CONSERVED SEQUENCES OF GENES CODING FOR THE MULTIDRUG RESISTANCE PUMP ACRAB-TOLC OF *ESCHERICHIA COLI* SUGGEST THEIR INVOLVEMENT INTO PERMANENT CELL "CLEANING"

Karakozova MV[1] ✉, Nazarov PA[2]

[1] Research Centre of Medical Genetics, Moscow

[2] Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow

Multidrug resistance pumps (MDR pumps) of bacteria confer protection against aggressive environmental factors. The genes coding for MDR pumps are thought to be variable. They belong to the group of the so-called contingency genes, i.e are necessary for bacterial adaptation to the changing environment. The aim of the present work was to establish how conserved are the sequences of genes coding for MDR pumps. We analyzed the sequences of AcrA, AcrB and TolC proteins of different *Escherichia coli* strains. Using sequence alignment tools, we demonstrated that strains originating in different countries and cultured in the labs for a long time are amazingly conserved in terms of AcrAB-TolC sequences. They resemble housekeeping genes, suggesting the involvement of the AcrAB-TolC pump into permanent "cleaning" of various biotic and abiotic agents.

Keywords: multidrug resistance, antibiotic, AcrAB-TolC, sequence alignment, *Escherichia coli*, pump, transporter, biocide, contingency genes, housekeeping genes

Acknowledgement: the authors wish to thank students (Dzhafarova T, Mamedova D, Maushev F, Miranian K, Omarova D and Rodionova E) and the Department of Morphology and Pathology of Medical Institute REAVIZ for the provided data, their analysis and discussion.

✉ **Correspondence should be addressed:** Marina Karakozova
Microrayon 1, d.15, kv. 43, Zaraisk, Zaraiski r-n, Moscow region, 140601; mvk752002@gmail.com

Received: 19.02.2018 **Accepted:** 30.03.2018

DOI: 10.24075/brsmu.2018.024

КОНСЕРВАТИВНОСТЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ГЕНОВ ПОМПЫ МНОЖЕСТВЕННОЙ ЛЕКАРСТВЕННОЙ УСТОЙЧИВОСТИ ACRAB-TOLC *ESCHERICHIA COLI* КАК ПРИЗНАК ВОВЛЕЧЕННОСТИ В ПЕРМАНЕНТНУЮ «УБОРКУ» БАКТЕРИАЛЬНОЙ КЛЕТКИ

М. В. Каракозова[1] ✉, П. А. Назаров[2]

[1] Медико-генетический научный центр, Москва

[2] Научно-исследовательский институт физико-химической биологии имени А. Н. Белозерского, Московский государственный университет имени М. В. Ломоносова, Москва

Помпы множественной лекарственной устойчивости (МЛУ) помогают бактериям защищаться от неблагоприятного воздействия окружающей среды. Считается, что гены, кодирующие помпы МЛУ, вариабельны и относятся к так называемым генам «роскоши», т. е. предназначены для адаптации бактерий к изменению окружающих условий. Целью работы было проверить насколько консервативны последовательности генов помпы МЛУ. Для этого проводили анализ последовательностей белков AcrA, AcrB и TolC для различных лабораторных штаммов *Escherichia coli*. Методом выравнивания последовательностей было показано, что штаммы из разных стран, культивируемые в лабораториях уже долгое время, имеют удивительную консервативность последовательностей белков помпы AcrAB-TolC. Она напоминает консервативность генов «домашнего хозяйства», что, по-видимому, говорит о вовлеченности помпы МЛУ AcrAB-TolC в перманентную «уборку» клетки от различных веществ биотического и абиотического происхождения.

Ключевые слова: множественная лекарственная устойчивость, антибиотик, AcrAB-TolC, выравнивание последовательностей, *Escherichia coli*, помпа, транспортер, биоцид, гены «роскоши», гены «домашнего хозяйства»

Благодарности: авторы благодарят студентов (Джафарову Т., Мамедову Д., Маушева Ф., Миранян К., Омарову Д. и Родионову Э.) и сотрудников кафедры морфологии и патологии Московского медицинского университета «Реавиз» за предварительные данные, их анализ и дискуссию.

