Authors: Vignesh Rao

Abstract:

In order to identify the genetic basis of antibiotic resistance, bacteria was evolved under antibiotic usage and their genomes were sequenced. The sequenced genomes were then compared to the sequenced genomes of a control group that were evolved under normal conditions. To compare the genomes of the control group and the treatment group, the single nucleotide polymorphisms (SNPs) in the treatment group and control group were identified using widely used bioinformatic tools. The SNPs that were only found in the treatment group were used to find genes that were directly impacted by evolution under antibiotic usage. Finally, the NCBI genome browser was used to find the pathways of impacted genes. A SNP unique to the treatment group was found on the ftsl gene, which produces the enzyme peptidoglycan DD-transpeptidase Ftsl.

Methodology:

Bacterium Escherichia coli K-12 MG1655 model strain was sequenced into six pairs of FASTQ files. Three of the files correspond to a treatment group which evolved under antibiotic usage while the other three files correspond to the control group, which evolved under normal conditions.

To analyze the files, a function was created called in order to process the files and convert the files to the right format. First, the files had to be trimmed in order to remove irrelevant data. The function trimmed the files using Trimmomatic V0.39 with the parameters of keepBothReads, LEADING: 3, TRAILING: 3, and MINLEN: 36. The adapter that we used for the IlluminaClip parameter was the TruSeq3 adapter. Next, Bowtie2 is used to align forward read and reverse read files for each respective sequence. The output of the files is a SAM file, which is then converted into a BAM file using samtools V1.11. The BAM format is a compressed binary version of the sequence. The BAM file is then sorted using samtools V1.11.

Once all the files are processed. The bam files that correspond to the treatment group are put into mpileup on SAMtools. This tool takes the data from the file and maps it against the reference genome, allowing for the identification of any SNPs in the sequence. Using mpileup outputs BCF files, which were analyzed using BCFtools.

In order to find SNPs in the BCF files another function was created.This function uses bcftools to call SNPs and index the BCF files.After we created the function, all the bam files belonging corresponding to the control groups were put through this the function outputting all the SNPs found on the treatment groups. Next, the control groups were put through the same function, which output the SNPs found in the control groups. To identify the SNPs that were unique to the treatment groups, bcftools was used to remove all the SNPs that were present in both treatment groups and control groups. The remaining SNPs were identified and their relevant genes were found using the NCBI genome browser. The NCBI genome browser was then used to identify the respective pathway that each gene belonged to.

Results:

SNPs Identified	Only 1 SNP identified ID: 93019 (ftsI gene)
Corresponding Mutated Genes	**FtsI gene (93019)** ● The FtsI gene is responsible for the production of PBP3.
Impacted Pathways	**Peptidoglycan maturation** The FtsI gene belongs to the pathway involving peptidoglycan maturation. Peptidoglycan is a huge macromolecule that surrounds the cell membrane Peptidoglycan helps prevent the pressure within the cell from rupturing the cell. Interrupting the production of peptidoglycan would be lethal to the cell.

After processing each of the FASTQ files and identifying SNPs in both the treatment group and the control group. One SNP unique to the treatment group was found. The SNP unique to the treatment group is located on the gene ftsI. This provides evidence that the ftsI gene belonging to bacteria in the treatment group had mutated due to the presence of antibiotics. ftsI is responsible for the production of PBP3 (penicillin-binding protein 3). PBP3 is an essential cell division protein, but it is found in very small amounts throughout the cell; at any point there are around 100 molecules of PBP3 in the cell. Inhibition of the production of PBP3 is lethal to the cell.

The ftsI gene is involved in the pathway: peptidoglycan maturation. Peptidoglycan, also known as periplasmic murein sacculus, is a macromolecule that encases the entire cytoplasmic membrane of the cell. The peptidoglycan helps protect the cell from turgor pressure. Turgor pressure also called hydrostatic pressure is the pressure the fluid within the cell applies onto the cell membrane. The peptidoglycan prevents the cell from rupturing due to turgor pressure (Vollmer, 2008). Peptidoglycan is also present in cell division, the peptidoglycan of the parent cell is distributed to the daughter cells.

Since the peptidoglycan is essential to the survival of the cell, targeting the process involving peptidoglycan maturation would be lethal to the cell. The inactivation of the ftsl gene by binding of beta-lactam antibiotics has been tested. However, it has been shown that the cell has a procedure in the event of the inactivation of the ftsl gene, allowing it to survive an otherwise lethal antibiotic (Miller, 2004). Although the antibiotic used on the treatment group is unknown, there is a possibility that the antibiotic used targeted the Ftsl gene, causing the treatment group to mutate in order to continue to survive.

Author contributions:

Vignesh was responsible for trimming the FASTQ files and determining quality control using FASTQC and was responsible for generating the genomic index of e.coli and using samtools and we both used these modules to map and determine the SNPs and regions.

Citations:

Miller, C., Thomsen, L. E., Gaggero, C., Mosseri, R., Ingmer, H., & Cohen, S. N. (2004). SOS response induction by beta-lactams and bacterial defense against antibiotic lethality. *Science (New York, N.Y.)*, *305*(5690), 1629–1631. https://doi.org/10.1126/science.1101630

Vollmer, W., & Bertsche, U. (2008). Murein (peptidoglycan) structure, architecture and biosynthesis in Escherichia coli. *Biochimica Et Biophysica Acta (BBA) - Biomembranes, 1778*(9), 1714-1734. doi:10.1016/j.bbamem.2007.06.007