

# Zastosowania nowych metod bioinformatycznych do opisu sekwencji genomu wirusów

Autor: mgr Damian Panas  
Promotor: dr hab. Piotr Wąż

## Streszczenie

Niniejsza praca poświęcona jest w głównej mierze zastosowaniom Dynamicznej Reprezentacji Sekwencji DNA/RNA – bioinformatycznym sposobie charakteryzowania kwasów nukleinowych, umożliwiającym tak zwaną analizę sekwencji bez dopasowania. Metoda ta została wykorzystana odrębnie – na potrzeby badania genomu wirusa Zika (D. Panas, P. Wąż, D. Bielińska-Wąż, A. Nandy, S.C Basak, 2D-Dynamic Representation of DNA/RNA Sequences as a Characterization Tool of the Zika Virus Genome, *MATCH Commun Math Comput Chem* **77**:321–332, 2017) – oraz w połączeniu z algorytmami uczenia maszynowego. W pracy (D. Panas, P. Wąż, D. Bielińska-Wąż, A. Nandy, S.C. Basak, An Application of the 2D-Dynamic Representation of DNA/RNA Sequences to the Prediction of Influenza A Virus Subtypes, *MATCH Commun Math Comput Chem* **80**:295–310, 2018) oraz bezpośrednio w samej rozprawie pokazano, że połączenia tego rodzaju umożliwiają dokładne identyfikowanie nieznanymi szczepów wirusa grypy.

Słowa kluczowe: bioinformatyka, uczenie maszynowe, analiza sekwencji.

# Applications of new methods of bioinformatics to a description of the genome sequences of viruses

Author: mgr Damian Panas  
Supervisor: dr hab. Piotr Wąż

## Abstract

The Thesis is focused mainly on applications of the bioinformatics method called Dynamic Representation of DNA/RNA Sequences – an alignment-free characterization tool of arbitrary nucleic acid sequence. The method was used single-handedly – within analysis of the Zika virus genome (D. Panas, P. Wąż, D. Bielińska-Wąż, A. Nandy, S.C Basak, 2D-Dynamic Representation of DNA/RNA Sequences as a Characterization Tool of the Zika Virus Genome, *MATCH Commun Math Comput Chem* **77**:321–332, 2017) – but also coupled with machine learning algorithms. In (D. Panas, P. Wąż, D. Bielińska-Wąż, A. Nandy, S.C. Basak, An Application of the 2D-Dynamic Representation of DNA/RNA Sequences to the Prediction of Influenza A Virus Subtypes, *MATCH Commun Math Comput Chem* **80**:295–310, 2018) and directly in the Thesis it was shown that such couplings can yield high accuracy of the identification of the unknown influenza virus strains.

Keywords: bioinformatics, machine learning, sequence analysis.