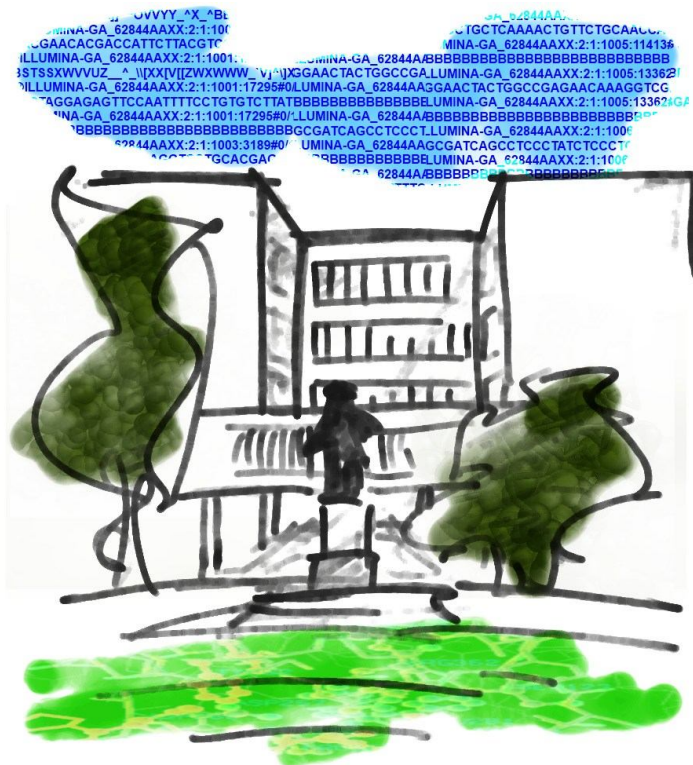


Група за биоинформатику



Семинар Биоинформатика

2. новембар 2011.

Универзитет у Београду
Математички факултет

План рада

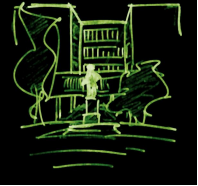


- Представљање Групе за биоинформатику и преглед досадашњих резултата
 - Саша Малков
- Текуће активности
 - Ненад Митић
- Планиране активности у наредном периоду
 - Гордана Павловић-Лажетић

Група за биоинформатику



Представљање групе за биоинформатику Досадашњи резултати



- Биоинформатика је област рачунарских наука у којој се применом метода рачунарских и математичких наука врше истраживања у домену био-медицинских наука

Проблеми



- Проблеми су најчешће из области молекуларне биологије
 - геномика
 - протеомика

Проблеми (2)



- Постоји и већи број проблема из других области
 - моделирање биолошких система
 - обрада биомедицинских слика и сигнала
 - уочавање и моделирање корелација између макро својстава и генског или протеинског састава
 - биохемија и хемија
 - и друго



- Блиске области рачунарства
 - моделирање информација и процеса
 - алгоритмика
 - базе података
 - истраживање података
 - машинско учење
- Блиске области математике
 - статистика и вероватноћа
 - анализа, геометрија, алгебра

Основне карактеристике



- Релативно велика количина података
- Релативно сложена обрада података
- Велика динамика прилива информација и знања

Група за биоинформатику



- Група за биоинформатику на Математичком факултету основана је у јесен 2002.
 - др Милош Бељански са Института за општу и физичку хемију
 - неколико наставника и сарадника Математичког факултета је успоставило сарадњу са



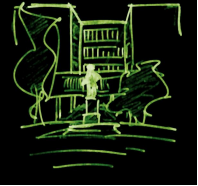
- Основни почетни ресурси су били
 - добра воља и спремност за рад
 - велики подстрек и неисцрпан извор тема
 - др Милош Бељански
- У почетку је група функционисала у облику семинара.
 - Циљеви:
 - у релативно кратком року прикупити информације из различитих области БИ истраживања
 - препознати теме у којима се могу остварити резултати

Организација рада



- Данас група функционише по мањим тимовима
 - тимови се праве према темама
 - баве се различитим проблемима
 - постоје преклапања (тимови нису дисјунктни)
- Рад у оквиру пројеката надлежног министарства
 - 1858 (2003-2006)
 - 144030 (2007-2010)
 - 174021 (2011-2014)
 - руководилац др Предраг Јаничић

Састав групе



- Основу групе чине наставници и сарадници Математичког факултета и истраживачи Института за општу и физичку хемију:
 - др Гордана Павловић-Лажетић
 - др Милош Бељански
 - др Мирјана Павловић
 - др Миодраг Живковић
 - др Ненад Митић
 - др Саша Малков
 - Јована Ковачевић

Састав групе (2)



- У раду групе су учествовали или повремено учествују и:

- др Новица Блажић
- др Снежана Зарић
- мр Милена Вујошевић-Јаничић
- мр Јелена Граовац
- мр Јелена Хаџи-Пурић
- мр Сана Стојановић
- др Андрија Томовић
- Горан Предовић
- мр Ана Манола
- мр Милан Јовановић
- мр Ивана Божић
- мр Даворка Јандрлић
- мр Горана Дацић
- мр Слободанка Марјановић
- мр Милена Шошић
- Невена Петровић
- Мирјана Маљковић
- Александар Зељић
- Милан Тодоровић
- Александар Стефановић
- Дарко Живановић
- Мирјана Вуковић
- Марко Стојићевић
- Ана Јеловић
- Милица Кнежевић
-

Сарадња



- Сарадња је остваривана са истраживачима из Србије и иностранства
 - *Center for information science and technology, Temple University, Philadelphia*
 - др Зоран Обрадовић
 - *VIB Department of Structural Biology, Vrije University, Brussel*
 - др Петер Томпа
 - Институт за физику
 - др Бранко Драговић и његов тим
 - Институт за биолошка истраживања “Синиша Станковић”
 - др Ана Симоновић
 - Институт за молекуларну генетику и генетичко инжењерство
 - др Јелена Беговић, др Наташа Голић, др Бранко Јовчић
 - Хемијски факултет
 - др Снежана Зарић, др Срђан Стојановић
 - и други



