Differential Gene Expressed Analysis of Antibiotic treated *E. Coli* Used to Discern Universal Antimicrobial Response Pathways

Abstract

Differential gene expression analysis of *Escherichia coli* bacterial cells treated with antibiotics novobiocin, rifampicin, and tetracycline reveals consistently increased expression in the aminoacyl-tRNA biosynthesis pathway, which is involved in protein synthesis. Other DEGs identified are somewhat related to the antibiotic mode of action and anticipated cellular response. Identification of differentially expressed genes in bacterial cells in response to various antibiotics may provide insightful information that can potentially be used in antibiotic development.

Introduction

Antibiotic resistance is a natural phenomenon in which bacteria spontaneously develop mutations that decrease a drug's efficacy against the bacteria. The probability of resistance is significantly increased by the excess usage of antibiotics, which allows for greater exposure of the bacteria to the antibiotic, which increases the prevalence of resistant strains.[1] Bacterial antibiotic resistance is inevitable and a growing concern in modern medicine as the threat of common infections becoming more challenging to treat increases.[1] Therefore, exploring mechanisms involved in the maintenance of cells under stress conditions and potential pathways of resistance onset may be advantageous for the development of further antibiotics.

In order to study antibiotic resistance, it is helpful to probe Gram-negative bacterial cells due to their unique double-membrane structure. Gram-negative bacteria cell envelope consists of an outer membrane, which is the first line of defense for the cell, and an inner membrane, which is the site of various biological processes.[2] Both membranes are separated by a cell wall, which contributes to the rigidity of the cell. These characteristics of Gram-negative bacteria serve as a potent barrier against harmful environmental agents and contribute to survival.[2] Thus, exploring the effects of antibiotics in a Gram-negative bacterium, like *Escherichia coli,* can help to understand the mechanisms and pathways impacted by the antibiotics and those involved in trying to help the cell survive.

This analysis utilizes a transcriptomic dataset of RNASeq counts of *E. coli* bacteria cells treated with antibiotics (GSE215300)[3]. The samples used in the dataset are *E. coli* (MG1655 strain) treated with no antibiotic (control), novobiocin (nov), rifampicin (Rif), or tetracycline (tet). The reads

are obtained with these cultures using three library preparation methods: standard 5' directional preparation, enrichment for VCE, or non-enrichment. Each sample has three biological replicates per combination of culture and library preparation methods, resulting in a total of 36 samples. The dataset has expression data of 4,464 genes for each sample.

To anticipate potential differentially expressed genes between each sample, it is crucial to be cognizant of the mode of action of the antibiotics used in the dataset. Firstly, novobiocin impacts the cell's DNA replication and damage repair mechanisms by targeting DNA gyrase.[4] Next, rifampicin functions by targeting RNA polymerase, impairing transcription in the cell.[5] Further, tetracycline alters translation in the cell by targeting the 30s ribosomal subunit.[6] The three antibiotics' general function is to impact a bacteria cell by impairing its ability to grow and replicate. However, one notable difference is that *E. Coli* is not susceptible to novobiocin. The resistance to Novobiocin is due to the robust envelope of *E. coli* (and other Gram-negative bacteria).[4] Instead, novobiocin is generally effective in gram-positive bacteria (like Staphylococcus). Although *E. coli* is technically susceptible to rifampicin and tetracycline antibiotics, studies have previously observed an emergence of some resistant *E. coli* strains.[5,6] Therefore, it is imperative to understand the impact of the antibiotics intracellularly, and observing patterns in the gene expression may be helpful.

The goal of this analysis is to understand which genes are up or down-regulated after administration of specific antibiotics and to discern the pathways/mechanisms impacted, which may be those involved in bacteria acquiring resistance to antibiotics. It can be hypothesized that the antibiotic-treated samples (in comparison to the control samples) will demonstrate an upregulation in genes involved in general stress response. This should be observed by some consistency in the differentially expressed genes in the experimental groups.

Methods

The raw count data was obtained from the Gene Expression Omnibus (series id: GSE215300). The original count dataset includes two additional control samples and the 36 samples described on GEO. To maintain consistency between the samples (3 biological replicates), the additional control samples labeled K4 are removed. Additionally, counts are put into a data frame structure with the given GenBank geneIDs as the row names, as required for analysis by the DESeq2 tool. The metadata was also defined to categorize the samples' condition – the antibiotic treatment given. For this analysis, the library preparation method is disregarded.

