Case report: Investigation of Escherichia coli ampicillin resistance using NGS data

Ogurtsova A. Bagrova O.

October 27, 2023

Abstract

This study focuses on the case of pathogenic E. coli k12 strain that has developed resistance to ampicillin. Through analysis of next-generation sequencing (NGS) data, we revealed that resistance have been caused by mutations in the genes ftsI and acrB. They encode enzymes: transpeptidase responsible for synthesizing peptidoglycans and multidrug pump, respectively. Finally, we propose potential strategy to increase treatment effectiveness.

1. Introduction

E.coli is one of the main causes of hospital-acquired infections in humans. Additionally, E. coli is a common resident in both human and animal gastrointestinal tracts, serving as an indicator of fecal contamination in food. Beta-lactam antibiotics find extensive application in both human and veterinary medicine for treating infections [1]. Ampicillin belongs to the family of beta-lactam antibiotics and it inhibits E.coli cell wall synthesis [2]. This widespread use of antibiotics could be associated with the selection of antibiotic resistance mechanisms in pathogenic and nonpathogenic isolates of E. coli [1]. Resistance to beta-lactam antimicrobial agents like ampicillin in E. coli primarily arises through the action of beta-lactamases, enzymes that facilitate the hydrolysis of the ampicillin's beta-lactam ring, rendering the antibiotic inactive. The other possible mechanisms are typical of the majority bacteria and include changes in membrane permeability, plasmid-mediated resistance, efflux pumps, protein expression and modification of ampicillin-binding proteins [3].

In our study, we use contemporary bioinformatic tools to analyse pathogenic E. coli k12 strain shotgun sequencing data. The aim of our study was to find the cause of resistance of the bacterium.

2. Methods

In this case report, we performed we dentification of possible causes of antibiotic resistance. The overall pipeline included reads filtering, aligning sequences, variant calling and automatic

SNP annotation.

Briefly,reference sequence of the parental (not resistant to antibiotics) E. coli strain and fastq files of resistant E. coli strain were given. Quality control was performed with fastq [4], Sample reads were filtered by Trimmomatic PE tool [5] with threshold coverage 20 and phred33 flag, then they were aligned to indexed reference by BWA tool[6] with mem algorithm. Aligned reads were sorted and indexed by Samtools[7], variant calling was performed by VarScan[8] with p-value threshold 0.01 and minimal variant frequency 0.4 (mpileup file made with Samtools). Automatic SNP annotation was performed by snpEff. For variant effect prediction IGV browser was used[9]. Information on altered proteins was obtained from the UniProt and RCSB PDB. All used commands and tools versions can be viewed on github (lab journal).

3. Results

A total of 455876 pairs of reads were initially obtained. After trimming with TrimomaticPE, 97.89% of reads were preserved, which were used for further analysis. Sequence length was in range from 20 to 101 nucleotides (Figure 1). GC-content was 50% as expected (50.5% in reference GCF_000005845.2). The quality (Q score) of the reads ranged from 24 to 42 points. After quality control paired reads were mapped to the reference and scanned with VarScan. As a result, five positions were found that contained SNP (Table 1). Four of them missense mutations (fstI, acrB, mntP, envZ) and one upstream gene variant (glnH).

Table 1. Genes and characteristics of the mutations revealed using NGS.

Gene	Protein	Genomic position	Variant (nucl)	Variant (prot)	Effect
fstI	Peptidoglycan D,D-transpeptidase FtsI	93043	C>G	Ala544Gly	missense
acrB	Multidrug efflux pump subunit AcrB	482698	T>A	Gln569Leu	missense
glnH	Glutamine-binding periplasmic protein	852762	A>G		upstream gene variant
rsgA	Maturation of the 30S subunit of the ribosome	4390754	G>T	Ala252Ala	synonymous
envZ	Sensor histidine kinase EnvZ	3535147	A>C	Val241Gly	missense

4. Discussion

The very first provided by SNP's annotation is a substitution in the ftsI gene (C>G) which is annotated as a missense variant with a moderate impact on protein's function and leads to the substitution of alanine for glycine in position 544.