- Група је радила / ради / остварила резултате у различитим областима:
 - Протеомика
 - Геномика
 - Алгоритми за рад са нискама
 - Истраживање података

Преглед резултата



- Представићемо укратко неке од проблема којима смо се бавили и најважније остварене резултате
1. Биоинформатичка анализа виралних генома
 2. Биоинформатичка анализа секундарне структуре протеина
 3. Биоинформатичка анализа бактеријских генома
 4. Биоинформатичка анализа неуређености протеина код прокариота
 5. Истраживање података у биоинформатици
 6. Реконструкција геномских секвенци

1. БИ анализа виралних генома

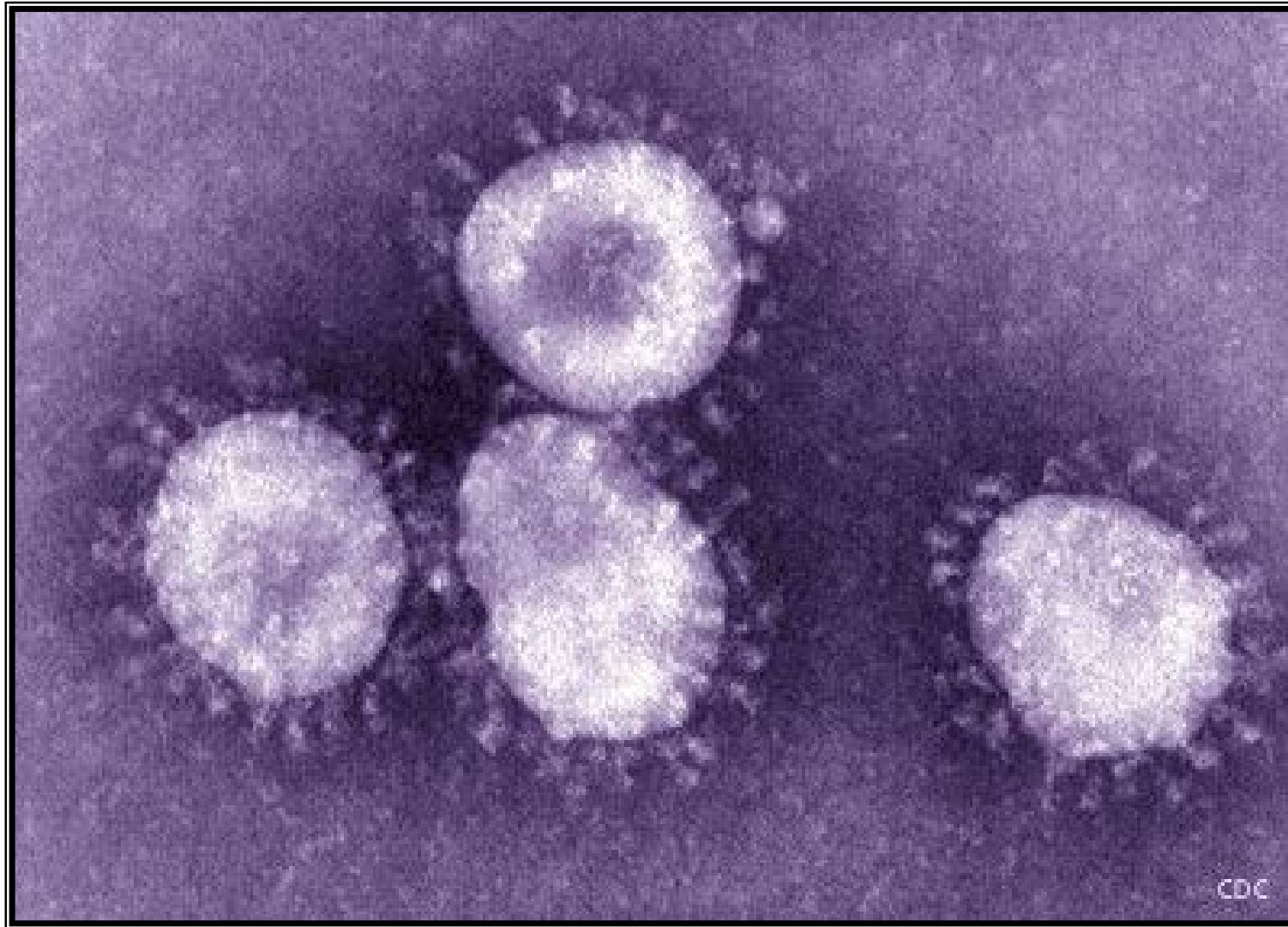


- Предмет:
 - Компаративна студија изолата *SARS CoV* (*Severe acute respiratory syndrome*, 2002.)
 - Вероватно настао мутацијама и/или генетском разменом (рекомбинације) између вируса са различитим својствима домаћина (сисари, птице)
 - Материјал: 103 изолата

Биоинф. анализа виралних генома (2)



- Циљ истраживања:
 - Установити геномски полиморфизам (разлике у структури изолата)
 - Установити еволуцију варијанти, кроз:
 - анализу и поређење нуклеотидних секвенци
 - груписање изолата по сличности секвенци: основ за филогенију овог вируса



