Bioinformatics Institute

What causes antibiotic resistance?

Rubinova V., Tarabonda G.

Abstract One of the biggest problems in medicine nowadays is antibiotic resistance, and so many studies are trying to find the reasons for this harmful effect. This lab work has taken several steps to find out the mutations in E. coli strain K-12 substrain MG1655 and the relationship between these and the antibiotic resistance in this E. coli strain. We were identified 3 genes, which products (proteins) are most likely involved in the development of antibiotic resistance in the studied strain of E.coli: FtsI (Peptidoglycan D,D-transpeptidase FtsI), AcrB (Multidrug efflux pump subunit AcrB), EnvZ (Sensor histidine kinase EnvZ).

1 Introduction

Antibiotic resistance is a special case of antimicrobial resistance, when bacteria become resistant to antibiotics. Resistant microbes are more difficult to treat, require higher doses, or alternative drugs that may be more toxic. These approaches can also be more expensive [1].

Resistance occurs in one of several ways: natural resistance in certain types of bacteria through random mutations and/or exposure to the antibiotic; genetic mutation; or by acquiring resistance by some bacterial species from others (through horizontal gene transfer). In addition, antibiotic resistance of microorganisms can be created artificially by genetic transformation. For example, the introduction of artificial genes into the genome of a microorganism [2]. Resistance may appear spontaneously due to arbitrary mutations; or most often as a result of gradual accumulation over time, and due to the misuse of antibiotics or antimicrobials [3].

The study of mutations in genes in bacteria with antibiotic resistance is an important task for finding ways to overcome antibiotic resistance. The mutations found help to understand which proteins are affected, what role they play in the body, and the mechanism of antibiotic resistance.

In our work, we were looking for genes associated with antibiotic resistance in E.coli to identify the mechanism of antibiotic resistance in each case, and it will be able to make recommendations for using alternative antibiotics to treat each strain.

2 Methods

This section provides information about the methods that have been used in this study.

2.1 Filtering raw data

First, we obtained raw Illumina sequencing reads from shotgun sequencing of an ampicillin-resistant `E. coli` strain and loaded the reference sequence of the parental `E. coli` strain from NCBI FTP.

Before the filtering of the reads, we used the FastQC tool to inspect the raw sequencing data before we filtered the reads. However, we required high-quality results, so we turned to Trimmomatic, a trimming application. We have used this program to trim reads using a sliding window technique with window size 10 and average quality within the window 20, drop the read if it is below length 20, and chop bases off the start and end of a read if quality is below 20. Also we made these actions, but with the quality 30 and than used this data in further steps.

2.2 Aligning sequences to reference

There are many alignment programs available that were developed to meet these objectives. However, we used an aligner named BWA-MEM from the huge number of these that makes use of the Burrows-Wheeler Transform. Because BWA-MEM is designed for 'long' next-generation sequencing reads of 100 bp or more, which may contain several mutations, insertions, or deletions.

This program returns the SAM (Sequence Alignment Map) file, and to work with this type of file we used the samtools program, which can compress, sort, and index SAM files.

2.3 Mutation prediction

The next step is to determine which readings actually contain mutations and which ones are just sequencing mistakes. We employed the VarScan tool to accomplish this purpose.

Finally, we utilized the snpEff tool, which returns the vcf file listing all the impacts for each SNP, to comprehend the mutation effect.

3 Results

3.1 Data information

Raw Illumina sequencing data contains 455876 reads with a length of 101, but this data has bad mean quality. After using Trimmomatic with minimum quality 20, we got 444961 reads from 20 to 101. And after using Trimmomatic with a minimum quality of 30, we got 369988 reads of length 20-101.

After running FastQC, we get the statics for each metric. Below is the table that represents the statistics. A ✓ indicates that the data passed the statistic test; a × indicates that the data did not pass the statistic test; and ! indicates that the data passed the statistic test but the results were not satisfactory.

