

HORIZON
2020

New algorithms for inference and optimization from large-scale biological data

Wyniki w skrócie

Przełom w projektowaniu i inżynierii białek, który zmieni reguły gry

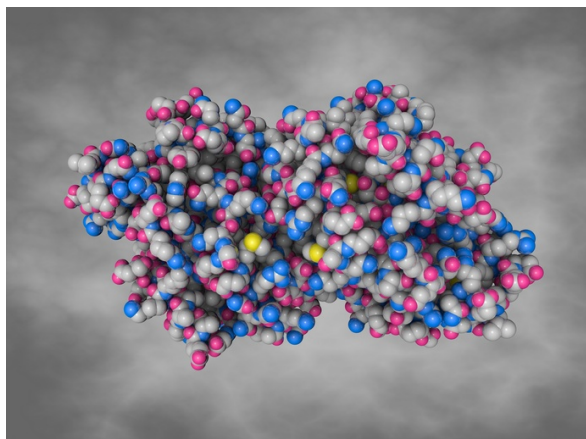
Wykorzystanie modeli statystycznych i algorytmów uczenia maszynowego może pomóc w projektowaniu i inżynierii nowych białek o udoskonalonych funkcjach.



GOSPODARKA
CYFROWA



ZDROWIE



© Maryna Olyak/stock.adobe.com

Projektowanie nowych białek o pożądanych funkcjach jest zadaniem skomplikowanym, lecz może przynieść niezwykle korzyści branży farmaceutycznej i biomedycznej oraz przemysłowi. Choć to zastosowania medyczne tworzą obecnie najbogatszy rynek dla produktów białkowych wytwarzanych metodami inżynieryjnymi, syntetyczne enzymy wykorzystywane są również w przemyśle spożywczym do celów przetwórstwa. Co więcej, sztuczne enzymy znalazły zastosowanie w branżach związanych z

ochroną środowiska, gdzie używa się ich do detoksykacji zanieczyszczeń lub do projektowania zmodyfikowanych mikroorganizmów, których zadaniem jest eliminacja zanieczyszczeń środowiskowych, takich jak tworzywa sztuczne.

Uproszczenie metod projektowania nowych białek

Projektowanie nowych białek o udoskonalonych funkcjach docelowych jest trudnym zadaniem z uwagi na dużą przestrzeń sekwencyjną oraz liczne ograniczenia strukturalne, które trzeba uwzględnić i pokonać. Na przykład, niewielkie białko zbudowane ze 100 aminokwasów posiada około 10^{130} możliwych wariantów – więcej niż atomów we wszechświecie – przy czym zdecydowana większość nie spełnia żadnych funkcji.

Coraz częściej badacze przekonują się, że odszukanie wariantu o sekwencji najlepiej służącej określonej roli wymaga zastosowania zaawansowanych rozwiązań doświadczalnych w połączeniu z zaawansowanymi metodami obliczeniowymi. W związku z tym badacze skupieni wokół projektu [INFERNET](#) opracowali skuteczne narzędzia do celów inferencji i optymalizacji danych wielkoskalowych. Badania przeprowadzono dzięki wsparciu otrzymanemu w ramach [działania „Maria Skłodowska-Curie”](#) (MSCA).

„Aby móc wyciągać wnioski lub formułować przewidywania w oparciu o zaobserwowane schematy i trendy, stworzyliśmy modele statystyczne i algorytmy uczenia maszynowego, które pomogły nam w analizie danych i identyfikacji związków i korelacji pomiędzy zmiennymi”, wyjaśnia stypendysta MSCA, Andrea Pagnani.

Modelowanie związków pomiędzy genotypem i fenotypem

Opracowanie dokładnych testów biochemicznych o wysokiej przepustowości za pomocą technik sekwencjonowania przyczyniło się do stworzenia wielkoskalowej metody skriningu genetycznego jako podstawowego narzędzia do badania związku pomiędzy ewolucją, przystosowaniem a innymi koncepcjami biologicznymi leżącymi u podstaw badań doświadczalnych. Umożliwiło to szczegółowe badanie związków pomiędzy genotypem i fenotypem w warunkach kontrolowanej selektywnej presji ze strony czynników zewnętrznych. Tego typu metody są standardowo stosowane do selekcji molekuł o specyficznych właściwościach.