Вирус *SARS*

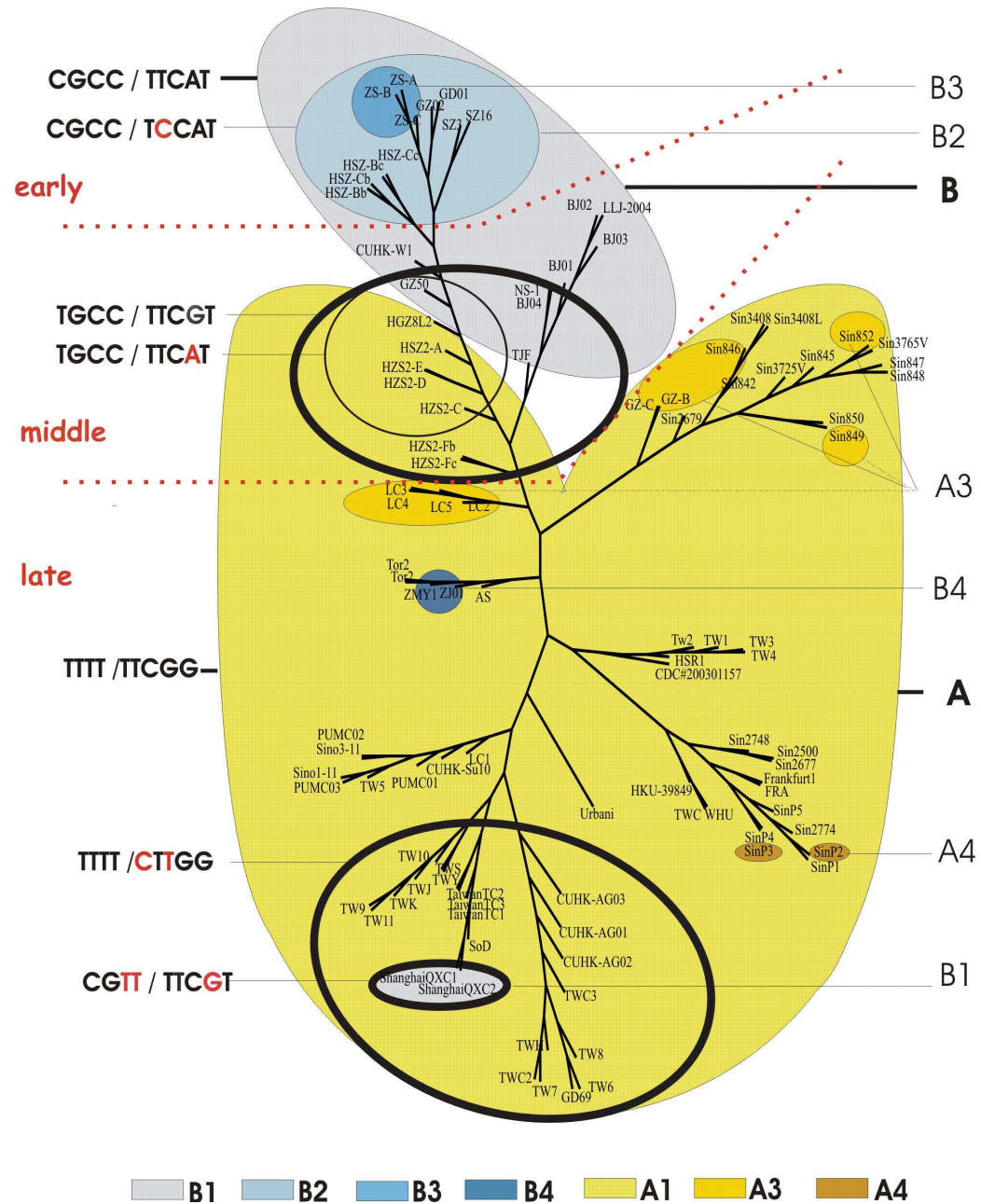
- тачке по ободу су протеини *S*, одговорни за везивање за домаћина

Биоинф. анализа виралних генома (4)



- Резултати:
 - Алгоритам анализе сличности секвенци
 - Класификација
 - *N*-грамска анализа
 - Мутациона анализа
 - Анотација и анализа *S* протеина

- Класификација
 - према степену и врсти одступања од “просечног” изолата
 - жуто/плаво
 - филогенетска сличност
 - стабло изолата



Research article

Open Access

Bioinformatics analysis of SARS coronavirus genome polymorphism

Gordana M Pavlović-Lažetić¹, Nenad S Mitić*¹ and Miloš V Beljanski²

Address: ¹Faculty of Mathematics, University of Belgrade, P.O.B. 550, Studentski trg 16, 11001 Belgrade, Serbia and Montenegro and ²Institute of General and Physical Chemistry, P.O.B. 551, Studentski trg 16, 11001 Belgrade, Serbia and Montenegro

Email: Gordana M Pavlović-Lažetić - gordana@matf.bg.ac.yu; Nenad S Mitić* - nenad@matf.bg.ac.yu; Miloš V Beljanski - mbel@matf.bg.ac.yu

* Corresponding author

Published: 25 May 2004

Received: 24 December 2003

BMC Bioinformatics 2004, 5:65

Accepted: 25 May 2004

This article is available from: <http://www.biomedcentral.com/1471-2105/5/65>

© 2004 Pavlović-Lažetić et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: We have compared 38 isolates of the SARS-CoV complete genome. The main goal was twofold: first, to analyze and compare nucleotide sequences and to identify positions of single nucleotide polymorphism (SNP), insertions and deletions, and second, to group them according to sequence similarity, eventually pointing to phylogeny of SARS-CoV isolates. The comparison is

SARS-CoV Genome Polymorphism: A Bioinformatics Study

Gordana M. Pavlović-Lažetić^{1*}, Nenad S. Mitić¹, Andrija M. Tomović², Mirjana D. Pavlović³, and Miloš V. Beljanski³

¹ Faculty of Mathematics, University of Belgrade, 11001 Belgrade, Serbia and Montenegro; ² Friedrich Miescher Institute for Biomedical Research, CH-4058 Basel, Switzerland; ³ Institute of General and Physical Chemistry, 11001 Belgrade, Serbia and Montenegro.

