Which genes provide resistance to ampicillin in *Escherichia coli*?

Nikita Vyatkin, Polina Guseva

October 28, 2022

Abstract

As antibiotic resistance becomes more common, usual drugs stop working. In this paper, Illumina shotgun sequenced reads from the ampicillin-resistant strain of *Escherichia coli* are selected by quality and aligned to the K-12 substrain reference genome. Established five mutations with acceptable coverage are coupled to the *E. coli* genes: missense (*ftsI*, *acrB*, *envZ*), synonymous (*rsgA*), upstream gene variant (*rybA*). Modifications in 1) *ftsI* could decrease its affinity to ampicillin preventing cell wall degradation and osmolysis; 2) *acrB* could efflux with high affinity; 3) *envZ* could affect membrane permeability. To avoid these mechanisms, azithromycin might be prescribed as an alternative.

Introduction

Antibiotic resistance among bacteria is a significant problem in treating infectious diseases, especially with last-resort options. It is crucial for researchers and medical professionals to understand resistance mechanisms such as 1) alternation of a target molecule, 2) destruction of a drug, 3) decreasing its concentration by pumping out and/or lowering permeability [1].

Beta-lactams and penicillins in particular are frequently used against *Escherichia coli* infections. As all of them, ampicillin inhibits the production of the bacterial cell wall by blocking transpeptidase (PBP, or penicillin-binding protein) [2]. This enzyme creates cross-linkages between peptidoglycans, so its dysfunction leads to cell wall breakages and death by osmolysis [3].

Our main hypothesis is that resistant *E. coli* are unaffected by ampicillin by changing target PBP, destroying the drug, and/or tailoring pumps with lowering membrane permeability. In this project, we investigate: 1) aligning of reads from the resistant strain to the reference *E. coli* genome, 2) locations of mutations with sufficient coverage, 3) and their role. The current research has great importance in the identification of relevant mutations and advising working antibiotics to mitigate resistant bacteria as well as designing more effective medications in future.

Methods

The raw data is obtained by Illumina shotgun sequencing of an *E. coli* strain that is resistant to the antibiotic ampicillin [4]. As for the reference, the genome of *E. coli* strain K-12 substrain MG1655 (Genbank, ID #167) is used.

To avoid false positive results, the reads are filtered by quality. For that, the poor-quality ($<$ 20 by autodetected Phred-33 score) ends are cut in the paired-end mode (`TrimmomaticPE`) using the 10-nucleotide sliding window with mean quality 20 by Trimmomatic 0.36 [5]. Following, processed reads less than 20 bp are also removed.

To recognise the impact of mutations, data is aligned to the reference genome and proven variants are linked to possible changes in the bacteria's functioning. Aligning to the indexed by the Burrows-Wheeler transform (`bwa index`) genome is performed (`bwa mem`) by bwa 0.7.17 [6]. The obtained results are compressed and sorted (`samtools view -b`, `samtools flagstat`, `samtools sort`, `samtools index`, `samtools mpileup -f`) for further quicker search by samtools 1.7-1 [7]. To identify actual mutations, single nucleotide polymorphisms (SNPs) are chosen with at least 50% coverage (`varscan mpileup2snp`) using VarScan 2.4.3 [8]. Automatic SNP annotation is performed (`snpEff build -genbank -v`, `snpEff ann`) with all effects described by snpEff 5.1d [9]. If the opposite is not stated, the default settings are used.

Results

In total, there are 455 876 raw reads each of 101 bp length both for forwards and reverse sequences 1 with the reference genome of 4.64 Mb. Generally, the quality of reads is above 30 by the Phred-33 scale with slight deterioration at the ends (see the FastQC report 1 [10]). A few tiles mostly in the forward reads have insufficient sequence quality. Yet after removing low-quality reads the number of them is reduced by less than 3%, while the rest is completely mapped to the reference 1. The piling up of all mutations has 100% coverage, i.e. all reads confirm the found transformations.

