BIOINFORMATICS INSTITUTE

Point missense mutations in the *ftsI*, *envZ* and *acrB* genes probably causes ampicillin resistance in E. coli

October 27, 2023

Homework number 1

by Kirill Petrikov and Ilia Popov

The study of the mechanisms of antibiotic resistance is one of the directions to solve the problem of widespread antibiotic-resistant bacteria. In this work variant calling and annotation was performed for reads from genome sequencing of an ampicillin resistant E. coli strain. Six-point mutations were revealed. The resistance is probably due to two key missense SNPs. One leads to the Ala544Gly mutation in penicillin-binding protein 3, involved in the synthesis of peptidoglycan, the other leads to Gln569Leu mutation in AcrB protein, a component of a drug efflux pump. Val241Gly mutation in sensor histidine kinase EnvZ also can play important role.

Introduction

Bacterial resistance to antibiotics is one of the major problems of our time [1]. Under constant selective pressure, mutations arise very quickly and are stably maintained, one way or another neutralizing the effect of the antibiotic and allowing the microorganisms to grow. Mutant alleles can actively spread through horizontal transfer [2]. In the modern world, with the increasing consumption of antibiotics, infections with resistant and multi-resistant pathogenic strains are an increasingly likely event. Studying possible mechanisms of antibiotic resistance is one of the directions to solve this problem. *Escherichia coli* is not only a fairly convenient microorganism to work with, but also a common pathogen, which is characterized by frequent resistance to a wide variety of antibiotics [2,3]. On the other hand, beta-lactam antibiotics are still among the most widely and universally used, including ampicillin among the most popular [4,5].
This work is devoted to the analysis of E. coli sequencing data to identify the mutations responsible for the emergence of resistance to ampicillin.

Materials and methods

Data accession

The genome of *Escherichia coli* strain K-12 substr. MG1655 (GenBank No GCF_000005845.2) was utilized as a reference genome.
Raw Illumina reads from shotgun genome sequencing of an ampicillin resistant *E. coli* strain obtained from 10.6084/m9.figshare. 10006541.v3.

Data preprocessing

FastQC v0.12.1 [6] was used to control the quality of raw reads.
Reads were trimmed using Trimmomatic 0.39 [7] (with parameters SLIDINGWINDOW:10:20 LEADING:20 TRAILING:20 MINLEN:20).

Variant calling and annotation

Trimmed reads mapping against a large reference genome was performed by BWA 0.7.17-r1188 with BWA-MEM algorithm [8].
Mpileup file was generated using samtools 1.18 [9]. Subsequent variant calling were performed by VarScan v2.3 with parameter --min-var-freq 0.70 [10].
Variants annotation were performed by SnpEff 5.2 [11].

Results

Quality assessment

The file amp_res_1.fastq contains 1,823,504 lines, that is, 455,876 reads. The amp_res_2.fastq file contains the same. These results coincide with the output of the FastQC program, which also showed that there are no duplicates, GC composition is normal.
When both files were trimmed with a quality index of 20, the number of reads decreased to 446,259.
When trimming with a quality score of 30, the number of reads decreased to 360,209.
When the reads were aligned using samtools flagstat, all reads (892,776) were aligned. 99.87% of the reads were mapped to the reference.
Table 1. Number of reads on the start, after trimming, aligned and mapped. *30 quality score trimming was used just as test. In the workflow 20 quality score trimming was used.

	Start	Trimming quality score		Aligned	Mapped
		20	30*		
Forward	455,876	446,259	360,209	892,776	891,649 (99.87%)
Reverse	455,876	446,259	360,209		

Variant calling

VarScan detects the same results at 50% and 70% non-reference bases in the position required to recognise it as a mutation in the sample. 70% were used.

Annotation

Table 2. Annotated SNPs with their description.

Position	Mutation				Gene product	NCBI Reference
	Type of mutation	Reference	Alternative	Protein		
93043	Missense variant	C	G	Ala544Gly	Peptidoglycan DD-transpeptidase *FtsI*	NP_414626.1
482698	Missense variant	T	A	Gln569Leu	Multidrug efflux pump RND permease *AcrB*	NP_414995.1
852762	Upstream gene variant	A	G	—	Small RNA *RybA*	*rybA*
1905761	Missense variant	G	A	Gly25Asp	Manganese efflux pump *MntP*	NP_416335.4
3535147	Missense variant	T	G	Val241Gly	Sensor histidine kinase *EnvZ*	NP_417863.1
4390754	Synonymous variant	C	A	Ala252Ala	Ribosome small subunit-dependent GTPase A*RsgA*	NP_418585.4

Discussion

Ampicillin acts as an irreversible inhibitor of the enzyme transpeptidase, which is needed by bacteria to make the cell wall [5].Our resistant bacterium has a mutation in the *ftsI* gene, which is an essential cell division protein that catalyses the cross-linking of the peptidoglycan cell wall at the division septum. This protein is inhibited by beta-lactam antibiotics such as penicillin, moenomycin, macarbomycin, furazlocillin, piperacillin and ampicillin [12]. Antibiotics inhibit the activity by binding to the catalytic serine, *ftsI* has a penicillin-binding domain at the C-terminus, and the mutation is in that region. Hence, ampicillin cannot bind to *ftsI* protein.