The cleaned count data was processed by the DESeq2 package in R with the design formula design = ~ condition, indicating the comparison of the samples by their antibiotic treatment alone. DESeq2 filters low-count genes and normalizes the data using built-in estimation functions. DEGs were identified by filtering for genes from the DESeq2 results that had differential expression values (log2 foldchange <-1 or >1) and were significant (had a p-value<0.05).

To gauge the similarity in genes between the samples, PCA was conducted on the DESeq2 results to visualize variance in the sample's expression patterns. Additionally, DEGs were clustered using a hierarchical clustering method to identify expression patterns. Clusters were visualized on a heatmap.

With the identified DEGs, gene set functional enrichment analysis was conducted with the ClusterProfiler package in R. ClusterProfiler requires Entrez IDs of the genes in order to extract information on the gene's functional information. For this, the rentrez package in R (specifically the esearch command) was used to fetch Entrez IDs from the given Genbank IDs. Once Entrez IDs were obtained, the gse commands of ClusterProfiler were used to retrieve the DEG's Gene Ontology (GO-terms) and KEGG pathway information using the *E. coli* genome annotation (org.EcK12.eg.db). Both were visualized on a series of graphs to identify the interactions of all pathways and GO-terms.

The group comparisons used in this analysis are each experimental antibiotic-treated sample against the control sample (novobiocin vs. control, rifampicin vs. control, and tetracycline vs. control). These three groups will be analyzed to discern the cell's response to each antibiotic and any consistencies in the DEGs that may be considered a universal antibiotic response.

Results

The PCA of the DEGs shows a dispersal of the data as a whole but with small localized clusters within each group, and samples are localized in the same regions. The heatmap visualizing the hierarchical clustering of the DEGs shows similar patterns of gene expression across biological replicated of the same groups, confirming consistency in the expression data without outliers. Some overlap can also be observed in the novobiocin samples with the rifampicin samples and tetracycline samples with the rifampicin samples. However, further analysis is required to probe similarities in gene expression across groups.

Differentially Expressed Genes:

Of the 4,464 genes in the original data set, 3,173 genes were identified to be differentially expressed across the three experimental groups. Novobiocin samples had 955 DEGs (520 upregulated genes and 435 downregulated genes). Rifampicin samples had 1015 DEGs (566 upregulated and 449 downregulated). Tetracycline samples had 1203 DEGs (593 upregulated and 610 downregulated). The ratio of up-and-down-regulated genes is about half. The DEGs shared 69 genes upregulated and 75 genes downregulated. This suggests some similar pattern in the DEGs and potentially a general response to antibiotics.

GO term enrichment

In novobiocin samples, the most activated GO terms involve carbohydrate processing, and those greatly suppressed involve iron transport. In rifampicin samples, many of the activated GO terms are related to oxidative stress, and the most common suppressed GO terms were in RNA metabolism regulation. These terms are consistent with rifampicin's mode of action as it targets RNA polymerase and impacting transcription can result in oxidative stress.[6] In the tetracycline samples, the activated GO terms were primarily involved in biofilm production, and the suppressed terms were in varied metabolism, located in the pilus, and biofilm production.

KEGG Pathways:

In novobiocin samples, notably activated pathways are lipopolysaccharide (LPS) metabolism and aminoacyl-tRNA biosynthesis, and the most suppressed pathways are involved in flagellar assembly. In rifampicin, the primary activated pathway is in aminoacyl-tRNA biosynthesis, and the most suppressed pathway is in sulfur metabolism. The primary activated pathway in tetracycline is in aminoacyl-tRNA biosynthesis, and suppressed pathways involve various metabolic pathways.