Table 1: Number of reads

Raw data	**After trimming**	**Aligned**	**Total loss, %**
455 876	446 388	444 279	2.54

Table 2: Description of the located single nucleotide polymorphisms (SNPs)

Gene	Codon changes	Amino acid changes	Mutation	Location, bp	Coding	Product
ftsI	GCC → GGC	A → G	missense	93 043	protein	peptidoglycan D,D-transpeptidase
acrB	CAG → CTG	Q → L	missense	482 698	protein	multidrug efflux pump RND permease
rybA	—	—	upstream gene variant	852 762	ncRNA	small RNA
envZ	GTA → GGA	V → G	missense	3 535 147	protein	sensor histidine kinase
rsgA	GCC → GCA	A → A	synonymous variant	4 390 754	protein	ribosome small subunit-dependent GTPase A

Overall, 5 mutations are detected: 3 missense, 1 synonymous, and 1 upstream gene variant 2. All of them are single nucleotide polymorphisms (SNPs). Moreover, these mutations can have not only a direct impact but indirectly influence other genes (see the snpEff report 3).

Discussion

All three missense mutations are associated with resistant *E. coli* to ampicillin [2] and other beta-lactams [11]. The synonymous one may potentially affect indirectly [12]. Yet it can also be just a founder effect. Whereas, small RNAs regulate gene expression contributing to antimicrobial resistance [13].

ftsI encodes a variant of D-alanyl-D-alanine transpeptidase also known as PBP_3. Penicillin antibiotics block this enzyme by making a direct bound in its active site [14], [15]. We assume that altering the gene may lead to the production of a protein with a modified active site that does not bind to ampicillin. ***acrB*** codes an efflux transporter that has increased affinity to ampicillin [16] as well as to other antibiotics [17]. So, its modification could tailor the enzyme binding more precisely. ***rybA*** controls the TyrR regulon which metabolises among others aromatic amino acids [18]. Furthermore, they act as precursors to enterobactin and ubiquinone protecting from oxidative stress. The last one is proven to be triggered by beta-lactams [19]. ***envZ*** regulates the permeability of the outer membrane by controlling the translation of ompC and ompF genes and by that causing porin deficiency [20]. Thus, the kinase with altered ATPase activity affects membrane permeability to beta-lactams [21]. ***rsgA*** (*yjeQ*) involves in ribosome biogenesis of 30S subunit [22], [23]. Therefore, it can regulate protein production. This synonymous mutation does not change the protein itself, yet it can alter the transcription assembly of the product. Without this gene *E. coli* grows slower in the presence of beta-lactams [24]. All described characteristics could be a reason for the evolutionary selection of such mutations for ampicillin resistance and have a combined effect.

Usually the other options for treating *E. coli* infections are either carbapenems or fluoroquinolones, yet they are also affected by mutations in the AcrAB pump and permeability proteins [25]. Therefore, our recommendation will be to use azithromycin as another one of the most common antibiotics [26], since the mentioned genes are not linked to its resistance [27], [28]. Unfortunately, there is no guarantee that in this case bacteria will not adapt to this drug too.

References

[1] B. Alberts, *Molecular biology of the cell*. 6 ed., 2017.

[2] M. Li, Q. Liu, Y. Teng, L. Ou, Y. Xi, S. Chen, and G. Duan, "The resistance mechanism of Escherichia coli induced by ampicillin in laboratory," *Infection and drug resistance*, vol. 12, p. 2853, 2019.

[3] M. Nguyen-Distèche, M. Leyh-Bouille, and J.-M. Ghuysen, "Isolation of the membrane-bound 26 000-M r penicillin-binding protein of Streptomyces strain K15 in the form of a penicillin-sensitive d-alanyl-d-alanine-cleaving transpeptidase," *Biochemical Journal*, vol. 207, no. 1, pp. 109–115, 1982.

[4] M. Raiko, "Amp_res_Ecoli_data," 7 2020. doi:10.6084/m9.figshare.10006541.v3.

[5] A. M. Bolger, M. Lohse, and B. Usadel, "Trimmomatic: a flexible trimmer for Illumina sequence data," *Bioinformatics*, vol. 30, no. 15, pp. 2114–2120, 2014.

