

# Interpretacja funkcji białek w ludzkim mikrobiomie jelitowym w skali ewolucyjnej

Ludzkie jelita są domem dla szerokiej gamy mikroskopijnych organizmów, które tworzą naszą mikroflorę jelitową. Te bakterie, wirusy i inne organizmy prokariotyczne odgrywają istotną rolę w regulacji procesów fizjologicznych i mogą wpływać na zdrowie człowieka. W ciągu ostatniej dekady naukowcy wykorzystali wielkoskalowe sekwencjonowanie metagenomiczne do zbadania różnorodności tych organizmów w czasie i przestrzeni oraz tego, w jaki sposób różnice w ich populacjach odnoszą się do ludzkiego zdrowia i zachowania. W ramach tych badań odkryto ponad 600 milionów genów kodujących białka oraz prawie 5000 gatunków prokariotycznych. Jednak tylko połowę tych genów można zrozumieć na podstawie ich podobieństwa do innych dobrze zbadanych, pozostawiając wiele „ciemnej materii” białek o nieznanym funkcjach.

W ostatnim czasie głębokie uczenie spowodowało rewolucję w dziedzinie biologii obliczeniowej. Ucząc się na podstawie bogactwa informacji o białkach zdeponowanych w różnych bazach danych, umożliwiło to opracowanie wielu narzędzi obliczeniowych do analizy takich białek. Stosując metody głębokiego uczenia się do przewidywania trójwymiarowych struktur białek z niespotykaną dotąd dokładnością, naukowcy mogą teraz badać funkcje i interakcje białek, które nie mają znanych homologów. W przestrzeni sekwencji białek, traktując białka jako ciągi słów, modele głębokiego uczenia mogą wykrywać relacje ewolucyjne, które wcześniej były poza zasięgiem, umożliwiając adnotację odległych homologów i genów sierocych. Aby lepiej zrozumieć, co białka robią funkcjonalnie, łącząc informacje o sekwencji i trójwymiarowej strukturze białek przy użyciu algorytmów głębokiego uczenia się, opisano ogromny funkcjonalny repertuar zakodowany w białkach. Jest to zatem właściwy czas na przeprowadzenie analizy tych białek na dużą skalę, łącząc metody oparte na głębokim uczeniu się do przewidywania struktury białek, wykrywania bardzo odległych relacji ewolucyjnych i przewidywania funkcji, wykraczających poza możliwości standardowych podejść.

W ramach tego projektu stworzymy atlas struktur białkowych ludzkiego jelita z adnotacjami funkcji i mapę wszechświata białek, aby pomóc nam poruszać się po tej ogromnej przestrzeni. Wykorzystamy głębokie uczenie się i modelowanie ewolucyjne na dużą skalę, aby lepiej zrozumieć funkcjonalnie ciemne białka metagenomu ludzkiego jelita. Badając białka i ich kontekst genomowy, będziemy dążyć do zidentyfikowania unikalnych białek i ich potencjalnej roli w zdrowiu i zachowaniu człowieka. Nasze odkrycia pomogą ustalić priorytety przyszłych badań i zapewnią pełniejszy obraz białek znajdujących się w jelitach człowieka.