

Математички факултет  
Универзитета у Београду

УНИВЕРЗИТЕТ У БЕОГРАДУ  
МАТЕМАТИЧКИ ФАКУЛТЕТ  
СТУДЕНТСКА СЛУЖБА

Бр. 6/47  
19.06-2013. год.  
Београд, Студентски трг 16  
Тел. 20 27 801

МОЛБА  
ЗА ОДОБРАВАЊЕ ТЕМЕ МАСТЕР РАДА

Молим да се одобри израда мастер рада под насловом:

„Интеграција резултата различитих програма за склапање генома”

**Значај теме и области:**

Познавање генома различитих организама у великој мери унапређује разумевање живог света на Земљи. Секвенца ДНК је са рачунарске стране посматрано ниска над азбуком {A,C,G,T} која представља низ нуклеотида од којих је сачињен један ланац молекула ДНК. Технологијом секвенцирања нове генерације (NGS) помоћу специјализованих машина читавају се кратки делови секвенце, који се затим повезују да би се целокупна ДНК секвенца програмски реконструисала. NGS производи велики број кратких читавања које покривају сваки појединачни ДНК сегмент и од изузетног је значаја да програми за њихово повезивање буду ефикасни и да обезбеде одговарајућу тачност. С обзиром да дужине ДНК секвенци могу износити и неколико милијарди базних парова и да је NGS технологија још увек непрецизна, као и да су читавања вишеструко редундантна, проблем склапања генома је рачунарски веома комплексан.

**Специфични циљ рада:**

Данас је у употреби на десетине програма за склапање генома, асемблера, који су често засновани на Де Бруиновим графовима. Циљ овог мастер рада је интеграција резултата различитих асемблера ради добијања што квалитетније реконструкције ДНК секвенце. Практични део рада укључује конфигурацију алата за интеграцију GAM NGS[1] који ће бити прилагођен одабраним асемблерима: Velvet[2] и Abyss[3]. Перформансе комбинованог асемблера биће упоређене са појединачним алатима. Тестирање ће се извести на сировим подацима преузетим из репозиторијума секвенци SRA[4] и резултати ће бити упоређени са одговарајућим референтним геномом. За приказ резултата биће имплементирана графичка репрезентација. Апликација ће бити развијена употребом програмског језика Python.

**Референце:**

1. Vicedomini, Riccardo, et al. "GAM-NGS: genomic assemblies merger for next generation sequencing." BMC bioinformatics 14.7 (2013): S6.
2. Zerbino, Daniel R., and Ewan Birney. "Velvet: algorithms for de novo short read assembly using de Bruijn graphs." Genome research 18.5 (2008): 821-829.
3. Simpson, Jared T., et al. "ABYSS: a parallel assembler for short read sequence data." Genome research 19.6 (2009): 1117-1123.
4. <https://www.ncbi.nlm.nih.gov/sra/>

