



HAL
open science

Bridges instead of boats? The Mla system of diderm Firmicute *Veillonella parvula* reveals an ancestral transenvelope core of phospholipid trafficking

Kyrie Grasekamp, Basile Beaud, Najwa Taib, Bianca Audrain, Benjamin Bardiaux, Yannick Rossez, Nadia Izadi-Pruneyre, Maylis Lejeune, Xavier Trivelli, Zina Chouit, et al.

► To cite this version:

Kyrie Grasekamp, Basile Beaud, Najwa Taib, Bianca Audrain, Benjamin Bardiaux, et al.. Bridges instead of boats? The Mla system of diderm Firmicute *Veillonella parvula* reveals an ancestral transenvelope core of phospholipid trafficking. 2023. hal-04235315

HAL Id: hal-04235315

<https://hal.science/hal-04235315v1>

Preprint submitted on 23 Oct 2023

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.

Copyright

1 **Bridges instead of boats? The Mla system of diderm Firmicute *Veillonella parvula***
2 **reveals an ancestral transenvelope core of phospholipid trafficking.**

3

4 Kyrie P. GRASEKAMP¹, Basile BEAUD^{2#}, Najwa TAIB^{2,3#}, Bianca AUDRAIN^{1#}, Benjamin
5 BARDIAUX^{#4,5}, Yannick ROSSEZ⁶, Nadia IZADI-PRUNEYRE^{4,5}, Maylis LEJEUNE^{4,5}, Xavier
6 TRIVELLI⁷, Zina CHOUIT⁶, Yann GUERARDEL^{6,8}, Jean-Marc GHIGO¹, Simonetta GRIBALDO^{2*},
7 Christophe BELOIN^{1*}

8

9 ¹ Institut Pasteur, Université Paris Cité, CNRS UMR6047, Genetics of Biofilms Laboratory,
10 Paris, France.

11 ² Institut Pasteur, Evolutionary Biology of the Microbial Cell Laboratory, Paris, France.

12 ³ Institut Pasteur, Université Paris Cité, Bioinformatics and Biostatistics Hub, F-75015 Paris,
13 France.

14 ⁴ Institut Pasteur, Université Paris Cité, Structural Bioinformatics Unit, CNRS UMR3528,
15 Paris, France.

16 ⁵ Institut Pasteur, Université Paris Cité, Bacterial Transmembrane Systems Unit, CNRS
17 UMR3528, Paris, France.

18 ⁶ Université de Lille, CNRS, UMR 8576 – UGSF – Unité de Glycobiologie Structurale et
19 Fonctionnelle, Lille, France.

20 ⁷ Université de Lille, CNRS, INRAE, Centrale Lille, Université d'Artois, FR 2638 – IMEC –
21 Institut Michel-Eugène Chevreul, Lille 59000, France

22 ⁸ Institute for Glyco-core Research (iGCORE), Gifu University, Gifu, Japan

23

24 [#] These authors contributed equally to the manuscript

25 ^{*} Corresponding authors: christophe.beloin@pasteur.fr; simonetta.gribaldo@pasteur.fr

26

27 **Abstract**

28 Despite extensive characterisation of envelope biogenesis systems in diderm bacteria,
29 glycerophospholipid (GPL) trafficking remains poorly understood, and has only been studied in a
30 handful of model species. Within the Proteobacteria, the maintenance of lipid asymmetry (Mla) system
31 facilitates retrograde GPL trafficking via six proteins, MlaA-F. GPLs are extracted from the outer
32 leaflet of the outer membrane by the lipoprotein MlaA which associates with porin trimers, then
33 shipped through the periplasmic space by the chaperone MlaC, which finally delivers GPLs to the
34 inner membrane complex formed by MlaBDEF. Here, we investigate GPL trafficking in *Veillonella*
35 *parvula*, a diderm member of the Firmicutes which encodes an Mla system devoid of MlaA and MlaC.
36 *V. parvula* Δmla mutants display phenotypes characteristic of disrupted lipid asymmetry such as
37 hypervesiculation and detergent hypersensitivity, and lipid content analysis from outer membrane
38 vesicles reveals an enrichment for the major lipid component phosphatidylethanolamine. Interestingly,
39 suppressor analysis identifies mutations in *tamB* that rescue detergent hypersensitivity and
40 hypervesiculation of Δmla strains, supporting the involvement of these two systems in antagonistic
41 GPL trafficking functions across diverse bacterial lineages. A combination of structural modeling and
42 subcellular localisation assays shows that MlaD_{Vp} is longer than in classical diderm models and forms
43 a transenvelope bridge, encoding both an inner membrane-localised MCE domain and an outer
44 membrane β -barrel. These results strongly suggest that *V. parvula* possesses a minimal Mla system
45 for GPL trafficking, replacing the need for chaperones and outer membrane lipoproteins by directly
46 connecting the two membranes. Finally, phylogenomic analysis indicates that this MlaEFD self-
47 contained architecture is widely distributed in diderm bacteria and most likely represents the ancestral
48 functional core of the Mla system, which subsequently increased in complexity in Proteobacteria and
49 closely related phyla following the emergence of MlaABC. Our work broadens the diversity of current
50 models of GPL trafficking in diderm bacteria, challenging the paradigm set by classical models and
51 shedding light on the evolution of a crucial system in the biogenesis and maintenance of the bacterial
52 outer membrane.

53

54

55

56 **Keywords:** phospholipid trafficking; outer membrane biogenesis; Mla system, lipid asymmetry,
57 evolution

58

59

60 Introduction

61 The diderm bacterial envelope is a complex structure consisting of two membranes: a cytoplasmic -or
62 inner- membrane (IM) composed primarily of glycerophospholipids (GPLs), and an outer membrane
63 (OM) which exhibits lipid asymmetry, comprising GPLs in its inner leaflet and primarily
64 lipopolysaccharide (LPS) molecules in its outer leaflet. This asymmetry manifests as a robust
65 exclusion barrier against a range of antibiotics and host components, such as vancomycin and bile
66 salts^{1,2}. Although GPLs are one of the most ubiquitous amphipathic components of the diderm
67 envelope, their transport systems remain poorly understood, and most current information stems from
68 a handful of model bacterial species. Recent studies implicate large AsmA-like proteins, such as TamB
69 and YhdP, in the mediation of anterograde GPL trafficking³⁻⁵, whilst MCE proteins – namely MlaD
70 within the maintenance of lipid asymmetry (Mla) machinery - are proposed to facilitate retrograde
71 GPL trafficking^{6,7}. In *E. coli*, the Mla system is composed of six proteins (**Fig 1**): MlaBDEF form the
72 IM complex^{8,9}, MlaC is the periplasmic chaperone^{6,10}, and MlaA is the OM lipoprotein that associates
73 with porin trimers¹¹. Despite extensive research of the Mla system within model diderms for over a
74 decade, several mechanistic details – including substrate specificity and even directionality, until
75 recently - remain controversial^{9,12-14}.

76
77 Here, we investigate GPL trafficking in *Veillonella parvula*, a genetically tractable member of the
78 Firmicutes (Negativicutes) which presents a diderm envelope with LPS^{15,16}. *V. parvula* is
79 phylogenetically distant from classical models such as *E. coli* as it belongs to the Terrabacteria, one
80 of the two large clades into which Bacteria are divided (containing phyla such as Actinobacteria,
81 Cyanobacteria and Candidate Phyla Radiation (CPR)), and whose cell envelopes are largely
82 understudied^{16,17}. Remarkably, the Firmicutes is the only phylum so far identified to contain a mixture
83 of both monoderm and diderm clades, offering an ideal platform on which to study both the
84 monoderm-diderm transition¹⁸ and the diversity of OM biogenesis systems¹⁹. We previously
85 demonstrated that this diderm architecture is an ancestral feature, present in both the ancestor of the
86 Firmicutes and the Last Bacterial Common Ancestor (LBCA)^{15,17-19}, an inference also supported by a
87 recent study²⁰. Studying diderm Firmicutes can therefore shed light on the evolution of ancestral
88 systems in the biogenesis and maintenance of the bacterial OM. Interestingly, the *V. parvula* genome
89 contains an operon embedded in a large OM biogenesis and maintenance cluster^{15,18} with only three
90 homologues of the Mla system: the IM proteins MlaDEF, directly followed by a homologue of the
91 OM efflux protein TolC²¹, and TamB (**Fig 1**). This arrangement is conserved in all Negativicutes
92 (**Extended Fig 1**). However, no homologues of the periplasmic chaperone MlaC or the OM lipoprotein
93 MlaA were identifiable at the sequence level, consistent with the fact that no OM lipoproteins or
94 homologues of the Lol system have so far been identified in *V. parvula*^{16,22}. This raises the question
95 of how these *V. parvula* Mla homologues are involved in retrograde lipid trafficking, and how they
96 can accomplish this function in the absence of MlaABC.

97 We combine phenotypic characterisation, lipid content analysis, structural modeling and evolutionary
98 reconstructions to show that the three-component Mla system of *V. parvula* is involved in GPL
99 trafficking, and is strikingly different to what has been described so far in model diderms; MlaDEF
100 appear to form a functional ‘minimal’ system, in which a transenvelope MlaD abrogates the
101 requirement for the missing homologues MlaA and MlaC by directly bridging the IM and OM. Further,
102 we show that MlaDEF represent an ancestral core for GPL trafficking that is widely distributed across
103 Bacteria and likely dates back to the LBCA, whereas the presence of MlaABC are an exception,
104 emerging later in the Proteobacteria and other closely related phyla. Finally, and most strikingly, we
105 show that the majority of MlaD sequences across the bacterial kingdom are ‘long’, revealing that the
106 short MlaD employed by Proteobacteria is atypical. Together, our results uncover novel functional
107 information about GPL trafficking in a non-model organism, shedding light on the evolution of the
108 Mla system and challenging the dogma that OM biogenesis systems studied in *E. coli* are
109 representative of all diderms.

110 **Results**

111 **A "minimal" Mla system in *V. parvula* is involved in GPL trafficking**

112 During a screen to identify genes required for the generation and maintenance of the OM barrier in *V.*
113 *parvula* SKV38 (see methods), we identified a transposon insertion in a homologue of *mldA*, and a
114 mixed clone containing an insertion in both *mldE* and a gene of unknown function, suggesting a role
115 for these *mld* genes in envelope biogenesis. To further investigate the function of the *V. parvula*
116 MlaDEF system, we generated the corresponding single and triple *mldDEF* deletion mutants and
117 assessed their phenotypes.

118

119 Consistent with Mla-defective phenotypes of model diderms such as *E. coli* and *A. baumannii*^{7,12}, all
120 *V. parvula* Δmla strains — with the exception of the $\Delta mldE$ mutant that could not be obtained —
121 display hypersensitivity to SDS/EDTA and increased resistance to vancomycin (**Fig 2A**), with
122 otherwise no differences in viability, morphology, biofilm formation or aggregation (**Extended Fig**
123 **2**). Phenotypes were not additive across single and triple mutants, suggesting these three *mldDEF*
124 genes work together as part of a single system, further implied by the overlapping start and stop codons
125 of *mldF* and *mldA*. SDS/EDTA sensitivity could be complemented in both single and triple mutants,
126 and appeared slightly fuller for the triple mutant, suggesting that correct stoichiometry in the
127 expression of all three genes is important for phenotypic rescue. $\Delta mldEFD$ and $\Delta mldA$ strains also
128 displayed a ~17-fold increase in outer membrane vesicle (OMV) production as compared to the WT
129 (**Fig 2B**), consistent with previously described hypervesiculation phenotypes of Δmla strains in *Vibrio*
130 *cholerae*, *Haemophilus influenzae*, *Bordetella pertussis* and *Neisseria gonorrhoeae*^{23–25}, which could
131 be significantly rescued via complementation to a ~4-fold increase in OMV production as compared
132 to the WT strain. Together, these phenotypes are indicative of a loss of lipid asymmetry^{23,26}. To
133 determine if these changes in membrane permeability and stability are due to a change in quantity or
134 composition of LPS, silver-staining was used to assess relative LPS profiles of these Δmla mutants,
135 revealing no observable differences (**Extended Fig 2B**).

136

137 Previous work in model diderms has inferred the function of the Mla system by exploiting other
138 pathways that maintain lipid asymmetry, such as overexpression of the phospholipase A1 encoding
139 gene (*pldA*) to rescue detergent hypersensitivity⁷ and radiolabelling with PagP to follow the hepta-
140 acylation of lipid A^{7,27}. As homologues of these systems are not present in *V. parvula*, we relied on
141 lipid extraction and comparative thin layer chromatography (TLC) to infer the function of these *mld*
142 genes. However, to understand potential changes in lipid composition due to deletion of the Mla
143 system, we first had to determine the nature of the major lipids present in *V. parvula* SKV38. We
144 identified homologues of genes that may be involved in the biosynthesis of phosphatidylethanolamine
145 (PE), phosphatidylserine (PS) and phosphatidylglycerol (PG), as well as a homologue of the recently
146 identified hydrogenase / reductase operon required for plasmalogen biosynthesis in *Clostridium*

147 *perfringens*²⁸, encoded within a single gene (**Supplementary Fig S1A**). Using TLC, matrix-assisted
148 laser desorption/ionisation quadrupole ion trap time-of-flight (MALDI-QIT-TOF) and nuclear
149 magnetic resonance (NMR) spectroscopy on lipid extracts from the WT, we then revealed that the
150 major lipid species in the envelope of *V. parvula* are PE (~68%) and PG (~20%), along with two other
151 uncharacterised lipids representing a further ~9% and ~3% of the envelope (**Supplementary Fig S1**,
152 **Supplementary Table S1**). NMR analyses also revealed that ~15-35% of each lipid species present
153 in the envelope of *V. parvula* contains a vinyl ether bond, which could result from the activity of the
154 identified plasmalogen biosynthesis homologue (**Supplementary Fig S2**). No cardiolipin was
155 observed via MALDI-QIT-TOF, consistent with the absence of homologues of genes involved in the
156 synthesis of this lipid in the *V. parvula* genome.

157
158 Both whole-cell and membrane lipid extracts from WT and Δmla strains were used to perform TLC,
159 however no differences in GPL composition were observed, as previously described in *E. coli*⁷
160 (**Supplementary Fig S3A**). We posited that this lack of difference may be due to the observed
161 hypervesiculation of Δmla strains, compensating for OM lipid imbalance by shedding excess envelope
162 material. We therefore purified OMVs produced by both the WT and $\Delta mlaD$ strains, performed lipid
163 extraction and TLC, and revealed different classes of lipid species by staining with iodine vapor, Schiff
164 reagent and phosphomolybdic acid (PMA). Across all iodine vapour-stained plates, OMVs produced
165 by $\Delta mlaD$ strains display a ~40% enrichment for PE, whilst also containing a marked reduction in
166 another lipid species that is abundant in WT OMVs (**Fig 2C, Supplementary Fig S3B**). These results
167 suggest that Δmla mutants accumulate PE in their OM, which is subsequently shed from the envelope
168 via excessive vesicle formation. Taken together, this analysis strongly suggests that the minimal
169 MlaEFD system of *V. parvula* is involved in retrograde GPL trafficking.

170
171 **Suppressor analysis implicates Mla and TamB in coordination of GPL homeostasis in *V. parvula***

172 To further strengthen the proposed function of the Mla system in *V. parvula*, we investigated the genes
173 able to suppress detergent hypersensitivity of Δmla strains. By performing random transposon
174 mutagenesis in the $\Delta mlaD$ background, we identified ten chromosomal insertion mutants able to grow
175 in SK media supplemented with 0.004% SDS and 312.5 μ M EDTA. Efficiency of plating revealed
176 that only four of these mutants could significantly rescue the $\Delta mlaD$ phenotype (**Fig 3A**), of which
177 two were of particular interest: sequencing revealed two identical Tn insertions in the downstream
178 homologue of *tamB* (**Fig 3B**).

179
180 Considering the recent implications of TamB and other large AsmA-like proteins in anterograde GPL
181 trafficking^{3,4}, we were interested to identify this gene as a suppressor of $\Delta mlaD$ detergent sensitivity.
182 Indeed, we had previously identified a spontaneous suppressor of $\Delta mlaD$ with a SNP encoding a STOP
183 codon in *tamB*, though the phenotype was unstable and reverted (**Fig 3B**). We therefore generated

184 clean, stable deletion mutants to further probe the relationship between these two genes. As with the
185 suppressor screens, loss of *tamB* partially rescues SDS/EDTA sensitivity in $\Delta mlaD$ and reduces
186 vancomycin resistance (**Fig 3C**). Interestingly, the *tamB* mutant presents a permeability phenotype
187 which appears to be the opposite of Δmla mutants, with no change in SDS/EDTA sensitivity and an
188 increased sensitivity to vancomycin, as previously described in *E. coli*³ (**Fig 3C, Supplementary Fig**
189 **S4**). Much like $\Delta mlaD$, however, $\Delta tamB$ also displays a hypervesiculation phenotype, with a ~10-fold
190 increase in vesicle production as compared to the WT strain (Fig 3D). Strikingly, loss of *tamB* in the
191 $\Delta mlaD$ background significantly reduces hypervesiculation; the $\Delta tamB\Delta mlaD$ mutant produces only
192 ~5-fold more vesicles than the WT (**Fig 3D**). This rate of reduction in OMV production is comparable
193 to the phenotype of the fully-complemented $\Delta mlaEFD::mleEFD$ strain, suggesting that deletion of
194 *tamB* is as effective at rescuing this envelope defect as reintroducing the *mle* genes. The contrasting
195 OM permeability phenotypes of $\Delta mlaD$ and $\Delta tamB$, together with the striking complementation of the
196 double $\Delta mlaD\Delta tamB$ strain, hints to the antagonistic functions of these two genes in envelope
197 biogenesis and maintenance. Overall, these results suggest that MlaEFD and TamB may represent the
198 two major systems maintaining GPL homeostasis in the OM of *V. parvula* - likely retrograde and
199 anterograde, respectively.

200

201 **Structural modelling suggests a transenvelope Mla system in *V. parvula***

202 Though phenotypic and suppressor analyses elucidated the functional role of MlaEFD, we sought to
203 understand how three IM proteins could be structurally arranged within the envelope of *V. parvula* to
204 facilitate GPL trafficking in the absence of MlaABC. Whilst the *mleE* and *mleF* homologues closely
205 resemble those of model diderms such as the Proteobacteria, *mleD* in *V. parvula* is over double the
206 size of the one of *E. coli* (419 amino acids vs 183 amino acids respectively) (**Fig 4A**). Both versions
207 of this gene contain a transmembrane domain (TMD) that anchors the protein in the IM, and a
208 conserved MCE domain that binds lipids⁶. However, MlaD in *V. parvula* extends with a long predicted
209 α -helical region spanning ~130 residues, and a predicted C-terminal β -barrel (**Fig 4A**).

210

211 AlphaFold2²⁹ (AF2) was used to predict the structure of this unusually long MlaD protein, but was
212 unable to confidently model the elongated α -helical region, and indeed the positioning of this domain
213 relative to the N- and C-termini of the protein (**Extended Fig 3**). However, high confidence predictions
214 were achieved for the TMD and MCE domain, and for the C-terminal β -barrel. This C-terminal domain
215 displays characteristic features of a membrane-embedded β -barrel, composed of ten alternating
216 antiparallel β -strands with a hydrophobic and neutral surface for the β -sheets, and with ‘girdles’ of
217 negatively charged and aromatic residues which flank the top and bottom of the structure (**Extended**
218 **Fig 4**).

219

220 Considering all characterised MCE proteins to date form hexamers, we merged overlapping predicted
221 structures of all domains of MlaD_{Vp} with hexameric stoichiometry, generating a model for the full-
222 length protein (**Fig 4B**). Even when generated without templates, a 6:2:2 stoichiometry of the MlaDEF
223 complex is confidently predicted for *V. parvula*, with high similarity to the resolved structure of the
224 MlaDEF complex from *E. coli* (**Extended Fig 3**). Interestingly, two major models are generated for
225 the hexameric α -helical region of MlaD_{Vp}: one with a closed tunnel structure, and one with an open
226 pore or groove with little interaction between the first and last chains of the hexamer (**Extended Fig**
227 **5, Extended Table 1**). We suspect that the “open groove” configuration may represent an artifact of
228 the prediction, owing to the way AlphaFold-Multimer concatenates sequences of all chains prior to
229 structure inference. Instead, the closed tunnel model (**Extended Fig 5**) more closely reflects the
230 resolved crystal structure of PqiB, another MCE-domain protein from *E. coli*, which forms a closed
231 α -helical tunnel on top of three hexameric MCE rings⁶, and the recently resolved heterohexamer of
232 Mce1A-1F from *Mycobacterium tuberculosis*³⁰. Like these resolved structures, the predicted tunnel
233 formed by the α -helices generates a hydrophobic interior of ~13 Å (**Fig 4C**), supporting the possibility
234 of a role in hydrophobic substrate transport. Further, both the N- and C-termini of the full-length model
235 generate highly hydrophobic bands where both regions are predicted to embed within the IM and OM,
236 respectively. The rest of the structure, which would reside in the aqueous periplasm, displays a
237 hydrophilic exterior (**Fig 4C**). It is interesting to note the prediction of a small external cluster of
238 hydrophobic residues at the lowest confidence region of the tunnel – which coincides with the region
239 that forms a groove in the higher confidence model.

240
241 From these high-confidence structural predictions of a membrane-embedded C-terminal β -barrel and
242 an IM MCE-domain hexamer, we hypothesised that MlaD from *V. parvula* may have domains
243 anchored in both the IM and OM, forming a transenvelope bridge. This would remove the necessity
244 for a periplasmic chaperone boat and OM lipoprotein, as MlaD would be capable of connecting the
245 two membranes directly to facilitate transport. To verify this hypothesis experimentally, we localised
246 full-length MlaD and its separate domains by performing membrane fractionation by sedimentation
247 on sucrose density gradients. We used total membrane pellets (containing both IM and OM vesicles)
248 obtained from WT and $\Delta mlaD$ strains complemented with HA-tagged versions of the TMD-MCE
249 domain and the β -barrel. Localisation of these domains and of the full-length protein was examined
250 via immunoblotting with our developed MlaD antisera, and the separate domains with a commercial
251 anti-HA antibody. Whilst the full-length MlaD was distributed across almost all membrane fractions,
252 the TMD-MCE domain construct was most abundant in fractions F4-5, similar to the localisation
253 profile of the IM control SecA, and the β -barrel was located almost exclusively in the final few OM
254 fractions, F9-11, mirroring the localisation of TolC (**Fig 4D**). These results support the hypothesis that
255 MlaD can span the periplasm, with domains anchored in both the IM and OM.