✉ **Для корреспонденции:** Марина Викторовна Каракозова
Микрорайон 1-й, д. 15, кв. 43, г. Зарайск, Зарайский р-н, Московская обл., 140601; mvk752002@gmail.com

Статья получена: 19.02.2018 **Статья принята к печати:** 30.03.2018

DOI: 10.24075/vrgmu.2018.024

The gram-negative gammaproteobacterium *Escherichia coli* (*E. coli*) was first discovered by Theodor Escherich in the stool samples of healthy individuals in 1885 [1]. *E. coli* naturally inhabits the lower intestines of warm-blooded species and is an important object of research. Four strains of *E. coli*, including K-12, B, W and C, are now used as model organisms. Strain K-12 was first isolated at Stanford university in 1922 [2]. Strain B was described by d'Herelle at the Pasteur Institute in Paris in 1918 [3]. The other 2 strains are less common. Strain C was discovered by Margaret Lieb in 1951 [4, 5], and strain W was originally reported by Selman Waksman in 1943 [6]. Strains comprising groups K-12 and B are the most widespread and best known. Laboratory strains have "evolved" to lose some of their properties, such as the ability to form biofilms on abiotic surfaces, and therefore can be advantageously used in research studies, especially for the discovery of novel antibiotics [7].

The pressure of both natural and artificial selection existing in laboratories has produced numerous derivatives of K-12 and B that are now used all over the world (Table 1). Among the derivatives of strain B are BL21 and BL21(DE3); DH5α, JM109, W3110, XL-1 Blue, and MG1655 are examples of strain K-12 derivatives.

Discovery of novel antibiotics or their effective alternatives is a pressing challenge. One of the most promising areas of research is identification of multidrug resistance (MDR) pump inhibitors. MDR pumps are responsible for removing antibiotics from the bacterial cell. Studies of deletion mutants with knocked-out genes coding for MDR pumps demonstrate that minimum effective inhibitory concentrations of antibiotics in their case are several times lower than usual [8]. This may help to reduce both treatment costs and the toxic effect of antibiotic therapies on the patient. Although effects of MDR pumps on antibacterial agents are actively studied, there is an extensive list of objective factors preventing cross-study comparisons, such as different genetic backgrounds of the strains. Even for such closely related strains as W3110 and MG1655 [9], the number of differences at genomic sites can be over 200, impeding comparison. Because bacterial resistance to drugs depends on the presence or absence of efflux pumps, we hypothesized that *E. coli* strains with identical sequences of MDR pumps might have comparable or equal resistance. To check this supposition, we selected the AcrAB-TolC pump. We aimed to compare sequences of AcrA, AcrB and TolC proteins obtained from different laboratory strains of *E. coli* and to study the associations between drug resistance and possible mutations if such were present in a sequence.

METHODS

Selecting an object

For our study we selected a few K-12 strains: W3110, MG1655, NEB 5-alpha, MDS42, GM4792, AG100, MC4100, DH10B, ER3413, HMS174, BW2952, and BW25113, as well as strain BL21(DE3) from group B. Their *acrA*, *acrB* and *tolC* sequences are known and stored in databases (Table 2).

Selecting a reference sequence

When selecting a reference sequence, we bore in mind a large number of deletion mutants in *E. coli* K-12 BW25113. It is a parent strain for the Keio collection, which comprises *E. coli* strains with 3,985 deletions (of 4,288 total *E. coli* genes) [10]. Sequence AIN30961.1 was selected as a reference sequence for AcrA; AIN30960.1, as a reference sequence for AcrB, and AIN33386.1, as a reference sequence for TolC.

Sequence alignment

Sequences were analyzed using a standard local alignment tool NCBI BLASTp, which allows comparison of multiple alignments [11], and the STRING database [12]. Visual representation of the results was done in NCBI MSA Viewer [13]. Each protein sequence was aligned against its reference sequence.

RESULTS

It is known that bacterial resistance can be a product of: 1) accumulation of resistance genes in plasmids; 2) increased expression of genes coding for MDR pumps; 3) gene duplication; 4) accumulation of mutations [14, 15]. Increased expression and accumulation of mutations in the genes coding for MDR pumps can result in single nucleotide polymorphisms (SNPs) in the amino acid sequences of proteins. Therefore, bacterial resistance can be predicted by sequence analysis.

