Characterization of TelE, an LXG effector of *Streptococcus gallolyticus*, antagonized by a non-canonical immunity protein

Wooi Keong Teh[1], Yichen Ding[1,2], Francesca Gubellini[3], Alain Filloux[1,4], Claire Poyart[5], Shaynoor Dramsi[6,7,*], Michael Givskov[1,8,*]

[1] Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore

[2] Present address: Temasek Life Sciences Laboratory, Singapore

[3] Institut Pasteur, Unité de Microbiologie Structurale, Paris, France

[4] MRC Centre for Molecular Bacteriology and Infection, Department of Life Sciences, Imperial College London, London SW7 2AZ, UK

[5] Université de Paris, Assistance Publique Hôpitaux de Paris, Service de Bactériologie, Centre National de Référence des Streptocoques, Groupe Hospitalier Paris Centre site Cochin, Paris, France

[6] Institut Pasteur, Université Paris Cité, CNRS UMR6047, Biology of Gram-positive Pathogens Unit, Paris 75015 France.

[7] Centre National de la Recherche Scientifique (CNRS) UMR2001, Paris, France

[8] Costerton Biofilm Centre, Department of Immunology and Microbiology, University of Copenhagen, Denmark

* Correspondence:

Shaynoor Dramsi, shaynoor.dramsi@pasteur.fr;

Michael Givskov, mgivskov@sund.ku.dk

Abstract

Streptococcus gallolyticus subsp. *gallolyticus* (*Sgg*) is an opportunistic bacterial pathogen strongly associated with colorectal cancer. Here, through comparative genomics analysis, we demonstrated that the genetic locus encoding the Type VIIb Secretion System (T7SSb) machinery is uniquely present in *Sgg* in two different arrangements. *Sgg* UCN34 carrying the most prevalent T7SSb genetic arrangement was chosen as the reference strain. To identify the effectors secreted by T7SSb, we inactivated the *essC* gene encoding the motor of this machinery. Comparison of the proteins secreted by Sgg UCN34 WT and its isogenic Δ*essC* mutant revealed six T7SSb effector proteins, including an LXG-family toxin named herein TelE displaying pore-forming activity. Seven homologs of TelE harboring a conserved glycine zipper motif at the C-terminus were identified in *Sgg* isolates. Scanning mutagenesis of this motif showed that the glycine residue at position 470 was crucial for TelE pore-forming activity. Unlike other pore-forming toxins commonly antagonized by a membrane protein, TelE activity was antagonized by a cytosolic protein TipE belonging to the DUF5085 family. Finally, we demonstrated that purified TelE can permeabilize *E. coli* cells rapidly when introduced extracellularly. Overall, we report herein a unique *Sgg* T7SSb effector exhibiting a pore-forming activity against non-immune bacteria.

45 ## Author summary

46

47 In this study, 38 clinical isolates of *Streptococcus gallolyticus* subsp. *gallolyticus* (*Sgg*)
48 were sequenced and a genetic locus encoding the Type VIIb secretion system (T7SSb)
49 was found conserved in all 38 genomes. This locus absent from the closely related *S.*
50 *gallolyticus* subsp. *pasteurianus*, constitutes a pathogenicity island. T7SSb is a
51 specialized machinery involved in the secretion of multiple effectors that have
52 been implicated in bacterial virulence, interbacterial killing or host immunity modulation.
53 Here, we report that the model organism *Sgg* strain UCN34 secretes six T7SSb effectors.
54 One of the six effectors named TelE displayed a strong toxicity when expressed in Gram-
55 negative bacteria. Our results indicate that TelE is a pore forming toxin whose activity can
56 be antagonized by a non-canonical immunity protein named TipE. Overall, we report a
57 unique toxin-immunity protein pair and our data expand the range of effectors secreted
58 through T7SSb. As shown previously for gallocin, a bacteriocin giving a selective
59 advantage to *Sgg* in the tumoral microenvironment, we propose that T7SSb effectors
60 including TelE may help *Sgg* to alter the host microbiota contributing to colon colonization.

61

62 **Keywords:** *Streptococcus gallolyticus*; Type VIIb secretion system; pore-forming toxin;
63 LXG family toxin; glycine zipper

Introduction

Streptococcus gallolyticus subsp*. gallolyticus* (*Sgg*) is a commensal bacterium in the rumen of herbivores*.* The first complete genome of strain UCN34 provided clear insights for its adaptation to the rumen (Rusniok et al., 2010). *Sgg* is also an opportunistic pathogen causing septicemia and endocarditis in the elderly who often have concomitant colon tumours (Corredoira et al., 2014, Corredoira-Sanchez et al., 2012, Corredoira et al., 2015, Kim et al., 2010, Kwong et al., 2018, Wending et al., 2006, Boleij et al., 2011, Zhang et al., 2015). *Sgg* was formerly known as *S. bovis* biotype I belonging to the large *S. bovis* / *equinus* complex, which is comprised of several opportunistic pathogens and non-pathogenic bacteria. In this complex, *Sgg* is assigned into a clade with two other closely related subspecies named *S. gallolyticus* subsp. *pasteurianus* (referred to hereafter as *Sgp*) and *S. gallolyticus* subsp. *macedonicus* (referred to hereafter as *Sgm*) (Dekker and Lau, 2016, Schlegel et al., 2003). In general, *Sgp* is considered less pathogenic than *Sgg*, mainly causing neonatal meningitis (Onoyama et al., 2009, Chen et al., 2020), whereas *Sgm* is generally recognized as a safe non-pathogenic strain (Georgalaki et al., 2002, Maragkoudakis et al., 2009).

The Type VIIb Secretion System (T7SSb) is a specialized secretion system exploited by Firmicutes to secrete various substrates or effectors involved in host immune system modulation or bacterial competition (Cao et al., 2016, Anderson et al., 2017, Whitney et al., 2017, Anderson et al., 2011, Unnikrishnan et al., 2017). T7SSb is distantly related to the Type VII secretion system originally discovered in *Mycobacterium tuberculosis.* These two evolutionarily related secretion systems have

89 two components in common; a main FtsK/SpoEIII ATPase (EssC/EccC) constituting
90 the motor of the secretion machinery, and a secreted effector named EsxA/ESAT6
91 belonging to the WXG100 family (Abdallah et al., 2007). In Firmicutes, the T7SSb core
92 machinery is composed of 4 additional proteins namely EssA, EsaA, EsaB and EssB.
93 T7SSb is involved in the secretion of a variety of effectors including (i) WXG100
94 proteins of approximately 100 amino acids (aa) with a WXG motif in the center such
95 as EsxA and EsxB (Burts et al., 2005); (ii) proteins of approximately 100-aa without a
96 WXG motif (named WXG100-like proteins) such as EsxC and EsxD (Anderson et al.,
97 2013, Burts et al., 2008); (iii) protein with an N-terminal domain COG5444 such as
98 EsaD/EssD (Cao et al., 2016, Ohr et al., 2017); and (iv) polymorphic toxins containing
99 a N-terminal domain of about 200 aa with an LXG motif (also known as LXG family
100 proteins) such as TelA, TelB, TelC, TspA and EsxX (Ulhuq et al., 2020, Whitney et al.,
101 2017, Dai et al., 2017, Holberger et al., 2012). Previous studies revealed that these
102 effectors function as a NADase toxin (TelB) (Whitney et al., 2017), a lipid II
103 phosphatase (TelC) (Whitney et al., 2017), nucleases (EsaD, YobL,, YxiD) (Cao et al.,
104 2016, Ohr et al., 2017, Holberger et al., 2012), or a membrane depolarizing toxin (TspA)
105 (Ulhuq et al., 2020).
106
107 Importantly, T7SSb of *Sgg* strain TX20005 was previously shown to contribute to the
108 development of colon tumors in a murine model (Taylor et al., 2021). Here, using
109 comparative genomics, we identified that the T7SSb locus residing in a 26-kb genomic
110 island is uniquely present in *Sgg* but neither in *Sgp* nor in *Sgm*. Subsequently, we
111 found that six proteins were secreted through the T7SSb machinery of the reference
112 strain UCN34, including the WXG100 effector EsxA, two hypothetical proteins, and
113 three LXG family proteins including two TelC homologs. One of the unknown LXG

114 family protein was characterized herein and named TelE. We show that TelE is a pore-

115 forming protein whose activity is antagonized by a non-canonical immunity protein.

116 ## Results

117

118 **A T7SSb locus specific to *Sgg*.** In order to identify the specific pathogenic trait(s) of
119 *Sgg*, comparative genomics was performed on newly sequenced 13 *Sgp* and 38 *Sgg*
120 clinical isolates obtained from the National Reference Center of Streptococci in France.
121 Phylogenetic analysis of these newly sequenced genomes, together with the complete
122 genomes of *Sgg* and *Sgp* publicly available on NCBI GenBank assigned *Sgg* and *Sgp*
123 into two distinct clades, confirming their correct classification **(Figure S1)**. BLASTn-
124 based comparative genomic analysis of the reference genome UCN34 together with
125 the representative *Sgg* and *Sgp* genomes uncovered a genomic island unique to *Sgg*
126 **(Figure 1A)**. This genomic island contains the six genes (*esxA*, *essA*, *esaB*, *essB*,
127 *essC* and *esaA*) encoding the core components of the Type VIIb Secretion System
128 (T7SSb) machinery **(Figure 1A and 1B)**.

129

130 An extended protein homology search of each T7SSb core components across all the
131 newly sequenced genomes confirmed the conservation of T7SSb in *Sgg* genomes
132 and its absence in *Sgp* or in *Sgm* **(Figure 1C)**. A closer inspection of the T7SSb
133 genetic locus uncovered two different gene arrangements in the *Sgg* isolates
134 sequenced so far. In a vast majority of *Sgg* isolates, the T7SSb locus is arranged in
135 the order of *esxA/essA-esaB-essB-essC-esaA*, as in our reference strain UCN34.
136 Notably, *esxA* is transcribed in the opposite direction from the other five genes *essA-*
137 *esaB-essB-essC-esaA,* a unique feature distinct from other Firmicutes. In contrast, in
138 a small minority of *Sgg* isolates including the US isolate TX20005, the T7SSb genes
139 are arranged in a different order of *esxA-esaA-essA-esaB-essB-essC* with an evident
140 shuffling of *esaA*, and the six genes transcribed in the same direction **(Figure 1D)**.

141

142 ***Sgg* UCN34 secretes 6 T7SSb effectors.** In order to identify the T7SSb effectors, we
143 inactivated the T7SSb machinery, through in-frame deletion of *essC* gene encoding
144 the main ATPase, in the strain UCN34. We next compared the secretome of the
145 UCN34 WT strain to that of the Δ*essC* mutant. This approach identified 6 proteins that
146 were differentially expressed **(Figure 2A, Figure S2A)** including the ubiquitous
147 WXG100 effector (EsxA/Gallo_0553) that was found 4-fold less abundant in the Δ*essC*
148 mutant, confirming that the Δ*essC* mutant is a *bona fide* T7SSb-deficient strain. In
149 addition, five proteins namely Gallo_0559, Gallo_0560, Gallo_0562, Gallo_1068 and
150 Gallo_1574 were detected only in the secretome of the UCN34 WT but not in the
151 Δ*essC* mutant.

152

153 Prediction of the putative function of these effectors by *in silico* SMART analysis
154 revealed that Gallo_0559 contains a domain of unknown function DUF5082,
155 Gallo_0560 as a TIGR04197 family protein, whereas Gallo_0562, Gallo_1068 and
156 Gallo_1574 contain the prototypical N-terminal LXG domain (pfam04740) found in
157 other T7SSb effectors **(Figure 2B)**. Protein homology search indicated that
158 Gallo_1068 and Gallo_1574 are homologous (>98% coverage, >40% similarity) to the
159 *Streptococcus intermedius* TelC effector that functions as a lipid II phosphatase
160 (Whitney et al., 2017). Similar to TelC, both Gallo_1068 and Gallo_1574 contain an
161 intact C-terminal aspartate-rich motif critical for TelC enzymatic activity. We thus
162 propose to rename Gallo_1068 and Gallo_1574 as TelC1 and TelC2, respectively
163 **(Figure 2B)**. The third LXG-containing effector (Gallo_0562) is a protein of 514 amino
164 acids sharing no similarities with known proteins that we propose to rename TelE in
165 accordance with the gene nomenclature (Toxin exported by Esx with LXG domain)

166 proposed for *S. intermedius* (Whitney et al., 2017). Of note, 4 out of 6 proteins
167 identified in the secretome of UCN34 WT are encoded in the T7SSb locus and three
168 are encoded by genes likely transcribed as an operon, namely *gallo_0559*, *gallo_0560*
169 and *telE (gallo_0562)*.
170
171 **TelE is toxic when overexpressed in Gram-negative bacteria.** LXG family proteins
172 are polymorphic toxins implicated in interbacterial competition (Whitney et al., 2017).
173 However, we failed to observe any role of *Sgg* T7SSb in interbacterial competition
174 under standard laboratory conditions. We reasoned that this could be due to the low
175 expression of T7SSb effectors as reflected by the low total peptide counts in the
176 proteomics data **(Figure S2A)** and decided to test their potential antibacterial activity
177 through heterologous and inducible expression in the Gram-negative *Escherichia coli*
178 K-12. As shown in **Figure S2B**, overexpression of *telE*, but none of the other T7SSb
179 effectors, caused a strong toxicity in *E. coli* or in another Gram-negative bacterium
180 *Pseudomonas aeruginosa*, but not in the Gram-positive bacteria *Enterococcus*
181 *faecalis*, *Streptococcus agalactiae* and *Sgm* **(Figure 2C)**. Since TelC was shown
182 previously to be toxic to Gram-positive bacteria (Whitney et al., 2017), TelC1 and
183 TelC2 were then expressed in *Sgm* and *E. faecalis.* Although the survival of *Sgm* or
184 *E. faecalis* was not affected upon induced expression of TelC1 or TelC2, we observed
185 an altered colony morphology in *Sgm* **(Figure S2C)**. Of note, the lack of activity of
186 TelC1, TelC2 and TelE in Gram-positive bacteria could be due to the lower production
187 of these effectors in these cells by unknown mechanisms (data not shown).
188
189 TelE is encoded in the vicinity of the T7SSb core machinery locus (**Figure 3A**). This
190 prompted us to examine the prevalence and the conservation of TelE among the *Sgg*

