bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

1 # DNA methylation epigenetically regulates gene expression in
2 # *Burkholderia cenocepacia*

3

4 Ian Vandenbussche[a], Andrea Sass[a], Marta Pinto-Carbó[b], Olga Mannweiler[b], Leo Eberl[b],
5 Tom Coenye[a]#

6

7 [a] Laboratory of Pharmaceutical Microbiology, Ghent University, Ghent, Belgium
8 [b] Department of Plant and Microbial Microbiology, University of Zurich, Zurich,
9 Switzerland

10

11 Running Head: Methylases regulate gene expression in *B. cenocepacia*

12

13 #Address correspondence to Tom Coenye, tom.coenye@UGent.be

14

15 Word Count Abstract: 232
16 Word Count Text: 4997
17
18
19
20
21
22

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.


23 **Abstract**

24 Respiratory tract infections by the opportunistic pathogen *Burkholderia cenocepacia*
25 often lead to severe lung damage in cystic fibrosis (CF) patients. New insights in how to
26 tackle these infections might emerge from the field of epigenetics, as DNA methylation
27 has shown to be an important regulator of gene expression. The present study focused
28 on two DNA methyltransferases (MTases) in *B. cenocepacia* strains J2315 and K56-2,
29 and their role in regulating gene expression. *In silico* predicted DNA MTase genes
30 BCAL3494 and BCAM0992 were deleted in both strains, and the phenotypes of the
31 resulting deletion mutants were studied: deletion mutant ΔBCAL3494 showed changes
32 in biofilm structure and cell aggregation, ΔBCAM0992 was less motile. *B. cenocepacia*
33 wild type cultures treated with sinefungin, a known DNA MTase inhibitor, exhibited the
34 same phenotype as DNA MTase deletion mutants. Single-Molecule Real-Time
35 sequencing was used to characterize the methylome of *B. cenocepacia*, including
36 methylation at the origin of replication, and motifs CACAG and GTWWAC were
37 identified as targets of BCAL3494 and BCAM0992, respectively. All genes with
38 methylated motifs in their putative promoter region were identified and qPCR
39 experiments showed an upregulation of several genes, including biofilm and motility
40 related genes, in MTase deletion mutants with unmethylated motifs, explaining the
41 observed phenotypes in these mutants. In summary, our data confirm that DNA
42 methylation plays an important role in regulating the expression of *B. cenocepacia*
43 genes involved in biofilm formation and motility.


44

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

45 ## Importance

46 CF patients diagnosed with *B. cenocepacia* infections often experience rapid
47 deterioration of lung function, known as *cepacia syndrome*. *B. cenocepacia* has a large
48 multi-replicon genome and a lot remains to be learned about regulation of gene
49 expression in this organism. From studies in other (model) organisms, it is known that
50 epigenetic changes through DNA methylation play an important role in this regulation.
51 The identification of *B. cenocepacia* genes of which the expression is regulated by DNA
52 methylation and identification of the regulatory systems involved in this methylation are
53 likely to lead to new insights in how to tackle *B. cenocepacia* infections in CF patients.
54

55 ## Introduction

56 *Burkholderia cenocepacia*, a member of the *Burkholderia cepacia* complex (Bcc),
57 is an aerobic Gram-negative bacterium that can be isolated from soil and water (1–3).
58 *B. cenocepacia* is also known as an opportunistic pathogen in immunocompromised
59 patients (4–6). Infection of the upper airways in cystic fibrosis (CF) patients often leads
60 to severe illness, typically referred to as *cepacia syndrome* (1,7). CF patients diagnosed
61 with *cepacia syndrome* experience a progressive decrease in lung function, often
62 accompanied by bacteremia and sepsis. If left untreated, *cepacia syndrome* can lead to
63 death within weeks (8,9). The genome of *B. cenocepacia* is complex (with usually three
64 large replicons), with a high GC-content (67%) and large size, comprising approximately
65 8.06 Mb (10). The species has been classified into different phylogenetic clusters and
66 subdivided into lineages, including the highly transmissible ET-12 lineage that harbors
67 *B. cenocepacia* strains J2315 and K56-2 (11,12).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

68 Epigenetics is the study of heritable changes in gene expression without changes
69 in the actual genomic sequence. In bacterial genomes, epigenetic control is exerted by
70 DNA methyltransferase enzymes (MTases) (13–15). DNA MTases originate from
71 restriction-modification (RM) systems, early defense mechanisms in bacteria with an
72 active interplay between endonucleases and DNA MTases, which cleave foreign DNA
73 but protect the own genome. In addition, discovery of orphan DNA MTases, enzymes
74 without a cognate endonuclease, shows that DNA MTases are not exclusively
75 dependent on the presence of the restriction part to function as regulator of gene
76 expression (16).

77 DNA MTases interact with specific DNA recognition sites and transfer a $CH_3$-
78 group from a methyl donor, mostly S-adenosyl methionine, to a cytosine ($C_5$-methyl
79 cytosine or $N_4$-methyl cytosine) or adenine ($N_6$-methyl adenine) base (17,18). As
80 methylated bases change the binding affinity of DNA binding proteins, methylation at
81 regulatory regions allows bacteria to regulate gene expression at the level of
82 transcription (19,20). While both cytosine and adenine methylation occur in eukaryotic
83 and prokaryotic cells, $C_5$-methyl cytosine is the archetypal eukaryotic base methylation
84 signature (16,21). Conversely, in prokaryotes, $N_6$-methyl adenine is the most important
85 base modification involved in gene expression regulation (22). In addition to this, studies
86 with DNA MTases *Dam* (deoxyadenosine methyltransferase) and *Dcm* (DNA cytosine
87 MTase) in *Escherichia coli* have demonstrated that, besides having a regulatory
88 function, DNA MTases also take part in crucial cellular processes like DNA replication
89 initiation or methyl-directed mismatch repair (21,23).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

90 Detection of (genome-wide) DNA methylation patterns has been challenging in
91 the past. The use of specific restriction enzymes with affinity for methylated sites,
92 followed by a comparison of the resulting fragment lengths gave a good impression of
93 methylation of the treated DNA at one particular area, but global methylation analysis
94 was until recently difficult at best (21,24). The rise of Next-Generation Sequencing and
95 Single-Molecule Real-Time (SMRT) technologies tremendously improved quality of
96 methylome analyses, but it also made it much more accessible (25,26). SMRT
97 Sequencing uses a sequencing-by-synthesis approach with fluorescently labeled
98 nucleotides. Pulse width, the signal of nucleotide incorporation, and interpulse duration,
99 the time between two incorporations, allow to discriminate between incorporated bases
100 and their methylation status (27).

101 The purpose of the present study is to understand how DNA methylation
102 regulates gene expression in *B. cenocepacia*. To this end, a genome-wide methylome
103 analysis was carried out, and genes under DNA methylation regulation were identified.
104 Interpretation of the working mechanisms of these regulatory systems, might lead to
105 new insights in how to tackle *B. cenocepacia* infections in CF patients.

106

107 ## Results

108 ### Identification of *B. cenocepacia* DNA MTases

109 All predicted DNA MTase genes in the *B. cenocepacia* J2315 genome were
110 identified using REBASE (Table S1). DNA MTase genes BCAL3494 and BCAM0992,
111 widely distributed within the genus *Burkholderia*, were selected for further analysis.
112 Gene BCAL3494, located on the first replicon of *B. cenocepacia*, is a type III

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

113 methyltransferase that is part of a RM-system, together with a restriction enzyme
114 encoded by the neighbouring gene BCAL3493. Gene BCAM0992 is located on the
115 second replicon and apparently does not have any adjacent genes coding for restriction
116 enzymes. The gene is classified as coding for a type II methyltransferase, i.e. the
117 restriction and modification enzymes act separately and are not dependent on each
118 other (28). To investigate the influence of BCAL3494 and BCAM0992 on bacterial
119 physiology, deletion mutants were constructed (Figure S1). For the other DNA MTase
120 genes in *B. cenocepacia* J2315 identified with REBASE (Table S1), homologues in
121 different *Burkholderia* strains could not be found; these genes were not further
122 investigated in the present study.

123 **Phenotype of mutant strains**

124 BCAL3494 and BCAM0992 were deleted in two *B. cenocepacia* strains, J2315
125 and K56-2, and the phenotype of the deletion mutants was investigated in detail. No
126 differences in growth between wild type and mutant strains were observed when
127 cultured in phosphate buffered mineral medium (Figure S2). Microscopic analysis
128 clearly showed a different, more clustered biofilm morphology for both BCAL3494
129 deletion mutants (ΔBCAL3494) compared to wild type strains, whereas the biofilm
130 structure of the BCAM0992 deletion mutants (ΔBCAM0992) did not differ from wild type
131 (Figure 1A). Cell aggregation in planktonic cultures was investigated using flow
132 cytometry (Figure 1B). The degree of aggregation in the BCAL3494 mutant strains was
133 significantly higher (p-value J2315: 0.049, p-value K56-2: 0.001) than in the
134 corresponding wild type strains. Also, the ability to form a pellicle, a biofilm-like structure
135 at the air-liquid interface, was investigated (Figure 1C). Pellicle formation was clearly

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

136 increased for both ΔBCAL3494 mutants compared to wild type strains and to
137 ΔBCAM0992 mutants. Complemented mutant strains *c*ΔBCAL3494 and *c*ΔBCAM0992
138 did not differ significantly from wild type in these experiments (Figure S3).
139 Motility of all strains was assessed using a swimming motility assay on 0.3 %
140 agar plates. After 24h (strain K56-2), and 32h (strain J2315), plates were photographed
141 and diameters measured (Figure 2). Diameters were significantly smaller for both
142 ΔBCAM0992 mutants compared to wild type (p-value J2315: 0.002, p-value K56-2 <
143 0.001). Both ΔBCAL3494 mutants, as well as the complemented mutants, were
144 identical to wild type in this regard. We also investigated swarming motility, but no
145 significant differences between the different strains were observed (Figure S4).

146 **Effect of the DNA MTase inhibitor sinefungin on methylation-dependent**
147 **phenotypes**

148 Sinefungin, a structural analog of S-adenosyl methionine and known for blocking
149 base methylation in other bacteria such as *Streptococcus pneumoniae* (29), was used
150 as DNA MTase inhibitor. The minimum inhibitory concentration (MIC) of sinefungin in
151 *B. cenocepacia* J2315 and K56-2 was determined and was found to be higher than 200
152 µg/mL. Both strains were exposed to sinefungin concentrations below the MIC of
153 sinefungin (50 µg/mL) to assure any effect observed was not due to growth inhibition by
154 sinefungin, and the effect on biofilm formation, pellicle formation, cell aggregation and
155 motility was quantified (Figure 3). Bacteria exposed to sinefungin produced more pellicle
156 mass, showed a higher degree of cell aggregation (p-value: 0.003), had a different
157 biofilm morphology, and were less motile (p-value: 0.004). These findings indicate that

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

158 chemically blocking DNA methylation or deleting genes responsible for DNA methylation
159 lead to the same phenotypes in *B. cenocepacia* J2315 and K56-2.

160 **Methylome analysis**

161 Using SMRT sequencing (PacBio), the complete methylome of *B. cenocepacia*
162 J2315 and K56-2 was identified. Only data for strain J2315 are reported in the following
163 section, as data for strain K56-2 were highly comparable (Figures 4 and 5). Three
164 distinct methylated motifs were identified in the wild type strain: CACAG, GTWWAC,
165 and GCGGCCGC. The CACAG motif was methylated at the fourth position on the
166 forward strand, whereas the GTWWAC motif was methylated at the fifth position on
167 both the forward and reverse strand. Cytosine methylation of the GCGGCCGC motif
168 occurred at the fifth position on both the forward and reverse strand. Although all
169 CACAG and GTWWAC motifs were methylated in the wild type strains, methylation of
170 the CACAG motif was absent in the ΔBCAL3494 deletion mutants, and likewise, no
171 methylation of the GTWWAC motif was seen in the ΔBCAM0992 mutants (Table S2).
172 This demonstrates that MTase BCAL3494 recognizes the CACAG motif, while MTase
173 BCAM0992 recognizes the GTWWAC motif. In contrast, cytosine methylation of the
174 GCGGCCGC motif was observed in wild type and in mutant strains, but the extent of
175 methylation at this motif was highly variable in the four datasets. This suggests that this
176 cytosine methylation occurs randomly, and that the GCGGCCGC motif might present a
177 false positive result of motif analysis due to the high occurrence of short repeats of G
178 and C in the GC-rich *B. cenocepacia* genome. In addition, almost no methylated
179 GCGGCCGC motifs were found in regulatory regions, hinting at only a minor role for

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

180 cytosine methylation in regulation of gene expression. Therefore, cytosine methylation
181 in *B. cenocepacia* was not further studied.
182 The location of every methylated CACAG and GTWWAC motif was mapped
183 (Figures 4 and 5). A total of 6834 methylated CACAG motifs and 961 methylated
184 GTWWAC motifs was found, of which the majority was present on the first replicon
185 (CACAG: 45.6 %, GTWWAC: 49.9 %), followed by the second replicon (CACAG: 42.1
186 %, GTWWAC: 38.9 %), the third replicon (CACAG: 10.6 %, GTWWAC: 9.0 %), and the
187 plasmid (CACAG: 1.7 %, GTWWAC: 2.2 %). Subsequently, all genes with methylated
188 motifs in their promoter region, here defined as 60 bases upstream of the transcription
189 start site, were identified. 91 promoter regions contained methylated CACAG motifs,
190 and 80 promoter regions contained methylated GTWWAC motifs, with most of the
191 motifs being present on the first replicon (Figures 4 and 5). Functional classes of genes
192 found in the dataset of genes with methylated promoter include genes involved in
193 intermediary metabolism, regulation, and transport (Tables S3 and S4).
194 Virtual Footprint was used to elucidate to which transcription factor (TF) binding
195 sites the discovered methylation motifs CACAG and GTWWAC showed any similarity.
196 Data output of the analysis is listed in Table S5. Sequences that contain methylation
197 motif CACAG were similar to the binding site of *E. coli* K12 GlpR, while GTWWAC-
198 containing sequences were similar to binding sites of several other *E. coli* K12 TFs,
199 including ArcA, OxyR, Fis and Fur.

200 **Expression of genes with a methylated promoter**

201 The expression level of genes with methylated promoter regions was determined
202 in wild type and mutant strains, using qPCR. Expression data of genes with methylated

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

203 promoter regions are listed in Tables 1 and 2. Volcano plots (Figure S5, fold changes
204 plotted against corresponding p-values) show that most genes tested were upregulated
205 in the mutants compared to the wild type strains. Six of these genes were significantly
206 upregulated in mutants of both strain backgrounds: BCAL1515, BCAL2465, and
207 BCAM0820 were upregulated in ΔBCAL3494, whereas genes BCAL0079, BCAL2415,
208 and BCAM1362 were upregulated in ΔBCAM0992. Four additional genes were
209 upregulated in K56-2 mutants only: BCAL0423, BCAM2738, and BCAS0223 were
210 upregulated in ΔBCAL3494, BCAL1556 in ΔBCAM0992. Subsequently, the methylated
211 promoter regions of these genes were analyzed in detail (Figure 6). In most cases, the
212 methylated motif was in close proximity of the -10 or -30/35 element in bacterial
213 promoter regions.

214 To confirm that the presence of methylation close to the -10 or -30/35 element
215 influences transcription and therefore gene expression in *B. cenocepacia*, translational
216 eGFP reporter fusions were constructed and eGFP production was quantified. The
217 eGFP production in strains harboring different plasmids is shown in Figure 7. As
218 expected, the production of eGFP, driven by the promoters of genes BCAL1515,
219 BCAM0820, and BCAL0079, was significantly ($p = 0.001$, $p = 0.014$, $p = 0.002$,
220 respectively) increased in the deletion mutant for which an upregulation of these genes
221 was observed using qPCR experiments (Figure 7).

