Differential gene expression and alternation of patterns of DNA methylation in the multidrug resistant strain *Escherichia coli* ATCC BAA-196 caused by iodine-containing nano-micelle drug FS-1 that induces antibiotic resistance reversion

Ilya S. Korotetskiy[1*], Ardak B. Jumagaziyeva[2], Sergey V. Shilov[1], Tatyana V. Kuznetsova[1], Natalya A. Suldina[2], Sabina T. Kenesheva[2,4], Aleksandr I. Ilin[3], Monique Joubert[5], Setshaba Taukobong[5], Oleg N. Reva[5]

[1]Department of Virology, Scientific Center for Anti-Infectious Drugs, Almaty, Kazakhstan

[2]Department of Microbiology, Scientific Center for Anti-Infectious Drugs, Almaty, Kazakhstan

[3]Department of Chemistry, Scientific Center for Anti-Infectious Drugs, Almaty, Kazakhstan

[4]Faculty of Biology and Biotechnology, al-Farabi Kazakh National University, Almaty, Kazakhstan

[5]Centre for Bioinformatics and Computational Biology; Department of Biochemistry, Genetics and Microbiology; University of Pretoria, Pretoria, South Africa

*Corresponding author

E-mail: laeda81@gmail.com (Ilya S. Korotetskiy).

Abstract

Recent studies have demonstrated that antibiotic resistance in pathogenic bacteria could be reverted to the drug sensitive phenotype by the treatment with the newly introduced iodine containing nanomolecular complex, FS-1. Antibiotic resistance reversion has been verified as a promising therapeutic approach to combat multidrug resistant infections. The mechanisms of action, however, remain unclear. A collection strain, *Escherichia coli* ATCC BAA-196, showing an extended spectrum of resistance to beta-lactam and aminoglycoside antibiotics was used in this study as a model organism. The FS-1 treated culture and the negative control variant were sequenced by PacBio RS II System following the SMRTbell 20-kb library preparation protocol. Total RNA samples of these strains were sequenced by Ion Torrent. It was shown that the treatment with FS-1 caused a profound gene expression alternation switching the bacterial metabolism to anaerobic respiration, increased anabolism, oxidative/acidic stress response and an inhibition of many nutrient uptake systems. All this leads to an increase in the susceptibility to antibiotics even when FS-1 is removed from the medium. The later fact implies an involvement of epigenetic mechanisms in gene regulation and antibiotic resistance reversion. This hypothesis was investigated by base-call kinetic analysis in PacBio reads and DNA methylation profiling in the sequenced genomes. Several DNA motifs of adenosine and cytosine methylation were identified. While the numbers of methylated sites in chromosomes and plasmids of both genomes, NC and FS, were similar, the distribution of the methylated sites was different. It may explain the observed long lasting effect of the treatment of *E. coli* with FS-1 on antibiotic susceptibility of this model microorganism.

Author summary

The emergence of multidrug resistant bacteria is a great concern, since the misuse of antibiotics have caused a strong selective pressure for these resistant bacteria and various treatment options are becoming ineffective. Drug induced reversion of antibiotic resistant is considering a promising approach to address this problem. This study was set out to investigate genetic mechanisms of action of a new iodine-containing nano-micelle drug FS-1 that induces antibiotic resistance reversion in bacteria. *Escherichia coli* ATCC BAA-196 was used as a model of multidrug resistant microorganisms. With this purpose in mind, we have sequenced the genomic DNA and total RNA samples of *E. coli* cultivated on medium containing FS-1, served as an experimental (FS) variant, and bacteria cultivated on normal medium, served as negative control (NC). RNA sequencing showed a differential gene expression in the FS-1 treated strain that may explain the observed increase in susceptibility to gentamicin and ampicillin. Application of the 3rd generation sequencing technology, SMRT PacBio, allowed an unambiguous whole genome assembly of NC and FS variants, and comparative DNA methylation profiling of bacterial chromosomes and plasmids. This study exemplifies antibiotic resistance reversion mechanisms acting through an epigenetically controlled alternation of gene expression in treated bacteria.

Introduction

Controlling antibiotic resistance development in bacteria is an intriguing area of research. It is generally recognized that pathogens may gain resistance to antibiotics either due to mutations modifying the target proteins or by acquisition of antibiotic resistance genes. The latter includes efflux pumps, alternative variants of target proteins, or antibiotic modifying enzymes [1]. There are various approaches to combat antibiotic resistance, these approaches include: new antibiotic

63 design or drug repurposing to target alternative key molecular processes [2]; drug combination
64 therapy, which uses several antibiotics with alternative mechanisms of action [3]; efflux pump
65 inhibitors, antibiotic depleting enzymes, other mechanisms of persistence and adaptation [4]; and
66 drug-induced reversion of the selective pressure in bacterial populations towards antibiotic
67 sensitivity by imposing a direct cost on the resistance. The later approach was discussed by
68 Baym et al. [5], were antibiotic resistance reversion was explained by several possible
69 mechanisms. Firstly, co-administrated drugs may have suppressive interactions, which imply that
70 bacteria getting resistance to drug A can become more sensitive to drug B due to competitiveness
71 of the drug resistance mechanisms. This causes a shift in the bacterial population towards a
72 higher sensitivity of bacteria to drug B when drug A is present. Secondly, co-administrated drugs
73 may impose a direct fitness cost on resistance mechanisms creating a concentration regime
74 where sensitive bacteria can continue to grow, while the growth of resistant strains is inhibited.
75 Thirdly, resistance to supplementary drug A can decrease the minimal inhibitory concentration
76 (MIC) of drug B due to evolutionary trade-offs, allowing selection against the resistant strains
77 even when drug A was removed.

78 Clinical practice of the iodine-containing nano-molecular complex, FS-1, which induces
79 antibiotic resistance reversion *in vitro* and in clinical trials (www.clinicaltrials.gov, acc.
80 NCT02607449), has progressed the study of the antibiotic resistance reversion mechanisms
81 beyond the theoretical realm [6, 7]. FS-1 was primarily introduced as new therapeutics against
82 multidrug resistant tuberculosis [8, 9]. Further laboratory experiments demonstrated antibiotic
83 resistance reversion on several other multidrug-resistant bacteria including *Staphylococcus*
84 *aureus* [10] and *Escherichia coli* [11]. An implication of these results was that there should be
85 some level of commonality among the mechanisms of antibiotic resistance reversion induced by

86 FS-1 in these diverse bacteria. It was hypothesized that epigenetic modifications of DNA caused
87 by the treatment with iodine-containing nanoparticles may be behind these common
88 mechanisms. It should also be accepted that in many aspects, both *S. aureus* and *E. coli*
89 multidrug-resistant isolates are more convenient laboratory models to study antibiotic resistance
90 reversion rather than the extensively drug-resistant *M. tuberculosis* clinical isolates where this
91 phenomenon was reported first [6].

92 This work presents genetic, transcriptional and epigenetic modifications detected in the
93 population of the reference strain *E. coli* ATCC BAA-196. This strain was selected in 1988 in
94 Boston from a clinical isolate showing an extended spectrum of resistance to beta-lactam
95 antibiotics and first identified as *Klebsiella pneumonia* [12]. The selected strain represents
96 pathogenic *E. coli*, which are common agents of nosocomial infections [13-15]. Outbreaks of
97 foodborne pathogenic *E. coli* are being reported around the world with a growing frequency [16,
98 17]. In addition to the increased virulence of enteropathogenic (EPEC), endometrial pathogenic
99 (EnPEC), Shiga toxin-producing/enterohemorrhagic (STEC/EHEC), enteroinvasive (EIEC),
100 enteroaggregative (EAEC), and enterotoxigenic (ETEC) *E. coli*, these isolates often show an
101 extended spectrum of antibiotic resistance [18]. It is alarming that the drug resistant isolates of
102 *E. coli* are also frequent in environmental habitats including water resources [19] and agricultural
103 products [20]. Complete genome sequencing of the collection strain *E. coli* ATCC BAA-196
104 (CP042865) showed clustering of this microorganism with three previously sequenced variants
105 of *E. coli* K-12, NC_000913, W3110 (NC_007779) and DH10B (NC_010473), which are
106 representatives of the normal avirulent human gut microflora [21-23]. The only difference
107 between the multidrug resistant strain BAA-196 and the avirulent *E. coli* K-12 isolates is the
108 presence of a large plasmid (CP042866) comprising multiple drug resistance determinants. This

109 plasmid shows a significant sequence similarity to the virulence plasmid pKP64477b from
110 *Klebsiella pneumonia* [24] and some other similar plasmids frequent in different species of
111 Enterobacteria. This makes *E. coli* ATCC BAA-196 an ideal model organism to study antibiotic
112 resistance evolution and antibiotic resistance reversion mechanisms, as the K-12 strains are
113 among the best-studied bacteria in the world in terms of biology, metabolism, and gene
114 regulation [25, 26]. Moreover, the strain BAA-196 exemplifies multidrug resistance
115 development associated with an acquisition of a conjugative plasmid that contrasts it against the
116 previously sequenced antibiotic resistance reversion model organisms: *M. tuberculosis* SCAID
117 187.0 [9] and *S. aureus* ATCC BAA-39 [10]. In this paper, the results of induced antibiotic
118 resistance reversion in *E. coli* ATCC BAA-196 treated with FS-1 are reported. Ion Torrent
119 RNASeq and PacBio complete genome sequencing of the FS-1 treated culture are presented and
120 discussed.

121 # Results

122 ## Antibiotic resistance reversion

123 This study was performed to evaluate the effect of drug induced resistance reversion of the
124 multidrug resistant strain, *E. coli* ATCC BAA-196, which is resistant to the beta-lactam
125 antibiotic, ampicillin, and the aminoglycoside antibiotic, gentamicin. The experimental culture
126 was cultivated for 10 passages with the sub-bactericidal concentration of FS-1 (500 μg/ml). In
127 parallel, a negative control (NC) culture was cultivated for 10 passages in liquid MH medium
128 without FS-1 and antibiotics.
129 Minimal bactericidal concentrations (MBC) of the antibiotics were determined for the
130 experimental FS-1 treated (FS) and NC cultures by serial dilution method on 96-well plates [27].

131 Cultivation with FS-1 reduced MBC for ampicillin from 2 – 2.6 mg/ml recorded for the culture
132 NC to 1.6 – 2 mg/ml recorded for the culture FS. For gentamicin, these numbers were 0.3 – 0.35
133 mg/ml and 0.2 – 0.25 mg/ml for NC and FS cultures, respectively. This difference in
134 susceptibility to antibiotics between the initial and FS-1 treated cultures was statistically reliable
135 as illustrated by OD records for the plates with ampicillin and gentamicin in Fig 1.

136 **Complexation of FS-1 with bacterial DNA**

137 Epigenetic changes in bacterial chromosomes under the effect of the iodine-containing nano-
138 molecular complex FS-1 suggested a possible complexation of the DNA with nanoparticles
139 and/or with released iodine, which potentially can halogenate chromosomal nucleotides [28]. To
140 study the ability of FS-1 to penetrate bacterial cell walls and create complexes with the
141 chromosomal DNA, the drug was synthesized with the radioactive isotope ^{131}I (20 MBq/ml).
142 After 1 h cultivation with the radioactive labeled FS-1, bacterial cells were washed twice to
143 remove the remaining micelles of FS-1 and the residual radioactivity was measured. Then the
144 residual radioactivity was measured in DNA samples extracted from the washed cells. It was
145 found that radioactivity of the extracted DNA was 43.46 ± 13.895 Bq/ng that constituted 0.46%
146 ± 0.15 of the residual radioactivity of treated bacterial cells after washing. The residual
147 radioactivity can be associated with either complexing of FS-1 with the DNA or a direct
148 halogenation of nucleotides by iodine isotopes.

