

Streszczenie pracy doktorskiej z akceptacją promotora

Wirusy, zajmujące niszę na granicy między organizmami żywymi a materią, wywierają ogromny wpływ na życie na naszej planecie. Ich zdolność infekowania różnych kategorii organizmów daje im możliwości kształtowania ekosystemu, oddziaływania na gospodarkę rolno-spożywczą, wywoływania chorób wśród ludzi, lecz również umożliwia ich wykorzystywanie w medycynie. Tradycyjnie proces odkrywania nowych wirusów obejmował ich namnażanie w hodowlach komórkowych oraz późniejsze sekwencjonowanie. Jednak dzięki postępowi w technologiach sekwencjonowania, wirusy można obecnie odkrywać bez konieczności namnażania. Dzieje się to jednak kosztem utraty informacji o ich naturalnym gospodarzu.

Istnieje wiele technik laboratoryjnych umożliwiających identyfikację naturalnego gospodarza wirusa, rozwijane są w tym celu również metody *in silico*. Niniejsza rozprawa podejmuje problem ustalenia przybliżonego gospodarza dla nowo odkrytego wirusa przy pomocy nauczania maszynowego. W ramach opisywanego projektu rozpoczęto od opracowania metody, która umożliwia rozróżnienie między wirusami eukariotycznymi a fagami, na podstawie ich kompletnych sekwencji genomowych. Następnie, wypracowaną metodologię przetestowano na fragmentach sekwencji wirusowych imitujących dane uzyskane za pomocą sekwencjonowania następnej generacji. Na końcu pokazano zastosowanie opracowanej metody do rozróżniania wirusów infekujących różne grupy organizmów eukariotycznych. Rozróżnianie między bakteriofagami a wirusami eukariotycznymi na podstawie ich kompletnych sekwencji, jak i długich fragmentów, zakończyło się sukcesem. Z powodzeniem rozwiązano również problem odróżnienia wirusów infekujących kręgowce i rośliny nasienne. Klasyfikatory umożliwiające rozróżnianie między bakteriofagami a wirusami eukariotycznymi udostępniono w postaci oprogramowania Host Taxon Predictor.

*Masadeel'w'ia Aewala*