In addition, the annotated bacterial genome also contains a mutation in the *envZ* gene, which has a number of important functions [13]. Both of these mutations are of direct interest to us, because in a previous study it was demonstrated that mutations in the *ftsI* gene alone did not increase antibiotic resistance, whereas simultaneous mutations in the *ftsI* and *envZ* genes increased the minimum inhibitory concentration of antibiotics several-fold [14].

Also, a mutation in multi-drug efflux pump gene (*acrB*) was annotated. AcrA-AcrB-AcrZ-TolC is a drug efflux protein complex with broad substrate specificity that uses the proton motive force to export substrates [15].A previous study suggested that *E. coli* inducted by ampicillin activates the AcrAB-TolC expression system or activates more than one of the numerous efflux pump systems, and there are other resistance mechanisms other than the efflux mechanism [14].

Therefore, we consider that the resistance to ampicillin is realized by two mechanisms:

1. Target site alteration so antibiotic can't bind;
2. Reduction of amount of drug in cell.

We assume that therapy with ceftriaxone, which is a semi-synthetic, broad-spectrum, third-generation cephalosporin antibiotic, will be more effective.

References

1. Cassell, G.H.; Mekalanos, J. Development of Antimicrobial Agents in the Era of New and Reemerging Infectious Diseases and Increasing Antibiotic Resistance. *JAMA***2001**, *285*, 601–605, doi:10.1001/jama.285.5.601.
2. Poirel, L.; Madec, J.-Y.; Lupo, A.; Schink, A.-K.; Kieffer, N.; Nordmann, P.; Schwarz, S. Antimicrobial Resistance in Escherichia Coli. *Microbiology Spectrum***2018**, *6*, 10.1128/microbiolspec.arba-0026–2017, doi:10.1128/microbiolspec.arba-0026-2017.
3. Wu, D.; Ding, Y.; Yao, K.; Gao, W.; Wang, Y. Antimicrobial Resistance Analysis of Clinical Escherichia Coli Isolates in Neonatal Ward. *Frontiers in Pediatrics***2021**, *9*.
4. Thakuria, B. The Beta Lactam Antibiotics as an Empirical Therapy in a Developing Country: An Update on Their Current Status and Recommendations to Counter the Resistance against Them. *JCDR***2013**, doi:10.7860/JCDR/2013/5239.3052.
5. Swami, O.C. Ampicillin: Rise Fall & Resurgence. *JCDR***2014**, doi:10.7860/JCDR/2014/8777.4356.
6. Babraham Bioinformatics - FastQC A Quality Control Tool for High Throughput Sequence Data Available online: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed on 27 October 2023).
7. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics***2014**, *30*, 2114–2120, doi:10.1093/bioinformatics/btu170.
8. Li, H.; Durbin, R. Fast and Accurate Short Read Alignment with Burrows–Wheeler Transform. *Bioinformatics***2009**, *25*, 1754–1760, doi:10.1093/bioinformatics/btp324.
9. Danecek, P.; Bonfield, J.K.; Liddle, J.; Marshall, J.; Ohan, V.; Pollard, M.O.; Whitwham, A.; Keane, T.; McCarthy, S.A.; Davies, R.M.; et al. Twelve Years of SAMtools and BCFtools. *GigaScience***2021**, *10*, giab008, doi:10.1093/gigascience/giab008.
10. Koboldt, D.C.; Chen, K.; Wylie, T.; Larson, D.E.; McLellan, M.D.; Mardis, E.R.; Weinstock, G.M.; Wilson, R.K.; Ding, L. VarScan: Variant Detection in Massively Parallel Sequencing of Individual and Pooled Samples. *Bioinformatics***2009**, *25*, 2283–2285, doi:10.1093/bioinformatics/btp373.
11. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A Program for Annotating and Predicting the Effects of Single Nucleotide Polymorphisms, SnpEff. *Fly***2012**, *6*, 80–92, doi:10.4161/fly.19695.
12. ftsI - Peptidoglycan D,D-Transpeptidase FtsI - Escherichia Coli (Strain K12) | UniProtKB | UniProt Available online: https://www.uniprot.org/uniprotkb/P0AD68/entry (accessed on 27 October 2023).
13. envZ - Sensor Histidine Kinase EnvZ - Escherichia Coli (Strain K12) | UniProtKB | UniProt Available online: https://www.uniprot.org/uniprotkb/P0AEJ4/entry (accessed on 27 October 2023).
14. 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. *Infect Drug Resist***2019**, *12*, 2853–2863, doi:10.2147/IDR.S221212.
15. acrB - Multidrug Efflux Pump Subunit AcrB - Escherichia Coli (Strain K12) | UniProtKB | UniProt Available online: https://www.uniprot.org/uniprotkb/P31224/entry (accessed on 27 October 2023).