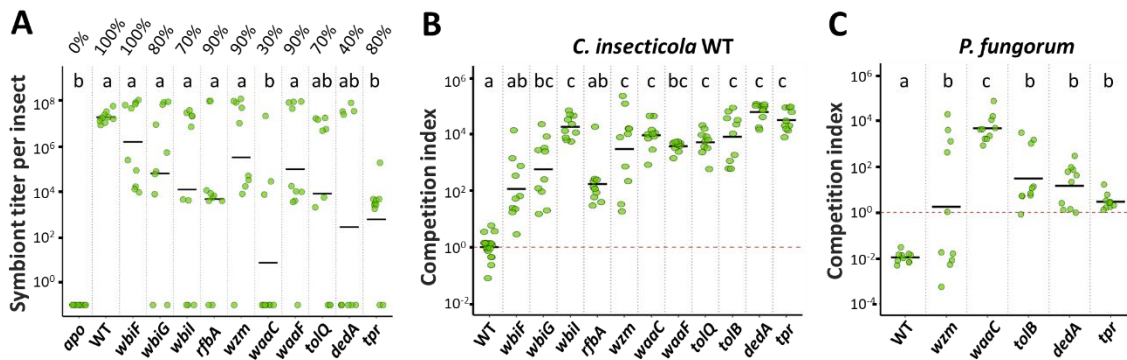
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667

Figure 4. Identification of AMP resistance genes by Tn-seq. (A) Heat map showing the level of depletion of transposon insertions in the indicated genes in the *C. insecticola* population grown in the presence of PMB at the indicated concentrations. The color-code scale indicates the log₂ fold change in the insertion abundance under the test conditions relative to the control conditions. (B) IGV view of Tn-seq sequencing data for a selected genomic region of *C. insecticola*. The histograms indicate the abundance of mutants in the Tn-seq population for the indicated samples. Genes whose products contribute to PMB resistance have a lower frequency of Tn insertions in peptide treatment screens than in the control. (C) Mutants in selected genes are hypersensitive to AMPs. Heat map and minimal concentrations of growth inhibition of the indicated wild-type and mutant strains by the listed peptides. Minimal concentrations are indicated in μM. The color key of the heat map is as indicated at the right. Grey cells indicate not tested. (D) Growth inhibition of the indicated strains by different concentrations of CCR1659. “Gut” in panels C and D indicates cryptocolonizing *C. insecticola* bacteria, directly isolated from dissected M4. Error bars are standard deviations (n=3).



669 **Figure 5. Surface properties of the AMP-sensitive mutants.** (A) Polyacrylamide gel
670 electrophoresis analysis of total LPS extracted from the indicated strains. The *waaC/waaC* strain
671 is the complemented mutant. Despite the altered core in the *waaC* mutant, an O-antigen ladder is
672 visible, that has a similar profile to the wild type, possibly corresponding to the O-antigen anchored
673 on an intermediate lipid carrier. (B) MS analysis of the lipid A molecule present in the indicated
674 mutants. Red arrows indicate the Ara4N carrying lipid A (SI Appendix, Fig. S7). (C) Fluorescence
675 microscopy of *C. insecticola* wild-type, *waaC*, *tpr* and *wbiF* cells treated with 50 µg/mL CCR1659-
676 FITC. All images are at the same magnification and the scalebar is 10 µm. (D) Flow cytometry
677 analysis of 50 µg/mL Polylysine-FITC (top) or 7.5 µM CCR1659-FITC binding by the indicated
678 strains. "Gut" is bacteria directly isolated from the midgut crypts. Light purple histograms are control
679 measurements without fluorescent label (-label); the dark purple histograms are in the presence of
680 the fluorescent label. Note the presence of a double peak in the Polylysine-FITC treated mutants
681 *waaC*, *waaF*, *wzm*, *rfbA*, *wbiFGI*, indicating of a heterogeneous bacterial population. (E) SEM
682 micrographs of untreated *C. insecticola* wild type and *waaC* and *tpr* mutant untreated cells or
683 treated with 7.5 µM CCR1659. Arrowheads indicate release of tiny amounts of cellular material in
684 intact cells. Double arrowheads indicate cellular material released from lysed cells. Arrows indicate
685 cell deformations. Scale bars are 1 µm for all images.

686



687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

Figure 6. AMP-sensitive mutants are impaired in gut colonization. (A) Single-strain infections of *R. pedestris* second instar nymphs with *C. insecticola* WT or indicated mutants or no bacteria (*apo*). Colonization of the M4 crypt region was determined at 5 dpi by dissection and microscopy observation of the guts and symbiont titer determination by qPCR in M4 total DNA extracts. The % above the dot plots indicate the proportion of insects that showed colonization by microscopy observation (n=10). The qPCR measurements for each individual insect are indicated by green dots and the mean per mutant is indicated by a horizontal black line. (B) Co-infections of *R. pedestris* with an equal mix of RFP-labelled *C. insecticola* WT and indicated GFP-labelled WT or mutant strains. Relative abundance of the two strains in the M4 midgut regions at 5 dpi was determined by flow cytometry on dissected intestines. The competition index expresses for all samples the ratio of RFP-labelled WT to the GFP-labelled WT or indicated mutant in the M4 midgut region, corrected by the ratio of the inoculum, which was in all cases close to 1. Each dot represents the competition index in an individual and the mean per mutant is indicated by a horizontal black line (n=10). (C) Co-infections of *R. pedestris* with a 1:1 mix of GFP-labelled *P. fungorum* and indicated mScarlett-l-labelled WT or mutant *C. insecticola*. Relative abundance of the two strains in the M4 midgut regions at 5 dpi was determined by flow cytometry on dissected intestines. The competition index expresses for all samples the ratio of *P. fungorum* to the indicated mutant, corrected by the ratio of the inoculum, which was in all cases close to 1. Each dot represents the competition index in an individual and the mean per mutant is indicated by a horizontal black line (n=10). In all panels, different letters indicate statistically significant differences ($P < 0.05$). Statistical significance was analyzed by Kruskal–Wallis test, Dunn *post hoc* test and Benjamini-Hochberg correction.

1



2

3

4

5

6

7

Supporting Information for

8

Hundreds of antimicrobial peptides create a selective barrier for insect gut symbionts

9

10

11

12

Joy Lachat, Gaëlle Lextrait, Romain Jouan, Amira Boukherissa, Aya Yokota, Seonghan Jang, Kota Ishigami, Ryo Futahashi, Raynald Cossard, Delphine Naquin, Vlad Costache, Luis Augusto, Pierre Tissières, Emanuele G. Biondi, Benoît Alunni, Tatiana Timchenko, Tsubasa Ohbayashi, Yoshitomo Kikuchi, Peter Mergaert

15

16

17

Peter Mergaert and Yoshitomo Kikuchi

18

Email: peter.mergaert@i2bc.paris-saclay.fr; y-kikuchi@aist.go.jp

19

20

21

This PDF file includes:

22

23

Supporting text

24

SI References

25

Figures S1 to S10

26

27

Other supporting materials for this manuscript include the following:

28

29

Table S1

30

Datasets S1 and S2

31

32 **Supplementary text**
33 **Supporting Materials and Methods**

34
35 **Bacterial strains and growth conditions**

36 *Caballeronia insecticola* and *Paraburkholderia fungorum* strains were routinely cultured in Yeast-
37 Glucose medium (YG: 5 g/L yeast extract; 1 g/L NaCl; 4 g/L glucose) at 28°C. *C. insecticola* RPE75
38 is a spontaneous rifampicin-resistant derivative of the wild-type strain RPE64 (1). *Sinorhizobium*
39 *meliloti* Sm1021 and *Bacillus subtilis* Bs168 were grown in YEB medium (0.5 % beef extract, 0.1 %
40 yeast extract, 0.5 % peptone, 0.5 % sucrose, 0.04 % MgSO₄·7H₂O, pH 7.5) at 28°C or LB medium
41 (5 g/L yeast extract, 10 g/L tryptone, 5 g/L NaCl) at 37°C, respectively.

42 For standard molecular microbiology purposes, *Escherichia coli* strains DH5 α , WM3064,
43 MFDpir, S17-1 λ pir and their derivatives were grown in LB medium at 37°C. Growth of the Δ dapA
44 MFDpir and WM3064 strains, which are auxotroph for diaminopimelic acid (DAP) synthesis,
45 required the supplement of 300 μ g/mL DAP to the medium. When appropriate, antibiotics were
46 added to the medium in the following concentrations: 50 μ g/mL kanamycin (Km) for *E. coli* and
47 30 μ g/mL for *C. insecticola*; 25 μ g/mL chloramphenicol (Cm); 30 μ g/mL rifampicin (Rif); 100 μ g/mL
48 ampicillin (Amp); 500 μ g/mL streptomycin (Sm) for *S. meliloti*. For solid agar plates, the media were
49 supplemented with 1.5 % agar. For motility assays of *C. insecticola* derivatives, 5 μ L of exponential
50 phase cultures at OD_{600nm} \approx 0.5 were injected in the centre of YG soft agar plates (0.3 % agar).
51 Swimming motility of the tested strains was observed by the spreading of the bacterial growth from
52 the inoculation point and was quantified by measuring the size of the formed halo using pictures of
53 the plates, taken at regular time points after inoculation.