149 **Complete genome assembly and annotation**

150 Complete sequences of the bacterial chromosome of the negative control (NC) and FS-1 treated
151 (FS) variants of *E. coli* ATCC BAA-196 and several plasmids were obtained by assembly of
152 PacBio SMRT reads. Chromosomal sequences of the variants NC and FS were 4,682,561 bp and
153 4,682,572 bp, respectively. Both genomes comprised of large plasmids 266,396 bp of NC and

154 279,992 bp of FS variants. The FS plasmid is larger due to an insertion of a genomic island in the
155 region [232952..258265] flanked with two copies of *insH* transposases comprising of 23 phage
156 and plasmid related genes (Fig 2A). Nine genes including a part of the plasmid conjugation
157 operon with *traG* and *traF* genes were lost in the plasmid FS compared to the plasmid NC. This
158 deletion potentially can affect the efficacy of conjugation of the plasmid FS.

159 The identified plasmids share 90-99% sequence similarity with several known pathogenicity and
160 drug resistance plasmids from *Klebsiella pneumonia* (pKP64477b), *Serratia marcescens*
161 (AP014611.1) and *E. coli* (p3521). Four broad range beta-lactamases of TEM, OXA and AmpC
162 families; five aminoglycoside inactivating enzymes of APH, ANT and AAC families; *catI*
163 chloramphenicol acetyltransferase; trimethoprim resistant dihydrofolate reductase of the DFR
164 family and heavy metal resistance genes *merA* and *terY* tellurium resistance operon were
165 identified on the plasmids NC and FS. Several other genes may additionally contribute to
166 antibiotic resistance, which include 23 efflux pump proteins of various families and three peptide
167 antibiotic resistance genes *bacA*, *pmrF* and *ugd* as predicted by RGI 5.1.0/CARD 3.0.5 Web
168 server.

169 Pathway Tools predictions showed that several genes located on the large plasmid, sucrose-6F-
170 phosphate phosphohydrolase, adenine/guanine phosphoribosyltransferase, carbamoylphosphate
171 synthase large subunit, starvation-induced cysteine protease and citrate lyase beta subunit, may
172 interfere with several central metabolic pathways.

173 Two additional smaller plasmids of 44,240 bp and 11,153 bp were assembled from PacBio reads
174 generated from *E. coli* ATCC BAA-196, variant FS. Alignment of the small plasmids against the
175 large ones showed that they were generated by fragmentation and rearrangement of the large

176 plasmid (Fig 2B-D). Middle size plasmid is constituted of two blocks linked by an integrase
177 *retA*. This plasmid contains conjugation genes *traDGFIHJ*, *trhH* and some other genes
178 associated with plasmid replication and gene integration. The smallest plasmid is a transposable
179 element with a cluster of antibiotic resistant genes aminoglycoside 3"-nucleotidyltransferase
180 *aadA*, small multidrug resistance (SMR) efflux transporter *qacE*, sulfonamide-resistant
181 dihydropteroate synthase *sull* and a GNAT family N-acetyltransferase, which are enclosed by
182 several copies of transposases and integrases. Plasmid rearrangements and fragmentation
183 observed in the genome of the variant FS suggests that cultivation of the strain with sub-
184 bactericidal concentrations of FS-1 may destabilize plasmids by activation of multiple
185 transposases and integrases found in the large plasmid genome.

186 **Differential gene regulation in *E. coli* cultivated with FS-1**

187 Gene expression regulation in the cultivar of *E. coli* ATCC BAA-196 that survived 10 passages
188 of cultivation with 500 μg/ml FS-1 in the medium (denoted as FS) was compared to the negative
189 control (NC) variant cultivated for 10 passages in the same medium without the drug. The
190 variant FS was characterized with the increased susceptibility to gentamicin and ampicillin
191 (Fig 1) that may result from an alternated gene regulation.

192 A volcano plot of regulated genes is shown in Fig 3. The differentially expressed genes are
193 plotted in the central part of Fig 3 according to the fold change difference of their expression in
194 the medium with FS-1 compared to NC. Panels on the right part of Fig 3 denoted as −Inf and Inf
195 present respectively the genes, which are expressed only in the NC medium or in the medium
196 supplemented with FS-1. Differentially expressed genes of the plasmid origin are highlighted in
197 red.

198 Analysis of the regulated genes showed that up- and down-regulation of these genes is strongly
199 associated with three conditions: strict anaerobic growth, low pH and the oxidative stress.
200 PheNetic network of transcriptional regulation of differentially expressed genes is shown in
201 Fig 4A. The genes are grouped in regulons controlled by the following transcriptional regulators:
202 RpoS, Crp, ArcA, RpoH, RpoN, GcvB, InfAB, Fnr and several others. RpoS is the master
203 regulator of the general stress response in *E. coli*. The expression of this sigma-factor and
204 controlled genes was strongly down-regulated, while the genes suppressed by *rpoS* were up-
205 regulated. RpoS related stressosome is expressed in response to sudden changes in the
206 environment to allow bacterial cells to adapt to this change [29]. Adaptation of *E. coli* to the
207 permanent presence of FS-1 in the medium involved a significant alternation of the central
208 metabolism of the bacterium (Fig 4B). First, and what was unexpected, the metabolic processes
209 switched completely to anaerobiosis using formate and glycerol-3-phosphate as electron donors
210 and nitrates and possibly fumarate as the terminal electron acceptor. The later reactions are
211 controlled by HydN electron carrier protein, Sn-glycerol-3-phosphate anaerobic dehydrogenase
212 complex GlpABC, selenocysteine containing formate dehydrogenase-H FdhF and formate
213 hydrogenlyase complex HycBCDEFG, which show no expression at the negative control
214 condition but are highly expressed with FS-1. All systems associated with the aerobic lifestyle,
215 for example the aerobic formate dehydrogenase complex FdoGHI-FdhE, are strongly inhibited
216 directly by ArcA repressor or by associated regulatory pathways (Fig 4A). Transcription from all
217 genes of cytochrome bo terminal oxidase complex, *cyoA*, *cyoB*, *cyoC* and *cyoE*, was halted
218 completely. It may be assumed that iodine released from FS-1 either damaged cytochrome
219 molecules in cell membrane, or the oxidative phosphorylation pathway was restrained to
220 decrease the oxidation stress caused by the halogen. A previous publication has shown that the

221 model organism, *Staphylococcus aureus*, experienced an oxidative stress after being exposed to
222 FS-1 for 5 minutes [30]. Many systems preventing cell damage by peroxides, which are
223 byproducts of the oxidative phosphorylation, were down-regulated in the culture cultivated with
224 FS-1, which include Tpx thiol peroxidase, OsmC peroxiredoxin, glutathione reductase Gor,
225 YodD, YchH and YgaM proteins. Other systems probably helping the cells to cope with halogen
226 oxidation were activated: periplasmic disulfide oxidoreductase DsbC and membrane bound
227 superoxide:ubiquinone oxidoreductase CybB. The activation of selenocysteine containing
228 formate dehydrogenase-H FdhF required an increased uptake of selenium ions from environment
229 and their intracellular activation by selenium-binding protein YdfZ, tRNA 2-selenouridine
230 synthase YbbB and SelB selenocysteinyl-tRNA-specific translation factor, which are
231 significantly activated in FS-1 treated *E. coli*. It suggests that the FS-1 treated pathogen may be
232 sensitive to the shortage of selenium.

233 All genes of the tricarboxylic acid cycle (TCA) are strongly inhibited in the FS-1 treated *E. coli*
234 except for MaeA/SfcA malate dehydrogenase that initiates the anabolic gluconeogenic pathway
235 starting from oxaloacetate. The key enzymes of the gluconeogenic pathway,
236 phosphoenolpyruvate carboxykinase PckA, phosphoglyceromutase GpmL, glyceraldehyde-3-
237 phosphate dehydrogenase GapA, triosephosphate isomerase TpiA, fructose 1,6-bisphosphatase
238 GlpX and glucose-6-phosphate isomerase Pgi, are significantly upregulated. Contrary, the
239 catabolic glycolysis pathway is halted by a strong downregulation of the genes 6-
240 phosphofructokinase *pfkB* and fructose-bisphosphate aldolase *fbaB*. Carbon sources of the
241 gluconeogenic anabolism are C4-dicarboxylates (fumarate and oxaloacetate) and polyhydric
242 alcohols. Uptake and intracellular production of these compounds are activated. In contrast to
243 this, all pathways of acetate uptake and metabolism through the glyoxylate pathway are

244 completely blocked including the intracellular production of acetyl-CoA through both aerobic
245 and anaerobic fatty acid β-oxidation pathways. Cell wall fatty acids can be targets for oxidation
246 by iodine that cause the cell to downregulate the internal fatty acid oxidation systems by
247 blocking the acetate metabolism. Accumulation of acetate and other organic acids may lead to an
248 acidic stress that is indicated by upregulation of acidic stress response genes: *yhjX*, *ymgD*, *yfiD*
249 and *adiABC*. Arginine decarboxylase AdiABC and lactate dehydrogenase LdhA are activated in
250 anaerobic conditions in response to an extremely acidic environment associated with the
251 accumulation of byproducts produced during carbohydrate fermentation [31, 32]. Particularly,
252 trehalose uptake by TreABC complex associated with aerobic or microaerophilic lifestyle [33]
253 was activated. Due to blocking of glycolysis, which is used FS-1 treated *E. coli* for
254 gluconeogenesis, carbohydrate catabolism goes through the pentose phosphate pathway that is
255 confirmed by an activation of 6-phosphogluconate dehydrogenase Gnd starting the oxidative
256 branch of the pentose phosphate pathway. The oxidative branch of the pentose phosphate
257 pathway is the starting points of many biosynthetic pathways of nucleotides and amino acids.
258 Protein and fatty acid biosynthesis are activated probably by the need to replace cell membrane
259 compounds damaged by iodine oxidation and halogenation. Proteins PrmC and YjiA controlling
260 the accuracy of protein translation [34, 35], and the elongation factor P (*efp*) are activated, while
261 the ribosomal inhibitors *ygiU*/*mqsR* and *rmf* [36, 37] are strongly downregulated. At the same
262 time, the uptake and transportation of many amino acids and polypeptides is halted. Iodine in FS-
263 1 is bound in polypeptide and oligosaccharide micelles [6]. Transmembrane polypeptide and
264 amino acid transporters, such as low affinity tryptophan transporter TnaAB, oligopeptide
265 transporter OppABCD, dipeptide transporter DppA, proline transporter PutP, glutamate/aspartate
266 ABC transporter GltI and glutamine ABC transporter GlnH can be sites of entry of bound iodine

267 atoms that caused a strong inhibition of all these genes in FS-1 treated *E. coli*. It may be
268 hypothesized that the activated synthesis of proteins and nucleotides with complications
269 associated with the transportation of necessary building blocks overloads the intracellular
270 biosynthetic pathways. The damage of bacterial cell wall structures by FS-1 leading to the
271 distruction of protective barriers and increased penetrability of the cell wall and cell membrane
272 was reported before on experiments with *Mycobacterium tuberculosis* [38]. In the current study
273 the hypothesis of an increased penetrability of bacterial cells is supported by an observed
274 activation of heavy metal homeostasis efflux pumps *cueO* and *copA*.

275 Genes located on the plasmid are transcriptionally silent when the bacterium is growing on the
276 nutrient medium at the negative control condition. An exception is several transposases showing
277 a permanent low-level expression. Cultivation of the strain with FS-1 activated most of the
278 plasmid genes. Among the activated genes, there are four β-lactamases shown in Fig 4A.
279 However, the highest level of activation was recorded for the genes associated with plasmid
280 conjugation and transformation from cell to cell. The repressor of the cell-to-cell plasmid transfer
281 *ybaV* [39] located on the chromosome is strongly downregulated. Hyperactivation of
282 transposases and integrases located on the plasmid may explain the fact that two smaller
283 plasmids were found in the *E. coli* population grown on the medium with FS-1. Gene
284 comparison showed that these plasmids originated from the large plasmid by genetic
285 rearrangements (Fig 2).

286 Despite of the activation of the plasmid born β-lactamases, viability of FS-1 treated *E. coli*
287 decreased and susceptibility to antibiotics increased (Fig 1). It may be caused by downregulation
288 of many other genes associated with drug resistance: *yejG* [40], *uspF* [41] and *ygaM* [42]. Strong
289 inhibition of genes involved in biofilm formation was observed in the FS-1 treated culture that

290 potentially can reduce pathogenicity of the bacterium. Expression of quorum sensing transporter
291 *lsrABC* and quorum sensing regulators *lsrR* and *csrB* [43-45] was strongly downregulated.
292 Directly or indirectly, the inhibition of regulatory elements has affected many other genes
293 involved in bacterial motility, adhesion and biofilm formation, which include the AI-2 quorum
294 sensing signal procession protein YgiU/MqsR [37], AriR transcriptional regulator [46], inducer
295 of biofilm formation BolA [47], stress induced protein *uspF* [41], biofilm-stress-motility
296 lipoprotein BsmA [48] and the regulator of biofilm through signal secretion protein BssS [49].