191 clinical isolates. As shown in **Figure S3,** TelE homologs were found in over 90% of
192 *Sgg* isolates. Based on their primary amino acid sequences, TelE homologs can be
193 categorized into 7 variants named TelE1 (initial TelE found in UCN34) through TelE7.
194 Notably, multiple copies of TelE homologs can be found in a few *Sgg* genomes. Some
195 of these TelE homologs are N-terminally truncated, resembling the orphan modules
196 commonly seen in polymorphic toxins (Jamet and Nassif, 2015) **(Figure S3).** To
197 assess the antimicrobial activity of these TelE homologs, each TelE variant was
198 expressed in *E. coli.* As shown in **Figure 3B,** all the TelE variants exhibited a strong
199 toxicity upon induction in *E. coli*. Interestingly, a multiple sequence alignment of these
200 seven TelE variants uncovered a highly conserved sequence motif at the C-terminus
201 (**Figure S4**). This motif, rich in small hydrophobic amino acids alternating with
202 conserved glycine residues, is similar to the glycine zipper motif (GxxxGxxxG) found
203 in many membrane proteins (Kim et al., 2005) **(Figure 3C)**. Three highly conserved
204 glycine residue were identified at positions 458, 470 and 474, and three less-
205 conserved glycine residues were found at positions 466, 478 and 480 **(Figure 3C and**
206 **3D)**. Single replacement of each glycine residue to a slightly bulkier valine residue
207 demonstrated that those at positions 466, 470 and 474 were key for TelE activity in *E.*
208 *coli* **(Figure 3E)**, with the replacement of the glycine residue at position 470 showing
209 the strongest impact on TelE activity (**Figure 3E**). Together, these data showed that
210 nearly all the *Sgg* isolates sequenced as to date encode a functional TelE homolog
211 whose primary sequence varies from one isolate to another, but all 7 TelE variants
212 contain a conserved glycine zipper motif at the C-terminus essential for TelE
213 antimicrobial activity.
214

215 **TelE is a pore-forming toxin.** Structural modelling of TelE did not show any
216 similarities with known proteins but indicated that TelE secondary structures are rich
217 in alpha-helices **(Figure 4A).** Alpha-helical structures with glycine zipper motifs are
218 the two main characteristics of membrane proteins with a high propensity to form
219 homo-oligomeric ion channels (Kim et al., 2005), suggesting that TelE might have a
220 pore-forming activity. To investigate this possibility, we used propidium iodide, a large
221 molecule (668 Da) that is conventionally used to assess the presence of membrane
222 pores or membrane disruption in bacterial cells (LaCourse et al., 2018, Fridman et al.,
223 2020). Propidium iodide can only enter the cellular cytoplasmic space through
224 membrane lesions or large membrane pores to subsequently bind to intracellular DNA
225 and emit a red fluorescence signal. Using time-lapse fluorescence microscopy, we
226 observed a gradual increase of red fluorescence signal in *E. coli* cells expressing TelE,
227 but not in control cells carrying the empty vector **(Figure 4B)**, indicating that TelE
228 expression resulted in an influx of propidium iodide into the cells. Intriguingly, this influx
229 was preceded immediately by shrinkage of the bacterial cell visible by phase contrast
230 microscopy **(Figure 4C)**. Similar morphological changes have been observed on cells
231 experiencing water loss or cytoplasmic content leakage (Sutterlin et al., 2016, Shi et
232 al., 2020), which can be due to the excessive presence of membrane pores. Our
233 microscopy observations thus indicate that TelE can form membrane pores in *E. coli*
234 cells.
235
236 Next, we probed the subcellular localization of TelE in *E. coli* by using TelE with an
237 hexahistidyl tag either at the N- or C-terminus. Notably, we found that TelE with a His-
238 tag at the C terminus (TelE-His) is significantly less active in intoxicating *E. coli* as
239 compared to the native or N-terminally tagged TelE. Since TelE-His was still functional

240 when overexpressed in strain BL21(DE3) **(Figure S5A-C)**, we proceeded with this less
241 toxic TelE-His as a means to monitor TelE subcellular localization. As shown in **Figure**
242 **4D**, TelE-His was predominantly found in the inner and outer membrane of *E. coli*. On
243 the basis that TelE localizes in the membrane and its expression induces cellular
244 uptake of propidium iodide from the extracellular medium, we propose that TelE is a
245 pore-forming toxin. We next investigated the subcellular localization of the inactive
246 $TelE_{G470V}$ mutant by adding a hexahistidyl tag at the C-terminus ($TelE_{G470V}$-His). As
247 shown in **Figure 4D**, this construct was stably expressed in *E. coli* cells and was able
248 to localize in the membrane as the wildtype TelE-His, indicating that the loss of activity
249 of the inactive $TelE_{G470V}$ variant was not due to protein instability or protein
250 delocalization in the cells.
251
252 **The cognate immunity protein TipE antagonizes TelE toxicity.** The activity of LXG
253 family toxins such as TelA, TelB and TelC can be antagonized by cognate immunity
254 proteins (Whitney et al., 2017). To identify the immunity protein able to counteract TelE
255 toxicity, each gene located downstream from *telE* was cloned and co-expressed in *E.*
256 *coli* cells using a second compatible plasmid (pJN105). We found that only co-
257 expression of the third gene downstream from *telE* (*gallo_0565*) rescued *E. coli* from
258 TelE intoxication **(Figure 2A, 5A)**. This gene encodes a 151 amino acids-long protein
259 renamed herein as TipE (Tel immunity protein), which harbor a domain of unknown
260 function DUF5085. An additional DUF5085 domain-containing protein (Gallo_0570),
261 sharing ~30% protein sequence similarity with TipE, was found five genes downstream
262 of TipE **(Figure 2A)**. However, co-expression of this protein did not protect *E. coli* from
263 TelE intoxication, suggesting a specific interaction between TipE and TelE **(Figure**
264 **5A)**.

265

266 To decipher how TipE confers immunity, we constructed a recombinant TipE tagged
267 with a 9-aa hemagglutinin tag (HA tag) at the N-terminus for the detection of the
268 recombinant TipE using a specific anti-HA monoclonal antibody. We showed that the
269 recombinant HA-TipE was functional and able to rescue *E. coli* from TelE intoxication
270 **(Figure S5B)**. Subcellular fractionation experiments indicates that TipE is
271 predominantly a soluble protein **(Figure 5B)**. To determine whether TipE is a
272 periplasmic or a cytosolic protein, we performed a separate extraction of periplasmic
273 proteins but did not detect TipE in this fraction (**Figure 5B**). Thus, we conclude that
274 TipE is a cytosolic protein. In contrast, when HA-TipE was co-expressed with TelE-
275 His, these two proteins were mostly localized in the insoluble fraction **(Figure S5C)**,
276 suggesting that TipE interacts with TelE to counteract its function.

277

278 To demonstrate that TipE interacts with TelE, a protein pulldown experiment was
279 carried out using nickel-immobilized resin. Incubation of TelE-bound resin with cellular
280 lysate containing HA-TipE captured a significant amount of TipE, which was detected
281 through immunoblot using anti-HA specific antibodies following TelE elution (**Figure**
282 **5C**). Importantly, TipE was not detected in similar amounts when the cellular lysate
283 was incubated with free nickel-immobilized resin (**Figure 5C**), indicating that TipE
284 indeed interacts with TelE. To monitor in more detail the impact of TipE on TelE in live
285 cells, we constructed a chimeric TelE variant C-terminally fused with superfolder green
286 fluorescent protein (TelE-sfGFP). This TelE chimera protein, upon excitation, allows
287 direct visualization of TelE in live cells using fluorescence microscopy. We first showed
288 that TelE-sfGFP is still toxic in *E. coli* **(Figure S5A**) and is able to cause cellular influx
289 of propidium iodide when expressed **(Figure 5D)**. As shown in **Figure 5D**, sfGFP was

290 found uniformly distributed throughout the cytoplasm in its native form **(Figure 5D)**.
291 However, in *E. coli* cells expressing TelE-sfGFP, TelE was found in small foci
292 distributed around the cell body **(Figure 5D)**. In contrast, in *E. coli* cells co-expressing
293 TelE-sfGFP and TipE, the GFP signal was mainly observed in the cytoplasm of these
294 cells **(Figure 5D)**, suggesting that TelE cellular localization was altered in the presence
295 of the immunity protein TipE. Taken together, these data indicate that TipE
296 antagonizes TelE activity by disrupting TelE cellular localization through protein-
297 protein interaction.
298
299 **Purified recombinant TelE protein is able to permeabilize live *E. coli* cells**. Since
300 TelE is a T7SSb effector found in the supernatant of *Sgg* UCN34, we wondered
301 whether TelE was able to induce pore-formation on *E. coli* cells when present in the
302 external milieu. To purify TelE, we turned to the TelE variant appended with 9-aa long
303 HA tag at the C-terminus (TelE-HA) that exhibits drastic reduction of toxicity in *E. coli*.
304 This allowed the purification of sufficient amount of TelE from the membrane fraction.
305 We next introduced the purified TelE onto *E. coli* cells labelled with the fluorescent dye
306 carboxyfluorescein diacetate succinimidyl ester (CFSE). This cell-tracker dye is
307 fluorescent only intracellularly upon cleavage by a non-specific esterase. It is long
308 lasting and well retained dye within labeled cells. TelE addition to fluorescent gut-
309 adapted *E. coli* pks+ induced a rapid leakage of the fluorescent dye in 10 min at 37°C
310 (**Figure 6**). A similar effect can be seen when the cells were treated with membrane-
311 permeabilizing detergent such as Triton X-100 or Tween 20 (**Figure 6A lower panel**
312 and **6B**). Importantly, detergent octyl glucoside used for TelE purification had little
313 effect on *E. coli* permeabilization. We further confirmed that the leakage of fluorescent
314 dye from TelE-treated *E. coli* cells was not due to the detergent octyl glucoside by

315 incubating the cells with TelE depleted of detergent either by beta-cyclodextrins (TelE1)
316 or Bio-Beads (TelE2 and TelE3) (**Figure 6C**). Taken together, we conclude that TelE
317 is able to permeabilize live *E. coli* cells when added in the external milieu.
318

319 ## Discussion

320

321 In this study, we show that a T7SSb cluster is a putative pathogenicity island
322 conserved in all the clinical isolates of *Streptococcus gallolyticus subsp. gallolyticus*
323 (*Sgg*). Interestingly, two different T7SSb genetic arrangements can be found in *Sgg*
324 clinical isolates. The most prevalent T7SS arrangement resembles those found in
325 *Streptococcus intermedius*, whereas the T7SS of a small minority of *Sgg* isolates is
326 identical to those of *Staphylococcus aureus* (Cao et al., 2016, Whitney et al., 2017)
327 **(Figure 1D)**. Six T7SSb effectors were identified in strain UCN34, including the
328 prototypical EsxA/ESAT-6, and three LXG-containing proteins. Amongst the three
329 LXG effectors, two encoded outside of the T7SSb locus are homologs of TelC, a lipid
330 II phosphatase shown to inhibit peptidoglycan biosynthesis in Gram-positive bacteria
331 (Whitney et al., 2017, Klein et al., 2018). The third LXG effector of unknown function
332 is unique to *Sgg* and is encoded in vicinity of the locus encoding T7SSb machinery.
333 Here, we showed that TelE exhibits pore-forming activity and its activity can be
334 neutralized by a cytoplasmic protein TipE.

335

336 TipE is an atypical immunity protein. In other pore-forming toxins similarly containing
337 glycine zipper motifs such as CdzC/D from *Caulobacter crescentus* and Tse4 from
338 *Pseudomonas aeruginosa*, the cognate immunity proteins are multi-pass
339 transmembrane proteins residing predominantly in the inner membrane (Garcia-
340 Bayona et al., 2017, LaCourse et al., 2018). Similarly, an inner membrane protein
341 protects the *E. coli* cells expressing the *S. aureus* membrane depolarizing T7SSb toxin
342 TspA (Ulhuq et al., 2020). Remarkably, the six-pass transmembrane protein
343 Gallo_0563 encoded by the gene located immediately downstream of *telE* did not

344 protect *E. coli* from TelE intoxication **(Figure 5A, 5B)**, and our data showed that only
345 TipE is able to block TelE activity through direct interaction and sequestration. Notably,
346 TipE is a DUF5085 family protein predicted related to T7SS in the Pfam database.
347 Based on our data, we propose herein that DUF5085 family proteins could serve as
348 immunity proteins to T7SSb toxins.
349
350 Interestingly, a diverse pool of TelE variants (TelE to TelE7) exist in the genomes of
351 *Sgg* isolates, all harboring the conserved C-terminal glycine zipper motif essential for
352 TelE activity. Glycine zipper motif is a common feature in membrane proteins
353 displaying pore-forming activity. In addition to the aforementioned CdzC/D and Tse4,
354 glycine zipper motif is also found in the vacuolating toxin VacA from *Helicobacter pylori*,
355 the prion protein (PrP) and the amyloid beta peptide (Aβ) (Kim et al., 2005), many of
356 which display pathological roles. For instance, VacA is involved in gastric ulceration
357 caused by *H. pylori*, whereas PrP and Aβ are associated with prion and Alzheimer's
358 disease, respectively[31]. In line with previous observations (LaCourse et al., 2018, Kim
359 et al., 2005, McClain et al., 2003), mutation of the key glycine residue in this motif
360 abolishes TelE activity, yet did not alter TelE membrane localization. Since the glycine
361 residues of the motif are usually positioned in the inner part of membrane pores, we
362 speculate that change of the small glycine to a bulkier valine residue may affect the
363 oligomerization and/or pore opening of TelE, as demonstrated for other glycine zipper
364 toxins (Kim et al., 2005).
365
366 In addition to TelE, we discovered two additional yet to be characterized T7SSb
367 effectors namely Gallo_0559 and Gallo_0560 located immediately upstream of TelE
368 and probably encoded in the same operon. Bioinformatic analysis showed that

369 Gallo_0559 is a DUF5082 family protein whereas Gallo_0560 is a TIGR01497 family
370 protein. Proteins belonging to these two families were previously identified in *S.*
371 *intermedius* as WxgA and WxgC, respectively (Whitney et al., 2017). WxgA and WxgC
372 are chaperone proteins essential for the secretion of LXG family proteins (Tel) in *S.*
373 *intermedius* (Whitney et al., 2017). These proteins were suggested as WXG100-like
374 proteins (Whitney et al., 2017), but WxgA and another chaperone protein WxgB are
375 40% longer than the canonical WXG100 proteins. These chaperones have not been
376 detected in the *S. intermedius* bacterial supernatant and were henceforth
377 hypothesized to disassociate from Tel proteins prior to or during secretion (Whitney et
378 al., 2017). Here we found that these probable Wxg-like chaperones were detected in
379 the supernatant of *Sgg* WT but not in the Δ*essC* mutant, indicating that the secretion
380 of these proteins is T7SSb-dependent, and that they may remain associated with LXG
381 toxins post secretion. Further investigation on the potential interaction of these two
382 proteins with TelE may provide insight into T7SSb effectors secretion and delivery.
383
384 The three LXG family toxins identified in this study expands our knowledge on the
385 antibacterial arsenal possessed by *Sgg*. As demonstrated here, TelE is a toxin able to
386 permeabilize Gram-negative bacteria, whereas the two other LXG family toxins TelC1
387 and TelC2 homologous to lipid II phosphatase TelC probably targets Gram-positive
388 bacteria (Whitney et al., 2017). It is noteworthy that *Sgg* also produces bacteriocins
389 such as gallocin exhibiting antimicrobial activity against gut commensals such as *E.*
390 *faecalis* (Aymeric et al., 2018), which together with the newly identified Tel toxins, are
391 likely the arsenal essential for *Sgg* survival in the highly dynamic and populated human
392 gut environment.
393

394 Finally, our findings may have further implications in understanding the pathogenicity
395 of *Sgg*. Indeed, it was shown in *Sgg* strain TX20005 that T7SSb is a key component
396 contributing to *Sgg* gastrointestinal tract colonization and cancer-promoting capability
397 (Taylor et al., 2021). Of note, TX20005 belongs to the small minority of *Sgg* isolates
398 displaying a different T7SSb gene locus arrangement as compared to UCN34, and
399 the impact of these differences on the pathogenic outcome is currently unclear. The
400 contribution of pathogenicity associated islands in bacterial virulence is well
401 established. For instance, the CagA toxin, VacA toxin and Type IV secretion system
402 encoded in the *Helicobacter pylori* pathogenicity island (Franco et al., 2008, Wen and
403 Moss, 2009), or the colibactin synthase genes encoded in the *E. coli* polyketide
404 synthase (*pks*) genotoxic island (Arthur et al., 2012, Cuevas-Ramos et al., 2010,
405 Pleguezuelos-Manzano et al., 2020) are key to the progression of gastric cancer and
406 colorectal cancer, respectively. Of particular interest is VacA that also contains a
407 glycine zipper motif essential for its function causing ulceration and gastric lesions in
408 murine model (Telford et al., 1994). Whether TelE could also alter host cells is a
409 perspective for future studies.