54 For growth inhibition assays, *C. insecticola* RPE75, *P. fungorum* JCM21562, *S. meliloti*
55 Sm1021, *B. subtilis* Bs168 were grown in the minimal medium MM9 (40 mM MOPS; 20 mM KOH;
56 19.2 mM NH₄Cl; 8.76 mM NaCl; 2 mM KH₂PO₄; 1 mM MgSO₄·7H₂O; 0.25 mM CaCl₂·2H₂O;
57 1 μ g/mL biotin; 42 nM CoCl₂; 38 μ M FeCl₃; 10 mM glucose) while *C. insecticola* RPE75 and derived
58 mutants were grown in the minimal medium MM (1 g/L KH₂PO₄; 2 g/L K₂HPO₄; 1 g/L (NH₄)₂SO₄;
59 0.2 g/L NaCl; 0.1 g/L MgSO₄·7H₂O; 2.46 mg/L FeSO₄·7H₂O; 3.31 mg/L Na₂EDTA·2H₂O; 50 mg/L
60 CaCl₂·2H₂O; 2 g/L glucose).

61

62 ***Riptortus pedestris* rearing**

63 A single pair of adult *Riptortus pedestris* bean bugs was originally collected from a single spot in a
64 soybean field in Tsukuba, Japan, in 2007 to establish the TKS1 inbred rearing population. For long-
65 term maintenance of the insect line in the laboratory, insects are reared in plastic boxes at 25°C
66 under a long-day regimen (16 h light, 8 h dark) and fed with soybean seeds and distilled water
67 containing 0.05 % ascorbic acid (DWA). Eggs are deposited on natural-fibre ropes present in the
68 cages. Ropes carrying eggs are regularly collected and placed in a separate container. New-born
69 hatchlings in these containers are collected daily and reared in fresh containers. At the second
70 instar, a suspension of wild-type *C. insecticola* RPE64 cells at 10⁷ cfu/mL is added to the drinking
71 water to keep the animals in the symbiotic state, which is optimal for reproduction (2).

72

73 ***Riptortus pedestris* colonization assays**

74 Freshly hatched first instars were transferred into sterile Petri dishes. Two days after hatching, at
75 the second larval stage, water was removed to make insects thirsty and facilitate the subsequent
76 ingestion of administered bacteria. After overnight water-starvation, a bacterial suspension of the
77 tested bacteria, grown in exponential phase at OD_{600nm} \approx 0.5-0.7 and adjusted to 10⁷ cfu/mL in sterile

78 distilled water, was provided for infection of the second instar nymphs insects. For co-inoculation
79 experiments, two bacterial strains were mixed together at a one to one ratio, each adjusted before
80 to 10^7 cfu/mL. At three and five days post inoculation (dpi), insects, at the stage of the end of the
81 second instar nymphs or the third instar, respectively, were dissected. Dissections were performed
82 under binocular microscope in sterile PBS (137 mM NaCl, 8.1 mM Na_2HPO_4 , 2.7 mM KCl, and
83 1.5 mM KH_2PO_4 , pH 7.5) containing 0.01 % of Tween 20. The M4 region of the midgut was
84 collected using fine forceps and assembled on glass slides for microscopy observations (Nikon
85 Eclipse 80i). The colonization rate of the inoculated insects was estimated by detection of the
86 presence or absence of fluorescent signal derived from colonizing bacteria expressing GFP or
87 mScarlett-I fluorescent proteins. Merged fluorescence pictures were obtained with GIMP version
88 2.10.32. For quantification, M4 region samples were homogenized in PBS solution and bacteria in
89 suspension were counted by plating on selective YG medium, by qPCR or by flow cytometry using
90 the fluorescent tags to determine the relative abundance of the two inoculated strains. For each
91 condition, 10 insects were tested within an experiment and each experiment was performed at least
92 twice.

93 To quantify the symbionts in the M4 after infection of the insects with single strains, DNA was
94 extracted from the dissected M4 by the QIAamp DNA Mini Kit (Qiagen), and real-time qPCR was
95 performed using primers *opcP-F* and *opcP-R* (SI Appendix, Table S1), which target the
96 *C. insecticola opcP* gene, the Fast Start Essential DNA Green Master qPCR Kit (Roche), and a
97 LightCycler 96 instrument (Roche). The number of gut symbionts was calculated based on a
98 calibration curve for the *opcP* gene consisting of ten-fold dilutions of 10^7 to 10 copies of the PCR
99 DNA fragment. Data were analysed with the instrument software and transferred to Microsoft Excel.
100 Statistical analysis, using a Kruskal-Wallis test, Dunn *post hoc* test and Benjamini-Hochberg
101 correction, was performed with R software (3).

102 For co-inoculation experiments, the relative abundance of the two tested strains in competition
103 was quantified by flow cytometry using a CytoFlex S instrument operated by CytExpert 2.4.0.28
104 software (Beckman Coulter). Gating by the forward-scatter (FSC) and side scatter (SSC) dot plot
105 permitted to collect signals specifically derived from bacteria. Doublets were discarded using the
106 SSC_Area-SSC_Height dot plot. GFP fluorescence was excited by a 488-nm laser and collected
107 through a 525/40 nm band pass filter; RFP fluorescence was excited by a 561-nm laser and
108 collected through a 610/20 nm band pass filter. Data acquisition for a total of 50,000-100,000
109 bacteria was performed for each sample. Thresholds for considering positive events for GFP and
110 RFP were determined using non-fluorescent control bacteria. Data was treated by CytExpert and
111 Microsoft Excel software. Statistical analysis, using a Kruskal-Wallis test, Dunn *post hoc* test and
112 Benjamini-Hochberg correction, was performed with R software.

113

114 **Transcriptome analysis**

115 For transcriptome analysis of the midgut, insects were reared and infected as above. For samples
116 of symbiotic insects (Sym), first day second instars were infected with *C. insecticola* RPE75. For
117 uninfected insect samples (Apo), the animals were reared continuously in the absence of bacteria.
118 Insects were harvested in the second instar at 1 dpi, 2 dpi and 3 dpi and in the fifth instar at 12 dpi.
119 Apo insects were harvested at the same time points. Insects were dissected and the M3, M4B and
120 M4 midgut regions were harvested in RNAiso Plus (Takara Bio). This resulted in 24 samples (Apo
121 and Sym insects \times 4 timepoints \times 3 midgut regions). For each condition, samples of 24 insects in
122 second instar and 12 insects in the fifth instar were pooled. Total RNA was extracted from the
123 pooled samples with the RNeasy mini kit (Qiagen). mRNA purification and fragmentation and the
124 preparation of cDNA libraries were conducted using TruSeq RNA Library Prep Kit v2 kit (Illumina).
125 The cDNA libraries were sequenced by Illumina HiSeq-2000. The obtained RNA-seq data were
126 analyzed by the bcl2fastq software (ver. 2-2.18.12, Illumina, Inc., San Diego, CA, USA) and FastQC
127 (ver. 0.11.5) (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) to keep only high-quality

128 data for *de novo* assembly and differential expression analysis. Raw read errors were corrected
129 using Rcorrector (4) (K-mer-based method), followed by removing read pairs where at least one
130 read had an unfixable error. Trim Galore
131 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) was used to remove adaptor
132 sequences, short reads, and low-quality reads. After the preprocessing, between 4 and 18 million
133 paired-end reads remained per sample with a total of 412 million paired-end reads. The RNA-seq
134 sequencing data were deposited in the Sequence Read Archive (SRA), BioProject accession no.
135 PRJNA1006624.

136 The preprocessed reads were provided to Trinity software (5) for *de novo* assembly. Three
137 methods were used to assess the quality of the Trinity assembly. First, assembly statistics were
138 calculated using TrinityStats.pl script from Trinity software: the assembly contained 135,346 contigs
139 (transcripts), N50 was 2541 bases and the total number of assembled bases in the contigs was
140 156,984,167. Those 135,346 Trinity transcripts represent splice variants of 83,840 Trinity single
141 genes. Second, the alignment rate of the preprocessed reads to the Trinity assembly was evaluated
142 using Bowtie 2 (6), which resulted in an overall alignment rate of 98,61%. Third, the completeness
143 of the *de novo* assembled transcriptome was assessed using BUSCO (7) with the OrthoDB v10
144 'Insecta' and 'Hemiptera' reference databases which resulted in >97% and >94% of the Insecta and
145 Hemiptera BUSCO genes identified as complete, respectively. In addition, transcripts encoding all
146 the 97 reference CCR (Crypt-specific Cysteine Rich) peptides were identified in the *de novo*
147 transcriptome using blast (8). The *de novo* assembled transcriptome was deposited in
148 Transcriptome Shotgun Assembly (TSA), BioProject accession no. PRJNA1006624.