Bacterial genes are subdivided into housekeeping genes, which support vital functions of the cell, and contingency genes, which play an important role in bacterial adaptation to the changing environment. Housekeeping genes usually have a low mutation rate, while contingency genes tend to demonstrate a high mutation rate [16]. It is believed that genes coding for multidrug efflux pumps are contingency genes; therefore, the proteins they encode are expected to have variable primary structures. Because laboratory strains are usually subject to the pressure of natural selection induced by various biocides and mutagens, the strains that have been cultured in world laboratories for over 100 years, as well as their derivatives, might be different in terms of their amino acid polymorphisms. The strains compared in this work originate from different countries and continents (Table 1), so we can infer the presence of mutations in one of the AcrAB-TolC-encoding genes.

However, the analysis of aligned sequences of AcrA (Fig. 1), AcrB (Fig. 2) and TolC (Fig. 3) proteins (substrain BW25113), those of strain K-12 (substrains W3110, MG1655, NEB 5-alpha, MDS42, GM4792, AG100, MC4100, DH10B,

Table 1. Geographic origin of *E. coli* strains used in this work

Strain	Institution	City, country
MG1655	University of Wisconsin	Milwaukee, USA
W3110	Nara Institute of Science and Technology	Ikoma, Japan
BL21(DE3)	Korea Research Institute of Bioscience and Biotechnology	Daejeon, South Korea
MDS42	Osaka University	Osaka, Japan
MC4100	University of Kiel, Germany	Kiel, Germany
BW25113	Universite de Sherbrooke, Canada	Sherbrooke, Canada
ER3413	New England Biolabs	Ipswich, USA
AG100	University of Exeter	Exeter, UK
NEB 5-alpha	New England Biolabs	Ipswich, USA
HMS174	Austrian Centre of Industrial Biotechnology	Graz, Austria
BW2952	Nankai University	Nankai, China
DH10B	University of Wisconsin-Madison	Madison, USA
GM4792	Beijing Normal University	Beijing, China

Table 2. Accession numbers for the stored protein sequences of *acrA*, *acrB* and *tolC* genes

Substrain	Strain	AcrA	AcrB	TolC
MG1655	*K-12*	NP_414996.1	NP_414995.1	NP_417507.2
W3110	*K-12*	BAE76242.1	BAE76241.1	BAE77091.1
NEB 5-alpha	*K-12*	AOO68785.1	AOO68784.1	AOO71261.1
MDS42	*K-12*	BAL37669.1	BAL37668.1	BAL39694.1
GM4792	*K-12*	AKK16793.1	AKK13611.1	AKK18828.1
AG100	*K-12*	CQR80062.1	CQR80061.1	CQR82466.1
MC4100	*K-12*	CDJ70932.1	CDJ70931.1	CDJ73817.1
DH10B	*K-12*	ACB01590.1	ACB01589.1	ACB04120.1
ER3413	*K-12*	AIZ54314.1	AIZ54313.1	AIZ52829.1
HMS174	*K-12*	CDY55568.1	CDY55565.1	CDY61615.1
BW2952	*K-12*	ACR63806.1	ACR63808.1	ACR65687.1
BW25113	*K-12*	AIN30961.1	AIN30960.1	AIN33386.1
BL21(DE3)	*B*	ACT42313.1	ACT42312.1	ACT44711.1

ER3413, HMS174, and BW2952) and those of strain B (substrain BL21(DE3)) reveals the absence of polymorphisms in all three proteins constituting the AcrAB-TolC efflux pump, regardless of whether the strain belongs to the derivatives of K-12 or B.

Considering the fact that *E. coli* mutation rate is ~1×10^{-3} per genome per generation [17] or even higher ($3–4\times10^{-3}$ per genome per generation) [18], we hypothesize that the AcrAB-TolC pump sequence is conserved. Given the same sequence coverage for all studied proteins (397 amino acid residues for AcrA, 1049 amino acid residues for AcrB and 493 amino acid residues for TolC), the sequence identity was 100%.

DISCUSSION

According to the currently existing classification, strains from group B and K-12 belong to phylogroup A [19], which may explain the similarity of amino acid sequences between all three proteins but not their identity. Our findings allow us to conclude the presence of a consensus sequence of a highly conserved AcrAB-TolC ensemble. Thus, the selected protein reference sequences AcrA (AIN30961.1 for AcrA, AIN30960.1 for AcrB and AIN33386.1 for TolC, respectively) are consensus for the studied *E. coli* strains.