410 **Methods**

411

412 **Bacterial culture conditions**

413 All bacterial strains and plasmids used in this study are listed in **Table S1**. Cultures of
414 *E. coli* and *P. aeruginosa* strains were prepared from single colonies in lysogeny broth
415 (LB) (BD) and incubated overnight at 37°C, 200 rpm, whereas cultures of *E. faecalis*,
416 *S. agalactiae* and *S. gallolyticus* strains were prepared from single colonies in brain
417 heart infusion broth (BHI) (Acumedia) and incubated statically overnight at 37°C.
418 When the use of growth medium with minimal amount of peptides were necessary,
419 overnight culture of *Sgg* were prepared from single colonies in M9YEG broth (1X M9
420 minimal salts supplemented with 0.5% yeast extract and 1.0% of glucose) (Teh et al.,
421 2019) and incubated statically at 37°C. When appropriate, antibiotics were
422 supplemented at the following concentrations: kanamycin, 50 µg/mL for Gram-
423 negative bacteria and 500 µg/mL for Gram-positive bacteria; gentamicin, 20 µg/mL for
424 *E. coli*, and 60 µg/mL for *P. aeruginosa*; erythromycin, 2 µg/mL for Gram-positive
425 bacteria.

426

427 **Genetic techniques**

428 Standard molecular cloning procedures as previously described (Green et al., 2012)
429 were followed to construct plasmids for gene deletion and gene expression. All primers
430 used in this study were synthesized by Integrated DNA Technology and listed in **Table**
431 **S1**. The enzymes Q5 DNA polymerase and restriction enzymes were obtained from
432 New England Biolabs, whereas T4 ligase was obtained from Thermo Fisher Scientific.
433 All genetic constructs were delivered into chemically transformed *E. coli* DH5α
434 prepared as described (Inoue et al., 1990) and verified by Sanger sequencing

435 performed at 1st BASE. For gene expression plasmids construction, the
436 corresponding open reading frames with the native ribosomal binding sites were
437 cloned downstream of inducible promoters such as P_{tetO} (anhydrotetracycline-
438 inducible) and P_{BAD} (arabinose-inducible).
439
440 To construct plasmids for gene deletion, two ~500bp DNA fragments corresponding
441 to the upstream and downstream region of the gene of interest were first PCR
442 amplified. The resulting two DNA fragments were subsequently spliced by overlap
443 extension PCR and ligated into pG1 plasmid at *Sma*I restriction site. The constructed
444 plasmids were transformed into *E. coli* and later transformed into *S. gallolyticus* as
445 described in the transformation section below. Subsequently, *Sgg* in-frame
446 chromosomal deletion mutants were generated by allelic exchange as previously
447 described (Danne et al., 2013).
448
449 To generate plasmids encoding TelE with a specific residue substitution, *telE* was
450 PCR-amplified into two fragments corresponding to the 5' and 3' regions by using
451 primers with the desired mutation incorporated. These two fragments were spliced into
452 a full-length *telE* open reading frame by overlap extension PCR and ligated into
453 pTCV*erm*-ptetO at the *Bam*HI/*Sph*I restriction site.
454
455 Plasmid encoding TelE with the C-terminus fused with superfolder GFP was
456 constructed by splicing the PCR amplicons corresponding to the *telE* open reading
457 frame without a stop codon and the superfolder GFP open reading frame that has
458 been codon-optimized for *Sgg* (gene synthesized at GenScript).
459

460 **Transformation**

461 *E. coli* DH5α was chemically transformed as described in (Inoue et al., 1990) whereas
462 *P. aeruginosa* was transformed by electroporation as previously described (Smith and
463 Iglewski, 1989). *E. faecalis* and *S. agalactiae* were transformed by electroporation as
464 previously described (Bae et al., 2002, Dunny et al., 1991, Biswas et al., 1993). *S.*
465 *agalactiae* carrying the respective plasmids were used for conjugal transfer into *Sgm*
466 as described previously (Danne et al., 2013). *Sgg* was naturally transformed using a
467 newly established method, by supplementing the log-phase culture grown in M9YEG
468 with 10 µM of the competence inducing peptides (ITGWWGL) outlined in the previous
469 study (Morrison et al., 2013), at least 5 µg of plasmid DNA, 10 µM of magnesium
470 chloride, 10 µM of magnesium sulfate and 1 µM of ferrous (II) chloride. The mixture
471 was incubated for an hour at 37°C, followed by three hours at 30°C, before being
472 plated onto BHI agar supplemented with erythromycin.

473

474 **Genomic DNA isolation**

475 *S. gallolyticus* genomic DNA was extracted using either Wizard Genomic DNA
476 Purification Kit or Wizard SV Genomic DNA Purification Kit (Promega). The extraction
477 was performed according to the manufacturer's instructions, except that 20U to 50U
478 of mutanolysin (Sigma-aldrich) was added at the lysozyme treatment step. When
479 necessary, an additional phenol:chloroform clean up step was performed to remove
480 the excessive impurities in the samples.

481

482 **Genomic DNA sequencing and assembly**

483 High integrity DNA were sequenced to at least a coverage of 100X on the Illumina
484 Miseq platform (reagent kit v3, 2 x 300bp) at the in-house next-generation sequencing

485 facility in the Singapore Centre for Environmental Life Sciences Engineering. The
486 paired-end sequencing reads were trimmed and assembled *de novo* into contigs on
487 the CLC Genomics Workbench 10.0 (Qiagen). These assembled contigs were
488 subsequently uploaded to the RAST server (Aziz et al., 2008, Brettin et al., 2015,
489 Overbeek et al., 2014) for open reading frame prediction and gene annotation.
490

491 **Phylogenetic analysis**

492 A core genome-based alignment was performed on Parsnp version 1.2 integrated in
493 the Harvest suite (Treangen et al., 2014), with *Sgg* UCN34 (GenBank accession no.
494 FN597254) served as the reference genome. The other *S. gallolyticus* complete
495 closed genomes available on NCBI, *i.e. Sgg* ATCC43143 (GenBank accession no.
496 NC_017576), ATCC BAA_2069 (GenBank accession no. NC_015215) and DSM
497 16831 (GenBank accession no. NZ_CP018822); *Sgp* ATCC43144 (GenBank
498 accession no. NC_015600), NCTC13784 (GenBank accession no. NZ_LS483462)
499 and WUSP067 (GenBank accession no. NZ_CP039457); and two additional draft
500 genomes – *Sgm* ACA-DC-206 (unpublished genome) and the *Sgg* TX20005 (NCBI
501 Assembly ID GCA_000146525) - were also included in the analysis for a clear
502 distinction of these closely related strains. The resulted alignment was subsequently
503 visualized on Gingr (integrated in the Harvest suite) and used for phylogenetic
504 inference on FigTree2 based on approximate-maximum-likelihood algorithm (Price et
505 al., 2010).
506

507 **Comparative genomics**

508 A genome-wide comparison among *Sgg* UCN34, three representative *Sgg* clinical
509 isolates, and three representative *Sgp* was performed using BLAST Ring Image

510 Generator v0.95 based on BLASTn search (Alikhan et al., 2011). Genomic islands of
511 the *Sgg* UCN34 were downloaded from IslandViewer 4(Bertelli et al., 2017). NCBI
512 tBLASTn program was used to identify the protein homologs of the T7SS core
513 components across all the *Sgg* and *Sgp* clinical isolates. Easyfig v2.2.2 was used to
514 visualize the region of interest and to perform pairwise comparisons on the protein
515 sequences using tBLASTx program (Sullivan et al., 2011).
516
517 **Comparative proteomics on the *Sgg* secretomes**
518 Single colonies of *Sgg* UCN34 WT and the Δ*essC* mutant were inoculated into 20 mL
519 of M9YEG and grown to OD_{600nm} 3.0 corresponding to the early stationary phase of
520 growth. The culture supernatant was collected and filtered through 0.22 µm filter units.
521 15 mL of the filtered culture supernatant was first concentrated using Amicon Ultra 15
522 mL Centrifugal Filters (cut-off 10 kDa, Merck), and the proteins were precipitated out
523 of the solution using trichloroacetic acid (TCA) as previously described (Flaugnatti and
524 Journet, 2017). The samples were further processed and analyzed at the Taplin Mass
525 Spectrometry Facility. Briefly, precipitated proteins were first treated with trypsin
526 (Promega) in 50 mM ammonium bicarbonate solution, dried in a speed-vac and were
527 subsequently reconstituted in HPLC solvent A (2.5% acetonitrile, 0.1% formic acid).
528 Peptides were loaded using a Famos auto sampler (LC Packings) into a nano-scale
529 reverse-phase HPLC fused silica capillary column (100 µm inner diameter x ~30 cm
530 length) packed with 2.6 µm C18 spherical silica beads and eluted with increasing
531 concentrations of solvent B (97.5% acetonitrile, 0.1% formic acid). Eluted peptides
532 were subjected to electrospray ionization and entered into an LTQ Orbitrap Velos Pro
533 ion-trap mass spectrometer (Thermo Fisher Scientific). Peptides were detected,
534 isolated, and fragmented to produce a tandem mass spectrum of specific fragment

535 ions for each peptide. Peptide sequences (and hence protein identity) were
536 determined by comparing the fragmentation pattern acquired by the software program
537 Sequest (Thermo Fisher Scientific) (Eng et al., 1994) to the protein sequences on
538 protein databases. All databases include a reversed version of all the sequences and
539 the data was filtered to between a one and two percent peptide false discovery rate.
540
541 Downstream data analysis was performed in R studio using the Differential
542 Enrichment analysis of Proteomics Data 1.10.0 package (Zhang et al., 2018). Proteins
543 that were not detected across all the three replicates of either the UCN34 WT or the
544 Δ*essC* mutant were excluded from the analysis. Next, the total peptide count of the
545 remaining proteins (n = 162) was scaled and variance stabilized (Huber et al., 2002)
546 before the missing values (*i.e.* peptide count = 0) were imputed with the assumption
547 that the missing values are missing not at random (MNAR), using the Deterministic
548 Minimum Imputation (MinDet) strategy (10^{-4} quantile) (Lazar et al., 2016). The MinDet-
549 imputed data were subsequently used for differential expression analysis, with a cut
550 off of false discovery rate of 5%, and a $\log_2$ fold change of 1. The derived data set was
551 plotted using the EnhancedVolcano package (Blighe, 2018).
552
553 **Bacterial toxicity experiments**
554 For an intracellular expression of the *Sgg* T7SSb effectors, *E. coli* DH5α, *E. faecalis*
555 OG1RF, *S. agalactiae* NEM316 and *Sgm* ACA-DC-206 were transformed with
556 pTCVerm-P_{tetO} plasmid series, whereas *P. aeruginosa* PAO1 was transformed with
557 pUCP22-P_{tetO} plasmid series (see **Table S1** for the plasmids used in this study).
558 Overnight cultures of these strains were serially diluted in tenfold increments and 5 µL
559 of each dilution was spotted onto LB or BHI agar supplemented with an appropriate

560 amount of antibiotics and inducer (kanamycin, 50 or 500 µg/mL; anhydrotetracycline,
561 100 or 500 ng/mL). Agar plates incubated overnight at 37°C were documented on the
562 Gel Doc XR+ Documentation System (Bio-Rad).
563
564 For the identification of the TelE cognate immunity protein, the *E. coli* DH5α carrying
565 pTCVerm-P_{tetO}::TelE was transformed with compatible plasmids of either pJN105
566 (vector control) or pJN105 containing each individual open reading frames found
567 downstream of *telE*. The expression of these open reading frames was driven by
568 arabinose-inducible promoter (P_{BAD}). Overnight cultures of these strains were first 50x-
569 diluted into sterile LB broth supplemented with antibiotics and incubated at 37°C with
570 agitation at 200 rpm for 1.5 hours. These bacterial cultures at the early log-phase of
571 growth were supplemented with arabinose to a final concentration of 0.5%, and
572 incubated further for 3 hours, before being subjected to serial dilution and spotted onto
573 LB agar supplemented with an appropriate amount of antibiotics and inducer
574 (kanamycin, 50 µg/mL; gentamicin, 20 µg/mL; anhydrotetracycline, 100 ng/mL;
575 arabinose, 0.5%). Agar plates incubated overnight at 37°C were documented on the
576 Gel Doc XR+ Documentation System (Bio-Rad).
577
578 ***In silico* TelE prediction and homologs identification**
579 The TelE amino acid sequence from the reference strain *Sgg* UCN34 (protein ID
580 CBI13054.1) was uploaded onto the web-based I-TASSER server (Roy et al., 2010,
581 Yang et al., 2015, Yang and Zhang, 2015) for protein structural prediction using the
582 default settings. The same amino acid sequence was also used to identify TelE
583 homologs among the clinical isolates using the NCBI tBLASTn program. Hit proteins

584 were considered as TelE homologs if significant similarities (e-value <10^{-11}) were
585 identified at the N terminus, C terminus, or the full-length protein.
586
587 **Multiple sequence alignment (MSA) and sequence logo generation**
588 MSA was performed on the online MAFFT server version 7
589 (https://mafft.cbrc.jp/alignment/server/) with E-INS-i iterative refinement method that
590 assumes all the sequences shared the same conserved motifs (Katoh et al., 2019,
591 Kuraku et al., 2013). Regional sequence logo (probability-weighted Kullback-Leibler
592 logo type) was generated on web-based Seq2Logo 2.0 server
593 (http://www.cbs.dtu.dk/biotools/Seq2Logo/), with the Heuristics clustering method
594 applied (Thomsen and Nielsen, 2012).
595
596 **Fluorescence microscopy imaging**
597 For time-lapse microscopy imaging, log-phase *E. coli* (OD_{600nm} ~ 1.0 Abs) cultured in
598 LB was mixed with propidium iodide (1:1000 diluted from 20 mM stock solution,
599 Thermo Fisher Scientific) and anhydrotetracycline at a final concentration of 500
600 ng/mL (Sigma Aldrich), and spotted onto an 1% (w/v) low melting agarose pad
601 (bioWORLD). Samples were immediately imaged with the microscope Axio Observer
602 7 fitted with Plan Apochromat 40x/0.95 dry lens for phase contrast imaging, or Plan
603 Apochromat 100x/1.4 oil lens for differential interference contrast imaging (Carl Zeiss),
604 with temperature maintained at 37°C, and aided with Definite Focus.2 for autofocus.
605 When necessary, samples were excited with filter wheel 555/30 to detect red
606 fluorescence signals.
607