149 Coding regions in the Trinity *de novo* assembly were predicted using the TransDecoder tool
150 (<https://github.com/TransDecoder/TransDecoder>), resulting in 122,466 protein-coding transcripts
151 representing 74,194 non-redundant protein-coding genes. Among the remaining 12,880
152 transcripts, 12,867 were predicted to be non-coding RNAs by the coding potential calculator
153 software CPC2 (9). Then, the CCR peptides were predicted from this assembled and translated
154 sequences using a Hidden Markov Model homology search against the 97 reference CCR peptides
155 of *R. pedestris* (10) with HMMER software (11). The jackhammer application was used with the
156 CCR reference sequences as input profile. In an iterative process by jackhammer, newly found
157 CCR sequences are added to the input data set and the search was repeated. The number of
158 iterations was 100 and additional iterations did not yield new sequences. Sequences with E-values
159 of less than 10^{-5} were selected. The final output sequences were further filtered using custom
160 scripts to remove those without a signal peptide and longer than the longest CCR peptide in the
161 reference CCR database used for the prediction. The process yielded a total of 310 transcripts,
162 including the reference data set, that were annotated as encoding CCR peptides. These transcripts
163 encode 217 different CCR peptides that represent 126 predicted genes.

164 The complete set of CCR peptides were aligned using Multalin
165 (<http://multalin.toulouse.inra.fr/multalin/>) (12) and the alignment obtained was adjusted manually to
166 keep only the well-aligned ones and to highlight the conserved cysteine residues. The adjusted
167 alignment was used in WebLogo (13) to generate a graphical representation of the consensus
168 sequence from the multiple sequence alignment.

169 Transcript and gene-expression levels of the transcripts encoding CCR peptides were estimated
170 for each sample using RSEM software (14) based on alignment and transcript abundance
171 estimation. The DEGs (Differentially Expressed Genes) were identified using edgeR (15) with FDR
172 < 0.05 and $\log_2(\text{FC}) \geq 1$ resulting in 305 (among 310) DEGs. The clustered heatmap, generated
173 by MEV software ver. 4.8.0, of CCR genes against samples was based on the normalized FPKM
174 (Fragments Per Kilobase of transcript per Million mapped reads) expression values matrix. FPKM
175 values were transformed according to $\text{Value} = [(\text{Value}) - \text{Mean}(\text{Row})] / [\text{Standard deviation}(\text{Row})]$
176 and clustered with hierarchical clustering using Pearson Correlation distance metric and average
177 linkage (16). The expression strength of CCR genes relative to the overall transcriptome was

178 estimated on the basis of the highest FPKM value among the 24 experimental conditions for each
179 transcript and the ranking of transcripts according to these highest FPKM values.

180 The following antimicrobial peptide *in silico* predictor tools were used to predict the AMP activity
181 potential of the 310 CCR peptide sequences: Collection of Anti-Microbial Peptides (CAMP_{R4}) (17),
182 using the Random Forest (RF), Support Vector Machine (SVM) and Artificial Neural Network (ANN)
183 algorithms; AMPpredictor in dbAMP (18); iAMPpred (19); Antimicrobial Peptide Scanner v2
184 (AMPscanner) (20); Antimicrobial Peptide Prediction at AxPEP (21). A consensus AMP prediction
185 was considered when 6 out of the 7 predictions were positive.

186

187 **Whole-mount *in situ* hybridization**

188 Whole-mount *in situ* hybridization on the *R. pedestris* midgut was performed with digoxigenin (DIG)-
189 labeled *CCR043* cRNA probe, using a method essentially as described (22). In brief, a DNA
190 fragment of the *CCR043* gene was amplified from M4 cDNA by PCR (SI Appendix, Table S1).
191 The amplified cDNA fragment was cloned into the pT7Blue T-vector (Novagen). Gene-specific
192 antisense or sense DIG-labeled cRNA probes were obtained by *in vitro* transcription with T7
193 polymerase using the Roche DIG RNA labeling kit (Roche Diagnostics) following the instructions
194 of the supplier. Midguts of symbiotic and aposymbiotic insects were obtained as above and the
195 tissues were fixed in 4 % paraformaldehyde, permeabilized with Proteinase K treatment and re-
196 fixed. DIG-labeled cRNA probe hybridization was performed in hybridization buffer (50 %
197 formamide, 5x SSC (0.75 M NaCl and 0.75 M Na-citrate, pH 7.4), 5x Denhardt's solution (0.2 %
198 bovine serum albumin, 0.2 % Ficoll, and 0.2 % polyvinylpyrrolidone), 25 mg/mL sonicated salmon
199 sperm DNA, and 0.1 % Tween 20) at 61°C for 16 h. Detection of mRNA was performed with anti-
200 DIG-alkaline-phosphatase-conjugated antibody (Roche Diagnostics) and the alkaline phosphatase
201 substrate 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and 4-nitroblue tetrazolium chloride (NBT)
202 (Roche Diagnostics). After washing, the gut samples were mounted between slide and cover slip
203 and observed by microscopy.

204

205 **Polymyxin B Tn-seq screening**

206 An aliquot of an available *Himar1* transposon mutant Tn-seq library (23) was diluted to obtain a
207 suspension of 2×10^8 cfu/mL. For each tested condition, 100 μ L of this dilution was inoculated into
208 20 mL of MM, supplemented with Rif and Km, to obtain an initial inoculum of 10^6 cfu/mL
209 ($OD_{600nm} \approx 0.0015$). In the control growth condition, no Polymyxin B (PMB) was added. For the test
210 conditions, the minimal media was supplemented with three different concentrations of PMB:
211 1.25 μ M, 5 μ M and 10 μ M. These concentrations were chosen after a sensitivity assay and
212 represent respectively half MIC, 1/4 MIC and 1/8 MIC of wild type *C. insecticola*. The resulting
213 cultures were incubated at 28°C, with shaking at 180 rpm. When the cultures reached an
214 $OD_{600nm} \approx 1$, corresponding to approximately 9 to 10 generations of multiplication, bacteria were
215 collected by centrifugation at 3300 rcf for 20 min at 4°C and the pellets were stored at -20°C until
216 DNA extraction. Each condition was performed in triplicates.

217 Genomic DNA extraction and processing to obtain Tn-seq sequencing libraries for Illumina
218 sequencing was performed as described (23) using materials provided in SI Appendix, Table S1.
219 The Tn-seq sequencing library samples were sequenced by an Illumina NextSeq 500 instrument
220 with 2 x 75 paired-end run at the I2BC sequencing platform (CNRS Gif-sur-Yvette, France). The
221 generated data were demultiplexed, trimmed, filtered and mapped as described (23). Tn-seq
222 sequencing data were deposited in SRA, BioProject accession no. PRJNA890438.

223 Tn-seq sequencing data was handled by TRANSIT Version 3.2.0 (24) using software
224 parameters as before and implementing statistics of the software (23). Each experimental condition
225 with PMB treatment was compared to the control condition without PMB as a reference. The
226 Integrative Genomics Viewer (IGV) software, an interactive tool for the visual exploration of
227 genomic data, was used to visualize the number of insertions per insertion sites in specific genes
228 of interest (25).

229

230 **Construction of *Caballeronia insecticola* RPE75 deletion mutants, fluorescent protein** 231 **tagged strains and complementation strains**

232 To create deletion mutants, regions of 400 bp flanking a gene of interest (Up and Down regions)
233 were obtained by gene synthesis (GenScript) as a fused 800 bp fragment (SI Appendix, Table S1)
234 and cloned into the *SmaI/HindIII* or *EcoRI/HindIII* sites of the pK18mobsacB vector. The constructs
235 were introduced in *E. coli* WM3064. Plasmids were introduced into *C. insecticola* RPE75 by bi-
236 parental mating, resulting in first recombinant clones with the plasmids integrated in the genome.
237 PCR with gene specific primers (SI Appendix, Table S1) and subsequent sequencing confirmed
238 insertions at the appropriate chromosomal positions. To obtain second recombinants in which the
239 gene of interest has been deleted, bacterial cultures of first recombinants were plated on YG
240 containing 10 % sucrose and Rif to counter-select the *sacB* gene located on the pK18mobsacB
241 plasmid. Candidate deletion mutants were verified by colony PCR (SI Appendix, Table S1).

242 GFP- or mScarlett-I-tagged strains of *C. insecticola* wild type and mutants were created by
243 introducing a Tn7-GFP or Tn7-Scarlet transposon using tri-parental mating with the Tn7-GFP donor
244 strain WM3064.pURR25 or the Tn7-Scarlet donor strain S17-1 λ pir.pMRE-Tn7-135, the helper
245 strain WM3064.pUX-BF13 and *C. insecticola* wild type or derivatives as acceptor strains.

246 For the construction of complemented strains, the plasmid pME6000 was digested by *HindIII*
247 and the PCR-amplified gene of interest was inserted by Gibson cloning (NEB). For this, DNA
248 fragments containing about 350-400 bp of upstream non-coding sequence, the open reading frame
249 and 300-350 bp of downstream non-coding sequence of a corresponding gene were amplified with
250 gene-specific primers (SI Appendix, Table S1). PCR products were used as template for second
251 PCR amplifications with Gibson-compatible primers (SI Appendix, Table S1). The constructed
252 plasmids, cloned in DH5 α , were introduced in *E. coli* WM3064 and transferred to the corresponding
253 mutant by bi-parental mating.

