Akceptuję
W ramach naszej witryny stosujemy pliki cookies w celu świadczenia państwu usług na najwyższym poziomie, w tym w sposób dostosowany do indywidualnych potrzeb. Korzystanie z witryny bez zmiany ustawień dotyczących cookies oznacza, że będą one zamieszczone w Państwa urządzeniu końcowym. Możecie Państwo dokonać w każdym czasie zmiany ustawień dotyczących cookies. Więcej szczegółów w naszej Polityce Prywatności

Polskie narzędzie do porządkowania wirusowego chaosu

Polscy badacze opracowali program komputerowy Vclust, który pozwala w ledwie kilka godzin porównać miliony sekwencji wirusów i uporządkować je według stopnia

podobieństwa. Analizowanie ogromnych zbiorów danych genetycznych tradycyjnymi metodami zajęłoby nawet kilka lat.

„Przy użyciu Vclust analiza zbioru 15 mln sekwencji zajmuje ok. czterech godzin, a najdokładniejsze narzędzia stosowane dotychczas potrzebowałyby na to ok. czterech lat. To istotny krok dla rozwoju wirusologii i metagenomiki, ponieważ ułatwi identyfikację i klasyfikację nowych wirusów, które w ostatnich latach są masowo odkrywane dzięki nowoczesnym technologiom sekwencjonowania" - podkreślili w rozmowie z PAP twórcy rozwiązania.

Na łamach czasopisma „Nature Methods" zespół naukowców z Wydziału Biologii Uniwersytetu im. Adama Mickiewicza w Poznaniu oraz Wydziału Automatyki, Elektroniki i Informatyki Politechniki Śląskiej, we współpracy ze specjalistą z Uniwersytet Friedricha Schillera w Jenie, opisał narzędzie, które pozwala odróżniać znane wirusy od nowych oraz analizować ich różnorodność w różnych środowiskach, co ma kluczowe znaczenie dla monitorowania nowych patogenów i badań nad mikrobiomem.

Badacze wyjaśnili, że współczesna mikrobiologia zmaga się problemem zalewu danych. Każdego roku odkrywa się nawet milion nowych wirusów, w efekcie czego powstają tak duże zbiory, że ich analiza oraz klasyfikacja staje się coraz większym wyzwaniem dla zespołów badawczych.

„Taka eksplozja danych to zasługa metagenomiki, czyli metody pozwalającej na odczytanie całego DNA obecnego w danej próbce środowiskowej, np. z oceanu, gleby czy jelita człowieka. Do tej pory brakowało narzędzi, które pozwalałoby efektywnie analizować i grupować tak dużą liczbę sekwencji. Istniały metody bardzo dokładne, ale nie radziły sobie one z taką skalą danych. Dlatego postanowiliśmy stworzyć program, który będzie równie precyzyjny, ale znacznie wydajniejszy i poradzi sobie z milionami genomów naraz" - wyjaśnił współautor publikacji dr hab. Andrzej Zieleziński z UAM.

Dlaczego wirusy są takie trudne?

Jak dodał, w biologii klasyfikacja organizmów - czyli taksonomia - opiera się zwykle na porównywaniu konkretnych genów obecnych u wszystkich przedstawicieli danej grupy. Dzięki temu można tworzyć drzewa filogenetyczne organizmów, grupować je, wyodrębniać rodziny czy gatunki i określać ich stopień pokrewieństwa. Z wirusami jest zupełnie inaczej.

„Wirusy, w przeciwieństwie chociażby do bakterii, nie mają jednego wspólnego genu, który można by porównywać. Różnią się od siebie zbyt mocno. Dlatego klasyczne metody filogenetyczne nie działają. Nie sprawdziło się też podejście oparte na ich morfologii, np. kształcie kapsydów, które okazało się zbyt powolne i mało skalowalne. Pozostało nam więc jedno - porównywać sekwencje całych genomów, litera po literze" - powiedział dr hab. Zieleziński.

Trudno to zrobić, kiedy takich genomów są miliony. Jak wyjaśnił kierujący projektem prof. Sebastian Deorowicz z Politechniki Śląskiej, istnieją już narzędzia pozwalające grupować te olbrzymie zbiory danych, jednak robią to ogromnym kosztem obliczeniowym, trudnym do powtórzenia w warunkach codziennej pracy badawczej. „Nie jest tak, że nikt wcześniej tego nie zrobił, ale wymagało to tak dużych zasobów (np. superkomputerów), że trudno byłoby powtarzać ten proces regularnie, zwłaszcza gdy mamy do czynienia z coraz mocniej rozrastającymi się zestawami danych" - zaznaczył.

„Dlatego my postawiliśmy na optymalizację, czyli zaprojektowanie jak najbardziej efektywnych algorytmów i możliwie najbardziej wydajnego kodu, który umożliwił redukcję czasu obliczeń o kilka rzędów wielkości. Wszystko po to, aby przenieść obliczenia z superkomputera na zwykłą stację

roboczą” - dodał.