256

257

258

259 **Phylogenomic analysis reveals an ancestral core of the Mla system**

260 The possibility of a minimal, transenvelope MlaEFD system was intriguing, and we sought to
261 understand if this type of pathway exists in other bacterial lineages by exploring the distribution of all
262 Mla genes. We therefore carried out an exhaustive search for homologues of the six components
263 (MlaABCDEF) in a local databank containing 1083 genomes, representing all current phyla. To
264 identify MlaA, MlaC, MlaD, and MlaE homologues, we used the Pfam domains PF04333, PF05494,
265 PF02470, and PF02405, respectively. As MlaB and MlaF belong to the generic STAS-domain and
266 ATPase families, respectively, we used MacSyFinder2³² to identify them when found in synteny with
267 at least one of the other Mla components in the genome. The distribution of all six components was
268 then mapped onto a reference phylogeny of Bacteria¹⁶ (**Fig 5, Extended Fig 6, Extended Table 2**).

269

270 In agreement with the fundamental role of the Mla system in maintaining OM asymmetry, Mla
271 homologues are largely found in diderm bacteria (**Fig 5**). Some diderm phyla in the Terrabacteria lack
272 any identifiable homologues, possibly due to their atypical envelope architectures: the Thermotogae
273 and their neighbor Candidate phyla Bipolaricaulota and Fraserbacteria have an OM detached at the
274 cell poles³⁴; Dictyoglomi and Deinococcus-Thermus both form rotund bodies where cell aggregates
275 share a single OM^{35,36}; some members of Deinococcus-Thermus have lost LPS; and cryo-electron
276 tomography of a member of Atribacteria (*Atribacter laminatus*) has recently revealed an atypical cell
277 envelope architecture with an intracytoplasmic lipidic bilayer surrounding the nucleoid³⁷. Within the
278 three diderm lineages belonging to the Firmicutes¹⁸, Mla components are found in the Negativicutes
279 and a few Limnochordia (**Extended Fig 1**), whilst they are completely absent from the
280 Halanaerobiales (**Extended Fig 6, Extended Table 2**).

281

282 Conversely, and consistently with its presence in *M. tuberculosis*, some monoderm phyla contain
283 MCE-domain proteins; we identified such homologues in some members of Actinobacteria, and also
284 within the Candidate Dormibacteraeota, usually in association with a Mce4_CUP1 domain (**Extended**
285 **Table 2**). Further, our analysis suggests the MCE homologues present in these monoderm phyla were
286 horizontally acquired from the Proteobacteria (**Extended Figs 6-8**).

287

288 More striking is the overall distribution of MlaDEF vs MlaABC; whilst MlaDEF are widely distributed
289 across almost all diderm phyla, with the exception of those mentioned above, the MlaABC
290 components have a much narrower distribution, present mostly within Proteobacteria, and totally
291 absent in diderm Terrabacteria (**Fig 5, Extended Fig 6, Extended Table 2**).

292

293 We then focused specifically on MlaD to investigate whether other bacteria possess the longer version,
294 similar to *V. parvula*. Surprisingly, we found that the short MlaD sequences – similar to that of *E. coli*
295 - are restricted to taxa within Proteobacteria, whereas 85% of MlaD homologues are of the longer
296 version (corresponding to an average of 374 residues) (**Extended Figs 6-8**). To understand if these
297 longer MlaD versions have a similar predicted structure to that of *V. parvula*, we screened them for
298 the presence of β -sheet structures using BOCTOPUS2³⁶ and AlphaFold2²⁹. We could indeed infer the
299 presence of a C-terminal β -barrel in 103 sequences, with a conserved fold and physicochemical
300 properties, some containing an additional helical extension (**Extended Fig 4F, 4H**). These long
301 versions of MlaD with a β -barrel are widely distributed in the diderm phyla belonging to the
302 Terrabacteria division, whilst in the Gracilicutes they are found mainly in some Proteobacteria and
303 closely related phyla (**Extended Fig 6**). The β -barrel is completely absent in the MlaD from the PVC
304 (Planctomycetes, Verrucomicrobia, Chlamydia) and FCB (Fibrobacteres, Chlorobi, Bacteroidetes)
305 groups (**Extended Fig 6, Extended Table 2**). Interestingly, many of the long MlaD sequences with
306 no identifiable complete β -barrel domain contain a disordered C-terminal region with some β -
307 structure, often short β -hairpins, which may represent the remnants or precursors of a β -barrel
308 (**Extended Fig 9**).

309
310 Finally, to infer the origin and evolutionary history of the Mla system, we concatenated MlaD and
311 MlaE sequences into a character supermatrix (538 leaves and 479 amino acid positions) and a
312 maximum likelihood tree was inferred. Only markers found in a conserved cluster arrangement
313 (MlaDEF and MlaDE) were used, due to the uncertainty of MlaF annotations. The resulting tree
314 (**Extended Fig 7**) is consistent with the reference phylogeny of Bacteria. It shows, in fact, a clear
315 separation between the Terrabacteria and the Gracilicutes (Ultra-Fast Bootstrap=95%), with
316 monophyly of phyla within Terrabacteria - indicating vertical inheritance. Contrastingly, the
317 evolutionary history of MlaDE within the Gracilicutes appears more complex, involving several
318 duplications and horizontal transfer events (**Extended Fig 7**).

319
320 Together, our phylogenomic analyses strongly indicate that the majority of diderms may employ a
321 transenvelope three-component Mla system for GPL homeostasis composed of MlaDEF, with the long
322 version of MlaD, like the one present in *V. parvula*. The six-component Mla system characterised in
323 Proteobacteria, containing MlaABC, may therefore represent an exception, rather than the rule.
324 Moreover, we show that this core three-component Mla system is ancestral and was present in the
325 diderm ancestor of all Bacteria. This system was then maintained in diderm Terrabacteria and many
326 Gracilicutes, while it underwent progressive complexification in Proteobacteria with the acquisition
327 of additional Mla components, forming the well-characterised Mla system that we recognise in *E. coli*
328 today.

329

330 Discussion

331 In this work, we investigated a novel Mla system in the non-model diderm Firmicute *Veillonella*
332 *parvula*, consisting only of MlaEFD, with no homologues of the OM lipoprotein MlaA or the
333 periplasmic chaperone MlaC. This minimal architecture is consistent with our recent in silico studies
334 suggesting that OM biogenesis systems present in *V. parvula* resemble streamlined versions of those
335 typically seen in Proteobacterial models. For example, the *V. parvula* genome encodes a homologue
336 of BamA, but none of the associated OM lipoproteins, BamB-E; a gene encoding TamB is present,
337 but no TamA^{15,18}. These bioinformatic findings spark a fundamental question: are these ‘missing’
338 components replaced by non-homologous proteins? Or do these systems function without them?

339
340 Our results show that MlaEFD likely perform a similar role in *V. parvula* as has been described in
341 model diderms; deletion of these genes results in a loss of OM lipid asymmetry, indicated by a
342 detergent hypersensitivity and hypervesiculation phenotype. Based on the enrichment of
343 phosphatidylethanolamine (PE) in lipid extracts from Δmla OMVs, we suggest the regular role of
344 MlaEFD in *V. parvula* is to perform retrograde PE trafficking. We also determined that ~15% of each
345 lipid species present in these bacteria are in their plasmalogen form, therefore it is likely that the
346 system shuttles plasmenylethanolamine (PlsPE) as well as typical diacyl PE.

347
348 Structural modelling and subcellular localisation helped us to understand how just three *mla* genes may
349 perform the role of GPL trafficking in *V. parvula*; MlaEFD likely form a minimal, periplasm-spanning
350 complex, removing the requirement for the ‘missing’ MlaABC components by directly connecting the
351 IM and OM with a transenvelope bridge. However, future work is required to understand the
352 stoichiometry of this protein complex, and to further explore the structure of MCE proteins across
353 diverse bacterial lineages. Interestingly, MCE proteins in *E. coli*⁶, *A. baumannii*¹³ and – recently - *M.*
354 *tuberculosis*³⁰ have been shown to form both hetero- and homohexamers, and for two of these structures
355 (PqiB and Mce1A-1F), the hexameric nature of the α -helical region results in a closed tunnel that spans
356 the periplasm, with a hydrophobic interior. These studies suggest the tunnel structures could facilitate
357 the transport of a hydrophobic substrate and, from homology, conservation and phenotypic
358 characterisation, the proposed substrate is a glycerophospholipid (or in the case of *M. tuberculosis*,
359 cholesterol). However, it is difficult to understand how such a structure could support the hydrophilic
360 head of GPLs. As phospholipid trafficking systems remain elusive even in the extensively characterised
361 model diderms, little comparison is currently available to understand if this MlaD_{vp} tunnel is a realistic
362 transporter of GPLs.

363
364 So far, the only structurally characterised protein that performs retrograde GPL trafficking is MlaC, the
365 periplasmic chaperone that can shield the acyl chains of a phospholipid whilst the head portion remains
366 solvent-exposed^{6,10}. This chaperone-based transport is a common theme across diderm OM biogenesis

367 systems, from LolA shielding the tails of nascent lipoproteins^{37,38} to Skp / DegP and SurA protecting
368 nascent polypeptides for delivery to the BAM system³⁹. However, our phylogenomic analysis shows
369 that the MlaC chaperone – along with MlaA and MlaB - is only sparsely distributed in diderm bacteria.
370 This suggests that the majority of diderms perform retrograde GPL trafficking in a very different way
371 from *E. coli*. Considering that the long version of MlaD is actually the most common one, and that the
372 short *E. coli*-like version of MlaD is almost entirely restricted to the Proteobacteria, it is possible that
373 periplasm-spanning components may be a widespread method of GPL transport in many diderms.
374 Whilst this challenges the paradigm set by extensive research in model diderms such as *E. coli* and *A.*
375 *baumannii*, it is also supported by recent work in these organisms; evidence now implicates two large,
376 periplasm-spanning AsmA-like proteins, TamB and YhdP, in anterograde lipid trafficking³⁻⁵. In these
377 models, large open β -sheet structures create a hydrophobic groove that is expected to shield the acyl
378 tails of phospholipids in a way analogous to LPS transport by the Lpt system, which employs a β -
379 jellyroll fold, leaving the hydrophilic portion of the lipid substrate solvent-exposed⁴⁰. Unexpectedly,
380 two different models were generated for the tunnel region of MlaD in *V. parvula*: one with an open
381 groove, and one closed. It is tempting to speculate that this open groove, that winds around the outside
382 of the α -helical structure, could support the tails of the PLs in a ‘helter skelter’-like fashion, allowing
383 the heads of the phospholipids to remain free in the aqueous environment. However, it is unlikely that
384 six identical α -helices could form a single asymmetrical groove in its structure and, further, the
385 hydrophobic grooves described above are composed of β -sheet structures. From the literature so far,
386 the closed tunnel model is better supported.

387
388 Structural characterisation of *V. parvula* MlaD should help to decipher both the stoichiometry and
389 arrangement of the MlaD complex, and its mechanism of GPL transport. This holds for the β -barrel
390 domain which could be reminiscent of the OmpC-trimers required for the functioning of the Mla
391 system in *E. coli*^{11,41,42}. Indeed, the presence of a β -barrel for substrate translocation across the OM is
392 a common theme across other OM biogenesis systems, such as the requirement for LptD to translocate
393 LPS to the outer leaflet of the OM⁴³. However, the predicted MlaD β -barrel is too narrow to allow the
394 passage of GPL, and a hexameric ring of β -barrels is yet to be described and seems unlikely. We posit
395 the possibility that several of the β -barrels could fuse to form a trimeric conformation, similar to the
396 trimeric form of TolC that results in a single, larger β -barrel²¹. Importantly, a C-terminal β -barrel was
397 only found present in ~10% of MlaD sequences, in both diderm Terrabacteria and a clade of
398 Proteobacteria. In the remaining Mla homologues, we observed a variety of structures: some proteins
399 contain disordered loops with no β -structures, whilst others contain a few β -hairpins in amongst
400 disordered loops. While we cannot speculate at this stage on the order of the evolutionary events that
401 led to this range of disordered, barrel-less MlaD proteins, it is interesting to consider that these
402 unstructured regions could represent either the remnants or even precursors of a β -barrel. Considering
403 the majority of long MlaD sequences are barrel-less, we could speculate that this structure is not

404 important for the putative lipid importer function across bacteria. However, the recently resolved
405 structure of *M. tuberculosis* Mce1A-1F – which do not contain complete C-terminal β -barrels - shows
406 that each monomer contributes a single β -strand, forming a six-stranded β -barrel as a heterohexamer.
407 Perhaps this structural organisation is repeated in other long MlaD sequences without complete barrel
408 structures, suggesting a functional convergence across diverse bacterial lineages. The function of these
409 barrels remains to be determined, and may differ across species; perhaps this domain is required for
410 stability or anchoring in some bacteria, and for substrate passage in others.

411
412 Regardless of the role of the β -barrel domain, or whether the extended MlaD in *V. parvula* is capable
413 of transporting GPLs via a winding fairground slide or a direct tunnel, the implications are the same:
414 this mechanism is one of higher throughput than a chaperone system. Indeed, this is an important
415 element to consider in the argument about the directionality of the Mla system. Whilst a wealth of
416 studies support the function of Mla in retrograde GPL trafficking^{7,9,12,44}, and several studies argue for
417 the anterograde direction^{13,14,45,46}, efficiency is a key factor to consider. In the case of anterograde GPL
418 trafficking to facilitate cell growth, GPL flow would have to be enormously high throughput,
419 considering the sheer volume of GPLs required to build a bilayer. In these cases, large periplasm-
420 spanning AsmA-like proteins, which can transport several GPLs at once, would likely facilitate the high
421 flow rate required. On the other hand, a chaperone-based system which can only offer a 1:1 ratio of
422 protein:GPL could not accommodate this high flow rate, and likely stands as a fine-tuning mechanism
423 – which fits with the proposed function of the Mla system, i.e. to remove surface-exposed GPLs from
424 the OM to restore lipid asymmetry. This idea is mirrored by the Lpt system that has evolved to transport
425 LPS into the outer leaflet of the OM, which requires a trans-envelope bridge - and not singular protein
426 chaperones. We do not yet understand where the long and likely transenvelope MlaD fits into these
427 scenarios; as MCE proteins are primarily implicated in importer functions, would long MlaD provide
428 high-throughput retrograde GPL trafficking? Or could this process be bi-directional with the same
429 transporter? Perhaps sites of hemifusion, along with accessory proteins, are necessary to control and
430 facilitate the functioning of these trans-envelope structures, as has been suggested for the functioning
431 of YhdP in an *mlaA** background⁵. Intriguingly, as the long MlaD version is widely distributed across
432 Bacteria, whatever role this protein performs is likely common across a wide diversity of lineages,
433 waiting to be uncovered.

434
435 Furthering our understanding of GPL homeostasis in *V. parvula* is the discovery that TamB is a major
436 suppressor of the Δmla phenotype. Given the recent implications of AsmA-like proteins in anterograde
437 GPL trafficking³⁻⁵, it is very interesting to discover that this protein likely performs a similar role in *V.*
438 *parvula*. These findings, in combination with the recent literature, strongly advocate for contrasting but
439 similar roles of TamB and MlaEFD in maintaining the OM in *V. parvula*, performing anterograde and

440 retrograde GPL trafficking respectively. However, as the double $\Delta mlaD\Delta tamB$ mutant is viable, we
441 propose that there are as yet undiscovered systems that perform lipid trafficking in *V. parvula*.

442

443 To conclude, our results reinforce the notion that the paradigms set by extensive studies within the
444 Proteobacteria can be challenged by using non-model and phylogenetically distant bacteria. Research
445 within diderm models such as *E. coli*, *A. baumannii* and *N. meningitidis* is crucial for our understanding
446 of the functioning and virulence of many pathogens, enabling us to expand research into effective
447 treatment options and aid identification of novel targets. However, learnings from fundamental
448 microbiology – including those from commensals such as *V. parvula* – are often transferable, and may
449 help us fill in the gaps regarding the structure, function and diversity of OM biogenesis and maintenance
450 systems across the Bacterial tree of life.

451 MATERIAL AND METHODS

452 Bacterial strains and growth conditions

453 *Veillonella parvula* SKV38 was grown in SK medium (10 g/L tryptone [Difco], 10 g/L yeast extract
454 [Difco], 0.4 g/L disodium phosphate, 2 g/L sodium chloride, and 10 ml/L 60% [wt/vol] sodium DL-
455 lactate; described in ⁴⁷. Cultures were incubated at 37°C in anaerobic conditions, either in anaerobic
456 bags (GENbag anaero; bioMérieux no. 45534) or in a C400M Ruskinn anaerobic-microaerophilic
457 station. *Escherichia coli* was grown in lysogeny broth (LB) (Corning) medium under aerobic conditions
458 at 37°C. Where required, *V. parvula* cultures were supplemented with 20 mg/L chloramphenicol (Cm),
459 200 mg/L erythromycin (Ery), or 2.5 mg/L tetracycline (Tc), and *E. coli* cultures were supplemented
460 with 100 mg/L ampicillin (Amp). 100 µg/L anhydrotetracycline (aTc) was added to induce the P_{tet}
461 promoter and 1mM isopropyl-β-d-thiogalactopyranoside (IPTG) was added to induce the T7 promoter.
462 All chemicals were purchased from Sigma-Aldrich unless stated otherwise. All chemicals were
463 purchased from Sigma-Aldrich unless stated otherwise. Primers and strains used in this chapter are
464 listed in **Supplementary Tables S2-3**.

465 Generation & Screening of Transposon mutagenesis library

466 A random transposon mutagenesis library was generated in *V. parvula* SKV38 WT, or the $\Delta mlaD$
467 mutant, using pRPF215, a plasmid previously used in *Clostridium difficile* that contains an aTc-
468 inducible transposase and mariner-based transposon (Addgene 106377) ⁴⁸. pRPF215 was transformed
469 into *V. parvula* SKV38 by natural transformation and selected on SK agar supplemented with Cm
470 (20µg/mL). Several independent overnight cultures of *V. parvula* SKV38-pRPF215 were diluted to
471 OD₆₀₀ = 0.1 in SK medium supplemented with aTc (100ng/µL) and incubated for 5 h to induce the
472 transposase. Following induction, cultures were diluted and plated onto SK supplemented with
473 erythromycin (200ug/µL) and aTc (100ng/µL) for selection, then incubated in anaerobic conditions for
474 48h. Resultant colonies were used to inoculate Greiner Bio-one polystyrene flat-bottom 96-well plates
475 (655101) in SK supplemented with either Ery and aTc or Cm to confirm both the presence of the
476 transposon and the loss of pRPF215. After 24h incubation, aliquots were taken from each well to
477 inoculate fresh 96-well plates filled with SK and SK supplemented with detergents and antibiotics for
478 the screening process. Original transposon mutagenesis plates were stored in 15% glycerol at -80°C,
479 whilst the screening plates were incubated for 24h then assessed for growth and OM permeability
480 defects. Mutants of interest were tested for stability of phenotype, then harvested for genomic DNA
481 using the Wizard® Genomic DNA Purification Kit (Promega). Genomic DNA was sent for whole-
482 genome sequencing at the Mutualized Platform for Microbiology at Institut Pasteur.

483 **Efficiency of plating / serial dilution assays**

484 Sensibility to different detergents was performed on SK agar plates supplemented with detergents and
485 antibiotics at different concentrations. Serial dilutions of strains (10^0 - 10^{-5}) were inoculated with a
486 starting $OD_{600nm} = 0.5$. Plates were incubated in anaerobic conditions for 48h and *cfu* were calculated.

487 **Chromosomal Mutagenesis by Allelic Exchange**

488 To generate deletion strains of *V. parvula* SKV38, site-directed mutagenesis was performed as
489 previously described⁴⁷. Briefly, 1kb regions upstream and downstream of the target sequence were
490 PCR amplified using Phusion Flash high-fidelity PCR master mix (Thermo Scientific, F548). For the
491 selection process, three main resistance cassettes were PCR amplified with overlapping primers for the
492 upstream and downstream regions of the target gene: the *V. atypica* tetracycline resistance cassette (*tetM*
493 in pBSJL2), the *C. difficile* chloramphenicol resistance cassette (*catP* in pRPF185; Addgene 106367⁴⁹
494 and the *Saccharopolyspora erythraea* erythromycin resistance cassette (*ermE*). PCR products were
495 ligated to generate deletion cassettes via Gibson cloning or used as templates in a second PCR reaction
496 to generate linear dsDNA with the resistance cassette flanked by the upstream and downstream
497 sequences. These deletion constructs were transformed into *V. parvula* by natural transformation
498 (detailed below), and its genomic integration was selected by plating on tetracycline, chloramphenicol
499 or erythromycin-supplemented medium. Positive candidates were further confirmed by PCR and
500 sequencing.

501 **Natural Transformation**

502 Recipient strains were grown on SK agar for 48h in anaerobic conditions at 37°C. Biomass was
503 resuspended in 1 ml SK medium adjusted to $OD_{600} = 0.4$ - 0.8 . 10 μ L aliquots were spotted onto SK
504 agar, then 0.5 - 1 μ g plasmid or 75 - 200 ng/ μ L linear double-stranded DNA (dsDNA) PCR product was
505 added, using distilled water as a negative control. Following 48h incubation, the biomass was
506 resuspended fresh SK medium, plated onto SK agar supplemented with the relevant antibiotic, and
507 incubated for a further 48h. Colonies were streaked onto fresh selective plates, and correct integration
508 of the construct was confirmed by PCR and sequencing.

509 **Growth kinetics**

510 Overnight cultures were diluted to 0.05 OD_{600} in 200 μ L SK that had previously been incubated in
511 anaerobic conditions overnight to remove dissolved oxygen in Greiner flat-bottom 96-well plates. To
512 maintain anaerobic conditions, a plastic adhesive film (adhesive sealing sheet, Thermo Scientific,
513 AB0558) was used to seal the plate whilst inside the anaerobic station. The sealed plates were then
514 incubated in a TECAN Infinite M200 Pro spectrophotometer for 24h at 37°C. OD_{600} was measured
515 every 30m after 900 seconds orbital shaking of 2 mm amplitude.

516 **LPS silver staining (AgNO₃)**

517 ***Sample Preparation***

518 Overnight cultures were concentrated to OD₆₀₀ = 10 in Tricine Sample Buffer (BioRad 1610739: 200
519 mM Tris-HCl, pH 6.8, 40% glycerol, 2% SDS, 0.04% Coomassie Blue) and incubated at 100°C for
520 10min. After cooling to room temperature, Proteinase K was added to a final concentration of 1mg/mL
521 and incubated at 37°C for 1h. 10 µL of samples were loaded onto tris-tricine gels (BioRad 4563063:
522 16.5% Mini-PROTEAN® Tris-Tricine Gel) and migrated in TTS buffer (BioRad 161-0744: 100 mM
523 Tris-HCl pH 8.3, 100 mM Tricine, 0.1% SDS) for 2h at 50mA.