254

255 **Antimicrobial peptide activity assays**

256 Precultures of tested strains were grown in MM9 or MM medium. Overnight grown cultures were
257 diluted to an OD_{600nm}=0.3 in fresh medium and grown until they reached OD_{600nm}≈1. The cells were
258 pelleted by centrifugation, resuspended in fresh medium and diluted to OD_{600nm}=0.05. For testing
259 crypt-colonizing *C. insecticola*, 100 insects at 5 dpi were dissected, the M4 region collected and
260 the bacteria extracted as described above. The obtained suspensions were diluted to OD_{600nm}=0.05
261 in MM. These cell suspensions were dispatched in a 96-well plate, one column per tested strain
262 and at 150 μ L per well, except for the wells of the first row, which contained 300 μ L. Peptides used
263 in this study are listed in SI Appendix, Table S1. The CCR peptides were selected on the basis of
264 the following criteria: 1) consistent prediction of AMP activity; 2) an independently confirmed
265 sequence of the assembled transcript sequence by cloned cDNA Sanger sequencing (10); 3)
266 diversity of expression patterns, including peptides expressed in apo and/or sym insects and in the
267 M3, M4B and/or M4, with high or medium expression levels; 4) favouring smaller peptides to
268 increase feasibility of successful peptide synthesis; 5) successful synthesis by commercial
269 suppliers (synthesis attempts for several peptides failed). AMPs, dissolved in water, were added to

270 the first row to reach a maximal tested final concentration. Two-fold serial dilutions in the
271 subsequent rows were obtained by the serial transfer of 150 μL to the next row and mixing by
272 pipetting up and down. No peptide was added to the last row of the 96-well plate. This setup
273 resulted in the testing of seven peptide concentrations forming a two-fold concentration series in
274 the range of 120 μM to 0.5 μM and a control sample without peptide. The 96-well plates were
275 incubated in a SPECTROstar Nano plate incubator (BMG Labtech). The growth of the cultures in
276 the wells was monitored by measuring the $\text{OD}_{600\text{nm}}$ and data points were collected every hour for
277 48 h. Plates were incubated at 28°C with double orbital shaking at 200 rpm. Data and growth curves
278 were analyzed using Microsoft Excel. For the comparison of *C. insecticola*, *P. fungorum*, *S. meliloti*
279 and *B. subtilis*, the minimal peptide concentration of growth inhibition was determined. For the
280 comparison of *C. insecticola* wild type and derived mutants, the minimal concentration was
281 determined at which growth was diminished compared to the untreated control. The assays were
282 performed in biological triplicates for all peptides.

283 To determine bacterial survival after peptide treatment by cfu counting, *S. meliloti* was grown in
284 MM9 until $\text{OD}_{600\text{nm}} < 1$. Bacteria were resuspended in fresh medium to $\text{OD}_{600\text{nm}} = 0.01$. CCR1659
285 peptide, Proteinase K-inactivated CCR1659 or PMB were added to the suspension to a final
286 concentration of 25 μM . To inactivate peptide CCR1659, a peptide stock at 500 μM was treated
287 with 0.1 mg/mL Proteinase K for 2h at 37°C. The treatments were made in the absence or presence
288 of 5 mM CaCl_2 . Suspensions were incubated at 28°C and 10 μL samples were withdrawn at various
289 time-intervals. 5 μL of ten-fold dilution series of these samples were spotted on plates with YEB
290 medium for cfu counting. The assays were performed in triplicates.

291

292 Membrane stability measurements

293 For evaluation of the membrane permeabilization of *S. meliloti* in response to CCR1659 or PMB,
294 bacteria were grown in MM9 medium until $\text{OD}_{600\text{nm}} < 1$. Bacteria were resuspended to $\text{OD}_{600\text{nm}} = 0.1$
295 in fresh MM9, containing either 10 μM 1-N-phenyl-naphthylamine (NPN) (26, 27) or 1 $\mu\text{g}/\text{mL}$
296 Propidium Iodide (PI). Suspensions were dispatched per 100 μL in a black 96-well plate with
297 transparent bottom (Falcon). CCR1659 peptide or PMB were added to the suspensions to a final
298 concentration of 10 μM ; control treatments were without peptide. All treatments were made in the
299 absence or presence of 5 mM CaCl_2 . Fluorescence was measured at 28°C every 2 min for a total
300 of 100 min incubation in an infinite M1000 PRO fluorescence plate reader (Tecan), operated with
301 Tecan i-control, version 1.11.1.0 software. NPN fluorescence was excited at 340 nm and emission
302 detection was at 420 nm. PI fluorescence was excited at 535 nm and emission detection was at
303 617 nm. Data was exported to Microsoft Excel for analysis.

304 Membrane stability properties in *C. insecticola* wild type and mutants were assessed by NPN
305 fluorescence and sensitivity to detergents SDS, Triton X-100 and CTAB. The strains were grown
306 in MM medium until $\text{OD}_{600\text{nm}} < 1$. For NPN fluorescence measurements, bacteria were resuspended
307 to $\text{OD}_{600\text{nm}} = 0.5$ in fresh MM, containing 10 μM NPN. Fluorescence was measured as above after
308 10 min incubation in the NPN-containing medium. Detergent sensitivity assays were performed
309 using maximum detergent concentrations of 0.02 % for SDS, 0.05 % for Triton X-100 or 0.001 %
310 for CTAB in the microplate two-fold-dilution setup as above for the AMP sensitivity assays.

311

312 Peptide binding to bacterial cells

313 Precultures of *S. meliloti* or *C. insecticola* strains were grown in MM9 or MM medium, respectively.
314 Overnight grown cultures were diluted to an $\text{OD}_{600\text{nm}} = 0.3$ in fresh medium and grown until they
315 reached $\text{OD}_{600\text{nm}} \approx 1$. The cells were pelleted by centrifugation, resuspended in fresh medium and
316 diluted to $\text{OD}_{600\text{nm}} = 0.1$. Crypt-colonizing bacteria were obtained by dissecting 100 insects at 5 dpi,

317 collecting the M4 region and extracting the bacteria by homogenizing the M4 tissues as described
318 above. The obtained suspensions were diluted to $OD_{600nm}=0.1$ in MM. FITC-CCR1659 or Poly-L-
319 lysine-FITC (SI Appendix, Table S1) were added to a final concentration of respectively 1.5 μ M or
320 10 μ g/mL for *S. meliloti* and of 7.5 μ M or 50 μ g/mL for *C. insecticola* strains. Labelling of cells was
321 analysed immediately after addition of the peptides by microscopy and flow cytometry. Labelled
322 bacteria were spotted on glass slides covered with an agar pad for fluorescence microscopy
323 observations (Nikon Eclipse 80i). Labelling was quantified by flow cytometry using a CytoFlex S
324 instrument operated by CytExpert 2.4.0.28 software (Beckman Coulter). Gating by the forward-
325 scatter (FSC) and side scatter (SSC) dot plot permitted to collect signals specifically derived from
326 bacteria. Doublets were discarded using the SSC_Area-SSC_Height dot plot. FITC fluorescence
327 was excited by a 488 nm laser and collected through a 525/40 nm band pass filter. Data acquisition
328 for a total of 20.000 bacteria was performed for each sample. Control samples without addition of
329 FITC-labelled peptide were used to determine the background signal of the bacteria. Data was
330 treated by CytExpert and Microsoft Excel software.

331

332 Scanning electron microscopy

333 *S. meliloti* was grown in MM9 and *C. insecticola* wild type and mutants in MM. Bacteria in
334 exponential-phase ($OD_{600nm}<1$) were resuspended in fresh medium to $OD_{600nm}=0.2$. One mL of
335 *S. meliloti* suspensions were treated with 1.5 μ M CCR1659 or 3.6 μ M PMB or without peptide. One
336 mL of *C. insecticola* suspensions were treated with 1.5 M CCR1659 or without peptide. After 1 h of
337 incubation at room temperature, bacteria were sedimented by soft centrifugation (2 min, 2000 g),
338 the bulk of the supernatant was removed and the remaining 50 μ L was directly fixed in 2 mL 2 %
339 glutaraldehyde buffered with sodium cacodylate 0.2 M, 2 h at room temperature then overnight at
340 4°C on glass slides. Samples attached to the glass slides were rinsed 10 min in 0.2 M sodium
341 cacodylate buffer, dehydrated in successive baths of ethanol (50, 70, 90, 100, and anhydrous
342 100 %), and then dried using a Leica EM300 critical point apparatus with slow 20 exchange cycles,
343 with a 2 min delay. Samples were mounted on aluminum stubs with adhesive carbon and coated
344 with 6 nm of Au/Pd using a Quorum SC7620, 50 Pa of Ar, 180 s of sputtering at 3.5 mA. Samples
345 were observed using the SE detector of a FEG SEM Hitachi SU5000, 2 KeV, 30 spot size, 5 mm
346 working distance (facilities located on the MIMA2 platform, INRAE, Jouy-en-Josas, France;
347 <https://doi.org/10.15454/1.5572348210007727E12>). Scanning electron microscopy (SEM) images
348 were analyzed with FIJI software and calibrated using the printed scale bar on the image during
349 the acquisitions.

350

351 LPS and lipid A extraction and analysis

352 Bacteria were grown in 100 mL YG at 28°C until $OD_{600nm}\approx 2$ and LPSs were isolated by the
353 phenol/water method of Westphal and Jann (28). Briefly, the wet pellet of bacteria was stirred in
354 30 mL 50% aqueous phenol at 65°C for 15 min, insoluble material was removed from the cooled
355 water phase by centrifugation and the clear extract was dialyzed under running tap water until free
356 of phenol, then dialyzed against distilled water. The samples were subsequently lyophilized.