222 **DNA methylation in the origin of replication**

223 DNA methylation was detected in all origins of replication of *B. cenocepacia*
224 (Figure 8). Similar methylation patterns were observed in the origins of the different
225 replicons. A previously discovered 7-mer (CTGTGCA) that can be found in all

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

226 replication origins (30), contains a CACAG methylation motif on the antisense strand.
227 This motif was also found at the 3'-end of almost every DnaA box. These boxes are
228 bound by DnaA proteins, essential for DNA unwinding and chromosome replication
229 initiation (31). Also, the GTWWAC motif was found in proximity of the replication origins,
230 consequently the origins in *B. cenocepacia* represent methylation-rich regions. Whereas
231 methylated CACAG motifs were found throughout the origins of replication, the position
232 of the GTWWAC methylation was unique in all replicons and at least two GTWWAC
233 motifs were found in between two CACAG methylated DnaA boxes. In contrast to the
234 origins of the three larger replicons, the origin of replication of the plasmid contained
235 only one CACAG methylated DnaA box.

## 236 Discussion

237 Despite the growing knowledge of DNA methylation in prokaryotes (15), the role
238 of DNA MTases in regulating gene expression in *B. cenocepacia* remains to be
239 revealed. In the present study, we identified two DNA MTases (BCAL3494 and
240 BCAM0992), and mutants in which these genes were deleted, showed differences in
241 biofilm formation and motility. In addition, when methylation was blocked by the DNA
242 MTase inhibitor sinefungin (32), the same phenotypic differences were observed. These
243 findings demonstrate that epigenetic control of gene expression by MTases play an
244 important role in controlling certain phenotypes. Similar results have been reported in
245 *Salmonella enterica*, where DNA methylation is crucial for optimal pellicle and biofilm
246 production (33).

247 Methylome analysis showed that mutants in which MTase ΔBCAL3494 or
248 ΔBCAM0992 were inactivated, lacked adenine methylation in specific motifs. MTase

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

249 BCAL3494 was specifically linked to methylation of the CACAG motif, MTase
250 BCAM0992 to methylation of the GTWWAC motif. This strategy of DNA methylation
251 analysis, in which the methylome of strains lacking MTases is determined, has been
252 used in various bacteria, as it is an effective way to find associations between predicted
253 MTases and genome-wide methylation motifs (34,35). For example, several methylation
254 motifs were identified in *Burkholderia pseudomallei*, including motifs CACAG and
255 GTWWAC (36). Two of the *B. pseudomallei* MTases (M.BpsI and M.BpsII) are
256 homologous to the *B. cenocepacia* MTases BCAL3494 and BCAM0992. In *Ralstonia*
257 *solanacearum*, an important plant pathogen that is phylogenetically related to *B.*
258 *cenocepacia*, the GTWWAC methylation motif co-occurs with the respective homolog of
259 the BCAM0992 MTase, whereas a BCAL3494 MTase homolog and methylation of
260 CACAG are absent (37). As in *B. cenocepacia*, the BCAM0992 homolog in *R.*
261 *solanacearum* is an orphan DNA MTase. Analysis of cytosine methylation suggests that
262 cytosine is more likely to be methylated at random instead of at specific motifs, and is
263 likely not having a major regulatory function. Also, GC-rich genomes complicate the
264 search for specific cytosine motifs.

265 Previous epigenetic research demonstrated that there is a negative correlation
266 between methylation in promoters and transcription (38). To uncover the role of DNA
267 methylation in regulation of *B. cenocepacia* gene expression, all methylated motifs in
268 promoter regions were identified. The data obtained in the present study indicates that
269 gene expression was upregulated in DNA MTase mutants, suggesting that adenine
270 DNA methylation in *B. cenocepacia* affects gene expression by a mechanism inhibiting
271 transcription. In both prokaryotes and eukaryotes, adenine and cytosine methylation are

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

272 involved in blocking (or enhancing) the binding of RNA polymerase to DNA (15,21,39),
273 and especially methylation near the -10 and -30/35 elements in the promoter region
274 seems to be important for affecting RNA polymerase binding (40). We found that, also
275 in *B. cenocepacia*, methylated motifs (CACAG and GTWWAC) are found close to, or in
276 these elements.

277 BCAM0820, upregulated in the J2315 and K56-2 ΔBCAL3494 mutant, is a two-
278 component response regulator, the first gene of an operon homologous to the Wsp
279 chemosensory system involved in biofilm formation in *Pseudomonas aeruginosa* (41).
280 BCAM0820 is homologous to WspR, but lacks the diguanylate cyclase domain. During
281 an experimental evolution study in which *B. cenocepacia* HI2424 biofilms were grown
282 on beads, mutations within the *wsp* gene cluster occurred in different clones; these
283 were associated with increased pellicle formation and increased biofilm formation on
284 beads. This demonstrates that the Wsp cluster is involved in pellicle formation in
285 *B. cenocepacia* (42,43), and the upregulation of BCAM0820 could explain the
286 differences in pellicle and biofilm formation between the wild type strains and the
287 ΔBCAL3494 deletion mutants observed in the present study. Interestingly, BCAL1515,
288 encoding 2-oxoglutarate dehydrogenase (SucA) and upregulated in ΔBCAL3494, also
289 acquired mutations in the course of the experimental evolution study (43), but the role of
290 this gene in biofilm formation has not been further explored. BCAL0079, upregulated in
291 the ΔBCAM0992 mutants, is annotated as a DNA helicase gene (*rep*). Besides
292 unwinding DNA during DNA replication, Rep plays a role in swimming motility in *E. coli*
293 (44). The reduced motility observed in the ΔBCAM0992 mutants suggests that Rep may
294 also affect motility in *B. cenocepacia*, although this remains to be confirmed.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

295 Measurement of eGFP production in translational fusion mutants revealed that
296 mutants with constructs containing the BCAL1515, BCAM0820, or BCAL0079 promoter,
297 showed a significant increase in eGFP production compared to wild type, thereby
298 supporting our hypothesis of gene expression regulation by DNA methylation. *In silico*
299 analyses predict that sequences containing methylation motifs are similar to binding
300 sites of TF in *E. coli* K12, and it is plausible that these sequences are also part of TF
301 binding sites in *B. cenocepacia*, allowing us to propose a possible mechanism of gene
302 expression regulation (Figure 9). TFs that bind close to the -10 and -35 region often act
303 as transcriptional repressors (45). Therefore, a methylated promoter region could
304 promote binding of a repressor (46), and sterically hinder RNA polymerase (OFF state),
305 whereas an absence of methylation would allow binding of the initiation factor sigma to
306 the promoter, which in turn could lead to binding of RNA polymerase and initiation of
307 transcription (ON state).

308 The role of DNA methylation in prokaryotes in multifaceted. Besides gene
309 expression regulation and a role in DNA mismatch repair in Gram-positive bacteria (47),
310 DNA methylation has also been implicated in the coordination of replication initiation.
311 Results of the present study seem to confirm this, as the *rep* gene, necessary for
312 replication, was found to be under epigenetic control by DNA methylation. In *E. coli*,
313 GATC motifs, omnipresent in the replication origin, are prone to adenine methylation.
314 The motifs are found within DnaA boxes, essential for binding of the DnaA protein and
315 initiation of replication. The methylation state of each of these GATC motifs changes the
316 affinity of DnaA and sequestering-protein SeqA for the DnaA box. Immediately after
317 replication, GATC motifs are hemi-methylated, which leads to sequestration of the DnaA

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

318 boxes by SeqA and prevents the process of replication to be reinitiated (21). The
319 occurrence of methylated motifs in the vicinity of the origins of replication of the four
320 replicons in *B. cenocepacia* was studied to check for a link between DNA methylation
321 and coordination of the replication process. An enrichment of the CACAG motif was
322 observed in the origin of replication of all replicons. The motif was found to be part of a
323 bigger sequence that has previously been reported as a recurring 7-mer (30), without
324 known function. In addition, the origin of replication of the different replicons showed
325 high similarities in methylation patterns, raising the possibility of replication coordination
326 by DNA methylation.

327 In conclusion, we have demonstrated that DNA methylation plays a role in
328 regulation of gene expression in *B. cenocepacia*. DNA MTases BCAL3494 and
329 BCAM0992 are essential for methylation of the *B. cenocepacia* genome, and are
330 responsible for methylation of base motifs CACAG and GTWWAC, respectively. In
331 absence of methylation, expression of certain genes in affected and this results in
332 altered phenotypes (including cell clustering, biofilm formation, and motility). Finally,
333 recurrent methylation patterns were detected in all origins of replication, which suggests
334 an additional role of DNA methylation in replication regulation.

## 335 Materials and Methods

### 336 Strains and culture conditions

337 All strains and plasmids used in this study are listed in Table 3. *B. cenocepacia*
338 strains were cultivated in phosphate buffered mineral medium (2.00 g/L $NH_4Cl$, 4.25 g/L
339 $K_2HPO_4.3H_2O$ (ChemLab), 1.00 g/L $NaH_2PO_4.H_2O$, 0.10 g/L nitriloacetic acid, 0.0030
340 g/L $MnSO_4.H_2O$, 0.0030 g/L $ZnSO_4.7H_2O$, 0.0010 g/L $CoSO_4.7H_2O$, 0.20 g/L

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

341 $MgSO_4.7H_2O$, 0.012 g/L $FeSO_4.7H_2O$ (Sigma-Aldrich), 5 g/L Yeast Extract (Lab M), 2
342 g/L Casamino Acids (BD Biosciences), and 5 g/L glycerol (Scharlab)). LB medium (Luria
343 Bertani medium with 5g/L NaCl, Sigma-Aldrich) was used for maintenance of *E. coli*
344 strains and during specific stages of the gene deletion procedure (see below) where
345 antibiotic selection with tetracycline (250 µg/mL) (Sigma-Aldrich) was desired. Prior to
346 phenotypic experiments, liquid overnight (ON) cultures were grown in a shaker
347 incubator (100 rpm) at 37 °C.

348 **Selection of DNA MTase genes – *in silico***

349 The REBASE Genome database was used to allocate all known DNA MTase
350 genes in the *B. cenocepacia* J2315 and K56-2 genomes (48). The Artemis Genome
351 Browser and Annotation Tool (Sanger) allowed to visualize the genomic context of
352 these genes (49). NCBI BLAST was used to screen for conservation of the genes within
353 the *Burkholderia* genus using default search parameters (50) (search mode: BLASTn, E
354 cut-off value: < 1E-5).

355 **Construction of deletion mutants**

356 All primers used for construction and complementation of the deletion mutants
357 are listed in Table S6. The procedure is an adapted allelic replacement approach, using
358 a suicide plasmid with a SceI endonuclease recognition site (51,52). The suicide
359 plasmid, containing DNA fragments of regions flanking the target gene, is integrated into
360 the *B. cenocepacia* genome by homologous recombination. Introducing a second
361 plasmid that carries SceI endonuclease genes into *B. cenocepacia*, results in a lethal
362 genomic strand break. Another homologous recombination event allows the bacteria to
363 repair the break with a 50 % chance of resulting in a gene deletion. Deletion mutants

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

364 ΔBCAL3494 and ΔBCAM0992 were constructed in both *B. cenocepacia* J2315 and
365 K56-2.

366 BCAL3494 was deleted together with neighboring gene BCAL3493, as well as
367 BCAL3488 to BCAL3492 (encoding hypothetical proteins). Targeting BCAL3494 alone
368 was not feasible because regions flanking BCAL3494 contain multiple recognition sites
369 for endonucleases used during construction of the deletion mutants, and digestion of
370 these regions would be inevitable (Figure S1).

371 *E. coli* One Shot PIR2 cells (Thermo Fisher), expressing λ *pir*, were used for
372 transformation, replication, and maintenance of the suicide plasmid during construction
373 of deletion mutants. Thawed cells were immediately exposed to a heat shock
374 transformation procedure, after which they were transferred to SOC medium for
375 recovery. For plasmid selection, the phosphate buffered mineral medium was
376 supplemented with one or more of following antibiotics: trimethoprim (Ludeco; 50 µg/mL
377 for initial screening in *E. coli*, 200 µg/mL when plasmid is introduced in *B. cenocepacia*),
378 chloramphenicol (400 µg/mL), gentamicin (50 µg/mL), kanamycin (50 µg/mL), and
379 ampicillin (200 µg/mL) (Sigma-Aldrich).

380 **Construction of plasmids for complementation**

381 To ensure that phenotypes were solely caused by the deletion of DNA MTases,
382 deletion mutants were complemented. The primers used for construction of plasmids
383 used for complementation are listed in Table S6. Plasmids pJH2 and pSCrhaB2 were
384 used for complementation of ΔBCAL3494 (*c*ΔBCAL3494) and ΔBCAM0992
385 (*c*ΔBCAM0992), respectively. The genomic sequences of the DNA MTase genes were
386 PCR-amplified and subsequently cloned into the plasmids. BCAL3494 was amplified

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

387 including its own regulatory region (approx. 250 nucleotides upstream of the
388 transcription start site) into pJH2, which does not have a promoter associated with its
389 multiple cloning site (53). BCAM0992 does not have its own upstream promoter,
390 therefore it was cloned into pSCrhaB2, which contains a rhamnose-inducible promoter
391 (54). Complemented mutant strains were subjected to the same phenotypic tests as the
392 deletion mutants and wild type *B. cenocepacia*. For strains *c*ΔBCAM0992, the
393 phosphate buffered mineral medium was supplemented with 0.2 % rhamnose.

394 **Biofilm and clustering experiments**

395 Biofilms were grown in plastic U-shaped 96-well microtiter plates in phosphate
396 buffered medium at 37 °C, starting from 200 µL/well planktonic overnight cultures with
397 an optical density (OD) of 0.05 (590 nm). After 4 h static incubation, all wells were
398 rinsed with physiological saline (PS, 0.9 % NaCl in water), thereby removing all
399 unattached planktonic cells. Wells were re-filled with 200 µL medium and incubated for
400 an additional 20 h. Where appropriate, biofilms were stained with LIVE/DEAD
401 (SYTO9/propidium iodide, Invitrogen) to visualize the bacteria and distinguish live and
402 dead cells (55). Pellicle formation was determined in glass tubes. Cultures were grown
403 statically for 24 h, after which adhering pellicles were stained and quantified with crystal
404 violet (56). Cell clustering, already shown to be correlated with pellicle formation, was
405 determined with flow cytometry (Attune NxT Flow Cytometer, Thermo Fisher) (57).
406 Forward scatter (FCS), a value for particle size, and side scatter (SSC), a value for
407 particle complexity, were measured for each particle present in the bacterial suspension
408 and visualized in scatter plots. After analysis of these graphs, the main cell population

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

409 was gated (gate ranging from approx. $10^3$ to $10^5$ for both FSC and SSC), and detected
410 events larger and more complex than the gate, were considered clustered (Figure S6).

411 **Motility experiments**

412 Petri dishes containing phosphate buffered mineral medium with agar
413 concentrations of 0.3 % and 0.5 % were used for assessment of swimming and
414 swarming motility, respectively. 1 µL of cultures with OD 0.1 was spotted on the agar
415 plates. Diameters were measured after 24 h (strain K56-2) or 32 h (strain J2315).

416 **DNA MTase inhibition with sinefungin**

417 A stock solution of the DNA MTase inhibitor sinefungin (Sigma-Aldrich) was
418 prepared (10 mg/mL) (29), aliquoted, and immediately frozen at -20 °C to prevent
419 degradation. Cells were grown for 24 h in sinefungin-supplemented medium (50 µg/mL)
420 and used as inoculum for an overnight culture, also in sinefungin supplemented
421 medium. This allowed the DNA MTase inhibitor to have an effect during several growth
422 cycles. Then, biofilm formation and motility of sinefungin-treated cells was assessed as
423 described above in medium supplemented with 50 µg/mL sinefungin.