524 ***Staining***

525 To visualise LPS profiles after gel electrophoresis separation, all development solutions were prepared
526 immediately before use. Gels were fixed in ethanol 30%, acetic acid 10% for 1h, then washed 3 times
527 in dH₂O. Gels were incubated for 10m with metaperiodate 0.7%, again washed 3 times in dH₂O, and
528 incubated in thiosulfate 0.02% for 1m. Following a further 3 washes in dH₂O, gels were incubated in
529 AgNO₃ 25mM for 10-15m, washed in dH₂O for 15s, then incubated in development solution (K₂CO₃
530 35mg/ml, formaldehyde 0.03% (w/v), thiosulfate 0.00125%) until bands appeared. To stop the reaction,
531 gels were finally incubated in 4% Tris Base, 2% acetic acid for 30min, then imaged.

532

533 **Spontaneous suppressor selection**

534 Spontaneous suppressor mutations were selected by passaging overnight cultures of *ΔmlaEFD* and
535 *ΔmlaD*, either in liquid media or on agar, supplemented with SDS 0.004% and EDTA 312.5 µM, until
536 growth occurred. Colonies were restreaked onto the same SDS / EDTA concentration, or directly grown
537 in a liquid culture overnight, then genomic DNA was harvested and sent for whole genome sequencing
538 to identify the corresponding SNPs that enabled growth of the mutants in SDS / EDTA.

539 **Lipid extraction and Thin Layer Chromatography (TLC)**

540 ***Sample Preparation (Lipid Extraction)***

541 To extract lipids, typical Bligh-Dyer extraction methods were used: briefly, overnight cultures were
542 pelleted and concentrated to OD₆₀₀ = 15 in 1mL of distilled water, then incubated on ice for 5-10min.
543 Using glass pipettes, a monophasic mixture of 4mL methanol and 2mL chloroform were added to the
544 resuspended cells and vortexed vigorously for 5min. Following a further 10min incubation on ice, 4ml
545 chloroform and 2ml of distilled water were added and samples were vortexed before centrifugation at
546 1500 x g for 5min. This centrifugation step allowed separation of the two phases: the upper water phase,
547 containing solutes, ions, sugars, DNA and RNA, and the lower chloroform phase, containing the lipids.
548 The upper phase was discarded then 1mL sodium chloride solution (0.5M) was added and the samples
549 were vortexed. Following a second centrifugation at 1500 x g for 5min, the lower phase was carefully

550 extracted using glass micropipettes and dried in a vacuum centrifuge or under a flow of nitrogen gas
551 (N₂). If the initial purity was not satisfactory, a secondary sodium chloride step (and centrifugation)
552 could be repeated. Dried lipid extracts were resuspended in 30µL chloroform.

553 ***Thin Layer Chromatography (TLC) and Staining***

554 To visualise lipids, TLC plates (Silica gel 60, non-fluorescent, 0.25mm thick, Merck Millipore) were
555 cut to a width that allowed ~1cm between each lipid sample. Plates were pre-run in acetone before use
556 and dried at 100°C for 30min. 10µL of each sample was loaded with 5µL glass microcapillary tubes
557 (Sigma Aldrich) 1cm from the bottom of the TLC plate. The plate was then placed in a solvent system
558 optimised for the separation of phospholipids by head group polarity: 65:25:4 (v/v/v) or 80:15:2.5
559 chloroform/methanol/water. For separation of lipids from OMVs, we used the ratio of 80:20:2.5. When
560 the solvent front reached ~2cm from the top of the plate, the plate was removed from the solvent system
561 and dried at 100°C for 30m or air-dried for 10min. To reveal phospholipid species, the dried plate was
562 incubated in a glass tank saturated with iodine vapour from iodine crystals until lipid spots were clearly
563 visible (~5min). After imaging this stained plate, the iodine was allowed to decolourise and the same
564 plate was re-stained with phosphomolybdic acid (PMA, 10%), followed by vigorous heating to develop
565 spots, or Schiff reagent (Sigma-Aldrich). Amines were detected by spraying TLC plates with 0.2g
566 Ninhydrin (Sigma-Aldrich) dissolved in 100ml ethanol, followed by vigorous heating to develop spots.
567 For the relative quantification of each lipid class, ImageJ was used.

569 **Lipid identification**

570 ***Nuclear Magnetic Resonance (NMR)***

571 After labile proton to deuteron exchange with CD₃OD, phospholipids were dried under vacuum and
572 dissolved in a mixture of CDCl₃/CD₃OD (2:1, v/v). The samples were then introduced into a 3mm NMR
573 tube. Solution NMR recordings were conducted at 10 or 20 °C on Bruker AVANCE NEO 400 and 900
574 spectrometers equipped with a 5mm TBI and 5mm cryo-TCI probe, respectively. Standard experiments
575 were run: 1D-¹H, 2D-¹H-COSY, 2D-¹H-TOCSY, 2D-¹H-ROESY, 2D-¹H-¹³C-HSQC-DEPT, 2D-¹H-
576 ¹³C-HSQC-TOCSY. After acquisition and phase correction, chemical shifts calibration on CHD₂OD
577 signals were performed for ¹H (3.31 ppm) and ¹³C (49.29 ppm). The data were analysed using the
578 TopSpin software (Bruker).

579 ***Mass spectrometry (MS)***

580 MALDI-QIT-TOF Shimadzu AXIMA Resonance mass spectrometer (Shimadzu Europe, Manchester,
581 UK) in the positive mode was used to identify the different lipids. All lipids were dissolved in
582 CHCl₃/CH₃OH (1:1, v/v) and mixed with the same volume (10µL) of 2,4,6-trihydroxyacetophenone
583 (THAP) dissolved in methanol.

584 **Outer Membrane Vesicle (OMV) Extraction and Purification**

585 Overnight cultures were used to generate 1 litre cultures of the relevant *V. parvula* strains (WT SKV38
586 and Δmla mutants). Bacterial cells were pelleted by centrifugation at 5000 x g for 10min, and
587 supernatants were passed through 0.45 μ m and 0.22 μ m pore-size filters to remove any leftover cells or
588 debris (Sarstedt). The filtered supernatant was then concentrated using Centricon Plus-70 centrifugal
589 filters (100kDa cutoff, Millipore) centrifuged at 3,500 x g 4°C for 30min, 50mL at a time. ~3mL of
590 concentrated OMV solution was purified from 1 litre of supernatant. This OMV solution was then
591 pelleted by ultracentrifugation at 100,000 x g for 3h at 4°C (Beckman Coulter Optima L-80 XP
592 Ultracentrifuge, Type 50.2 Ti rotor). The mass of the OMV pellets was recorded and used for
593 subsequent lipid extraction and analysis via Thin Layer Chromatography (TLC). Purified OMVs were
594 analysed via NanoFCM confirming an average size of ~60 nm, matching the average size of vesicles in
595 whole cell cultures observed via both NanoFCM and cryo-electron tomography (**Supplementary Fig**
596 **S5**).

597 **OMV quantification**

598 Overnight cultures were adjusted for OD₆₀₀ and diluted 1/25 and 1/50 in filtered PBS solution. The size
599 distribution and particle concentration within the samples were analysed by nano-flow cytometry
600 (nFCM) (NanoFCM, Inc., Xiamen, China). First, the instrument was calibrated for particle
601 concentration using 200 nm PE and AF488 fluorophore-conjugated polystyrene beads, and for size
602 distribution using Silica Nanosphere Cocktail (NanoFCM, Inc., S16M-Exo and S17M-MV). The flow
603 rate and side scattering intensity were then calculated based on the calibration curve, and used to infer
604 the size and concentration of the large and small events present in each sample using the NanoFCM
605 software (NanoFCM Profession V2.0). All samples were diluted to ensure the particle count fell within
606 the optimal range of 2000-12,000/min, and all particles that passed by the detector during time intervals
607 of 60s were recorded for each test.

608 **Membrane fractionation by sucrose gradient sedimentation**

609 1.5L overnight cultures were pelleted at 5,000 x g for 15min and resuspended in ~10mL HEPES buffer.
610 Resuspended pellets were supplemented with 100 μ L benzonase and a small amount of lysozyme prior
611 to lysis by two passages through a high-pressure French press (French Press G-M, Glen Mills)
612 homogenizer at 20,000 psi. Following lysis, cell debris was pelleted at 15,000 x g for 90min. Resulting
613 supernatant was spun in an ultracentrifuge (Beckman Coulter Optima L-80 XP Ultracentrifuge, Type
614 50.2 Ti rotor) at 35,000 x g for 90min to obtain the membrane pellet, containing both IM and OM
615 vesicles. Membrane pellets were resuspended in 750 μ l of HEPES buffer using a Dounce homogeniser,
616 then 500 μ L was loaded onto a discontinuous sucrose gradient comprising a bottom 2mL layer of 51%,
617 a 3ml layer of 37%, a 4mL layer of 31%, and a 2.5mL top layer of 23% (w/w) sucrose. Sucrose gradient
618 columns were centrifuged at 35,000 x g for 40h (Beckman Coulter Optima L-80 XP Ultracentrifuge,

619 Type SW41 Ti rotor). At the end of the sucrose gradient sedimentation, 1mL by 1mL was extracted
620 from the top of the gradient column to the bottom, resulting in 1ml membrane fractions.

621 **Structural modelling**

622 For the prediction of protein domains, SMART (Simple Modular Architecture Research Tool)⁵⁰ and I-
623 TASSER⁵¹ were used. AlphaFold2²⁹ was used to generate structural predictions of the Mla proteins in
624 *V. parvula*. To specifically analyse the presence of β -barrels within our target sequences, the
625 BOCTOPUS2 database was used³⁶. As another transmembrane site-determining web-based tool,
626 TMHMM 2.0 (Transmembrane Hidden Markov Model) was also used to analyse the transmembrane
627 sites of proteins⁵².

628
629 All structure predictions with AlphaFold2 were performed through the ColabFold⁵³ implementation.
630 Output models of monomeric full-length MlaD (5 models) were ranked according to the pLDDT score.
631 MlaEFD₁₋₁₃₀ predictions (10 models) were obtained without structural templates and sorted by predicted
632 TM-scores (*pTM*) since it predates the release of AlphaFold-multimer⁵⁴. MCE, α -helical and β -barrel
633 domains in Chimera⁵⁵ and later refined with Modeller⁵⁶ to correct the topology of the rotated hinges.
634 MlaEFD₁₋₁₃₀ predictions (10 models) were obtained without structural templates and sorted by predicted
635 TM-scores (*pTM*) since it predates the release of AlphaFold-multimer⁵⁴. To impose the 2:2:6
636 stoichiometry of the complex, 2 copies of the MlaE and MlaF sequences were concatenated together
637 with 6 copies of the MlaD₁₋₁₃₀ sequence. Model of the hexameric full-length MlaD complex was
638 obtained by stitching together predictions for the MlaD₁₋₁₃₀, MlaD₃₆₋₂₆₃ and MlaD₂₃₉₋₄₁₉ domains, via
639 structural superimposition of the overlapping regions. Predictions for the MlaD₃₆₋₂₆₃ and MlaD₂₃₉₋₄₁₉
640 domains were first obtained using the AlphaFold2-multimer v2.3 parameters with templates found in
641 the PDB70 database⁵⁷ and ranked by *multimer* score⁵⁴.

642
643 Full-length monomeric models of long MlaD sequences were obtained with AlphaFold2 as described
644 above. Secondary structure class of all positions in the best model (highest pLDDT) for each sequence
645 were obtained using DSSP⁵⁸ and PROSS⁵⁹. To automatically assess the presence of residues in β -strand
646 conformation after the MCE domain, the start and end positions of the MCE domain in each sequence
647 was obtained using HMMSEARCH⁶⁰ with the “MlaD protein” pfam profile (PF02470).

649 **Phylogenetic Analysis**

650 We assembled a databank of 1,083 genomes representing all bacterial phyla present at the National
651 Center for Biotechnology (NCBI). We selected three species per order for each phylum. The number
652 of genomes per phylum therefore reflects their taxonomical diversity (Extended Data Table 2). We
653 chose preferably genomes from reference species and the most complete assemblies. We then queried
654 this databank for the presence of MlaA, MlaB, MlaC, MlaD, MlaE, MlaF, and TamB, using the pfam

655 domains PF04333, PF05494, PF02470, PF02405 and PF04357 respectively. The alignments of protein
656 families NF033618 (MlaB) and PRK11831 (MlaF) were downloaded from NCBI (Conserved Domain
657 Database) and used to build HMM profiles using HMMBUILD from the HMMER3.3.2 package⁶⁰
658 MlaA, MlaC, MlaD, MlaE and TamB homologs were searched using HMMSEARCH from the
659 HMMER3.3.2 package. As MlaB and MlaF belong to large protein families, we used MacSyFinder2³¹
660 to identify their occurrences when they are in genomic synteny with at least one of the other components
661 (*mlaACDE*) of the Mla system, with no more than five genes separating them. This also allowed the
662 identification of the occurrences of *mlaACDE* when they were in genomic cluster. All the retrieved hits
663 were curated using functional annotations, domains organization, alignments, and phylogeny.
664 We identified 1119 MCE-domain containing sequences (Extended Data Table 2). Among these,
665 homologues of PqiB and LetB (containing more than one MCE domain) were identified only within
666 Proteobacteria and closely related phyla.

667
668 For each protein, curated sequences were aligned with MAFFT v7.407⁶¹ with the L-INS-I option and
669 trimmed using BMGE-1.12⁶² with the BLOSUM30 substitution matrix. Maximum likelihood trees were
670 generated for each protein using IQ-TREE.2.0.6⁶³ with the best evolutionary model assessed with
671 ModelFinder⁶⁴ according to the Bayesian Information Criterion, and ultrafast bootstrap supports
672 computed on 1000 replicates of the original dataset. Finally, the presence or absence of each protein
673 was mapped onto a reference tree of Bacteria taken from¹⁶. All tree figures were generated using custom
674 made scripts and iTOL⁶⁵.

675
676

677 **STATISTICS AND REPRODUCIBILITY**

678 Data are presented as the mean \pm standard deviation (SD) and with individual data points from at least
679 three independent biological experiments. Statistical analyses were performed using Prism 9.5.0
680 (GraphPad Software Inc.). The distribution of data was analysed by Shapiro-Wilk test and their variance
681 using Bartlett test. Since some of the data generated did not follow a normal distribution or had two
682 high variance we performed non-parametric two-tailed Mann-Whitney test. P-values <0.05 were
683 defined as the level of statistical significance.

684
685

686 **DATA AVAILABILITY STATEMENT**

687 The authors declare that all data supporting the findings of this study are either available within the
688 paper and its supplementary information files or, otherwise, are available from the corresponding author
689 upon request.

690
691

692 **ACKNOWLEDGEMENTS**

693 We gratefully thank Noa Guzzi, and members of the UGB and EBMC units, for their assistance in
694 generating the transposon mutant libraries used in this study. We also thank Robert Smith for his help
695 in scaling up large anaerobic cultures for *V. parvula*. We thank Pierre-Henri COMMERE for his help
696 with the use of the NanoFCM machine. We also thank Philippe Delepelaire for generously providing
697 us with anti-SecA and anti-TolC serum, and Jean-Michel Betton for his help in developing the
698 membrane fractionation protocol. This work was supported by funding from the French National
699 Research Agency (ANR) (Fir-OM ANR-16-CE12-0010) and (OM-LipAsy-CE44-008), by the French
700 government's Investissement d'Avenir Program, Laboratoire d'Excellence "Integrative Biology of
701 Emerging Infectious Diseases" (grant n°ANR-10-LABX-62-IBEID) and by the Fondation pour la
702 Recherche Médicale (grant DEQ20180339185). KG was supported by the Pasteur Paris University
703 program (PPU). B-Beaud was supported by a MENESR (Ministère Français de l'Éducation Nationale,
704 de l'Enseignement Supérieur et de la Recherche) fellowship. Financial support from the IR
705 INFRANALYTICS FR2054 for conducting the research is gratefully acknowledged. The authors
706 acknowledge the IT department at Institut Pasteur, Paris, for providing computational and storage
707 services (TARS cluster).

708

709 **AUTHORS CONTRIBUTIONS STATEMENT**

710 KG performed all molecular biology and microbiology experiments with the assistance of BA. B-Beaud
711 developed and optimised membrane fractionation protocols for *Veillonella parvula* and lipid extraction
712 together with KG. NT performed all evolutionary and phylogenetic analyses. B-Bardiaux performed
713 the structural modeling analysis. YR, XT, ZC and YG performed lipidomic analyses. JMG provided
714 lab facilities. NIP provided lab facilities and expertise for protein purification, assisted by ML. CB and
715 SG supervised the study. KG, CB, and SG wrote the paper with contributions from NT, B-Bardiaux
716 and JMG. All authors contributed to the final version of the manuscript.

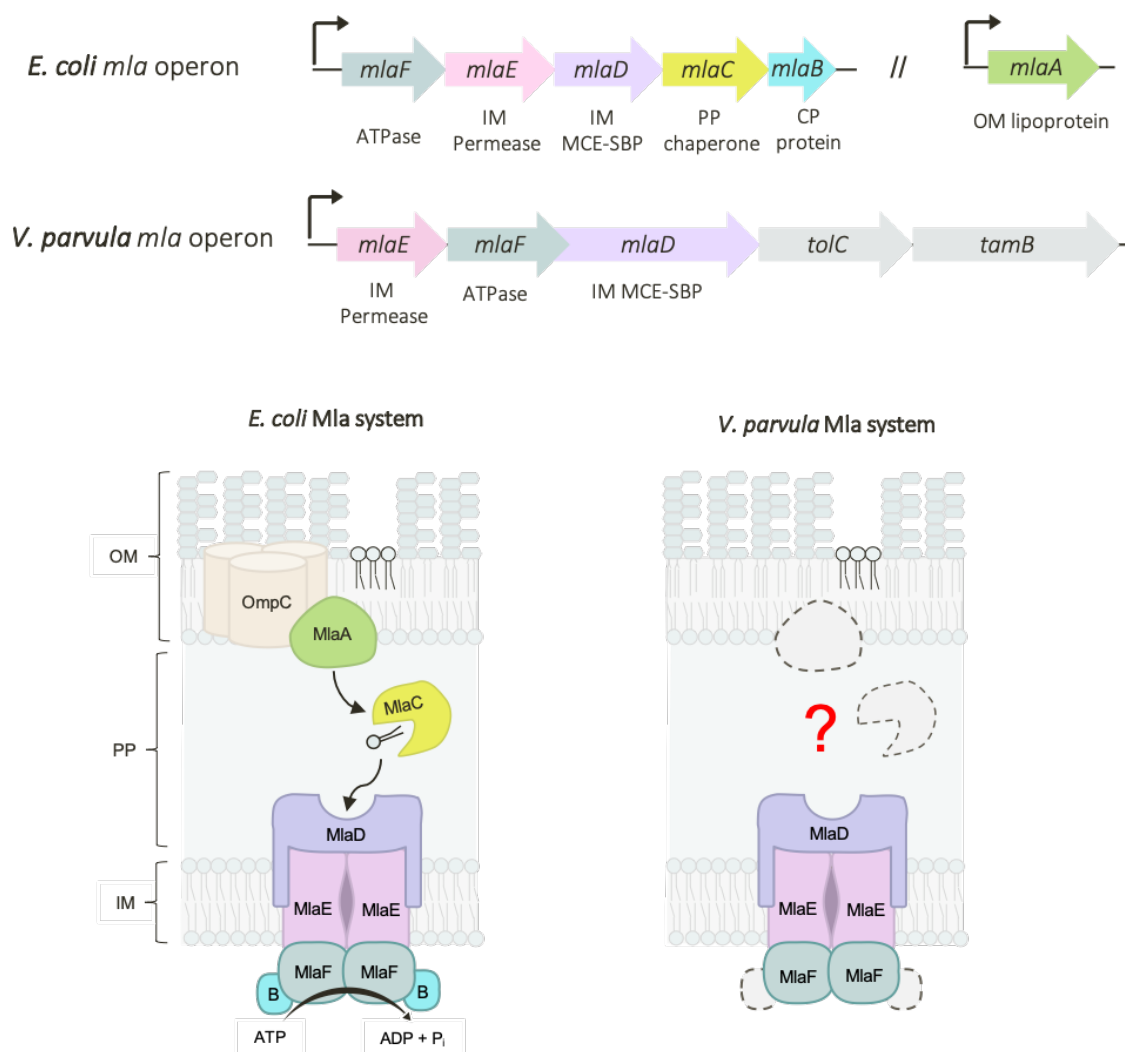
717

718 **COMPETING INTEREST STATEMENT**

719 All the authors declare that they have no competing interests.

720 **MAIN TEXT FIGURES + LEGENDS**

721

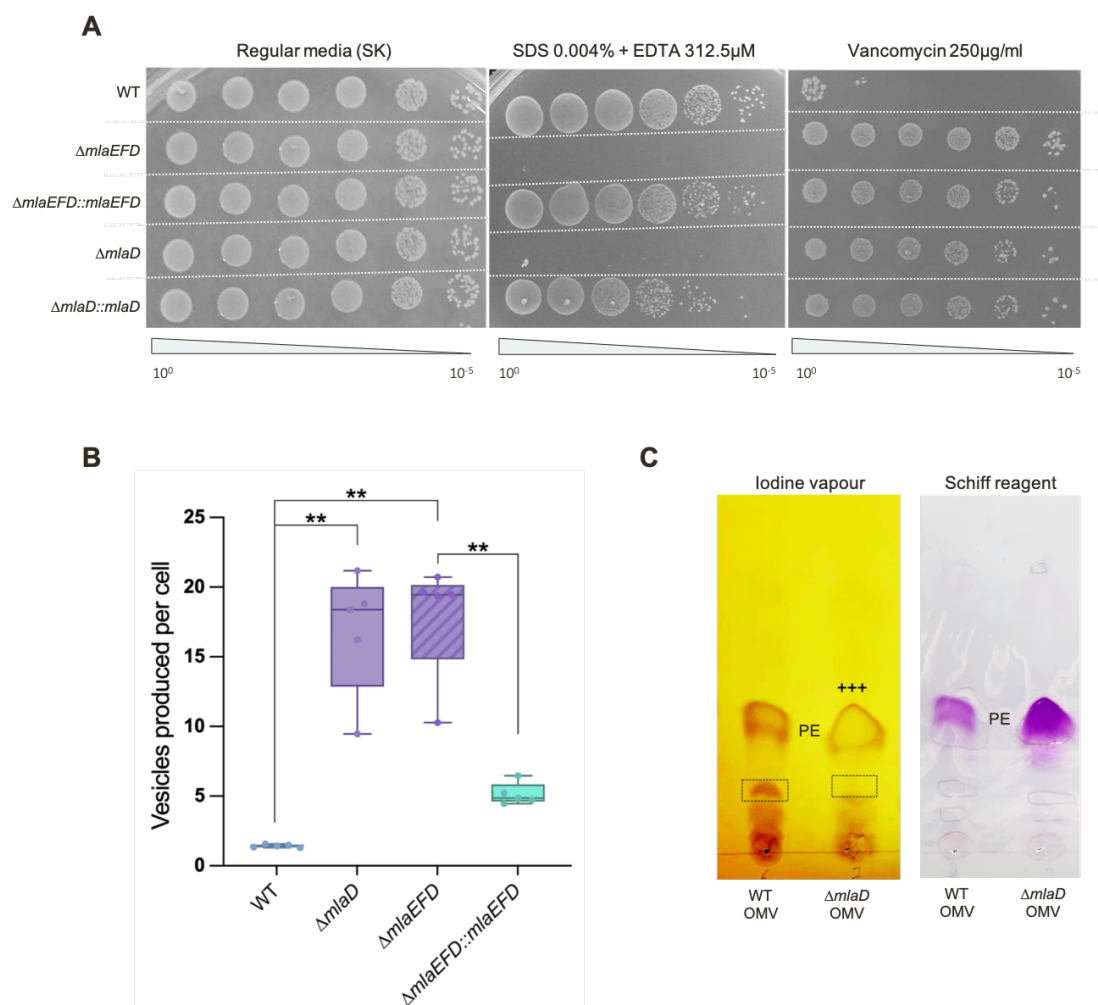


722

723 **Figure 1: Comparison of Mla operons in *V. parvula* and *E. coli*.**

724

725 *In E. coli*, the Mla system is encoded by *mlaB-F*, present in a single operon, and *mlaA* at a separate
 726 locus. MlaBDEF form the IM complex, an ABC transporter in which MlaD is the MCE-domain
 727 substrate binding protein (SBP); MlaC is the periplasmic chaperone, and MlaA is the OM lipoprotein
 728 that associates with OmpC trimers. In *V. parvula*, only homologues of *mlaDEF* were identified. These
 729 genes are encoded together in an operon, immediately upstream of a homologue of *tolC* and *tamB*. OM
 730 = outer membrane; PP = periplasm; IM = inner membrane; CP = cytoplasm.