357 LPS sample analysis by SDS-PAGE was performed on 15 % polyacrylamide separation gels
358 layered with 4 % polyacrylamide stacking gel, using standard electrophoresis buffers and migration
359 conditions (29). About 0.5 μ g LPS per sample was loaded. Gels were stained with the silver nitrate
360 method essentially as described before (30).

361 Lipid A was prepared by the triethylamine-citrate method (31). Briefly, the LPS samples were
362 suspended at a concentration of 10 μ g/ μ l in a 0.01 M triethylamine-citrate solution (1:1 molar ratio,
363 pH 3.6) and heated for 1 h at 100°C. The samples were then lyophilized and suspended in

364 methanol. After centrifugation (7000× g for 10 min at 4°C), lipid A fractions were extracted with a
365 mixture of chloroform: methanol: water (3:1.5:0.25, v:v:v) at a concentration of 10 µg/µL.

366 The molecular species present in this preparation were analyzed using an AXIMA performance
367 (Shimadzu Biotech) matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass
368 spectrometer. A suspension of lipid A (1 µg/µL) in chloroform: methanol: water (3:1.5:0.25, v:v:v),
369 1 µL was deposited on the target mixed with 1 µL of a gentisic acid (2,5-dihydroxybenzoic acid)
370 matrix (DHB from Fluka) suspended at 10 µg/µL in the same solvent, and dried. Analyte ions were
371 desorbed from the matrix with pulses from a 337 nm nitrogen laser. Spectra were obtained in the
372 negative-ion mode at 20 kV, with the linear detector. Mass calibration was performed with a *peptide*
373 *mass standards* kit (AB SCIEX) or with a purified and structurally characterized LPS sample from
374 *Bordetella pertussis*.

375

376 **SI References**

- 377 1. H. Itoh, S. Jang, K. Takeshita, T. Ohbayashi, N. Ohnishi, X. Y. Meng, Y. Mitani, Y. Kikuchi,
378 Host-symbiont specificity determined by microbe-microbe competition in an insect gut.
379 *Proceedings of the National Academy of Sciences USA* 116, 22673-22682 (2019). doi:
380 10.1073/pnas.1912397116.
- 381 2. Y. Kikuchi, T. Hosokawa, T. Fukatsu, Insect-microbe mutualism without vertical transmission:
382 a stinkbug acquires a beneficial gut symbiont from the environment every generation. *Applied*
383 *and Environmental Microbiology* 73, 4308-4316 (2007). doi: 10.1128/AEM.00067-07.
- 384 3. R Core Team, R: A language and environment for statistical computing. R Foundation for
385 Statistical Computing, Vienna, Austria (2021). <https://www.R-project.org/>.
- 386 4. L. Song, L. Florea, Rcorrector: efficient and accurate error correction for Illumina RNA-seq
387 reads. *Gigascience* 4, 48 (2015). doi: 10.1186/s13742-015-0089-y.
- 388 5. M. G. Grabherr, B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L.
389 Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di
390 Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, A. Regev, Full-length
391 transcriptome assembly from RNA-Seq data without a reference genome. *Nature*
392 *Biotechnology* 29, 644-652 (2011). doi: 10.1038/nbt.1883.
- 393 6. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9,
394 357-359 (2012). doi: 10.1038/nmeth.1923.
- 395 7. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO:
396 assessing genome assembly and annotation completeness with single-copy orthologs.
397 *Bioinformatics* 31, 3210-3212 (2015). doi: 10.1093/bioinformatics/btv351.
- 398 8. C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T. L. Madden,
399 BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421 (2009). doi:
400 10.1186/1471-2105-10-421.
- 401 9. Y. J. Kang, D. C. Yang, L. Kong, M. Hou, Y. Q. Meng, L. Wei, G. Gao, CPC2: a fast and
402 accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids*
403 *Research* 45(W1), W12-W16 (2017). doi: 10.1093/nar/gkx428.
- 404 10. R. Futahashi, K. Tanaka, M. Tanahashi, N. Nikoh, Y. Kikuchi, B. L. Lee, T. Fukatsu, Gene
405 expression in gut symbiotic organ of stinkbug affected by extracellular bacterial symbiont.
406 *PLoS One* 8, e64557 (2013). doi: 10.1371/journal.pone.0064557.
- 407 11. L. S. Johnson, S. R. Eddy, E. Portugaly, Hidden Markov model speed heuristic and iterative
408 HMM search procedure. *BMC Bioinformatics* 11, 431 (2010). doi: 10.1186/1471-2105-11-431.
- 409 12. F. Corpet, Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research*
410 16, 10881-10890 (1988). doi: 10.1093/nar/16.22.10881.
- 411 13. G. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, WebLogo: a sequence logo generator.
412 *Genome Research* 14, 1188-1190 (2004). doi: 10.1101/gr.849004.

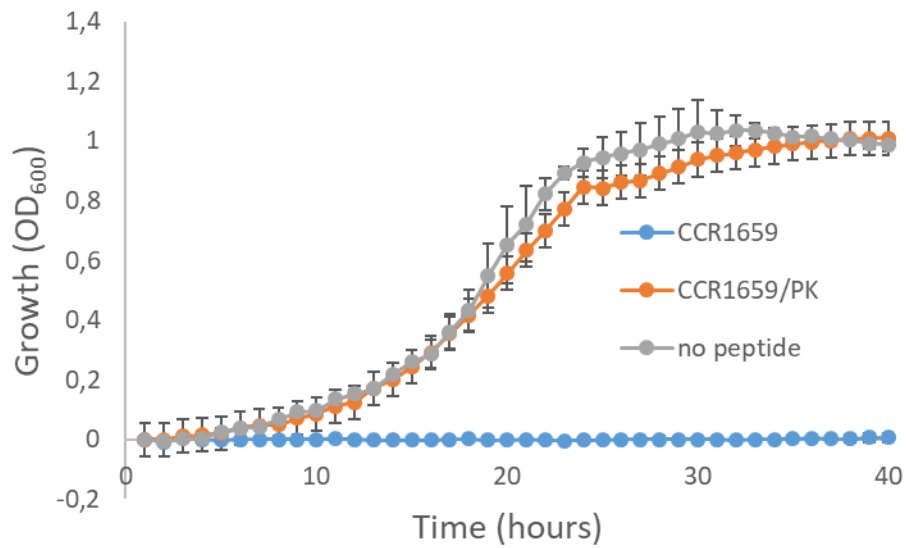
- 413 14. B. Li, C. N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or
414 without a reference genome. *BMC Bioinformatics* 12, 323 (2011). doi: 10.1186/1471-2105-12-
415 323.
- 416 15. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential
417 expression analysis of digital gene expression data. *Bioinformatics* 26, 139-140 (2010). doi:
418 10.1093/bioinformatics/btp616.
- 419 16. A. I. Saeed, V. Sharov, J. White, J. Li, W. Liang, N. Bhagabati, J. Braisted, M. Klapa, T. Carrier,
420 M. Thiagarajan, A. Stum, M. Snuffin, A. Rezantsev, D. Popov, A. Ryltsov, E. Kostukovich, I.
421 Borisovsky, Z. Liu, A. Vinsavich, V. Trush, J. Quackenbush, TM4: a free, open-source system
422 for microarray data management and analysis. *Biotechniques* 34, 374-378 (2003). doi:
423 10.2144/03342mt01.
- 424 17. U. Gawde, S. Chakraborty, F. H. Wagh, R. S. Barai, A. Khandekar, R. Indraguru, T. Shirsat,
425 S. Idicula-Thomas, CAMPR4: a database of natural and synthetic antimicrobial peptides.
426 *Nucleic Acids Research* 51(D1), D377-D383 (2023). doi: 10.1093/nar/gkac933.
- 427 18. J. H. Jhong, Y. H. Chi, W. C. Li, T. H. Lin, K. Y. Huang, T. Y. Lee, dbAMP: an integrated
428 resource for exploring antimicrobial peptides with functional activities and physicochemical
429 properties on transcriptome and proteome data. *Nucleic Acids Research* 47(D1), D285-D297
430 (2019). doi: 10.1093/nar/gky1030.
- 431 19. P. K. Meher, T. K. Sahu, V. Saini, A. R. Rao, Predicting antimicrobial peptides with improved
432 accuracy by incorporating the compositional, physico-chemical and structural features into
433 Chou's general PseAAC. *Scientific Reports* 7, 42362 (2017). doi: 10.1038/srep42362.
- 434 20. D. Veltri, U. Kamath, A. Shehu, Deep learning improves antimicrobial peptide recognition.
435 *Bioinformatics* 34, 2740-2747 (2018). doi: 10.1093/bioinformatics/bty179.
- 436 21. P. Bhadra, J. Yan, J. Li, S. Fong, S. W. I. Siu, AmPEP: Sequence-based prediction of
437 antimicrobial peptides using distribution patterns of amino acid properties and random forest.
438 *Scientific Reports* 8, 1697 (2018). doi: 10.1038/s41598-018-19752-w.
- 439 22. R. Futahashi, Whole-mount in situ hybridization of sectioned tissues of species hybrids to
440 detect cis-regulatory changes in gene expression pattern. *Methods in Molecular Biology* 772,
441 319-328 (2011). doi: 10.1007/978-1-61779-228-1_19.
- 442 23. R. Jouan, G. Lextrait, J. Lachat, A. Yokota, R. Cossard, D. Naquin, T. Timtchenko, Y. Kikuchi,
443 T. Ohbayashi, P. Mergaert, Transposon sequencing reveals the essential gene set and genes
444 enabling gut symbiosis in the insect symbiont *Caballeronia insecticola*. *ISME Communications*
445 (2024). doi: 10.1093/ismeco/ycad001.
- 446 24. M. A. DeJesus, C. Ambadipudi, R. Baker, C. Sasseti, T. R. Ioerger, TRANSIT - A software
447 tool for Himar1 TnSeq analysis. *PLOS Computational Biology* 11, e1004401 (2015). doi:
448 10.1371/journal.pcbi.1004401.
- 449 25. J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, J. P.
450 Mesirov, Integrative genomics viewer. *Nature Biotechnology* 29, 24-26 (2011). doi:
451 10.1038/nbt.1754.
- 452 26. M. Sobota, P. N. Rodilla Ramirez, A. Cambré, A. Rocker, J. Mortier, T. Gervais, T. Haas, D.
453 Cornillet, D. Chauvin, I. Hug, T. Julou, A. Aertsen, M. Diard, The expression of virulence genes
454 increases membrane permeability and sensitivity to envelope stress in *Salmonella*
455 Typhimurium. *PLoS Biology* 20, e3001608 (2022). doi: 10.1371/journal.pbio.3001608.
- 456 27. T. Tsuchido, I. Aoki, M. Takano, Interaction of the fluorescent dye 1-N-phenyl-naphthylamine
457 with *Escherichia coli* cells during heat stress and recovery from heat stress. *The Journal of*
458 *General Microbiology* 135, 1941-1947 (1989). doi: 10.1099/00221287-135-7-1941.
- 459 28. O. Westphal, K. Jann, Bacterial lipopolysaccharides extraction with phenol-water and further
460 applications of the procedure. *Methods in carbohydrate chemistry* 5, 83-91 (1965).
- 461 29. UK Laemmli. Cleavage of structural proteins during the assembly of the head of bacteriophage
462 T4. *Nature* 227, 680-685 (1970). doi: 10.1038/227680a0.
- 463 30. C. M. Tsai, C. E. Frasch. A, Sensitive silver stain for detecting lipopolysaccharides in
464 polyacrylamide gels. *Analytical Biochemistry* 119, 115-119 (1982). doi: 10.1016/0003-
465 2697(82)90673-x.