424 **Genomic DNA extraction**

425 Prior to DNA extraction, planktonic strains were grown overnight in a shaker
426 incubator (100 rpm) at 37 °C. Biofilm cells were grown as described above. Next, cells
427 were harvested and genomic DNA (gDNA) was extracted using the Wizard Genomic
428 DNA Purification Kit (Promega). Quantification was performed with a BioDrop µLITE
429 (BioDrop) spectrophotometer.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

430 **SMRT sequencing**

431 To determine the methylome of *B. cenocepacia*, gDNA extracts were analyzed
432 with Single Molecule Real-Time (SMRT) Sequencing technology. gDNA samples of
433 both wild type and mutant strains were run on a Pacific Biosciences Sequel System
434 (250x coverage) according to the manufacturer's guidelines. Library preparations were
435 multiplexed as data output of approximately 2 Gb per genome was expected, and a
436 single SMRT Sequel cell provides up to 6 Gb data. Initial data output was processed
437 with SMRT Link software (Pacific Biosciences). Identification of the modified bases and
438 analysis of the methylated motifs was performed with the Base Modification and Motif
439 Analysis application (SMRT Link v6.0, Pacific Biosciences). In depth data analysis was
440 performed with CLC Workbench Genomics (v11.0.1, Qiagen). Differential analysis
441 between wild type and mutants was performed to identify methylation motifs specifically
442 associated with certain DNA MTases. Previously predicted promoter regions and
443 transcription start sites of *B. cenocepacia* were used to determine the methylation
444 profile of regulatory regions (58). Virtual Footprint software (promoter analysis mode,
445 default search parameters) was used to assess similarity of the methylation motifs to
446 known TF binding sites (59).

447 **qPCR**

448 To evaluate the impact of DNA methylation in promoter regions on gene
449 expression, qPCR was performed on all genes that had a methylated promoter region in
450 wild type *B. cenocepacia*, but an absence of methylation in promoter region in one of
451 the deletion mutants. All hypothetical genes and genes with unknown function, as well
452 as genes with low innate expression level, were excluded from testing. The primers

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

453 used for qPCR are listed in Table S7. First, all strains were grown to an OD of 0.6 in
454 phosphate buffered medium, after which they were pelleted by centrifugation and frozen
455 at -80 °C. Next, RNA was extracted using the RiboPure – Bacteria extraction kit
456 (Invitrogen), followed by a DNase treatment to remove trace quantities of gDNA.
457 Quantification and measurement of RNA purity of the extracts was performed with a
458 BioDrop µLITE (BioDrop). Subsequently, cDNA was synthesized, using 500 ng RNA per
459 reaction, with a Reverse Transcriptase kit (High-Capacity cDNA RT Kit, Applied
460 Biosystems). Per qPCR reaction, 2 µL template cDNA was mixed with 10 µL GoTaq
461 qPCR Master Mix, 0.6 µL qPCR primer mix (10 µg/mL), and 7.4 µL nuclease-free water
462 according to the GoTaq qPCR Master Mix (Promega) protocol. Samples were run on a
463 CFX96 Real-Time System C1000 Thermal Cycler (Bio-Rad) and output data was
464 processed with Bio-Rad CFX Manager 3.1 software. The baseline threshold was set to
465 a defined 100 RFU. Obtained Cq values were normalized to reference gene *rpo*D
466 (BCAM0918) of which the expression was stable across all samples, differences to wild
467 type were calculated (ΔΔCq) and log-transformed. Volcano plots were used to plot the
468 negative logarithm of statistical p-values against log 2-fold changes (Figure S5).

469 **Construction of translational eGFP reporter fusions and measurement of eGFP**
470 **production**

471 Genes with methylated promoter regions that showed a significant upregulation
472 of gene expression in one of both mutant strains, were selected for eGFP experiments.
473 Translational eGFP reporter fusion plasmids were constructed by cloning the regulatory
474 regions of the genes, comprising 60 to 390 nucleotides upstream of the transcription
475 start site, into vector pJH2. The insert is cloned right in front of the eGFP gene and

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

476 contains an ATG start codon at the 3'-end, in frame with the codon sequence of the
477 gene. All primers used for amplification of the regulatory regions and screening of pJH2
478 with correct insert length, are listed in Table S8. The plasmids were transferred to *B.*
479 *cenocepacia* J2315 and K56-2 by triparental mating. Exconjugants were grown on
480 selective plates (LB medium supplemented with 200 µg/mL chloramphenicol and 50
481 µg/mL gentamicin) and PCR-screened to confirm the presence of the insert. Constructs
482 carrying genes BCAL2415, BCAL2465 and BCAM1362 repeatedly failed to be
483 transformed into *B. cenocepacia* and were not included in further experiments.
484 Fluorescent signals of eGFP production in wild type and mutant strains were measured
485 by flow cytometry (Attune NxT Flow Cytometer, Thermo Fisher) (53).

486 **Data analysis and statistics**

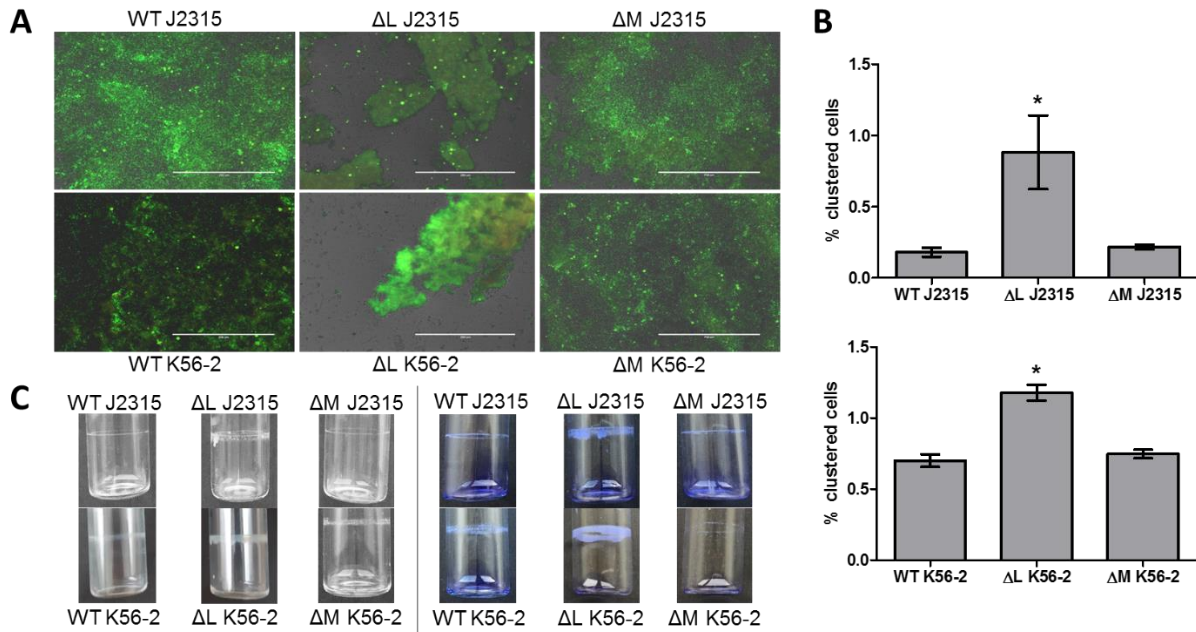
487 Statistical analysis was performed using SPSS Statistics v. 25 software. All tests
488 and experiments were run in triplicate unless otherwise mentioned. Normality of data
489 was verified with a Shapiro-Wilk test. To check for significant differences between data,
490 normally distributed data was subjected to a T-test or One-way ANOVA test, not
491 normally distributed data to a non-parametric Mann-Whitney U-test. Resulting p-values
492 smaller than 0.05 were reported as statistically significant.

## 493 Acknowledgements

494 This work was funded by the Special Research Fund of Ghent University
495 (Bijzonder Onderzoeksfonds, BOF, grant number BOFDOC2016001301) and the Swiss
496 National Science Foundation (31003A_169307).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

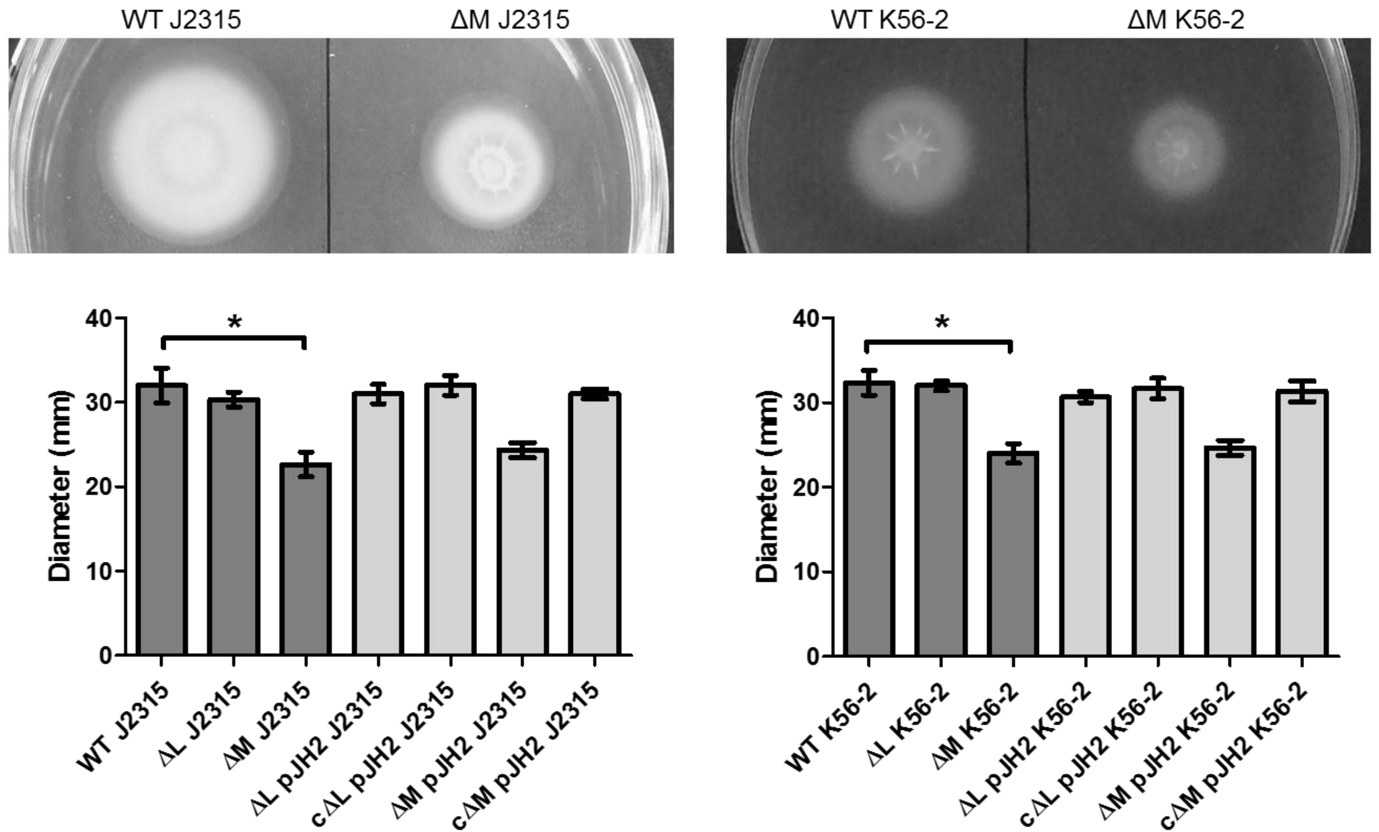
497 # Figures




498

499 **FIGURE 1** Effect of DNA MTase deletion on biofilm structure, cell aggregation, and
500 pellicle formation in *B. cenocepacia* J2315 and K56-2. (A) Microscopic images of
501 LIVE/DEAD stained biofilms, grown in microtiter plate wells for 24 h. White bar (200 μm)
502 for scale. (B) Clustering of cells in planktonic cultures, quantified with flow cytometry.
503 (C) Pellicle formation inside glass tubes after 24 h of static incubation. Left pictures
504 represent unstained samples, right pictures display pellicles stained with crystal violet.
505 (n=3, * $p < 0.05$ compared to wild type, error bars represent the standard error of the
506 mean (SEM)). WT: wild type, ΔL: deletion mutant ΔBCAL3494, ΔM: deletion mutant
507 ΔBCAM0992)

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.




508

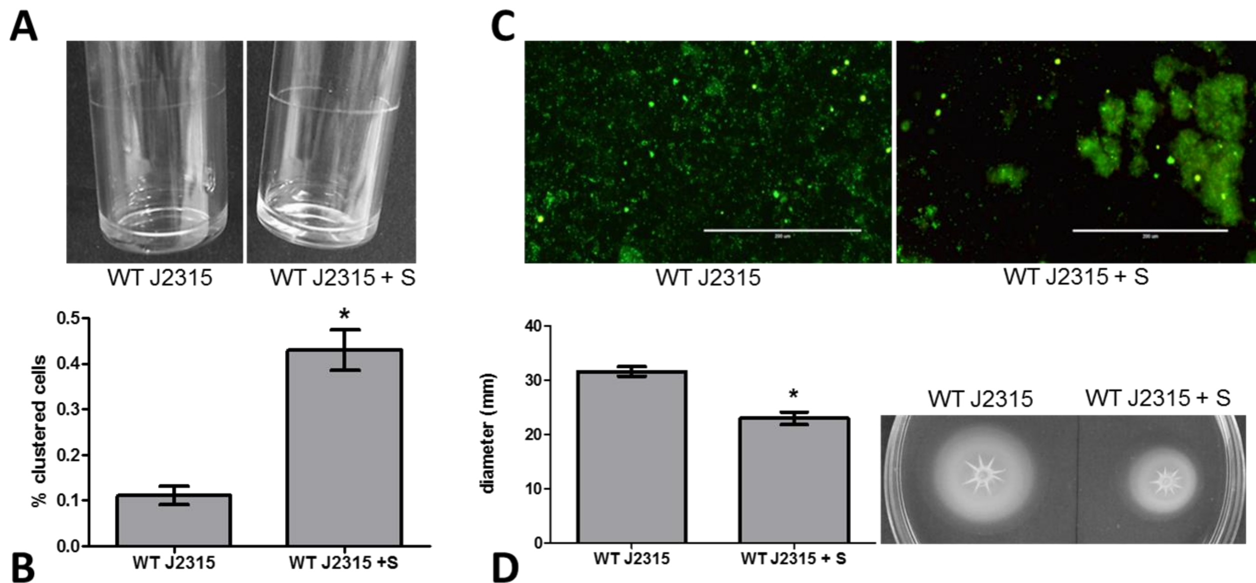
509 **FIGURE 2** Swimming motility of DNA MTase deletion mutants. Diameters were
510 measured after 24 h (K56-2) or 32 h (J2315). (n=3, * $p < 0.05$ compared to wild type,
511 error bars represent the SEM. WT: wild type, ΔL: deletion mutant ΔBCAL3494, ΔM:
512 deletion mutant ΔBCAM0992, ΔL pJH2 and ΔM pJH2: mutant strains with empty vector
513 pJH2 (vector control), *c*ΔL pJH2 and *c*ΔM pJH2: deletion mutants complemented with
514 genes BCAL3494 and BCAM0992)

515

516

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

517




518

519 **FIGURE 3** Effect of DNA MTase inhibitor sinefungin on biofilm and pellicle formation,
520 cell aggregation and motility. (A) Pellicle formation inside glass tubes after 24 h of static
521 incubation. (B) Clustering of planktonic cultures analyzed with flow cytometry. (C)
522 Microscopic images of LIVE/DEAD stained biofilms, grown on plastic surfaces in
523 microtiter plates for 24 h. (D) Swimming motility of treated and untreated samples. (n=3,
524 * $p < 0.05$ compared to wild type, error bars represent the SEM). WT: wild type, +S:
525 medium supplemented with 50 µg/mL sinefungin)