608 To image *E. coli* expressing TelE-sfGFP with or without TipE, overnight cultures were
609 diluted 50x into sterile LB broth, incubated at 37°C with 200 rpm for 1.5 hours to grow
610 to early log-phase. Samples were first induced with 0.5% arabinose for 1.5 hours and
611 were subsequently supplemented with anhydrotetracycline to a final concentration of
612 500 ng/mL and propidium iodide (1:1000 diluted from 20 mM stock solution, Thermo
613 Fisher Scientific), and incubated further for 2 hours at 37°C without agitation. Following
614 incubation, samples were spotted onto an 1% low melting agarose pad (bioWORLD),
615 and imaged with the microscope Axio Observer Z1 fitted with Plan Neofluar 100x/1.3
616 oil lens, with excitation at 488 nm (GFP) and 535 nm (PI).
617
618 **Protein expression, localization, purification and pulldown**
619 Protein samples were in general prepared as follow: overnight cultures of *E. coli*
620 BL21(DE3) harboring the respective protein expression pET28b plasmids were 50x
621 diluted into sterile LB broth supplemented with 50 µg/mL of kanamycin and grown to
622 an optical density of ~0.7 Abs measured at 600 nm. Subsequently, cultures were
623 added with isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of
624 100 µM and incubated for an additional 3 hours at 37°C, 200 rpm. Following incubation,
625 cells were harvested by centrifugation at 8,000 xg for 10 minutes, and the resulted cell
626 pellets were stored at -80°C (for protein purification) or processed immediately (for
627 protein localization). When needed, cells were resuspended in 100 mM HEPES pH
628 7.5 and lysed by sonication (10s on-off pulse cycles, at 20 or 40% amplitude, Sonics
629 Vibracell VCX750). Unbroken cells were removed by centrifugation at 8,000 xg for 10
630 minutes at 4°C, and the resulted supernatant was ultracentrifuged (Beckman Coulter
631 XPN100) at 35,000 rpm for 1 hour at 4°C to separate the soluble (supernatant) and
632 membrane proteins (pellet).

633

634 For protein localization study, the membrane pellets were resuspended in 100 mM
635 HEPES pH 7.5 containing 1% sarkosyl and were ultracentrifuged further to separate
636 the soluble inner membrane proteins and the insoluble outer membrane proteins. Both
637 the soluble proteins and inner membrane fractions were subjected to TCA precipitation
638 prior to SDS-PAGE analysis. Periplasmic proteins were extracted as described
639 previously(Quan et al., 2013), except that the Tris buffer was replaced with HEPES
640 buffer.

641

642 For protein purification, the membrane pellets were resuspended to a final
643 concentration of 100 mg/mL in Buffer A (100 mM HEPES pH 7.5, 500 mM NaCl and
644 1% n-dodecyl-β-D-maltoside (DDM)) and incubated for 1 hour at 4°C aided with top-
645 down rotation. The solubilized membrane proteins were subsequently supplemented
646 with imidazole to a final concentration of 50 mM and incubated with nickel-immobilized
647 HisLink resins (Promega) that had been pre-calibrated in Buffer B (100 mM HEPES
648 pH 7.5, 500 mM NaCl, 50 mM imidazole and 0.1% DDM), for 30 minutes at 4°C with
649 top-down rotation. Proteins-bound HisLink resins were washed with 30X resin volume
650 of ice-cold Buffer B, and finally rinsed with ice-cold buffer containing 100 mM HEPES
651 pH 7.5, 500 mM NaCl, 500 mM imidazole and 0.1% DDM to elute the proteins.

652

653 For protein pulldown, the proteins-bound HisLink resins were incubated with cellular
654 lysate containing HA-TipE prepared in Buffer B for 20 minutes at 4°C aided with top-
655 down rotation. The resins were subsequently washed and the bound proteins were
656 eluted as described above.

657

658 **Purification of TelE-HA**

659 TelE-HA was expressed in *E. coli* cells BL21(DE3) transformed with the pET28b/TelE-
660 HA plasmid. At OD_{600} = 0.8, TelE-HA production was induced with 0.5 mM IPTG in
661 Terrific Broth media (Sigma) supplemented with Kanamycin (50 µg/ml). After overnight
662 incubation at 18 °C, 160 rpm, cells were collected by centrifugation at 10 000 xg for
663 20 min at 4°C and the media was discarded. Cell pellet was resuspended in Lysis
664 buffer containing 50 mM HEPES pH 8, 1 mg/ml lysozyme, 4 tablets of Protease
665 Inhibitors (Roche), 20X-diluted benzonase. Cells were lysed by sonication (6 bursts of
666 1 min each at 50%, with 1 min gaps). Unbroken cells were discarded by centrifugation
667 at 6000 xg for 15 min and membranes were isolated after the cellular lysate was
668 ultracentrifuged at 100,000 xg for 1 h at 4 °C. Total membrane pellets were
669 resuspended in solubilization buffer containing 50 mM NaCl, 50 mM Tris pH8, 1 tablet
670 of Protease Inhibitors, 1 mM TCEP, 5% n-Octyl-β-D-Glucopyranoside (OG, Anatrace),
671 aided with gentle stirring for 30 min at 6 °C. The resuspended membrane was again
672 ultracentrifuged at 195 000 xg for 45 min, and the non-solubilized membranes were
673 removed. Solubilized membrane in the supernatant was loaded overnight onto an
674 Hitrap SP HP 5ml column (Pharmacia) pre-equilibrated in solubilization buffer. TelE-
675 HA was eluted using a 100 mM to 2 M NaCl gradient in buffer containing 50 mM Tris
676 pH8, 0.5 mM TCEP, 0.7% OG. Fractions were analysed by SDS-PAGE. Fractions
677 containing TelE-HA were pooled and dialyzed overnight against 50 mM Tris pH8, 0.5
678 mM TCEP, 0.7% OG, 150 mM NaCl. TelE-HA was concentrated 10 times on a
679 Microcon-10kDa Centrifugal Filter (Millipore). Final protein concentration was
680 determined to be 1.1 mg/ml by Bradford assay, using BSA in the presence of 0.7 %
681 OG as a standard. Glycerol was added at 8 % final and the aliquoted protein was flash-
682 frozen in liquid nitrogen. The protein identity was confirmed using mass-spectrometry.
683

684 **SDS-PAGE and Western blot**

685 Protein samples were mixed with an equal amount of 2X Laemmli reducing buffer and
686 boiled for 10 minutes before being subjected to SDS-PAGE analysis on 12%
687 polyacrylamide gels. Proteins were stained with InstantBlue solution (Expedeon) for
688 visualization or were transferred onto PVDF membranes using iBlot Dry Blotting
689 System (Thermo Fisher Scientific) at 20V for 7 minutes. Following protein transfer, the
690 PVDF membranes were immediately blocked with 1X casein blocking buffer (Sigma-
691 Aldrich), and were probed with either horseradish peroxidase (HRP)-conjugated
692 primary antibodies against His-tag (Qiagen) or with rabbit primary antibodies against
693 HA-tag (Sigma-Aldrich) followed by anti-rabbit HRP-conjugated secondary antibodies.
694 Membranes were thoroughly rinsed with PBS containing 0.1% Tween 20 between the
695 incubations. Chemiluminescent signals were detected using Amersham ImageQuant
696 800 System (GE Healthcare) after a brief incubation of the PVDF membranes with
697 Immobilon Forte Western Chemiluminescent HRP Substrates (EMD Millipore). For
698 reblotting with a different antibody after image acquisition, the membranes were
699 incubated with Restore PLUS Western Blot Stripping Buffer (Thermo Fisher Scientific)
700 for 30 minutes at room temperature, rinsed thoroughly before being blocked with
701 casein blocking buffer and probed with antibodies as described above.

702 **Acknowledgement**

703 We are very grateful to Alexandra Doloy and Nicolas Dmytruk for providing us with the
704 collection of S. *gallolyticus* clinical isolates. We thank Ross Tomaino from Taplin Mass
705 Spectrometry Facility for the mass-spectrometry analysis. We thank Tarek Msadek for
706 critical reading of the manuscript. This work was supported by the National Research
707 Foundation and Ministry of Education Singapore under its Research Centre of
708 Excellence Program (SCELSE). S. Dramsi would like to acknowledge the support of
709 the Institut National contre le Cancer (INCA, grant PLBIO16-025) and from the French
710 Government's Investissement d'Avenir program, Laboratoire d'Excellence Integrative
711 Biology of Emerging Infectious Diseases (grant no. ANR-10-LABX-62-IBEID).

712

713 **Author contributions**

714 Conceptualization, W.K.T.; Methodology, W.K.T. and Y.D.; Investigation, W.K.T. and
715 S. D.; Protein purification, F. G., Resources – clinical isolates, C.P.; Writing –
716 Original Draft, W.K.T.; Writing – Review & Editing, W.K.T., Y.D., C.P., A.F., S.D. and
717 M.G.; Visualization, W.K.T.; Supervision, A.F., S.D. and M.G.; Funding Acquisition,
718 S.D. and M.G.

719 **Figure Legends**

720

721 **Figure 1 - A T7SSb locus specific to *Sgg*.** (A) An overview of the BLASTn-based
722 comparative genomics analysis. From the innermost circle, the reference genome *Sgg*
723 UCN34 (black ring); followed by *Sgp* isolates (Gallo1, Gallo4, Gallo6, blue rings); and
724 *Sgg* isolates (Gallo37, Gallo43 and Gallo51, red rings). The outermost black blocks
725 represent the genomic islands in UCN34 genomes predicted by IslandViewer 4. White
726 spaces within each ring represent a ≤50% similarities of the respective genomic
727 regions to UCN34. (B) A pairwise nucleotide sequence comparison of the genomic
728 region of interest. Each of the T7SSb genes (arranged in the order of *esxA, essA,*
729 *esaB, essB, essC, esaA*) was represented by colored arrows. Light gray arrows
730 represent other open reading frames in this genomic region. Gray shading shows the
731 similarities between the sequences. (C) Summary of BLAST search using NCBI
732 tBLASTn program, queried against the protein sequences of each T7SSb components
733 (EsxA, EssA, EsaB, EssB, EssC, EsaA) from the reference genomes UCN34. The
734 identification of homologs was based on RAST annotation, or significant hits from the
735 BLAST search. (D) The general gene arrangement of the T7SS gene locus in *Sgg*, *S.*
736 *intermedius* and *S. aureus*. Each T7SSb genes was color-coded and labelled
737 accordingly. The open reading frames flanked by *esxA* and *essA* in the *S. intermedius*
738 B196 were excluded in this figure for a clear presentation of the T7SSb gene locus.

739

740 **Figure 2. A *Sgg* T7SSb effector exhibits specific toxicity in Gram-negative**
741 **bacteria.** (A) A volcano plot of the relative abundance of all proteins (n=162) detected
742 in the secretomes of the *Sgg* UCN34 wildtype versus the Δ*essC* mutant, ranked by
743 false discovery rate (adjusted *p*-value). Each dot represents an individual protein. The

744 vertical dotted line marked the cut-off value of $\log_2$ fold change of 1, and the horizontal
745 dotted line marked the cut-off value of adjusted *p*-value of 0.05. The proteins that met
746 the cut-off requirements were represented by red dots and labelled with the protein
747 identity. (B) The domain architecture of the six proteins labelled in (A). Both
748 Gallo_1068 and Gallo_1574 are homologous to the lipid II phosphatase TelC and the
749 conserved aspartate-rich motifs essential for the protein function were indicated. (C)
750 Viability of the bacterial cells harboring either an empty vector or a vector encoding
751 the full-length *telE* with the expression inducible by anhydrotetracycline (aTc). Bar
752 chart on the right showed the logarithm-transformed CFU count of the *telE*-expressing
753 cells in relative to the control. Error bars represent mean + standard deviation (n = 3).
754
755 **Figure 3. T7SSb Effector TelE contains a C-terminal glycine zipper motif**
756 **essential for its activity.** (A) The genomic region encoding the T7SSb core
757 components, TelE and TipE in *Sgg* UCN34. Black, T7SSb core components; blue,
758 TelE (LXG domain-containing protein); red, DUF5085 domain-containing protein);
759 white, hypothetical proteins. (B) Viability of the *E. coli* cells harboring either an empty
760 vector or a vector encoding various *telE* variants (*telE1* to *telE7*) with the expression
761 inducible by anhydrotetracycline (aTc). Bar chart on the right showed the logarithm-
762 transformed CFU count of the *telE*-expressing cells in relative to the control. Error bars
763 represent mean + standard deviation (n = 3). (C) Alignment of the C-terminus
764 sequence of the *telE* variants expressed in (B) revealed a conserved region
765 (underlined) enriched with hydrophobic residues. (D) Sequence logo of the underlined
766 region in (C) generated from all aligned TelE variants identified a sequence motif
767 highly similar to glycine zipper motif (GxxxGxxxG). The number at the bottom of the
768 sequence logo represents the position of the respective amino acid residue in

769 reference to the TelE sequence from UCN34. Shrunk residues (*i.e.* residues at position
770 462 to 465) reflected the presence of gaps in the sequence alignment. (E) Viability of
771 the *E. coli* cells harboring either an empty vector or a vector encoding the mutated *telE*
772 with the expression inducible by anhydrotetracycline (aTc). Bar chart on the right
773 showed the logarithm-transformed CFU count of the *telE*-expressing cells in relative
774 to the control. Error bars represent mean + standard deviation (n = 3).
775
776 **Figure 4. TelE is a pore-forming toxin.** (A) The amino acid sequence of TelE was
777 color coded to indicate the predicted secondary structure. (B) Differential interference
778 contrast (DIC) micrograph of *E. coli* cells harboring an empty vector or a vector
779 encoding *telE* with the expression inducible by anhydrotetracycline (aTc). The
780 physiological changes of the *E. coli* cells following protein induction were documented
781 at the indicated time. Cells with membrane lesions were stained red by propidium
782 iodide. Arrows indicate the region with visible cell shrinkage. Scale bar, 2 μm. (C)
783 Phase contrast micrograph of *telE*-expressing *E. coli* documented at the indicated time.
784 Arrows indicate the region with visible cell shrinkage. Scale bar, 2 μm. (D) Total
785 proteins collected from the respective cellular fractions were resolved by SDS-PAGE
786 and stained with Coomassie blue (upper), or transferred onto a PVDF membrane for
787 anti-His immunoblot (lower). The two most intense proteins bands in the outer
788 membrane fractions are the outer membrane proteins OmpF/C (~37 kDa) and OmpA
789 (~35 kDa) serving as the control to approximate the purity of the soluble and inner
790 membrane fractions. T, total cellular lysate; L, cell-free lysate; S, soluble proteins; IM,
791 sarkosyl-souble inner membrane proteins; OM, sarkosyl-insoluble outer membrane
792 proteins.
793