A dataset of 103 SARS-CoV isolates (101 human patients and 2 palm civets) was investigated on different aspects of genome polymorphism and isolate classification. The number and the distribution of single nucleotide variations (SNVs) and insertions and deletions, with respect to a “profile”, were determined and discussed (“profile” being a sequence containing the most represented letter per position). Distribution of substitution categories per codon positions, as well as synonymous and non-synonymous substitutions in coding regions of annotated isolates, was determined, along with amino acid (a.a.) property changes. Similar analysis was performed for the spike (S) protein in all the isolates (55 of them being predicted for the first time). The ratio K_a/K_s confirmed that the S gene was subjected to the Darwinian selection during virus transmission from animals to humans. Isolates from the dataset were classified according to genome polymorphism and genotypes. Genome polymorphism yields to two groups, one with a small number of SNVs and another with a large number of SNVs, with up to four subgroups with respect to insertions and deletions. We identified three basic nine-locus genotypes: TTTT/TTCGG, CGCC/TTCAT, and TGCC/TTCGT, with four subgenotypes. Both classifications proposed are in accordance with the new insights into possible epidemiological spread, both in space and time.

2. Секундарна структура протеина



- Како на основу примарне структуре предвидети секундарну структуру протеина?
- Проблем није сасвим решив:
 - један исти протеин у различитим околностима може имати различиту просторну структуру
 - мерења просторне структуре нису сасвим тачна
- Методи:
 - биохемијски – минимизација енергије структуре
 - статистички – примена учених зависности
 - машинско учење

Секундарна структура протеина (2)



- Циљ истраживања:
 - Установити зависност типа СС од врсте АК
- Претходни резултати су из периода 1976-1983.
 - тада је било много мање познатих структура
 - примењен математички метод има алтернативе

Секундарна структура протеина (3)



- Резултати:
 - Уочена значајна статистичка зависност за 18 од 20 АК
 - разликује се на неколико значајних места од старих резултата
 - неке разлике су због већег узорка, а неке због метода
 - По први пут уочена непосредна веза између добијених статистичких зависности и структурних особина АК
 - веза је изузетно природна и логична
 - штавише, неки истраживачи су раније претпостављали неке елементе ових правила али то није било потврђено резултатима

MIODRAG ŽIVKOVIĆ¹
SAŠA MALKOV^{1*}
SNEŽANA ZARIĆ²
MILENA VUJOŠEVIĆ-JANIČIĆ¹
JELENA TOMAŠEVIĆ¹
GORAN PREDOVIĆ¹
NOVICA BLAŽIĆ¹
MILOŠ V. BELJANSKI³

¹Department of Mathematics,
University of Belgrade, Beograd,
Serbia and Montenegro

²Department of Chemistry,
University of Belgrade, Beograd,
Serbia and Montenegro

³Institute of General and
Physical Chemistry, Beograd,
Serbia and Montenegro

SCIENTIFIC PAPER

547.96:547.466:519.23

STATISTICAL DEPENDENCE OF PROTEIN SECONDARY STRUCTURE ON AMINO ACID BIGRAMS

The statistical dependence of protein secondary structure on amino acid bigram frequencies was studied.

Proteins in the PDBSELECT subset of the Protein Data Bank database were investigated. Protein secondary structures were determined using DSSP software. The conditional probabilities of protein secondary structures were calculated and presented. The results on bigrams show the frequencies of all the possible bigrams in all secondary structure types. These results elucidate some factors important for the prediction of the secondary structures of proteins based on the amino acid sequence.

Key words: Amino–acid pairs, Protein secondary structure, Bigram, Bigram frequencies.

Pharmacology and biotechnologies using protein engineering strongly depend on the prediction of the generic protein function and structure, since the protein function is related to its structure. Because detailed protein 3D structure determination is a very costly process, there is extensive development of methods for the prediction of protein structure based on the amino acid sequence. In many of these methods, the first step is the prediction of the protein secondary structure.

bigrams can be used as a tool for efficient protein comparison and classification [14].

The occurrence of amino acids and amino acid pairs in different positions in α -helices has been studied by Goliaei et al. [15] and Engel et al. [16]. Exceptional amino acid pairs are identified as pairs, the frequencies of which substantially differ from the expected values. Statistical data on the amino acid pair compatibility of both spatially nearest neighbors and adjacent residues

A reexamination of the propensities of amino acids towards a particular secondary structure: classification of amino acids based on their chemical structure

Saša N. Malkov · Miodrag V. Živković ·
Miloš V. Beljanski · Michael B. Hall · Snežana D. Zarić

Received: 7 December 2007 / Accepted: 8 April 2008 / Published online: 27 May 2008
© Springer-Verlag 2008

Abstract The correlation between the primary and secondary structures of proteins was analysed using a large data set from the Protein Data Bank. Clear preferences of amino acids towards certain secondary structures classify amino acids into four groups: α -helix preferrers, strand preferrers, turn and bend preferrers, and His and Cys (the latter two amino acids show no clear preference for any secondary structure). Amino acids in the same group have similar structural characteristics at their $C\beta$ and $C\gamma$ atoms that predicts their preference for a particular secondary structure. All α -helix preferrers have neither polar heteroatoms on $C\beta$ and $C\gamma$ atoms, nor branching or aromatic group on

Keywords Amino acid · Protein ·
Protein secondary structure · Statistical correlation

Introduction

The conformational preferences of amino acids are very important for understanding conformational interactions in proteins. Moreover, when used as propensities they can be helpful in predicting secondary and tertiary structures of proteins. There are many methods of addressing protein folding, and many of them use information regarding the

A Reexamination of Correlations of Amino Acids with Particular Secondary Structures

Saša N. Malkov · Miodrag V. Živković ·
Miloš V. Beljanski · Srđan Đ. Stojanović ·
Snežana D. Zarić

© Springer Science+Business Media, LLC 2009

Abstract Using the data from Protein Data Bank the correlations of primary and secondary structures of proteins were analyzed. The correlation values of the amino acids and the eight secondary structure types were calculated, where the position of the amino acid and the position in sequence with the particular secondary structure differ at most 25. The diagrams describing these results indicate that correlations are significant at distances between -9 and 10 . The results show that the substituents on $C\beta$ or $C\gamma$ atoms of amino acid play major role in their preference for particular secondary structure at the same position in the sequence, while the polarity of amino acid has significant influence

on α -helices and strands at some distance in the sequence. The diagrams corresponding to polar amino acids are noticeably asymmetric. The diagrams point out the exchangeability of residues in the proteins; the amino acids with similar diagrams have similar local folding requirements.