The ftsl gene encodes a protein named essential cell division protein FtsI or penicillin-binding protein 3. According to EcoCyc, FtsI functions as peptidoglycan transpeptidase and is essential in the cell division process. Binding of beta-lactam antibiotics to FtsI inhibits FtsI activity and is lethal.

It has been previously shown that the substitution of S to A or T in position 307 causes the protein to stop binding ampicillin [10].

Even though there is no information on ftsI mutagenesis in position 544 neither on EcoCyc nor in UniProt or in literature, it is known that the the C-terminal 349 amino acids domainncontains the penicillin-binding region and the position 544 is also a part of penicillin-binding region, which is relatively not far from position 307, mutations in which lead to ampicillin resistance.

Therefore, **one of the possible mechanisms of ftsI - mediated ampicillin resistance** is that the substitution p.Ala544Gly may somehow change the spatial structure of the target protein so the ampicillin can't bind anymore.

Another possible mechanism of ftsI - mediated ampicillin resistance raises from the information in Marlen Adler's study, where it has been shown that mutations in the ftsI gene alone did not increase antibiotic resistance, whereas ftsI and envZ gene mutations increased the MIC of antibiotics multiple times [11]. And, as will be seen further from the information in the discussion section, this is exactly the case when mutations in these two genes occur simultaneously. It is also important to note that the accumulation of mutations in ftsI can also lead to Ceftriaxone resistance, so it wouldn't be suggested to the doctor to switch to cephalosporin antibiotics.

The following SNP occurred in gene envZ (T>G) which is classified as a missense variant and according to SnpEff has a moderate impact on protein structure. However the information on EcoCyc reveals that there exists a EnvZ/OmrR signal transduction system, which plays a crucial role in differential regulation of the outer membrane diffusion pores OmpF and OmpC and their permeability.

Many of the alterations in outer membrane permeability are often associated with increased levels of antibiotic efflux [12]. Molecular dynamics simulations also demonstrated the translocation of ampicillin through OmpF from E. coli [13].

Thus the potential mechanism of envZ - mediated ampicillin resistance is that SNP in encvZ leads to disruption in outer membrane permeability for ampicillin and increases ampicillin efflux. A similar mechanism has been already demonstrated for S. Enteritidis [14].

Finally consider a mutation in the acrB gene. The acrB gene encodes the protein from AcrA-AcrB-AcrZ-TolC, which is a drug efflux complex with broad substrate specificity. It uses the proton motive force to export substrates. Gene acrB have been shown to be associated with ampicillin resistance in E.coli [15 - 17]. Although ampicillin does not bind directly

to the amino acid at position 569 [18], due to mutations in this gene, resistance by the efflux pumpsmechanism may occur. In this case it is acrB - mediated ampicillin resistance. The found position (Gln569Leu) is located near the surface of the protein. Replacing a polar amino acid with a hydrophobic one can change the properties of the protein.

This study was provided to determine the possible mechanisms of ampicillin resistance in E.coli based on NGS-sequence. As a result, **we have proposed 3 possible mechanisms: changes in target (ftsI), decreased permeability (envZ) and efflux pumps (acrB)**.

References

[1] L. Briñas, M. Zarazaga, Y. Sáenz, F. Ruiz-Larrea, and C. Torres, "Beta-lactamases in ampicillin-resistant Escherichia coli isolates from foods, humans, and healthy animals," *Antimicrob Agents Chemother*, vol. 46, no. 10, pp. 3156-3163, Oct. 2002. doi: 10.1128/AAC.46.10.3156-3163.2002.

[2] K. Bush and P.A. Bradford, "-Lactams and -Lactamase Inhibitors: An Overview," *Cold Spring Harb Perspect Med*, vol. 6, no. 8, a025247, Aug. 2016. doi: 10.1101/cshperspect.a025247.

[3] J.M. Munita and C.A. Arias, "Mechanisms of Antibiotic Resistance," *Microbiol Spectr*, vol. 4, no. 2, Apr. 2016. doi: 10.1128/microbiolspec.VMBF-0016-2015. PMID: 27227291.

