## **Essential mutations in Escherichia coli strain K-12 substrain MG1655 genome associated with antibiotic resistance** *Evgenia Khokhlova, Eugenia Ivanova (Strebulaeva)*

## **1. Abstract**

This paper demonstrates the results of studying single nucleotide polymorphisms (SNP) in the genome of Escherichia coli and their contribution to antibiotic resistance. Using sequencing data from Escherichia coli strain K-12 substrain MG1655 with resistance to ampicillin we found five single nucleotide polymorphisms, which then were examined as prospective mutations. It has been shown that three mutations were responsible for amino acid replacement in proteins associated with several well-known bacterial resistance mechanisms. For each of these mutations we assume its role in the genesis of antibiotic resistance. In addition, we make some treatment recommendations for patients infected by the strain of Escherichia coli under the study.

**2. Introduction**

In recent years microbial resistance to antibiotics has become a world-wide threat. Overuse and misuse of antibiotics in treatment of infectious diseases and in agriculture have led to emergence of resistant bacterial strains. Diseases caused by these bacterial strains are intractable and require development of new drugs and sophisticated therapeutic approaches.  
Microbes perfectly mastered four major mechanisms of resistance to antibiotics. Bacteria are able to change a target site for antibiotics which means alterations in structure of particular bacterial proteins (Munita and Arias 2016). The molecules of antibiotics can be inactivated by special bacterial enzymes. Sometimes bacteria can even alter metabolic pathways to compensate for the effect of antibiotics (Egorov, Ulyashova, and Rubtsova 2018). Finally, bacteria can affect the transport of antibiotics by more effective pumping of hostile molecules outside of the cell or blocking the penetration of antibiotics inside. Understanding of those mechanisms can help to devise thoughtful treatment strategies using newly creating and currently accessible drugs.

Nowadays scientists exploit different bacteria to study microbial resistance. Escherichia coli is a widely used model bacterium found in the intestine. Many strains of E. coli are representatives of normal microbiota but some strains cause serious clinical statements such as severe digestive disorders and sepsis (Davies and Davies 2010). This research is based on sequencing data of Escherichia coli strain K-12 substrain MG1655 resistant to antibiotic ampicillin. Ampicillin is an inhibitor of transpeptidase, the enzyme that plays the key role in cell wall synthesis during fission. Therefore, the resistance to ampicillin may occur due to mutations in genes, associated with cell wall synthesis or membrane transport. Whole-genome sequencing with analysis of single nucleotide polymorphisms can be used to elucidate genes that are probably responsible for ampicillin resistance in this particular strain and suggest the molecular mechanisms of its resistance. On this basis new methods of treatment can be created to prevent spreading of multi-drug resistant bacteria.

**3. Methods**

**Alignment to reference**

For our analysis we used the reference genome of E.coli strain K-12 substrain MG1655 downloaded from NCBI

ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/bacteria/Escherichia\_coli/all\_assembly\_versions/GCA\_000005845.2\_ASM584v2/.

And raw Illumina sequencing reads from shotgun sequencing of an *E. coli* strain that is resistant to the antibiotic ampicillin downloaded from

<http://public.dobzhanskycenter.ru/mrayko/amp_res_1.fastq.zip>, <http://public.dobzhanskycenter.ru/mrayko/amp_res_2.fastq.zip>.

**Filtering the reads**

To check the quality of the row sequence data we used the FASTQC program.

To improve the quality of the data raw reads were trimmed by the Trimmomatic tool. Trimmomatic parameters are shown in Table 1.

Table 1. - Parameters of trimming by Trimmomatic.

|  |  |
| --- | --- |
| Step | Description |
| LEADING:20 | 20 is a quality threshold |
| TRAILING:20 | 20 is a quality threshold |
| SLIDINGWINDOW:10:20 | 10-base wide window and 20 is a quality threshold |
| MINLEN:20 | 20 is a length threshold |

**Mapping sequences to reference**

BWA-MEM aligner was used to map our sequence file to the reference genome.