794 **Figure 5. The cognate immunity protein TipE antagonizes TelE toxicity by**
795 **preventing TelE membrane insertion.** (A) Viability of *E. coli* cells harboring two
796 vectors for single expression of *telE* inducible by anhydrotetracycline (aTc), or co-
797 expression of *telE* and the indicated gene candidates with the expression inducible by
798 arabinose. Bar chart on the right showed the logarithm transformed CFU count of the
799 cells co-expressing two genes in relative to the control. Error bars represent mean +
800 standard deviation (n = 3). (B) Total proteins collected from the respective cellular
801 fractions were resolved by SDS-PAGE and stained with Coomassie blue (upper) or
802 transferred onto a PVDF membrane for anti-HA immunoblot (lower). The two most
803 intense protein bands in the outer membrane fractions are the outer membrane
804 proteins OmpF/C (~37 kDa) and OmpA (~35 kDa) serving as the control to
805 approximate the purity of the soluble and inner membrane fractions. OmpF/C (~37
806 kDa) also found in the periplasm was used as the positive control for the periplasmic
807 extract. T, total cellular lysate; L, cell-free lysate; S, soluble proteins; IM, sarkosyl-
808 souble inner membrane proteins; OM, sarkosyl-insoluble outer membrane proteins; P,
809 periplasmic proteins. (C) Immunoblot with anti-His or anti-HA on the proteins eluted
810 from the TelE-bound nickel-immobilized resins incubated with cellular lysate
811 containing HA-TipE. The TelE-free resins subjected to the same treatment were used
812 as the control. (D) Fluorescence micrograph of *E. coli* cells expressing superfolder
813 GFP (sfGFP), TelE-sfGFP, or co-expressing TipE and TelE-sfGFP. Cells with
814 membrane lesions were stained red by propidium iodide. Scale bar 10 μm; 2 μm for
815 inlets. Bar chart (bottom) showed the quantification of cells fluoresced green or red (n
816 ≈ 300).
817

818 **Figure 6. Purified TelE readily permeabilize *E. coli* pks+ cells when added in the**
819 **external milieu.** (A) Flow-cytometry analysis of the CFSE-stained *E. coli* cells showing
820 the fluorescence intensity of the *E. coli* cell population treated with various detergents
821 or purified TelE protein for 10 minutes at 37°C. Upper panel, *E. coli* cells treated with
822 n-Octyl-β-D-Glucopyranoside (OG) at the concentration of 0.7% (present in TelE
823 elution buffer), 2.1% and 8.4%. Middle panel, *E. coli* cells treated with TelE depleted
824 of OG either using β-cyclodextrin (TelE1) or BioBeads (TelE2 and TelE3). Lower panel,
825 *E. coli* cells treated with pore-forming detergent TritonX100 and Tween20 at 0.05% as
826 positive control. (B) Representation of the flow-cytometry results as median
827 fluorescence ± CV. (C) Fluorescence images of live CFSE-stained *E. coli* cells treated
828 with TelE1, TelE2 and TelE3. Images were acquired with Zeiss microscope fitted with
829 100X objective lens. Scale bar 2.5 µm.
830
831 **Figure S1. Phylogenetic tree of the clinical isolates sequenced in this study.**
832 Phylogenetic inference of the isolates was performed using approximate-maximum-
833 likelihood algorithm based on the core-genome alignment using Parsnp version 1.2.
834 Scale bar represents substitution per variant site within the core genome. The isolates
835 sequenced in this study were labeled in black font, whereas the reference strains *Sgp*
836 (NCTC13784, ATCC43144 and WUSP067) and *Sgg* (DSM16831, ATCC43143,
837 BAA2069 and UCN34) with complete genomes available on NCBI GenBank, and a
838 draft genome of *Sgm* (strain ACA-DC-206) were labeled in blue font. Genomes used
839 for comparative genomics analysis as shown in Figure 1A were highlighted with darker
840 color shades. Bootstrap value for all the major branches ranges from 0.97 to 1.
841 Genomes labelled with asterisk contain a different T7SSb gene locus arrangement
842 than the UCN34 reference strain.

843

844 **Figure S2. Antibacterial activity of *Sgg* T7SSb effectors.** (A) Shortlisted proteins
845 that were significantly more abundant in the secretome of the UCN34 wildtype than
846 the Δ*essC* mutant. The total peptide count of each protein was expressed as means
847 ± standard deviation from three biological replicates. N.D., not detected; N. A., not
848 applicable. (B) Viability of *E. coli* cells harboring an empty vector or a vector encoding
849 *Sgg* T7SSb effectors with the expression inducible by anhydrotetracycline (aTc). Bar
850 chart on the right showed the logarithm-transformed CFU count of the cells expressing
851 the T7SSb effectors in relative to the control. Error bars represent mean + standard
852 deviation (n = 3). (C) Viability of *Sgm* ACA-DC-206 and *E. faecalis* OG1RF cells
853 harboring an empty vector or a vector encoding *telC1* and *telC2* with the expression
854 inducible by anhydrotetracycline (aTc). Bar chart on the right showed the logarithm-
855 transformed CFU count of the cells expressing *telC1* or *telC2* in relative to the control.
856 Error bars represent mean + standard deviation (n = 3).

857

858 **Figure S3. The prevalence of TelE homologs among *Sgg* isolates.** The
859 phylogenetic inference of the *Sgg* isolates was shown on the left. Genomes in black
860 font represent the clinical isolates sequenced in this study, whereas the genomes in
861 blue font represent the complete genomes publicly available. Scale bar represents
862 substitution per variant site within the core genome. Table on the right summarized
863 the distribution of TelE homologs among *Sgg* isolates. Based on the C-terminal
864 sequence (shown in Figure 3B), the identified TelE homologs were broadly
865 categorized into 7 TelE variants and named as TelE1 to TelE7, where TelE1 refers to
866 the reference TelE from strain UCN34. Bootstrap values of all the major branches are
867 1.

868 **Figure S4. Multiple sequence alignment of various TelE subtypes.** Amino acid
869 sequences of TelE1 to TelE7 subtypes were aligned using online MAFFT server
870 version 7 with E-INS-i iterative refinement method that assumes all the sequences
871 shared the same conserved motifs. The aligned sequences were viewed using MView
872 (Brown et al., 1998) available on EMBL-EBI.
873
874 **Figure S5. Impact of epitope tags or protein fusion on TelE activity.** (A) Viability
875 of *E. coli* cells harboring an empty vector or a vector encoding *telE* fused with epitope
876 tags or superfolder GFP with the expression inducible by anhydrotetracycline (aTc).
877 Bar chart on the right showed the logarithm-transformed CFU count of the cells
878 expressing the *telE* variants in relative to the control. Error bars represent mean +
879 standard deviation (n ≥ 3). (B) Growth kinetics of the *E. coli* BL21(DE3) harboring
880 either the empty vector or a vector for IPTG-inducible expression of *telE*-His or co-
881 expression of *telE*-His and HA-*tipE.* Arrow indicates the time point when the gene
882 expression was induced. (C) Total protein collected from the respective lysates was
883 resolved by SDS-PAGE and stained with Coomassie blue. I, insoluble lysate; S,
884 soluble lysate. Arrow indicates the protein band corresponding to TelE and TipE,
885 respectively.

886 ## References

887 Abdallah, A. M., Gey Van Pittius, N. C., Champion, P. A., Cox, J., Luirink, J., Vandenbroucke-
888 Grauls, C. M., Appelmelk, B. J. & Bitter, W. 2007. Type VII secretion--mycobacteria
889 show the way. *Nat Rev Microbiol,* 5**,** 883-91.
890 Alikhan, N. F., Petty, N. K., Ben Zakour, N. L. & Beatson, S. A. 2011. BLAST Ring Image
891 Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics,* 12**,** 402.
892 Anderson, M., Aly, K. A., Chen, Y. H. & Missiakas, D. 2013. Secretion of atypical protein
893 substrates by the ESAT-6 secretion system of Staphylococcus aureus. *Mol Microbiol,*
894 90**,** 734-43.
895 Anderson, M., Chen, Y. H., Butler, E. K. & Missiakas, D. M. 2011. EsaD, a secretion factor for
896 the Ess pathway in Staphylococcus aureus. *J Bacteriol,* 193**,** 1583-9.
897 Anderson, M., Ohr, R. J., Aly, K. A., Nocadello, S., Kim, H. K., Schneewind, C. E., Schneewind,
898 O. & Missiakas, D. 2017. EssE Promotes Staphylococcus aureus ESS-Dependent Protein
899 Secretion To Modify Host Immune Responses during Infection. *J Bacteriol,* 199.
900 Arthur, J. C., Perez-Chanona, E., Muhlbauer, M., Tomkovich, S., Uronis, J. M., Fan, T. J.,
901 Campbell, B. J., Abujamel, T., Dogan, B., Rogers, A. B., Rhodes, J. M., Stintzi, A.,
902 Simpson, K. W., Hansen, J. J., Keku, T. O., Fodor, A. A. & Jobin, C. 2012. Intestinal
903 inflammation targets cancer-inducing activity of the microbiota. *Science,* 338**,** 120-3.
904 Aymeric, L., Donnadieu, F., Mulet, C., Du Merle, L., Nigro, G., Saffarian, A., Berard, M., Poyart,
905 C., Robine, S., Regnault, B., Trieu-Cuot, P., Sansonetti, P. J. & Dramsi, S. 2018.
906 Colorectal cancer specific conditions promote Streptococcus gallolyticus gut
907 colonization. *Proc Natl Acad Sci U S A,* 115**,** E283-E291.
908 Aziz, R. K., Bartels, D., Best, A. A., Dejongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes,
909 S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek,
910 R. A., Mcneil, L. K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G. D., Reich, C.,
911 Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. & Zagnitko, O. 2008. The RAST Server:
912 rapid annotations using subsystems technology. *BMC Genomics,* 9**,** 75.
913 Bae, T., Kozlowicz, B. & Dunny, G. M. 2002. Two targets in pCF10 DNA for PrgX binding: their
914 role in production of Qa and prgX mRNA and in regulation of pheromone-inducible
915 conjugation. *J Mol Biol,* 315**,** 995-1007.
916 Bertelli, C., Laird, M. R., Williams, K. P., Simon Fraser University Research Computing, G., Lau,
917 B. Y., Hoad, G., Winsor, G. L. & Brinkman, F. S. L. 2017. IslandViewer 4: expanded
918 prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res,* 45**,** W30-
919 W35.
920 Biswas, I., Gruss, A., Ehrlich, S. D. & Maguin, E. 1993. High-efficiency gene inactivation and
921 replacement system for gram-positive bacteria. *J Bacteriol,* 175**,** 3628-35.
922 Blighe, K., S Rana, And M Lewis 2018. EnhancedVolcano: Publication-ready volcano plots with
923 enhanced colouring and labeling.
924 Boleij, A., Van Gelder, M. M. H. J., Swinkels, D. W. & Tjalsma, H. 2011. Clinical Importance of
925 Streptococcus gallolyticus Infection Among Colorectal Cancer Patients: Systematic
926 Review and Meta-analysis. *Clinical Infectious Diseases,* 53**,** 870-878.
927 Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., Olson, R., Overbeek, R.,
928 Parrello, B., Pusch, G. D., Shukla, M., Thomason, J. A., 3rd, Stevens, R., Vonstein, V.,
929 Wattam, A. R. & Xia, F. 2015. RASTtk: a modular and extensible implementation of the
930 RAST algorithm for building custom annotation pipelines and annotating batches of
931 genomes. *Sci Rep,* 5**,** 8365.

932 Brown, N. P., Leroy, C. & Sander, C. 1998. MView: a web-compatible database search or
933 multiple alignment viewer. *Bioinformatics,* 14**,** 380-1.
934 Burts, M. L., Dedent, A. C. & Missiakas, D. M. 2008. EsaC substrate for the ESAT-6 secretion
935 pathway and its role in persistent infections of Staphylococcus aureus. *Mol Microbiol,*
936 69**,** 736-46.
937 Burts, M. L., Williams, W. A., Debord, K. & Missiakas, D. M. 2005. EsxA and EsxB are secreted
938 by an ESAT-6-like system that is required for the pathogenesis of Staphylococcus
939 aureus infections. *Proc Natl Acad Sci U S A,* 102**,** 1169-74.
940 Cao, Z., Casabona, M. G., Kneuper, H., Chalmers, J. D. & Palmer, T. 2016. The type VII secretion
941 system of Staphylococcus aureus secretes a nuclease toxin that targets competitor
942 bacteria. *Nat Microbiol,* 2**,** 16183.
943 Chen, W. C., Lee, P. I., Lin, H. C., Chang, L. Y., Lee, T. F., Chen, J. M. & Hsueh, P. R. 2020.
944 Clustering of Streptococcus gallolyticus subspecies pasteurianus bacteremia and
945 meningitis in neonates. *J Microbiol Immunol Infect*.
946 Corredoira, J., Grau, I., Garcia-Rodriguez, J. F., Alonso-Garcia, P., Garcia-Pais, M. J., Rabunal,
947 R., Garcia-Garrote, F., Ardanuy, C., Coira, A., Lopez-Alvarez, M. J. & Pallares, R. 2015.
948 The clinical epidemiology and malignancies associated with Streptococcus bovis
949 biotypes in 506 cases of bloodstream infections. *J Infect,* 71**,** 317-25.
950 Corredoira, J. C., Alonso, M. P., Garcia-Pais, M. J., Rabunal, R., Garcia-Garrote, F., Lopez-Roses,
951 L., Lancho, A., Coira, A., Pita, J., Velasco, D., Lopez-Alvarez, M. J., Tjalsma, H. & Varela,
952 J. 2014. Is colonoscopy necessary in cases of infection by Streptococcus bovis biotype
953 II? *Eur J Clin Microbiol Infect Dis,* 33**,** 171-7.
954 Corredoira-Sanchez, J., Garcia-Garrote, F., Rabunal, R., Lopez-Roses, L., Garcia-Pais, M. J.,
955 Castro, E., Gonzalez-Soler, R., Coira, A., Pita, J., Lopez-Alvarez, M. J., Alonso, M. P. &
956 Varela, J. 2012. Association between bacteremia due to Streptococcus gallolyticus
957 subsp. gallolyticus (Streptococcus bovis I) and colorectal neoplasia: a case-control
958 study. *Clin Infect Dis,* 55**,** 491-6.
959 Cuevas-Ramos, G., Petit, C. R., Marcq, I., Boury, M., Oswald, E. & Nougayrede, J. P. 2010.
960 Escherichia coli induces DNA damage in vivo and triggers genomic instability in
961 mammalian cells. *Proc Natl Acad Sci U S A,* 107**,** 11537-42.
962 Dai, Y., Wang, Y., Liu, Q., Gao, Q., Lu, H., Meng, H., Qin, J., Hu, M. & Li, M. 2017. A Novel Esat-
963 6 Secretion System-Secreted Protein EsxX of Community-Associated Staphylococcus
964 aureus Lineage ST398 Contributes to Immune Evasion and Virulence. *Front Microbiol,*
965 8**,** 819.
966 Danne, C., Guerillot, R., Glaser, P., Trieu-Cuot, P. & Dramsi, S. 2013. Construction of isogenic
967 mutants in Streptococcus gallolyticus based on the development of new mobilizable
968 vectors. *Research in Microbiology,* 164**,** 973-978.
969 Dekker, J. P. & Lau, A. F. 2016. An Update on the Streptococcus bovis Group: Classification,
970 Identification, and Disease Associations. *J Clin Microbiol,* 54**,** 1694-1699.
971 Dunny, G. M., Lee, L. N. & Leblanc, D. J. 1991. Improved electroporation and cloning vector
972 system for gram-positive bacteria. *Appl Environ Microbiol,* 57**,** 1194-201.
973 Eng, J. K., Mccormack, A. L. & Yates, J. R. 1994. An approach to correlate tandem mass spectral
974 data of peptides with amino acid sequences in a protein database. *J Am Soc Mass*
975 *Spectrom,* 5**,** 976-89.
976 Flaugnatti, N. & Journet, L. 2017. Identification of Effectors: Precipitation of Supernatant
977 Material. *Methods Mol Biol,* 1615**,** 459-464.