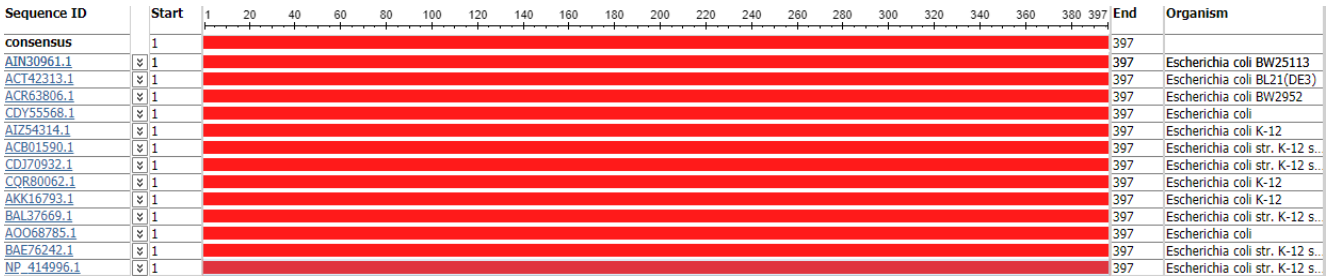

Fig. 1. Alignment of AcrA sequences for strains K-12 and B against the reference AcrA sequence of substrain BW25113

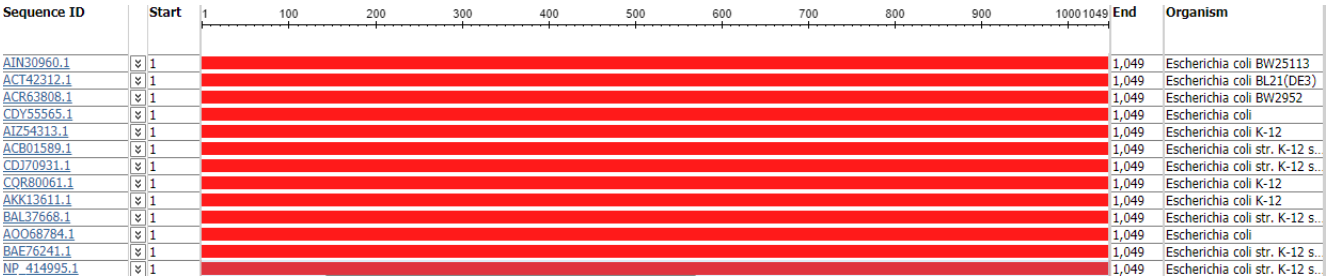

Fig. 2. Alignment of AcrB sequences for strains K-12 and B against the reference AcrB sequence of substrain BW25113

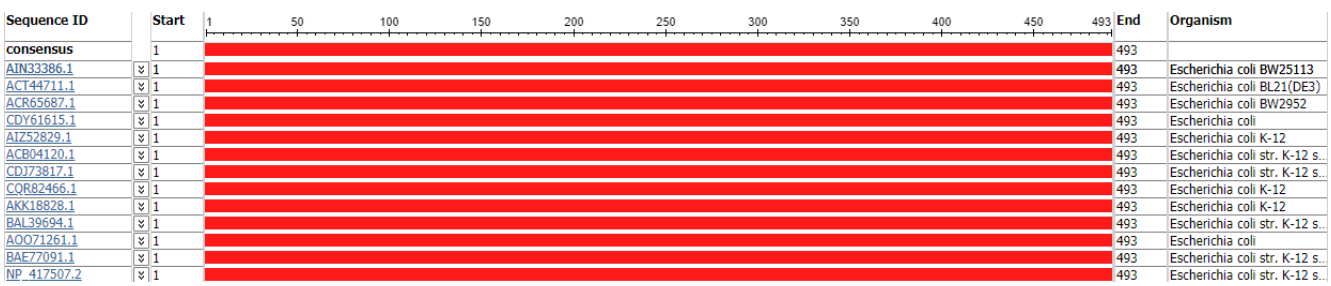

Fig. 3. Alignment of TolC sequences for strains K-12 and B against the reference TolC sequence of substrain BW25113

The discovered sequences are consensus for all representatives of group A and possibly other phylogroups, including B1, B2, D, and E, which can facilitate normalization of sequences against their consensus counterparts.

The absence of point mutations in the genes coding for protein components of the AcrAB-TolC pump in all studied strains is indicative of the strict selection control, as is the case with housekeeping genes. Such control is particularly important for the major multidrug efflux pump of *E. coli* (AcrAB-TolC) responsible for removing benzalkonium chloride, ethidium bromide, indole, hexane, antibiotics (erythromycin, ciprofloxacin, etc.), rhodamine, berberine and also triphenylphosphonium and its derivatives from the cell [20–21].