Table 1. The statistics on various data

Statistics name	FR	RR	F20	R20	F30	R30
Per base sequence quality	✓	✓	✓	✓	✓	✓
Per tile sequence quality	×	×	×	!	×	×
Per sequence quality scores	×	!	✓	✓	✓	✓
Per base sequence content	✓	✓	!	!	!	!
Per sequence GC content	!	!	!	!	!	✓
Per base N content	!	!	✓	✓	✓	✓
Sequence length distribution	✓	✓	!	!	!	!
Sequence duplication levels	✓	✓	✓	✓	✓	✓
Overrepresented sequences	✓	✓	✓	✓	✓	✓
Adapter content	✓	✓	✓	✓	✓	✓

FR - forward read data, RR - reverse read data, F20 - forward read after using Trimmomatic with minimum quality 20, R20 - reverse read after using Trimmomatic with minimum quality 20, F30 - forward read after using Trimmomatic with minimum quality 30, R30 - reverse read after using Trimmomatic with minimum quality 30

3.2 Mutation prediction

After aligning reads and using VarScan and snpEff tools, we got several mutations that may cause different changes in proteins. These mutations are shown in the table below.

Table 2. The mutations in *E. Coli* strain K-12

Mutation position	Reference nucleotide	Reference nucleotide
93043	C	G
482698	T	A
852762	A	G
1905761	G	A
3535147	A	C
4390754	G	T

4 Discussion

Sequencing data from a strain of E. coli resistant to the antibiotic ampicillin were analyzed. Reads obtained after sequencing were aligned to the reference bacterial genome. And several mutations were found, presumably involved in the formation of antibiotic resistance in the strain. After that, we studied the genes in which

the mutations occurred, explored their products (proteins), the function that they perform in the cell and identified 3 genes (and their main proteins), which are most likely involved in the development of antibiotic resistance in the studied strain of E.coli.

FtsI. Peptidoglycan D,D-transpeptidase FtsI is an essential cell division protein that catalyzes cross-linking of the peptidoglycan cell wall at the division septum. Activity regulation: inhibited by beta-lactam antibiotics such as penicillin, moenomycin, macarbomycin, furazlocillin and piperacillin. Antibiotics inhibit the activity by binding to the catalytic serine [4]. So, if a mutation occurs in the gene encoding this protein (FtsI), the mutation changed the antibiotic's target so the antibiotic couldn't bind it anymore. Therefore, bacterial cells divide well and are not sensitive to the antibiotic.

AcrB. AcrA-AcrB-AcrZ-TolC is a drug efflux protein complex with broad substrate specificity that uses the proton motive force to export substrates. Multidrug efflux pump subunit AcrB involved in contact-dependent growth inhibition (CDI), acts downstream of BamA, the receptor for CDI. Its role in CDI is independent of the AcrA-AcrB-TolC efflux pump complex [5]. A mutation occurs in the gene AcrB and the antibiotic that binds to this protein stops binding to the target. Therefore, it ceases to inhibit the outflow of drugs from the cell. As a result, all the drug leaves the cell and the bacterium does not react to it and becomes resistant. Th[6]is mutation plays an important role in the formation of antibiotic resistance in this strain.

EnvZ. Sensor histidine kinase EnvZ is a member of the two-component regulatory system EnvZ/OmpR involved in osmoregulation (particularly of genes ompF and ompC) as well as other genes. EnvZ functions as a membrane-associated protein kinase that phosphorylates OmpR in response to environmental signals; at low osmolarity OmpR activates ompF transcription, while at high osmolarity it represses ompF and activates ompC transcription [7] Bacterial histidine kinases (HKs) are part of the bacterial two-component system (TCS), the key pathway for signal transduction in bacteria that regulate several processes like secretion systems, virulence, and antibiotic resistance [8]. The antibiotic does not bind to the modified protein (the product of the gene in which the mutation has occurred), so it does not block signaling pathways in bacteria. Thus, the strain with the mutation becomes antibiotic resistant. Also there is a research where authors determined that resistance to antimicrobial peptide arises through mutations in the histidine kinase EnvZ [9].