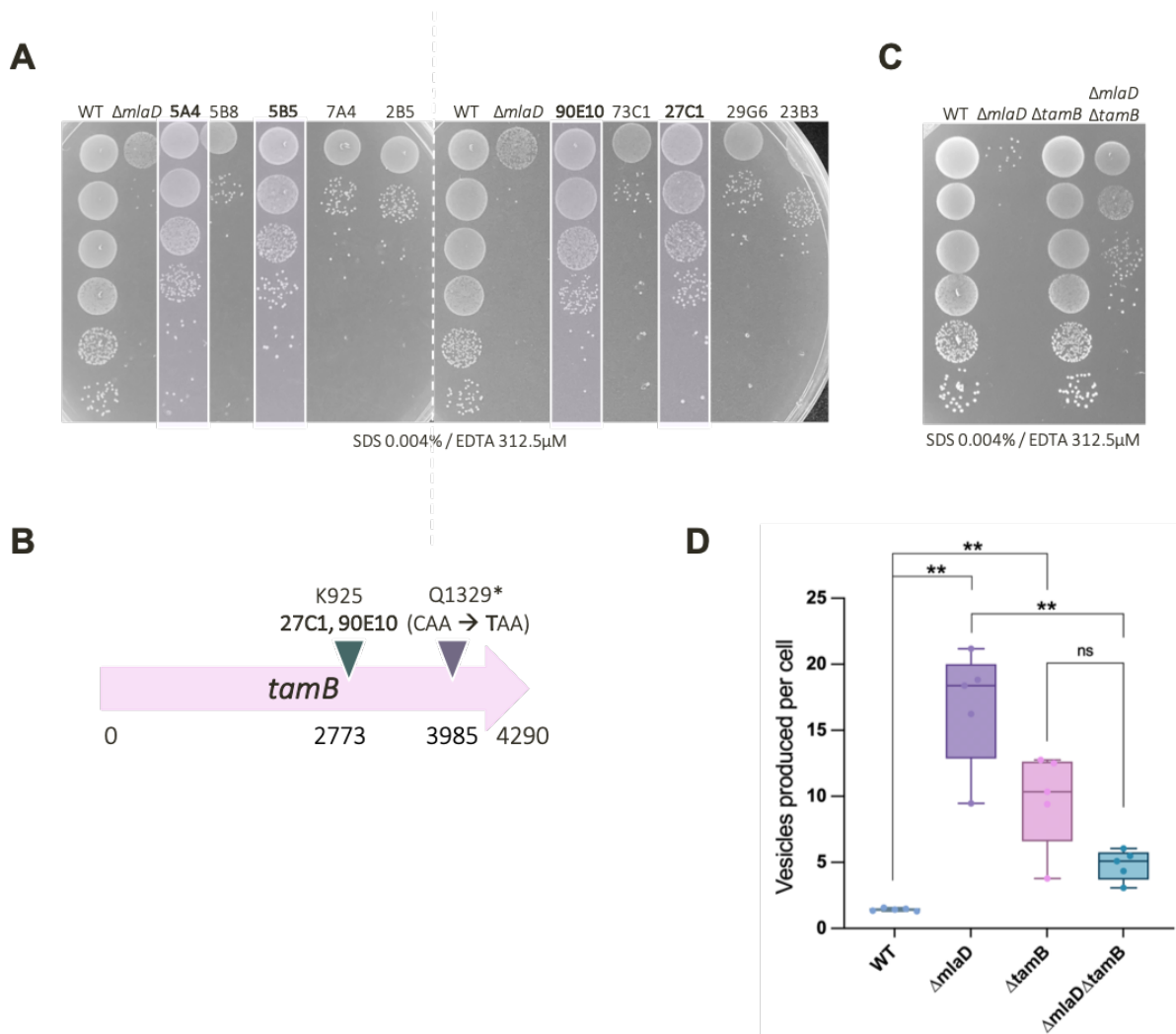


731

732 **Figure 1: Phenotypic characterisation of Δmla mutants reveals detergent hypersensitivity and**
 733 **hypervesiculation.**

734

735 **A) Efficiency of plating assay shows SDS / EDTA sensitivity of Δmla deletion strains.** Overnight
 736 cultures were adjusted for OD₆₀₀ and serially diluted onto SK media supplemented with SDS / EDTA
 737 and vancomycin. Δmla mutants display hypersensitivity to SDS / EDTA, and a strong resistance to
 738 vancomycin. Complementation of the detergent hypersensitivity was performed with an aTc-inducible
 739 vector pRPF185. **B) Hypervesiculation of Δmla deletion strains.** Outer membrane vesicles (OMVs)
 740 produced by the WT, $\Delta mlaD$, $\Delta mlaEFD$ and complemented strains were quantified via NanoFCM
 741 (see methods) and used to generate a ratio of hypervesiculation. The WT produces a similar number
 742 of vesicles to cells, resulting in a ~1:1 ratio, whilst the Δmla deletion strains produce ~17-fold more
 743 vesicles than the WT. This phenotype can be significantly rescued, reducing vesicle production to ~5-
 744 fold in the $\Delta mlaEFD::mleEFD$ strain. 5 biological replicates and 3 technical replicates were tested per
 745 strain; n = 5; significance calculated by Mann-Whitney U, p = 0.0079; two-tailed p-value. **C)**
 746 **Enrichment of phosphatidylethanolamine (PE) in OMVs produced by $\Delta mlaD$.** Thin layer
 747 chromatography (TLC) of lipid extracts from OMVs produced by both the WT and $\Delta mlaD$ strains
 748 show a ~40% (+/- 16%) enrichment for PE in the vesicles produced by the Δmla mutants when stained
 749 with iodine vapour (labelled +++, quantified by ImageJ over 3 biological replicates). We also notice
 750 a decrease in the abundance of another as yet unidentified lipid species, highlighted in the black box.
 751 Schiff reagent staining shows a ~130% increase in the abundance of plasmemyl PE in $\Delta mlaD$ OMV
 752 lipid extracts.

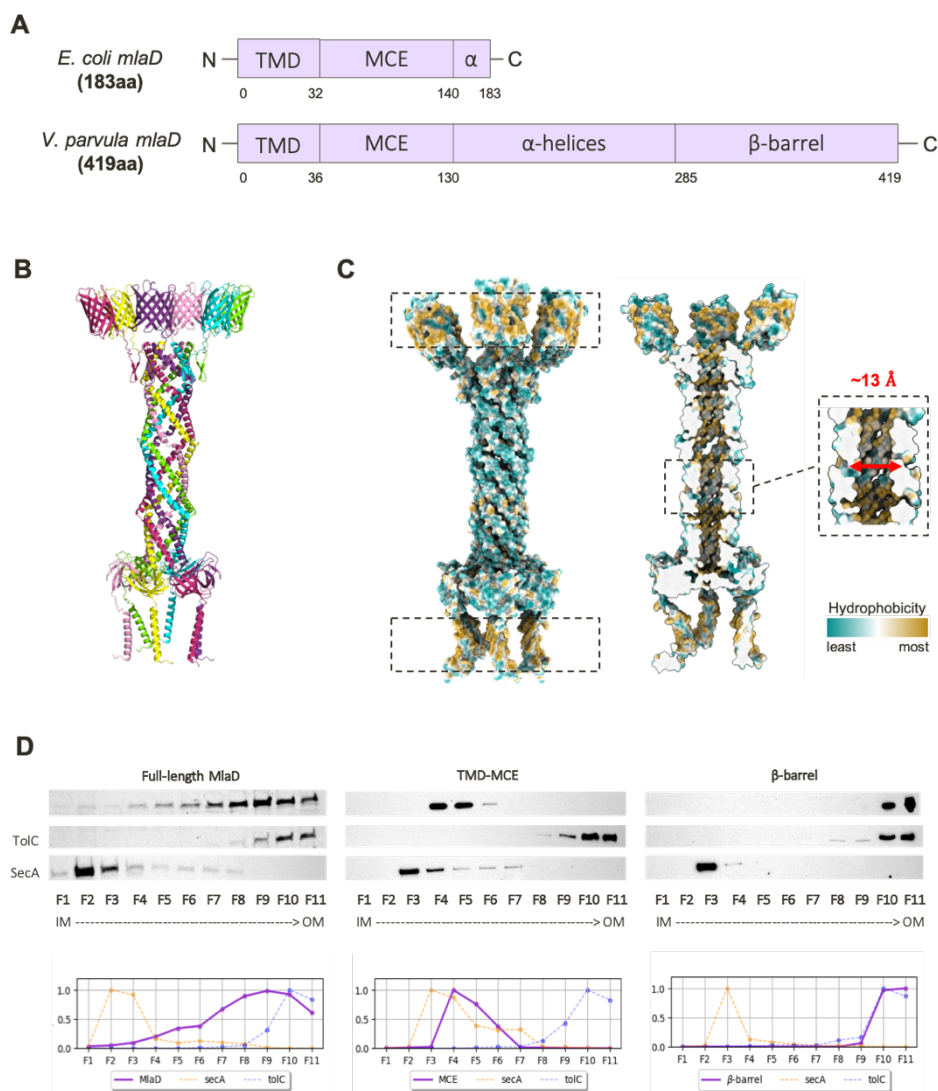


753

754 **Figure 2: TamB is a suppressor of the $\Delta mldA$ OM permeability and hypervesiculation**
 755 **phenotype.**

756

757 **A) Extent of phenotypic rescue of $\Delta mldA$ across 10 Tn mutants.** Efficiency of plating assays on
 758 SDS 0.004% + EDTA 312.5 μ M were used to determine the extent of phenotypic rescue by Tn
 759 insertions. Only 4 of the 10 Tn insertions could significantly rescue the detergent sensitivity of $\Delta mldA$:
 760 5A4, 5B5, 27C1 and 90E10. **B) Suppressor mutations identified in *tamB*.** From both spontaneous
 761 suppressor analysis, and random Tn insertion suppression analysis, we identified mutations in a
 762 homologue of *tamB*. The two identical Tn insertions were at K925, whilst the SNP in the spontaneous
 763 suppressor was mapped to Q1329*, generating a STOP codon (CAA \rightarrow TAA).
 764 **C) OM permeability of WT, $\Delta tamB$ and $\Delta mldA$ mutants.** Clean deletion mutants were generated for
 765 $\Delta tamB$, $\Delta mldA$ and $\Delta mldA \Delta tamB$, then overnight cultures of these strains and the WT were serially
 766 diluted onto SDS / EDTA to assess OM permeability. $\Delta tamB$ does not display a hypersensitivity to
 767 SDS 0.004% / EDTA 312.5 μ M, whilst $\Delta mldA$ is almost unable to grow at this concentration. Deletion
 768 of *tamB* in this $\Delta mldA$ background partially rescues the detergent hypersensitivity of $\Delta mldA$. **D) Outer**
 769 **Membrane Vesicle (OMV) production of WT, $\Delta tamB$ and $\Delta mldA$ mutants.** OMV production of
 770 WT, $\Delta tamB$, $\Delta mldA$ and $\Delta mldA \Delta tamB$ strains were analysed via NanoFCM quantification. At least 5
 771 biological replicates and 2-3 technical replicates were tested per strain. Whilst $\Delta mldA$ and $\Delta tamB$
 772 display hypervesiculation, the double mutant ($\Delta mldA \Delta tamB$) shows a significant reduction in OMV
 773 production as compared to $\Delta mldA$; n = 5; significance calculated by Mann Whitney U, p = 0.0079;
 774 two-tailed p-value. This reduction in hypervesiculation is non-significant between $\Delta tamB$ and
 775 $\Delta mldA \Delta tamB$ (p = 0.0952). WT and $\Delta mldA$ OMV production data previously shown in Fig 2 above.

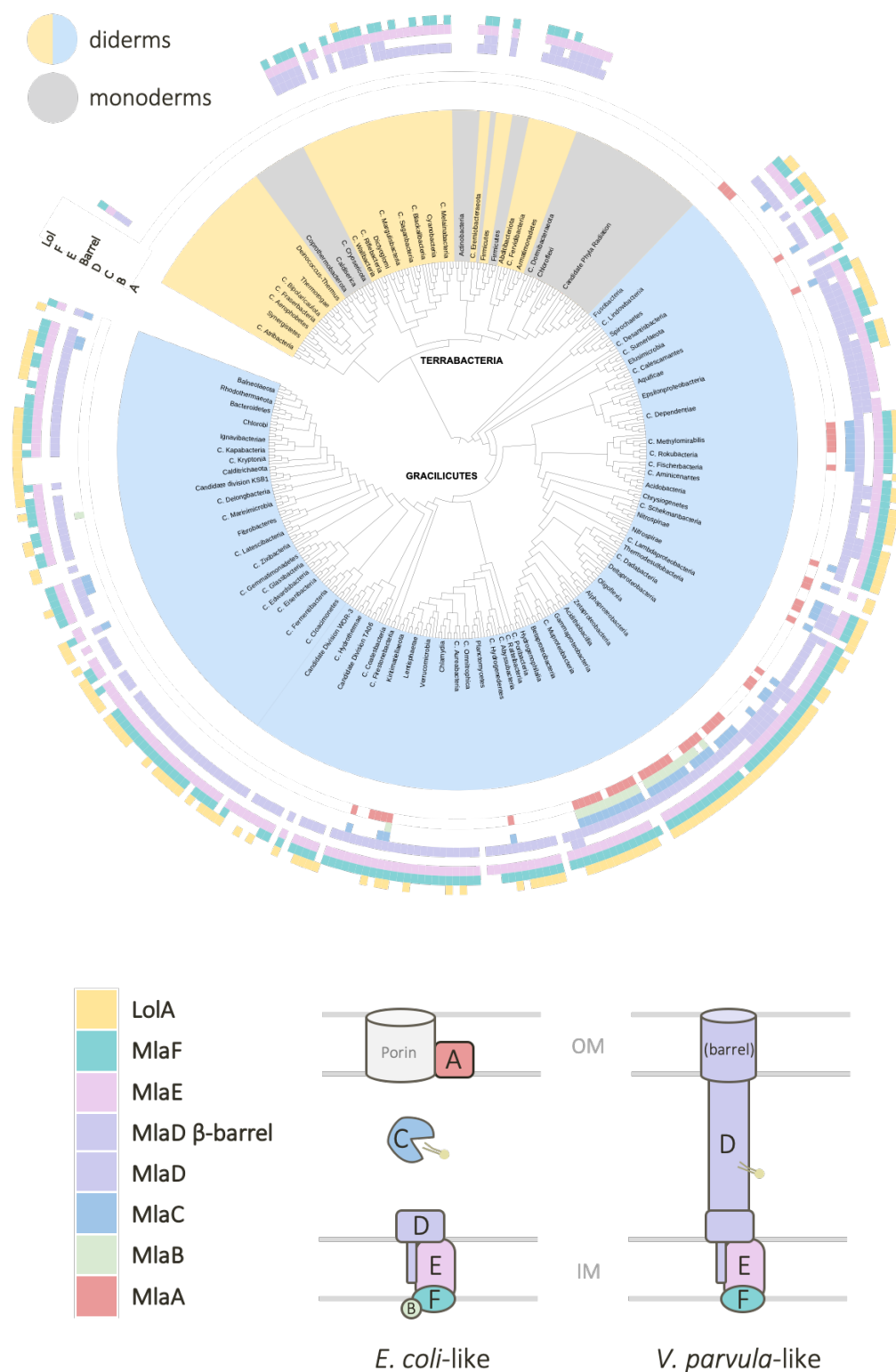


776

777 **Figure 3: Structural modelling and experimental localisation of a transenvelope MlaD in *V.***
 778 ***parvula*.**

779

780 **A) Domain comparison of MlaD in *E. coli* and *V. parvula*.** MlaD in *E. coli* consists of 183 residues
 781 and two major domains: a transmembrane domain (TMD) and an MCE domain, with a small alpha-
 782 helical region at its C-terminus. MlaD in *V. parvula* is 419 residues, consisting of a TMD, MCE
 783 domain, extended alpha-helical region and a β-barrel at its C-terminus. **B) Merged hexameric model**
 784 **of full-length MlaD.** Separate AF2 models of each domain of MlaD were overlapped to create a
 785 merged, full-length hexameric model of MlaD. **C) Hydrophobicity of the full-length hexameric**
 786 **MlaD model.** The full-length MlaD model displays two strikingly hydrophobic regions, at both the
 787 N- and C-termini, which are predicted to both be membrane-embedded domains. The tunnel of this
 788 model displays a hydrophobic interior, with a diameter of around ~13Å. **D) Subcellular localisation**
 789 **of MlaD domains via sucrose-density gradient membrane fractionation.** The WT strain, and
 790 Δ*mldD* complemented with the TMD-MCE domain / β-barrel domain with a C-ter HA-tag, were used
 791 for membrane fractionation via sucrose gradient sedimentation. Fractions were visualised with the
 792 native MlaD antisera (WT) or anti-HA antibody (MCE / β-barrel strains). The full-length protein is
 793 present in all membrane fractions, and most abundant in F7-10. The TMD-MCE construct localises to
 794 the IM, most abundant in F4-5, whilst the barrel localises almost exclusively to the final fractions,
 795 F10-11, mirroring the localisation profile of the OM control, TolC. Relative fluorescence intensity of
 796 these localisation profiles are shown below the immunoblot.



797

798 **Figure 5: Taxonomic distribution of MlaABCDEF and LolA across Tree of Bacteria.**

799

800 The distribution of all six Mla components and of LolA as a proxy for the Lol system was mapped
 801 onto a reference phylogeny of Bacteria taken from ¹⁶ (for a more detailed version of the tree, see
 802 Extended Fig 6). In the inner ring of the tree, diderms are represented by light blue and yellow, whilst
 803 the monoderms are colored in grey. On the outer ring of the tree, each colored block corresponds to

804 the presence or absence of the six Mla components, A-F, and of LolA. Two purple blocks are used to
805 represent MlaD; the first block represents presence / absence of the gene itself, and the second block
806 represents the presence / absence of the β -barrel domain. From this tree, we can see MlaEFD are
807 widely distributed across the bacteria kingdom and were likely present in the LBCA. In contrast,
808 MlaABC are sparsely distributed and restricted mostly to the Proteobacteria, suggesting these
809 components evolved later (and notably the lipoprotein MlaA with the emergence of the Lol system).
810 We also see that, in general, MlaD versions with a β -barrel tend to co-occur with the absence of MlaA
811 and MlaC. (Yellow / Grey = Terrabacteria; Blue = Gracilicutes). Tree generated using custom made
812 scripts and iTOL⁶⁵. For detailed species names, see Extended Fig 6.

813 **References**

814

815 1. Nikaido, H. Molecular Basis of Bacterial Outer Membrane Permeability Revisited. *Microbiol. Mol.*
816 *Biol. Rev.* **67**, 593–656 (2003).

817 2. Silhavy, T. J., Kahne, D. & Walker, S. The Bacterial Cell Envelope. *Cold Spring Harb. Perspect. Biol.*
818 **2**, a000414–a000414 (2010).

819 3. Ruiz, N., Davis, R. M. & Kumar, S. YhdP, TamB, and YdbH Are Redundant but Essential for
820 Growth and Lipid Homeostasis of the Gram-Negative Outer Membrane. *mBio* **12**, e02714-21 (2021).

821 4. Douglass, M. V., McLean, A. B. & Trent, M. S. Absence of YhdP, TamB, and YdbH leads to defects
822 in glycerophospholipid transport and cell morphology in Gram-negative bacteria. *PLOS Genet.* **18**,
823 e1010096 (2022).

824 5. Grimm, J. *et al.* The inner membrane protein YhdP modulates the rate of anterograde phospholipid
825 flow in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **117**, 26907–26914 (2020).

826 6. Ekiert, D. C. *et al.* Architectures of Lipid Transport Systems for the Bacterial Outer Membrane. *Cell*
827 **169**, 273-285.e17 (2017).

828 7. Malinverni, J. C. & Silhavy, T. J. An ABC transport system that maintains lipid asymmetry in the
829 Gram-negative outer membrane. *Proc. Natl. Acad. Sci.* **106**, 8009–8014 (2009).

830 8. Thong, S. *et al.* Defining key roles for auxiliary proteins in an ABC transporter that maintains
831 bacterial outer membrane lipid asymmetry. *eLife* **5**, e19042 (2016).

832 9. Tang, X. *et al.* Structural insights into outer membrane asymmetry maintenance in Gram-negative
833 bacteria by MlaFEDB. *Nat. Struct. Mol. Biol.* **28**, 81–91 (2021).

834 10. Huang, Y. M. *et al.* Molecular dynamic study of MlaC protein in Gram-negative bacteria:
835 conformational flexibility, solvent effect and protein-phospholipid binding: Molecular Dynamic
836 Study of MlaC Protein. *Protein Sci.* **25**, 1430–1437 (2016).