Discussion

The analysis demonstrated that treating the E. coli cell with either antibiotic resulted in approximately 1000 differentially expressed genes. Approximately 4000 plus genes have been identified in the E. coli genome, which suggests that exposure to antibiotics causes a significant impact on the cell as a quarter of its genome's expression is altered.[7] The KEGG pathway analysis revealed that the aminoacyl-tRNA biosynthesis pathway is activated in all three antibiotic-treated samples. This pathway is involved in protein synthesis, which is a critical function needed for all cells to grow and survive.[8] In order for the cell to signal for stress or foreign agents, it requires signaling molecules which are often proteins. Another point that may explain the upregulation of this pathway is any mechanism used by the cell in an effort to protect the cell requires the increased production of proteins. For example, novobiocin samples saw an increase in LPS production for the cell wall, and tetracycline samples had increased biofilm production; both these mechanisms likely require increased tRNA production. Additionally, in tetracycline samples where the ribosome is impacted, aminoacyl-tRNA biosynthesis may be upregulated because the stunting of protein translation may cause the cell to overproduce tRNAs in an attempt to push for some protein synthesis. Although there was not a recognized stress-response pathway that was upregulated as hypothesized, the upregulation of aminoacyl-tRNA biosynthesis almost goes a step further by impacting a primary function of the cell. Furthermore, it is possible that a general stress response gene like RpoS is not a DEG because it is regularly expressed in control cells in response to other processes, including nutrient scavenging, osmotic stress, etc.[9]

Studies have been conducted to explore the possibility of inhibitors of the aminoacyl-tRNA biosynthesis pathway to be used as antibiotics.[8] This confirms the design and ideology of this analysis in that identifying DEGs of antibiotic-treated cells may be useful in developing novel antibiotics. In the analysis above, it is observed that some enriched GO terms and pathways are relevant to the antibiotic's mechanism and the anticipated cellular response. However, a number of identified GO terms and pathways do not have an obvious reason for activation/suppression and may be significant points of study. Additionally, for future analysis, it may be worth comparing DEGs of antibiotic samples against each other (instead of comparison to the control) to exacerbate alterations in gene expression that are specific to each antibiotic and may give information into how cells may develop resistance and how current antibiotics can be modified to overcome the bacteria's use of the resistance mechanisms.

References

1. Akram, F., Imtiaz, M., & Haq, I. U. (2023). Emergent crisis of antibiotic resistance: A silent pandemic threat to 21st century. Microbial pathogenesis, 174, 105923. https://doi.org/10.1016/j.micpath.2022.105923
2. Silhavy, T. J., Kahne, D., & Walker, S. (2010). The bacterial cell envelope. *Cold Spring Harbor perspectives in biology*, *2*(5), a000414. https://doi.org/10.1101/cshperspect.a000414
3. Balkin, Alexander S. (2023). Transcriptomic data of Escherichia coli K-12 control and antibiotic-treated cultures. *Institute for Cellular and Intracellular Symbiosis.* https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE215300
4. May, J. M., Owens, T. W., Mandler, M. D., Simpson, B. W., Lazarus, M. B., Sherman, D. J., Davis, R. M., Okuda, S., Massefski, W., Ruiz, N., & Kahne, D. (2017). The Antibiotic Novobiocin Binds and Activates the ATPase That Powers Lipopolysaccharide Transport. *Journal of the American Chemical Society*, *139*(48), 17221–17224. https://doi.org/10.1021/jacs.7b07736
5. Chopra, I., & Roberts, M. (2001). Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. Microbiology and molecular biology reviews: MMBR, 65(2), 232–260. https://doi.org/10.1128/MMBR.65.2.232-260.2001
6. Weinstein, Z. B., & Zaman, M. H. (2018). Evolution of Rifampin Resistance in Escherichia coli and Mycobacterium smegmatis Due to Substandard Drugs. Antimicrobial agents and chemotherapy, 63(1), e01243-18. https://doi.org/10.1128/AAC.01243-18
7. Keseler IM, Gama-Castro S, Mackie A, Billington R, Bonavides-Martinez C, Caspi R, Kothari A, Krummenacker M, Midford PE, Muniz-Rascado L, Ong WK, Paley S, Santos-Zavaleta A, Subhraveti P, Tierrafria VH, Wolfe AJ, Collado-Vides J, Paulsen IT, Karp PD (2021). "The EcoCyc Database in 2021." Front Microbiol 12;711077. PMID: 34394059
8. Pang, L., Weeks, S. D., & Van Aerschot, A. (2021). Aminoacyl-tRNA Synthetases as Valuable Targets for Antimicrobial Drug Discovery. International journal of molecular sciences, 22(4), 1750. https://doi.org/10.3390/ijms22041750
9. Schellhorn H. E. (2020). Function, Evolution, and Composition of the RpoS Regulon in Escherichia coli. Frontiers in microbiology, 11, 560099. https://doi.org/10.3389/fmicb.2020.560099