297 It should be noted that the increased susceptibility of FS-1 treated culture of *E. coli* to gentamicin
298 and ampicillin (Fig 1) was observed after washing the bacterial cells from the drug FS-1 when
299 the direct antibacterial effect of the drug was removed. It may be supposed that the cultivation of
300 *E. coli* with FS-1 caused profound epistatic and epigenetic adaptive changes in the bacterial
301 population affecting its susceptibility to other antibiotics, which may persist in several
302 generations. To investigate this hypothesis, epigenetic modification in chromosomal DNA of FS-
303 1 treated *E. coli* were studied.

304 **Epigenetic modifications of chromosomal nucleotides**

305 Epigenetic modifications of nucleotides of chromosomal and plasmid sequences of NC and FS
306 variants of the strain *E. coli* ATCC BAA-196 were predicted by the kinetic analysis of base
307 calling during SMRT PacBio sequencing. The statistical reliability of epigenetic modifications
308 was calculated as Phred-like quality values termed QV scores. The value above 21 corresponds
309 to the confidence *p*-value 0.01 or smaller. PacBio sequencing of the control and FS-treated
310 bacterial samples was performed in tree repetitions followed by average score estimation for
311 every nucleotide site in the genome. Rank diagrams of distribution QV scores in NC and FS
312 genomes calculated for different nucleosides are shown in Fig 5. SMRT Link program predicted

313 that the peaks of high scored adenine residues in both genomes (QV > 80) pertained to an
314 intensive global methylation of the chromosomal DNA at these sites.

315 Methylation motifs and types of methylations were predicted by the program *motifMaker*.
316 Frequencies of methylation motifs in the sequenced genomes are shown in Table 1. Three
317 adenine methylation motifs were identified: G<u>A</u>*T*C, A<u>A</u>C(N6)G*T*GC and GC<u>A</u>C(N6)G*T*T. The
318 latter two motifs pertain to the same palindromic sequences. GATC and AAC(N6)GTGC /
319 GC<u>A</u>C(N6)G*T*T methylation is exhaustive in the sequenced genomes covering 100% of all
320 available sites; however, only 73-76% of the available motifs are methylated at both DNA
321 strands. G<u>A</u>*T*C sites were often associated with a more complex pattern of nucleotide
322 modifications, <u>C</u>RGKG<u>A</u>*T*C, where the first cytosine residue was also methylated. While this
323 type of cytosine methylation was not the most abundant, these sites of modified cytosines
324 showed the highest QV scores. Notably, while the total numbers of <u>C</u>RGKG<u>A</u>*T*C sites in NC and
325 FS genomes remain similar (respectively 105 and 106 out of 745 available CRGKGATC motifs
326 on both chromosomes), they were found at different loci on the chromosome (Fig 6A-B).

327 **Table 1. Methylation motifs predicted by the program *motifMaker*.**

Methylation motifs	***E. coli* BAA-196 NC**		***E. coli* BAA-196 FS**	
	Chromosome	**Plasmid**	**Chromosome**	**Plasmid**
G<u>A</u>*T*C*	28426 / 38442 (74%)†	1151 / 1854 (62%)	28408 / 38442 (74%)	1218 / 1950 (62%)
A<u>A</u>C(N6)G*T*GC	437 / 602 (73%)	14 / 20 (70%)	437 / 602 (73%)	16 / 21 (76%)
GC<u>A</u>C(N6)G*T*T	447 / 602 (74%)	13 / 20 (65%)	447 / 602 (74%)	14 / 21 (66%)
C<u>C</u>AGG	622 / 12113 (5%)	28 / 566 (5%)	651 / 12113	30 / 596 (5%)

			(5%)	
<u>C</u>RGKG<u>A</u>*T*C	105 / 745 (14%)	5 / 39 (13%)	106 / 745 (14%)	3 / 42 (7%)
W<u>C</u>CCTGGYR	109 / 407 (27%)	3 / 21 (14%)	103 / 407 (25%)	1 / 36 (3%)

[*]In the sequences of motifs, the methylated nucleotide is underlined and, in the case of a double strand methylation, the nucleotide opposing the methylated nucleotide on the reverse complement strand is shown in italic typeface;

[†]Numbers of methylated sites is displayed in the following format: *methylated_motif_namber / total_number_of_motif_sequences (fraction of methylated motifs)*.

Other predicted cytosine methylation motifs are C<u>C</u>AGGRAH and WCCCTGGYR. Methylation at C<u>C</u>AGG is typical for *E. coli* isolates; however, *motifMaker* predicted a more complex methylation motif C<u>C</u>AGGRAH. Only a fraction of methylated sites C<u>C</u>AGG was associated with 3'-sequence conservation; however, it may explain why this palindromic sequence was methylated almost always only at one DNA strand. Only 5% of available CCAGG sites were methylated on the chromosomes and plasmids. While the number was constant, different sites were methylated in NC and FS genomes. The distribution of methylated C<u>C</u>AGG motifs in plasmid sequences of two genomes is shown in Fig 6C-D. In sequences WCCCTGGYR, despite of the obvious palindromic nature of this motif, bipartite methylation at both DNA strands took place only in 10% of the methylated sites and was asymmetric: W<u>C</u>CCT*G*GYR. Interestingly, this type of methylation was more common for the chromosomal rather than for the plasmid DNA. Fractions of methylation of all other motifs were similar on the chromosomes and plasmids (Table 1).

347 An important notion of this study is that while the numbers of methylated sites in different motifs
348 remained similar in the NC and FS genomes, the locations of methylated nucleotides may be
349 different. On the chromosomes NC and FS, the total numbers of adenosine residues with QV
350 score > 21 including sporadic sites not associated with the recognized motifs were 43,361 and
351 43,760, respectively. Among them, 41,352 sites (95%) were the same loci in both genomes. The
352 numbers of modified adenosine residues with QV > 80 were 39,098 in NC and 39,083 in FS, and
353 39,045 (99%) were the same loci in both genomes. Only 53 high scored modified adenosine
354 residues were found in NC and 38 in FS. NC-specific adenosine methylation sites were found
355 within sequences of 16 genes, but only one of them was significantly downregulated – *yebV*
356 encoding an uncharacterized protein. Seven genes in the genome FS contained genome-specific
357 high QV scored adenine residues. One of these genes, nitrite reductase *nitB*, was significantly
358 upregulated.

359 There were 10,915 methylated cytosine residues in the genome NC and 11,220 in the genome
360 FS, 6001 methylated sites (54%) were shared by two genomes. Numbers of modified cytosine
361 residues with QV > 80 were 201 and 203 respectively, 119 sites (59%) were common for both
362 genomes. NC-specific modified cytosine residues were found in 33 genes. Two of these genes,
363 citrate-ACP transferase *citF* and DNA-binding transcriptional activator *gadW*, were significantly
364 upregulated. One CRGKGA*T*C site with NC-specific cytosine methylation was near the
365 promoter region of the NADP-dependent malic enzyme *maeB*, which is strongly downregulated
366 after the treatment with FS-1. FS-specific cytosine methylation was found in 39 genes with three
367 of them upregulated: biodegradative arginine decarboxylase *adiA*, DNA-directed RNA
368 polymerase *rpoC* and glucose-6-phosphate isomerase *pgi*; and two genes strongly
369 downregulated: bifunctional isocitrate dehydrogenase *aceK* and acyl-CoA dehydrogenase *fadE*.

370 Iodine may halogenate DNA nucleotides, mostly thymidine residues [28]. The number of
371 thymidine residues with QV > 21 increased insignificantly from 3199 sites in NC to 3306 in FS.
372 The analysis of expression of genes responsible for DNA reparation showed that the genes of the
373 SOS-response, *recA* and *lexA*, were downregulated. Strong upregulation was detected for the
374 gene *yjiA*, which is expressed in response to mitomycin C treatment causing DNA damage [50].
375 Genes *iraD*/*yjiD* inhibitor of σ^S activity in response to DNA damage [51] also were upregulated.
376 Genes of the DNA repair RecFOR complex were differentially regulated. RecF was strongly
377 upregulated while RecO and RecR were downregulated. RecFOR independent double-strand
378 break repair protein YegP was downregulated. It indicates that damaging of the chromosomal
379 DNA of *E. coli* most likely was not the major target of the drug FS-1 and the DNA of this
380 bacterium was protected from a direct iodination. The activated DNA repair genes may be
381 responsible for the DNA protection from halogens.

382 # Discussion

383 Cultivation of the multidrug-resistant strain *E. coli* ATCC BAA-196 with a sub-bactericidal
384 concentration of the iodine-containing nano-molecular complex FS-1 leads to several adaptive
385 changes in the phenotype leading to an increased sensitivity to antibiotics gentamicin and
386 ampicillin (Fig 1). The antibiotic sensitivity trials were performed on the media containing no
387 FS-1 that excludes any direct synergetic effect of the drugs and suggests some genetic, epistatic
388 and epigenetic changes in the bacterial population associated with the adaptation to the presence
389 of FS-1 in the medium. The reversion of the drug resistance to the antibiotic sensitive phenotype
390 may be considered as a promising therapeutic approach to combat multidrug resistant infections
391 [5, 6].

392 Complete genome sequencing of the negative control (NC) and FS-1 treated (FS) cultures did not
393 reveal any significant mutations in these genomes to explain the difference in the sensitivity to
394 antibiotics. The more sensitive FS strain contains all genetic determinants conferring the
395 resistance to aminoglycoside and beta-lactam antibiotics represented mostly by the plasmid
396 associated genes present also in the drug resistant variant NC. It may be hypothesized that the
397 increased antibiotic sensitivity may result from an alternative gene transcription regulation and
398 due to population phase changes controlled by epigenetic modifications and/or epistatic
399 arrangements. Transcriptional analysis showed dramatic changes in the metabolism of bacteria
400 treated with FS-1. The bacterial population switched completely to anaerobic respiration and
401 predominant gluconeogenesis instead of glycolysis. It was hypothesized that these metabolic
402 alternations were caused by a need to cope with an oxidative stress, damaging of cell membrane,
403 periplasmic proteins and fatty acids by iodine containing in FS-1 nanoparticles. In association
404 with a strong inhibition of many transmembrane peptide and amino acid transporters, the
405 activation of bacterial cell anabolism is realized at the expense of weakening of cell defense
406 systems that makes the bacteria more sensitive to antibiotics despite of the observed activation of
407 the plasmid born beta-lactamases. The FS-1 treated *E. coli* remained sensitive to antibiotics even
408 after removal of the drug from the medium. It may be explained by an involvement in gene
409 regulation of epigenetic and/or epistatic factors, which may persist several bacterial generations
410 keeping the window open for ridding pathogens off by the treatment with the antibiotics that the
411 pathogen was resistant to.

412 This study showed that the treatment of *E. coli* with FS-1 alternated the pattern of methylation of
413 the chromosomal DNA. In the strain *E. coli* BAA-196, the dominant type of DNA modification
414 is bipartite adenosine methylation in sequence motifs G<u>A</u>*T*C controlled by S-adenosylmethionine

415 depended Dam methylase that is common in *E. coli* [52]. Another type of bipartite methylation at
416 palindromic motifs AAC(N6)G*T*GC and GCAC(N6)G*T*T is under control of *Eco*KI methylase
417 [53]. In the sequenced genomes these two methylases are encoded by the genes *dam* and *hsdM*
418 located on the chromosomes. Expression of these genes is not alternated by the treatment with
419 FS-1. Both enzymes are part of the DNA restriction/modification systems associated with
420 cognate restriction endonucleases cleaving DNA. This is the mechanism of prevention of
421 bacteria from infection by foreign DNA, particularly by phages and destructive conjugative
422 plasmids. Methylation of the corresponding sites protects the host DNA from the cleavage by
423 cognate endonucleases. Other authors report that partial methylation of these sites naturally
424 observed in long-term stationary phase [54] may prevent DNA replication [55] and reduces
425 resistance to antibiotics [56]. Approximately 25% of these motifs are methylated only at one
426 DNA strand that may be of importance at the methylated nucleotides more likely interfere with
427 transcription of gene, which are transcribed from the same DNA strand. An example of such
428 regulation is controlling of the expression from a pyelonephritis-associated pili (pap) operon in
429 uropathogenic *E. coli* by methylation of a GATC site in the promoter region reported by Hernday
430 et al. [57]. The operon is transcribed when the methylation occurs proximal to the promoter, and
431 vice versa. Another role of GATC partially methylated sites is switching between alternative
432 promoters of an operon as it was shown on an example the *sci1* gene cluster in enteroaggregative
433 *E. coli* [58].