837 11. Chong, Z.-S., Woo, W.-F. & Chng, S.-S. Osmoporin OmpC forms a complex with MlaA to maintain
838 outer membrane lipid asymmetry in *Escherichia coli*: An outer membrane complex important for
839 lipid asymmetry. *Mol. Microbiol.* **98**, 1133–1146 (2015).

840 12. Powers, M. J., Simpson, B. W. & Trent, M. S. The Mla pathway in *Acinetobacter baumannii* has no
841 demonstrable role in anterograde lipid transport. *eLife* **9**, e56571 (2020).

842 13. Kamischke, C. *et al.* The *Acinetobacter baumannii* Mla system and glycerophospholipid transport
843 to the outer membrane. *eLife* **8**, e40171 (2019).

844 14. Hughes, G. W. *et al.* Evidence for phospholipid export from the bacterial inner membrane by the
845 Mla ABC transport system. *Nat. Microbiol.* **4**, 1692–1705 (2019).

846 15. Antunes, L. C. *et al.* Phylogenomic analysis supports the ancestral presence of LPS-outer
847 membranes in the Firmicutes. *eLife* **5**, e14589 (2016).

848 16. Witwinowski, J. *et al.* An ancient divide in outer membrane tethering systems in bacteria suggests
849 a mechanism for the diderm-to-monoderm transition. *Nat. Microbiol.* **7**, 411–422 (2022).

850 17. Megrian, D., Taib, N., Witwinowski, J., Beloin, C. & Gribaldo, S. One or two membranes? Diderm
851 Firmicutes challenge the Gram-positive/Gram-negative divide. *Mol. Microbiol.* **113**, 659–671 (2020).

852 18. Taib, N. *et al.* Genome-wide analysis of the Firmicutes illuminates the diderm/monoderm
853 transition. *Nat. Ecol. Evol.* **4**, 1661–1672 (2020).

854 19. Witwinowski, J. *et al.* An ancient divide in outer membrane tethering systems in bacteria suggests
855 a mechanism for the diderm-to-monoderm transition. *Nat. Microbiol.* **7**, 411–422 (2022).

856 20. Coleman, G. A. *et al.* A rooted phylogeny resolves early bacterial evolution. *Science* **372**, eabe0511
857 (2021).

858 21. Bavro, V. N. *et al.* Assembly and channel opening in a bacterial drug efflux machine. *Mol. Cell* **30**,
859 114–121 (2008).

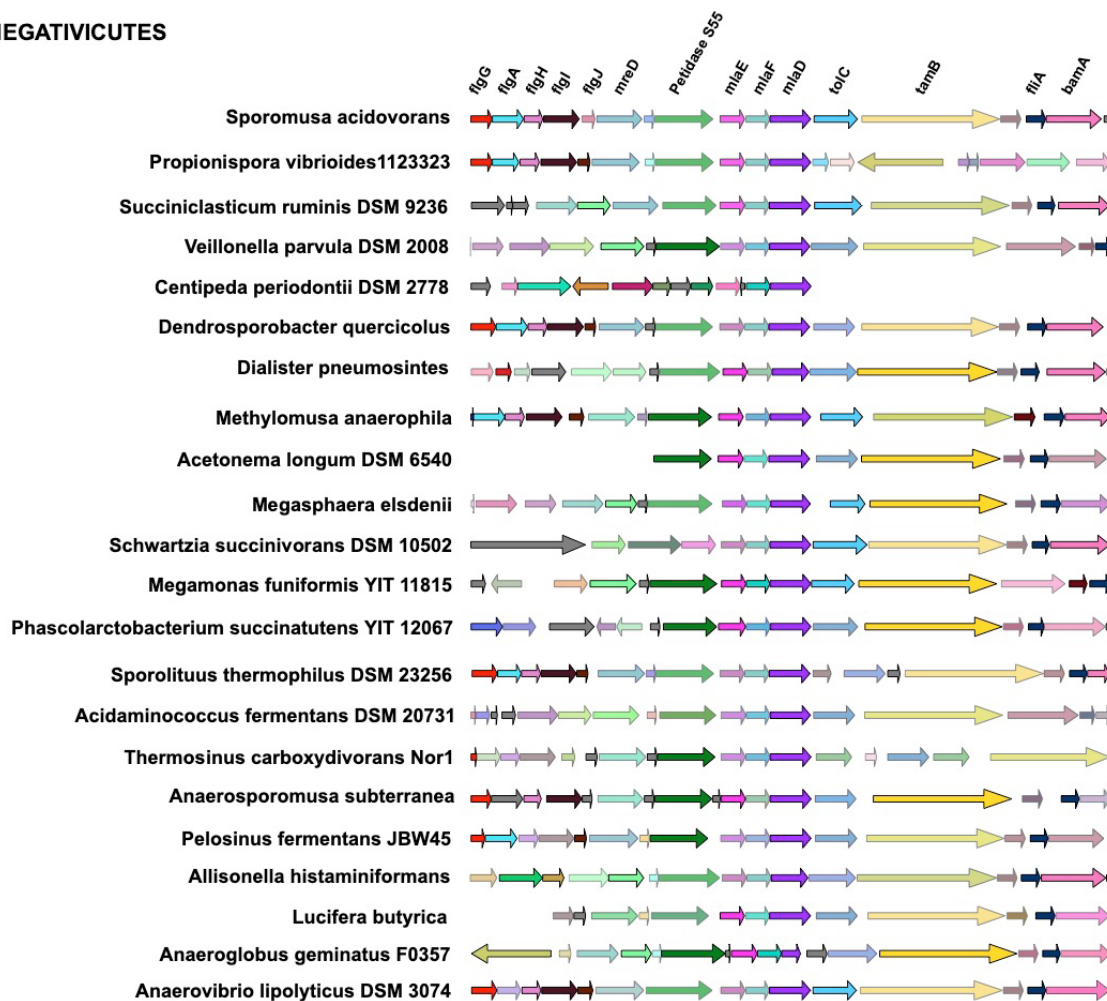
860 22. Poppleton, D. I. *et al.* Outer Membrane Proteome of *Veillonella parvula*: A Diderm Firmicute of the
861 Human Microbiome. *Front. Microbiol.* **8**, 1215 (2017).

- 862 23. Baarda, B. I., Zielke, R. A., Le Van, A., Jerse, A. E. & Sikora, A. E. Neisseria gonorrhoeae MlaA
863 influences gonococcal virulence and membrane vesicle production. *PLOS Pathog.* **15**, e1007385
864 (2019).
- 865 24. de Jonge, E. F., Vogrinec, L., van Boxtel, R. & Tommassen, J. Inactivation of the Mla system and
866 outer-membrane phospholipase A results in disrupted outer-membrane lipid asymmetry and
867 hypervesiculation in Bordetella pertussis. *Curr. Res. Microb. Sci.* 100172 (2022)
868 doi:10.1016/j.crmicr.2022.100172.
- 869 25. Roier, S. *et al.* A novel mechanism for the biogenesis of outer membrane vesicles in Gram-negative
870 bacteria. *Nat. Commun.* **7**, 10515 (2016).
- 871 26. Davies, C. *et al.* Sodium Taurocholate Stimulates Campylobacter jejuni Outer Membrane Vesicle
872 Production via Down-Regulation of the Maintenance of Lipid Asymmetry Pathway. *Front. Cell.*
873 *Infect. Microbiol.* **9**, 177 (2019).
- 874 27. Bishop, R. E. *et al.* Transfer of palmitate from phospholipids to lipid A in outer membranes of Gram-
875 negative bacteria. *EMBO J.* **19**, 5071–5080 (2000).
- 876 28. Jackson, D. R. *et al.* Plasmalogen Biosynthesis by Anaerobic Bacteria: Identification of a Two-Gene
877 Operon Responsible for Plasmalogen Production in *Clostridium perfringens*. *ACS Chem. Biol.* **16**, 6–
878 13 (2021).
- 879 29. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589
880 (2021).
- 881 30. Chen, J. *et al.* Structure of an endogenous mycobacterial MCE lipid transporter. *bioRxiv* (2022).
- 882 31. Neron, B. *et al.* MacSyFinder v2: Improved modelling and search engine to identify molecular systems in
883 genomes. <http://biorxiv.org/lookup/doi/10.1101/2022.09.02.506364> (2022)
884 doi:10.1101/2022.09.02.506364.
- 885 32. Hao, L. *et al.* Novel prosthecate bacteria from the candidate phylum Acetothermia. *ISME J.* **12**,
886 2225–2237 (2018).
- 887 33. Hoppert, M. *et al.* Structure–functional analysis of the Dictyoglomus cell envelope. *Syst. Appl.*
888 *Microbiol.* **35**, 279–290 (2012).
- 889 34. Brock, T. D. & Edwards, M. R. Fine Structure of *Thermus aquaticus*, an Extreme Thermophile. *J.*
890 *Bacteriol.* **104**, 509–517 (1970).
- 891 35. Katayama, T. *et al.* Isolation of a member of the candidate phylum ‘Atribacteria’ reveals a unique
892 cell membrane structure. *Nat. Commun.* **11**, 6381 (2020).
- 893 36. Hayat, S., Peters, C., Shu, N., Tsigos, K. D. & Elofsson, A. Inclusion of dyad-repeat pattern
894 improves topology prediction of transmembrane β -barrel proteins. *Bioinformatics* **32**, 1571–1573
895 (2016).
- 896 37. Konovalova, A. & Silhavy, T. J. Outer membrane lipoprotein biogenesis: Lol is not the end. *Philos.*
897 *Trans. R. Soc. B Biol. Sci.* **370**, 20150030 (2015).
- 898 38. Matsuyama, S., Tajima, T. & Tokuda, H. A novel periplasmic carrier protein involved in the sorting
899 and transport of Escherichia coli lipoproteins destined for the outer membrane. *EMBO J.* **14**, 3365–
900 3372 (1995).
- 901 39. Kim, K. H., Aulakh, S. & Paetzel, M. The bacterial outer membrane β -barrel assembly machinery.
902 *Protein Sci.* **21**, 751–768 (2012).
- 903 40. Josts, I. *et al.* The Structure of a Conserved Domain of TamB Reveals a Hydrophobic β Taco Fold.
904 *Structure* **25**, 1898-1906.e5 (2017).
- 905 41. Abellón-Ruiz, J. *et al.* Structural basis for maintenance of bacterial outer membrane lipid
906 asymmetry. *Nat. Microbiol.* **2**, 1616–1623 (2017).
- 907 42. Yeow, J. *et al.* The architecture of the OmpC–MlaA complex sheds light on the maintenance of outer
908 membrane lipid asymmetry in Escherichia coli. *J. Biol. Chem.* **293**, 11325–11340 (2018).
- 909 43. Sperandio, P., Martorana, A. M. & Polissi, A. The lipopolysaccharide transport (Lpt) machinery: A
910 nonconventional transporter for lipopolysaccharide assembly at the outer membrane of Gram-
911 negative bacteria. *J. Biol. Chem.* **292**, 17981–17990 (2017).

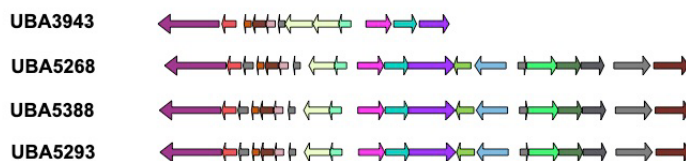
- 912 44. Low, W.-Y., Thong, S. & Chng, S.-S. ATP disrupts lipid-binding equilibrium to drive retrograde
913 transport critical for bacterial outer membrane asymmetry. *Proc. Natl. Acad. Sci.* **118**, e2110055118
914 (2021).
- 915 45. Ercan, B., Low, W.-Y., Liu, X. & Chng, S.-S. Characterization of Interactions and Phospholipid
916 Transfer between Substrate Binding Proteins of the OmpC-Mla System. *Biochemistry* **58**, 114–119
917 (2019).
- 918 46. Mann, D. *et al.* Structure and lipid dynamics in the maintenance of lipid asymmetry inner
919 membrane complex of *A. baumannii*. *Commun. Biol.* **4**, 817 (2021).
- 920 47. Knapp, S. *et al.* Natural Competence Is Common among Clinical Isolates of *Veillonella parvula* and
921 Is Useful for Genetic Manipulation of This Key Member of the Oral Microbiome. *Front. Cell. Infect.*
922 *Microbiol.* **7**, (2017).
- 923 48. Dembek, M. *et al.* High-throughput analysis of gene essentiality and sporulation in *Clostridium*
924 *difficile*. *MBio* **6**, e02383-14 (2015).
- 925 49. Fagan, R. P. & Fairweather, N. F. *Clostridium difficile* has two parallel and essential Sec secretion
926 systems. *J. Biol. Chem.* **286**, 27483–27493 (2011).
- 927 50. Letunic, I., Khedkar, S. & Bork, P. SMART: recent updates, new developments and status in 2020.
928 *Nucleic Acids Res.* **49**, D458–D460 (2021).
- 929 51. Zhang, Y. I-TASSER: Fully automated protein structure prediction in CASP8. *Proteins Struct. Funct.*
930 *Bioinforma.* **77**, 100–113 (2009).
- 931 52. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. Predicting transmembrane protein
932 topology with a hidden markov model: application to complete genomes¹¹Edited by F. Cohen. *J.*
933 *Mol. Biol.* **305**, 567–580 (2001).
- 934 53. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682
935 (2022).
- 936 54. Evans, R. *et al.* Protein complex prediction with AlphaFold-Multimer.
937 <http://biorxiv.org/lookup/doi/10.1101/2021.10.04.463034> (2021) doi:10.1101/2021.10.04.463034.
- 938 55. Pettersen, E. F. *et al.* UCSF Chimera?A visualization system for exploratory research and analysis.
939 *J. Comput. Chem.* **25**, 1605–1612 (2004).
- 940 56. Webb, B. & Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc.*
941 *Bioinforma.* **54**, (2016).
- 942 57. Steinegger, M. *et al.* HH-suite3 for fast remote homology detection and deep protein annotation.
943 *BMC Bioinformatics* **20**, 473 (2019).
- 944 58. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: Pattern recognition of
945 hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983).
- 946 59. Srinivasan, R. & Rose, G. D. A physical basis for protein secondary structure. *Proc. Natl. Acad. Sci.*
947 **96**, 14258–14263 (1999).
- 948 60. Johnson, L. S., Eddy, S. R. & Portugaly, E. Hidden Markov model speed heuristic and iterative
949 HMM search procedure. *BMC Bioinformatics* **11**, 1–8 (2010).
- 950 61. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7:
951 improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
- 952 62. Criscuolo, A. & Gribaldo, S. BMGE (Block Mapping and Gathering with Entropy): a new software
953 for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol.*
954 *Biol.* **10**, 1–21 (2010).
- 955 63. Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective
956 stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274
957 (2015).
- 958 64. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A. & Jermini, L. S. ModelFinder: fast
959 model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
- 960 65. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments.
961 *Nucleic Acids Res.* **47**, W256–W259 (2019).
- 962

Extended Data

NEGATIVICUTES

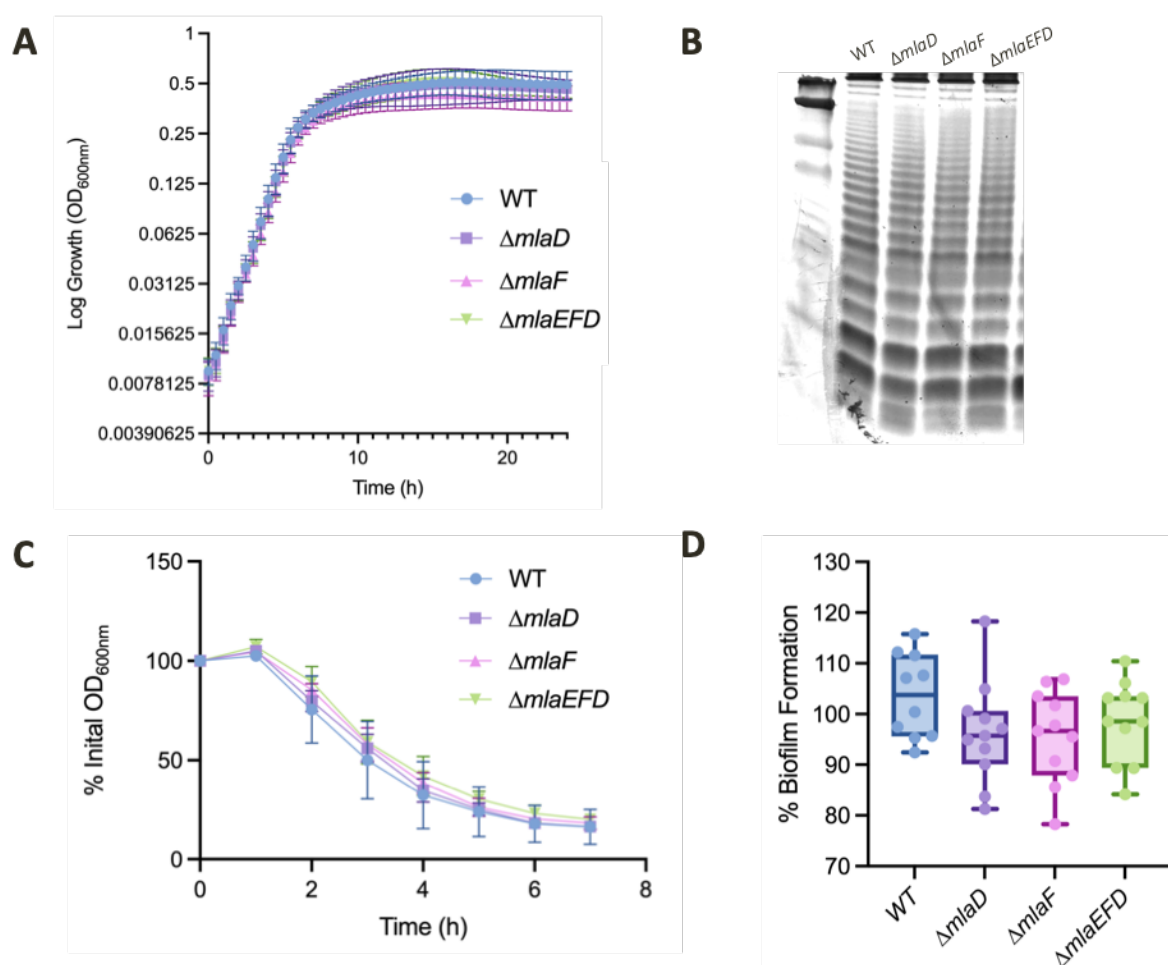


LIMNOCHORDIA



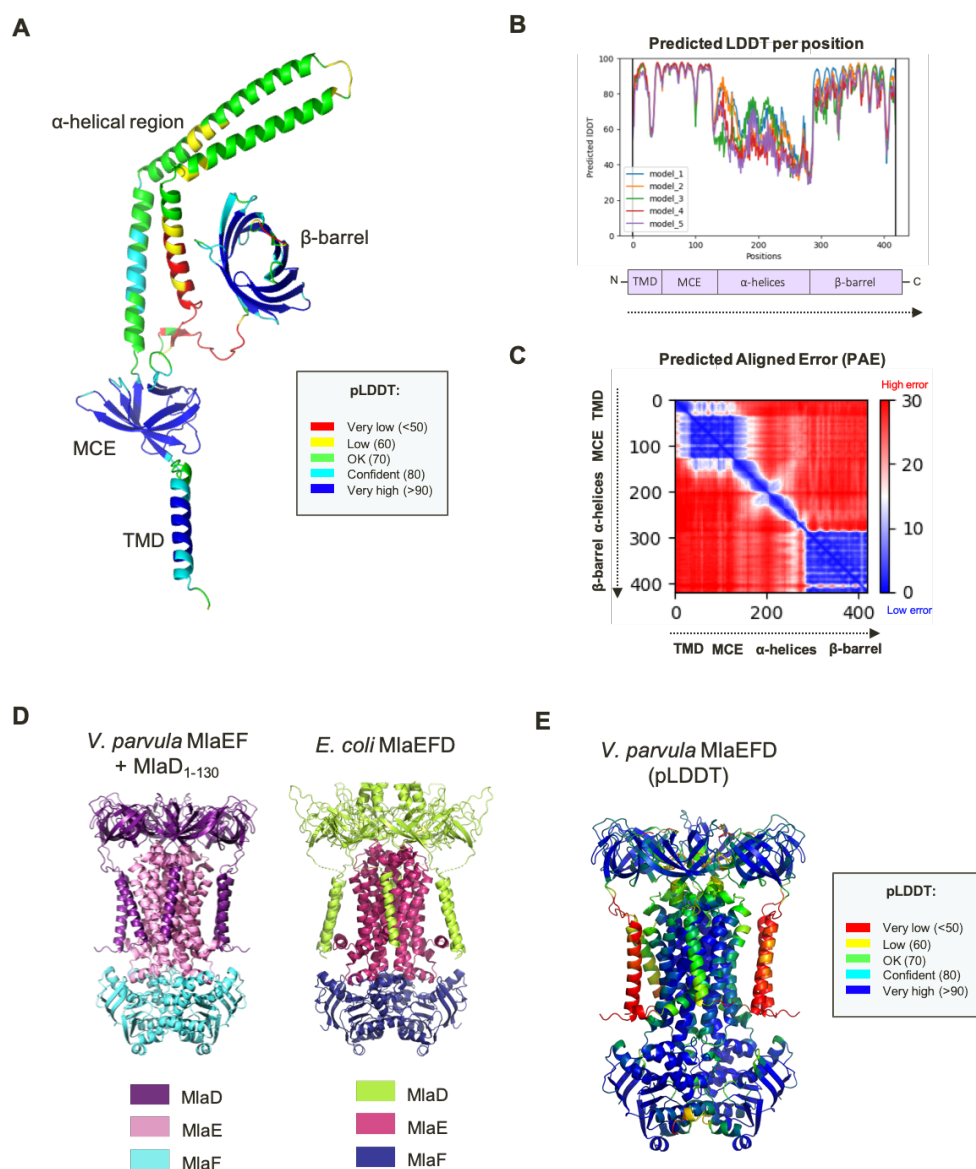
Extended Figure 1: Conserved synteny of *mla* genes with *tamB* and *tolC* in the Negativicutes, and presence of *mla* genes in Limnochordia

Within the Negativicutes, the three *mla* homologues, *mlaEFD*, are directly followed by a homologue of *tolC* and a homologue of *tamB*. This synteny is not conserved within the other diderm Firmicutes; no *mla* homologues were identified in the Halanaerobiales, whilst *mla* genes in the Limnochordia are spare and not in synteny with *tolC* and *tamB* homologues.



