

## STRESZCZENIE

Praca doktorska: Jetka, Tomasz, *Quantitative methods to describe information flow in biochemical signalling pathways*, 2020.

Tytuł w języku polskim: *Ilościowe metody opisu przetwarzania informacji przez biochemiczne szlaki sygnałowe*

Aby złożony organizm mógł przetrwać, funkcjonować oraz rozwijać się, każda komórka, która jest jego częścią, musi podejmować decyzję adekwatną do warunków panujących w otaczającym je środowisku zewnętrznym. W tym celu, kluczowe jest, by komórki dysponowały jak najbardziej dokładną wiedzą na temat stanu tego środowiska. Z tego powodu wyewoluował zaawansowany system komunikacji pomiędzy nawet najmniejszymi składowymi organizmu. Jednymi z podstawowych mechanizmów, które spełniają funkcję przekazywania informacji są tzw. biochemiczne szlaki sygnałowe. Stanowią one schematyczny model tego, jak przesyłane i przetwarzane są informacje wymieniane pomiędzy: komórkami, komórką i organizmem oraz komórką i środowiskiem zewnętrznym. W najbardziej typowym przykładzie, nagromadzone w przestrzeni międzykomórkowej cząsteczki reprezentują informację o pewnej zmianie warunków w środowisku, np. koncentracja molekuł TNF- $\alpha$  zwiększa się w wyniku pojawienia się niepożądanego patogenu. Informacja ta jest przechwytywana przez komórkę za pomocą receptorów na błonie komórkowej. Takie wiązanie ligandu przez receptor wywołuje w następstwie szereg interakcji pomiędzy białkami komórkowymi, co w konsekwencji prowadzi do aktywacji jednego z czynników transkrypcyjnych, np. NF- $\kappa$ B, który inicjuje funkcjonalną odpowiedź komórkową np. pobudzenie lokalnego zapalenia poprzez ekspresję odpowiednich genów.

Właściwy i dokładny opis tego, jak komórki przekazują informację stanowi istotny krok, by pogłębić nasze rozumienie zachowania organizmu w różnych warunkach zarówno homeostazy jak i stanach chorobowych. Niestety, pomimo ogromnej wiedzy jakościowej na temat procesów oraz reakcji leżących u podstaw działania szlaków sygnałowych, nasze rozumienie zależności ilościowych, które są z nimi związane, jest nadal znikome. Tym samym, przedstawienie wiarygodnych predykcji dotyczących zachowania procesów molekularnych jest w dużym stopniu wciąż poza naszym zasięgiem. Powody takiego stanu rzeczy są wielorakie i można je zawrzeć w trzech głównych aspektach: i) aktywność szlaków sygnałowych jest nierozzerwalnie związana z losowością, która wynika zarówno ze stochastyczności reakcji na poziomie molekularnym, jak i niepewności co do środowiska i bodźców zewnętrznych, ii) mimo ciągłego postępu, dostępne metody eksperymentalne są wciąż wysoce niedoskonałe i generują dane z relatywnie dużym poziomem niedokładności, iii) brakuje metod statystycznych, które byłyby dostosowane do złożoności problemów biologii molekularnej a zarazem charaktery-

zowały się efektywnością obliczeniową.

Biorąc pod uwagę powyższe powody, w niniejszej pracy skupiłem się na badaniu i rozwoju metodologii ilościowej analizy szlaków sygnałowych. Przede wszystkim, przyjęliśmy perspektywę matematycznej teorii informacji, która w sposób formalny pozwala traktować szlaki sygnałowe jako systemy przekazujące i przetwarzające informacje. Zastosowanie takiego, alternatywnego, podejścia pozwoliło dotychczas uzyskiwać atrakcyjną interpretację oraz wyciągnąć istotne wnioski, które rzuciły nowe światło na rozumienie funkcjonowania komórek. Początkowe sukcesy w tej dziedzinie zostały jednak zahamowane poprzez brak odpowiednich metod, algorytmów i narzędzi numerycznych umożliwiających analizę złożonych zagadnień związanych ze szlakami sygnałowymi, które reprezentują ciekawe praktyczne problemy.

Kierując się potrzebami środowiska biologii systemów, aby rozszerzyć stosowalność teorii informacji do problemów biologii molekularnej, opracowałem oryginalne rozwiązania teoretyczne oraz praktyczne narzędzia, które stanowią znaczący postęp w stosunku do dotychczasowej metodologii. Po pierwsze, przedstawiam całościową i ogólną teorię, jak określać centralną miarę stosowaną w teorii informacji, tzw. channel capacity, nawet w przypadku, gdy szlak sygnałowy odbiera wiele różnych sygnałów jednocześnie, aktywując wielokierunkową odpowiedź. Z drugiej strony, jako komplementarną metodę opisuję nowatorski algorytm, który pozwala mierzyć informację systemów biologicznych bezpośrednio na podstawie danych eksperymentalnych. Jego implementacja, dostępna publicznie jako biblioteka dla środowiska obliczeniowego R, pozwala w efektywny sposób stosować perspektywę teorii informacji nawet tym badaczom, którzy nie posiadają zaawansowanej wiedzy nt. statystyki czy bioinformatyki. Co więcej, na przykładzie kilku istotnych biologicznie szlaków sygnałowych, pokazałem również, w jaki sposób opracowana przeze mnie metodologia może zostać zastosowana w praktyce, by pogłębić wiedzę na temat mechanizmów leżących u podstaw komunikacji między komórkami.

Ostatecznie, niniejsza praca łączy teorię, eksperymenty, analizę statystyczną oraz narzędzia obliczeniowe do stworzenia całościowej metodologii, która umożliwi stosowanie teorii informacji do złożonych szlaków sygnałowych. Przedstawione wyniki zdają się być wyczekiwany zestawem metod, których użyteczność potwierdzono we współpracy z innymi grupami badawczymi. Co więcej, z dużym prawdopodobieństwem zaprezentowany wkład teoretyczny jest nie bez znaczenia dla innych dziedzin takich jak matematyczne projektowanie eksperymentów, neurobiologia obliczeniowa czy też teoretyczne podstawy uczenia maszynowego.