Jan Grzegorz Kosiński

Identification and characteristics of the intrinsic transcription terminators in bacteria and their contribution to the antibiotic stress response

Identyfikacja i charakterystyka rho-niezależnych terminatorów transkrypcji u bakterii oraz ich udziału w odpowiedzi na stres antybiotykowy

Abstract

Transcription is an essential process allowing the propagation of genetic information from DNA to RNA. The full spectrum of RNA roles in cells is yet to be uncovered. However, it is already known that the variability of mechanisms and processes in which RNA molecules are involved is extraordinary. Despite acting as a template for the synthesis of proteins, RNAs perform a wide array of regulatory functions within the cell. Therefore, it is crucial to fully understand how RNA molecules regulate gene expression in bacteria and how people can use this knowledge to their advantage, e.g., in treating infectious diseases.

In bacteria, RNA molecules can modulate cellular responses to various environmental or physiological stimuli via several well-understood mechanisms. RNAs can fold into sophisticated and actively changing secondary structures forming through base-pairing with other adjacent or distant parts of the same molecule. RNA secondary structures can modulate a range of processes involved in gene expression, usually by affecting the work of RNA polymerases, ribosomes, and other crucial enzymes or complexes. One of the most common mechanisms of regulating gene expression is via changes in the RNA secondary structure leading to premature transcription termination.

This thesis introduces innovative bioinformatic tools developed to detect and annotate intrinsic transcription terminators in bacterial organisms, utilizing 3'-end RNA sequencing (Term-seq) and total RNA sequencing (RNA-seq) data.

The first tool aims to identify stable 3’RNA ends within the signal represented by the genomic coverage of 5' Term-seq read ends. A novel algorithm, based on Irreproducible Discovery Rates, is employed to effectively distinguish signals being reproducible between

biological replicates from others, likely attributed to intermediate products of RNA processing or degradation. Moreover, the provided software enables the annotation of various features associated with the 3' ends of RNA molecules, specifically identifying those resulting from intrinsic transcription termination.

The second algorithm introduced as a part of this thesis is a machine-learning method for accurate prediction of intrinsic termination events based on RNA-seq data, leveraging publicly available datasets from RNA-seq and Term-seq protocols. The development of this toolkit facilitated the creation of a comprehensive atlas of intrinsic bacterial terminators, validated through high-throughput sequencing techniques.

The role of the third presented software is to annotate genomic or transcriptomic intervals using information from GTF or GFF files. These file formats are used to store annotations of genes, transcripts, and other genomic features identified within a given genome. The application of this tool allowed for the completion of the annotations presented in the atlas of intrinsic bacterial terminators.

This study's author has comprehensively compared terminators across the bacterial kingdom through computational analysis, revealing substantial variability between different taxa. This finding challenges the prevailing assumption of their sequence and structural similarity that has been primarily derived from research conducted on model organisms like *E. coli* and *B. subtilis*. The thesis extensively discusses the notable differences observed in intrinsic terminators' secondary structure, nucleotide composition, and genomic location.

Significantly, the author presents a meticulous atlas of intrinsic termination events in *S. aureus*, a highly clinically relevant bacterium known for its multidrug resistance. Based on the analysis of 140 publicly available datasets, this atlas represents, to the best of the author's knowledge, the most comprehensive compilation of *S. aureus* terminators compiled to date.

Furthermore, the thesis introduces additional results unveiling several context-specific transcription termination events governing gene expression in response to antibiotic-induced stress in *S. aureus*. These findings allow us to better understand the therapeutic implications of specific antibiotic treatments for bacteria of high clinical importance.