Moreover, there is research that claims that combinations of mutations in envZ, ftsI, mrdA, acrB and acrR can cause high-level carbapenem resistance in Escherichia coli [10].

The studied strain is resistant to the antibiotic ampicillin, so there is no point in prescribing this antibiotic to the patient infected with this strain of E coli. In addition to ampicillin, E.coli is sensitive to antibiotics amoxicillin and tetracycline. So we can suggest these alternative antibiotics with other targets that may be useful in treating a patient with this strain of E.coli.

References

[1] Dadgostar P. Antimicrobial resistance: implications and costs // Infection and Drug Resistance. 2019. Vol. 12.

[2] Cesur S., Demiroz A.P.Antibiotics and the Mechanisms of Resistance to Antibiotics // Medical Journal of Islamic World Academy of Sciences. 2013. Vol. 21, 4.

[3] Chinedum I.E. Microbial resistance to antibiotics // African Journal of Biotechnology. 2005. Vol. 4, 13 SPEC. ISS.

[4] ftsI - Peptidoglycan D,D-transpeptidase FtsI - Escherichia coli (strain K12) — UniProtKB — UniProt [Electronic resource]. URL: https://www.uniprot.org/uniprotkb/P0AD68/entry (accessed: 28.10.2022).

[5] acrB - Multidrug efflux pump subunit AcrB - Escherichia coli (strain K12) — UniProtKB — UniProt [Electronic resource]. URL: https://www.uniprot.org/uniprotkb/P31224/entry (accessed: 28.10.2022).

[6] Chetri S. et al. AcrAB-TolC efflux pump system plays a role in carbapenem non-susceptibility in Escherichia coli // BMC Microbiol. 2019. Vol. 19, 1.

[7] envZ - Sensor histidine kinase EnvZ - Escherichia coli (strain K12) — UniProtKB — UniProt [Electronic resource]. URL: https://www.uniprot.org/uniprotkb/P0AEJ4/entry (accessed: 28.10.2022).

[8] Matsushita M., Janda K.D. Histidine kinases as targets for new antimicrobial agents // Bioorganic and Medicinal Chemistry. 2002. Vol. 10, 4.

[9] Georgieva M, Heinonen T, Hargraves S, Pillonel T, Widmann C, Jacquier N. The EnvZ/OmpR Two-Component System Regulates the Antimicrobial Activity of TAT-RasGAP317-326 and the Collateral Sensitivity to Other Antibacterial Agents. // Microbiol Spectr. 2022.

[10] Adler M. et al. Combinations of mutations in envZ, ftsI, mrdA, acrB and acrR can cause high-level carbapenem resistance in Escherichia coli // Journal of Antimicrobial Chemotherapy. 2016. Vol. 71, 5.