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

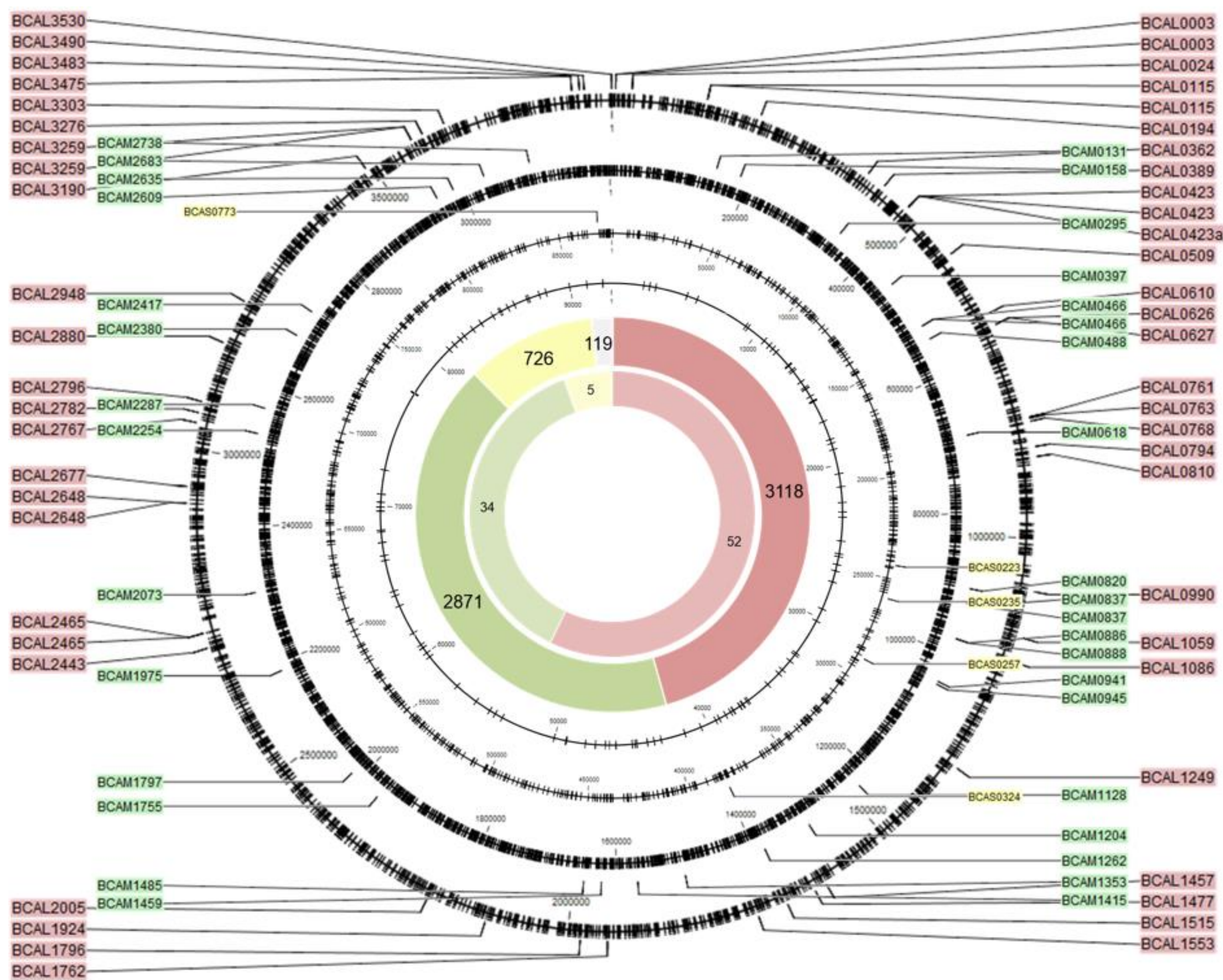



**FIGURE 4** Genomic position of all methylated CACAG motifs. Black circles represent the four replicons of *B. cenocepacia*, black ticks mark the motif locations. The total number of methylated CACAG motifs and methylated CACAG motifs in promoter regions, per replicon (red: replicon 1, green: replicon 2, yellow: replicon 3, and grey: plasmid), is shown on the large and small inner circle, respectively. The position and names of genes with methylated promoter regions is indicated with colored labels (same color code).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

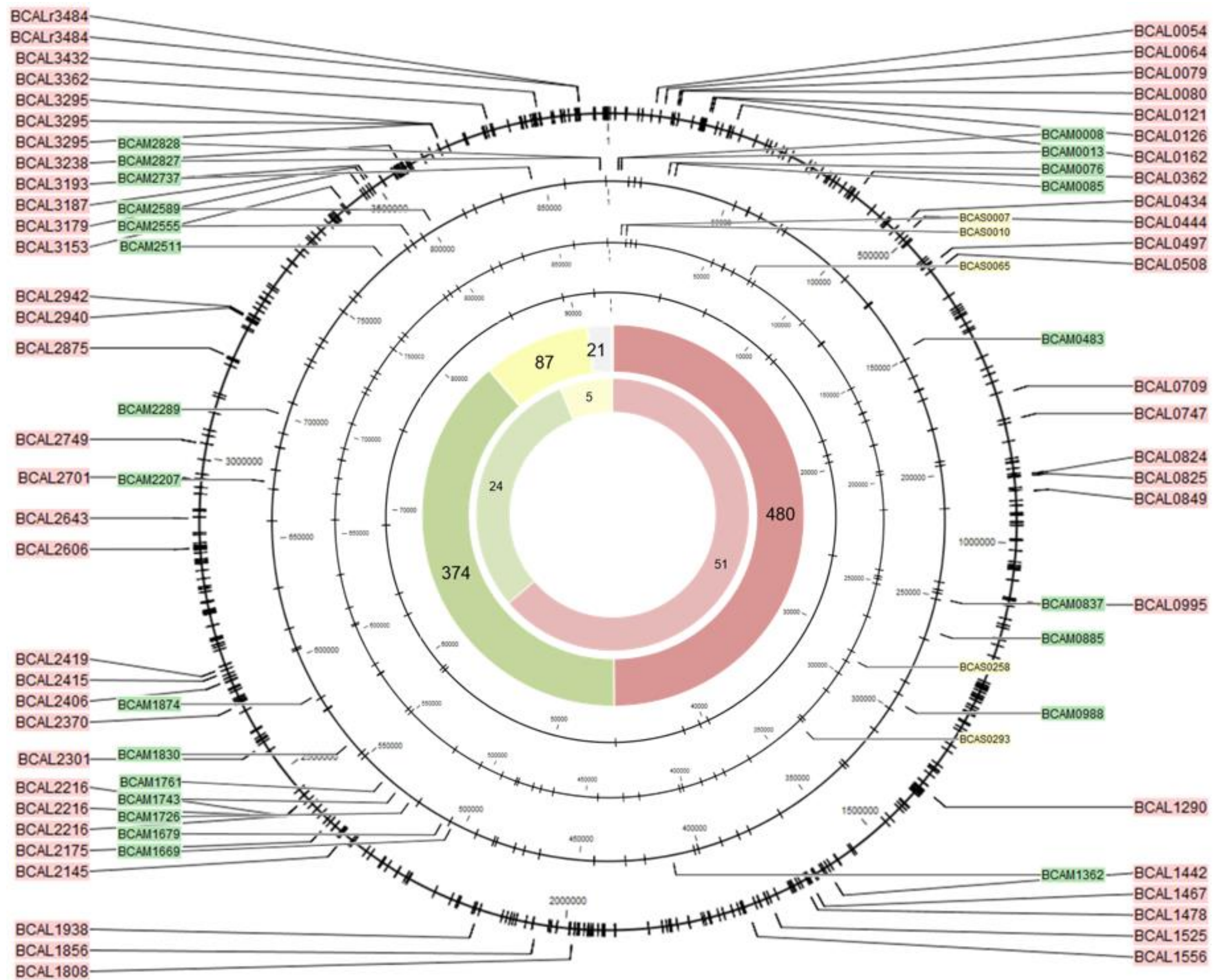



**FIGURE 5** Genomic position of all methylated GTWWAC motifs. Black circles represent the four replicons of *B. cenocepacia*, black ticks mark the motif locations. The total number of methylated GTWWAC motifs and methylated GTWWAC motifs in promoter regions, per replicon (red: replicon 1, green: replicon 2, yellow: replicon 3, and grey: plasmid), is shown on the large and small inner circle, respectively. The position and names of genes with methylated promoter regions is indicated with colored labels (same color code).

| | Gene | Sequence (-10 … -30) | Strain |
|---|---|---|---|
| **A** | BCAL1515 | TGCGCGCAG[CATAT**C]ACAG**AACCGATACC[GGAGT]TCACGGCGAAATGTGCCTGAAAGCT | J2315 + K56-2 |
| | BCAL2465 | GATCGATAA[AATAAA]TCAATTGACTGTGC[GGCGC]AACTTGCCACATT**CTGTG**TCGGTCG | J2315 + K56-2 |
| | BCAM0820 | CCGGCTGAT[TCTAAA]TGAAATTACAA**CAC[AG**CAT]CTATCTCCTTCCGGTTTAGTTTTGC | J2315 + K56-2 |
| | BCAL0423 | GGCGAGATT[CTAACG]CCAAACGCGGC**CAC[AG**CGG]GACTTATC**CACAG**GCCGGTTGTGTG | K56-2 |
| | BCAM2738 | ATCGCGCAT[TTTAAC]CCTATCGGGCCGTC[CGTGC]TTTGGTG**CTGTG**TCGGCACCGTTGC | K56-2 |
| | BCAS0223 | GGTTGCAAT[TCTT**CT]GTG**AGTCGATGTTT[CGGAC]AAACGAATTTTCCGAACGCATTGCA | K56-2 |
| **B** | BCAL0079 | GACGCGCAT[**GTTAAC**]ACGTCGGCGGCGGG[TATTT]CGGGCGCGCCGCCGGACGGGGCGCG | J2315 + K56-2 |
| | BCAL2415 | GGCTTCGAT[**GTTAAC**]ACGGGCATCCGGGC[CGGTT]TGGGCGGTTTGGTCGCGCCCACGCC | J2315 + K56-2 |
| | BCAM1362 | AATGCAATT[GTAATA]AATGA**GTAAAC**ACC[GTGTG]GGGAAAGTATGTTCCGGCCCCGCAA | J2315 + K56-2 |
| | BCAL1556 | ACCGCCA**GT[TTAC**CG]GAGTCGGCGGACGC[CCCCG]CGTTTTTCGGCGTGCGGGGCGTCGC | K56-2 |

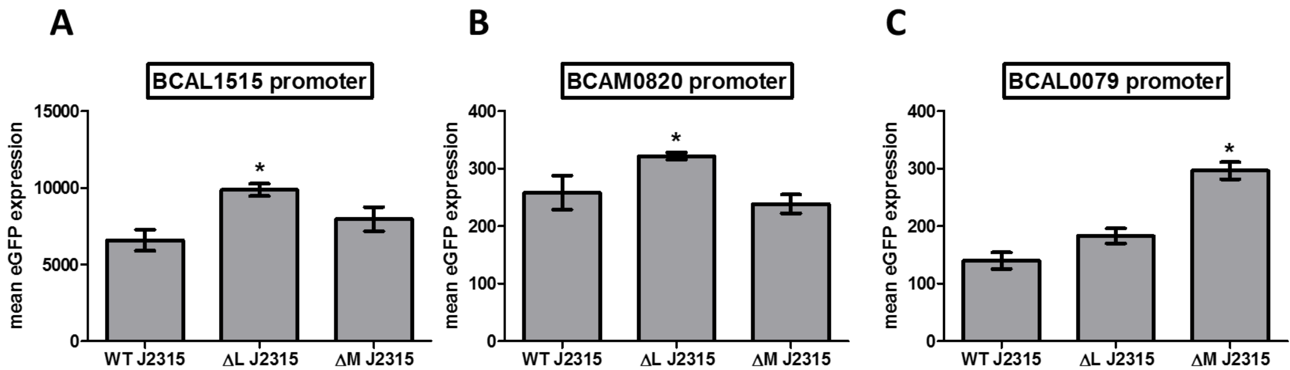
bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

545 **FIGURE 6** Position of methylated motifs relative to gene start for genes of which the
546 expression is upregulated in DNA MTase deletion mutants. (A) Genes with methylated
547 CACAG motifs in their corresponding promoter region. (B) Genes with methylated
548 GTWWAC motifs in their corresponding promoter region. The motifs are marked in bold,
549 the position of -10 and -30/35 elements in bacterial promoters are framed. ('J2315 +
550 K56-2': upregulation in both strains, 'K56-2': upregulation in strain K56-2 only)

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

551

552

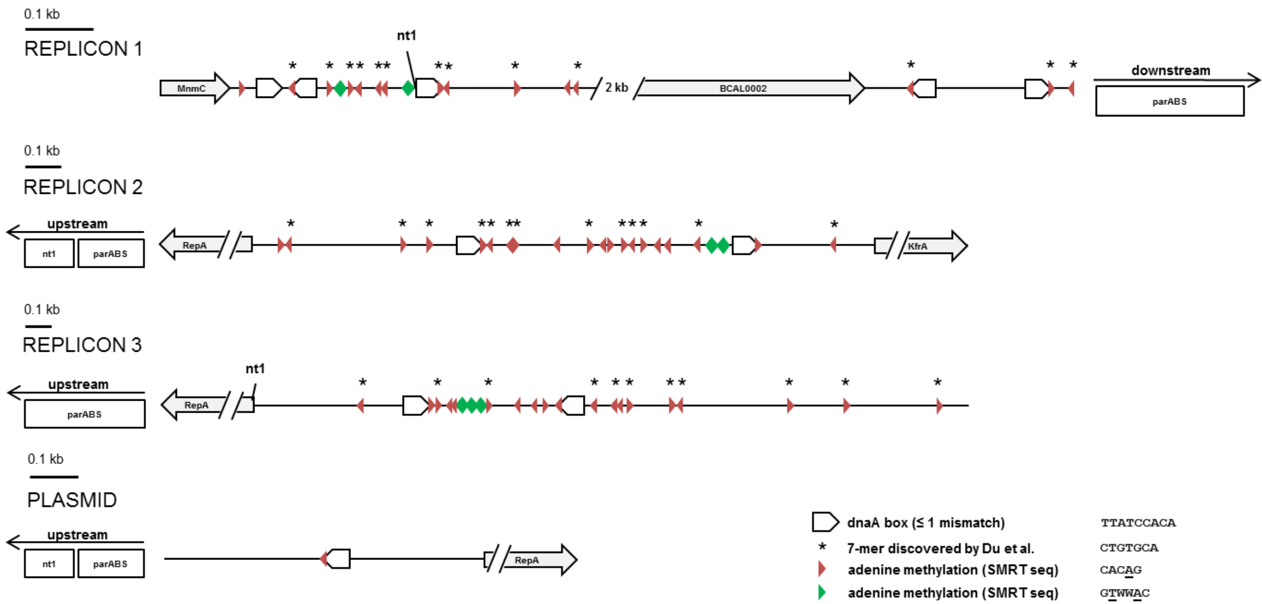



553

554 **FIGURE 7** eGFP production in *B. cenocepacia* J2315 strains harboring a pJH2 plasmid
555 that contains a BCAL1515 promoter-eGFP construct (A), a BCAM0820 promoter-eGFP
556 construct (B), or a BCAL0079 promoter-eGFP construct (C). BCAL 1515 and
557 BCAM0820 are associated with methylation of the CACAG motif by DNA MTase
558 BCAL3494, BCAL0079 is associated with methylation of the GTWWAC motif by DNA
559 MTase BCAL0992 (n=3, * $p < 0.05$ compared to wild type, error bars represent the
560 SEM). WT: wild type, ΔL: deletion mutant ΔBCAL3494, ΔM: deletion mutant
561 ΔBCAM0992).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