434 The global pattern of cytosine methylation at motifs CRGKGA*T*C, CCAGGRAH and
435 WCCCT*G*GYR shows more variations between the NC and FS chromosomes and plasmids.
436 CCAGGRAH and WCCCTGGYR sequences share the same central motif CCWGG.
437 Nevertheless, they show different patterns of distribution and are likely controlled by different

438 methyltransferases. Methylation of CCWGG motifs is controlled by Dcm cytosine methylase
439 also referred to as Mec methylase in early publications [59]. Two alleles of *dcm* genes associated
440 with *Eco*RII restriction-modification systems were found in the genome of *E. coli* ATCC BAA-
441 196 with chromosomal and plasmid locations: BAA-196NC_1741 and BAA-196NC_RS24085.
442 Both these genes are expressed at NC and FS conditions. The role of CCWGG-methylation in
443 gene regulation in both, eukaryotes and prokaryotes, has been reported in many publications.
444 This is the mechanism by which Sry protein controls sex development in mammals [60]. Other
445 authors reported a crucial role of CCWGG-methylation control in cancer development [61]. In
446 *E. coli*, CCWGG-methylation is a regulator of stationary phase processes and stress response
447 [59] that also was confirmed in experiments on *Vibrio cholerae* [62]. CCWGG-methylation also
448 was shown to increase the rate of spontaneous mutations [63]. However, no phenotypic changes
449 were reported for *E. coli* cells with under- and over-production of the Dcm methyltransferase
450 suggesting that the amount of methyltransferase molecules does not affect the methylation
451 pattern [52]. This doze independence of the Dcm methylation pattern may be explained by the
452 specificity of *Eco*RII restriction-modification systems in *E. coli*. *Eco*RII restriction enzymes do
453 not cleave DNA at a single recognition site but require binding to an additional target site serving
454 as an allosteric effector [64]. Thus, the restriction of CCWGG sites in *E. coli* may depend on the
455 spatial conformation of the chromosomal and plasmid chromatin, which may be guided by
456 methylation of CCWGG sites.

457 The alternative distribution of methylated sites under the effect of FS-1 may contribute to the
458 specific gene transcription regulation fitting to survival with the drug but to the detriment of the
459 resistance to antibiotics. The comparison of alternations in the global DNA methylation patterns
460 of the control and FS-1 treated *E. coli* cultures with the differential gene expression did not allow

461 deducing any reliable links between presence/absence of methylated sites and up- or
462 downregulation of the genes. Depending on the location of a methylated nucleotide within
463 genetic regions recognized by transcriptional activators, inhibitors and/or DNA-dependent RNA
464 polymerases, the effect of methylation may be ambivalent. Switching between alternative
465 promoter regions and transcriptional start points may have no effect on the expression level but
466 on properties and activities of the encoded protein. Lower QV scores of cytosine methylation
467 compared to adenine methylation may indicate that these sites are methylated only in a fraction
468 of the bacterial population. It may complicate finding links between the alternative cytosine
469 methylation and gene regulation. On the other hand, it suggests a possible role of cytosine
470 methylation in supporting necessary phase-variations and metabolic heterogeneity of bacterial
471 populations. The highest QV scores of cytosine methylation were observed in sequences
472 CRGKGA*T*C associated with bipartite adenine methylation. This type of methylation has never
473 been described before and the involved cytosine methylase is unknown. Repetitive adenine and
474 cytosine methylation of alternative loci in the NC and FS genomes discovered in three repetitions
475 of PacBio sequencing demonstrates an obvious importance of the DNA methylation for the
476 adaptation of the bacterium to the presence of iodine-containing nanoparticles in the medium and
477 may be associated with the observed increase in susceptibility to gentamicin and ampicillin.

478 Conclusion

479 Iodine is one of the oldest antibacterial medicines used by humankind. Acquired resistance to
480 iodine, in contrast to antibiotics, has never been reported. Application of iodine-containing nano-
481 molecular complexes, such as FS-1, allows a broader use of iodine against antibiotic resistance
482 pathogens. This study showed a profound alternation of the gene expression profile in the

483 antibiotic resistant strain *E. coli* ATCC BAA-196 cultivated with a sub-bactericidal
484 concentration of FS-1 that led to an increase in susceptibility to gentamicin and ampicillin. The
485 analysis of the differential gene regulation suggested that the possible targets of iodine-
486 containing particles are cell membrane fatty acids and proteins, particularly the cytochrome
487 molecules, that leads to an oxidative and acidic stress on the bacterial cells. Damaging of the cell
488 wall and membrane structures increases penetrability of bacterial cells by toxic compounds and
489 heavy metal ions. Bacteria responded by an increased activity of anion ion pumps and inhibiting
490 of permeases that complicates the uptake of nutrient and energy source compounds into cells,
491 which are already stressed by the necessity to increase the cell anabolism to replace iodine
492 damaged proteins and fatty acids. All this led to the increased susceptibility to antibiotics which
493 remains even after the removal of the drug from the medium. Inability of a quick restoration of
494 the antibiotic resistance by bacteria indicates the involvement of epigenetic and epistatic
495 mechanisms in adaptation to FS-1. This hypothesis was confirmed by discovering of a steadily
496 alternated DNA methylation pattern in FS-1 treated *E. coli*. Another hypothesis that iodine may
497 halogenate the bacterial DNA cannot be ruled out; however, the current results do not confirm
498 this supposition. Nevertheless, FS-1 suspended antibiotic resistance allows an effective treatment
499 of drug resistant pathogens that was demonstrated in three rounds of clinical evaluation of FS-1
500 as a new drug against antibiotic-resistant tuberculosis (www.clinicaltrials.gov, acc.
501 NCT02607449).

502 # Methods

503 **Bacterial cultures**

504 Model multidrug resistant microorganism *Escherichia coli* ATCC BAA-196™ producing
505 extended spectrum beta-lactamase was obtained from the ATCC collection
506 (https://www.lgcstandards-atcc.org/en.aspx) and kept in freezer at minus 80°C. Lyophilized
507 culture obtained from ATCC was reactivated in two passages in Mueller-Hinton (MH) liquid
508 medium (Himedia, India) supplemented with ceftazidime 10 μg/ml as recommended in the
509 ATCC product sheet.

510 **Cultivation with FS-1**

511 Bacteria were inoculated into test-tubes with 10 ml of liquid MH medium supplemented with FS-
512 1 (500 μg/ml) that corresponds to ½ minimal bactericidal concentration (MBC) of the drug
513 calculated previously for *E. coli*. As a negative control, the culture was cultivated in the same
514 medium without FS-1. Test-tubes were incubated at 37°C for 24 h and then 0.1 ml aliquots of the
515 cultures were transferred to fresh tubes with the corresponding media. After 10 passages of daily
516 recultivations, the experimental and control bacteria were cross-inoculated in three repetitions
517 into tubes with FS-1 containing and FS-1 free media for further overnight incubation, followed
518 by DNA extraction. Sensitivity of bacterial cultures to ampicillin and gentamicin was evaluated
519 by serial dilution in 100 ml liquid MH medium with 0.05% resazurin (Sigma-Aldrich) using 96
520 well plates. Plates were inoculated with 20 μl overnight bacterial cultures diluted to 1.5×10^6
521 CFU/ml. Plates were incubated at 37°C for 24 h. Optical density (OD) was measured at
522 540/620 nm using Multiscan Ascent and compared to the negative control well with resazurin.

523 **Radiological study of FS-1complexing with chromosomal DNA**

524 To carry out the radiological study, isotope ^{131}I was used for the synthesis of FS-1. The
525 documented radioactivity of the isotope 20 MBq/ml was controlled at the beginning of the
526 synthesis. Radioactivity was measured by β-spectrometer Hidex 300 SL (Finland).

527 Bacteria were incubated in a thermo-shaker overnight on MH medium at 37°C. The culture
528 growth was centrifuged at 5,000 g and the cells were resuspended in saline to achieve
529 1×10^9 CFU/ml. Aliquots of 0.4 ml of the marked FS-1 solution (total radioactivity of an aliquot
530 was 300 Kbq) were added to 7.6 ml of the bacterial suspension and rigorously shaken. Samples
531 were incubated for 1 h at 37°C in a thermo-shaker, Thermomixer comfort (Germany). After
532 incubation, bacterial cells were washed from FS-1 by two rounds of centrifugation at 5,000 g and
533 re-suspended in sterile saline. DNA samples were extracted from bacterial cells using PureLink
534 Genomic DNA Kits (Publication Number: MAN0000601, Revision 2.0), following the
535 manufacturer's recommendations. DNA purity and concentration were controlled by NanoDrop
536 2000C (USA) at wavelengths 260 and 280 nm. Then the DNA samples were resolved in
537 scintillation solute (ULTIMA GOLD LLT, PerkinElmer) to a final volume of 5 ml. The residual
538 radiation was measured by β-spectrometer Hidex 300 SL (Finland).

539 **DNA extraction and PacBio sequencing**

540 DNA samples were extracted from bacterial cells using PureLink Genomic DNA Kits
541 (MAN0000601, Revision 2.0) following the manufacturer's recommendations.

542 Samples were prepared according to a guide of preparation of SMRTbell templates for
543 sequencing on the PacBio RS II System. The samples were sequenced at Macrogene (South
544 Korea) on SMRT Cell 8Pac V3 cells using the DNA Polymerase Binding Kit P6 and DNA

545 Sequencing Reagent 4.0 v2 following the SMRTbell 10-kb library preparation protocol. PacBio
546 reads generated in three repeats from the NC and FS genomes were deposited at NCBI SRA
547 database under accession numbers (SRX6842003, SRX6842011, SRX6842012) and
548 (SRX6842007 – SRX6842009), respectively.

549 **Genome assembly and annotation**

550 The complete genome assembly was performed using SMRT Link 6.0.0 pipeline as published
551 previously [11]. Two genome variants, NC and FS, were obtained. The NC (CP042865) and FS
552 (CP042867) chromosomes were 4,682,561 and 4,682,572 bp, respectively. The genomes also
553 comprise large plasmids of 266,396 and 279,992 bp, for NC (CP042866) and FS (CP042868)
554 variants, respectively. Moreover, FS-1 treated strain contains two smaller plasmids CP042869
555 and CP042870 of 44,240 bp and 11,153 bp, respectively. The completeness of the final
556 assemblies was evaluated using the benchmarking universal single-copy orthologous (BUSCO)
557 software [65]. Genome annotation was performed using the RAST Server (http://rast.nmpdr.org/;
558 [66]) and then manually corrected. Involvement of predicted genes in metabolic pathways was
559 predicted by Pathway Tools [67] and The EcoCyc Database [25]. Genes involved in antibiotic
560 resistance were predicted by RGI 5.1.0/CARD 3.0.5 Web-portal [68]. Locations of horizontally
561 transferred genomic islands were identified and general DNA sequence parameters such as GC-
562 content and GC-skew in 8 kbp sliding windows stepping 2 kbp along the genome sequence were
563 calculated by the program SeqWord Genome Island Sniffer [69]. The same program was used to
564 identify the replication origin and terminus on bacterial chromosomes by GC-skew between the
565 leading and lagging strands [70]. The genomes were deposited at NCBI with the accession
566 numbers CP042865-CP042866 for the NC variant of the chromosome and large plasmid of this
567 strain; and CP042867-CP042870 for the chromosome and three plasmids of the FS variant. Raw

568 PacBio sequences are available from the NCBI BioProject Website under the accession number
569 PRJNA557356.

570 **RNA extraction and sequencing**

571 Isolation of total RNA from the NC and FS cultures after overnight cultivation (the mid
572 logarithmic growth phase controlled by OD) on the respective media was performed with the
573 RiboPure Bacteria Kit (Ambion, Lithuania) according to the developer's guidelines. The quality
574 and quantity of the resulting RNA were determined using the Qubit 2.0 Fluorometer (Thermo
575 Scientific, USA) and Qubit RNA Assay Kit (Invitrogen, USA). Purification of total RNA from
576 ribosomal 16S and 23S RNA was carried out using the MICROBExpress Bacterial mRNA
577 Purification Kit (Ambion, Lithuania), as recommended by the developer. The efficiency of the
578 template-RNA purification was determined on the Bioanalyzer 2100 (Agilent, Germany) with
579 the RNA 6000 Nano LabChip Kit (Agilent Technologies, Lithuania).
580 The RNA fragment library was prepared by enzymatic fragmentation with the Ion Total RNA
581 Seq Kit V2 (Life Technologies, USA). Barcoding of the formed library was carried out with the
582 Ion Xpress RNA-Seq Barcode 01-16 Kit (Life Technologies, USA), according to the
583 manufacturer's instructions. RNA sequencing was performed in three repetitions for each
584 condition using the Ion Torrent PGM sequencer (Life Technologies, USA) with the Ion 318 Chip
585 Kit V2. Three repetitions of RNA reads generated from the NC and FS variants were deposited
586 at NCBI SRA database under accession numbers SRX6842006 and SRX6842002, respectively.