[6] H. Li and R. Durbin, "Fast and accurate long-read alignment with Burrows–Wheeler transform," *Bioinformatics*, vol. 26, no. 5, pp. 589–595, 2010.

[7] P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, and H. Li, "Twelve years of SAMtools and BCFtools," *GigaScience*, vol. 10, 02 2021. giab008.

[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 research*, vol. 22, no. 3, pp. 568–576, 2012.

[9] P. Cingolani, A. Platts, M. Coon, T. Nguyen, L. Wang, S. Land, X. Lu, and D. 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] S. Andrews, "FASTQC. A quality control tool for high throughput sequence data," 2010.

[11] M. Adler, M. Anjum, D. I. Andersson, and L. Sandegren, "Combinations of mutations in envZ, ftsI, mrdA, acrB and acrR can cause high-level carbapenem resistance in Escherichia coli," *Journal of Antimicrobial Chemotherapy*, vol. 71, no. 5, pp. 1188–1198, 2016.

[12] A. Ballard, S. Bieniek, and D. B. Carlini, "The fitness consequences of synonymous mutations in Escherichia coli: Experimental evidence for a pleiotropic effect of translational selection," *Gene*, vol. 694, pp. 111–120, 2019.

[13] P. Dersch, M. A. Khan, S. Mühlen, and B. Görke, "Roles of regulatory RNAs for antibiotic resistance in bacteria and their potential value as novel drug targets," *Frontiers in microbiology*, vol. 8, p. 803, 2017.

[14] S. J. Aedo, M. A. Orman, and M. P. Brynildsen, "Stationary phase persister formation in Escherichia coli can be suppressed by piperacillin and PBP3 inhibition," *BMC microbiology*, vol. 19, no. 1, pp. 1–12, 2019.

[15] Y. Zhang, A. Kashikar, C. A. Brown, G. Denys, and K. Bush, "Unusual Escherichia coli PBP 3 insertion sequence identified from a collection of carbapenem-resistant Enterobacteriaceae tested in vitro with a combination of ceftazidime-, ceftaroline-, or aztreonam-avibactam," *Antimicrobial Agents and Chemotherapy*, vol. 61, no. 8, pp. e00389–17, 2017.

[16] H. Okusu, D. Ma, and H. Nikaido, "AcrAB efflux pump plays a major role in the antibiotic resistance phenotype of Escherichia coli multiple-antibiotic-resistance (Mar) mutants," *Journal of bacteriology*, vol. 178, no. 1, pp. 306–308, 1996.

[17] E. W. Yu, J. R. Aires, and H. Nikaido, "AcrB multidrug efflux pump of Escherichia coli: composite substrate-binding cavity of exceptional flexibility generates its extremely wide substrate specificity," *Journal of bacteriology*, vol. 185, no. 19, pp. 5657–5664, 2003.

[18] K. Gerstle, K. Klätschke, U. Hahn, and N. Piganeau, "The small RNA RybA regulates key-genes in the biosynthesis of aromatic amino acids under peroxide stress in E. coli," *RNA biology*, vol. 9, no. 4, pp. 458–468, 2012.

[19] L. Léger, A. Budin-Verneuil, M. Cacaci, A. Benachour, A. Hartke, and N. Verneuil, "β-lactam exposure triggers reactive oxygen species formation in enterococcus faecalis via the respiratory chain component DMK," *Cell Reports*, vol. 29, no. 8, pp. 2184–2191, 2019.

[20] H. Nikaido, "Molecular basis of bacterial outer membrane permeability revisited," *Microbiology and molecular biology reviews*, vol. 67, no. 4, pp. 593–656, 2003.

[21] E. Trampari, C. Zhang, K. Gotts, G. M. Savva, V. N. Bavro, and M. Webber, "Cefotaxime Exposure Selects Mutations within the CA-Domain of envZ Which Promote Antibiotic Resistance but Repress Biofilm Formation in Salmonella," *Microbiology Spectrum*, pp. e02145–21, 2022.

[22] V. Leong, M. Kent, A. Jomaa, and J. Ortega, "Escherichia coli rimM and yjeQ null strains accumulate immature 30S subunits of similar structure and protein complement," *Rna*, vol. 19, no. 6, pp. 789–802, 2013.

