Causes of ampicillin resistance of *E.coli* K-12 substrain

A. Zhurakovskaya, V. Zvezdin

October 2022

Abstract

The antibiotic resistance acquired by bacteria might interfere with the treatment of many infectious diseases. Hence, it is important to discover its causes. In this study, we used the NGS Illumina data to identify the SNP locations of the resistant E.coli strain and discover the function of impacted genes. We found out mutations in several genes directly involved in penicillin resistance (FtsI, AcrB) as well as those of metabolic regulators (MntP, small RNA RybA), and suggested their possible role in maintaining resistance.

Introduction

Bacteria utilize many different mechanisms to defend themselves against the antibiotics. Among them, there are enhanced efflux of the drug or the lowered intake of it in the first place, drug molecule utilisation in the cell, modification of the drug target, overproduction of competitive binders etc. Identifying its underlying causes might allow to create new drugs and modify the existing ones or create the supplements to make them more efficient. Such phenomena can be effectively studied on non-hazardous objects such as some E.Coli strains. The one which was used in this study, E. Coli K-12, has a reference genome of high quality and is a well established model organism otherwise.

The involvement of many resistance mechanisms may be traced to mutations in the genome that alter functional proteins or regulatory sequences. Using the data acquired from the NGS reads of resistant E.coli substrain, we identified several indels and SNPs when compared to a reference genome. We decided not to analyze the indels because they often cause the frame shifts and are rarely, if ever, beneficial. Thus, the goal of our study was to annotate the SNPs and propose their role in increasing the resistance to the antibiotics.

Methods

Reference genome and annotation of E.coli K-12 strain were obtained from NCBI databases. Raw Illumina sequencing reads of ampicillin resistant strain of E.coli were obtained from figshare [1].

FastQC program was used for inspecting quality of reads [2]. Poor per base sequence quality was reported for both the forward and reverse reads. To counter this problem, the reads were filtered using Trimmomatic tool [3]. For filtering, the following parameters were used: paired reads, phred33 quality scale, removal of leading and trailing bases with quality below 20, sliding window with size 10 for quality below 20, removal of reads below length 20. FastQC analysis was redone on filtered reads.

Then, the reference genome was indexed and filtered reads were aligned to it using BWA [4] and compressed into .bam file with Samtools [5]. After that the .bam file was also sorted and indexed with Samtools.
Next, the .mpileup file was created with Samtools, and the variant calling was performed on it using the VarScan program [6]. The threshold for variation frequency in reads was set at 75% and positions filtered accordingly (–min-var-frequency = 0.75, –variants). The output VCF file was used for automatic SNP annotation with SnpEff program [7].

Results

The number of raw reads, as well as those of them successfully filtered is presented in the table 1 with the respective per base sequence quality.

Table 1. The amount of reads and their average quality.

	Number of reads	Per base sequence quality
Forward	455 876	2 - 40
Reverse	455 876	2 - 40
Forward filtered	446 259	28 - 40
Reverse filtered	446 259	28 - 40

891649 of these reads (99.87%) were mapped to a genome.
The results of the annotations are given in the table 2. Here we also present the products of the genes where we found the SNPs, as well as the supposed function of these genes (see the discussion).:))))))))))))))))

Table 2. Genes with SNP mutations and their functions

Gene	Mutation	Position	Substitution	Coding	Product	Function
FtsI	missense	93043	Ala554Gly	protein	peptidoglycan transpeptidase	Cross-linking of the peptidoglycan cell wall at the division septum
AcrB	missense	482698	Gln569Leu	protein	drug efflux protein complex	Binds and expels drugs
RybA	upstream gene variant	852762	—	sRNA	small RNA	Regulator of GlnH (ABC transporter gene, glutamine upake)
MntP	missense	1905761	Gly25Asp	protein	manganese efflux pump	Imports or exports manganese
EnvZ	missense	3535147	Val241Gly	protein	histidine kinase	Sensor-transmittor involved in osmoregulation
RsgA	synonymous variant	4390754	Ala252Ala	protein	small GTFase	Maturation of the 30s ribosome subunit

Discussion

We could identify 6 SNP sites, one of which occurs in a small RNA gene. In databases, this small RNA gene is described as the regulator of metabolic processes, in particular in the amino acid synthesis. The annotation with the highest effect impact predicts it to be a regulator of GlnH gene, involved in glutamine uptake, expression of which is higher in resistant bacteria [8]. The cited study discusses the fitness cost of acquiring the antibiotic resistance. Therefore it may be theorised that a mutation in such regulatory sRNA can lead to an enhanced production of proteins which helps bacteria survive under the stress. Additionally, some proteins are directly involved in providing resistance and bacteria need to produce them in higher quantities which might be limited by the uptake or synthesis of a particular amino acid.

In fact, all the other SNPs are located in the protein-coding sequences (CDS) and 4 out of 5 SNPs in proteins are non-synonymous, being the sites of missense mutations. It might hint at their importance for obtaining resistance.