It would be wrong to see genes coding for MDR pumps as responsible for biocide resistance only. They have a role in bacterial colonization and persistence [22], so it is not limited to defense against antibiotics. It appears that proteins produced by MDR pump-encoding genes routinely protect bacterial cells from various biotic and abiotic agents and can be regarded as housekeeping genes engaged in permanent cell "cleaning", unlike contingency genes that get involved only at certain times.

CONCLUSION

Our findings suggest a unique role of the AcrAB-TolC multidrug resistance pump in *E. coli*. The protein sequence of AcrAB-TolC has turned to be surprisingly conserved. This provides a fresh look at AcrAB-TolC from a different angle: this pump ensures permanent protection against aggressive environment, determines bacterial resistance to antibiotics or their alternatives and even ensures bacterial survival.

References

1. Escherich T. Die Darmbakterien des Neugeborenen und Säuglinge. Fortschr. Med. 1885; 3: 515–522.
2. Daegelen P, Studier FW, Lenski RE, Cure S, Kim JF. Tracing ancestors and relatives of Escherichia coli B, and the derivation of B strains REL606 and BL21(DE3). J Mol Biol. 2009; 394 (4): 634–43.
3. Bachmann BJ. Pedigrees of Some Mutant Strains of Escherichia coli K-12. Bacteriological Reviews, 1972; 36 (4): 525–557.
4. Lieb M, Weigle JJ, Kellenberger E. A study of hybrids between two strains of Escherichia coli. Journal of Bacteriology. 1955; 69 (4): 468–471.
5. Lieb M. Forward and Reverse Mutation in a Histidine-Requiring Strain of Escherichia Coli Genetics. 1951; 36 (5): 460–477.
6. Archer CT, Kim JF, Jeong H, Park JH, Vickers CE, Lee SY, et al. The genome sequence of E. coli W (ATCC 9637): comparative genome analysis and an improved genome-scale reconstruction of E. coli. BMC Genomics. 2011; 12: 9.
7. Vidal O, Longin R, Prigent-Combaret C, Dorel C, Hooreman M, Lejeune P. Isolation of an Escherichia coli K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new ompR allele that increases curli expression. J Bacteriol. 1998; 180 (9): 2442–2449.
8. Sulavik MC, Houseweart C, Cramer C, Jiwani N, Murgolo N, Greene J, et al. Antibiotic susceptibility profiles of Escherichia coli strains lacking multidrug efflux pump genes. Antimicrob Agents Chemother. 2001; 45 (4): 1126–1136.
9. Hayashi K, Morooka N, Yamamoto Y, Fujita K, Isono K, Choi S, et al. Highly accurate genome sequences of Escherichia coli K-12 strains MG1655 and W3110. Mol Syst Biol. 2006; 2: 2006.0007.
10. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol. 2006; 2: 2006.0008.
11. BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)
12. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017; 45: D362–68.
13. Multiple Sequence Alignment Viewer v. 1.7.7 (https://www.ncbi.nlm.nih.gov/tools/msaviewer/about/)
14. Nikaido H. Multidrug Resistance in Bacteria. Annu Rev Biochem. 2009; 78: 119–146.
15. Martinez JL, Baquero F Mutation frequencies and antibiotic resistance. Antimicrob Agents Chemother. 2000; 44 (7): 1771–1777.
16. Moxon ER, Rainey PB. Nowak MA. Lenski RE. Adaptive evolution of highly mutable loci in pathogenic bacteria. Curr Biol. 1994; 4: 24–33.
17. Lee H, Popodi E, Tang H, Foster PL. Rate and molecular spectrum of spontaneous mutations in the bacterium Escherichia coli as determined by whole-genome sequencing. Proc Natl Acad Sci USA, 2012; 109 (41): E2774–E2783.
18. Drake JW. A constant rate of spontaneous mutation in DNA-based microbes. Proc Natl Acad Sci USA, 1991; 88: 7160–7164.
19. Sims GE, Kim S.-H. Whole-genome phylogeny of Escherichia coli/Shigella group by feature frequency profiles (FFPs) Proc Natl Acad Sci U S A. 2011; 108 (20): 8329–8334.
20. Pos KM. Drug transport mechanism of the AcrB efflux pump. Biochim Biophys Acta. 2009; 1794 (5): 782–793.
21. Nazarov PA, Osterman IA, Tokarchuk AV, Karakozova MV, Korshunova GA, Lyamzaev KG et al. Mitochondria-targeted antioxidants as highly effective antibiotics. Sci Rep. 2017; 7 (1): 1394.
22. Piddock LJ Multidrug-resistance efflux pumps — not just for resistance. Nat Rev Microbiol. 2006; 4 (8): 629–36.