Supplementary

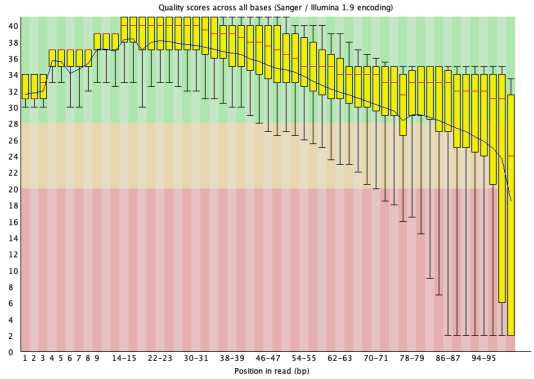

Fig.1a. Quality of forward read data according to FastQC

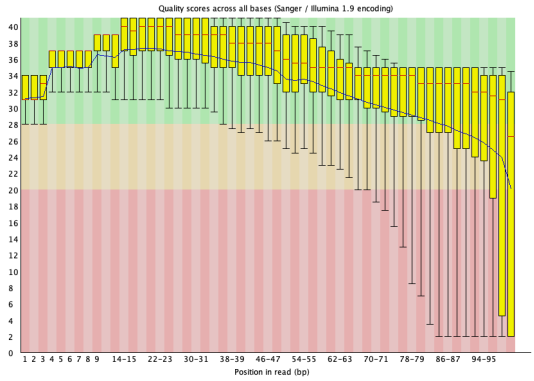

Fig.1b. Quality of reverse read data according to FastQC

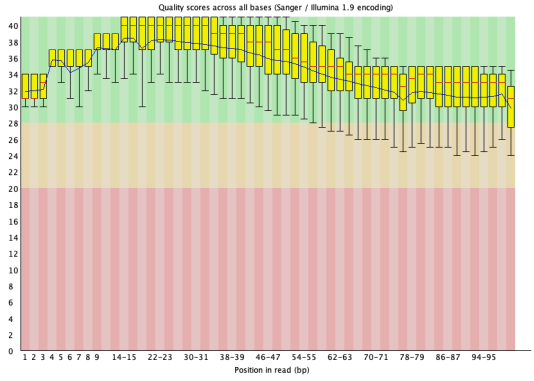

Fig.2a. Quality of forward read data after using Trimmomatic with quality 20 according to FastQC

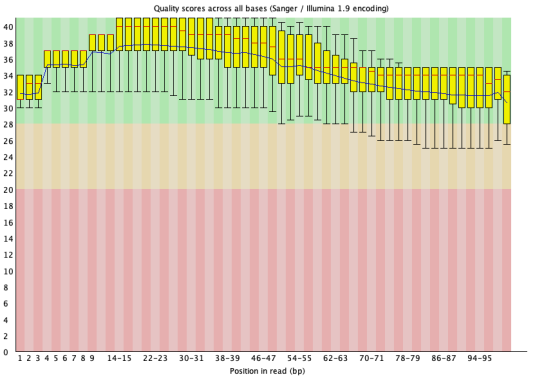

Fig.2b. Quality of reverse read data after using Trimmomatic with quality 20 according to FastQC

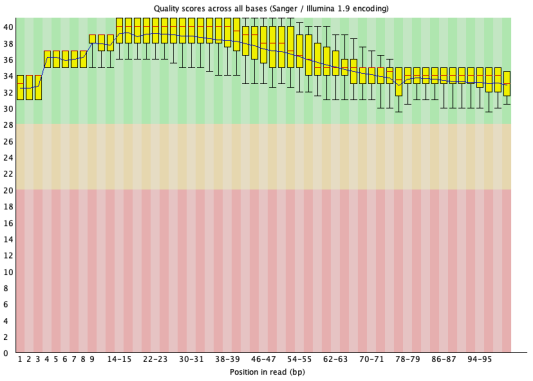

Fig.3a. Quality of forward read data after using Trimmomatic with quality 30 according to FastQC

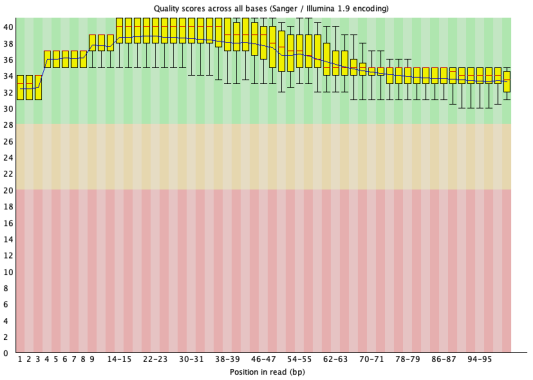

Fig.3b. Quality of reverse read data after using Trimmomatic with quality 30 according to FastQC