587 **Differential gene expression analysis**

588 The differential expression was done using the R-3.4.4 software. Firstly, a reference index was
589 built for each reference genome using the *buildindex* function available in the *Rsubreads*
590 package (Bioconducter). For each bacterium, the obtained RNA fragments were aligned to the

591 relevant reference genome with the use of the "align" function. The aligned BAM files and
592 relevant GFF annotation files were then used as input for the *featureCounts* function to obtain
593 gene counts. The R packages *DESeq2* (Bioconducter) and *GenomicFeatures* was then used in R
594 studio for the differential expression analyses. The visualization of the volcano plots of
595 differential expression in the studied genomes was performed by an in-house Python script using
596 the output excel files generated by R package *DESeq2*. Networks of regulated genes were
597 constructed using the Web based tool PheNetic [71].

598 **Profiling of epigenetic modifications**

599 SMRT Link 6.0.0 tools (https://www.pacb.com/support/software-downloads/) pipelined by an in-
600 house Python script was used for the base call kinetic analysis in PacBio reads generated from
601 chromosomal and plasmid DNA. The pipeline consists of the following steps. The complete
602 genome consensus sequence in FASTA format obtained by assembling of PacBio reads is
603 indexed by the program *samtools* to be used as the reference sequence for the alignment of
604 PacBio reads. PacBio reads are converted from the original BAX.H5 format to BAM format by
605 the program *bax2bam*. Reads stored in BAM files are aligned against the indexed reference
606 sequence by the program *blasr*. Aligned reads in BAM format are sorted by locations on the
607 replicon and indexed by *samtools sort* and *index* functions. On the next step, the sorted and
608 indexed BAM files of read alignments against the reference genome are analyzed by the program
609 *ipdSummary*, which estimates the base call kinetics for every nucleotide in the reference genome
610 (the output file **_kinetics.csv*). Also, for every nucleotide position significantly delaying the base
611 calling in multiple overlapping repeats, the program calculates several statistical parameters such
612 as IPDRatio of exceeding of the base call delay over the expectation and the QV identification
613 score representing the likelihood of modification of the nucleotide at the given location. QV

614 score 14 corresponds to p-value 0.05, and QV score 21 – p-value 0.01. The program stores all
615 estimated parameters together with context sequences into an output file *_*basemods.gff*.
616 Contextual motifs of base modifications are searched by the program *motifMaker*. Then the
617 visualization of profiles of epigenetic modifications in the studied genomes was performed by an
618 in-house Python script using the output files of the above-mentioned pipeline.

619

620 **References**

621 1. Reygaert WC. An overview of the antimicrobial resistance mechanisms of bacteria.
622 AIMS Microbiol. 2018; 4:482-501. doi: 10.3934/microbiol.2018.3.482.
623 2. López-Jácome E, Franco-Cendejas R, Quezada H, Morales-Espinosa R, Castillo-Juárez I,
624 González-Pedrajo B, et al, García-Contreras R. The race between drug introduction and
625 appearance of microbial resistance. Current balance and alternative approaches. Curr.
626 Opin. Pharmacol. 2019; 48:48-56. doi: 10.1016/j.coph.2019.04.016.
627 3. Brooks BD, Brooks AE. Therapeutic strategies to combat antibiotic resistance. Adv.
628 Drug Deliv. Rev. 2014; 78:14-27. doi: 10.1016/j.addr.2014.10.027.
629 4. Smith PA, Romesberg FE. Combating bacteria and drug resistance by inhibiting
630 mechanisms of persistence and adaptation. Nat. Chem. Biol. 2007; 3:549-556. doi:
631 10.1038/nchembio.2007.27.
632 5. Baym M, Stone LK, Kishony R. Multidrug evolutionary strategy to reverse antibiotic
633 resistance. Science. 2016; 351:aad3292. doi: 10.1126/science.aad3292.
634 6. Ilin AI, Kulmanov ME, Korotetskiy IS, Islamov RA, Akhmetova GK, Lankina MV, Reva
635 ON. Genomic insight into mechanisms of reversion of antibiotic resistance in multidrug

636 resistant *Mycobacterium tuberculosis* induced by a nanomolecular iodine-containing
637 complex FS-1. Front. Cellular and Infect. Microbiol. 2017; 7:151. doi:
638 10.3389/fcimb.2017.00151.
639 7. Islamov R, Kerimzhanova B, Ilin A. New antituberculosis drug FS-1 in Vašková J,
640 Vaško, L. (Eds.), Medicinal Chemistry. IntechOpen, London, UK, 2018; 103-116.
641 http://dx.doi.org/10.5772/intechopen.80795.
642 8. Ilin AI, Kulmanov ME. Antibacterial agent for treating infectious diseases of bacterial
643 origin. Patent US20140010782 A1—2014-09-03. 2014;
644 https://www.google.com/patents/US20140010782.
645 9. Ilin AI, Kulmanov ME, Korotetskiy IS, Akhmetova GK, Lankina MV, Shvidko SV, Reva
646 ON. Complete genome sequence of multidrug-resistant clinical isolate *Mycobacterium*
647 *tuberculosis* 187.0, used to study the effect of drug susceptibility reversion by the new
648 medicinal drug FS-1. Genome Announc. 2015; 3:e01272-15. doi:
649 10.1128/genomeA.01272-15.
650 10. Joubert M, Reva O.N, Korotetskiy IS, Shvidko SV, Shilov SV, Jumagaziyeva AB, et al.
651 Assembly of complete genome sequences of negative-control and experimental strain
652 variants of *Staphylococcus aureus* ATCC BAA-39 selected under the effect of the drug
653 FS-1, which induces antibiotic resistance reversion. Microbiol. Resour. Announc. 2019;
654 8:e00579-19. doi: 10.1128/MRA.00579-19.
655 11. Korotetskiy IS, Joubert M, Setshaba T, Jumagaziyeva AB, Shilov SV, Shvidko SV, et al.
656 Complete genome sequence of a multidrug resistant strain *Escherichia coli* ATCC BAA-
657 196 as a model to study the induced antibiotic resistance reversion. Microbiol. Resour.
658 Announc. 2019; 8:e01118-19. doi: 10.1128/MRA.01118-1.

659 12. Rice LB, Willey SH, Papanicolaou GA, Medeiros AA, Eliopoulos GM, Moellering RC.
660 Jr Jacoby GA. Outbreak of ceftazidime resistance caused by extended-spectrum beta-
661 lactamases at a Massachusetts chronic-care facility. Antimicrob. Agents Chemother.
662 1990; 34:2193-2199. https://doi.org/0.1128/aac.34.11.2193.
663 13. Neu HC. Infections due to gram-negative bacteria: an overview. Rev. Infect. Dis. 1985; 7
664 Suppl 4:S778-S782. doi: 10.1093/clinids/7.supplement_4.s778.
665 14. Guerrant RL, Hughes JM, Lima NL, Crane J. Diarrhea in developed and developing
666 countries: magnitude, special settings, and etiologies. Rev. Infect. Dis. 1990; 12 Suppl
667 1:S41-50. doi: 10.1093/clinids/12.supplement_1.s41.
668 15. Emori TG, Gaynes RP. An overview of nosocomial infections, including the role of the
669 microbiology laboratory. Clin. Microbiol. Rev. 1993; 6:428-442. doi:
670 10.1128/cmr.6.4.428.
671 16. Bezuidt O, Pierneef R, Mncube K, Lima-Mendez G, Reva ON. Mainstreams of
672 horizontal gene exchange in enterobacteria: consideration of the outbreak of
673 enterohemorrhagic *E. coli* O104:H4 in Germany in 2011. PLoS One. 2011; 6:e25702.
674 doi: 10.1371/journal.pone.0025702
675 17. Yang SC, Lin CH, Aljuffali IA, Fang JY. Current pathogenic *Escherichia coli* foodborne
676 outbreak cases and therapy development. Arch. Microbiol. 2017; 199:811-825. doi:
677 10.1007/s00203-017-1393-y.
678 18. Downing T. Tackling drug resistant infection outbreaks of global pandemic *Escherichia*
679 *coli* ST131 using evolutionary and epidemiological genomics. Microorganisms. 2015;
680 3:236-267. doi: 10.3390/microorganisms3020236.

681 19. Odonkor ST, Addo KK. Prevalence of multidrug-resistant *Escherichia coli* isolated from
682 drinking water sources. Int. J. Microbiol. 2018; 2018:7204013. doi:
683 10.1155/2018/7204013.
684 20. O'Flaherty E, Solimini AG, Pantanella F, De Giusti M, Cummins E. Human exposure to
685 antibiotic resistant-*Escherichia coli* through irrigated lettuce. Environ. Int. 2019;
686 122:270-280. doi: 10.1016/j.envint.2018.11.022.
687 21. Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, et al. The
688 complete genome sequence of *Escherichia coli* K-12. Science. 1997; 277:1453-1462. doi:
689 10.1126/science.277.5331.1453.
690 22. Hayashi K, Morooka N, Yamamoto Y, Fujita K, Isono K, Choi S, et al. Highly accurate
691 genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110. Mol. Syst.
692 Biol. 2006; 2:2006.0007. doi: 10.1038/msb4100049.
693 23. Durfee T, Nelson R, Baldwin S, Plunkett G 3rd, Burland V, Mau B, et al. The complete
694 genome sequence of *Escherichia coli* DH10B: insights into the biology of a laboratory
695 workhorse. J. Bacteriol. 2008; 190:2597-2606. doi: 10.1128/JB.01695-07.
696 24. Cassu-Corsi D, Martins WM, Nicoletti AG, Almeida LG, Vasconcelos AT, Gales AC.
697 Characterisation of plasmid-mediated *rmtB*-1 in Enterobacteriaceae clinical isolates from
698 São Paulo, Brazil. Mem. Inst. Oswaldo Cruz. 2018; 113:e180392. doi: 10.1590/0074-
699 02760180392.
700 25. Karp PD, Ong WK, Paley S, Billington R, Caspi R, Fulcher C, et al. The EcoCyc
701 Database. EcoSal Plus. 2018; 8(1). doi: 10.1128/ecosalplus.ESP-0006-2018.

