

# POSUDEK VEDOUCÍHO DIPLOMOVÉ PRÁCE

**Autor práce:** Bc. Anh Vu Le

**Název práce:** Hledání sekundárních struktur v primárních strukturách nukleových kyselin

**Vedoucí:** Doc. Ing. Jiří Kléma, Ph.D.

Diplomová práce pana Le se zabývá vyhledáváním kandidátských oblastí lidského genomu s předem danou funkcí. Protože funkce není často přímo rozpoznatelná v samotné posloupnosti nukleotidů v DNA, tedy primární sekvenci, ale je svázána především s konzervovanými strukturálními motivy produktů, jež jsou z DNA odvozeny, musí vyhledávání zohlednit právě tyto strukturální motivy. Jde i o případ tzv. IRES, tedy míst přítomných v oblasti před začátkem genu, která umožňují nasednutí ribozomu na mRNA a spuštění translace. Diplomová práce navrhuje obecný pracovní tok pro detekci takových míst a aplikuje jej pro vyhledání kandidátských výskytů IRES virové žloutenky typu C (HCV).

Z pohledu bioinformatika jde o složitou úlohu čekající na efektivní řešení. IRES nelze redukovat na primární sekvenci a problém řešit jejím přímým zarovnáním vůči genomu kvůli její rychlé evoluci a variabilitě. Řešení založené na srovnání dobře konzervované sekundární struktury IRES se sekundárními strukturami RNA odvozenými z genomu je výpočetně příliš náročné. Současné postupy balancují mezi těmito dvěma extrémy a snaží se najít kompromis mezi vyjadřovací silou reprezentace použité v jednotlivých krocích algoritmu a jejich efektivitou. Diplomant zvolil postup, který nejprve ze sekundární struktury IRES pomocí inverzního skládání odvodí velkou množinu možných primárních sekvencí. Ty pak zarovnává vůči genomu. Oblasti s nadprahovou shodou skládá do sekundární struktury a tuto strukturu porovnává s výchozí sekundární strukturou IRES. Na základě porovnání sekundárních struktur pak rozhoduje o vhodnosti kandidátů.

Jde o práci, která má integrační charakter. Všechny dílčí algoritmy pracovního toku již existovaly před započítím práce. Hlavním úkolem bylo je nalézt, seznámit se s jejich parametry a vlastnostmi, zvolit pro ně správné vstupy a vhodně je kombinovat. Přesto šlo o těžký úkol, který se v několika chvílích uplynulého roku jevil jako technicky neřešitelný (nedařilo se nalézt dostupné implementace), popřípadě výpočetně nezvladatelný (navržený pracovní tok jako celek nedával možnost prohledat genom v dostatečné šíři). Jednoduché nebylo ani rozhodnutí, v jaké formě poměrně rozsáhlou IRES HCV strukturu použít. Velmi oceňuji, že navržený pracovní tok je funkční, není jen demem a podařilo se nalézt kandidátskou oblast výskytu IRES HCV v 5'UTR oblasti genu DRC3.

Diplomant řešil své téma dlouhodobě a zodpovědně. Aktivně se podílel na formulaci dílčích úkolů, při rešerši molekulárně biologické literatury a dalších zdrojů prokázal velkou samostatnost. Konzultoval úlohu s molekulárním biologem. Půl roku pracoval nezávisle na studijním pobytu v Jižní Koreji a udělal tam znatelný pokrok. Rád bych zmínil, že řada pro nás původně nadějných pokusů skončila nezdarem. Některé se v závěrečné práci ani neobjevily, i tak má práce přes 80 stran textu. Řadu věcí jsem se musel naučit i já sám. Bylo důležité, že student nerezignoval a vytrval. Diplomová práce má publikační potenciál, pan Le bude pokračovat v doktorském studiu.

Souhrnně lze říci, že práce splňuje všechny cíle, které jsme si na počátku vytkli. V případě tohoto zadání to není samozřejmost. Práci doporučuji k obhajobě a hodnotím ji známkou

**A — výborně.**

V Praze, dne 24. května 2019

Doc. Ing. Jiří Kléma, Ph.D., vedoucí