466 31. I. Chafchaoui-Moussaoui, A. Novikov, F. Bhrada, M. B. Perry, A. Filali-Maltouf, M. Caroff. A,
467 New rapid and micro-scale hydrolysis, using triethylamine citrate, for lipopolysaccharide
468 characterization by mass spectrometry. *Rapid Communications in Mass Spectrometry* 25,
469 2043-2048 (2011). doi: 10.1002/rcm.5084.
470

471 **Supplementary figures**

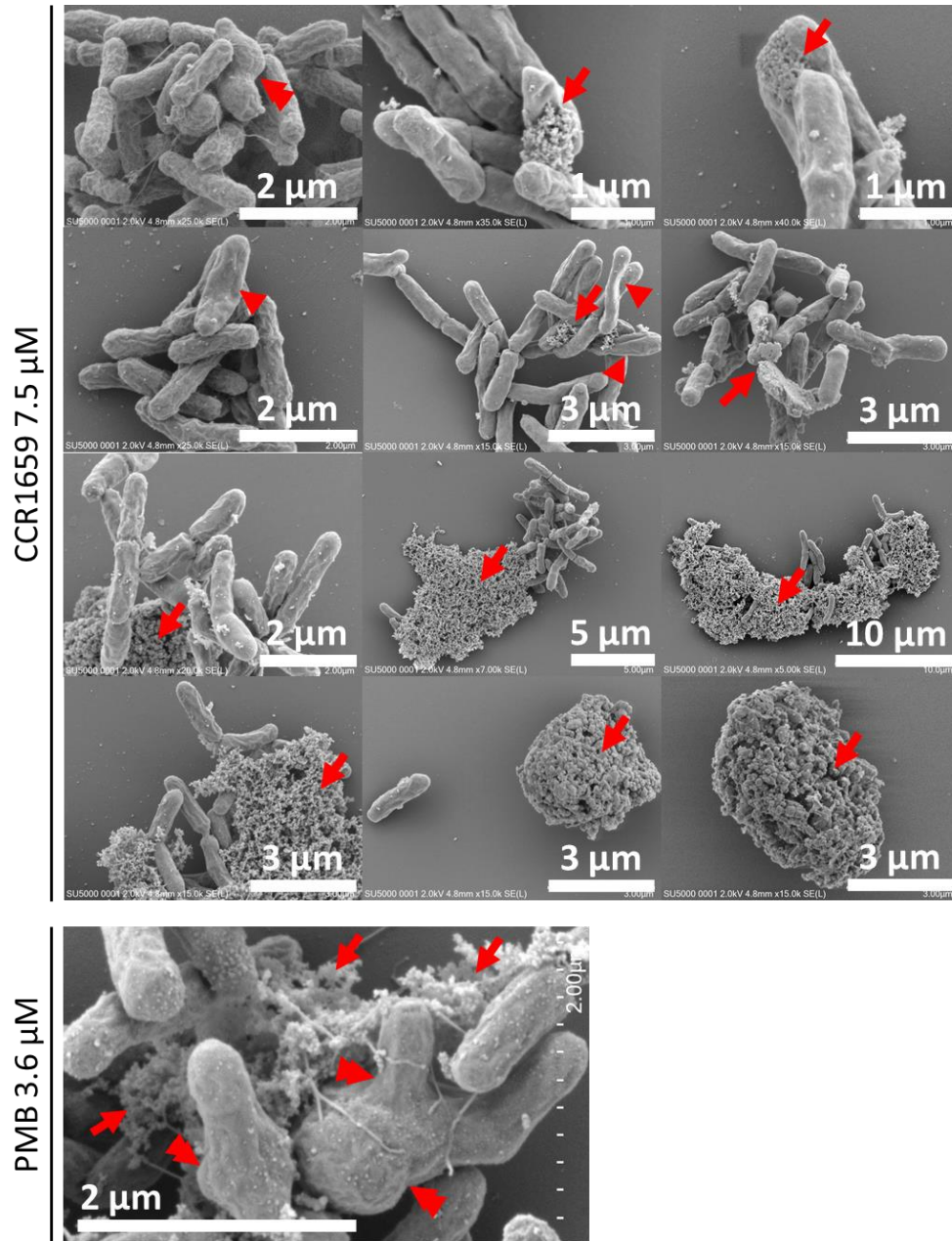


472
473 **Figure S1. Whole-mount *in situ* hybridization controls.** (A) Hybridization of a 3 dpi midgut with
474 a CCR0043 sense probe revealing the absence of signal. (B) Hybridization of an aposymbiotic
475 midgut with a CCR0043 antisense probe revealing a signal at the base of crypts. (C) Detail of crypts
476 as in panel B.



478
 479
 480
 481
 482
 483
 484
 485

Figure S2. Suppression of CCR1659 activity by Proteinase K treatment. Growth curves of *S. meliloti* control or in the presence of 18 μM CCR1659 or an equivalent amount of CCR1659 treated prior with Proteinase K (PK). A peptide stock at 500 μM was treated with 0.1 mg/mL Proteinase K for 2h at 37°C before use in the growth assay. The control samples (no peptide) were grown in the presence of a similar amount of Proteinase K (2.5 $\mu\text{g}/\text{mL}$). Error bars are standard deviations (n=3).