Keywords Protein · Amino acid ·
Protein secondary structure · Statistical correlation

Abbreviations

DSSP Define secondary structure of proteins the

3. БИ анализа бактеријских генома



- Геномска острва – *Genomic Islands (GI)*
 - Специфични региони у геному многих бактерија
 - Настају хоризонталним трансфером гена (*HGT*) између бактерија
 - Доприносе прилагодљивости
- *GI* гени имају различите функције:
 - додатне метаболичке активности
 - способност симбиозе
 - антибиотичка резистенција
 - и друге

БИ анализа бактеријских генома (2)

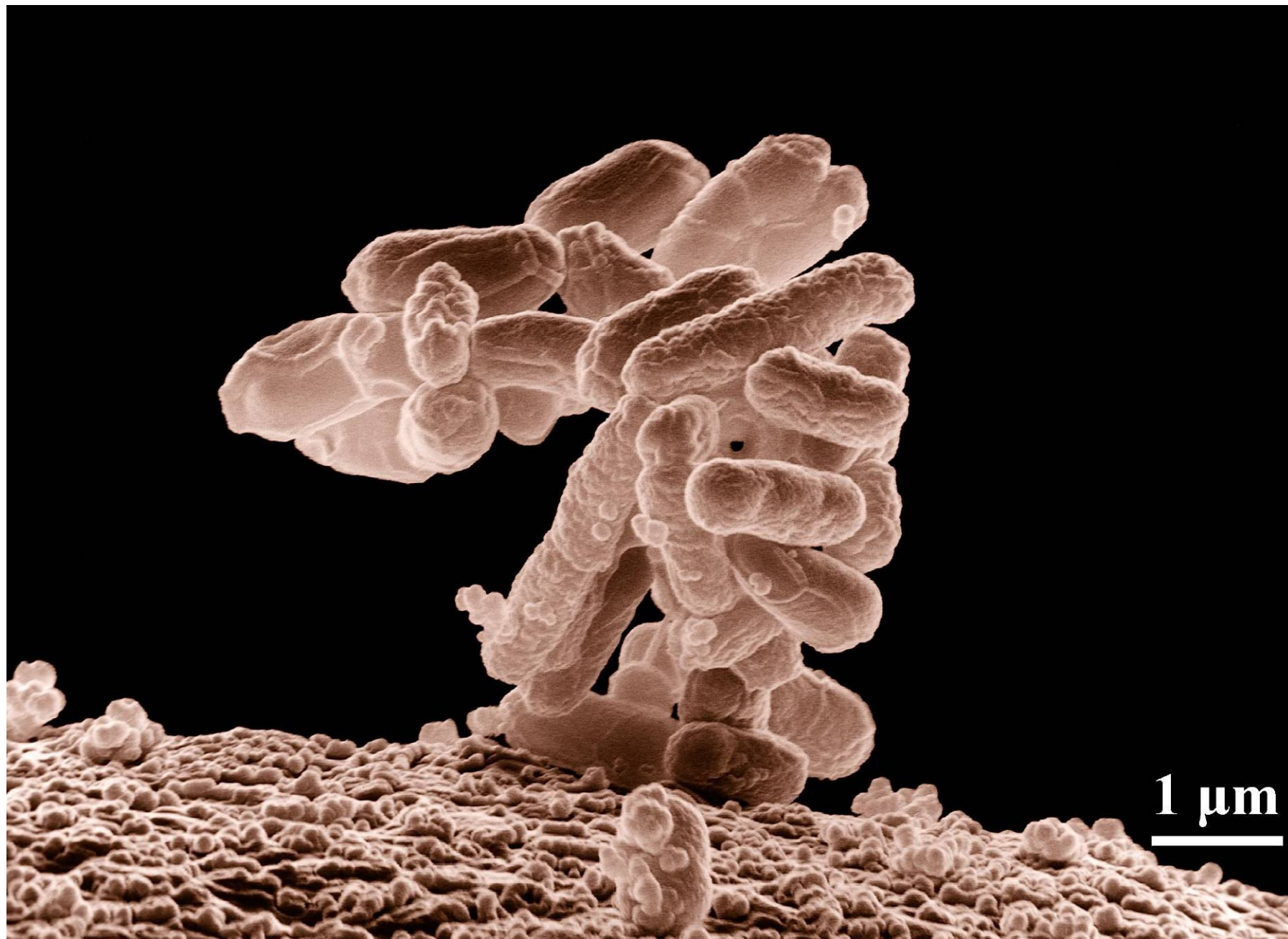


- Патогена острва (*pathogenicity islands, PAIs*)
 - *GI* са вирулентним факторима
 - Способна да заобиђу одбрамбени систем домаћина
- Карактеристике *PAI*:
 - Различита **композициона** својства, нпр. *GC* садржај
 - Различита **функционална** својства

БИ анализа бактеријских генома (3)



- Циљ истраживања
 - Прецизнија карактеризација и боље разумевање *GI*
 - Предикција *GI*
- Методологија
 - исцрпна *n*-грамска анализа већ аотираних острва
 - лингвистичка метода
- Илустрација методе
 - Геном *Escherichia coli* O157:H7 EDL933
 - Познат и значајан експериментални, медицински и биотехнолошки организам



БИ анализа бактеријских генома (5)



- Материјал:
 - Комплетан геном *E-coli O157:H7 EDL933*
 - 14 изолата из породице *Enterobacteriaceae* (*Escherichia* и *Shigella*)
 - 21 случајно изабран бактеријски геном
- Хипотеза:
 - *N*-грамска анализа може да допринесе карактеризацији, класификацији и предикцији *GI*
- Методе:
 - статистичка анализа *N*-грама

БИ анализа бактеријских генома (6)



- Резултати:
 - Бинарна класификација острва:
 - кандидати за *GI* / остало
 - Предикција острва помоћу модела
 - Провера модела – *PAI-HGT* база



doi:10.1016/j.jbi.2008.03.007 | How to Cite or Link Using DOI

Copyright © 2008 Elsevier Inc. All rights reserved.

