

## **Dirbtinio intelekto panaudojimas funkciniam prokariotinių virusų charakterizavimui**

### **Harnessing the powers of AI for functional characterization of prokaryotic viromes**

#### **Aprašymas:**

Prokariotiniai virusai (fagai) yra gausiausi organizmai žemėje (Clokie ir kt., 2011). Nobelio premijos laureatų Maxo Delbrücko, Salvadoro Lurijos, Alfredo Hershey ir Francois Jacobo T4 ir  $\lambda$  fagų tyrimai padėjo pagrindus molekulinei biologijai ir šiuolaikinei genetikai. Fagų tyrimai ne tik naudingi fundamentiniams tyrimams, bet ir yra svarbūs mikrobų ekosistemose nuo mūsų žarnyno iki vandenynų (Naureen ir kt., 2020), padeda kovoti su atsparumu antibiotikams (Dedrick ir kt., 2022) ir pritaikomi biotechnologijoje (Abril ir kt., 2022). Norėdami sėkmingai panaudoti fagus, turime žinoti, kaip jie veikia. Nepaisant beveik šimtmetį trukusių fagų tyrimų ir jų svarbos, 83% prokariotinių virusų baltymų šeimų funkcijos yra nežinomos (Pappas ir Dutilh, 2021).

AlphaFold2 – dirbtiniu intelektu (DI) pagrįsto struktūrų nusakymo įrankio sukūrimas padarė revoliuciją struktūrinėje biologijoje ir leido gauti baltymų struktūrą, kurios kokybė gali prilygti eksperimentinėms struktūroms. Be to, DI pagrįsti baltymų kalbos modeliai tampa pagrindine pažangiausios bioinformatikos programinės įrangos, skirtos baltymų funkcinei anotacijai, dalimi.

Mes planuojame panaudoti pažangiausius DI įrankius prokariotinių virusų baltymų charakterizavimui. Mes tikimės rasti naujų baltymų šeimų ir jas apibūdinti, pirmenybę teikdami nukleorūgštis jungiantiems baltymams, nes jie gali būti pritaikyti biotechnologijoje ir genų redagavime. Įdomiausiems baltymams, nustatytiems kompiuterinės analizės metu, bus atliekama eksperimentinė analizė ir sprendžiamos krioEM struktūros.

#### **Abstract:**

Prokaryotic viruses (phages) are the most abundant organisms on earth. Studies of Nobel laureates Max Delbrück, Salvador Luria, Alfred Hershey and Francois Jacob on T4 and  $\lambda$  phages laid foundations for molecular biology and modern genetics. Phage studies not only help to broaden our fundamental research but are important in microbial ecosystems from our gut to the oceans, help to combat antibiotic resistance and unlock biotechnological applications. For successful phage applications we need to know how they function. Despite almost a century of phage studies and their importance, functions of 83% of prokaryotic virus protein families remain unknown.

Development of the AlphaFold2 – an Artificial Intelligence (AI)-based tool for structure prediction, revolutionized structural biology and made it possible to obtain a protein structure the quality of which can be on par with the experimental structures. Moreover, AI-based protein language models are becoming the fundamental part of state-of-art bioinformatics software for protein functional annotation.

Here, we propose to use AI tools for characterization of prokaryotic virus proteins. We expect to find novel protein families and characterize them by prioritizing nucleic acids binding proteins, due to their potential for translational applications in the regulation and modification of nucleic acids. Most intriguing proteins will be subjected to experimental studies and cryoEM analysis.

**Mokslinis vadovas / supervisor: Darius Kazlauskas**