[23] B. Thurlow, J. H. Davis, V. Leong, T. F. Moraes, J. R. Williamson, and J. Ortega, "Binding properties of YjeQ (RsgA), RbfA, RimM and Era to assembly intermediates of the 30S subunit," *Nucleic acids research*, vol. 44, no. 20, pp. 9918–9932, 2016.

[24] S. Rocchio, D. Santorelli, S. Rinaldo, M. Franceschini, F. Malatesta, F. Imperi, L. Federici, C. Travaglini-Allocatelli, and A. Di Matteo, "Structural and functional investigation of the Small Ribosomal Subunit Biogenesis GTP ase A (RsgA) from Pseudomonas aeruginosa," *The FEBS Journal*, vol. 286, no. 21, pp. 4245–4260, 2019.

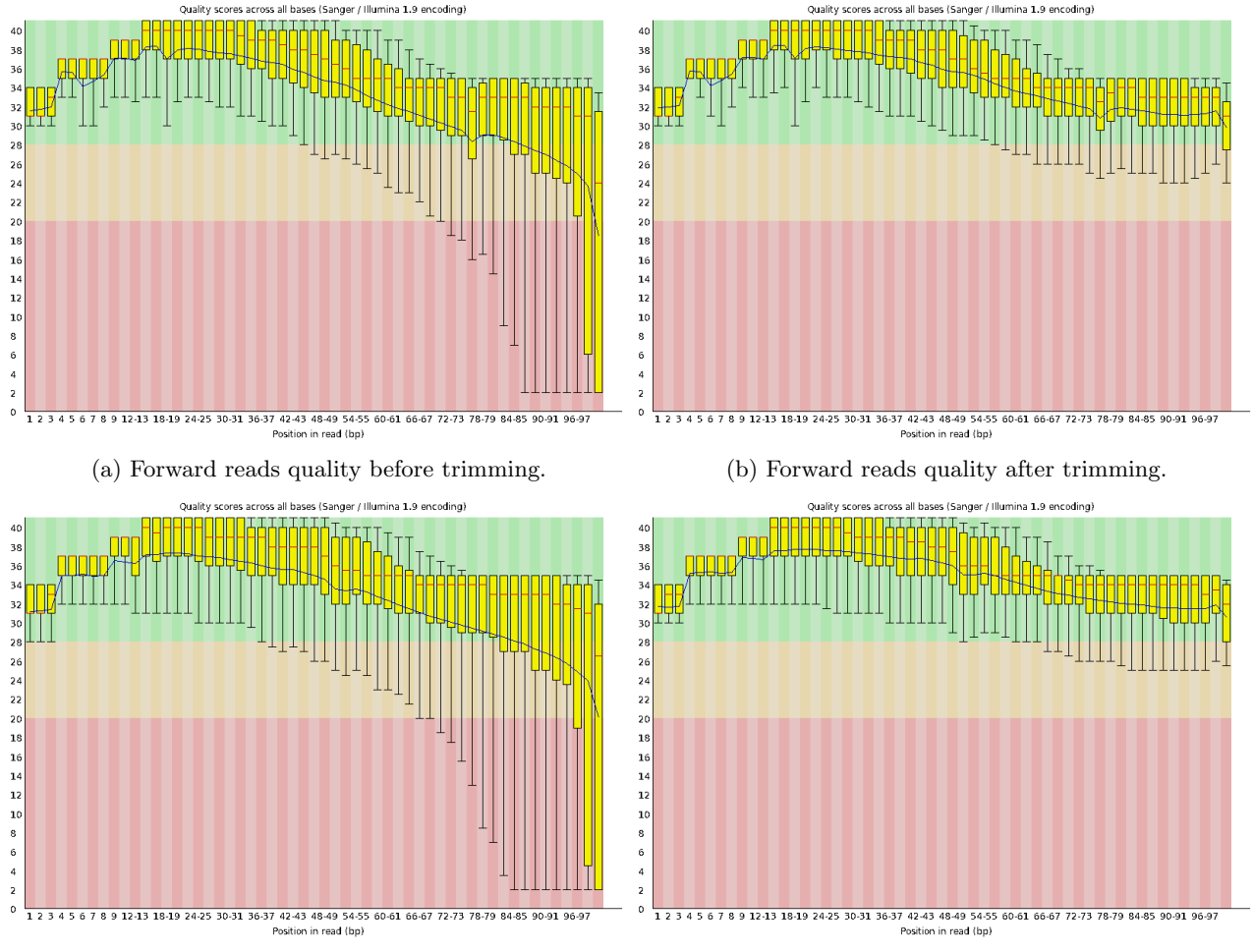
[25] A. Fàbrega, S. Madurga, E. Giralt, and J. Vila, "Mechanism of action of and resistance to quinolones," *Microbial biotechnology*, vol. 2, no. 1, pp. 40–61, 2009.