702 26. Freddolino PL, Amini S, Tavazoie S. Newly identified genetic variations in common
703 *Escherichia coli* MG1655 stock cultures. J. Bacteriol. 2012; 194:303-306. doi:
704 10.1128/JB.06087-11
705 27. Eloff JN. A sensitive and quick microplate method to determine the minimal inhibitory
706 concentration of plant extracts for bacteria. Planta Med. 1998; 64:711-713. doi:
707 10.1055/s-2006-957563.
708 28. Ilin AI, Parsadanyan GG, Nersesyan AK. Interaction of halogens with nucleic acids and
709 its consequence. The New Armenian Medical Journal. 2013; 7:33-43.
710 29. Dong T, Yu R, Schellhorn H. Antagonistic regulation of motility and transcriptome
711 expression by RpoN and RpoS in *Escherichia coli*. Mol. Microbiol. 2011; 79(2):375-386.
712 doi: 10.1111/j.1365-2958.2010.07449.x.
713 30. Korotetskiy IS, Shilov SV, Shvidko SV, Jumagaziyeva AB, Suldina NA, Korotetskaya
714 NV, et al. Transcriptional response of the multidrug resistant *Staphylococcus aureus*
715 following FS-1 exposure. Eurasian Journal of Applied Biotechnology. 2017; 2017:43-48.
716 31. Stim KP, Bennett GN. Nucleotide sequence of the adi gene, which encodes the
717 biodegradative acid-induced arginine decarboxylase of *Escherichia coli*. J. Bacteriol.
718 1993; 175(5):1221-1234.
719 32. Jiang GR, Nikolova S, Clark DP. Regulation of the *ldhA* gene, encoding the fermentative
720 lactate dehydrogenase of *Escherichia coli*. Microbiology. 2001; 147(Pt 9):2437-2446.
721 doi: 10.1099/00221287-147-9-2437.
722 33. Horlacher R, Boos W. Characterization of TreR, the major regulator of the *Escherichia*
723 *coli* trehalose system. J. Biol. Chem. 1997; 272(20):13026-13032.

724 34. Mora L, Heurgué-Hamard V, de Zamaroczy M, Kervestin S, Buckingham RH.
725 Methylation of bacterial release factors RF1 and RF2 is required for normal translation
726 termination *in vivo*. J. Biol. Chem. 2007; 282(49):35638-35645.
727 35. Gagarinova A, Stewart G, Samanfar B, Phanse S, White CA, Aoki H, et al. Systematic
728 genetic screens reveal the dynamic global functional organization of the bacterial
729 translation machinery. Cell Rep. 2016; 17(3):904-916. doi: 10.1016/j.celrep.2016.09.040.
730 36. Wada A, Igarashi K, Yoshimura S, Aimoto S, Ishihama A. Ribosome modulation factor:
731 stationary growth phase-specific inhibitor of ribosome functions from *Escherichia coli*.
732 Biochem. Biophys. Res. Commun. 1995; 214(2):410-417.
733 37. Yamaguchi Y, Park JH, Inouye M. MqsR, a crucial regulator for quorum sensing and
734 biofilm formation, is a GCU-specific mRNA interferase in *Escherichia coli*. J. Biol.
735 Chem. 2009; 284(42):28746-28753. doi: 10.1074/jbc.M109.032904.
736 38. Kerimzhanova BF, Jumagaziyeva A, Ilin AI. Influence of FS-1 drug on resistant
737 *Mycobacterium tuberculosis*. Int. Conf. Clinical and Pharmaceutical Microbiology. 2019;
738 Oct. 23 – 25, Rome, Italy.
739 39. Matsuda A, Kurono N, Kawano C, Shirota K, Hirabayashi A, Horino M, et al. Genome-
740 wide screen for *Escherichia coli* genes involved in repressing cell-to-cell transfer of non-
741 conjugative plasmids. Biochem. Biophys. Res. Commun. 2012; 428(4):445-450. doi:
742 10.1016/j.bbrc.2012.10.098.
743 40. Soo VW, Hanson-Manful P, Patrick WM. Artificial gene amplification reveals an
744 abundance of promiscuous resistance determinants in *Escherichia coli*. Proc. Natl. Acad.
745 Sci. U.S.A. 2011; 108(4):1484-1489. doi: 10.1073/pnas.1012108108.

746 41. Nachin L, Nannmark U, Nyström T. Differential roles of the universal stress proteins of
747 *Escherichia coli* in oxidative stress resistance, adhesion, and motility. J. Bacteriol. 2005;
748 187(18):6265-6272.
749 42. Pomposiello PJ, Bennik MH, Demple B. Genome-wide transcriptional profiling of the
750 *Escherichia coli* responses to superoxide stress and sodium salicylate. J. Bacteriol. 2001;
751 183(13):3890-3902.
752 43. Wang X, Dubey AK, Suzuki K, Baker CS, Babitzke P, Romeo T. CsrA post-
753 transcriptionally represses *pgaABCD*, responsible for synthesis of a biofilm
754 polysaccharide adhesin of *Escherichia coli*. Mol. Microbiol. 2005; 56(6):1648-1663.
755 44. Jørgensen MG, Nielsen JS, Boysen A, Franch T, Møller-Jensen J, Valentin-Hansen P.
756 Small regulatory RNAs control the multi-cellular adhesive lifestyle of *Escherichia coli*.
757 Mol. Microbiol. 2012; 84(1):36-50. doi: 10.1111/j.1365-2958.2012.07976.x.
758 45. De Lay N, Gottesman S. A complex network of small non-coding RNAs regulate motility
759 in *Escherichia coli*. Mol. Microbiol. 2012; 86(3):524-538. doi: 10.1111/j.1365-
760 2958.2012.08209.x.
761 46. Attila C, Ueda A, Wood TK. 5-Fluorouracil reduces biofilm formation in *Escherichia*
762 *coli* K-12 through global regulator AriR as an antivirulence compound. Appl. Microbiol.
763 Biotechnol. 2009; 82(3):525-533. doi: 10.1007/s00253-009-1860-8.
764 47. Dressaire C, Moreira RN, Barahona S, Alves de Matos AP, Arraiano CM. BolA is a
765 transcriptional switch that turns off motility and turns on biofilm development. mBio.
766 2015; 6(1):e02352-14. doi: 10.1128/mBio.02352-14.

767 48. Weber MM, French CL, Barnes MB, Siegele DA, McLean RJ. A previously
768 uncharacterized gene, *yjfO* (*bsmA*), influences *Escherichia coli* biofilm formation and
769 stress response. Microbiology. 2010; 156(Pt 1):139-147. doi: 10.1099/mic.0.031468-0.
770 49. Domka J, Lee J, Wood TK. YliH (BssR) and YceP (BssS) regulate *Escherichia coli* K-12
771 biofilm formation by influencing cell signaling. Appl. Environ. Microbiol. 2006;
772 72(4):2449-2459.
773 50. Khil PP, Camerini-Otero RD. Over 1000 genes are involved in the DNA damage
774 response of *Escherichia coli*. Mol. Microbiol. 2002; 44(1):89-105.
775 51. Bougdour A, Cunning C, Baptiste PJ, Elliott T, Gottesman S. Multiple pathways for
776 regulation of sigmaS (RpoS) stability in *Escherichia coli* via the action of multiple anti-
777 adaptors. Mol. Microbiol. 2008; 68(2):298-313.
778 52. Marinus MG, Løbner-Olesen A. DNA methylation. EcoSal Plus. 2014; 6. doi:
779 10.1128/ecosalplus.ESP-0003-2013.
780 53. Loenen WA. Tracking *Eco*KI and DNA fifty years on: a golden story full of surprises.
781 Nucleic Acids Res. 2003; 31:7059-7069. doi: 10.1093/nar/gkg944.
782 54. Westphal LL, Sauvey P, Champion MM, Ehrenreich IM, Finkel SE. Genomewide Dam
783 methylation in *Escherichia coli* during long-term stationary phase. mSystems. 2016;
784 1:e00130-16. doi: 10.1128/mSystems.00130-16.
785 55. Russell DW, Zinder ND. Hemimethylation prevents DNA replication in *E. coli*. Cell.
786 1987; 50:1071-1079. doi: 10.1016/0092-8674(87)90173-5.
787 56. Cohen NR, Ross CA, Jain S, Shapiro RS, Gutierrez A, Belenky P, et al. A role for the
788 bacterial GATC methylome in antibiotic stress survival. Nat. Genet. 2016; 48:581-586.
789 doi: 10.1038/ng.3530.

790 57. Hernday AD, Braaten BA, Low DA. The mechanism by which DNA adenine methylase
791 and PapI activate the pap epigenetic switch. Mol. Cell. 2003;12(4):947-957. doi:
792 10.1016/s1097-2765(03)00383-6.
793 58. Brunet YR, Bernard CS, Cascales E. Fur-Dam regulatory interplay at an internal
794 promoter of the enteroaggregative *Escherichia coli* type VI secretion *sci1* gene cluster. J.
795 Bacteriol. 2020; 202(10) pii:e00075-20. doi: 10.1128/JB.00075-20.
796 59. Kahramanoglou C, Prieto AI, Khedkar S, Haase B, Gupta A, Benes V, Fraser GM,
797 Luscombe N.M, Seshasayee A.S. Genomics of DNA cytosine methylation in *Escherichia*
798 *coli* reveals its role in stationary phase transcription. Nat. Commun. 2012; 3:886. doi:
799 10.1038/ncomms1878.
800 60. Nishino K, Hattori N, Sato S, Arai Y, Tanaka S, Nagy A, Shiota K. Non-CpG
801 methylation occurs in the regulatory region of the Sry gene. J. Reprod. Dev. 2011;
802 57:586-593. doi: 10.1262/jrd.11-033a.
803 61. Agirre X, Vizmanos JL, Calasanz MJ, García-Delgado M, Larráyoz M, Novo FJ.
804 Methylation of CpG dinucleotides and/or CCWGG motifs at the promoterof TP53
805 correlates with decreased gene expression in a subset of acutelymphoblastic leukemia
806 patients. Oncogene. 2003; 22:1070-1072. doi: 10.1038/sj.onc.1206236.
807 62. Chao MC, Zhu S, Kimura S, Davis BM, Schadt EE, Fang G, Waldor MK. A cytosine
808 methytransferase modulates the cell envelope stress response in the cholera pathogen.
809 PLOS Genet. 2015; 11:e1005666. doi: 10.1371/journal.pgen.1005739.
810 63. Krokan HE, Drabløs F, Slupphaug G. Uracil in DNA – occurrence, consequences and
811 repair. Oncogene. 2002; 21:8935-8948. doi: 10.1038/sj.onc.1205996.

812 64. Golovenko D, Manakova E, Tamulaitiene G, Grazulis S, Siksnys V. Structural
813 mechanisms for the 5'-CCWGG sequence recognition by the N- and C-terminal domains
814 of *Eco*RII. Nucleic Acids Res. 2009; 37:6613-6624. doi: 10.1093/nar/gkp699.
815 65. Seppey M, Manni M, Zdobnov EM. BUSCO: assessing genome assembly and annotation
816 completeness. Methods Mol. Biol. 2019; 1962:227-245. doi: 10.1007/978-1-4939-9173-
817 0_14.
818 66. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST
819 Server: Rapid Annotations using Subsystems Technology. BMC Genomics. 2008; 9:75.
820 doi: 10.1186/1471-2164-9-75.
821 67. Karp PD, Latendresse M, Paley SM, Krummenacker M, Ong QD, Billington R, et al.
822 Pathway Tools version 19.0 update: software for pathway/genome informatics and
823 systems biology. Brief Bioinform. 2016; 17:877-890. doi: 10.1093/bib/bbv079.
824 68. Jia B, Raphenya A.R, Alcock B, Waglechner N, Guo P, Tsang K.K, et al. CARD 2017:
825 expansion and model-centric curation of the comprehensive antibiotic resistance
826 database. Nucleic Acids Res. 2017; 45(D1):D566-D573. doi: 10.1093/nar/gkw1004.
827 69. Bezuidt O, Lima-Mendez G, Reva ON. SEQWord Gene Island Sniffer: a program to
828 study the lateral genetic exchange among bacteria. World Academy of Science,
829 Engineering and Technology. 2009; 58:1169-1274.
830 70. Freeman JM, Plasterer TN, Smith TF, Mohr SC. Patterns of genome organization in
831 bacteria. Science. 1998; 279:1827-1827. doi: 10.1126/science.279.5358.1827a.
832 71. De Maeyer D, Renkens J, Cloots L, De Raedt L, Marchal K. PheNetic: network-based
833 interpretation of unstructured gene lists in *E. coli*. Molecular bioSystems. 2013, 9
834 (7):1594-1603. doi:10.1039/c3mb25551d.