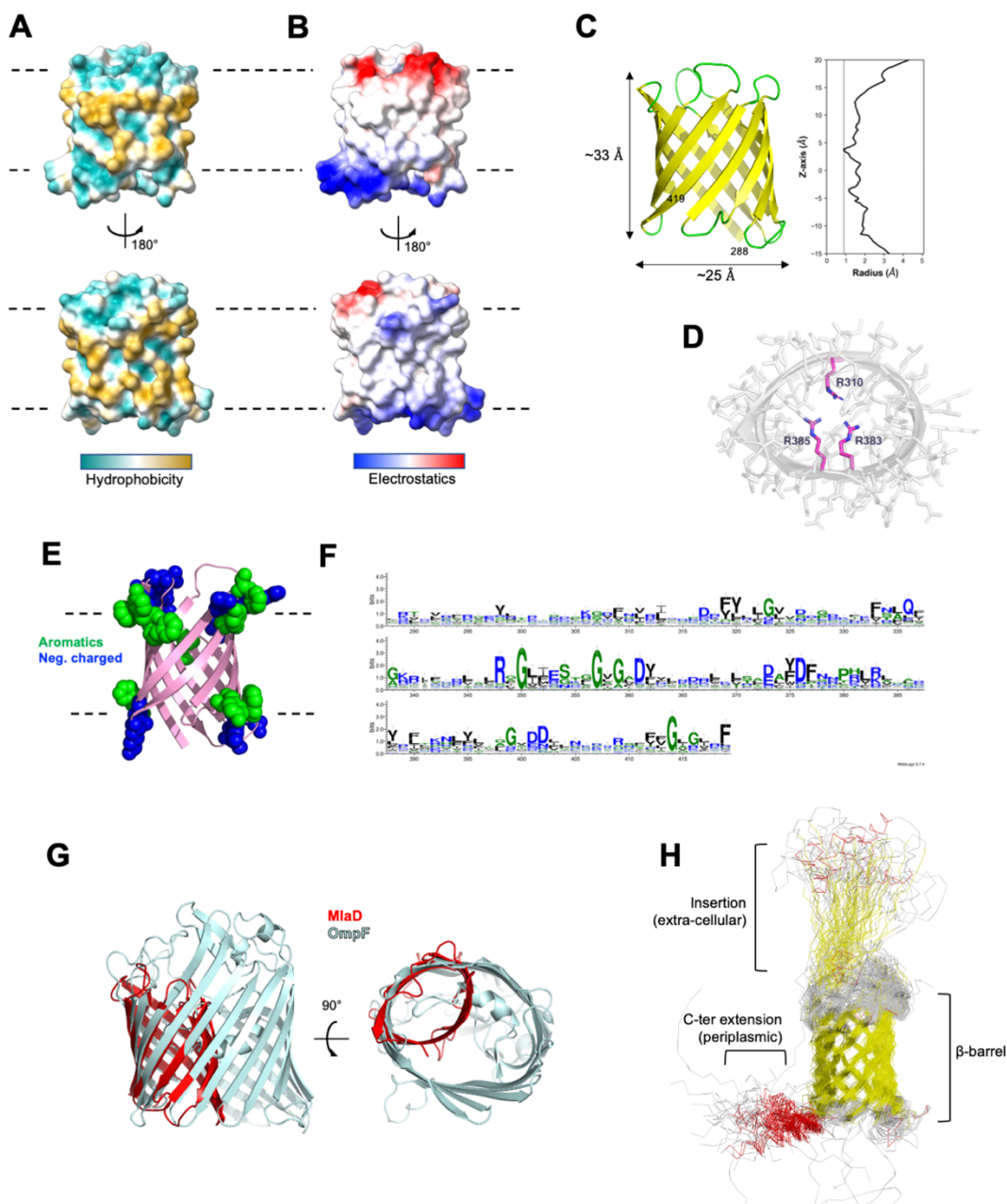
Extended Figure 2: Phenotypes unchanged by deletion of *mla* genes.

A) Comparison of growth kinetics of WT and Δmla mutants. All strains were assessed for growth kinetics via OD₆₀₀ measurements via TECAN spectrophotometer (see methods). $\Delta mlaD$, $\Delta mlaF$ and $\Delta mlaEFD$ were found to display similar growth rates to the WT, with no defects. At least 3 biological replicates and 3 technical replicates were performed per strain. **B) Comparison of LPS profiles.** LPS was extracted from WT and all Δmla cultures, in triplicate, normalised by OD₆₀₀, then observed with silver staining. All strains display similar LPS profiles, in both banding pattern and quantity. **C) Comparison of autoaggregation capabilities.** OD₆₀₀ was recorded for WT and Δmla strains once an hour for a period of seven hours. All strains display a similar decrease in OD₆₀₀ over time, suggesting a similar capability of autoaggregation. At least 3 biological replicates and 3 technical replicates were performed per strain. **D) Comparison of biofilm formation.** Biofilm formation was assessed via crystal violet staining of biofilms formed over 24h in 96-well plates. Δmla strains display a small but non-significant decrease in biofilm formation capability, as assessed by a decrease in fluorescence signal at A575nm. Due to variability in this technique, 10 biological replicates with 3 technical replicates each were tested per strain.



Extended Figure 3: AlphaFold2 modelling of MlaD in *V. parvula*

A) Predicted structure of full-length MlaD. When the whole sequence is processed by AlphaFold2, the TMD, MCE and β -barrel domain are of very high confidence, and the α -helical domain is of lower confidence. The confidence rating is indicated by Predicted Local Distance Different Test (pLDDT) colouring. **B) Predicted Local Distance Difference Test (pLDDT) per position of full-length MlaD.** Graphical depiction of the pLDDT per position of the full-length models shows low confidence for the α -helical region, and high confidence for the TMD, MCE and β -barrel domains. **C) Predicted Aligned Error (PAE) of full-length MlaD.** Graphical depiction of the PAE shows the confidence in the relative positioning of the domains. The relative positioning of main chain atoms within the TMD and MCE domains and within the connecting α -helical region are low error, suggesting the predicted local conformation of these individual domains to be likely. However, the high predicted alignment error of atoms across these domain regions indicates that the global conformation of full-length MlaD is likely inaccurate. **D) Modelling of MlaEFD complex in *V. parvula* with a 2:2:6 stoichiometry.** MlaEF modelled as dimers in complex with hexameric MlaD (residues 1-130) closely resemble the resolved structure of the MlaEFD complex in *E. coli*. **E) pLDDT per position of MlaEFD in *V. parvula*** shows high confidence with the exception of 2 pairs of TM helices.



Extended Figure 4: The C-terminal domain of MlaD₂₈₈₋₄₁₉ folds as a membrane β -barrel.

A) Hydrophobicity potential of the MlaD C-terminal β -barrel surface. The bottom view is rotated 180°. Dotted lines represent the possible position of GPL polar heads when inserted in a lipid bilayer.

B) Electrostatics potential of the MlaD C-terminal β -barrel surface. The bottom view is rotated 180°.

C) Dimensions of the β -barrel (right) and radius profile (calculated using the program HOLE) along the Z-axis (membrane normal). The height of the barrel matches the thickness of a lipid bilayer.

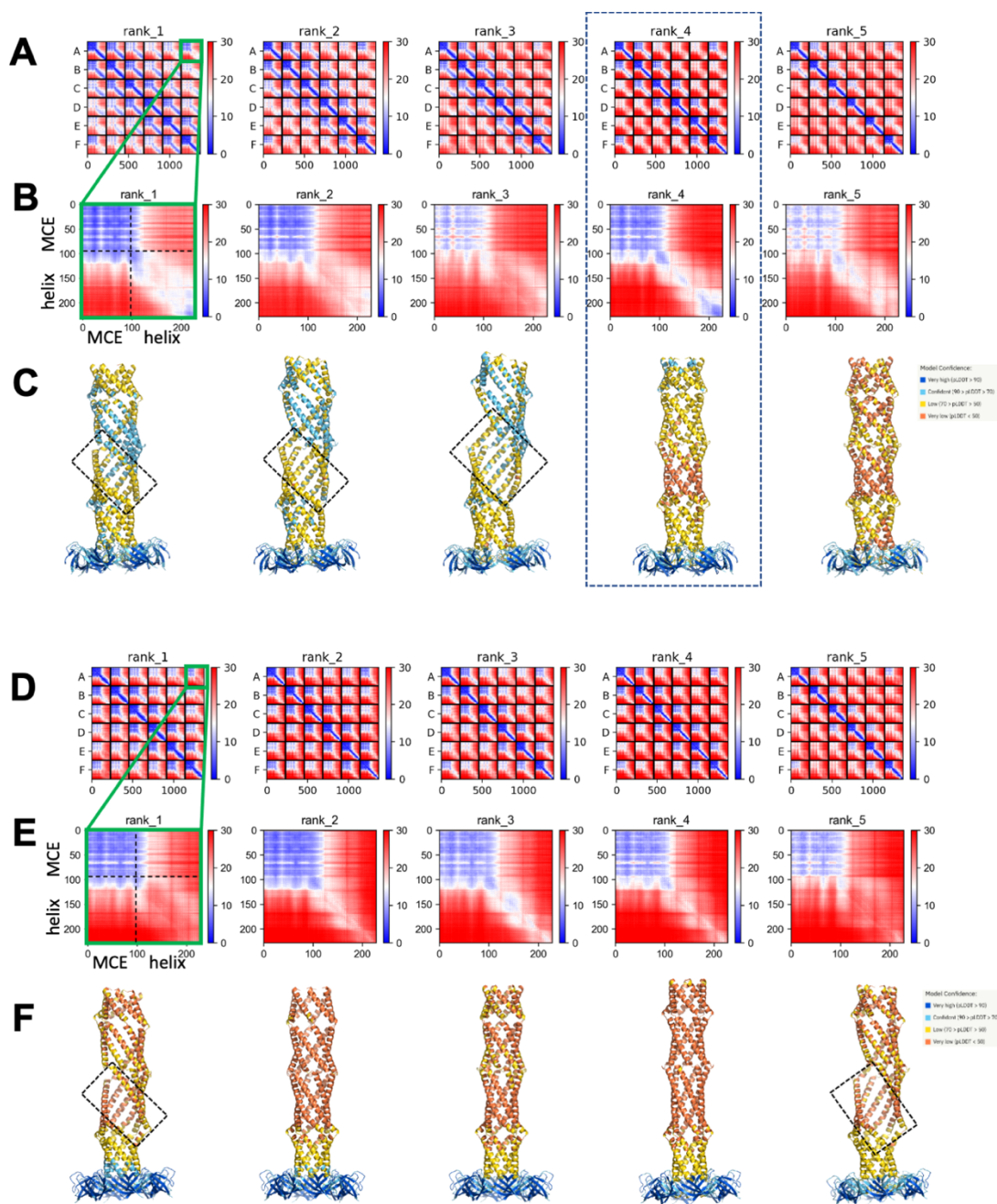
D) Position of side-chains of R310, R383 et R385 at the constriction point of the β -barrel.

E) Rings of surface accessible negatively charged and aromatics residues at each side of the β -barrel.

F) Weblogo of MlaD C-terminal β -barrel sequences coloured by hydrophobicity (Hydrophilic in blue, neutral in green and hydrophobic in black).

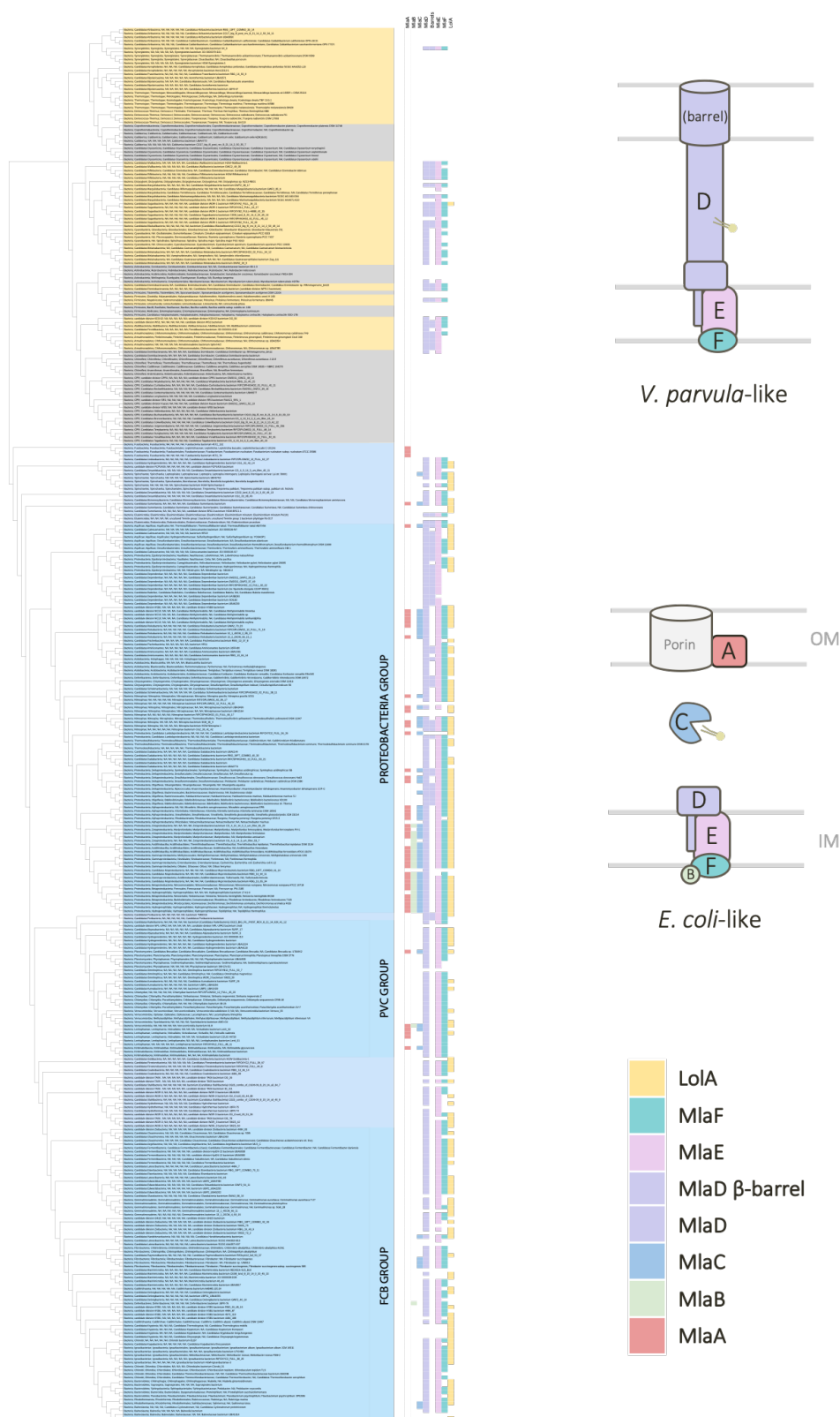
G) Superimposition of the *V. parvula* MlaD C-terminal β -barrel model (red) onto the *S. marcescens* OmpF structure (light cyan). Its small diameter, compared to large porins such as OmpF, and very narrow inner radius (~ 1 Å) would likely prevent the passage of any small molecule.

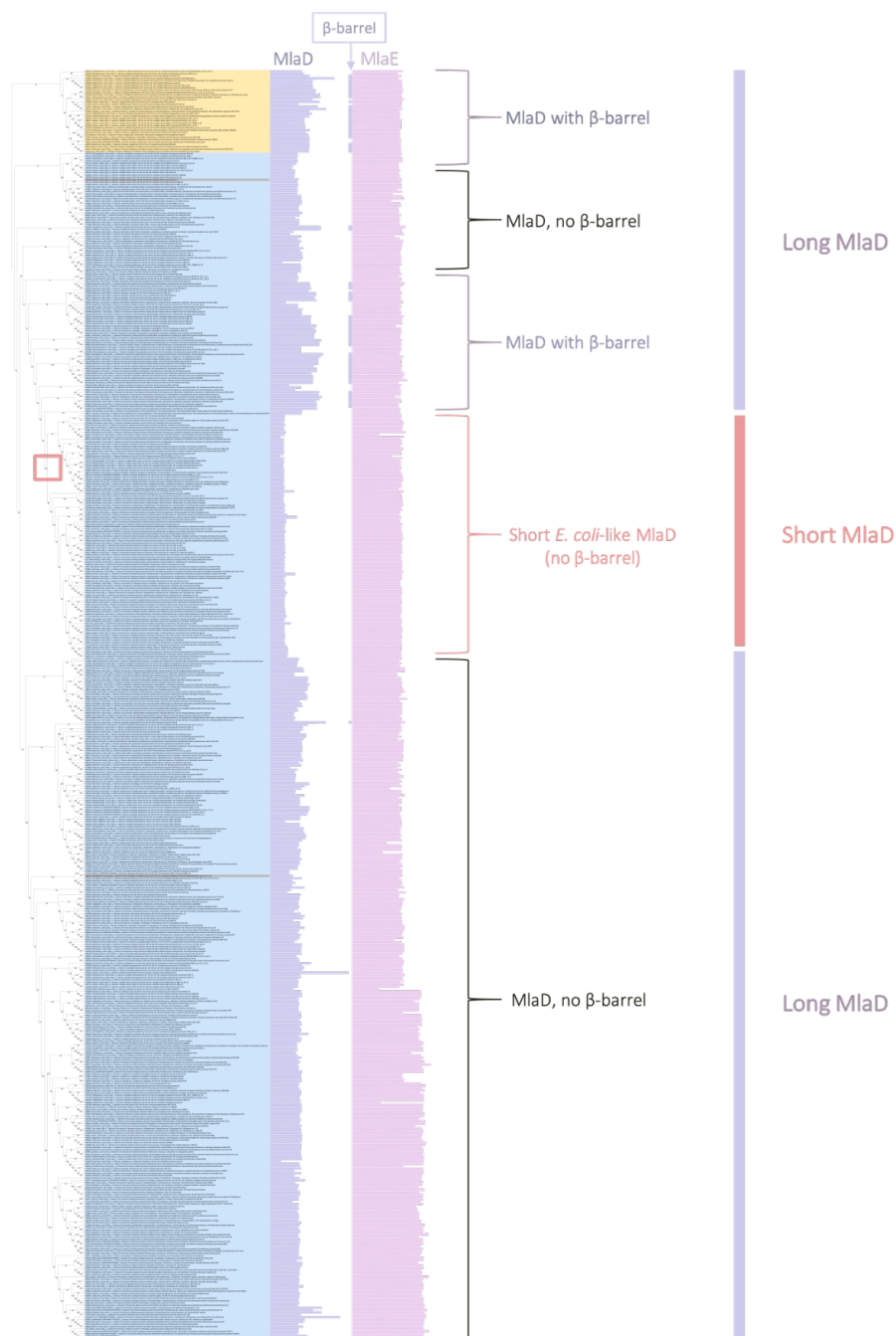
H) Superimposed ensembles of MlaD C-terminal β -barrel models for long MlaD sequences predicted with AlphaFold2. Helices are coloured in red and β -strands in yellow. Insertions in the barrel and C-terminal helical extensions found in some sequences are labelled.



Extended Figure 5: AlphaFold2 models for the alpha-helical domain of MlaD₃₆₋₂₆₃ in hexameric configuration.

A) Predicted Aligned error (PAE) for the 5 ranked models obtained with AlphaFold-Multimer v2.2. The regions corresponding to each chain in the hexamer (from A to G) are labeled. **B)** PAE between the first (A) and last (G) chain for the 5 ranked models obtained with AlphaFold-Multimer v2.2. The regions corresponding the MCE and helical domains are labeled. **C)** Models obtained AlphaFold-Multimer v2.2, colored by predicted local distance difference test (pLDDT). **D)** Predicted Aligned error (PAE) for the 5 ranked models obtained with AlphaFold-Multimer v2.3. The regions corresponding to each chain in the hexamer (from A to G) are labeled. **E)** PAE between the first (A) and last (G) chain for the 5 ranked models obtained with AlphaFold-Multimer v2.3. The regions corresponding the MCE and helical domains are labeled. **F)** Models obtained AlphaFold-Multimer v2.3 colored by predicted local distance difference test (pLDDT).