486

487

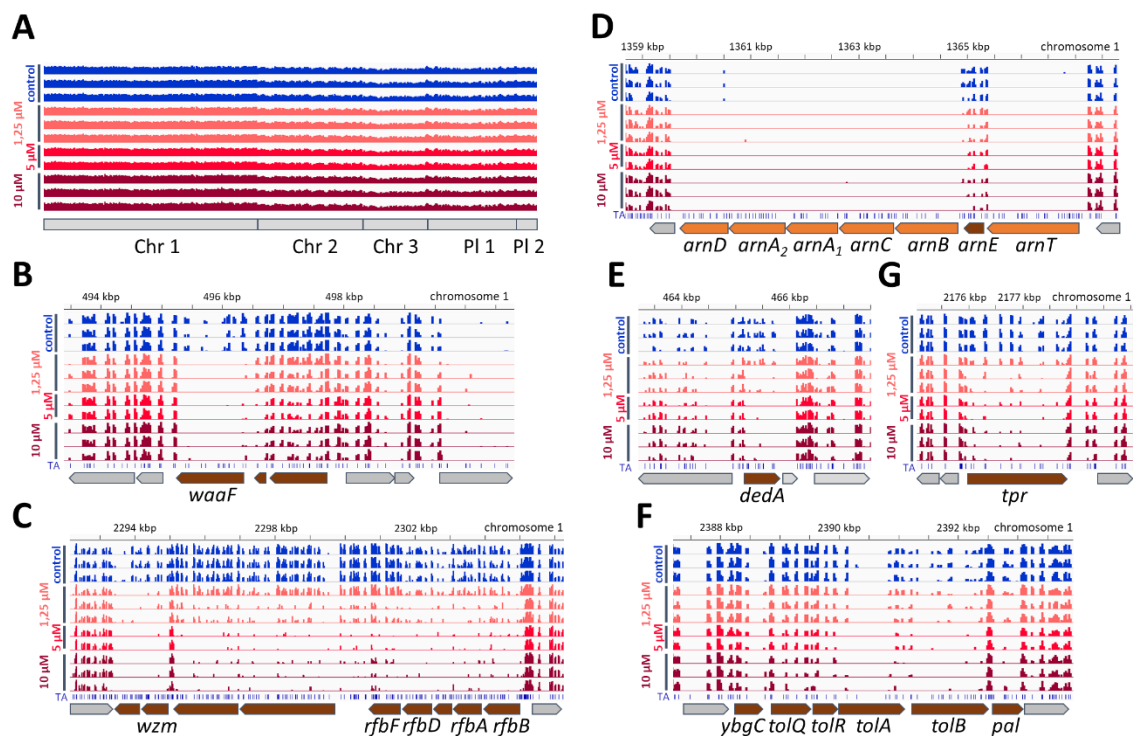
488

489

490

491

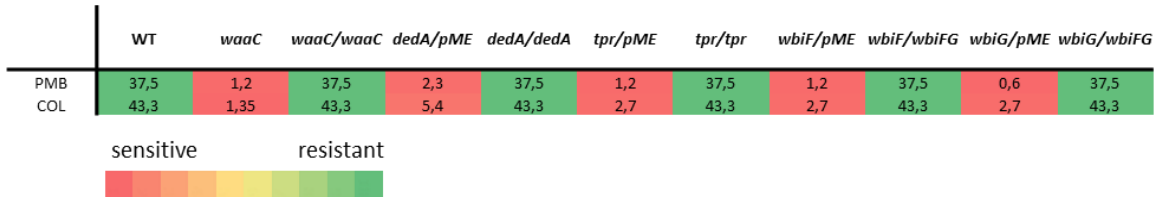
Figure S3. SEM micrographs *S. meliloti* treated with CCR1659 or PMB. Cells were treated with 7.5 μM CCR1659 or with 3.6 μM PMB. The arrows indicate cellular material released from cells. Arrowheads indicate cells with lost turgor. The double arrowheads indicate morphologically aberrant cells.



492

493 **Figure S4. IGV views of Tn-seq sequencing data of *Caballeronia insecticola* genome regions.**
 494 (A) Whole genome view showing the reproducible profiles in the control and PMB conditions. (B-
 495 F) Views of selected regions corresponding to the genes that were further characterized in this
 496 study.

497



498

499

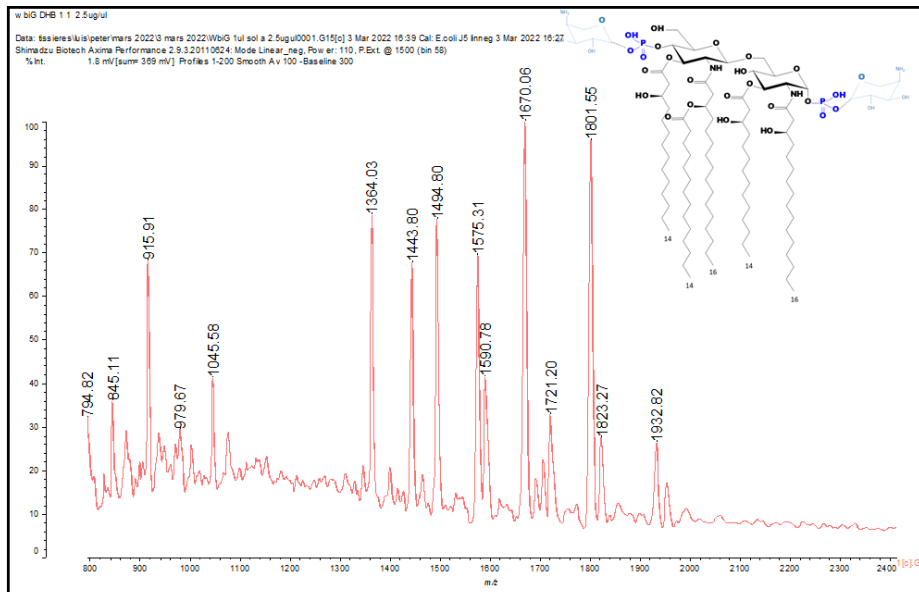
500

501

502

503

Figure S5. Complementation of PMB sensitivity. Minimal concentrations of growth inhibition of the indicated wild-type, mutant and complemented mutant strains by PMB and COL. Minimal concentrations are indicated in μM . Complementing genes were introduced on the plasmid pME6000. pME indicates the empty plasmid pME6000.



- Peak 1364.03 = 2.GlcN + 1.P + 1.C14OH + 2.C16OH + 1.C14
- Peak 1443.8 = 2.GlcN + 2.P + 1.C14OH + 2.C16OH + 1.C14
- Peak 1494.8 = 2.GlcN + 1.P + 1.C14OH + 2.C16OH + 1.C14 + 1.AraN
- Peak 1575.31 = 2.GlcN + 2.P + 1.C14OH + 2.C16OH + 1.C14 + 1.AraN
- Peak 1590.78 = 2.GlcN + 1.P + 2.C14OH + 2.C16OH + 1.C14
- Peak 1670.06 = 2.GlcN + 2.P + 2.C14OH + 2.C16OH + 1.C14
- Peak 1721.0 = 2.GlcN + 1.P + 2.C14OH + 2.C16OH + 1.C14 + 1.AraN
- Peak 1801.55 = 2.GlcN + 1.P + 2.C14OH + 2.C16OH + 2.C14
- Peak 1801.55 = 2.GlcN + 2.P + 2.C14OH + 2.C16OH + 1.C14 + 1.AraN
- Peak 1932.82 = 2.GlcN + 2.P + 2.C14OH + 2.C16OH + 1.C14 + 2.AraN

504

505 **Figure S6. MS profile of *C. insecticola* lipid A.** The MS spectrum is shown for the lipid A molecule
 506 of *C. insecticola*. The masses of the peaks are indicated and their corresponding chemical
 507 composition is listed below. The inset shows the structure of the lipid A molecule with m/s= 1932,82
 508 with its two Ara4N modifications shown in light blue.