835

836 # Figure legends

837

838 **Fig 1. Growth of *Escherichia coli* ATCC BAA-196 in plate wells with different**
839 **concentration of antibiotics measured in 3 repetitions.** X axis depicts antibiotic concentration
840 and Y axis shows OD values at 540/620 nm light wave normalized by OD values in negative
841 control well with no bacterial growth.

842

843 **Fig 2. Mauve alignments of genome sequences of the plasmids of *E. coli* ATCC BAA-196.**
844 Sequence identity of homologous blocks is depicted by colored histograms. Gene locations are
845 represented by open blocks. (A) Alignment of two large plasmids of NC and FS variants. (B)
846 Alignment of the plasmid 2 of the variant FS against the large plasmid of the variant NC. (C)
847 Alignment of the plasmid 2 of the variant FS against the large plasmid of the variant FS. (D)
848 Alignment of the plasmid 3 of the variant FS against the large plasmid of the variant FS.

849

850 **Fig 3. Volcano plot of gene regulation in the FS-1 treated variant (FS) of *E. coli* ATCC**
851 **BAA-196 in comparison to the negative control variant (NC).** Dots on the plots depict the
852 negative (blue – CDS, green - ncRNA) and positive (red – CDS, yellow - ncRNA) gene
853 regulation. Plasmid born genes are depicted by red cycles and beta-lactamase genes located on
854 the plasmids are highlighted by a thick yellow outline. Log2 fold gene expression changes are on
855 the X axes, with the common p-values logarithms of gene expression estimations on the Y axes.

856 P-values were calculated by DESeq2 algorithm on three repeats of gene expression counts.
857 Genes expressed only in the variant FS are plotted along the axis of expression counts in the
858 positive infinity columns (Inf), and the genes expressed only in the NC variant are in the
859 negative infinity columns (-Inf). Size of the cycles and the color intensity depict estimated p-
860 value above 0.05 and above 0.01 (bigger and brighter cycles). Numbers of regulated CDS and
861 ncRNA genes of different categories are indicated. Highly regulated genes are labelled.

862

863 **Fig 4. PheNetic networks of regulated genes.** (A) Network of genes clustered according to the
864 common regulation of these genes by higher level transcriptional regulators; (B) Network of
865 regulated genes functionally link to each other. Several functional groups of genes are labelled.
866 Upregulated genes are depicted by yellow nodes and downregulated genes – by blue nodes
867 (vertices). Color intensity indicates the level of the regulation. Grey nodes are genes of
868 transcriptional regulators involved in the network, which expression was not reliable changed at
869 the aforementioned conditions. Green edges show activation relations; blue edges – inhibitory
870 relations; and brown edges – ambivalent or neutral relations. Direct regulations by transcriptional
871 regulators are indicated by arrowheads.

872

873 **Fig 5. Distribution of QV scores of modified nucleotides.** Plot (A) shows the distribution in
874 the negative control (NC) variant of *E. coli* ATCC BAA-196; plot (B) shows the distribution in
875 the FS-1 treated (FS) variant. Axis X depicts score ranks stepping 5 score units. Axis Y indicates
876 decimal logarithms of rank numbers.

877

878 **Fig 6. Alternative cytosine methylation.** Patterns associated with CRGKGA*T*C motifs in the
879 chromosomes (A-B) and CCAGG motifs in the large plasmids (C-D) of the NC and FS variants
880 of *E. coli* ATCC BAA-196 are shown. Triangle marks indicate locations of methylated sites on
881 the direct (blue marks) and complement (red marks) DNA strands. Fluctuations of GC-content
882 and GC-skew parameters calculated for 8 kbp sliding window stepping 2 kbp are shown by
883 corresponding histograms. Locations of identified genomic islands in bacterial chromosomes are
884 indicated by yellow blocks.

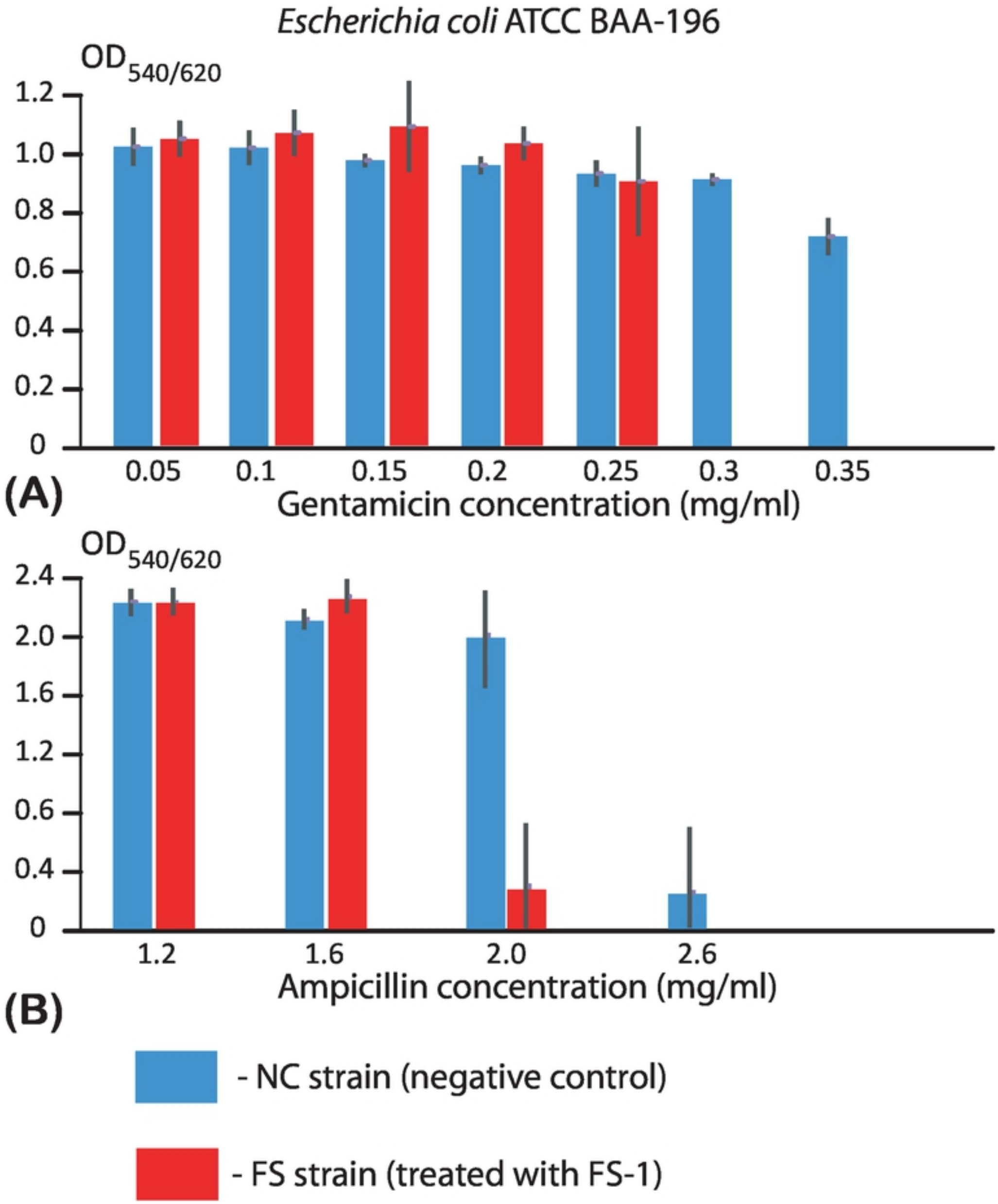
Escherichia coli ATCC BAA-196
OD540/620
1.2
1.0
0.8
0.6
0.4
0.2
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
(A)
Gentamicin concentration (mg/ml)
OD540/620
2.4
2.0
1.6
1.2
0.6
0.4
0
1.2
1.6
2.0
2.6
Ampicillin concentration (mg/ml)
(B)
- NC strain (negative control)
- FS strain (treated with FS-1)

Figure 1

Log2(baseMean+1)
-Log10(p-value)
Log2FoldChange
103 cds+8 nc
217 cds+13 nc
63 cds+1 nc
154 cds+1 nc
44 cds+1 nc
51 cds
-Inf
Inf
gltA
rmf
icd
osmY
sucA
cstA
yliH
fadE
sdhA
gltI
cspD
ompF
fadA
sucD
tpx
fdoG
sucC
ttdA
ecnB
ychH
rybA
ryeE
ryeA
csrB
fdhF
pflB
glpQ
glpT
adiA
copA
yhjX
hycI
purC
yjiY
yfiD
ydfZ
narG
focA
hypF
adiC
ymgG
cueO
aldA
phoH
sdhB
ter
cafA
trhA
hydN
aacA6
glpA
penP

Figure 3

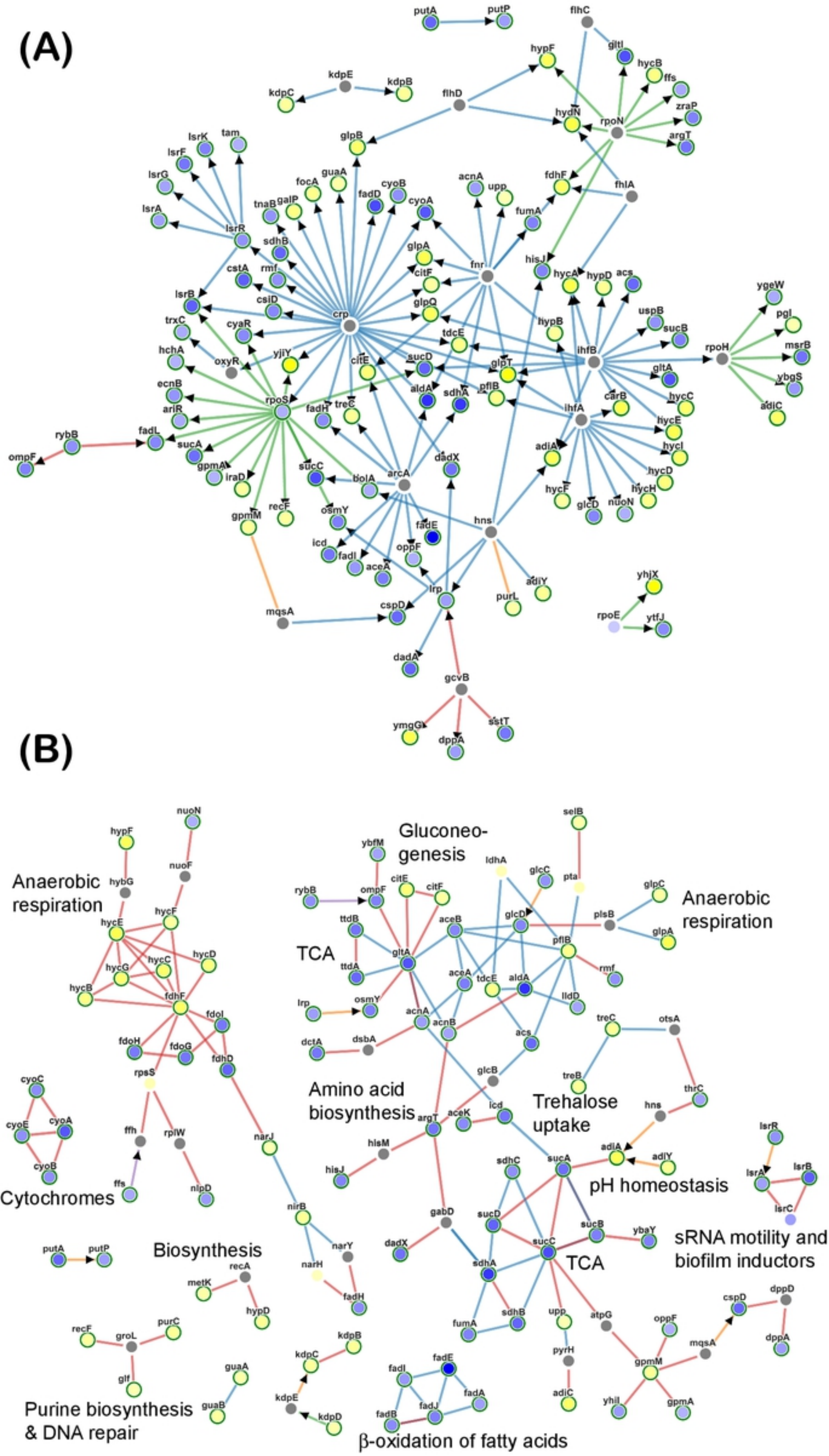
(A)
putA
putP
flhC
hypF
gltI
hycB
ffs
zraP
argT
kdpE
kdpC
kdpB
flhD
hydN
rpoN
glpB
lsrK
tam
lsrF
lsrG
lsrA
lsrR
focA
guaA
tnaB
galP
fadD
cyoB
cyoA
acnA
upp
fumA
fdhF
flhA
sdhB
glpA
citF
fnr
hisJ
hycA
hypD
acs
cstA
rmf
csiD
glpQ
crp
lsrB
trxC
cyaR
hchA
oxyR
yjiY
citE
sucD
tdcE
hypB
uspB
sucB
ygeW
pgl
rpoH
msrB
ybgS
adiC
ihfB
glpT
gltA
aldA
sdhA
pflB
carB
hycC
hycE
hycI
hycD
hycH
nuoN
glcD
ecnB
arlR
rpoS
fadH
treC
ihfA
adiA
rybB
fadL
ompF
sucA
gpmA
iraD
sucC
bolA
arcA
dadX
recF
osmY
gpmM
fadE
hns
icd
fadI
oppF
aceA
lrp
adiY
purL
yhjX
rpoE
ytfJ
mqsA
cspD
dadA
gcvB
ymgG
dppA
sstT
(B)
nuoN
hypF
nuoF
Anaerobic respiration
hybG
hycF
hycE
hycC
hycD
hycG
hycB
fdhF
fdoI
fdoH
fdoG
fdhD
rpsS
cyoC
cyoE
cyoA
cyoB
Cytochromes
ffh
rplW
ffs
nlpD
putA
putP
Biosynthesis
recA
metK
hypD
recF
groL
purC
glf
Purine biosynthesis & DNA repair
guaA
guaB
kdpB
kdpC
kdpE
kdpD
β-oxidation of fatty acids
fadE
fadI
fadA
fadJ
fadB
ybfM
Gluconeo-genesis
ldhA
glcC
pta
selB
rybB
ompF
citE
citF
ttdB
aceB
glcD
plsB
glpC
glpA
Anaerobic respiration
TCA
gltA
ttdA
pflB
rmf
aceA
tdcE
aldA
lldD
lrp
osmY
acnA
acnB
acs
dctA
dsbA
glcB
treC
otsA
treB
thrC
Amino acid biosynthesis
argT
aceK
icd
Trehalose uptake
hns
hisM
hisJ
adiA
adiY
sdhC
sucA
pH homeostasis
lsrR
lsrA
lsrB
lsrC
sRNA motility and biofilm inductors
nirB
narJ
gabD
sucD
sucB
ybaY
dadX
sucC
TCA
narY
narH
sdhA
fadH
fumA
sdhB
upp
atpG
oppF
cspD
dppD
dppA
mqsA
pyrH
gpmM
adiC
yhiI
gpmA