**Variant calling**

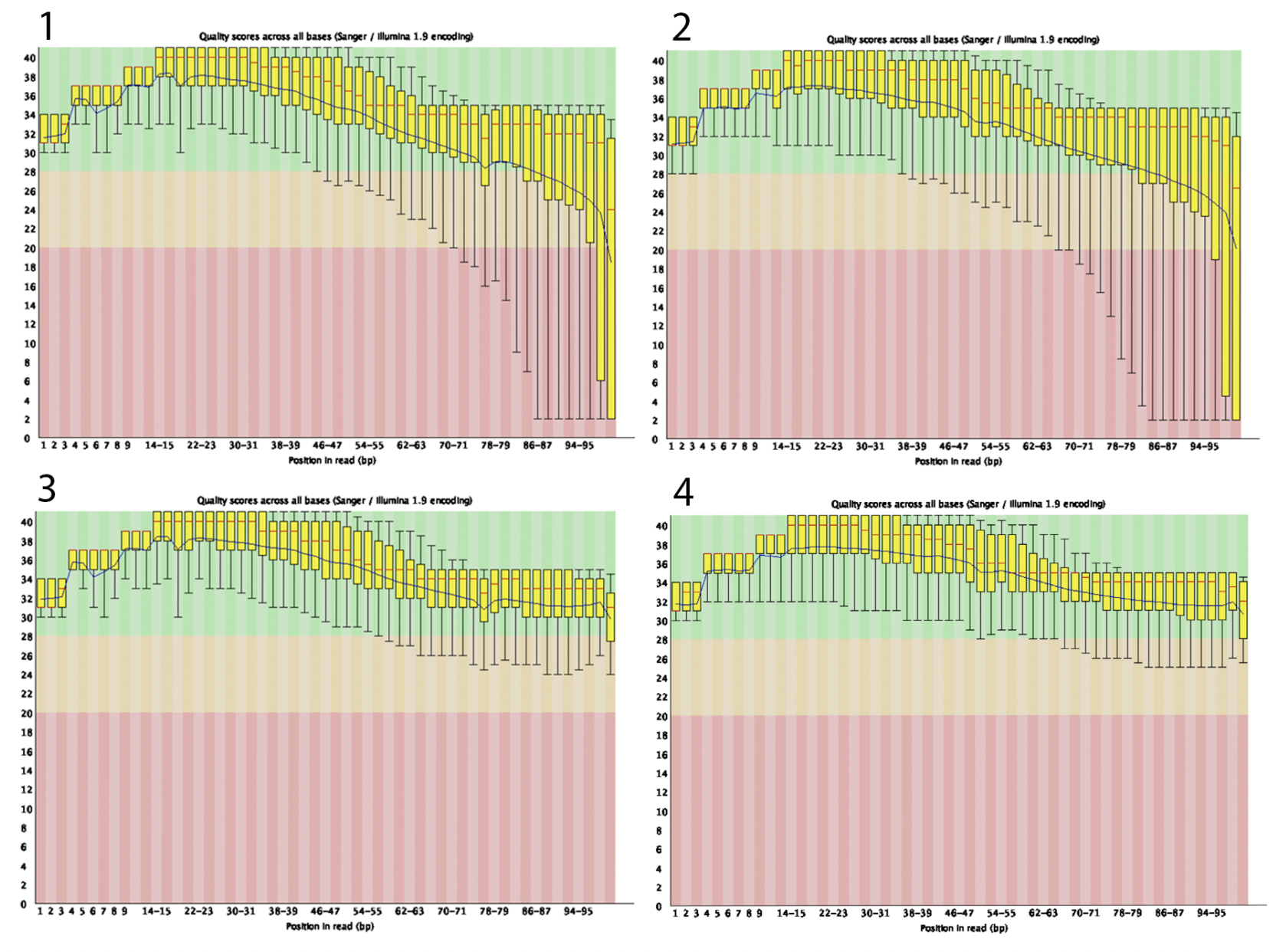
To see how many reads have a mutation at the same position we used a program called VarScan (variant scanner).

**Variant effect prediction**

To find out where mutations are and whether they actually change any proteins we visualize our data with IGV browser, finding out the position of SNP. And then we check this position in the UniProt database to find out what it means.

**4. Results**

In our study we’ve analyzed 4455876 reads of row data, after visualizing quality of reads with FastQC and using the Trimmomatic tool there were 446259 reads (Picture 1). 446259 reads were mapped to the reference genome.



Picture 1. Per-base read quality. 1, 2 - before trimming; 3,4 - after trimming.

We find out 5 SNPs described in table 2. Three mutations have been associated with amino acid substitutions in genes ftsl, acrB, envZ, the mutation in a gene rsgA was synonymous and did not lead to an amino acid substitution. The mutation in gene rybA was in the non-coding region.

Table 2. The mutation in E. coli resistant strain

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Gene** | **Protein** | **Function** | **Base position in genome** | **Amino acid**  **Change, position in genome** |
| **flsl** | Peptidoglycan DD-transpeptidase (PBP3 | Essential cell division protein that catalyzes cross-linking of the peptidoglycan cell wall at the division septum (Nguyen-Distèche et al. 1998) | 93043 | G**C**C - G**G**C (alanine- glycine)  A545G |
| **acrB** | Multi-drag efflux pump RND permease | AcrA-AcrB-AcrZ-TolC is a drug efflux protein complex with broad substrate specificity that uses the proton motive force to export substrates (Hobbs et al. 2012). | 482698 | C**A**G - C**T**G  (glutamine- leucine)  Q569L |
| **rybA** | Small RNA | Regulates synthesis of aromatic amino acids under peroxide stress conditions (Gerstle et al. 2012). | 852762 | T**T**T - T**C**T (noncoding - noncoding) |
| **envZ** | Sensory histidine kinase | Member of the two-component regulatory system EnvZ/OmpR involved in osmoregulation (Kenney and Anand 2020). | 3535147 | G**T**A - G**G**T (valine - glycine)  V241L |
| **rsgA** | Ribosome small subunit-dependent GTPase A | One of at least 4 proteins (Era, RbfA, RimM and RsgA/YjeQ) that assist in the late maturation steps of the functional core of the 30S ribosomal subunit. Binds the 30S subunit contacting the head, platform, and rRNA helix 44, which may assist the last maturation stages (Goto et al. 2011). | 4390754 | GC**C -** GC**A** (alanine - alanine)  A252A |