Trzy kroki do uporządkowania wirusów

Vclust działa w trzech etapach. Pierwszy z nich polega na wstępnym filtrowaniu, w którym program błyskawicznie identyfikuje pary sekwencji wykazujące choć minimalne podobieństwo. Dzięki temu zamiast porównywać każdą sekwencję z każdą inną - co oznaczałoby biliony możliwych kombinacji - algorytm ogranicza analizę do znacznie mniejszej liczby, rzędu setek milionów najbardziej obiecujących par.

Drugi etap to precyzyjne porównanie wybranych sekwencji. Wykorzystywany jest tu autorski algorytm LZ-ANI, oparty na technikach inspirowanych algorytmami kompresji danych używanych w formatach ZIP czy RAR. Zasada jego działania jest prosta: im bardziej dwie sekwencje są do siebie podobne, tym lepiej „kompresują się” razem, czyli zajmują mniej miejsca po przetworzeniu. Ten efekt jest wykorzystywany jako miara podobieństwa.

W ostatnim etapie następuje klastrowanie, czyli grupowanie sekwencji na podstawie ich podobieństwa. Wirusy, których genomy są do siebie najbardziej zbliżone, trafiają do tej samej grupy. Dzięki temu można łatwiej określić, które z nich są ze sobą spokrewnione i tworzą „rodziny”, a które są zupełnie odrębne. To pozwala lepiej zrozumieć różnorodność wirusów i ich ewolucyjne powiązania.

„Dzięki temu program wykorzystuje moc komputera do maksimum. Każdy, kto testował Vclust, był pełen zdumienia co do jego szybkości” - podkreślił dr hab. Zieleziński.

Twórcy Vclust zadbali o to, by narzędzie było w pełni darmowe i ogólnodostępne. Można je pobrać z internetu i uruchomić na własnym komputerze. Dla tych, którzy nie mają zaawansowanego sprzętu, przygotowano wersję przeglądarkową: vclust.org.

Narzędzie działa w bardzo prosty sposób: użytkownik może wkleić własne sekwencje, uruchomić analizę i po krótkim czasie otrzymać wynik - bez potrzeby logowania czy rejestracji. Aktualnie wersja przeglądarkowa pozwala na analizę do tysiąca sekwencji jednocześnie, co w wielu przypadkach okazuje się w zupełności wystarczające.

Prof. Deorowicz i dr hab. Zieleziński zapewniają, że projekt będzie rozwijany. „Planujemy dodać więcej funkcji, a w przyszłości chcielibyśmy rozszerzyć Vclust także o możliwość analizy genomów bakterii” — zapowiedzieli.

Źródło: pap.pl
http://laboratoria.net/aktualnosci/32482.html

12-06-2025

Zaawansowany rak szyjki macicy – przełamać tabu

Kobiety z rakiem szyjki macicy wciąż pozostają w cieniu.

12-06-2025

39 projektów z dofinansowaniem - Welcome to Poland 2024

39 projektów z 37 instytucji na łączną kwotę 18,8 mln zł.

12-06-2025

Pasożyty mogą pomóc w leczeniu chorób autoimmunologicznych

Zidentyfikowała dr hab. Katarzyna Donskow-Łysoniewska z Uczelni Łazarskiego.

12-06-2025

Pierwszy w Polsce zabieg termoablacji

pierwotnego raka płuca

W Instytucie Gruźlicy i Chorób Płuc w Warszawie.

12-06-2025

Rodzaj treningu może wpływać na skład mikrobiomu jelitowego

Nasze jelita będą zasiedlały nieco inne, prozdrowotne bakterie.

12-06-2025

Krystaliczne gąbki zmniejszają ryzyko pracy nad niebezpiecznymi...

Naukowcy opracowali metodę zabezpieczania związków metaloorganicznych.

12-06-2025

Terapia genowa choroby Alzheimera

Komentują badacze w piśmie „Signal Transduction and Targeted Therapy”.

12-06-2025

Zbyt restrykcyjna dieta może skutkować depresją

Wynika z najnowszych badań.

Informacje dnia: Zaawansowany rak szyjki macicy – przełamać tabu 39 projektów z dofinansowaniem - Welcome to Poland 2024 Polscy naukowcy zajmą się certyfikacją technologii kwantowych Pasożyty mogą pomóc w leczeniu chorób autoimmunologicznych Pierwszy w Polsce zabieg termoablacji pierwotnego raka płuca Rodzaj treningu może wpływać na skład mikrobiomu jelitowego Zaawansowany rak szyjki macicy – przełamać tabu 39 projektów z dofinansowaniem - Welcome to Poland 2024 Polscy naukowcy zajmą się certyfikacją technologii kwantowych Pasożyty mogą pomóc w leczeniu chorób autoimmunologicznych Pierwszy w Polsce zabieg termoablacji pierwotnego raka płuca Rodzaj treningu może wpływać na skład mikrobiomu jelitowego Zaawansowany rak szyjki macicy – przełamać tabu 39 projektów z dofinansowaniem - Welcome to Poland 2024 Polscy naukowcy zajmą się certyfikacją technologii kwantowych Pasożyty mogą pomóc w leczeniu chorób autoimmunologicznych Pierwszy w Polsce zabieg termoablacji pierwotnego raka płuca Rodzaj treningu może wpływać na skład mikrobiomu jelitowego

Partnerzy