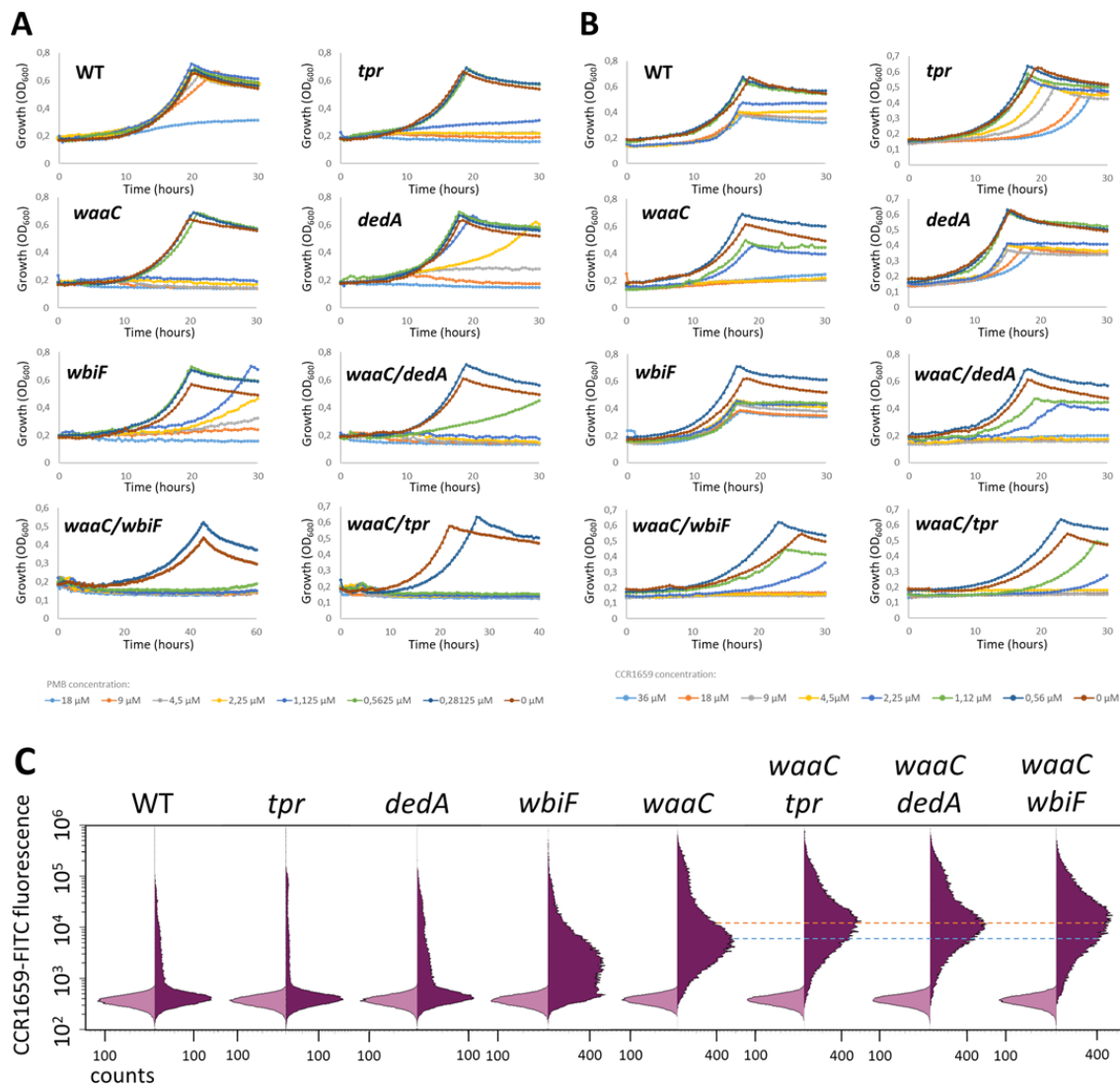
509

	WT	<i>tpr</i>	<i>dedA</i>	<i>tolB</i>	<i>tolQ</i>	<i>waaC</i>	<i>waaF</i>	<i>rfbA</i>	<i>wbiF</i>	<i>wbiG</i>	<i>wbiI</i>	<i>wzm</i>
NPN (au)	1079	1231	1387	1142	1206	2407	2569	1479	1528	1462	863	1163
SDS (%)	0,02	0,02	0,02	0,01	0,01	0,02	0,02	0,02	0,02	0,02	0,02	0,02
Triton X100 (%)	0,1	0,1	0,1	0,1	0,1	0,025	0,025	0,1	0,1	0,1	0,1	0,1
CTAB (%)	0,00025	0,00025	0,00025	0,00025	0,00025	0,000125	0,000125	0,00025	0,00025	0,00025	0,00025	0,00025

510

511 **Figure S7. Membrane integrity.** Steady state membrane integrity was measured in the wild type
512 (RPE75) and indicated mutant strains, grown in MM medium in exponential growth phase, by NPN
513 fluorescence (au, arbitrary units of fluorescence) and sensitivity to anionic (SDS), neutral (Triton
514 X100) or cationic (CTAB) detergents. The indicated sensitivities (% detergent) indicate the minimal
515 concentration to inhibit or diminish growth.

516



517

518

Figure S8. Synthetic phenotype in double mutants. (A) Growth inhibition of *C. insecticola* wild

519

type, single mutants and double mutants by different concentrations of PMB. (B) Growth inhibition

520

of *C. insecticola* wild type, single mutants and double mutants by different concentrations of

521

CCR1659. (C) Flow cytometry analysis of 7.5 μM CCR1659-FITC binding by *C. insecticola* wild

522

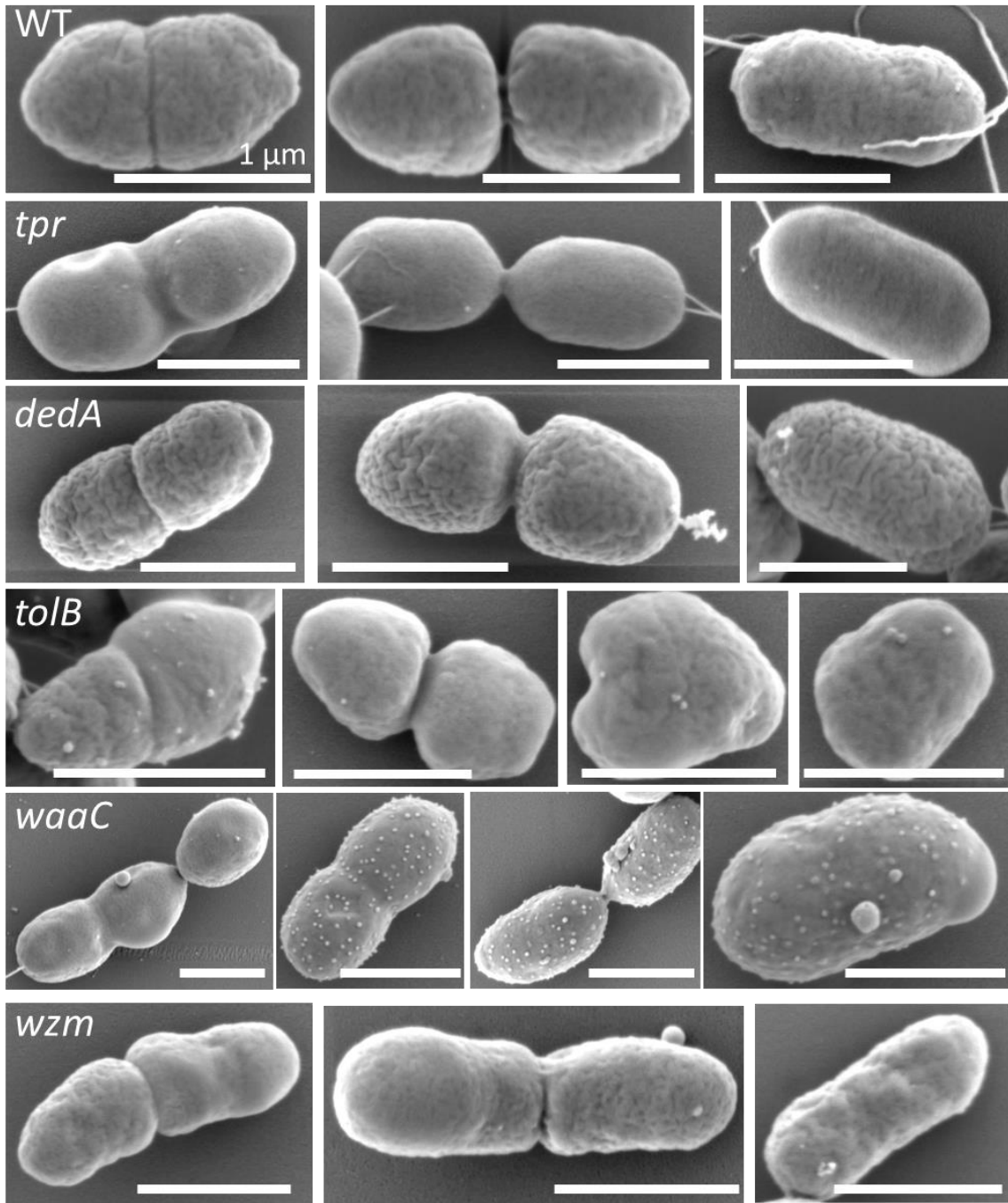
type, single mutants and double mutants. Light purple histograms are control measurements

523

without fluorescent label; the dark purple histograms are in the presence of the fluorescent label.

524

524



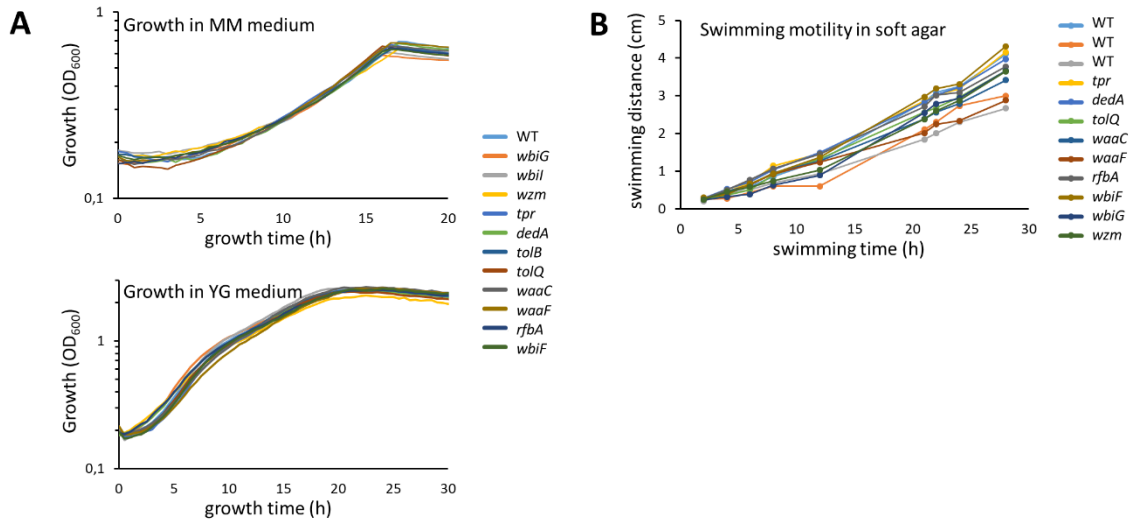
525

526

527

528

Figure S9. AMP-sensitivity mutants have altered envelope and morphology properties. SEM micrographs of untreated *C. caballeronia* wild-type and indicated mutant cells. Scale bars are 1 μm.



529

530

531 **Figure S10. Growth characteristics of the *C. insecticola* wild type and mutants.** (A) Growth

532 curves of *C. insecticola* wild type (RPE75) and the indicated mutants in MM medium (top) and YG

533 medium (bottom). (B) Swimming activity of *C. insecticola* wild type (RPE75) and the indicated

534 mutants in YG soft agar 0,3 %. The swimming distance corresponds to the diameter of the growth

534 halo on the plate.