Permissions & Reprints

Cited By in Scopus (1)

Could *n*-gram analysis contribute to genomic island determination?

Nenad S. Mitić^a, Gordana M. Pavlović-Lažetić^a and Miloš V. Beljanski^b

^aFaculty of Mathematics, University of Belgrade, Studentski trg 16, 11001 Belgrade, Serbia and Montenegro

^bIGPC, Studentski trg 16, 11001 Belgrade, Serbia and Montenegro

Received 17 September 2007. Available online 26 March 2008.

Purchase the full-text article 

- ▶ PDF and HTML
- ▶ All references
- ▶ All images
- ▶ All tables



Abstract

There are two approaches to identifying genomic and pathogenesis islands (GIPAs) in bacterial genomes: the compositional and the functional, based on DNA or protein level composition and gene function, respectively. We applied *n*-gram analysis in addition to other compositional features, combined them by *union* and *intersection* and defined two measures for evaluating the results —*recall* and *precision*. Using the best criteria (by training on the *Escherichia coli* O157:H7 EDL933 genome), we predicted GIs for 14 Enterobacteriaceae family members and for 21 randomly selected bacterial genomes. These predictions were compared with results obtained from HGT DB (based on the compositional approach) and PAIDB (based on the combined approach). The results obtained show that intersecting *n*-grams with other compositional features improves relative precision by up to 10% in case of HGT DB and up to 60% in case of PAIDB. In addition, it was demonstrated that the union of all compositional features results in maximum recall (up to 37%). Thus, the application of *n*-gram analysis alongside existing or newly developed methods may improve the prediction of GIPAs.

Keywords: *n*-Grams; Genomic islands; Bacteria; *Escherichia coli*

***n*-Gram characterization of genomic islands in bacterial genomes**

Gordana M. Pavlović-Lažetić^a✉, Nenad S. Mitić^a✉, Miloš V. Beljanski^b✉

Received 20 April 2008; received in revised form 10 September 2008; accepted 21 October 2008.

Abstract

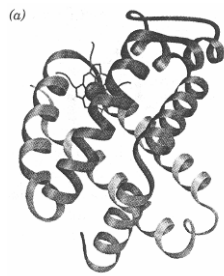
The paper presents a novel, *n*-gram-based method for analysis of bacterial genome segments known as *genomic islands* (GIs). Identification of GIs in bacterial genomes is an important task since many of them represent inserts that may contribute to bacterial evolution and pathogenesis. In order to characterize and distinguish GIs from rest of the genome, binary classification of islands based on *n*-gram frequency distribution have been performed. It consists of testing the agreement of islands *n*-gram frequency distributions with the complete genome and backbone sequence. In addition, a statistic based on the maximal order Markov model is used to identify significantly overrepresented and underrepresented *n*-grams in islands. The results may be used as a basis for Zipf-like analysis suggesting that some of the *n*-grams are overrepresented in a subset of islands and underrepresented in the backbone, or vice versa, thus complementing the binary classification. The method is applied to strain-specific regions in the *Escherichia coli* O157:H7 EDL933 genome (O-islands), resulting in two groups of O-islands with different *n*-gram characteristics. It refines a characterization based on other compositional features such as G + C content and codon usage, and may help in identification of GIs, and also in research and development of adequate drugs targeting virulence genes in them.

Abbreviations: HGT, horizontal gene transfer, GI, genomic island, OI, O-island, PAI, pathogenesis island, CU, codon usage

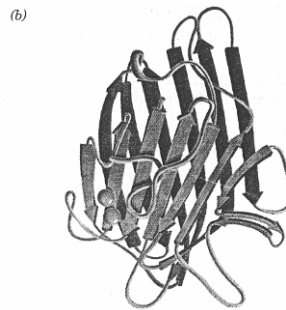
4. Би анализа неуређености протеина код прокариота



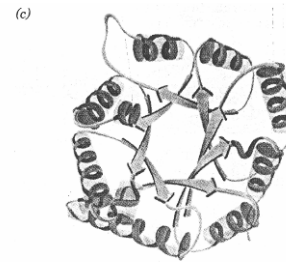
- Структуралне класе протеина:
 - Глобуларни (сферични)



Hemoglobin α

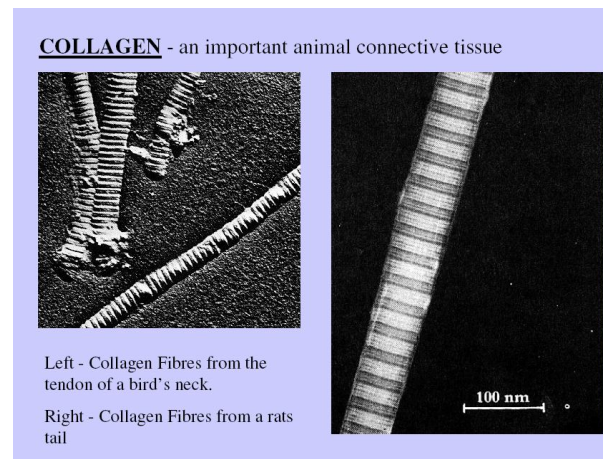


Concanavalin A



Triose Phosphate isomerase

- Фибриле (прутићи)