978 Franco, A. T., Johnston, E., Krishna, U., Yamaoka, Y., Israel, D. A., Nagy, T. A., Wroblewski, L.
979 E., Piazuelo, M. B., Correa, P. & Peek, R. M., Jr. 2008. Regulation of gastric
980 carcinogenesis by Helicobacter pylori virulence factors. *Cancer Res,* 68**,** 379-87.
981 Fridman, C. M., Keppel, K., Gerlic, M., Bosis, E. & Salomon, D. 2020. A comparative genomics
982 methodology reveals a widespread family of membrane-disrupting T6SS effectors. *Nat*
983 *Commun,* 11**,** 1085.
984 Garcia-Bayona, L., Guo, M. S. & Laub, M. T. 2017. Contact-dependent killing by Caulobacter
985 crescentus via cell surface-associated, glycine zipper proteins. *Elife,* 6.
986 Georgalaki, M. D., Van Den Berghe, E., Kritikos, D., Devreese, B., Van Beeumen, J.,
987 Kalantzopoulos, G., De Vuyst, L. & Tsakalidou, E. 2002. Macedocin, a food-grade
988 lantibiotic produced by Streptococcus macedonicus ACA-DC 198. *Appl Environ*
989 *Microbiol,* 68**,** 5891-903.
990 Green, M. R., Sambrook, J. & Sambrook, J. 2012. *Molecular cloning : a laboratory manual,*
991 Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press.
992 Holberger, L. E., Garza-Sanchez, F., Lamoureux, J., Low, D. A. & Hayes, C. S. 2012. A novel
993 family of toxin/antitoxin proteins in Bacillus species. *FEBS Lett,* 586**,** 132-6.
994 Huber, W., Von Heydebreck, A., Sultmann, H., Poustka, A. & Vingron, M. 2002. Variance
995 stabilization applied to microarray data calibration and to the quantification of
996 differential expression. *Bioinformatics,* 18 Suppl 1**,** S96-104.
997 Inoue, H., Nojima, H. & Okayama, H. 1990. High-Efficiency Transformation of Escherichia-Coli
998 with Plasmids. *Gene,* 96**,** 23-28.
999 Jamet, A. & Nassif, X. 2015. New players in the toxin field: polymorphic toxin systems in
1000 bacteria. *mBio,* 6**,** e00285-15.
1001 Katoh, K., Rozewicki, J. & Yamada, K. D. 2019. MAFFT online service: multiple sequence
1002 alignment, interactive sequence choice and visualization. *Brief Bioinform,* 20**,** 1160-
1003 1166.
1004 Kim, S., Jeon, T. J., Oberai, A., Yang, D., Schmidt, J. J. & Bowie, J. U. 2005. Transmembrane
1005 glycine zippers: physiological and pathological roles in membrane proteins. *Proc Natl*
1006 *Acad Sci U S A,* 102**,** 14278-83.
1007 Kim, S. Y., Joo, S. I., Yi, J. & Kim, E. C. 2010. A case of Streptococcus gallolyticus subsp.
1008 gallolyticus infective endocarditis with colon cancer: identification by 16S ribosomal
1009 DNA sequencing. *Korean J Lab Med,* 30**,** 160-5.
1010 Klein, T. A., Pazos, M., Surette, M. G., Vollmer, W. & Whitney, J. C. 2018. Molecular Basis for
1011 Immunity Protein Recognition of a Type VII Secretion System Exported Antibacterial
1012 Toxin. *J Mol Biol,* 430**,** 4344-4358.
1013 Kuraku, S., Zmasek, C. M., Nishimura, O. & Katoh, K. 2013. aLeaves facilitates on-demand
1014 exploration of metazoan gene family trees on MAFFT sequence alignment server with
1015 enhanced interactivity. *Nucleic Acids Res,* 41**,** W22-8.
1016 Kwong, T. N. Y., Wang, X. S., Nakatsu, G., Chow, T. C., Tipoe, T., Dai, R. Z. W., Tsoi, K. K. K.,
1017 Wong, M. C. S., Tse, G., Chan, M. T. V., Chan, F. K. L., Ng, S. C., Wu, J. C. Y., Wu, W. K.
1018 K., Yu, J., Sung, J. J. Y. & Wong, S. H. 2018. Association Between Bacteremia From
1019 Specific Microbes and Subsequent Diagnosis of Colorectal Cancer. *Gastroenterology,*
1020 155**,** 383-+.
1021 Lacourse, K. D., Peterson, S. B., Kulasekara, H. D., Radey, M. C., Kim, J. & Mougous, J. D. 2018.
1022 Conditional toxicity and synergy drive diversity among antibacterial effectors. *Nat*
1023 *Microbiol,* 3**,** 440-446.

1024 Lazar, C., Gatto, L., Ferro, M., Bruley, C. & Burger, T. 2016. Accounting for the Multiple Natures
1025 of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare
1026 Imputation Strategies. *J Proteome Res,* 15**,** 1116-25.
1027 Maragkoudakis, P. A., Papadelli, M., Georgalaki, M., Panayotopoulou, E. G., Martinez-
1028 Gonzalez, B., Mentis, A. F., Petraki, K., Sgouras, D. N. & Tsakalidou, E. 2009. In vitro
1029 and in vivo safety evaluation of the bacteriocin producer Streptococcus macedonicus
1030 ACA-DC 198. *Int J Food Microbiol,* 133**,** 141-7.
1031 Mcclain, M. S., Iwamoto, H., Cao, P., Vinion-Dubiel, A. D., Li, Y., Szabo, G., Shao, Z. & Cover, T.
1032 L. 2003. Essential role of a GXXXG motif for membrane channel formation by
1033 Helicobacter pylori vacuolating toxin. *J Biol Chem,* 278**,** 12101-8.
1034 Morrison, D. A., Guedon, E. & Renault, P. 2013. Competence for natural genetic
1035 transformation in the Streptococcus bovis group streptococci S. infantarius and S.
1036 macedonicus. *J Bacteriol,* 195**,** 2612-20.
1037 Ohr, R. J., Anderson, M., Shi, M., Schneewind, O. & Missiakas, D. 2017. EssD, a Nuclease
1038 Effector of the Staphylococcus aureus ESS Pathway. *J Bacteriol,* 199.
1039 Onoyama, S., Ogata, R., Wada, A., Saito, M., Okada, K. & Harada, T. 2009. Neonatal bacterial
1040 meningitis caused by Streptococcus gallolyticus subsp. pasteurianus. *J Med Microbiol,*
1041 58**,** 1252-1254.
1042 Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., Edwards, R. A., Gerdes,
1043 S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A. R., Xia, F. & Stevens, R. 2014. The
1044 SEED and the Rapid Annotation of microbial genomes using Subsystems Technology
1045 (RAST). *Nucleic Acids Res,* 42**,** D206-14.
1046 Pleguezuelos-Manzano, C., Puschhof, J., Rosendahl Huber, A., Van Hoeck, A., Wood, H. M.,
1047 Nomburg, J., Gurjao, C., Manders, F., Dalmasso, G., Stege, P. B., Paganelli, F. L., Geurts,
1048 M. H., Beumer, J., Mizutani, T., Miao, Y., Van Der Linden, R., Van Der Elst, S., Genomics
1049 England Research, C., Garcia, K. C., Top, J., Willems, R. J. L., Giannakis, M., Bonnet, R.,
1050 Quirke, P., Meyerson, M., Cuppen, E., Van Boxtel, R. & Clevers, H. 2020. Mutational
1051 signature in colorectal cancer caused by genotoxic pks(+) E. coli. *Nature,* 580**,** 269-273.
1052 Price, M. N., Dehal, P. S. & Arkin, A. P. 2010. FastTree 2--approximately maximum-likelihood
1053 trees for large alignments. *PLoS One,* 5**,** e9490.
1054 Quan, S., Hiniker, A., Collet, J. F. & Bardwell, J. C. 2013. Isolation of bacteria envelope proteins.
1055 *Methods Mol Biol,* 966**,** 359-66.
1056 Roy, A., Kucukural, A. & Zhang, Y. 2010. I-TASSER: a unified platform for automated protein
1057 structure and function prediction. *Nat Protoc,* 5**,** 725-38.
1058 Rusniok, C., Couve, E., Da Cunha, V., El Gana, R., Zidane, N., Bouchier, C., Poyart, C., Leclercq,
1059 R., Trieu-Cuot, P. & Glaser, P. 2010. Genome sequence of Streptococcus gallolyticus:
1060 insights into its adaptation to the bovine rumen and its ability to cause endocarditis. *J*
1061 *Bacteriol,* 192**,** 2266-76.
1062 Schlegel, L., Grimont, F., Ageron, E., Grimont, P. A. & Bouvet, A. 2003. Reappraisal of the
1063 taxonomy of the Streptococcus bovis/Streptococcus equinus complex and related
1064 species: description of Streptococcus gallolyticus subsp. gallolyticus subsp. nov., S.
1065 gallolyticus subsp. macedonicus subsp. nov. and S. gallolyticus subsp. pasteurianus
1066 subsp. nov. *Int J Syst Evol Microbiol,* 53**,** 631-45.
1067 Shi, H., Westfall, C. S., Kao, J., Odermatt, P. D., Cesar, S., Anderson, S., Sievert, M., Moore, J.,
1068 Gonzalez, C. G., Zhang, L., Elias, J. E., Chang, F., Huang, K. C. & Levin, P. A. 2020.
1069 Starvation induces shrinkage of the bacterial cytoplasm. *bioRxiv***,** 2020.12.06.413849.

1070 Smith, A. W. & Iglewski, B. H. 1989. Transformation of Pseudomonas aeruginosa by
1071 electroporation. *Nucleic Acids Res,* 17**,** 10509.
1072 Sullivan, M. J., Petty, N. K. & Beatson, S. A. 2011. Easyfig: a genome comparison visualizer.
1073 *Bioinformatics,* 27**,** 1009-10.
1074 Sutterlin, H. A., Shi, H., May, K. L., Miguel, A., Khare, S., Huang, K. C. & Silhavy, T. J. 2016.
1075 Disruption of lipid homeostasis in the Gram-negative cell envelope activates a novel
1076 cell death pathway. *Proc Natl Acad Sci U S A,* 113**,** E1565-74.
1077 Taylor, J. C., Gao, X., Xu, J., Holder, M., Petrosino, J., Kumar, R., Liu, W., Hook, M., Mackenzie,
1078 C., Hillhouse, A., Brashear, W., Nunez, M. P. & Xu, Y. 2021. A type VII secretion system
1079 of Streptococcus gallolyticus subsp. gallolyticus contributes to gut colonization and
1080 the development of colon tumors. *PLoS Pathog,* 17**,** e1009182.
1081 Teh, W. K., Dramsi, S., Tolker-Nielsen, T., Yang, L. & Givskov, M. 2019. Increased Intracellular
1082 Cyclic di-AMP Levels Sensitize Streptococcus gallolyticus subsp. gallolyticus to Osmotic
1083 Stress and Reduce Biofilm Formation and Adherence on Intestinal Cells. *J Bacteriol,*
1084 201.
1085 Telford, J. L., Ghiara, P., Dell'orco, M., Comanducci, M., Burroni, D., Bugnoli, M., Tecce, M. F.,
1086 Censini, S., Covacci, A., Xiang, Z. & Et Al. 1994. Gene structure of the Helicobacter
1087 pylori cytotoxin and evidence of its key role in gastric disease. *J Exp Med,* 179**,** 1653-
1088 58.
1089 Thomsen, M. C. & Nielsen, M. 2012. Seq2Logo: a method for construction and visualization
1090 of amino acid binding motifs and sequence profiles including sequence weighting,
1091 pseudo counts and two-sided representation of amino acid enrichment and depletion.
1092 *Nucleic Acids Res,* 40**,** W281-7.
1093 Treangen, T. J., Ondov, B. D., Koren, S. & Phillippy, A. M. 2014. The Harvest suite for rapid
1094 core-genome alignment and visualization of thousands of intraspecific microbial
1095 genomes. *Genome Biol,* 15**,** 524.
1096 Ulhuq, F. R., Gomes, M. C., Duggan, G. M., Guo, M., Mendonca, C., Buchanan, G., Chalmers, J.
1097 D., Cao, Z., Kneuper, H., Murdoch, S., Thomson, S., Strahl, H., Trost, M., Mostowy, S.
1098 & Palmer, T. 2020. A membrane-depolarizing toxin substrate of the Staphylococcus
1099 aureus type VII secretion system mediates intraspecies competition. *Proc Natl Acad*
1100 *Sci U S A,* 117**,** 20836-20847.
1101 Unnikrishnan, M., Constantinidou, C., Palmer, T. & Pallen, M. J. 2017. The Enigmatic Esx
1102 Proteins: Looking Beyond Mycobacteria. *Trends Microbiol,* 25**,** 192-204.
1103 Wen, S. & Moss, S. F. 2009. Helicobacter pylori virulence factors in gastric carcinogenesis.
1104 *Cancer Lett,* 282**,** 1-8.
1105 Wending, G. K., Metzger, P. P., Dozois, E. J., Chua, H. K. & Krishna, M. 2006. Unusual bacterial
1106 infections and colorectal carcinoma - Streptococcus bovis and Clostridium septicum:
1107 Report of three cases. *Diseases of the Colon & Rectum,* 49**,** 1223-1227.
1108 Whitney, J. C., Peterson, S. B., Kim, J., Pazos, M., Verster, A. J., Radey, M. C., Kulasekara, H. D.,
1109 Ching, M. Q., Bullen, N. P., Bryant, D., Goo, Y. A., Surette, M. G., Borenstein, E., Vollmer,
1110 W. & Mougous, J. D. 2017. A broadly distributed toxin family mediates contact-
1111 dependent antagonism between gram-positive bacteria. *Elife,* 6.
1112 Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. & Zhang, Y. 2015. The I-TASSER Suite: protein
1113 structure and function prediction. *Nat Methods,* 12**,** 7-8.
1114 Yang, J. & Zhang, Y. 2015. I-TASSER server: new development for protein structure and
1115 function predictions. *Nucleic Acids Res,* 43**,** W174-81.