Figure 4

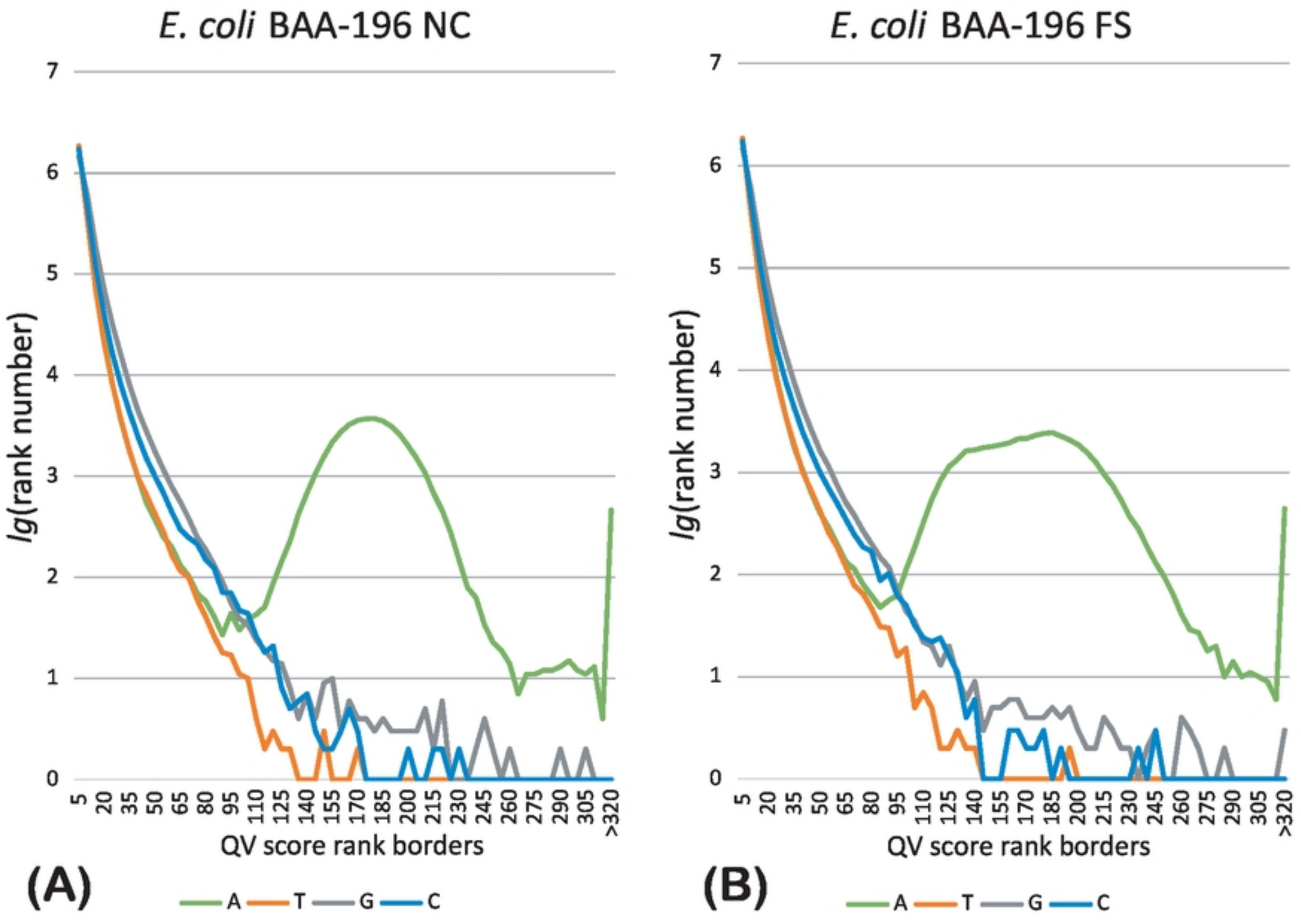
E. coli BAA-196 NC
E. coli BAA-196 FS
lg(rank number)
QV score rank borders
0
1
2
3
4
5
6
7
5
20
35
50
65
80
95
110
125
140
155
170
185
200
215
230
245
260
275
290
305
>320
(A)
(B)
A
T
G
C

Figure 5

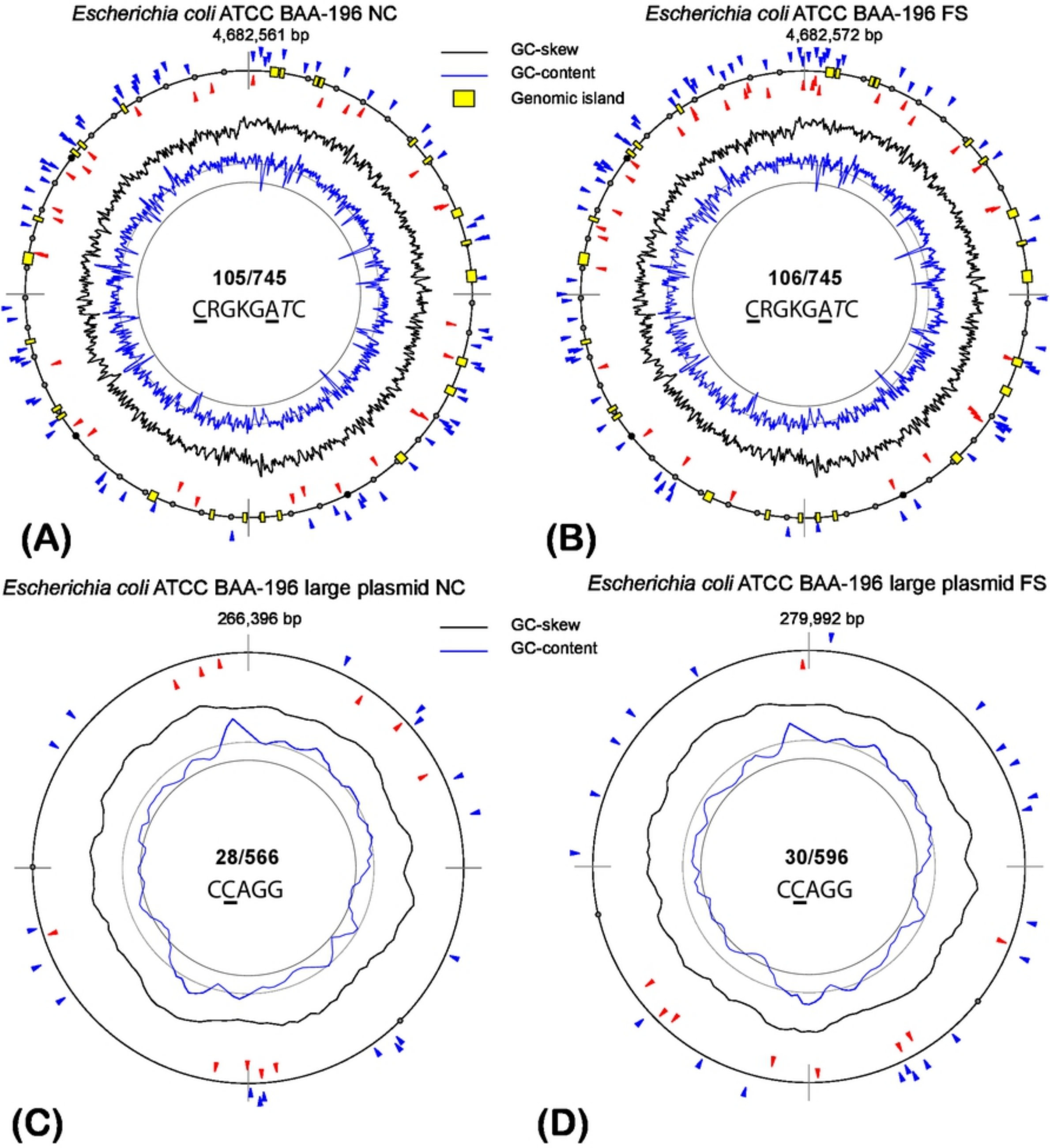
Escherichia coli ATCC BAA-196 NC
4,682,561 bp
GC-skew
GC-content
Genomic island
105/745
CRGKGATC
(A)
Escherichia coli ATCC BAA-196 FS
4,682,572 bp
106/745
CRGKGATC
(B)
Escherichia coli ATCC BAA-196 large plasmid NC
266,396 bp
GC-skew
GC-content
28/566
CCAGG
(C)
Escherichia coli ATCC BAA-196 large plasmid FS
279,992 bp
30/596
CCAGG
(D)

Figure 6

(A)
E. coli ATCC BAA-196 NC pl 1
E. coli ATCC BAA-196 FS pl 1
(B)
E. coli ATCC BAA-196 NC pl 1
E. coli ATCC BAA-196 FS pl 2
(C)
E. coli ATCC BAA-196 FS pl 1
E. coli ATCC BAA-196 FS pl 2
(D)
E. coli ATCC BAA-196 FS pl 1
E. coli ATCC BAA-196 FS pl 2

Figure 2