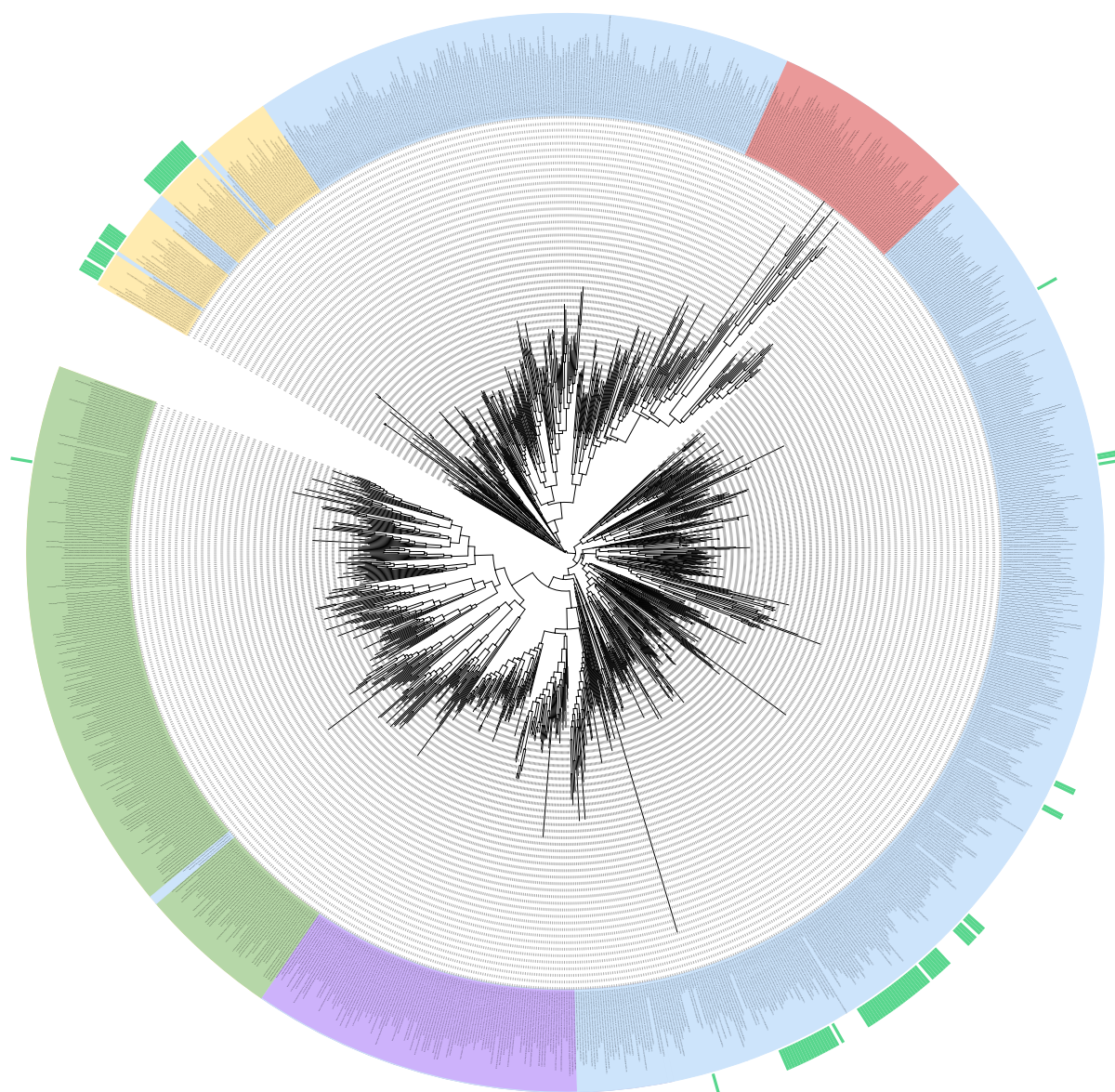




Extended Figure 7: Phylogeny concatenation of MlaE with the three types of MlaD: short, long and long with β -barrel

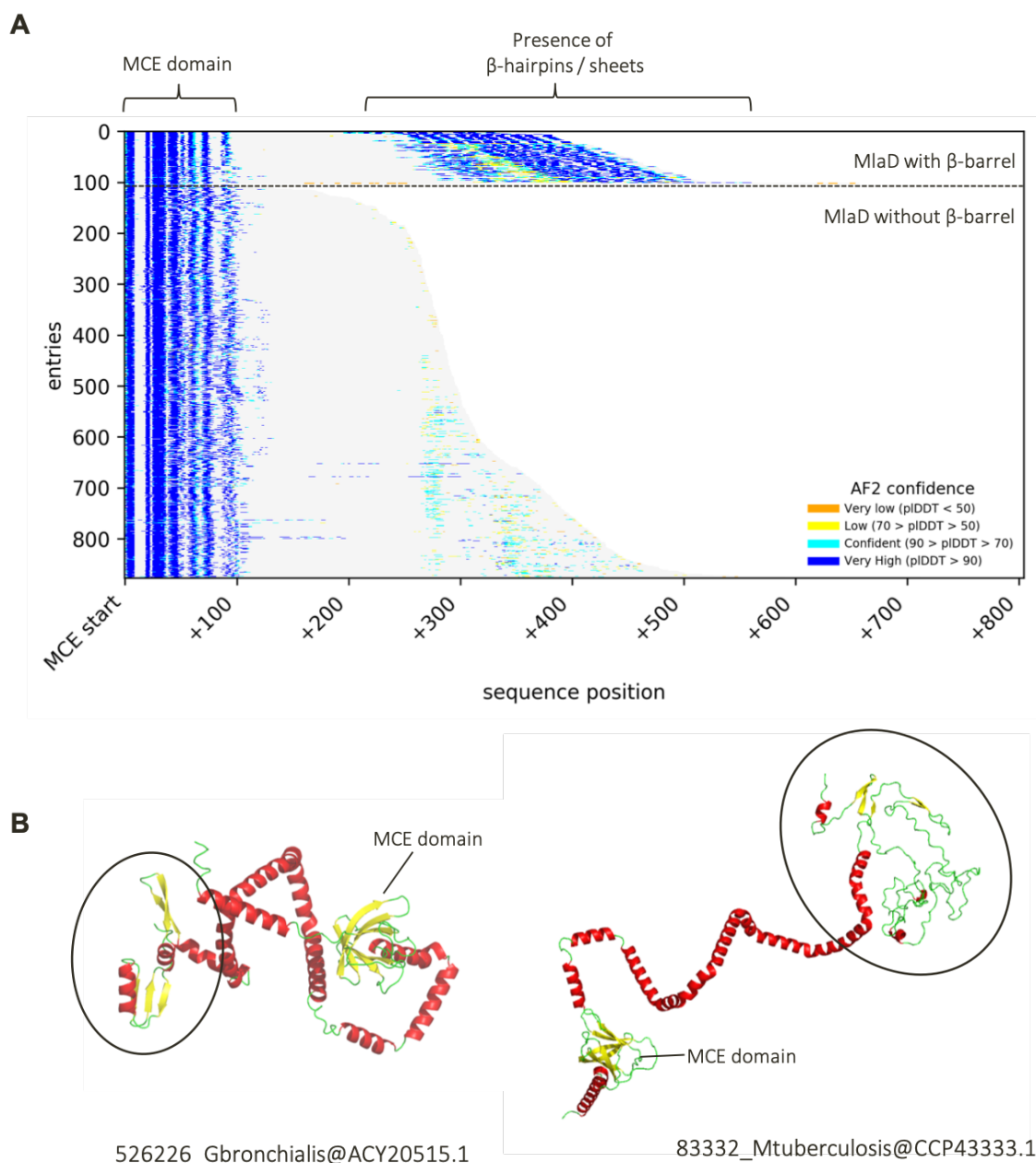
Sequences of MlaE and MlaD were concatenated into a character supermatrix (538 sequences and 479 amino acid positions) and a maximum likelihood tree was inferred. The length of MlaD sequence is indicated by the length of the purple bars, and the presence / absence of the β -barrel is depicted with a small purple at the end of the sequence. MlaE sequences are represented in pink. We clearly see from this analysis that the majority of MlaD sequences are long, and that the short MlaD sequences are restricted mostly to the Proteobacteria. We also see a clear separation between the Terrabacteria (yellow) and the Gracilicutes (blue), with monophyly of phyla within the Terrabacteria, indicating vertical inheritance of these Mla genes. Highlighted in the red box is the branch at which loss of this elongated form of MlaD seems to have occurred.

Tree scale: 1 



Extended Figure 8: Distribution of MCE-containing sequences across Bacteria

Maximum likelihood tree built from an alignment of 1119 sequences and 194 amino acid positions. The scale bar corresponds to the average number of substitutions per site. MlaD sequences are marked in green for Actinobacteria and C. Dormibacteraeota, yellow for other Terrabacteria and blue for Gracilicutes. We highlighted the short version of MlaD in purple, and PqiB and LetB sequences in red. The presence of a β -barrel is highlighted by a green bar at the outer ring of the tree.



Extended Figure 9: Presence of disordered β -structures in long MlaD sequences with no β -barrel

A) Presence of β -structures in long MlaD sequences with no β -barrel. The start of the sequence represents the presence of the MCE domain, to align all long MlaD sequences. The total length of each sequence is represented by the grey shaded area. Later, in some of these long sequences, we see the presence of β -hairpins / sheets, but no real folded β -barrel domain. For comparison, we also show the ~ 100 sequences of long MlaD that do possess a predicted folded β -barrel. Predicted Local Distance Difference Test (pLDDT) values, corresponding to the confidence of these predictions, are coloured and labelled in the key above. **B) AlphaFold modelling of long MlaD sequences with no β -barrel.** As seen from these two predicted structures, these long MlaD sequences do not possess a β -barrel, but do possess unstructured, disordered regions at their C-termini that contain fragments of β -hairpins / sheets.

Extended Table 1: Statistics of AF2-Multimer models generated for MlaD36-263 hexamers

Model rank	AF-multimer v2.2					AF -multimer v2.3				
	1	2	3	4	5	1	2	3	4	5
Open/Closed	Open	Open	Open	Closed	Closed	Open	Closed	Closed	Closed	Open
pLDDT	75.38	75.57	73.79	68.07	62.59	68.23	64.88	65.96	64.24	64.73
pTM	0.56	0.55	0.50	0.47	0.46	0.58	0.54	0.53	0.49	0.45
ipTM	0.53	0.52	0.48	0.44	0.43	0.56	0.52	0.50	0.46	0.42
multimer	0.54	0.53	0.48	0.45	0.44	0.56	0.52	0.51	0.47	0.43
Contacts*	187/ 20	180/ 0	190/ 0	139/ 86	126/ 83	185/ 3	101/ 99	143/ 151	134/ 126	154/ 5
PAE < 8 Å [#]	3431/ 72	1378/ 0	1633/ 0	350/ 102	16/ 0	180/ 0	95/ 24	43/ 0	4/ 0	0/ 0

* Number of contacts ($C\beta-C\beta < 8 \text{ \AA}$) for the helical regions (MlaD₁₃₂₋₂₆₃) between neighboring chains. The first number is the average of all $i, i+1$ pairs of chains except between the 1st and the last chain (i being the index of the chain in the hexamer). The second number (in bold) is between the 1st and the last chain in the hexamer.

[#] Number of residue pairs with a Predicted Aligned Error (PAE) < 8 Å for the helical regions (MlaD₁₃₂₋₂₆₃) between neighboring chains. The first number is the average of all $i, i+1$ pairs of chains except between the 1st and the last chain (i being the index of the chain in the hexamer). The second number (in bold) is between the 1st and the last chain in the hexamer.

Extended Table 2: Distribution and accession numbers of Mla components homologs among 1083 genomes representing the bacterial diversity.

The presence of MlaB and MlaF is only indicated when they are in cluster with at least one of the other components of the Mla system.

Table attached separately.

References associated with Extended Data

1. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments.

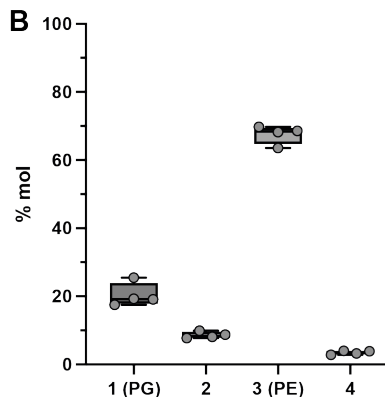
Nucleic Acids Res. **47**, W256–W259 (2019).

Supplementary Data

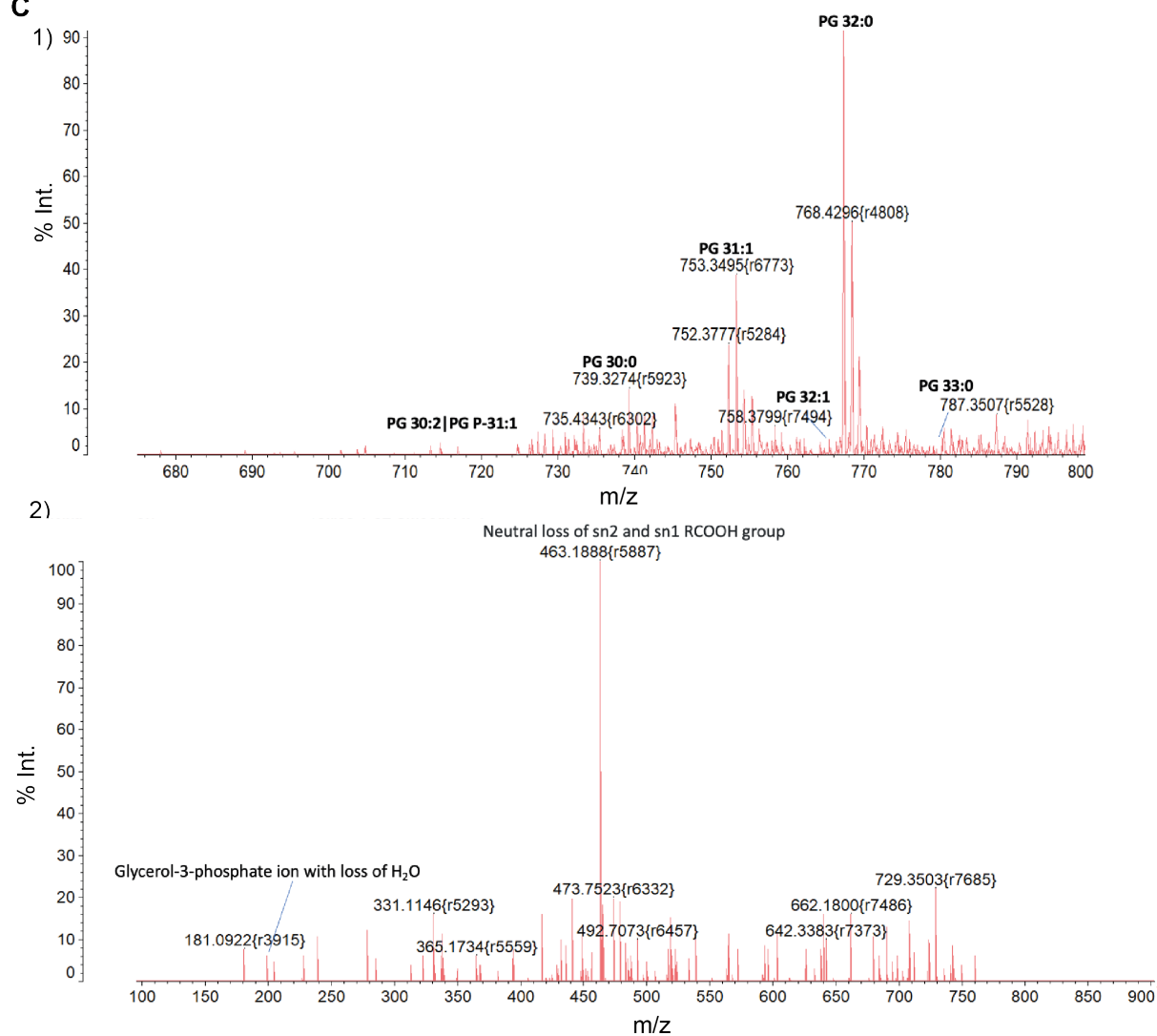
A

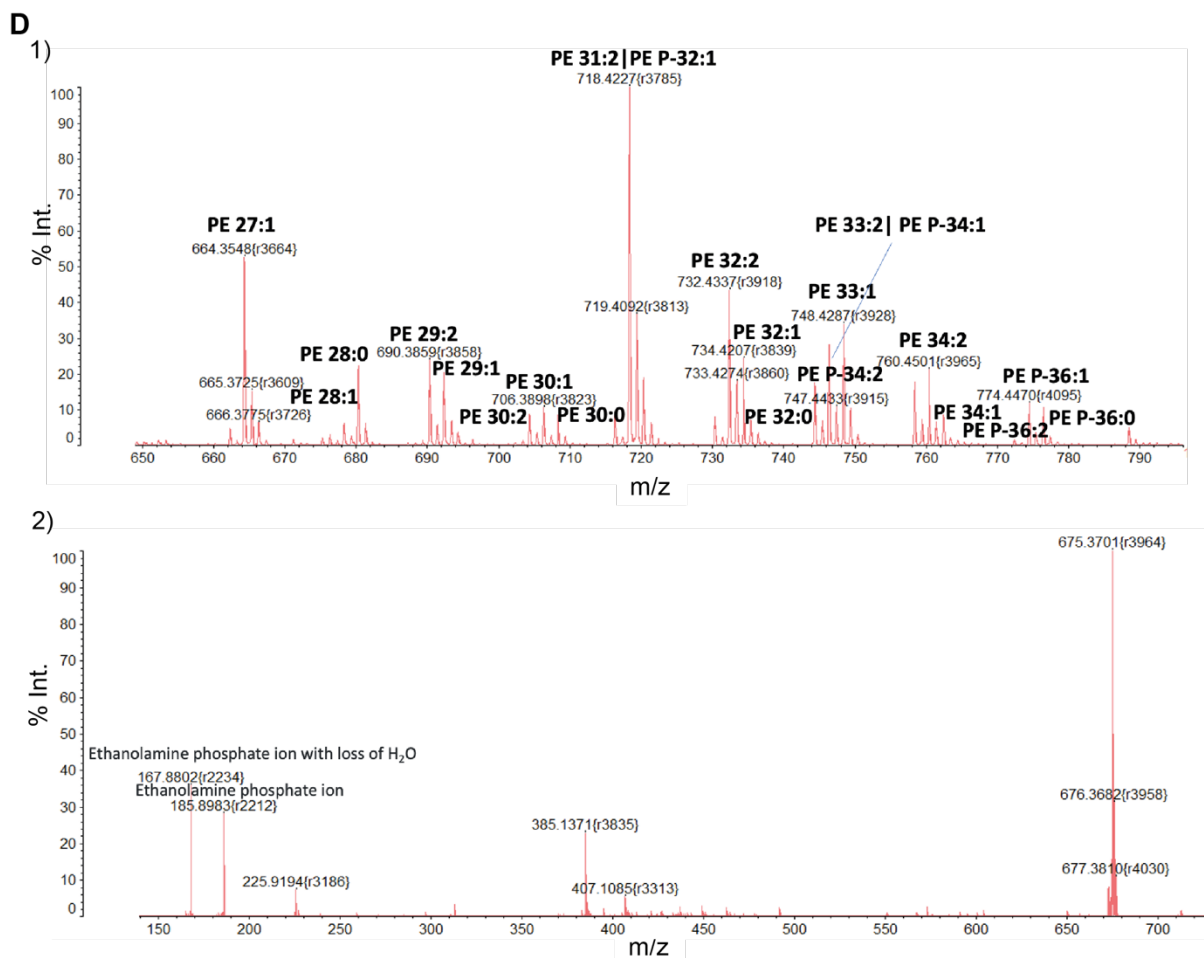
<i>V. parvula</i> homologue	PL biosynthesis gene	Function
FNLLGLLA_00888	Phosphatidylserine decarboxylase (<i>psd</i>)	PS biosynthesis
FNLLGLLA_00889	Phosphatidylserine synthase (<i>ps</i>)	PE biosynthesis
FNLLGLLA_01855	Phosphatidylglycerophosphatase (<i>pgp/yutG</i>)	PG biosynthesis
FNLLGLLA_01719	<i>plsA/R</i>	Plasmalogen biosynthesis

B



C



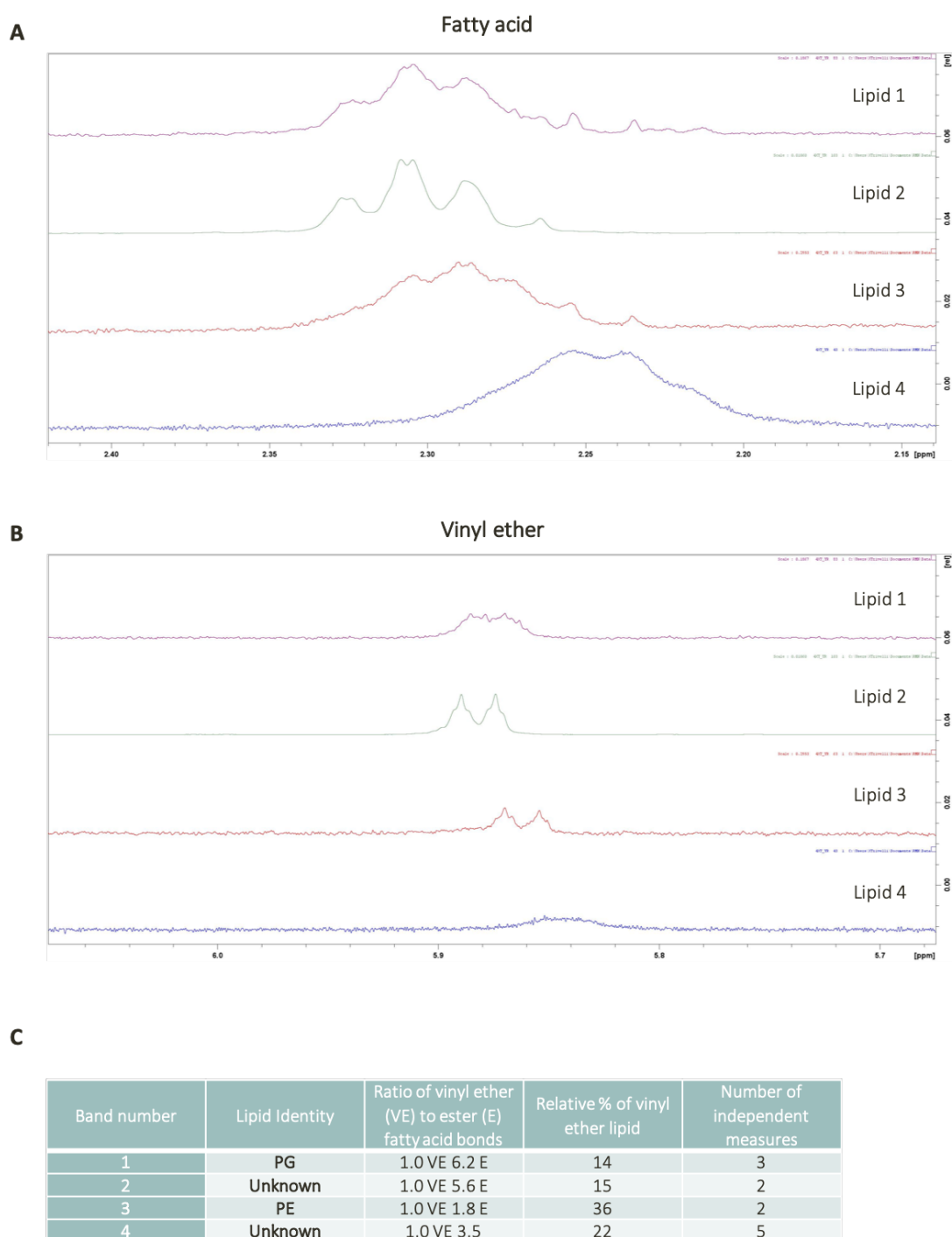


Supplementary Figure S1: GPL biosynthesis homologues in *V. parvula* and identification of PE and PG.

A) Identification of GPL biosynthesis homologues in *V. parvula*. Homologues of four GPL biosynthesis genes were identified in *V. parvula*, including a homologue for the recently identified operon involved in plasmalogen biosynthesis, *plsAR*. **B) Relative proportion of each lipid species in WT SKV38.** 4 biological replicates were tested, stained with phosphomolybdic acid and quantified by ImageJ; n = 4. **C) Mass spectrometry data – PG identification (band 1).** 1) MALDI-QIT-TOF MS analyses of lipid band 1 (positive ion mode) are shown to illustrate phosphatidylglycerol (PG) diversity from lipid extracts of WT *V. parvula* SKV38. 2) MS/MS spectrum of the main species of PG, PG 32:0 at m/z 767.408 with diagnostic ions for PG of 198.9 corresponding to the Glycerol-3-phosphate ion with loss of H_2O . Only one fatty acid neutral loss was found suggesting the presence of two C16:0. **D) Mass spectrometry data – PE identification (band 3).** 1) MALDI-QIT-TOF MS analyses of lipid band 3 (positive ion mode) are shown to illustrate phosphatidylethanolamine (PE) diversity from lipid extracts of WT *V. parvula* SKV38. 2) MS/MS spectrum of the main species observed by MS, PE 31:2 at m/z 718.4 with diagnostic ions for PE of 43 (neutral loss from the precursor ion) and 185 and 167 for respectively the ethanolamine phosphate ion and the ethanolamine phosphate ion with loss of H_2O .