1116 Zhang, X., Smits, A. H., Van Tilburg, G. B., Ovaa, H., Huber, W. & Vermeulen, M. 2018.
1117 Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nat Protoc,* 13**,**
1118 530-550.
1119 Zhang, Y., Li, S., Zhu, C. G., Guo, Y. L., Wu, N. Q., Xu, R. X., Dong, Q., Liu, G. & Li, J. J. 2015. Risk
1120 Factors, Coronary Severity, Outcome and ABO Blood Group: A Large Chinese Han
1121 Cohort Study. *Medicine (Baltimore),* 94**,** e1708.
1122

Figure 1

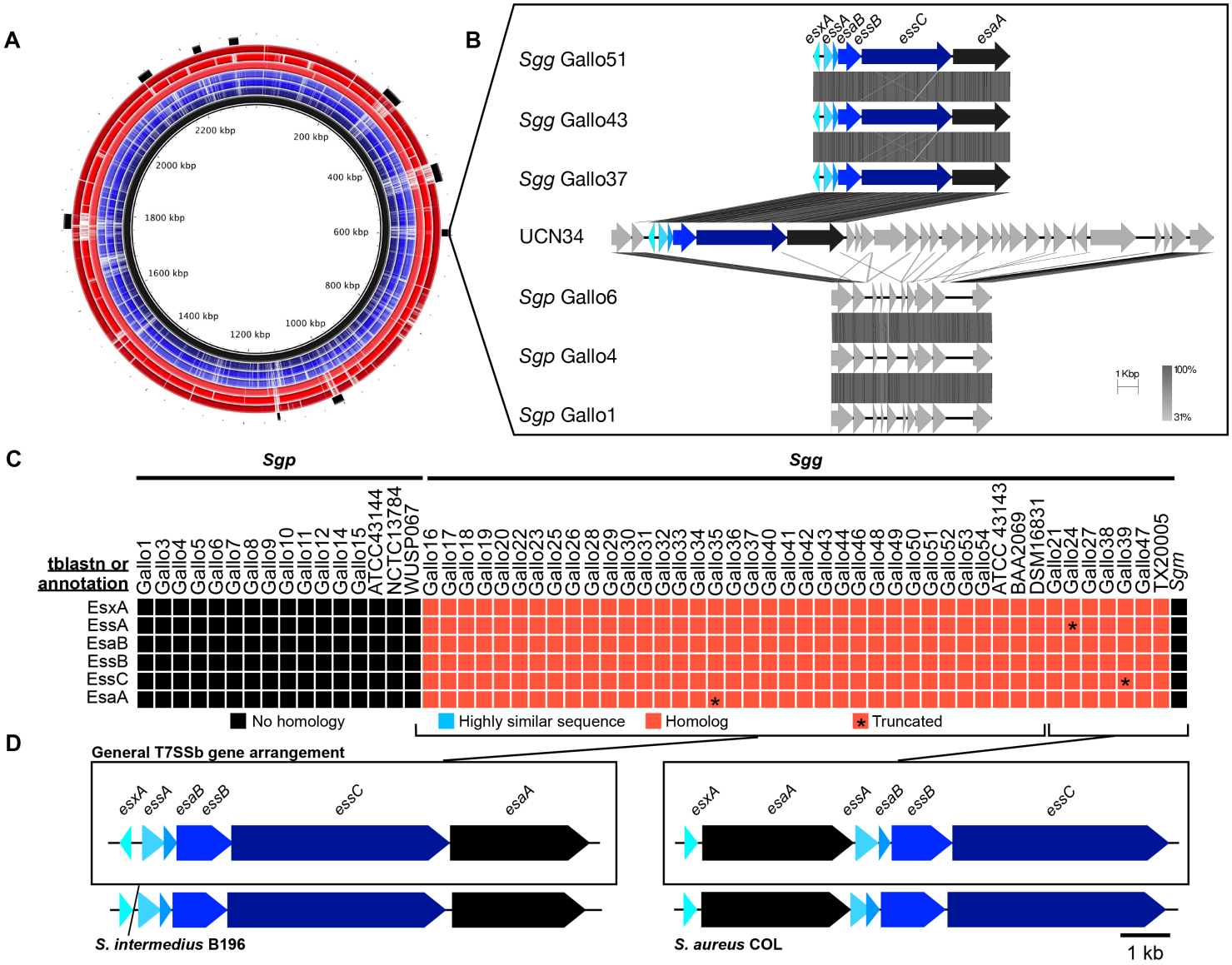

A
B
2200 kbp
200 kbp
2000 kbp
400 kbp
1800 kbp
600 kbp
1600 kbp
800 kbp
1400 kbp
1000 kbp
1200 kbp
esxA
essA
esaB
essB
essC
esaA
Sgg Gallo51
Sgg Gallo43
Sgg Gallo37
UCN34
Sgp Gallo6
Sgp Gallo4
Sgp Gallo1
1 Kbp
100%
31%
C
Sgp
Sgg
tblastn or annotation
Gallo1
Gallo3
Gallo4
Gallo5
Gallo6
Gallo7
Gallo8
Gallo9
Gallo10
Gallo11
Gallo12
Gallo14
Gallo15
ATCC43144
NCTC13784
WUSP067
Gallo16
Gallo17
Gallo18
Gallo19
Gallo20
Gallo22
Gallo23
Gallo25
Gallo26
Gallo28
Gallo29
Gallo30
Gallo31
Gallo32
Gallo33
Gallo34
Gallo35
Gallo36
Gallo37
Gallo40
Gallo41
Gallo42
Gallo43
Gallo44
Gallo46
Gallo48
Gallo49
Gallo50
Gallo51
Gallo52
Gallo53
Gallo54
ATCC 43143
BAA2069
DSM16831
Gallo21
Gallo24
Gallo27
Gallo38
Gallo39
Gallo47
TX20005
Sgm
EsxA
EssA
EsaB
EssB
EssC
EsaA
No homology
Highly similar sequence
Homolog
Truncated
D
General T7SSb gene arrangement
esxA
essA
esaB
essB
essC
esaA
S. intermedius B196
S. aureus COL
1 kb

Figure 2

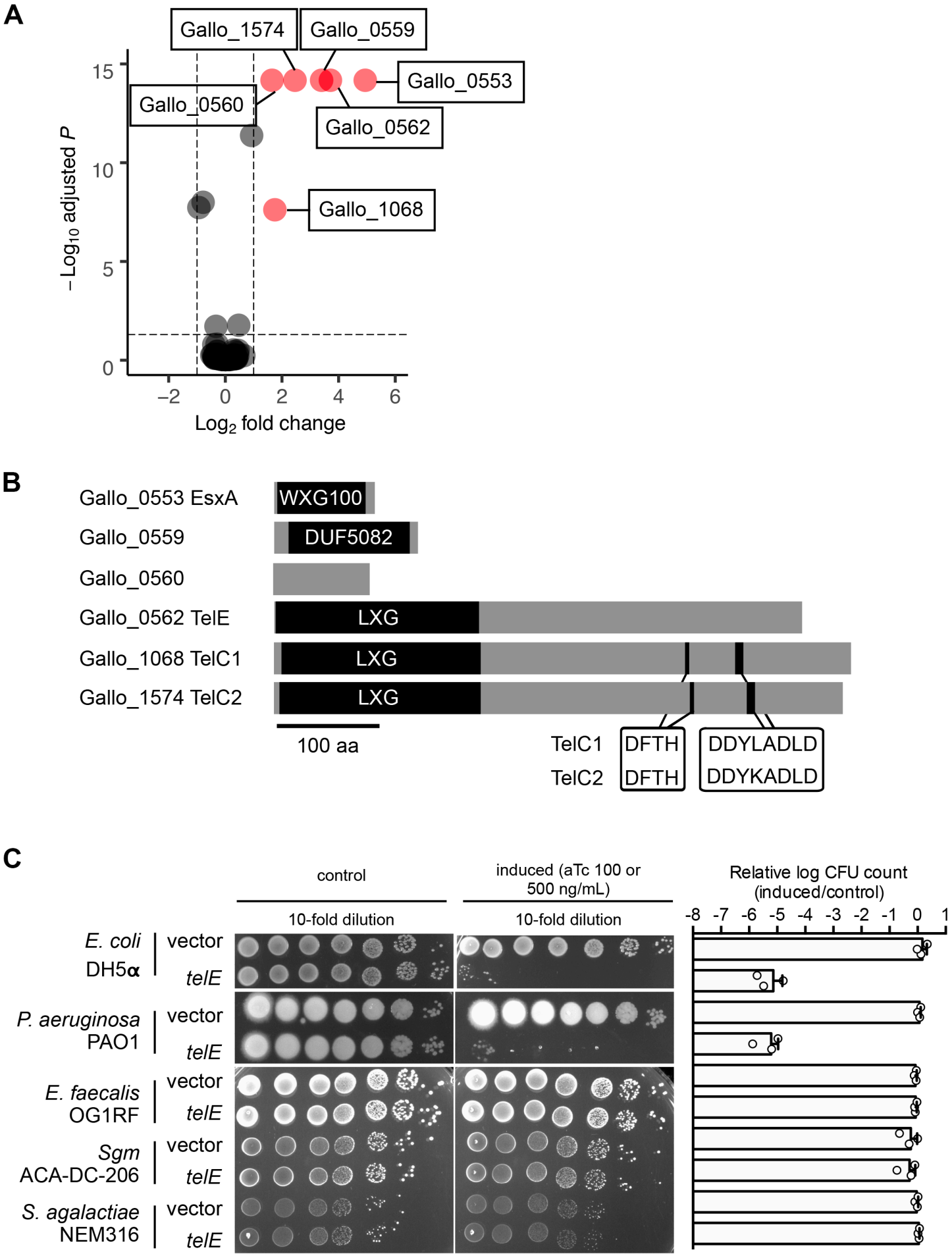

A
Gallo_1574
Gallo_0559
Gallo_0553
Gallo_0560
Gallo_0562
Gallo_1068
15
10
5
0
−Log10 adjusted P
−2
0
2
4
6
Log2 fold change
B
Gallo_0553 EsxA
WXG100
Gallo_0559
DUF5082
Gallo_0560
Gallo_0562 TelE
LXG
Gallo_1068 TelC1
LXG
Gallo_1574 TelC2
LXG
100 aa
TelC1
DFTH
DDYLADLD
TelC2
DFTH
DDYKADLD
C
control
induced (aTc 100 or 500 ng/mL)
10-fold dilution
10-fold dilution
Relative log CFU count (induced/control)
-8 -7 -6 -5 -4 -3 -2 -1 0 1
E. coli DH5α
vector
telE
P. aeruginosa PAO1
vector
telE
E. faecalis OG1RF
vector
telE
Sgm ACA-DC-206
vector
telE
S. agalactiae NEM316
vector
telE

Figure 3

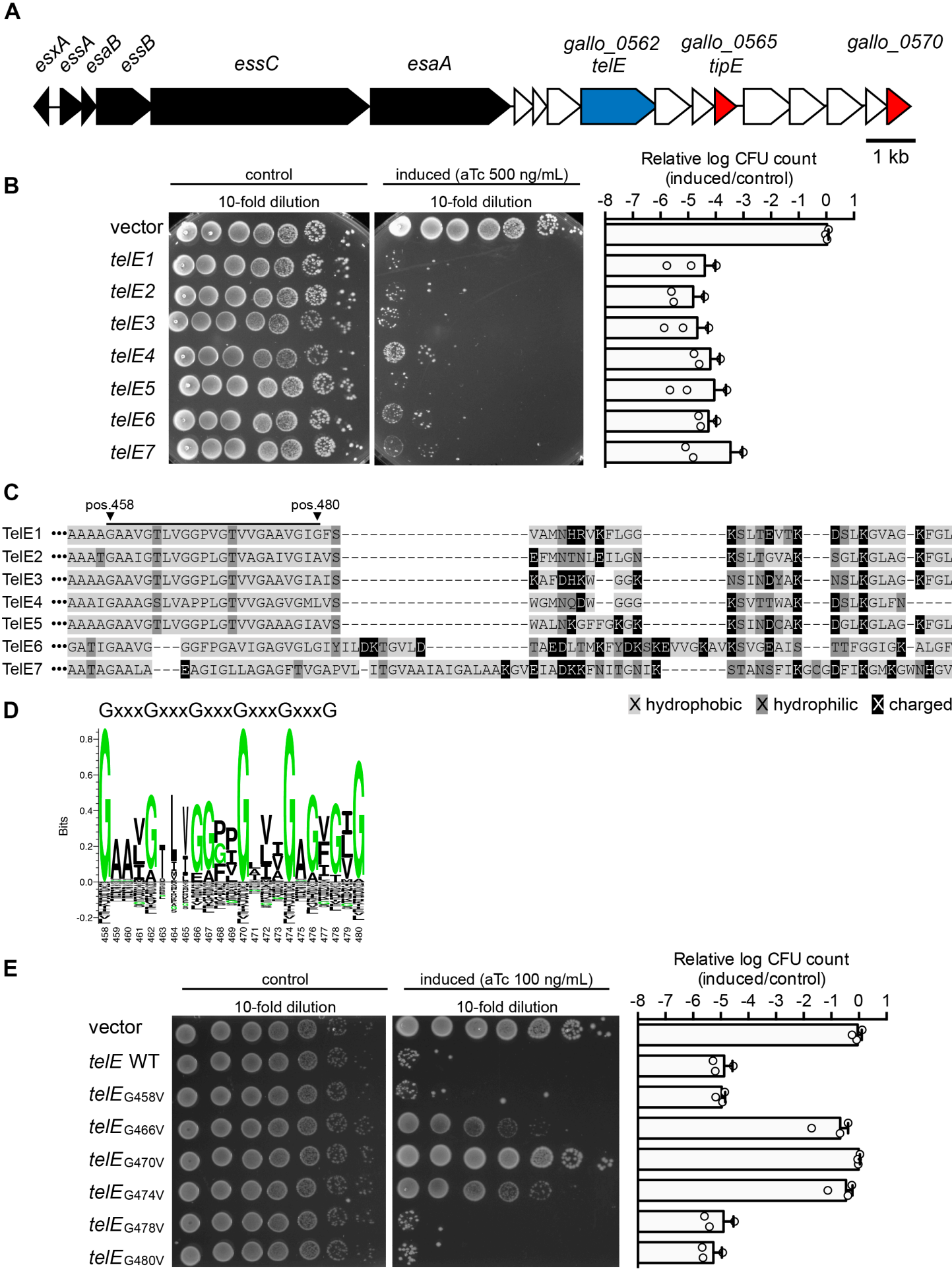

A
esxA
essA
esaB
essB
essC
esaA
gallo_0562
telE
gallo_0565
tipE
gallo_0570
1 kb
B
control
induced (aTc 500 ng/mL)
10-fold dilution
10-fold dilution
Relative log CFU count
(induced/control)
-8 -7 -6 -5 -4 -3 -2 -1 0 1
vector
telE1
telE2
telE3
telE4
telE5
telE6
telE7
C
pos.458
pos.480
TelE1
TelE2
TelE3
TelE4
TelE5
TelE6
TelE7
D
GxxxGxxxGxxxGxxxGxxxG
Bits
X hydrophobic
X hydrophilic
X charged
E
control
induced (aTc 100 ng/mL)
10-fold dilution
10-fold dilution
Relative log CFU count
(induced/control)
-8 -7 -6 -5 -4 -3 -2 -1 0 1
vector
telE WT
telE G458V
telE G466V
telE G470V
telE G474V
telE G478V
telE G480V