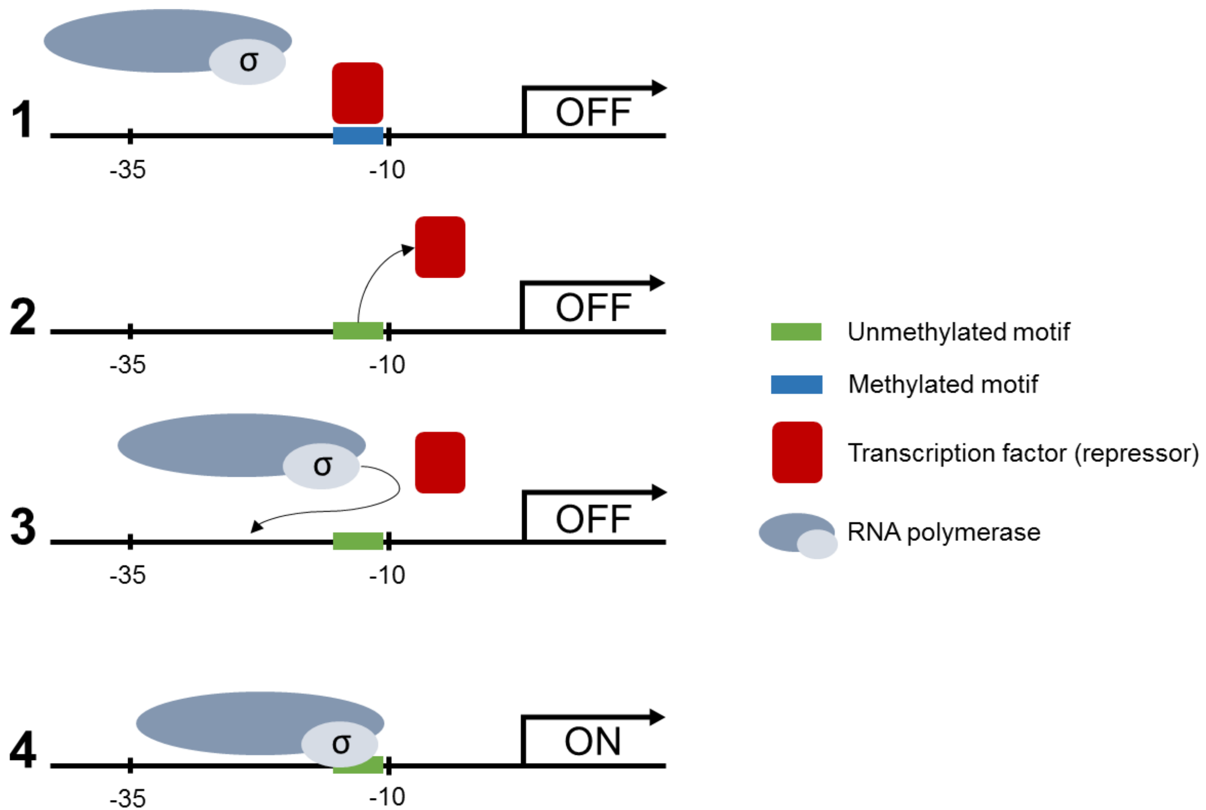
562




563

564 **FIGURE 8** Methylation in the origin of replication of the different replicons in
565 *B. cenocepacia* J2315. SMRT sequencing was used to detect methylated CACAG (red
566 triangles) and GTWWAC (green triangles) motifs within these regions. DnaA boxes
567 (TTATCCACA, consensus sequence of DnaA boxes in *E. coli*) are indicated in the
568 figure. CACAG motifs were frequently found to be part of a previously discovered 7-mer
569 (sense: CTGTGCA, antisense: TGCACAG) (30). The positions of these 7-mers are
570 indicated with an asterisk. (nt1, nucleotide 1; parABS genes, responsible for
571 chromosome segregation in *B. cenocepacia*)

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.




572

573 **FIGURE 9** Proposed mechanism of regulation of gene expression in *B. cenocepacia*.
574 (1) Methylated motifs in the promoter region of the gene are bound by a TF, acting as
575 repressor (OFF state). (2) In absence of methylation in the promoter region, the TF
576 dissociates from the motif and vacates the promoter region. (3) The sigma factor is no
577 longer sterically hindered by a repressor and is able to bind to the promoter region. (4)
578 RNA polymerase can access the promoter region and start transcription of the gene
579 (ON state).

580

581

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

582 # Tables

583 **TABLE 1** Expression changes of genes with a methylated CACAG motif in their
584 promoter region in deletion mutants compared to wild type.

| Locus tag | J2315 | | K56-2 | | Gene function |
|---|---|---|---|---|---|
| | Fold change | p-value | Fold change | p-value | |
| **BCAL0003** | 0.954 | 0.791 | 1.242 | 0.214 | MarR family regulatory protein |
| **BCAL0024** | 1.477 | 0.143 | 0.909 | 0.678 | GidA tRNA uridine 5-carboxymethylaminomethyl modification enzyme |
| **BCAL0423** | 1.169 | 0.306 | 1.948 | 0.014 | DnaA chromosomal replication initiation protein |
| **BCAL0509** | 1.129 | 0.473 | 1.175 | 0.199 | MetK S-adenosylmethionine synthetase |
| **BCAL1059** | 1.129 | 0.662 | 0.767 | 0.457 | ArgD bifunctional N-succinyldiaminopimelate-aminotransferase/acetylornithine transaminase protein |
| **BCAL1457** | 1.343 | 0.309 | 1.793 | 0.056 | LysR family regulatory protein |
| **BCAL1515** | 1.790 | 0.032 | 1.869 | 0.012 | SucA 2-oxoglutarate dehydrogenase E1 component |
| **BCAL2465** | 1.277 | 0.047 | 2.042 | 0.014 | TetR family regulatory protein |
| **BCAL2767** | 1.281 | 0.270 | 1.397 | 0.382 | ArgF ornithine carbamoyltransferase |
| **BCAL2782** | 1.373 | 0.237 | 1.166 | 0.668 | PdxH pyridoxamine 5'-phosphate oxidase |
| **BCAL3303** | 1.048 | 0.845 | 1.093 | 0.071 | QueA S-adenosylmethionine:tRNA ribosyltransferase-isomerase |
| **BCAM0820** | 2.621 | 0.004 | 2.253 | 0.002 | hybrid two-component system kinase-response regulator protein |
| **BCAM0941** | 1.240 | 0.448 | 1.761 | 0.050 | *gnd* 6-phosphogluconate dehydrogenase |
| **BCAM1262** | 1.237 | 0.397 | 1.163 | 0.445 | dihydroxy-acid dehydratase |
| **BCAM1415** | 1.183 | 0.665 | 1.315 | 0.177 | AraC family regulatory protein |
| **BCAM2738** | 1.213 | 0.147 | 1.649 | 0.022 | IspH 4-hydroxy-3-methylbut-2-enyl diphosphate reductase |
| **BCAS0223** | 1.251 | 0.202 | 1.993 | 0.030 | AfcC fatty acid desaturase |

585

586

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

587 **TABLE 2** Expression changes of genes with a methylated GTWWAC motif in their
588 promoter region in deletion mutants compared to wild type.

| Locus tag | J2315 | | K56-2 | | Gene function |
|---|---|---|---|---|---|
| | Fold change | p-value | Fold change | p-value | |
| **BCAL0054** | 0.487 | 0.127 | 0.641 | 0.107 | MerR family regulatory protein |
| **BCAL0079** | 2.838 | 0.020 | 3.074 | 0.005 | *rep* ATP-dependent DNA helicase |
| **BCAL0126** | 0.845 | 0.356 | 0.801 | 0.607 | MotA chemotaxis protein |
| **BCAL0162** | 0.139 | 0.634 | 0.133 | 0.478 | GmhA phosphoheptose isomerase |
| **BCAL0508** | 1.015 | 0.934 | 1.721 | 0.137 | LpxL lipid A biosynthesis myristoyl acyltransferase |
| **BCAL0709** | 1.599 | 0.104 | 0.763 | 0.430 | LipB lipoate-protein ligase B |
| **BCAL1556** | 1.611 | 0.171 | 1.690 | 0.006 | RpiA ribose-5-phosphate isomerase A |
| **BCAL2406** | 1.693 | 0.273 | 1.009 | 0.938 | WabR putative glycosyltransferase |
| **BCAL2415** | 2.819 | 0.006 | 6.029 | 0.001 | PurT phosphoribosylglycinamide formyltransferase 2 |
| **BCAL2701** | 0.613 | 0.076 | 1.519 | 0.170 | A*rg*D acetylornithine transaminase protein |
| **BCAL2942** | 1.143 | 0.474 | 1.451 | 0.274 | CysM cysteine synthase B |
| **BCAM0076** | 1.630 | 0.112 | 1.358 | 0.051 | TetR family regulatory protein |
| **BCAM1362** | 1.959 | 0.025 | 1.516 | 0.004 | putative penicillin-binding protein |
| **BCAS0258** | 1.247 | 0.451 | 1.141 | 0.438 | GntR family regulatory protein |

589

590

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

591 **TABLE 3** Bacterial strains and plasmids.

| Strain/Plasmid | Description | Abbreviation | Source |
|---|---|---|---|
| ***B. cenocepacia*** | | | |
| J2315 | CF sputum isolate | WT J2315 | LMG16656 |
| J2315 ΔBCAL3494 | BCAL3494 MTase deletion mutant | ΔL J2315 | this study |
| J2315 ΔBCAM0992 | BCAM0992 MTase deletion mutant | ΔM J2315 | this study |
| J2315 ΔBCAL3494 pJH2 | BCAL3494 MTase mutant with empty pJH2 vector | ΔL pJH2 J2315 | this study |
| J2315 ΔBCAL3494 pJH2 + BCAL3494 | BCAL3494 MTase complemented deletion mutant | *c*ΔL pJH2 J2315 | this study |
| J2315 ΔBCAM0992 pJH2 | BCAM0992 MTase mutant with empty pJH2 vector | ΔM pJH2 J2315 | this study |
| J2315 ΔBCAM0992 pJH2 + BCAM0992 | BCAM0992 MTase complemented deletion mutant | *c*ΔM pJH2 J2315 | this study |
| K56-2 | CF sputum isolate | WT K56-2 | LMG 18863 |
| K56-2 ΔBCAL3494 | BCAL3494 MTase deletion mutant | ΔL K56-2 | this study |
| K56-2 ΔBCAM0992 | BCAM0992 MTase deletion mutant | ΔM K56-2 | this study |
| K56-2 ΔBCAL3494 pJH2 | BCAL3494 MTase mutant with empty pJH2 vector | ΔL pJH2 K56-2 | this study |
| K56-2 ΔBCAL3494 pJH2 + BCAL3494 | BCAL3494 MTase complemented deletion mutant | *c*ΔL pJH2 K56-2 | this study |
| K56-2 ΔBCAM0992 pJH2 | BCAM0992 MTase mutant with empty pJH2 vector | ΔM pJH2 K56-2 | this study |
| K56-2 ΔBCAM0992 pJH2 + BCAM0992 | BCAM0992 MTase complemented deletion mutant | *c*ΔM pJH2 K56-2 | this study |
| ***E. coli*** | | | |
| DH5α | Maintenance of replicative plasmids | | lab stock |
| One Shot PIR2 | Maintenance of suicide plasmids with $ori_{R6K}$ | PIR2 | ThermoFisher |
| **Plasmids** | | | |
| pGPI-SceI-XCm | Suicide plasmid, $Tp^r$, $Cm^r$, I-SceI restriction site, $ori_{R6K}$ | pGPI | 44,45 |
| pDAI-SceI-SacB | Replicative plasmid, $Tet^r$, I-SceI nuclease, counterselectable marker SacB, $ori_{pBBR1}$ | pDAI | 44,45 |
| pRK2013 | Helper plasmid, $Km^r$, $ori_{colEI}$ | pRK | 44,45 |
| pJH2 | Broad-range translational fusion vector, $Cm^r$, fluorescent marker eGFP: complementation of ΔBCAL3494 | | 46 |
| pSCrhaB2 | Broad-range translational fusion vector, $Tp^r$, rhaR, rhaS-$P_{rhaB}$, $ori_{pBBr1}$: complementation of ΔBCAM0992 | | 47 |
| pGPI + BCAL3494 upstream sequence | pGPI-SceI-XCm with ligated upstream sequence BCAL3494, used during deletion | $pGPI_{UL}$ | this study |
| pGPI + BCAL3494 upstream and downstream sequence | pGPI-SceI-XCm with ligated upstream and downstream sequence BCAL3494, used during deletion | $pGPI_{UL-DL}$ | this study |
| pGPI + BCAM0992 upstream sequence | pGPI-SceI-XCm with ligated upstream sequence BCAM0992, used during deletion | $pGPI_{UM}$ | this study |
| pGPI + BCAM0992 upstream and downstream sequence | pGPI-SceI-XCm with ligated upstream and downstream sequence BCAM0992, used during deletion | $pGPI_{UM-DM}$ | this study |

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

| | | | |
|---|---|---|---|
| pJH2 + BCAL3494 sequence | Fusion vector with ligated BCAL3494 sequence, used for complementation | $pJH2_{L3494}$ | this study |
| pSCrhaB2 + BCAM0992 sequence | Fusion vector with ligated BCAM0992 sequence, used for complementation | $pSCrhaB2_{M0992}$ | this study |

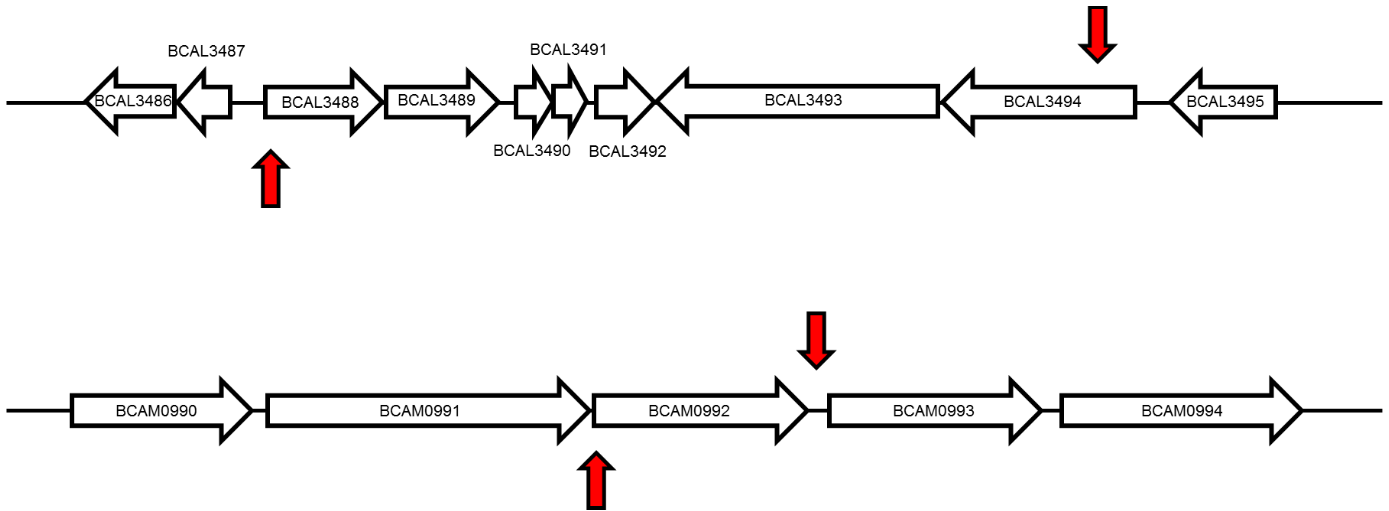
592

593

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

594 **Supplementary data**

595

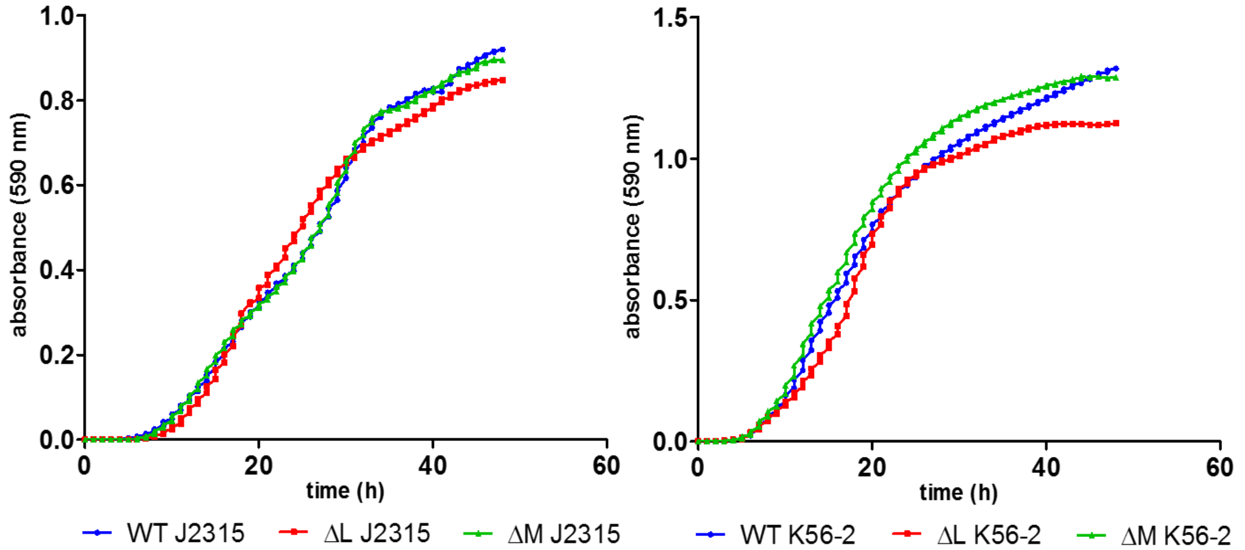



596

597 **FIGURE S1** Genome context of deleted DNA MTase genes BCAL3494 (top) and
598 BCAM0992 (bottom). The red arrows indicate the boundaries of the deleted part. For
599 gene BCAL3494, adjacent restriction gene BCAL3493, and hypothetical genes
600 BCAL3488 to BCAL3492 were deleted as well.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