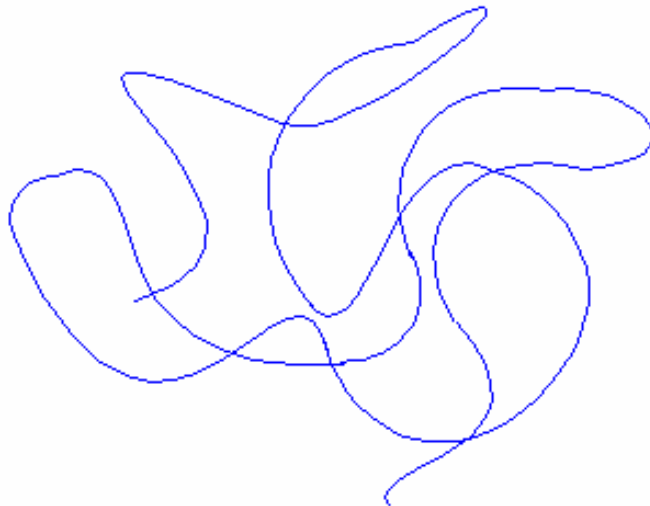


БИ анализа неуређености протеина код прокариота (2)

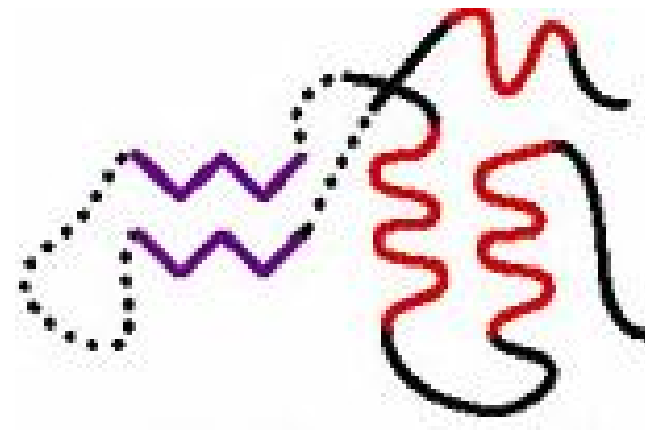


- Неуређеност протеина
 - Неки делови протеина (или цели протеини) могу бити мање / више без секундарне структуре, нпр.

насумичне петље



растопљена лоптица
(глобула)



БИ анализа неуређености протеина код прокариота (3)



- За неке неуређене протеине је
 - *експериментално или*
 - *биоинформатичком анализом*утврђено да су у вези са разним обољењима
 - канцер
 - дијабетес
 - кардиоваскуларне
 - неуродегенеративне болести

БИ анализа неуређености протеина код прокариота (4)



- Циљеви
 - Да се испита однос неуређености протеина и њихове функције код прокариота (археа и бактерија)
 - преко припадности протеина одређеним кластерима ортологих (по пореклу и функцији сличних) група протеина (*Clusters of Orthologous Groups - COG*)
 - Да се испита ниво неуређености протеина у односу на различите геномске, метаболичке и еколошке карактеристике организама

Bioinformatics analysis of disordered proteins in prokaryotes

Gordana M Pavlović-Lažetić¹ ✉, Nenad S Mitić¹ ✉, Jovana J Kovačević¹ ✉, Zoran Obradović² ✉, Saša N Malkov¹ ✉ and Miloš V Beljanski³ ✉

¹ Faculty of Mathematics, University of Belgrade, P.O.B. 550, Studentski trg 16, 11001 Belgrade, Serbia

² Center for Information Science and Technology, Temple University, 303 Wachman Hall (038-24), 1805 N. Broad St., Philadelphia, PA 19122, USA

³ Institute of General and Physical Chemistry, P.O.B. 551 Studentski trg 16, 11001 Belgrade, Serbia

✉ author email ✉ corresponding author email

BMC Bioinformatics 2011, **12**:66 doi:10.1186/1471-2105-12-66

The electronic version of this article is the complete one and can be found online at: <http://www.biomedcentral.com/1471-2105/12/66>

Received: 10 July 2010

Accepted: 2 March 2011

Published: 2 March 2011

© 2011 Pavlović-Lažetić et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background

A significant number of proteins have been shown to be intrinsically disordered, meaning that they lack a fixed 3 D structure or contain regions that do not possess a well defined 3 D structure. It has also been proven that a protein's disorder content is related to its function. We have performed an exhaustive analysis and comparison of the disorder content of proteins from prokaryotic organisms (i.e., superkingdoms Archaea and Bacteria) with respect to functional categories they belong to, i.e., Clusters of Orthologous Groups of proteins (COGs) and groups of COGs-Cellular processes (Cp), Information storage and processing (Isp), Metabolism (Me) and Poorly characterized (Pc).

5. БИ истраживање података



- Примена техника истраживања података (*Data Mining*) на биолошке податке
 - Протеомика и имуноинформатика
 - Геномика
- Акценат је на укључивању студената у рад
 - магистарски радови
 - мастер радови

БИ истраживање података (2)



- Истраживања података – протеомика
 - Одређивање растојања аминокиселина у протеинима на основу описа у бази *PDB*
 - Невена Петровић, мастер рад
 - Примена техника ИП-а у циљу успостављања корелације између антигених региона и неуређених делова протеина
 - мр Даворка Јандрлић, магистарски рад
 - Одређивање понављајућих секвенци у протеинима
 - Мирјана Вуковић, мастер рад



- Истраживања података – геномика
 - Груписање генома помоћу понављајућих секвенци и палиндрома
 - мр Горана Дацић, магистарски рад
 - Кластер анализа бактеријских геномских острва
 - мр Слободанка Марјановић, магистарски рад
 - Примена класификације на н-грамску анализу генома
 - мр Милена Шошић, магистарски рад