Литература

1. Escherich T. Die Darmbakterien des Neugeborenen und Säuglinge. Fortschr. Med. 1885; 3: 515–522.
2. Daegelen P, Studier FW, Lenski RE, Cure S, Kim JF. Tracing ancestors and relatives of Escherichia coli B, and the derivation of B strains REL606 and BL21(DE3). J Mol Biol. 2009; 394 (4): 634–43.
3. Bachmann BJ. Pedigrees of Some Mutant Strains of Escherichia coli K-12. Bacteriological Reviews, 1972; 36 (4): 525–557.
4. Lieb M, Weigle JJ, Kellenberger E. A study of hybrids between two strains of E. coli. Journal of Bacteriology. 1955; 69 (4): 468–471.
5. Lieb M. Forward and Reverse Mutation in a Histidine-Requiring Strain of Escherichia Coli Genetics. 1951; 36 (5): 460–477.
6. Archer CT, Kim JF, Jeong H, Park JH, Vickers CE, Lee SY, et al. The genome sequence of E. coli W (ATCC 9637): comparative genome analysis and an improved genome-scale reconstruction of E. coli. BMC Genomics. 2011; 12: 9.
7. Vidal O, Longin R, Prigent-Combaret C, Dorel C, Hooreman M, Lejeune P. Isolation of an Escherichia coli K-12 mutant strain able

to form biofilms on inert surfaces: involvement of a new ompR allele that increases curli expression. J Bacteriol. 1998; 180 (9): 2442–2449.
8. Sulavik MC, Houseweart C, Cramer C, Jiwani N, Murgolo N, Greene J, et al. Antibiotic susceptibility profiles of Escherichia coli strains lacking multidrug efflux pump genes. Antimicrob Agents Chemother. 2001; 45 (4): 1126–1136.
9. Hayashi K, Morooka N, Yamamoto Y, Fujita K, Isono K, Choi S, et al. Highly accurate genome sequences of Escherichia coli K-12 strains MG1655 and W3110. Mol Syst Biol. 2006; 2: 2006.0007.
10. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol. 2006; 2: 2006.0008.
11. BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)
12. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017; 45: D362–68.
13. Multiple Sequence Alignment Viewer v. 1.7.7 (https://www.ncbi.nlm.nih.gov/tools/msaviewer/about/)
14. Nikaido H. Multidrug Resistance in Bacteria. Annu Rev Biochem. 2009; 78: 119–146.
15. Martinez JL, Baquero F Mutation frequencies and antibiotic resistance. Antimicrob Agents Chemother. 2000; 44 (7): 1771–1777.
16. Moxon ER, Rainey PB. Nowak MA. Lenski RE. Adaptive evolution of highly mutable loci in pathogenic bacteria. Curr Biol. 1994; 4: 24–33.
17. Lee H, Popodi E, Tang H, Foster PL. Rate and molecular spectrum of spontaneous mutations in the bacterium Escherichia coli as determined by whole-genome sequencing. Proc Natl Acad Sci USA, 2012; 109 (41): E2774–E2783.
18. Drake JW. A constant rate of spontaneous mutation in DNA-based microbes. Proc Natl Acad Sci USA, 1991; 88: 7160–7164.
19. Sims GE, Kim S.-H. Whole-genome phylogeny of Escherichia coli/Shigella group by feature frequency profiles (FFPs) Proc Natl Acad Sci U S A. 2011; 108 (20): 8329–8334.
20. Pos KM. Drug transport mechanism of the AcrB efflux pump. Biochim Biophys Acta. 2009; 1794 (5): 782–793.
21. Nazarov PA, Osterman IA, Tokarchuk AV, Karakozova MV, Korshunova GA, Lyamzaev KG et al. Mitochondria-targeted antioxidants as highly effective antibiotics. Sci Rep. 2017; 7 (1): 1394.
22. Piddock LJ Multidrug-resistance efflux pumps — not just for resistance. Nat Rev Microbiol. 2006; 4 (8): 629–36.