601

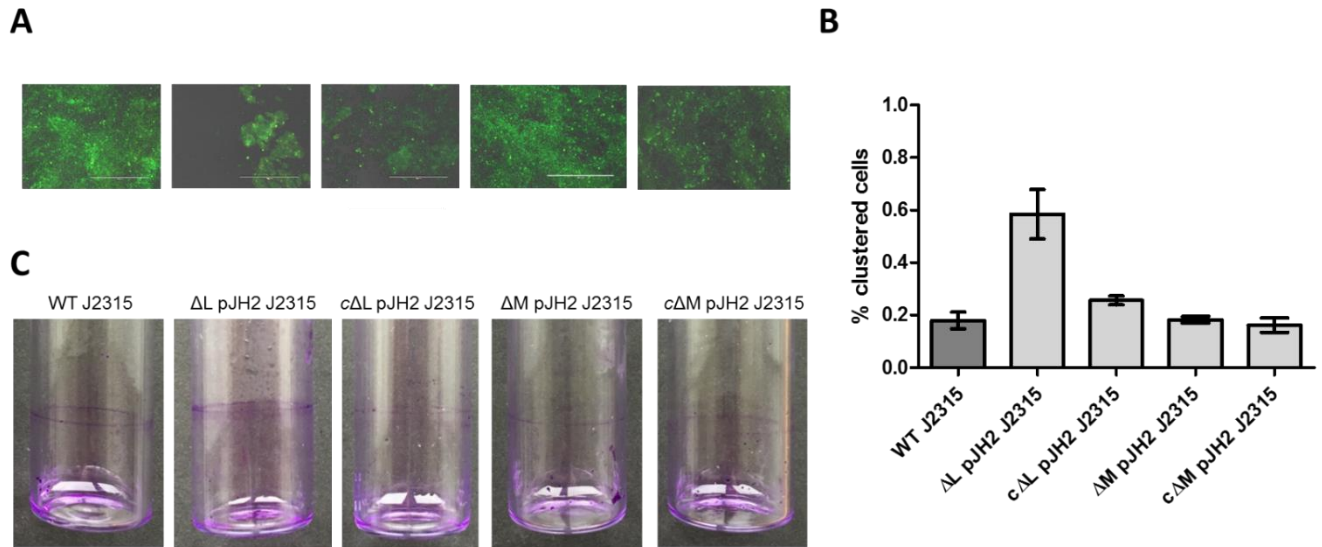



602

603 **FIGURE S2** Growth of *B. cenocepacia* J2315 (left) and K56-2 (right) in phosphate
604 buffered minimal medium. (WT: wild type, ΔL: deletion mutant ΔBCAL3494, ΔM:
605 deletion mutant ΔBCAM0992)

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

606

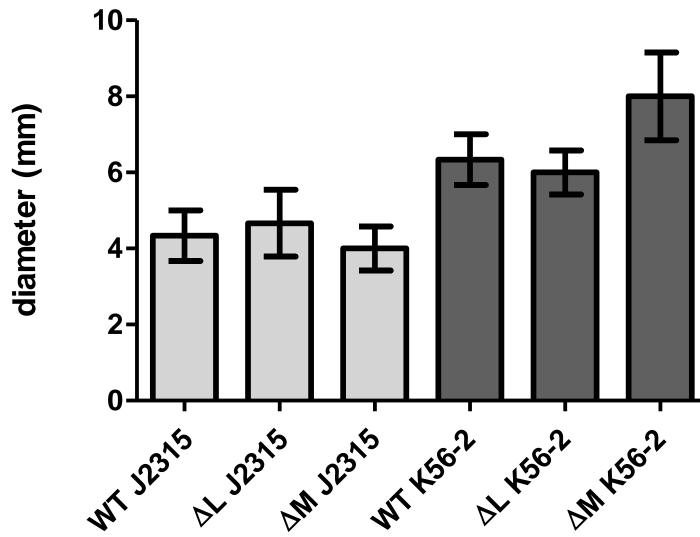



607

608 **FIGURE S3** Biofilm formation, cell aggregation, and pellicle formation of complemented
609 DNA MTase mutants in *B. cenocepacia* J2315. (A) Microscopic images of LIVE/DEAD
610 stained biofilms, grown in microtiter plate wells for 24 h. White bar (200 µm) for scale.
611 (B) Clustering of cells in planktonic cultures, analyzed with flow cytometry. (C) Pellicle
612 formation inside glass tubes after 24 h of static incubation, stained with crystal violet.
613 (n=3, error bars represent the Standard Error of the Mean (SEM). WT: wild type, ΔL
614 pJH2 and ΔM pJH2: mutant strains with empty vector pJH2 (vector control), *c*ΔL pJH2
615 and *c*ΔM pJH2: deletion mutants complemented with genes BCAL3494 and
616 BCAM0992)

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

617

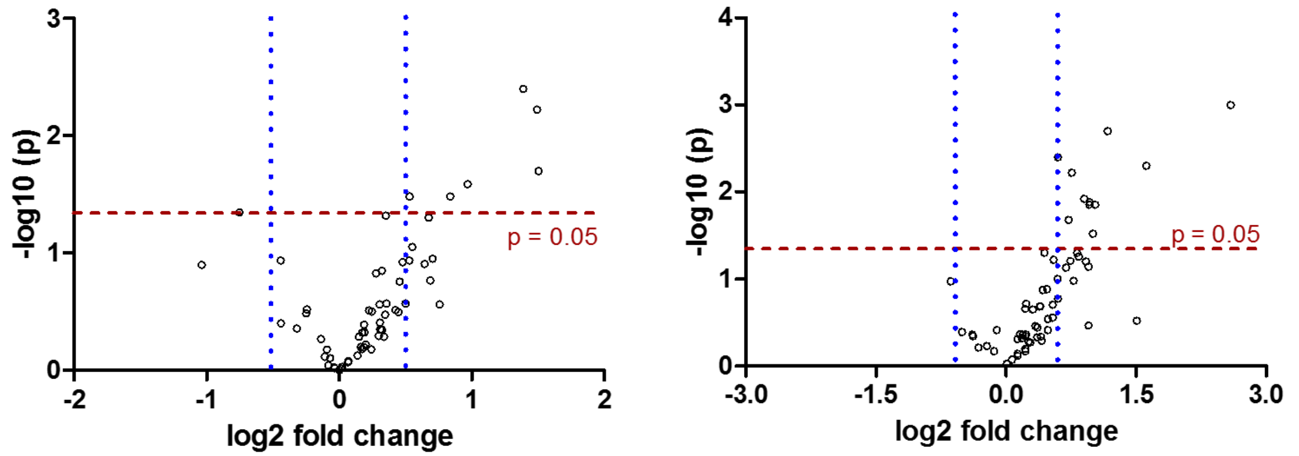



618

619 **FIGURE S4** Swarming motility of DNA MTase deletion mutants. Diameters were
620 measured after 24 h. (n=3, error bars represent the SEM. WT: wild type, ΔL: deletion
621 mutant ΔBCAL3494, ΔM: deletion mutant ΔBCAM0992)

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

622



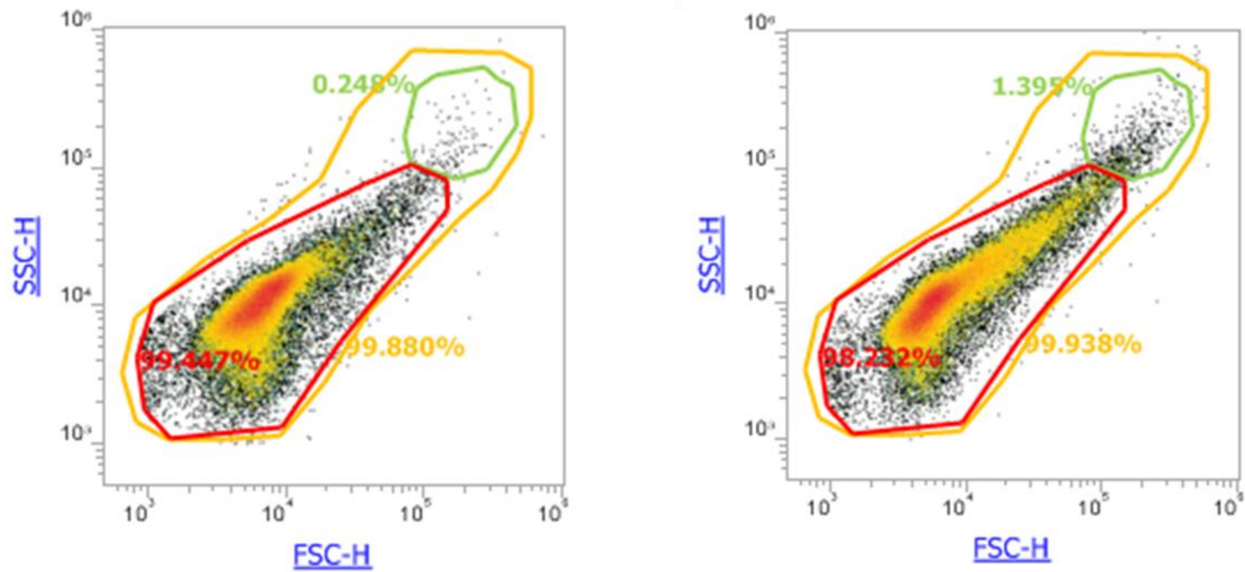

623

624 **FIGURE S5** Differential expression (Volcano plots) of all genes with methylated
625 promoter region in J2315 (left) and K56-2 (right) for which expression was quantified
626 using qPCR. Cut-offs were drawn at fold changes -1.5 and 1.5 (blue) and at p-value
627 0.05 (red). All genes outside of these cut-offs were considered significantly up- or
628 downregulated.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

629

630




631

632 **FIGURE S6** Quantification of the number of clusters in wild type J2315 (left) and
633 ΔBCAL3494 J2315 (right) (SSC: side scatter, FSC: forward scatter, green circle
634 indicates clusters, red circle indicates main population (ranging from approx. $10^3$ to $10^5$
635 for both FSC and SSC)).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

636

637 **TABLE S1** All DNA MTase genes in the *B. cenocepacia* J2315 genome identified by
638 REBASE (* BLASTn search against the genus *Burkholderia*, default screening
639 parameters were used).

| Locus Tag | Name | Location | Predicted type MTase (48) | Widely distributed in genus *Burkholderia* * |
|---|---|---|---|---|
| **BCAL3494** | M.BceJI | CHR 1 | Type III | ✓ |
| **BCAM1036** | M.BceJII | CHR 2 | | |
| **pBCA072** | M.BceJIII | PLASMID | | |
| **BCAM0992** | M.BceJIV | CHR 2 | | ✓ |
| **BCAL0178** | M.BceJ178P | CHR 1 | Type II | |
| **BCAL0414** | M.BceJ414P | CHR 1 | Type I | |

640

641 **TABLE S2** Methylation motifs in *B. cenocepacia* (methylated bases in bold and
642 underlined, 6mA: $N_6$-methyl adenine, 4mC: $N_4$-methyl cytosine, F: forward strand, R:
643 reverse strand). * wild type strain J2315 percentages.

| Motif | Methylation | Strand | Called modified motifs * (%) | Methylation of motif by |
|---|---|---|---|---|
| CAC**A**G | 6mA | F | 6834/6836 (99.9) | BCAL3494 |
| G**T**WW**A**C | 6mA | F+R | 961/982 (97.8) | BCAM0992 |
| GCG**GC**CGC | 4mC | F+R | 1738/6850 (25.3) | unknown |

644

645 **TABLE S3** Genes with methylated promoter region (CACAG motif) in J2315
646 (methylated promoter regions in K56-2 are indicated with '+', non-methylated promoter
647 regions with '-').

| | Locus tag | Gene function | Methylated in K56-2 |
|---|---|---|---|
| | | **Intermediary metabolism** | |
| **1** | BCAL0627 | putative hydrolase | + |
| **2** | BCAL1059 | ArgD bifunctional N-succinyldiaminopimelate-aminotransferase/acetylornithine transaminase protein | + |
| **3** | BCAL1086 | putative lipoprotein | + |
| **4** | BCAL1249 | putative PHB depolymerase | + |
| **5** | BCAL1515 | SucA 2-oxoglutarate dehydrogenase E1 component | + |
| **6** | BCAL1762 | acetyltransferase (GNAT) family protein | + |
| **7** | BCAL1796 | putative saccharopine dehydrogenase | + |
| **8** | BCAL1924 | MoeA3 molybdopterin biosynthesis protein | + |
| **9** | BCAL2767 | ArgF ornithine carbamoyltransferase | + |
| **10** | BCAL2782 | PdxH pyridoxamine 5'-phosphate oxidase | + |
| **11** | BCAL2796 | benzoylformate decarboxylase | + |
| **12** | BCAM0941 | *gnd* 6-phosphogluconate dehydrogenase | + |

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

| | | | |
|---|---|---|---|
| **13** | BCAM1128 | putative glycosyl transferase family protein | + |
| **14** | BCAM1204 | DadX alanine racemase | + |
| **15** | BCAM1262 | IlvD dihydroxy-acid dehydratase | + |
| **16** | BCAM1353 | *ald* alanine dehydrogenase | + |
| **17** | BCAM1459 | AtoE short-chain fatty acid transporter | + |
| **18** | BCAM1485 | ornithine cyclodeaminase | + |
| **19** | BCAM2380 | putative D-isomer specific 2-hydroxyacid dehydrogenase | + |
| **20** | BCAM2738 | IspH 4-hydroxy-3-methylbut-2-enyl diphosphate reductase | + |
| **21** | BCAS0223 | AfcC putative fatty acid desaturase | + |
| **22** | BCAS0257 | putative acetyltransferase | + |
| **23** | BCAS0324 | sugar ABC transporter ATP-binding protein | + |
| | | **Regulation** | |
| **1** | BCAL0003 | MarR family regulatory protein | + |
| **2** | BCAL0003 | MarR family regulatory protein | + |
| **3** | BCAL1457 | LysR family regulatory protein | + |
| **4** | BCAL1477 | LysR family regulatory protein | + |
| **5** | BCAL2443 | GntR family regulatory protein | + |
| **6** | BCAL2465 | TetR family regulatory protein | + |
| **7** | BCAL2465 | TetR family regulatory protein | + |
| **8** | BCAL3190 | IclR family regulatory protein | + |
| **9** | BCAM0466 | LysR family regulatory protein | + |
| **10** | BCAM0466 | LysR family regulatory protein | + |
| **11** | BCAM0618 | two-component regulatory system response regulator protein | + |
| **12** | BCAM0820 | hybrid two-component system kinase-response regulator protein | + |
| **13** | BCAM0886 | LysR family regulatory protein | + |
| **14** | BCAM1415 | AraC family regulatory protein | + |
| **15** | BCAM1755 | GntR family regulatory protein | + |
| **16** | BCAM1975 | AraC family regulatory protein | + |
| **17** | BCAS0235 | two-component regulatory system, response regulator protein | + |
| | | **Membrane and transport** | |
| **1** | BCAL2005 | putative membrane protein | + |
| **2** | BCAL2648 | putative outer membrane protein | + |
| **3** | BCAL2648 | putative outer membrane protein | + |
| **4** | BCAL2677 | putative permease protein | + |
| **5** | BCAL2948 | putative membrane protein | + |
| **6** | BCAL3490 | putative exported protein | + |
| **7** | BCAM0837 | putative membrane protein | + |
| **8** | BCAM0837 | putative membrane protein | + |
| **9** | BCAM0945 | putative membrane protein | + |
| **10** | BCAM1797 | putative ion channel protein | + |
| **11** | BCAM2683 | putative cation-transporting ATPase membrane protein | + |