БИ истраживање података (3)



- Истраживања података – геномика (наставак)
 - Асемблирање генома на основу скупа контига
 - Дарко Живановић, мастер рад
 - Одређивање позиције *riboswitch* секвенци у геному *Lactobacillus paracasei subsp. paracasei BGSJ2-8*
 - Александар Стефановић, мастер рад
 - Развој нове методе за одређивање елемената за контролу генске експресије код бактерија
 - Марко Стојићевић, мастер рад

Група за биоинформатику



Текуће активности

Реконструкција геномских секвенци



- Основни кораци
 - припрема материјала
 - колеге из Института за молекуларну генетику и генетичко инжењерство, ИБИСС
 - секвенцирање
 - поравнавање секвенци
 - у сарадњи са ИБИСС
 - спајање контига
 - Дарко Живановић, колеге из Института за молекуларну генетику и генетичко инжењерство
 - анотација гена
 - Одређивање елемената за контролу генске експресије код бактерија
 - Марко Стојићевић, Александар Стефановић, колеге из Института за молекуларну генетику и генетичко инжењерство



- Корелација употребе кодона са кластерима ортологих гена
 - Александар Зељић, Милан Тодоровић
- Карактеризација кластера ортологих гена помоћу н-грамске анализе
 - Улфета Маровац (нуклеотидни н-грами)
 - Самаира Алмуктар (протеински н-грами)
- Предвиђање позиција геномских острва класификацијом н-грама
 - Милена Шошић
- Предвиђање позиција геномских острва и генских-интегенских секвенци кластеровањем карактеристика н-грама
 - Слободанка Марјановић



- Интеракција између аминокиселина у протеинима
 - Мирјана Маљковић
- Груписање генома помоћу понављајућих секвенци и палиндрома
 - Ана Јеловић, Милица Кнежевић, Мирјана Вуковић
- Корелација између антигених региона и неуређених делова протеина
 - Даворка Јандрлић
- Анализа неуређености протеина на проширеном узорку и са проширеним скупом метода
- Истраживање еко-карактеристика бактеријских генома

Група за биоинформатику



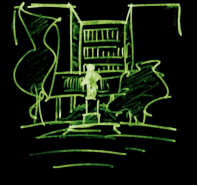
Планови за наредни период

Планови за наредни период



- Наставак истраживања у више смерова
 - рад у оквиру пројекта 174021
 - руководилац др Предраг Јаничић
- Школовање младих истраживача
 - докторске студије
 - завршни радови магистарских и мастер студија
 - настава на предмету Биоинформатика на докторским и мастер студијама
 - летње/зимске школе из Биоинформатике

Планови за наредни период



- Редовно одржавање данас отвореног Семинара за биоинформатику
- Одржавање скупа Истраживање података у биоинформатици
 - скуп планиран за јун 2012. године
- Часопис из биоинформатике

Докторске дисертације у изради



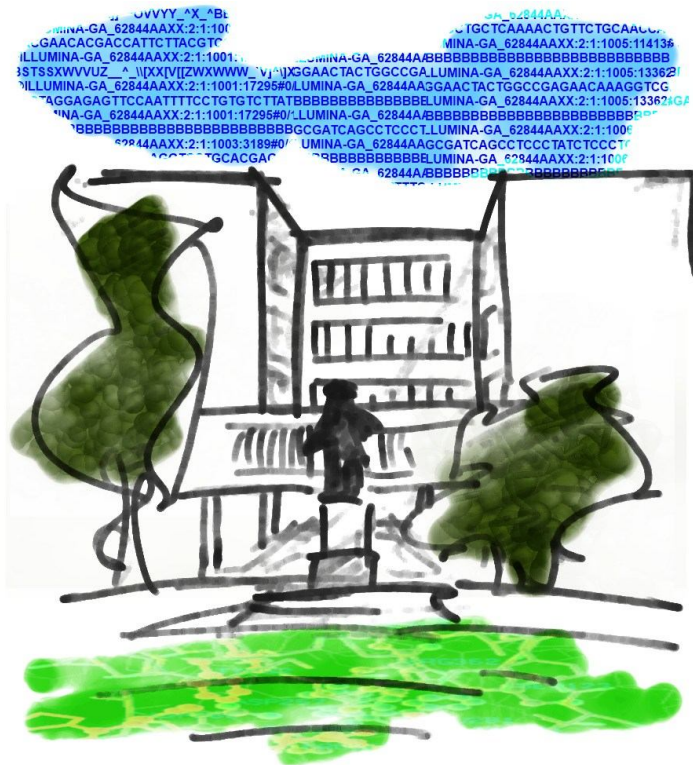
- мр Јелена Хаџи-Пурић
 - Истраживање података у биомедицини: математички модели и примене
 - класификација и предвиђање таргет-гена за одређене класе болести и лекова
- мр Јелена Граовац
 - Прилог методама класификације текста: математички модели и примене
 - један од задатака је класификовање говора (текстова) оболелих од неуродегенеративних болести
- мр Весна Пајић
 - Модели коначних стања у екстракцији информација
 - “Употреба научних енциклопедија као ресурса – креирање и допуна базе података о микроорганизмима” (радни наслов једног од поглавља)

Докторске дисертације у изради



- Јована Ковачевић
 - Модели предвиђања у биоинформатици – анализа, процена и примене (радни наслов)
 - анализа, карактеризација и предвиђање неуређених протеина и региона у различитим функционалним групама.
- Предраг Стефановић (мастер рад)
 - Дескриптивни модели истраживања података – примена у биоинформатици
 - корелација између степена неуређености аминокиселина и неких од њихових физичко-хемијских својстава

Хвала на пажњи!



Група за биоинформатику

Универзитет у Београду

Математички факултет