One of such missense mutations is the mutation in FtsI, a gene coding for penicillin-binding protein 3 (PBP3) that is involved in cross-linking of peptidoglycans in a cell wall of a dividing cell [9]. Mutations in this gene are known for causing alterations in this protein in *Haemophilus influenzae* [10]. We may suggest that such mutations impair the binding of a penicillin molecule to a protein. As such, the drug does not affect its cell-building function and the division of bacteria is stabilised.

Another missense mutation occurs in gene AcrB, whose product is a part of a complex that binds different xenobiotics and removes them from the cell [11]. If this mutation promotes the efflux then it directly improves the bacteria resistance to antibiotics. It might also be associated with multiple drug resistance, thus, it is interesting to check if the resistance to other types of antibiotics is increased in this strain too.

Mutation in MntP might also be significant for resistance of *E.coli* because it can lead to a higher manganese export, which lowers the level of manganese inside the cell. This ensures that the process of RybA production won't be suppressed, as shown on the Figure 1 [12]. Gene EnvZ, where we found yet another mutation, might be involved in a similar regulation, encoding a protein which functions as the sensor of osmotic concentration.

These findings suggest several mutations that utilize the variety of mechanisms that help bacteria to resist antibiotics. The observed resistance might be caused by a combination of them all, rather than any one mutation in particular.

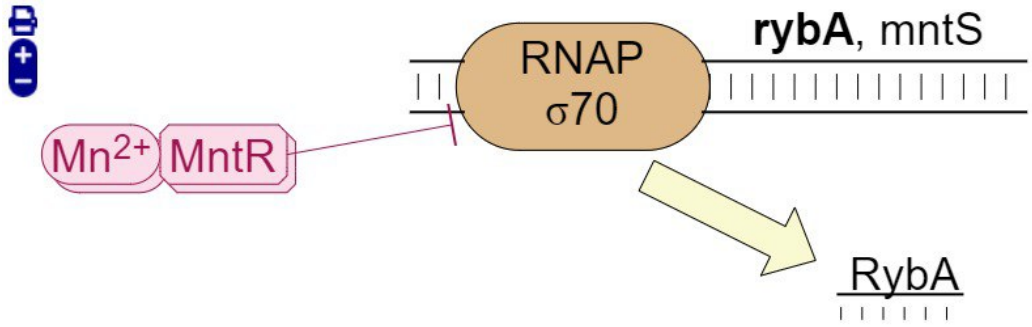

Figure 1: The regulation of RybA gene by manganese. If manganese is largely removed from the cell, this leads to RybA upregulation. RybA then upregulates metabolic genes, which positively affect the resistance.

References

[1] Mike Raiko. Amp res Ecoli data. 2020.

[2] Fastqc, Jun 2015.

[3] Anthony M. Bolger, Marc Lohse, and Bjoern Usadel. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15):2114–2120, 04 2014.

[4] Heng Li and Richard Durbin. Fast and accurate short read alignment with burrows–wheeler transform. *bioinformatics*, 25(14):1754–1760, 2009.

[5] Petr Danecek, James K Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin O Pollard, Andrew Whitwham, Thomas Keane, Shane A McCarthy, Robert M Davies, and Heng Li. Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2), 02 2021. giab008.

[6] Daniel Koboldt, Qunyuan Zhang, David Larson, Dong Shen, Michael Mclellan, Ling Lin, Christopher Miller, Elaine Mardis, Li Ding, and Richard Wilson. Varscan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome research*, 22:568–76, 03 2012.

[7] P. Cingolani, A. Platts, M. Coon, T. Nguyen, L. Wang, S.J. Land, X. Lu, and D.M. Ruden. A program for annotating and predicting the effects of single nucleotide polymorphisms, snpeff: Snps in the genome of drosophila melanogaster strain w1118; iso-2; iso-3. *Fly*, 6(2):80–92, 2012.

[8] Chang-Ro Lee, Jung Hun Lee, Kwang Park, Byeong Jeong, and Sang Lee. Quantitative proteomic view associated with resistance to clinically important antibiotics in gram-positive bacteria: A systematic review. *Frontiers in microbiology*, 6:828, 08 2015.

[9] Mohamed Attaibi and Tanneke Blaauwen. An updated model of the divisome: Regulation of the septal peptidoglycan synthesis machinery by the divisome. *International Journal of Molecular Sciences*, 23:3537, 03 2022.

[10] D. Skaare, A.-G. Allum, I.L. Anthonisen, A. Jenkins, A. Lia, L. Strand, Y. Tveten, and B.-E. Kristiansen. Mutant ftsi genes in the emergence of penicillin-binding proteinmediated -lactam resistance in haemophilus influenzae in norway. *Clinical Microbiology and Infection*, 16(8):1117–1124, 2010.

[11] Marlen Adler, Mehreen Anjum, Dan I. Andersson, and Linus Sandegren. Combinations of mutations in envZ, ftsI, mrdA, acrB and acrR can cause high-level carbapenem resistance in Escherichia coli. *Journal of Antimicrobial Chemotherapy*, 71(5):1188–1198, 02 2016.

[12] Julia Martin, Lauren Waters, Gisela Storz, and James Imlay. The escherichia coli small protein mnts and exporter mntp optimize the intracellular concentration of manganese. *PLoS genetics*, 11:e1004977, 03 2015.