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

| | | Protein synthesis | |
|---|---|---|---|
| 1 | BCAL0024 | GidA tRNA uridine 5-carboxymethylaminomethyl modification enzyme | + |
| 2 | BCAL0115 | RpsU 30S ribosomal protein S21 | + |
| 3 | BCAL0115 | RpsU 30S ribosomal protein S21 | + |
| 4 | BCAL0423a | RpmH 50S ribosomal protein L34 | + |
| 5 | BCAL0990 | RpmF 50S ribosomal protein L32 | + |
| 6 | BCAL1553 | putative ribonuclease | + |
| 7 | BCAL2880 | RpmF 50S ribosomal protein L32 | + |
| 8 | BCAL3303 | queA S-adenosylmethionine:tRNA ribosyltransferase-isomerase | + |
| 9 | BCAL3530 | HupA DNA-binding protein HU-alpha | + |
| 10 | BCAM0131 | HchA chaperone protein | + |
| 11 | BCAM0158 | putative diguanylate phosphodiesterase | + |
| | | **Electron transport and ATP synthesis** | |
| 1 | BCAL0194 | putative oxidoreductase | + |
| 2 | BCAL0389 | DsbC thiol:disulfide interchange protein | + |
| 3 | BCAL0626 | putative 2-nitropropane dioxygenase | + |
| 4 | BCAL3276 | PpnK NAD(+)/NADH kinase family protein | + |
| 5 | BCAL3475 | putative molybdopterin-containing oxidoreductase | + |
| | | **Chromosome replication** | |
| 1 | BCAL0423 | DnaA chromosomal replication initiation protein | + |
| 2 | BCAL0423 | DnaA chromosomal replication initiation protein | + |
| | | **Proteins involved in DNA methylation** | |
| 1 | BCAL0509 | MetK S-adenosylmethionine synthetase | + |
| | | **Hypothetical proteins and pseudogenes** | |
| 1 | BCAL0362 | conserved hypothetical protein | + |
| 2 | BCAL0610 | conserved hypothetical protein | + |
| 3 | BCAL0761 | conserved hypothetical protein | + |
| 4 | BCAL0763 | conserved exported protein | + |
| 5 | BCAL0768 | conserved hypothetical protein | + |
| 6 | BCAL0794 | conserved hypothetical protein | + |
| 7 | BCAL0810 | pseudogene | + |
| 8 | BCAL3259 | pseudogene | + |
| 9 | BCAL3259 | pseudogene | - |
| 10 | BCAL3483 | hypothetical protein | + |
| 11 | BCAM0295 | conserved hypothetical protein | + |
| 12 | BCAM0397 | conserved hypothetical protein | + |
| 13 | BCAM0488 | conserved hypothetical protein | + |
| 14 | BCAM0888 | conserved hypothetical protein | + |
| 15 | BCAM2073 | hypothetical protein | + |
| 16 | BCAM2254 | hypothetical protein | + |
| 17 | BCAM2287 | hypothetical protein | + |

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

| | | | |
|---|---|---|---|
| **18** | BCAM2417 | conserved hypothetical protein | + |
| **19** | BCAM2609 | hypothetical protein | + |
| **20** | BCAM2635 | hypothetical protein | + |
| **21** | BCAS0773 | hypothetical protein | + |

648

649 **TABLE S4** Genes with methylated promoter region (GTWWAC motif) in J2315
650 (methylated promoter regions in K56-2 are indicated with '+', non-methylated promoter
651 regions with '-').

| | Locus tag | Gene function | Methylated in K56-2 |
|---|---|---|---|
| | | **Intermediary metabolism** | |
| **1** | BCAL0064 | AcoD acetaldehyde dehydrogenase | + |
| **2** | BCAL0162 | GmhA phosphoheptose isomerase | + |
| **3** | BCAL0508 | LpxL lipid A biosynthesis myristoyl acyltransferase | + |
| **4** | BCAL0709 | LipB lipoate-protein ligase B | + |
| **5** | BCAL0995 | AcpP acyl carrier protein | + |
| **6** | BCAL1290 | undecaprenyl pyrophosphate phosphatase | + |
| **7** | BCAL1467 | AroC chorismate synthase | + |
| **8** | BCAL1478 | putative hydrolase | + |
| **9** | BCAL1556 | RpiA ribose-5-phosphate isomerase A | + |
| **10** | BCAL1938 | family C40 cysteine peptidase | + |
| **11** | BCAL2406 | WabR putative glycosyltransferase | + |
| **12** | BCAL2419 | glycosyl hydrolases family protein | + |
| **13** | BCAL2701 | ArgD acetylornithine transaminase protein | + |
| **14** | BCAL2875 | AcpP acyl carrier protein | + |
| **15** | BCAL2942 | CysM cysteine synthase B | + |
| **16** | BCAL3153 | putative lipoprotein | + |
| **17** | BCAL3179 | LdhA putative D-lactate dehydrogenase | + |
| **18** | BCAM0013 | putative acetyltransferase | + |
| **19** | BCAM1679 | putative lysylphosphatidylglycerol synthetase | + |
| **20** | BCAM1761 | putative lipoprotein | + |
| **21** | BCAM2511 | GarD putative D-galactarate dehydratase | + |
| **22** | BCAM2737 | putative glycosyl transferase | + |
| **23** | BCAS0065 | putative glutathione S-transferase | + |
| | | **Membrane and transport** | |
| **1** | BCAL0121 | AqpZ aquaporin Z | + |
| **2** | BCAL0126 | MotA chemotaxis protein | + |
| **3** | BCAL0824 | putative membrane protein | + |
| **4** | BCAL1525 | *flp* type pilus subunit | + |
| **5** | BCAL1808 | putative membrane protein | + |
| **6** | BCAL2301 | putative exported protein | + |

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

| | | | |
|---|---|---|---|
| **7** | BCAL2370 | putative membrane protein | + |
| **8** | BCAM0837 | putative membrane protein | + |
| **9** | BCAM0885 | putative membrane protein | + |
| **10** | BCAM0988 | putative exported protein | + |
| **11** | BCAM1669 | putative exported protein | + |
| **12** | BCAM1726 | putative exported protein | + |
| **13** | BCAM1743 | periplasmic solute-binding protein | + |
| **14** | BCAM1830 | putative exported protein | + |
| **15** | BCAM2555 | putative exported protein | + |
| **16** | BCAM2827 | putative exported protein | + |
| **17** | BCAM2828 | putative membrane protein | + |
| | | **Regulation** | |
| **1** | BCAL0054 | MerR family regulatory protein | + |
| **2** | BCAL0444 | GntR family regulatory protein | + |
| **3** | BCAL0497 | two-component regulatory system, sensor kinase protein | + |
| **4** | BCAL2606 | two-component regulatory system, response regulator protein | + |
| **5** | BCAM0076 | TetR family regulatory protein | + |
| **6** | BCAM0085 | TetR family regulatory protein | + |
| **7** | BCAM0483 | ADA-like AraC family regulatory protein | + |
| **8** | BCAM2589 | IclR family regulatory protein | + |
| **9** | BCAS0007 | TetR family regulatory protein | + |
| **10** | BCAS0258 | GntR family regulatory protein | + |
| | | **Electron transport and ATP synthesis** | |
| **1** | BCAL0080 | putative cytochrome | + |
| **2** | BCAL2145 | NADH-ubiquinone oxidoreductase subunit | + |
| **3** | BCAL2415 | PurT phosphoribosylglycinamide formyltransferase 2 | + |
| **4** | BCAL3187 | putative oxidoreductase | + |
| **5** | BCAL3362 | putative oxidoreductase | + |
| **6** | BCAL3432 | cytochrome c assembly protein | + |
| | | **DNA transposition** | |
| **1** | BCAL2216 | putative transposase | + |
| **2** | BCAL3238 | putative transposase | + |
| **3** | BCAL3295 | putative transposase | + |
| | | **Protein synthesis** | |
| **1** | BCAL1856 | RimO ribosomal protein S12 methylthiotransferase | + |
| **2** | BCALr3484 | tRNA-Val | + |
| | | **Chromosome replication** | |
| **1** | BCAL0079 | *rep* ATP-dependent DNA helicase | + |
| | | **Proteins involved in DNA methylation** | |
| **1** | BCAL0747 | putative methyltransferase | + |
| | | **Other proteins** | |

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

| | | | |
|---|---|---|---|
| **1** | BCAL0825 | UvrA excinuclease ABC subunit A | + |
| **2** | BCAL0849 | subfamily M48B metalopeptidase | + |
| **3** | BCAL2643 | SodC superoxide dismutase | + |
| **4** | BCAL2749 | putative diguanylate phosphodiesterase | + |
| **5** | BCAL2940 | putative histone deacetylase-family protein | + |
| **6** | BCAM1362 | putative penicillin-binding protein | + |
| **7** | BCAS0010 | putative activator of osmoprotectant transporter | + |
| **8** | BCAS0293 | AidA nematocidal protein | + |
| | | **Hypothetical proteins and pseudogenes** | |
| **1** | BCAL0362 | conserved hypothetical protein | + |
| **2** | BCAL0434 | putative exported protein | + |
| **3** | BCAL1442 | conserved hypothetical protein | + |
| **4** | BCAL2175 | conserved hypothetical protein | + |
| **5** | BCAL3193 | conserved hypothetical protein | + |
| **6** | BCAM0008 | conserved hypothetical protein | + |
| **7** | BCAM1874-2 | pseudo | + |
| **8** | BCAM2207 | conserved hypothetical protein | + |
| **9** | BCAM2289 | conserved hypothetical protein | + |

652

653 **TABLE S5** List of TFs that bind to methylation motifs CACAG and GTWWAC, predicted
654 by Virtual Footprint. Bold sequences represent methylation motifs. (Consensus
655 sequence based on TF binding in *E. coli* K12)

| Transcription Factor | Strand | Score | Consensus sequence |
|---|---|---|---|
| **CACAG motif** | | | |
| **GlpR** | - | 6.07 | **TGTG**TTCTAATTTCATTTAG |
| **GTWWAC motif** | | | |
| **ArcA** | + | 7.39 | T**GTTAAC**ATG |
| **ArcA** | - | 7.22 | T**GTTAAC**ACG |
| **OxyR** | - | 4.49 | CAT**GTTAAC**AC |
| **OxyR** | + | 4.11 | CGT**GTTAAC**AT |
| **Fis** | - | 3.49 | GACGCGCAT**GTTAAC** |
| **Fur** | - | 2.55 | AT**GTTAAC** |

656

657 **TABLE S6** Overview of all primers used for construction and complementation of gene
658 deletion mutants.

| Primers | Sequence (5' 3') | Abbreviation |
|---|---|---|
| | **ΔBCAL3494 primers** | |
| **Flanking sequences** | | |
| upstream sequence F | ATATGAATTCCCAACGGTTTCAAGGAGACG | UL3494-EcoRI |

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

| | | |
|---|---|---|
| upstream sequence R | ATATAGATCTGGCGGATCGATGTAGACGAG | UL3494-Bgl II |
| downstream sequence F | ATATAGATCTGGGATGCAAGAAGGCTCATC | DL3494-Bgl II |
| downstream sequence R | TTTACCCGGGATAGGTCTCGCGCTGGTGTC | DL3494-SmaI |
| **Control primers** | | |
| Overlapping sequence insert control F | ATGGAGAATCCCGGAAGAAG | joinL3494-F |
| Overlapping sequence insert control R | TGCTGTTTCATCTGGTGCTC | joinL3494-R |
| BCAL3494 gene control F | GGCAGCGATTTCGTCTATCC | geneL3494-F |
| BCAL3494 gene control R | CACTTCGTGCTCGTCGATGT | geneL3494-R |
| **Complementation** | | |
| BCAL3494 complementation F | TTTGGATCCTCGTCCTGTTTCAGCCTTTGAGC | L3493-4-ov-BamHI |
| BCAL3494 complementation R | TTTTCTAGAGCTTTCACGCGAATGACAGGATG | L3493-4-ov-XbaI |
| | **ΔBCAM0992 primers** | |
| **Flanking sequences** | | |
| upstream sequence F | ATATGAATTCGATCTACCTGAAGCGCGAAG | UM0992-EcoRI |
| upstream sequence R | ATATGCTAGCGGCTCTTCGATCAGGTCACG | UM0992-NheI |
| downstream sequence F | ATATGCTAGCCGTATGAGACCGGAGCAAGC | DM0992-NheI |
| downstream sequence R | ATATAGATCTCACTTGACCCACAGGCCTTC | DM0992-Bgl II |
| **Control primers** | | |
| Overlapping sequence insert control F | ATACCTCGGTGCAGCTGATC | joinM0992-F |
| Overlapping sequence insert control R | CAATGCTCGAAACATCCAGA | joinM0992-R |
| BCAM0992 gene control F | AACGATTCGGACAAGCGTTC | geneM0992-F |
| BCAM0992 gene control R | CGGTCCCAGATGATCTCGTT | geneM0992-R |
| **Complementation** | | |
| BCAM0992 complementation F | AATAATAATCATATGCGTGACCTGATCGAAGAG | M0992-ov-NdeI |
| BCAM0992 complementation R | TTTGGATCCCATACGATGTATGCGTTGCGTTC | M0992-ov-BamHI |
| | **pGPI-SceI-XCm MCS primers** | |
| MCS plasmid sequence insert control F | AACAAGCCAGGGATGTAACG | MCS-B-F |
| MCS plasmid sequence insert control R | TGTTCGGCCAGATAGAAACC | MCS-B-R |

659

660 **TABLE S7** Primers used in qPCR experiments.

| Primers | Sequence (5' 3') | Primers | Sequence (5' 3') |
|---|---|---|---|
| **M0918-F** | GAGATGAGCACCGATCACAC | **L0024-F** | TACAGGCGTGATCGAAGGTG |
| **M0918-R** | CCTTCGAGGAACGACTTCAG | **L0024-R** | GGAAGATCTGGTGCGATTCC |
| **L0003-F** | AATGGCCTGAATTCCTGACG | **L0509-F** | GGTGATGGTCAACACGTTCG |
| **L0003-R** | GTGATGCACGGTCTTCTTCG | **L0509-R** | CCGTAAGCTGCCGTCTTCTC |
| **L0423-F** | AGCTGGACTGGGTCAAGAGC | **L2465-F** | GGCTGTCTGATCGTGCTGTC |
| **L0423-R** | GGATCGAGGACGAACTGGAC | **L2465-R** | ATGCCCTGTTGAACCGTCAC |
| **M0820-F** | ACGTCTACCGGACCGAACAC | **L2767-F** | AGACCTATCACCCGCTGCAC |
| **M0820-R** | TCGAGCACGATTTCGTTGAG | **L2767-R** | ACGGGTGGTATTCGTTCGTC |