[26] L. M. King, M. C. Lovegrove, N. Shehab, S. Tsay, D. S. Budnitz, A. I. Geller, J. N. Lind, R. M. Roberts, L. A. Hicks, and S. Kabbani, "Trends in US outpatient antibiotic prescriptions during the coronavirus disease 2019 pandemic," *Clinical Infectious Diseases*, vol. 73, no. 3, pp. e652–e660, 2021.

[27] C. Gomes, L. Ruiz-Roldán, J. Mateu, T. J. Ochoa, and J. Ruiz, "Azithromycin resistance levels and mechanisms in Escherichia coli," *Scientific reports*, vol. 9, no. 1, pp. 1–10, 2019.

[28] A. M. A. Tabrizi, S. Kakhki, S. Kakhki, M. Foroughi, and M. H. A. Azqhandi, "Azithromycin resistance genes in Escherichia coli isolated from wastewater: characterization and modeling-based evaluation of factors affecting the prevalence," *Process Safety and Environmental Protection*, 2022.

Supplementary materials



(a) Forward reads quality before trimming.

(b) Forward reads quality after trimming.

(c) Backward reads quality before trimming.

(d) Backward reads quality after trimming.

Figure 1: Per base reads quality by the Phred-33 score (FastQC report)

Table 3: SNPs impact to protein-coding genes grouped per mutation (snpEff report)

Gene		**Variants impact**			**Variants effect**			
Name	**Id**	**LOW**	**MODERATE**	**MODIFIER**	**Downstream gene**	**Missense**	**Synonymous**	**Upstream gene**
ftsI	b0084		✓			✓		
murE	b0085			✓				✓
murF	b0086			✓				✓
mraY	b0087			✓				✓
murD	b0088			✓				✓
cra	b0080			✓	✓			
mraZ	b0081			✓	✓			
rsmH	b0082			✓	✓			
ftsL	b0083			✓	✓			
acrB	b0462		✓			✓		
pdeB	b0457			✓				✓
ylaC	b0458			✓				✓
maa	b0459			✓				✓
hha	b0460			✓				✓
tomB	b0461			✓				✓
acrR	b0464			✓				✓
mscK	b0465			✓				✓
acrA	b0463			✓	✓			
glnH	b0811			✓				✓
dps	b0812			✓				✓
rhtA	b0813			✓				✓
opgE	b0815			✓				✓
mntR	b0817			✓				✓
ybiR	b0818			✓				✓
ybiT	b0820			✓				✓
yliM	b4736			✓	✓			
ompX	b0814			✓	✓			
mntS	b4705			✓	✓			
ldtB	b0819			✓	✓			
envZ	b3404		✓			✓		
yhgE	b3402			✓				✓
greB	b3406			✓				✓
yhgF	b3407			✓				✓
hslO	b3401			✓	✓			
pck	b3403			✓	✓			
ompR	b3405			✓	✓			
rsgA	b4161	✓					✓	
mscM	b4159			✓				✓
psd	b4160			✓				✓
orn	b4162			✓				✓
yjeV	b4670			✓				✓
nnr	b4167			✓				✓
tsaE	b4168			✓				✓
yjeO	b4158			✓	✓			
queG	b4166			✓	✓			