**5. Discussion**

One of the mutations obtained after variant calling was in the region coding small RNA RybA, position 852762. This small RNA regulates synthesis of aromatic amino acids under peroxide stress conditions (Gerstle et al. 2012b). Hence, we suggested the RybA gene unlikely to be involved in protection from antibiotics. Mutation in the ftsI gene is of great interest because this gene encodes penicillin-binding protein 3 — the protein required for division septum assembly (Weiss et al. 1997). It was shown that binding of beta-lactam antibiotics inhibited activity of the ftsI and led to cell lysis (Curtis et al. 1985). Since ampicillin is beta-lactam antibiotic, ftsI gene can be its target site. The replacement of a single amino acid in antibiotic binding site can result in a decrease of the affinity to ampicillin. Mengchen Li at al. in their study of E. coli ampicillin resistance mechanism showed that drug resistant strain had nonsynonymous mutations in several genes including ftsI which is consistent with our results (Li et al. 2019). Replacement of glutamine by leucine in acrB, a component of proton-dependent drug efflux pump, could result in more efficient pumping of ampicillin outside of the cell. Several previous studies have shown the importance of acrB in appearance of multi-drug resistance in Escherichia coli (Okusu, Ma, and Nikaido 1996; Li et al. 2019). Thus, changing the structure of this protein may contribute to ampicillin resistance. Yet another mutated gene could be responsible for resistance. The envZ gene is a sensor histidine kinase which modulates the diffusion through porins (Delcour 2009). Potentially the mutation in envZ could decrease permeability of outer membrane for ampicillin molecules and reduce the amount of antibiotic to harmless level. In addition, mutation in this gene also was mentioned by Mengchen Li et al. as one of the contributors to ampicillin resistance.   
One more mutation observed in our study was in GTPase A of a small ribosomal subunit and seemed to be uninvolved in resistance formation. Additionally, this mutation didn’t result in amino acid replacement.

**6. Conclusion**

Three possible mechanisms of resistance may be used by ampicillin resistance Escherichia coli strain K-12 substrain MG1655: modification of the target protein (ftsI), improved pumping of the antibiotic (acrB) and changing permeability of the outer membrane for ampicillin (envZ). Further topological determination of changed amino acids in the protein and examination of mutant protein affinity to ampicillin should be performed to check whether these are crucial mutations for resistance. Nevertheless, we can formulate some treatment recommendations based on the obtained knowledge. Antibiotics from a group of fluoroquinolones can be used because of another inhibition mechanism. These antibiotics interfere in DNA replication by inhibiting topoisomerase IV and DNA gyrase.

**References**

1. Davies, J.; Davies, D. Origins and evolution of antibiotic resistance. *Microbiol Mol Biol Rev* **2010**, *74*, 417–433, doi:10.1128/MMBR.00016-10.

2. Gerstle, K.; Klätschke, K.; Hahn, U.; Piganeau, N. The small RNA RybA regulates key-genes in the biosynthesis of aromatic amino acids under peroxide stress in *E. coli*. *RNA Biology* **2012**, *9*, 458–468, doi:10.4161/rna.19065.

3. Weiss, D.S.; Pogliano, K.; Carson, M.; Guzman, L.; Fraipont, C.; Nguyen‐Distèche, M.; Losick, R.; Beckwith, J. Localization of the *Escherichia coli* cell division protein FtsI (PBP3) to the division site and cell pole. *Molecular Microbiology* **1997**, *25*, 671–681, doi:10.1046/j.1365-2958.1997.5041869.x.

4. Curtis, N.A.C.; Eisenstadt, R.L.; Turner, K.A.; White, A.J. Inhibition of penicillin-binding protein 3 of *Escherichia coli* K-12. Effects upon growth, viability and outer membrane barrier function. *J Antimicrob Chemother***1985**, *16*, 287–296, doi:10.1093/jac/16.3.287.

5. Li, M.; Liu, Q.; Teng, Y.; Ou, L.; Xi, Y.; Chen, S.; Duan, G. The resistance mechanism of Escherichia coli induced by ampicillin in laboratory. *IDR* **2019**, *Volume 12*, 2853–2863, doi:10.2147/IDR.S221212.

6. Okusu, H.; Ma, D.; Nikaido, H. AcrAB efflux pump plays a major role in the antibiotic resistance phenotype of Escherichia coli multiple-antibiotic-resistance (Mar) mutants. *Journal of bacteriology* **1996**, *178*, 306–308, doi:10.1128/JB.178.1.306-308.1996.

7. Delcour, A.H. Outer membrane permeability and antibiotic resistance. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2009**, *1794*, 808–816, doi:10.1016/j.bbapap.2008.11.005.