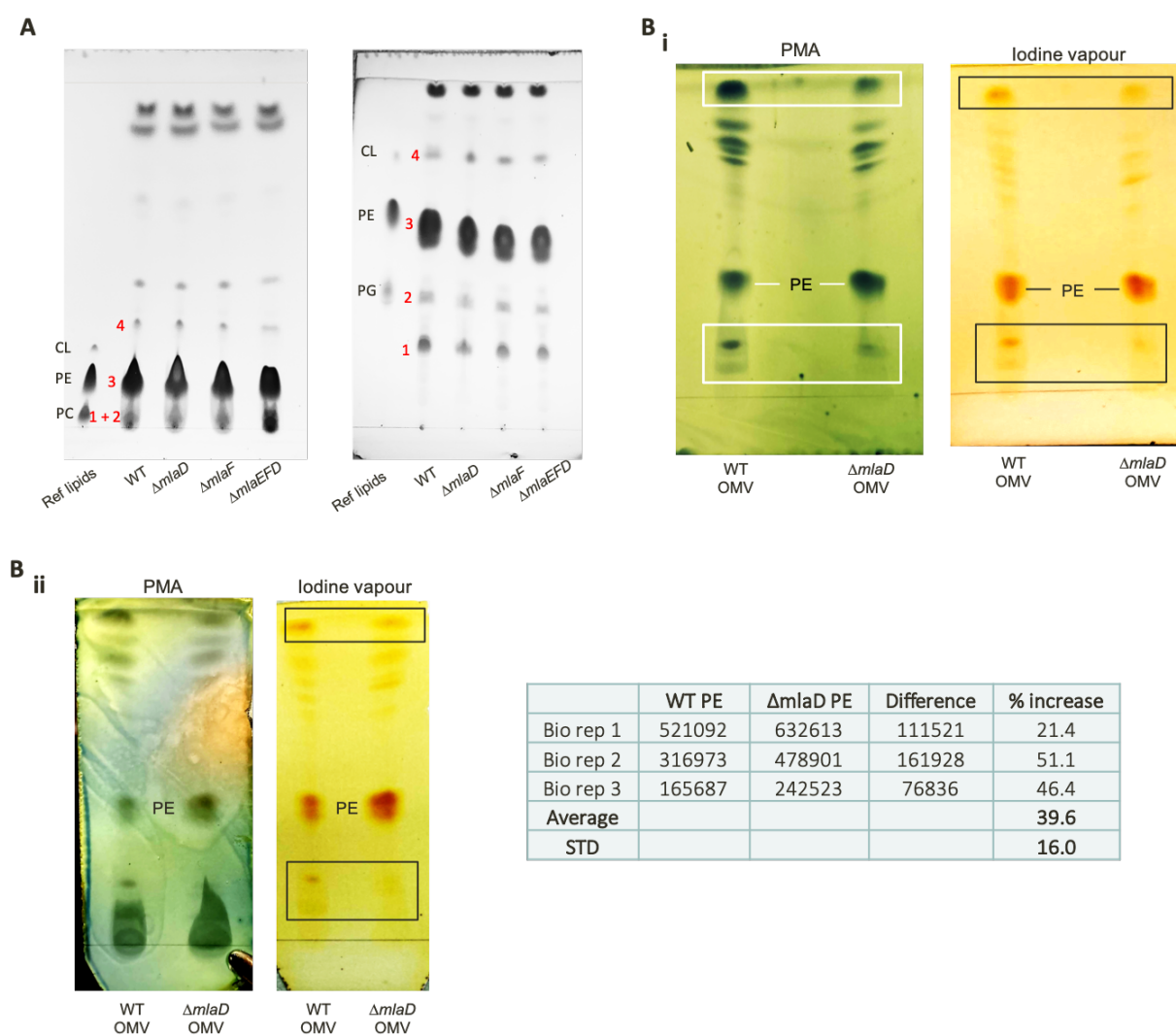
Lipids	<i>m/z</i>	Assignment
Phosphatidylglycerol (PG)	735.434	PG 30:2 PG P-31:1
	739.449	PG 30:0
	753.349	PG 31:0
	765.451	PG 32:1
	767.408	PG 32:0
	781.406	PG 33:0
Phosphatidylethanolamine (PE)	664.354	PE 27:1
	676.368	PE 28:2
	678.366	PE 28:1
	681.369	PE 28:0
	690.385	PE 29:2
	692.384	PE 29:1
	704.396	PE 30:2
	706.389	PE 30:1
	708.387	PE 30:0
	718.422	PE 31:2 PE P-32:1
	732.433	PE 32:2
	734.420	PE 32:1
	736.425	PE 32:0
	744.423	PE P-34:2
	746.440	PE 33:2 PE P-34:1
	748.428	PE 33:1
	760.450	PE 34:2
762.442	PE 34:1	
772.470	PE P-36-2	
774.447	PE P-36:1	
	776.457	PE P-36:0

Supplementary Table S1. Lipid assignments of the total lipid extract of *Veillonella parvula*. Analyses by MALDI-QIT-TOF Shimadzu AXIMA Resonance mass spectrometer in the positive mode. The adduct type for phosphoethanolamine (PE) and phosphatidylglycerol (PG) are [M+2Na-H]⁺.



Supplementary Figure S2: Identification of vinyl ether bonds in *V. parvula* phospholipids

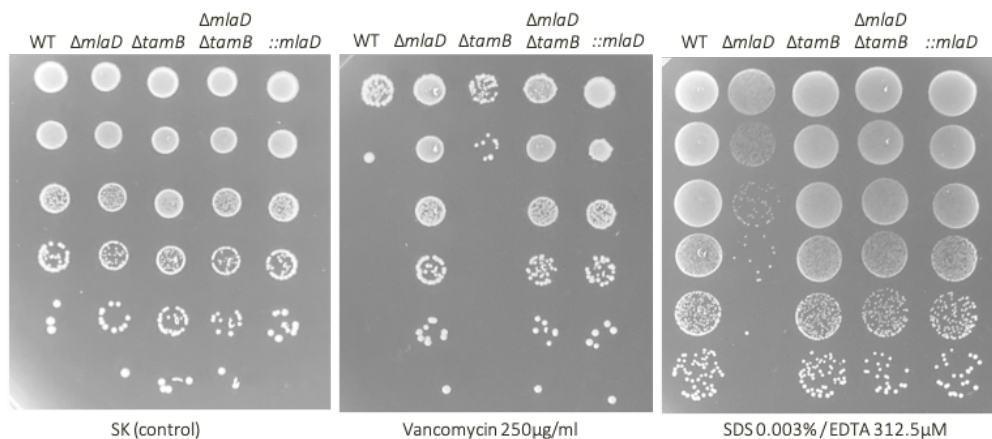
A) Proton Nuclear Magnetic Resonance (^1H -NMR) spectra of lipid fractions 1-4 used for the relative quantification of Ester moieties (FA). Relative quantification was made using signal area from the alpha-methylene of the fatty acids (FA) (two ^1H signals at ca. 2.30 ppm per FA). **B)** ^1H -NMR spectra for relative quantification of Vinyl Ether (VE) bonds. Relative quantification was made using signal area for the VE moiety (one ^1H at 5.88 ppm per VE chain). **C)** Relative proportion of plasmalogen form of each lipid. Based on the relative molar quantity of each of the four major lipids present in the WT, and the NMR data, we calculated the relative molar proportion of lipids containing ether bonds.



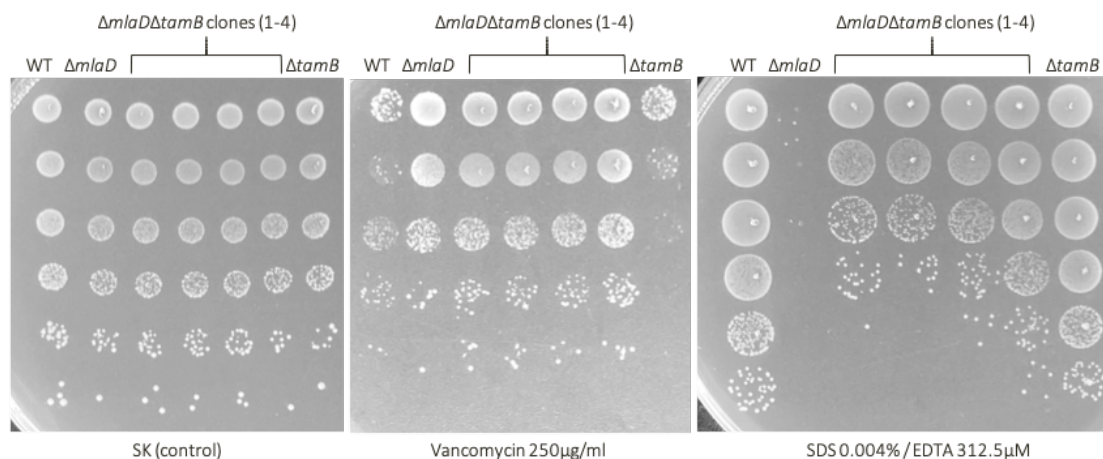
Supplementary Figure S3: Phospholipid compositions of WT and $\Delta ml a$ *V. parvula* mutants

A) Thin Layer Chromatography (TLC) of extracted lipids with iodine vapour staining. i) Solvent system was chloroform/methanol/water (v/v/v 80:20:2.5). A strong clustering of polar species was observed near the origin of the plate. ii) Solvent system was chloroform/methanol/water (v/v/v 65:25:4). This solvent system enabled a better separation of the more polar species, allowing the separation of PE as a double band, likely representative of two different PE species. The major lipid species in *V. parvula*, listed in the table below, are numbered 1 – 4. Commercial reference lipid standards (PC, PE and CL) are labelled. **B) TLC of lipids extracts from OMVs: Phosphomolybdic acid (PMA) and iodine vapour staining** (chloroform/methanol/water 80:20:2.5 v/v/v). Lipid extraction was performed from purified OMVs of WT and $\Delta ml a D$ strains in biological triplicate as described in the methods. 10 μ l of extracted lipids were run on TLC, revealing a ~40% (\pm 16%) enrichment of PE in all $\Delta ml a D$ OMV lipid extracts. PMA staining of lipid extracts from WT and $\Delta ml a$ OMVs shows a decrease in the relative quantities of two unknown lipid species, highlighted in white. This decrease was also visualised after iodine vapour staining, as highlighted in black. The relative decrease in these two species from the $\Delta ml a$ OMVs confirmed the relative enrichment of PE. **B ii)** Additional biological replicate of OMV lipid extraction and staining, including a table showing ImageJ quantifications of staining intensity of iodine vapour-stained lipid extracts across 3 biological replicates.

A

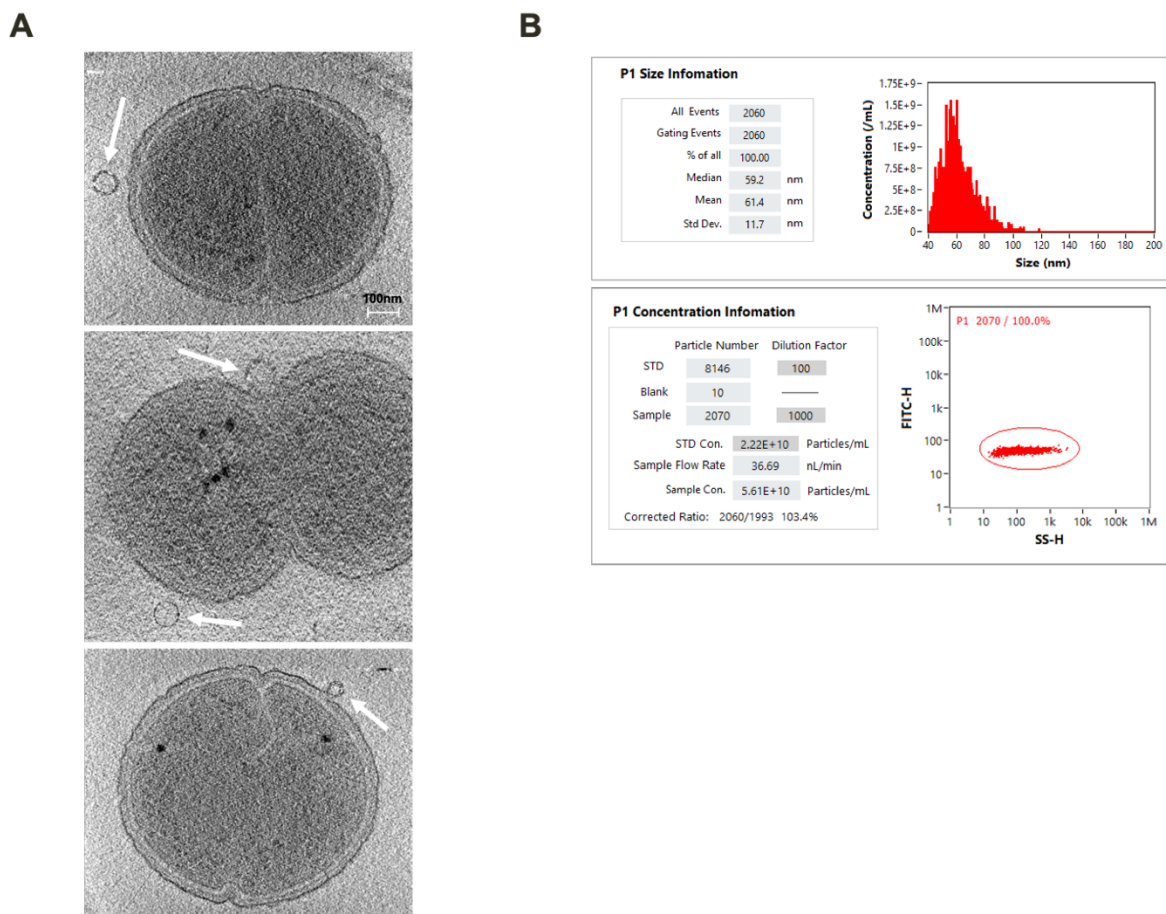


B



Supplementary Figure S4: OM permeability phenotypes of $\Delta mlaD$, $\Delta tamB$ and $\Delta mlaD \Delta tamB$ strains

A) Overnight cultures of WT, $\Delta mlaD$, $\Delta tamB$, $\Delta mlaD \Delta tamB$ and $\Delta mlaD :: mlaD$ were plated onto SDS / EDTA and vancomycin to assess OM permeability. At this concentration of SDS (0.003%) and EDTA (312.5 μ M), the rescue of detergent hypersensitivity of $\Delta mlaD$ by deletion of $tamB$ is striking, and similar to the rescue obtained from reintroducing the expression of $mldA$ itself. $\Delta tamB$ strains are also highly sensitive to vancomycin. **B)** Cultures of WT, $\Delta mlaD$, $\Delta tamB$ and 4 independent biological replicates of the double $\Delta mlaD \Delta tamB$ mutant were plated onto SDS 0.004% / EDTA 312.5 μ M and vancomycin 250 μ g/ml. The slight increase in vancomycin sensitivity of the $\Delta tamB$ can be observed as compared to the WT, which is in direct contrast to the high resistance of $\Delta mlaD$ to vancomycin.



Supplementary Figure S5: Average size of outer membrane vesicles (OMVs) is consistent across whole cell cultures and purified OMV samples.

A) Cryo-electron tomography images of $\Delta mldD$ with outer membrane vesicles (OMVs). Cultures of WT and $\Delta mldD$ strains were observed via cryo-electron tomography (cryo-ET). In all tomograms of $\Delta mldD$, outer membrane vesicles (OMVs) were observed either attached to the cell or externally (highlighted by white arrows). These vesicles were on average ~60nm in diameter, matching the average size observed of these vesicles (present both in whole-cell cultures and in purified samples) via NanoFCM. Scale bar = 100nm. **B) NanoFCM report showing size distribution of purified OMV sample.** OMVs were purified from large-scale supernatants of WT and $\Delta mldD$ cultures. Prior to concentration of these vesicles via ultracentrifugation, a small sample was extracted and processed via NanoFCM, confirming the average size of OMVs is ~60nm, matching the average size of OMVs present in full cell cultures of WT and $\Delta mldD$ strains, and also the average size observed via cryo-ET.

Supplementary Table S2: Primers used in this study.

Primer name	DNA sequence
tamB_3F	CTTCACTCGCCTGCATTTTACTAATGATACTCCCATACACATTATG
tamB_3R	GTTGGTTTGTAGTGGCGAGCTCCG
tamB_5F	GTATAGCCAGCTTGTACCTCAGC
tamB_5R	ATTGTGTCCTGTGGATCCACGAGTATGGCTAGAATGTGTAAAC
m1aD_3F_ery	ATTATTAAACGGGAGGAAATAATGACTAACACATTTACTCTATATGAGTATGGA GGT
m1aD_3R	AGTAACCATCTGGATCAC
m1aD_5F	CGTGTATGATTATGGTGC
m1aD_5R_ery	GAGAATATTTTATATTTTTGTTTCATTTATATATCTCCCATCGCAGATTCGTTAATC G
ery_F_mod	ATGAACAAAAATATAAAATATTCTCAAACTTTTTAACGAGTG
ery_R_mod	TTATTTCCCTCCCGTTAAATAATAGATAACTATTAATAATAGACAATACTTG
catP_F	GGCCTTTTGCTCACATGTTC
catP_R	CCTGAAGTAACTATTTATCAATTCCTGC
tetM_F	AGTAAAATGCAGGCGAGTGAAG
tetM_R	GTGGATCCACAGGACACAAT
m1aE_5F	GCAGCTACAGCCTTATAC
m1aE_5R	TCACTCGCCTGCATTTTACTCAGTGTCTCCCTTTCCC
m1aD_pRPF 185OH_F	AGCGTTAACAGATCTGAGCTGTACAACAATTTATTCATGGTGAAGCAGAGC
m1aD_pRPF 185OH_R	TCTCCTTACTGCAGGAGCTTTAGAATACTTGTTTAATACCGAAATAGGTACCTC CA
m1aE_pRPF 185OH_F	CGTTAACAGATCTGAGCTACTGTAAGGGAAAGGGAGACACTGTTG
m1aD_F_pR PF185_2	CTCTAAGAAAGAAGGAATTCACATGAAGTGGACGACGGAGG
MCE_R_link_pRPF185_2	TACCGCTGCCGCTACCACTATTGATAGACTGTAGCATCTTCTCAGTG
MCE_R_pRPF185_2	CCAGGAGAGTTGTTGATCACTATTGATAGACTGTAGCATCTTCTCAGTG
Barrel_F_OmpSP_pRPF185_2	CTAAGAAAGAAGGAATTCACATGAAAAACAATTCGCAACAATGTTAGCAGCA ACAGCAGTGTTAGGTGTAACAACAGCATTGCTCATAATACGGCACAGATTCA GCTCG

Supplementary Table S3: Strains and plasmids used in this study.

Strain or Plasmid	Description	Antibiotic Resistance	Origin
SKV38	Wild-type	/	1
SKV38 Δ <i>mld</i>	Wild-type with chromosomal <i>mld</i> deletion	Ery	This study
SKV38 Δ <i>mldF</i>	Wild-type with chromosomal <i>mldF</i> deletion	Tc	This study
SKV38 Δ <i>mldEFD</i>	Wild-type with chromosomal <i>mldEFD</i> deletion	Tc	This study
SKV38 Δ <i>mld::mld</i>	Δ <i>mld</i> complemented with pRPF185:: <i>mld</i>	Ery, Cm	This study
SKV38 Δ <i>mldEFD::mldEFD</i>	Δ <i>mldEFD</i> complemented with pRPF185:: <i>mldEFD</i>	Tc, Cm	This study
SKV38 Δ <i>mld::mld</i> -Barrel-HA	Δ <i>mld</i> complemented with pRPF185:: <i>mld</i> (Barrel: residues 250-419 with C-ter HA-tag with linker)	Ery, Cm	This study
SKV38 Δ <i>mld::mld</i> -MCE-HA	Δ <i>mld</i> complemented with pRPF185:: <i>mld</i> (TM + MCE domain: residues 1-155 with C-ter HA-tag with linker)	Ery, Cm	This study
SKV38 Δ <i>tamB</i>	Wild-type with chromosomal <i>tamB</i> deletion	Tc	This study
SKV38 Δ <i>mld</i> Δ <i>tamB</i>	Wild-type with chromosomal <i>mld</i> and <i>tamB</i> deletion	Ery, Tc	This study
SKV38 Δ <i>mld</i> (Tn)	Wild-type with transposon insertion in <i>mld</i> homologue	Ery	This study
Δ <i>mld</i> Δ <i>tamB</i> (Tn 27C1)	Δ <i>mld</i> with transposon insertion in <i>tamB</i>	Ery, Tc	This study
Δ <i>mld</i> Δ <i>tamB</i> (Tn 90E10)	Δ <i>mld</i> with transposon insertion in <i>tamB</i>	Ery, Tc	This study
Δ <i>mld</i> Δ <i>sstT</i> (Tn 5B5)	Δ <i>mld</i> with transposon insertion in <i>sstT</i>	Ery, Tc	This study
Δ <i>mld::inter</i> (Tn 5A4)	Δ <i>mld</i> with transposon insertion in intergenic region between <i>sodA</i> and FNLLGLLA_01854	Ery, Tc	This study
Plasmids			
pRPF215	mariner Tn delivery plasmid, P _{tet} :: <i>HimarI</i> ITR- <i>ermB</i> -ITR <i>catP tetR</i>	Cm	2
pRPF185	Tetracycline-inducible expression system fused with β -glucuronidase <i>gusA</i> Term(fdx)-P _{tet} - <i>gusA</i> -Term(slpA), <i>catP</i>	Cm	3
pRPF185 Δ <i>gusA_2</i>	Tetracycline-inducible expression system with β -glucuronidase <i>gusA</i> removed, <i>catP</i>	Cm	This study

References associated with Supplementary Data

1. Knapp, S. *et al.* Natural Competence Is Common among Clinical Isolates of *Veillonella parvula* and Is Useful for Genetic Manipulation of This Key Member of the Oral Microbiome. *Front. Cell. Infect. Microbiol.* **7**, (2017).
2. Dembek, M. *et al.* High-throughput analysis of gene essentiality and sporulation in *Clostridium difficile*. *MBio* **6**, e02383-14 (2015).
3. Fagan, R. P. & Fairweather, N. F. *Clostridium difficile* has two parallel and essential Sec secretion systems. *J. Biol. Chem.* **286**, 27483–27493 (2011).