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

| | | | |
|---|---|---|---|
| **L1059-F** | GATGCTGACGACGAACGAAC | **L2782-F** | AGCCGAACACGATGACACTC |
| **L1059-R** | GTCCTTGAAGATGCCGAAGC | **L2782-R** | CTTGCGGCTTTCGTAATTGG |
| **L1457-F** | CAGCAGATGAATTCGACCAC | **L3303-F** | GACGAGACGCGCTACCAGAC |
| **L1457-R** | TCGACGTAAGCGAGGATCTG | **L3303-R** | GGTCGTACCACTCGCTGTGC |
| **S0223-F** | ATGCTCGTGTCGTTTCATGC | **L0054-F** | AGCGCACCGATTCGAACTAC |
| **S0223-R** | ACTGGTCGCCGTAGTCGAAG | **L0054-R** | ACGTGTCCGATGTGATCGTC |
| **M2738-F** | GCTGAGCGAACAGGTTGACG | **L0162-F** | GCATCCACGAAGTCCATCTG |
| **M2738-R** | ACCATTCGGCCTTCACTTCC | **L0162-R** | AATCCTCACCCAGCAGCATC |
| **M1262-F** | ACCGCGAACTCGATGAACTG | **L1556-F** | ACTGCTTCATCGACGCACTC |
| **M1262-R** | GGTGCAGGATCGTGTTGGTC | **L1556-R** | CGTTCAGGTCGAACACCTTG |
| **L0079-F** | GTCAACCAGCTCACCGTCTG | **L2415-F** | GTGAAGCCCGTGATGTCGTC |
| **L0079-R** | CTTCCACAGCGAGATGATGC | **L2415-R** | GACCGGCTCGCAGAAGTAGG |
| **L0126-F** | TGATGGCGCTTCTTTACGTG | **S0258-F** | AAGATGCGGGAACTGATCGAC |
| **L0126-R** | ATTCGACGATGTGGTGATCG | **S0258-R** | ATGAAACACCCAGCCGATACG |
| **L0508-F** | TCGTCTGAGGGTGTTCAAGC | **M0941-F** | ACAAGCAATCGGTGTGATCG |
| **L0508-R** | ATCAGCGGAATCTGCTCCTC | **M0941-R** | AGCGTATAGGTCGGCACCAG |
| **L0709-F** | GCGGCGTATAATCTCGCTTC | **M1415-F** | AGACGACAACGCGAAACTCG |
| **L0709-R** | ATGTCGACGGTTTCCAGTCC | **M1415-R** | ATCAGGTACGACGGCGACAG |
| **L2701-F** | TCACGTTCGACCACAGCTTC | **M0076-F** | TGCCGCCTTTGTACTCATGG |
| **L2701-R** | GACGCGATGTTGTTCAGCTC | **M0076-R** | GCGACACGGAAATGATCTCG |
| **L2942-F** | AAGCCTACATGCCGACCATC | **M1362-F** | GATCGTGGTCGTCGTGTTCC |
| **L2942-R** | GATCGCAGACGATGAACACG | **M1362-R** | GTCTTGTCGTTGCCGAGACG |
| **L2406-F** | TGCCGAGATTGCTGTTCAAG | **L1515-F** | CGCAAGCAACCTGTACTTCG |
| **L2406-R** | AGCAACGGTGTCAGCAACAG | **L1515-R** | GTCAGGCGATTCAGGATGTG |

661

662 **TABLE S8** List of primers used for construction of translational eGFP reporter fusion
663 plasmids.

| Gene | Primers | Sequence (5' 3') |
|---|---|---|
| **BCAL0079** | F-eGFP-L0079-BamHI | ATATGGATCCTGCGTATTGTGTCCGATCA |
| | R-eGFP-L0079-EcoRI | ATATGAATTCCATGATGGCGGATGGTGTT |
| **BCAL1515** | F-eGFP-L1515-BamHI | ATATGGATCCGGTGCTTTCAGGCACATTTC |
| | R-eGFP-L1515-EcoRI | ATATGAATTCGCCGAACAGATAGGAGTTCAG |
| **BCAM0820** | F-eGFP-M0820-BamHI | ATATGGATCCCTGCCGATTCGGAGTATCTG |
| | R-eGFP-M0820-EcoRI | ATATGAATTCATCCGAGGCATTATCACTGCT |
| **Plasmid insert** | F-pJH2 | CGTAGAGGATCTGCTCATGTTTGAC |
| | R-pJH2 | GACGTAAACGGCCACAAGTTCA |

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

664

665 ## References

666 1. Mahenthiralingam E, Urban TA, Goldberg JB. 2005. The multifarious,
667 multireplicon *Burkholderia cepacia* complex. Nat Rev Microbiol 3:144–56.
668 2. Coenye T. 2010. Social interactions in the *Burkholderia cepacia* complex: biofilm
669 formation and quorum sensing. Future Microbiol 5:1087–99.
670 3. Mahenthiralingam E, Vandamme P. 2005. Taxonomy and pathogenesis of the
671 *Burkholderia cepacia* complex. Chron Respir Dis 2:209–17.
672 4. Lipuma JJ. 2010. The Changing Microbial Epidemiology in Cystic Fibrosis. Clin
673 Microbiol Rev 23:299–323.
674 5. Sousa SA, Feliciano JR, Pita T, Guerreiro SI, Leitão JH. 2017. *Burkholderia*
675 *cepacia* Complex Regulation of Virulence Gene Expression : A Review. Genes
676 19:8(1).
677 6. Drevinek P, Mahenthiralingam E. 2010. *Burkholderia cenocepacia* in cystic
678 fibrosis : epidemiology and molecular mechanisms of virulence. Clin Microbiol
679 Infect 16(7):821–30.
680 7. Coutinho CP. 2011. Long-term colonization of the cystic fibrosis lung by
681 *Burkholderia cepacia* complex bacteria: epidemiology, clonal variation, and
682 genome-wide expression alterations. Front Cell Infect Microbiol 1(12):1–11.
683 8. Courtney JM, Dunbar KEA, Mcdowell A, Moore JE, Warke TJ. 2004. Clinical
684 outcome of *Burkholderia cepacia* complex infection in cystic fibrosis adults. J Cyst
685 Fibros 3:93–8.
686 9. Hindo H, Sigley C, Karlson K. 2008. Case Report Cepacia Syndrome in an

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

687 Adolescent With Cystic Fibrosis. Infectious Diseases in Clinical Practice
688 16(3):198–200.
689 10. Holden MTG, Seth-smith HMB, Crossman LC, Sebaihia M, Bentley SD, Cerden
690 AM, et al. 2009. The Genome of *Burkholderia cenocepacia* J2315 , an Epidemic
691 Pathogen of Cystic Fibrosis Patients. J Bacteriol 191(1):261–77.
692 11. Coenye T, LiPuma JJ. 2003. Population structure analysis of *Burkholderia*
693 *cepacia* genomovar III: Varying degrees of genetic recombination characterize
694 major clonal complexes. Microbiology 149(1):77–88.
695 12. Vandamme P, Holmes B, Coenye T, Goris J, Mahenthiralingam E, LiPuma JJ,
696 Govan JR. 2003. *Burkholderia cenocepacia* sp. nov. - A new twist to an old story.
697 Res Microbiol 154(2):91–6.
698 13. Weinhold B. 2006. Epigenetics: the science of change. Environ Heal Perspect
699 114(3):160–7.
700 14. Sánchez-Romero MA, Cota I, Casadesús J. 2015. DNA methylation in bacteria :
701 from the methyl group to the methylome. Curr Opin Microbiol 25:9–16.
702 15. Sánchez-Romero MA, Casadesús J. 2020. The bacterial epigenome. Nat Rev
703 Microbiol 18(1):7-20.
704 16. Blow MJ, Clark TA, Daum CG, Deutschbauer AM, Fomenkov A, Fries R, et al.
705 2016. The Epigenomic Landscape of Prokaryotes. PLoS Genet 12(2):1–28.
706 17. Casadesús J, Low D. 2006. Epigenetic Gene Regulation in the Bacterial World.
707 Microbiol Mol Biol Rev 70(3):830–56.
708 18. Lacey L. Westphal, Peter Sauvey, Matthew M. Champion, Ian M. Ehrenreich,
709 Steven E. Finkel. 2016. Genomewide Dam Methylation in Stationary Phase.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

710 mSystems 1(6):1–13.

711 19. Attwood JT, Yung RL, Richardson BC. 2002. DNA methylation and the regulation
712 of gene transcription. Cell Mol Life Sci 59:241–57.

713 20. Low DA, Casadesús J. 2008. Clocks and switches: bacterial gene regulation by
714 DNA adenine methylation. Curr Opin Microbiol 11(2):106–12.

715 21. Adhikari S, Curtis PD. 2016. DNA methyltransferases and epigenetic regulation in
716 bacteria. FEMS Microbiol Rev 40(5):575-91.

717 22. Wion D, Casadesús J. 2006. N6-methyl-adenine: an epigenetic signal for DNA–
718 protein interactions. Nat Rev Microbiol 4(3):183–92.

719 23. Militello KT, Mandarano AH, Varechtchouk O, Simon RD. 2014. Cytosine DNA
720 methylation influences drug resistance in *Escherichia coli* through increased sugE
721 expression. FEMS Microbiol Lett 350:100–6.

722 24. Hermann A, Jeltsch A. 2003. Methylation Sensitivity of Restriction Enzymes
723 Interacting with GATC Sites. Biotechniques 34(5):30–3.

724 25. Flusberg BA, Webster D, Lee J, Travers K, Olivares E, Clark A, et al. 2010. Direct
725 detection of DNA methylation during single-molecule, real- time sequencing. Nat
726 Methods 7(6):461–5.

727 26. Fang G, Munera D, Friedman DI, Mandlik A, Chao MC, Banerjee O, et al. 2012.
728 Genome-wide mapping of methylated adenine residues in pathogenic *Escherichia*
729 *coli* using single-molecule real-time sequencing. Nat Biotechnol 30(12):1232–9.

730 27. Eid J. et al. 2009. Real-time DNA sequencing from single polymerase molecules.
731 Science 323:133–138.

732 28. Murphy J, Mahony J, Ainsworth S, Nauta A, Sinderen V. 2013. Bacteriophage

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

733 Orphan DNA Methyltransferases : Insights from Their Bacterial Origin , Function ,
734 and Occurrence. Appl Environ Microbiol 79(24):7547–55.
735 29. Yadav MK, Park S, Chae S, Song J. 2014. Sinefungin, a Natural Nucleoside
736 Analogue of S-Adenosylmethionine, Inhibits *Streptococcus pneumoniae* Biofilm
737 Growth. Biomed Res Int 2014:156987.
738 30. Du W, Dubarry N, Passot FM, Kamgoué A, Murray H. 2016. Orderly Replication
739 and Segregation of the Four Replicons of *Burkholderia cenocepacia*. PLoS Genet
740 12(7):1–27.
741 31. Roberts RJ, Vincze T, Posfai J, Macelis D. 2005. REBASE — restriction enzymes
742 and DNA methyltransferases. Nucleic Acids Res 33:230–2.
743 32. McCammon MT, Parks LW. 1981. Inhibition of Sterol Transmethylation by S-
744 adenosylhomocysteine analogs. J Bacteriol 145(1):106–12.
745 33. Aya R, Sarnacki SH, Llana MN, Gabriela A, Guerra L, Giacomodonato MN, et al.
746 2015. Dam methylation is required for efficient biofilm production in *Salmonella*
747 *enterica* serovar Enteritidis. Int J Food Microbiol 193:15–22.
748 34. Ouellette M, Gogarten JP. 2018. Characterizing the DNA Methyltransferases of
749 *Haloferax volcanii* via Bioinformatics, Gene Deletion, and SMRT Sequencing.
750 Genes 9(3).
751 35. Casselli T, Tourand Y, Scheidegger A, Arnold WK, Proulx A, Stevenson B, et al.
752 2018. DNA methylation by restriction modification systems affects the global
753 transcriptome profile in *Borrelia burgdorferi*. J Bacteriol 200(24):1–17.
754 36. Nandi T, Holden MTG, Didelot X, Mehershahi K, Boddey JA, Beacham I, et al.
755 2015. *Burkholderia pseudomallei* sequencing identifies genomic clades with

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

756 distinct recombination, accessory, and epigenetic profiles. Genome Res
757 25(1):129–41.
758 37. Erill I, Puigvert M, Legrand L, Guarischi-Sousa R, Vandecasteele C, Setubal JC,
759 et al. 2017. Comparative analysis of *Ralstonia solanacearum* methylomes. Front
760 Plant Sci 13;8:504.
761 38. Suzuki MM, Bird A. 2008. DNA methylation landscapes: provocative insights from
762 epigenomics. Nat Rev Genetics 9(6):465–76.
763 39. Tate PH, Bird P. 1993. Effects of DNA methylation on proteins and gene
764 expression. Curr Opin Genet Dev. 3(2):226-31.
765 40. Zachary F. Burton. 2018. Evolution Since Coding: Cradles, Halos, Barrels, and
766 Wings. General Transcription Factors and Promoters 27:115–22.
767 41. Hickman JW, Tifrea DF, Harwood CS. 2005. A chemosensory system that
768 regulates biofilm formation through modulation of cyclic diguanylate levels. Proc
769 Natl Acad Sci USA 102(40).
770 42. Cooper VS, Staples RK, Traverse CC, Ellis CN. 2014. Parallel evolution of small
771 colony variants in *Burkholderia cenocepacia* biofilms. Genomics 104(6, Part
772 A):447-52.
773 43. Traverse CC, Mayo-Smith LM, Poltak SR, Cooper VS. 2013. Tangled bank of
774 experimentally evolved *Burkholderia* biofilms reflects selection during chronic
775 infections. Proceedings of the National Academy of Sciences 110(3):E250–E9.
776 44. Girgis HS, Liu Y, Ryu WS, Tavazoie S. 2007. A Comprehensive Genetic
777 Characterization of Bacterial Motility. PLoS Genet 3(9):1644-60.
778 45. Balleza E, López-Bojorquez LN, Martínez-Antonio A, Resendis-Antonio O,

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

779 Lozada-Chávez I, Balderas-Martínez YI, et al. 2009. Regulation by transcription
780 factors in bacteria: Beyond description. FEMS Microbiol Rev 33(1):133–51.
781 46. Héberlé É, Bardet AF. 2019. Sensitivity of transcription factors to DNA
782 methylation. Essays Biochem 63(6):727–41.
783 47. Lenhart JS, Pillon MC, Guarné A, Biteen JS, Simmons LA. 2016. Mismatch repair
784 in Gram-positive bacteria. Res Microbiol 167(1):4–12.
785 48. Roberts, R. J., Vincze, T., Posfai, J., & Macelis, D. 2015. REBASE-a database for
786 DNA restriction and modification: enzymes, genes and genomes. Nucleic acids
787 research. 43(Database issue), D298–D299.
788 49. Carver T, Harris SR, Berriman M, Parkhill J, Mcquillan JA. 2012. Artemis : an
789 integrated platform for visualization and analysis of high-throughput sequence-
790 based experimental data. Bioinformatics 28(4):464–9.
791 50. Altschup SF, Gish W, Pennsylvania T, Park U. 1990. Basic Local Alignment
792 Search Tool. J Mol Biol 215(3):403-10.
793 51. Hamad MA, Skeldon AM, Valvano MA. 2010. Construction of Aminoglycoside-
794 Sensitive *Burkholderia cenocepacia* Strains for Use in Studies of Intracellular
795 Bacteria with the Gentamicin Protection Assay. Appl Environ Microbiol 76:3170–
796 3176
797 52. Flannagan RS, Linn T, Valvano MA. 2008. A system for the construction of
798 targeted unmarked gene deletions in the genus *Burkholderia*. Environ Microbiol
799 10(6):1652–60.
800 53. Van Acker H, Ostyn L, Sass A, Daled S, Dhaenens M, Deforce D, et al. 2019. The
801 role of small proteins in *Burkholderia cenocepacia* J2315 biofilm formation,

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.21.960518; this version posted February 25, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

802 persistence and intracellular growth. Biofilm. 2019;1:100001.

803 54. Cardona ST, Valvano MA. 2005. An expression vector containing a rhamnose-
804 inducible promoter provides tightly regulated gene expression in *Burkholderia*
805 *cenocepacia*. Plasmid 54(3):219–28.

806 55. Drago L, Agrappi S, Bortolin M, Toscano M, Romanò CL, Vecchi E De. 2016.
807 How to Study Biofilms after Microbial Colonization of Materials Used in
808 Orthopaedic Implants. Int J Mol Sci.17(3): 293.

809 56. Kong EF, Tsui C, Kucharíková S, Andes D. 2016. Commensal Protection of
810 *Staphylococcus aureus* against Antimicrobials by *Candida albicans* Biofilm Matrix.
811 7(5):1–12.

812 57. Kobayashi K. 2007. *Bacillus subtilis* Pellicle Formation Proceeds through
813 Genetically Defined Morphological Changes. J Bacteriol 189:4920–31.

814 58. Sass AM, Van Acker H, Förstner KU, Van Nieuwerburgh F, Deforce D, Vogel J, et
815 al. 2015. Genome-wide transcription start site profiling in biofilm-grown
816 *Burkholderia cenocepacia* J2315. BMC Genomics 16:775.

817 59. Münch R, Hiller K, Grote A, Scheer M, Klein J, Schobert M, Jahn D. 2005. Virtual
818 Footprint and PRODORIC: An integrative framework for regulon prediction in
819 prokaryotes. Bioinformatics 21(22):4187–9.

820