[4] S. Andrews, "FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]," available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/, 2010.

[5] A.M. Bolger, M. Lohse, and B. Usadel, "Trimmomatic: A flexible trimmer for Illumina Sequence Data," *Bioinformatics*, 2014, btu170.

[6] H. Li, "Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM," arXiv preprint arXiv:1303.3997v2 [q-bio.GN], 2013.

[7] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, et al., "The Sequence Alignment/Map format and SAMtools," *Bioinformatics*, vol. 25, no. 16, pp. 2078-2079, 2009.

[8] D.C. Koboldt, Q. Zhang, D.E. Larson, D. Shen, M.D. McLellan, L. Lin, C.A. Miller, E.R. Mardis, L. Ding, and R.K. Wilson, "VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing," *Genome Res*, vol. 22, no. 3, pp. 568-576, Mar. 2012. doi: 10.1101/gr.129684.111.

[9] P. Cingolani, A. Platts, L.L. Wang, M. Coon, T. Nguyen, L. Wang, 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*, vol. 6, no. 2, pp. 80-92, 2012.

[10] N. Houba-Hérin, H. Hara, M. Inouye, Y. Hirota, "Binding of penicillin to thiol-penicillin-binding protein 3 of Escherichia coli: identification of its active site," *Mol Gen Genet*, vol. 201, no. 3, pp. 499-504, 1985. doi: 10.1007/BF00331346. PMID: 3911028.

[11] M. Adler, M. Anjum, D.I. Andersson, et al., "Combinations of mutations in envZ, ftsI, mrdA, acrB and acrR can cause high-level carbapenem resistance in Escherichia coli," *J Antimicrob Chemother*, 2016; 71: 1188–98.

[12] A.H. Delcour, "Outer membrane permeability and antibiotic resistance," *Biochim Biophys Acta*, vol. 1794, no. 5, pp. 808-16, May 2009. doi: 10.1016/j.bbapap.2008.11.005. Epub 2008 Nov 27.

[13] A. Kumar, E. Hajjar, P. Ruggerone, M. Ceccarelli, "Molecular simulations reveal the mechanism and the determinants for ampicillin translocation through OmpF," *J Phys Chem B*, vol. 114, no. 29, pp. 9608-9616, 2010. doi: 10.1021/jp9110579.

[14] D. Ko, S.H. Choi, "Mechanistic understanding of antibiotic resistance mediated by EnvZ/OmpR two-component system in Salmonella enterica serovar Enteritidis," *J Antimicrob Chemother*, vol. 77, no. 9, pp. 2419-2428, Aug 2022. doi: 10.1093/jac/dkac223.

[15] M. Li, Q. Liu, Y. Teng, et al., "The resistance mechanism of Escherichia coli induced by ampicillin in laboratory," *Infection and Drug Resistance*, vol. 12, pp. 2853–2863, 2019. doi: 10.2147/IDR.S221212.

[16] F.S. Kaczmarek, T.D. Gootz, F. Dib-Hajj, et al., "Genetic and Molecular Characterization of -Lactamase-Negative Ampicillin-Resistant Haemophilus influenzae with Unusually High Resistance to Ampicillin," *Antimicrobial Agents and Chemotherapy*, vol. 48, no. 5, pp. 1630–1639, 2004. doi: 10.1128/AAC.48.5.1630-1639.2004.

[17] A. Marlen, M. Mehreen, D.I.A. Dan, L. S., "Combinations of mutations in envZ, ftsI, mrdA, acrB and acrR can cause high-level carbapenem resistance in Escherichia coli," *Journal of Antimicrobial Chemotherapy*, 2016; 71(5): 1188–1198. https://doi.org/10.1093/jac/dkv475.

[18] S. Törnroth-Horsefield, P. Gourdon, R. Horsefield, et al., "Crystal Structure of AcrB in Complex with a Single Transmembrane Subunit Reveals Another Twist," *Structure*, vol. 15, no. 12, pp. 1663–1673, 2007. doi: 10.1016/j.str.2007.09.023.