

Jan Grzegorz Kosiński

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

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

Streszczenie rozprawy doktorskiej

Transkrypcja jest procesem, który umożliwia propagację informacji genetycznej z poziomu DNA do RNA. Mimo że pełne spektrum funkcji pełnionych przez cząsteczki RNA nie zostało jeszcze w pełni poznane, wiadomo, że różnorodność mechanizmów i procesów, w których są one bezpośrednio zaangażowane, jest nadzwyczajna. Pomimo pełnienia roli matrycy niezbędnej do syntezy białek, cząsteczki RNA pełnią także szeroki zakres funkcji regulacyjnych wewnątrz komórki. Dlatego też kluczowe jest pełne zrozumienie sposobów regulacji tych procesów oraz praktyczne wykorzystanie tej wiedzy, na przykład w leczeniu chorób zakaźnych.

W bakteriach cząsteczki RNA mogą uczestniczyć w regulacji odpowiedzi komórkowych na różne bodźce środowiskowe lub fizjologiczne przy pomocy kilku dobrze poznanych mechanizmów. RNA może związać się w złożone i aktywnie zmieniające się struktury drugorzędowe, które tworzą się poprzez parowanie zasad z innymi przyległymi lub odległymi częściami tej samej cząsteczki. Zmiany struktur drugorzędowych RNA mogą modulować szereg procesów związanych z ekspresją genów, zazwyczaj poprzez wpływanie na aktywność polimeraz RNA, rybosomów i innych kluczowych enzymów lub kompleksów. Indukowane zmiany w strukturze drugorzędowej RNA mogą również w konsekwencji prowadzić do przedwczesnej terminacji transkrypcji.

Niniejsza rozprawa prezentuje szereg innowacyjnych narzędzi bioinformatycznych, opracowanych w celu wykrywania i adnotacji rho-niezależnych terminatorów transkrypcji u bakterii. Wykorzystują one dane pochodzące z wysokoprzepustowych technik sekwencjonowania 3' końców RNA (Term-seq) oraz całkowitego sekwencjonowania RNA (RNA-seq).

Pierwsze z prezentowanych w rozprawie narzędzi ma na celu identyfikację stabilnych 3' końców RNA w ramach sygnału reprezentowanego jako pokrycie genomu przez odczyty uzyskiwane metodą Term-seq. Opracowany algorytm, wykorzystując metodę IDR (ang. *Irreproducible Discovery Rate*), pozwala skutecznie różnicować sygnały zachowawcze pomiędzy replikatami biologicznymi, od pozostałych, w znacznej mierze reprezentujących pośrednie produkty przetwarzania RNA lub jego degradacji. Opracowane oprogramowanie umożliwia także precyzyjną adnotację szeregu cech związanych z 3' końcami cząsteczek RNA, w szczególności identyfikując te będące wynikiem rho-niezależnej terminacji transkrypcji.

Drugi z opracowanych algorytmów służy do przewidywania miejsc rho-niezależnej terminacji transkrypcji na podstawie danych z sekwencjonowania RNA (RNA-seq) oraz metod uczenia maszynowego. Do utworzenia modelu predykcyjnego wykorzystane zostały publicznie dostępne zbiory odczytów, uzyskanych przy pomocy protokołów RNA-seq oraz Term-seq. Zastosowanie wspomnianych narzędzi bioinformatycznych do analizy szerokiego zestawu publicznie dostępnych danych umożliwiło stworzenie niezwykle rozbudowanego atlasu rho-niezależnych terminatorów bakteryjnych, zidentyfikowanych przy pomocy technik wysokoprzepustowego sekwencjonowania.

Rolą trzeciego z przygotowanych programów jest adnotacja interwałów genomowych lub transkryptomowych przy pomocy informacji zawartych w plikach GTF lub GFF. Wymienione formaty plików służą do przechowywania adnotacji genów, transkryptów oraz innych cech genomowych poznanych w obrębie danego genomu. Zastosowanie niniejszego narzędzia pozwoliło uzupełnić zestaw adnotacji przedstawionych w ramach opracowanego atlasu rho-niezależnych terminatorów bakteryjnych.

Prezentowana praca przedstawia kompleksowe porównanie terminatorów występujących u wybranych gatunków bakterii, ujawniając ich znaczącą zmienność między różnymi taksonami. Obserwacja ta poddaje w wątpliwość dominujące przeświadczenie mówiące o znacznym podobieństwie sekwencji i struktur terminatorów występujących u bakterii. Przeświadczenie to jest głównie wynikiem badań przeprowadzonych na organizmach modelowych, takich jak *Escherichia coli* i *Bacillus subtilis*. Przedstawiona praca szczegółowo omawia zauważalne różnice w strukturze drugorzędowej, kompozycji nukleotydowej, a także lokalizacji genomowej zidentyfikowanych rho-niezależnych terminatorów transkrypcji.

Niezwykle istotnym elementem prezentowanej rozprawy jest starannie przygotowany atlas miejsc rho-niezależnej terminacji transkrypcji u *Staphylococcus aureus*, bakterii o wysokim znaczeniu klinicznym, charakteryzującej się szczepami wykazującymi oporność na

wiele znanych leków. Atlas ten został opracowany na podstawie analizy 140 publicznie dostępnych zbiorów danych, stanowiąc tym samym według najlepszej wiedzy autora, najbardziej wszechstronne zestawienie terminatorów opracowane do dziś dla *S. aureus*.

Ponadto, niniejsza praca dyplomowa prezentuje wyniki, które ujawniają kilka kontekstowo specyficznych zdarzeń terminacji transkrypcji, prowadzących do regulacji ekspresji genów w odpowiedzi na stres wywołany stosowaniem antybiotyków u *S. aureus*. Wyniki te mają potencjalne znaczenie terapeutyczne, pozwalając szczegółowo zrozumieć wpływ stosowanych antybiotyków na bakterie stanowiące wysokie zagrożenie kliniczne.