Figure 4

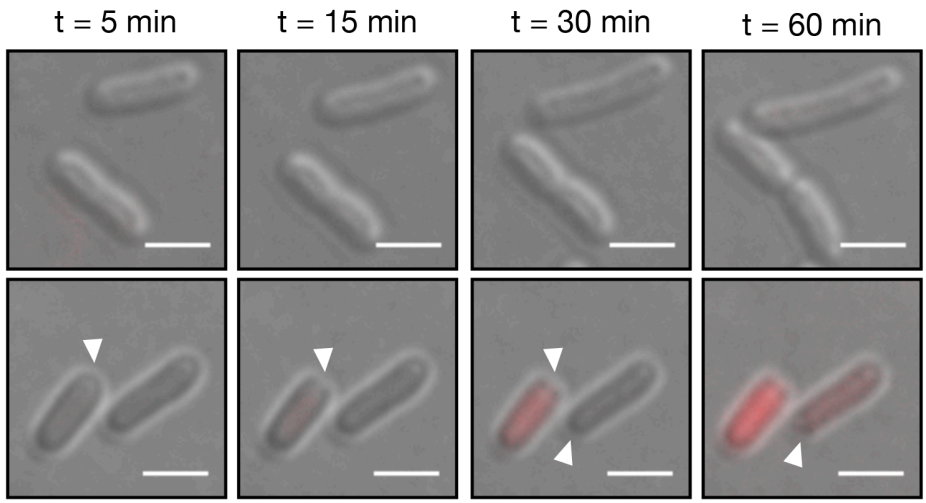
A

MKIKMSEVLSQKESLASSISSAKSQLESASRKLKSTADSDALQGNVKEAIDNKITNYQVPLLTNY
VNSLAVISQGYDNLISTFKETVSETSDSAIIDTDVLQQMVDKFDSPLEQLKTCSDEINKAIDEVAD
IVSLTKVTTADAISEFKSAKKVLTNTIKDMGIFNSTVFTSEAGDILEEQNTQLSSLSDLGSSSYTS
KKAKEFYKDSTFKAEVKEVKSVVNGKARRDQLKNALAKNLYDSKYSGYMLETSGLEVKAGILIT
STKKSIGKGKKVVKILDATDKALKQKKFIGGRELMVDSKGRVKVGERFLYKKPSKGKTKTVHLY
SKGTKEYKKTVGEDFSFKKTTLYNGRKLKANGKVGIDWKGVGTSGKEAFKSGLKDNFKSTINP
LNEFKGFKEASNFSKAGKVLGFAGTVMTIGSNIKTDFIDDRRSSTKEKVRNFAVDTTVDVVSSS
SAAAA**GAAVGTLVGGPVGTVVGAAVG**IGFSVAMNHRVKFLGGKSLTEVTKDSLKGVAGKFGL

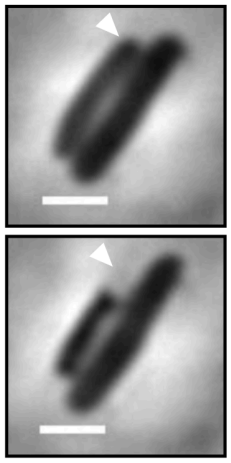
glycine zipper motif

X coil X sheet X helix

B



C



D

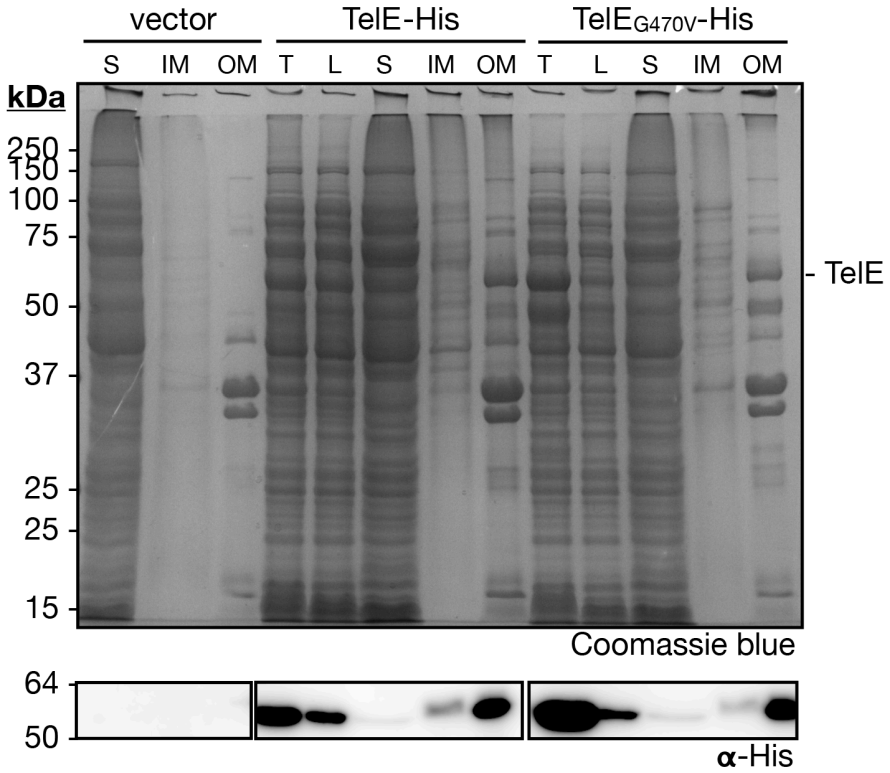


Figure 5

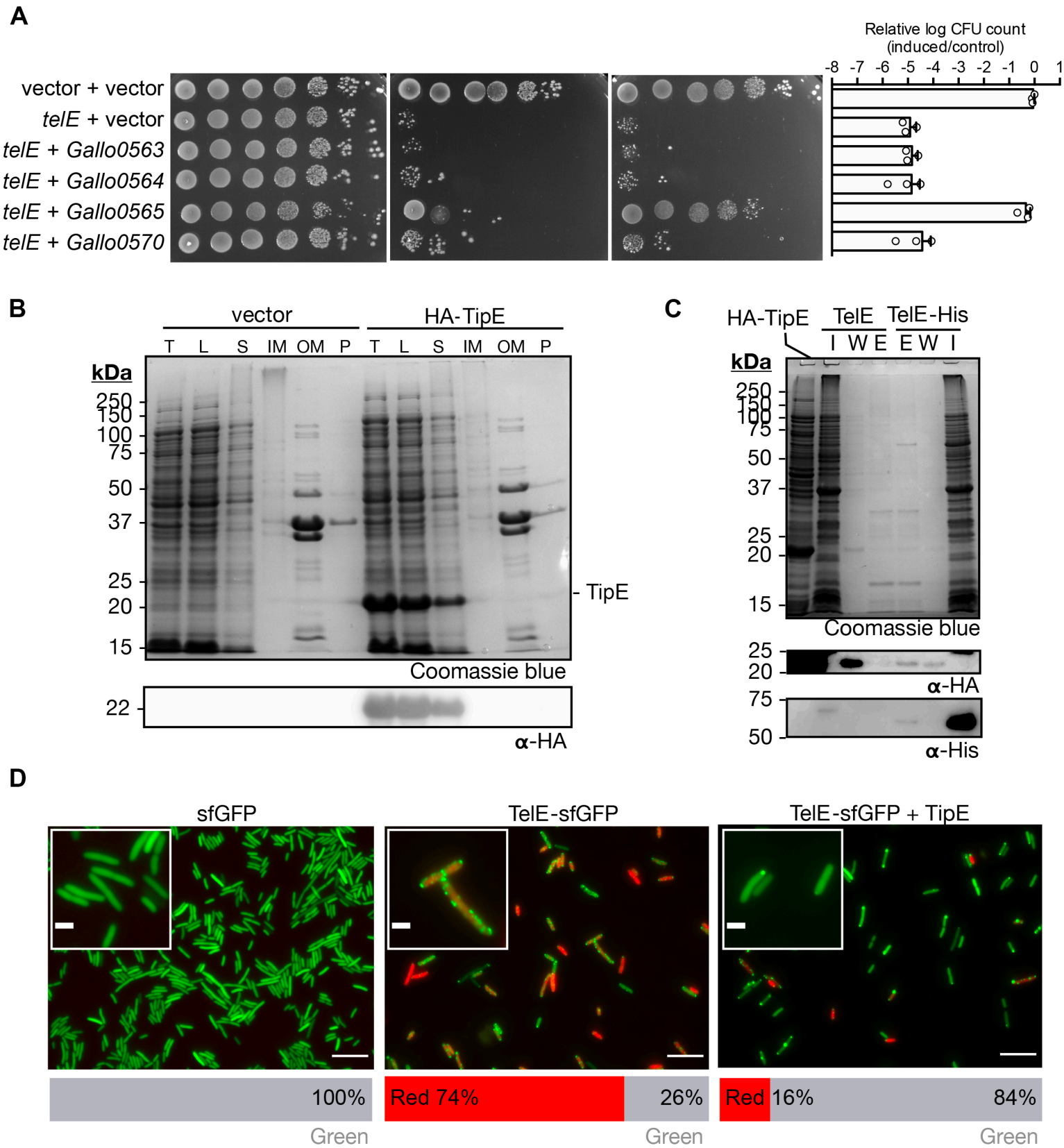
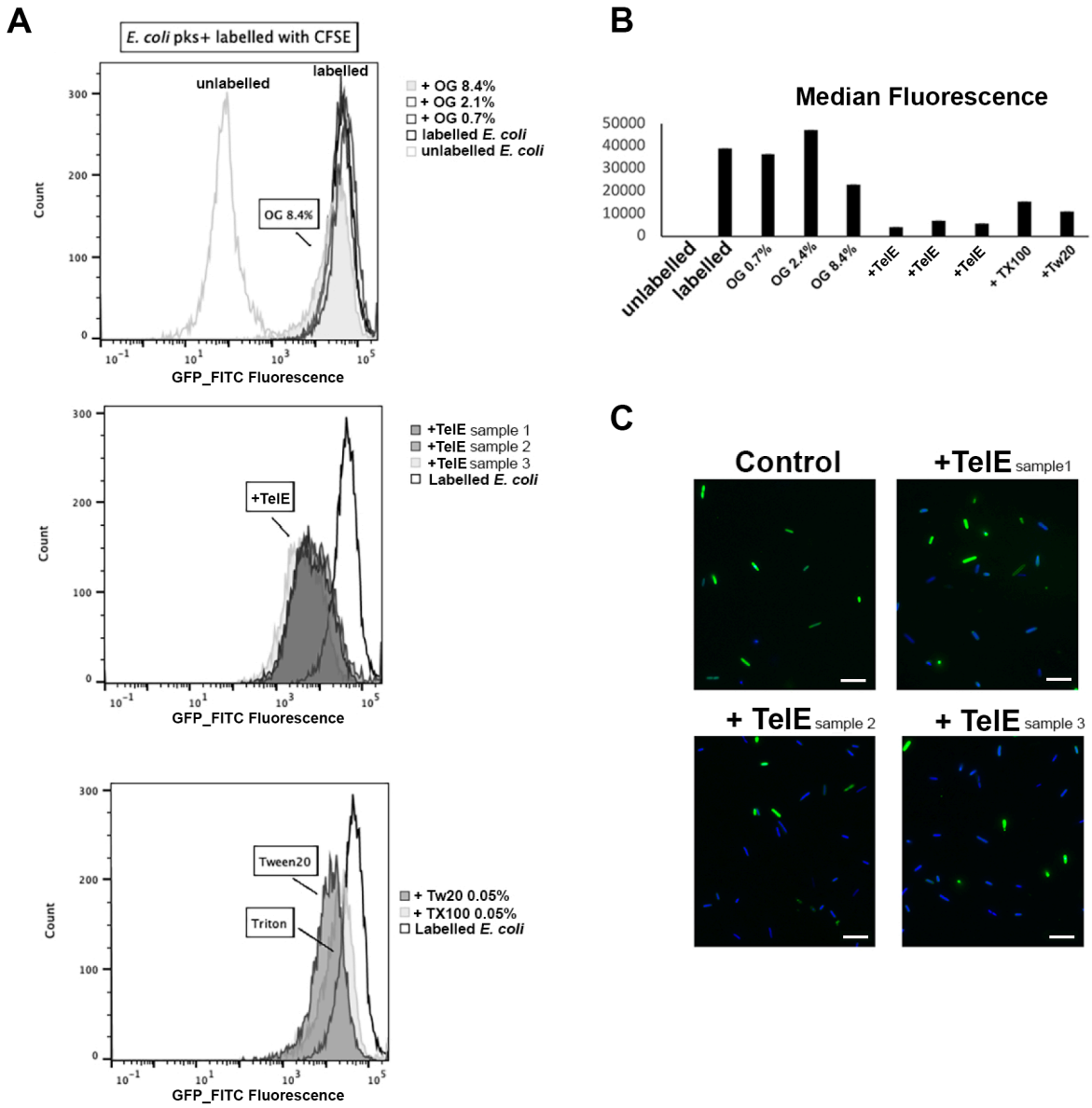

A
Relative log CFU count
(induced/control)
-8 -7 -6 -5 -4 -3 -2 -1 0 1
vector + vector
telE + vector
telE + Gallo0563
telE + Gallo0564
telE + Gallo0565
telE + Gallo0570
B
vector
HA-TipE
T L S IM OM P T L S IM OM P
kDa
250
150
100
75
50
37
25
20
15
- TipE
Coomassie blue
22
α-HA
C
HA-TipE
TelE
TelE-His
I W E E W I
kDa
250
150
100
75
50
37
25
20
15
Coomassie blue
25
20
α-HA
75
50
α-His
D
sfGFP
TelE-sfGFP
TelE-sfGFP + TipE
100%
Green
Red 74%
26%
Green
Red 16%
84%
Green

Figure 6


A
E. coli pks+ labelled with CFSE
unlabelled
labelled
OG 8.4%
+ OG 8.4%
+ OG 2.1%
+ OG 0.7%
labelled E. coli
unlabelled E. coli
Count
GFP_FITC Fluorescence
+TelE
+TelE sample 1
+TelE sample 2
+TelE sample 3
Labelled E. coli
Tween20
Triton
+ Tw20 0.05%
+ TX100 0.05%
Labelled E. coli
B
Median Fluorescence
unlabelled
labelled
OG 0.7%
OG 2.4%
OG 8.4%
+TelE
+TelE
+TelE
+ TX100
+Tw20
C
Control
+TelE sample1
+ TelE sample 2
+ TelE sample 3