Zespół projektu INFERNET opracował [oparte na danych podejście probabilistyczne](#) do modelowania związków pomiędzy genotypem i fenotypem wynikających z doświadczeń. Metodę tę można wykorzystać jako model generacyjny do znajdowania nowych wariantów genetycznych o wysokim stopniu przystosowania, a także można ją włączyć do procesów opartych na uczeniu maszynowym związanych z ewolucją ukierunkowaną.

Przewidywanie wystąpienia mutacji w toku ewolucji

Kluczową zdolnością związaną z przewidywaniem rozpowszechnienia i częstotliwości występowania mutacji genetycznych jest skuteczne generowanie

sztucznych sekwencji o określonej docelowej specyfice. W tym celu opracowano różne strategie obliczeniowe i specyficzne metody modelowania.

„Z naszego punktu widzenia, generowanie sztucznych sekwencji wiąże się z możliwością skutecznego wygenerowania zestawu sekwencji o nie dających się odróżnić cechach statystycznych z zestawu treningowego”, podkreśla Pagnani.

Zespół projektu INFERNET zaproponował [nową strategię obliczeniową](#) do wygenerowania sekwencji, które istotnie różnią się od tych naturalnych. Nowy system obliczeniowy musi zostać poddany walidacji doświadczalnej pod kątem aktywności biologicznej wyselekcjonowanego zestawu sztucznych sekwencji.

Metodologia INFERNET poprawi funkcjonalność białek

Kluczowe testy walidacyjne metodologii INFERNET dotyczyły [projektu sztucznej mutazy choryzmianowej](#), kluczowego enzymu w biosyntezie aromatycznych aminokwasów. Badaczom udało się zaprojektować nowe, przypominające naturalne warianty z zachowaną lub udoskonaloną funkcjonalnością.

Opracowane w ramach projektu modele statystyczne oparte na sekwencji okazały się skuteczne w określaniu białek i zapewnianiu dostępu do ogromnej przestrzeni sekwencji funkcjonalnych. To osiągnięcie położyło podwaliny pod rozwój bardziej ogólnego procesu związanego z projektowaniem sztucznych białek opartego na ewolucji.

„Podejścia statystyczne oparte na ewolucji mogą dostarczać cennych wskazówek badaczom poszukującym funkcjonalnych białek o udoskonalonej docelowej funkcjonalności”, podsumowuje Pagnani.

Słowa kluczowe

INFERNET, białka, ewolucja, modele statystyczne, algorytmy uczenia maszynowego, inżynieria, projektowanie białek, mutacje genetyczne, inferencja, mutaza choryzmianowa

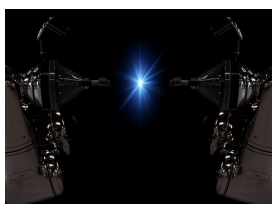
Znajdź inne artykuły w tej samej dziedzinie zastosowania



Przyszłość regeneracji ścięgien



Oprogramowanie typu open source ułatwia prowadzenie badań nad mózgiem



Farmaceutyczne linie napełniające zrewolucjonizowane przez ramiona robotów



Wnioski z badań genetycznych mogą doprowadzić do zmian w leczeniu niewydolności serca



Informacje na temat projektu

INFRNET

Finansowanie w ramach

Identyfikator umowy o grant: 734439

EXCELLENT SCIENCE - Marie Skłodowska-Curie
Actions

Strona internetowa projektu 

DOI

[10.3030/734439](https://doi.org/10.3030/734439) 

Projekt został zamknięty

Data podpisania przez KE

12 Grudnia 2016

Koszt całkowity

€ 900 000,00

Wkład UE

€ 900 000,00

Koordynowany przez

IIGM FOUNDATION



Italy

Data rozpoczęcia

1 Marca 2017

Data zakończenia

31 Grudnia 2022

Ostatnia aktualizacja: 2 Czerwca 2023

Permalink: <https://cordis.europa.eu/article/id/444053-game-changing-advances-in-protein-design-and-engineering/pl>

European Union, 2025