

Опис стручног курса

Алгоритамски алати у биоинформатици

II семестар

1) Циљ курса

Након одслушаног курса, студент ће бити упознат са основним алгоритмима који се користе у биоинформатичким алатима и њиховом имплементацијом. Овладаће основним појмовима везаним за биоинформатичку анализу и детаљније упознати софтверске алате који се најчешће користе у биоинформатичкој пракси.

2) Очекивана предзнања

- Основе програмског језика C/C++
- Основе програмског језика Python
- Основе програмског језика Java
- Основи алгоритмике

3) Технологије

- C/C++
- Python
- Java
- Bash/Unix

4) Теме курса

- **Кратак увод и историјат биоинформатичких алгоритама:**
 - Основни биолошки појмови и секвенцирање генома – 2 ч
 - Секундарна биоинформатичка анализа – 2 ч
- **Различити алати за поравнање (сличности и разлике)**
 - Мапирање генома – 1 ч
 - Графовски алгоритми – 1 ч
 - Динамичко програмирање – 1 ч
 - BWA-MEM и BWA-MEM2: сличности, разлике и упоређивање перформанси – 3 ч
- **Алати и алгоритми за позивање варијанти**
 - Позивање варијанти – основни појмови – 1 ч
 - Алгоритми алата: HaplotypeCaller – 2 ч
 - Алгоритми алата: FreeBayes (примена Бајесове теореме) – имплементација – 2 ч

- Упоређивање перформанси алата за позивање варијанти – 1 ч
- **Алати и алгоритми за позивање соматских варијанти**
 - Соматске варијанте и тумор – увод – 1 ч
 - Разлика у алгоритмима позивања регуларних и соматских варијанти – 1 ч
 - Алгоритми алата: Mutect2 и Strelka2 – упоређивање перформанси и разлика – 4 ч
- **Филогенетска анализа и метагеномика**
 - Увод у филогенетску анализу и метагеномику – 1 ч
 - Преглед алата: Kraken2 – имплементација хеш табеле – 2 ч
 - Ортолози и паралози – увод – 1 ч
 - Идентификација ортолога и базе података ортологија – преглед алгоритама – 2 ч

5) Литература (на енглеском језику)

- Phillip Compeau, Pavel Pevzner: Bioinformatics Algorithms: An Active Learning Approach, ISBN 0990374637, 9780990374633, web: [Bioinformatics Algorithms: Learn Computational Biology Online](#)
- Heng Li, Richard Durbin: Fast and accurate short read alignment with Burrows–Wheeler, *Bioinformatics*, Volume 25, Issue 14, July 2009, Pages 1754–1760, web: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2705234/>
- [GitHub - lh3/bwa: Burrow-Wheeler Aligner for short-read alignment \(see minimap2 for long-read alignment\)](#)
- [GitHub - bwa-mem2/bwa-mem2: The next version of bwa-mem](#)
- <https://gatk.broadinstitute.org/hc/en-us/articles/360035531412-HaplotypeCaller-in-a-nutshell>
- Erik Garrison, Gabor Marth, Haplotype-based variant detection from short-read sequencing, *Genomics (q-bio.GN)*; *Quantitative Methods (q-bio.QM)*, arXiv: 1207.3907, web: <https://github.com/freebayes/freebayes>
- Koboldt, D.C. Best practices for variant calling in clinical sequencing. *Genome Med* 12, 91 (2020). <https://doi.org/10.1186/s13073-020-00791-w>
- Chang Xu, A review of somatic single nucleotide variant calling algorithms for next-generation sequencing data, DOI: <https://doi.org/10.1016/j.csbj.2018.01.003>
- Chen, Z., Yuan, Y., Chen, X. et al. Systematic comparison of somatic variant calling performance among different sequencing depth and mutation frequency. *Sci Rep* 10, 3501 (2020). <https://doi.org/10.1038/s41598-020-60559-5>
- <https://gatk.broadinstitute.org/hc/en-us/articles/360037593851-Mutect2>
- Kim, S., Scheffler, K. et al. (2018) Strelka2: fast and accurate calling of germline and somatic variants. *Nature Methods*, 15, 591-594. doi:10.1038/s41592-018-0051-x, GitHub: <https://github.com/Illumina/strelka/blob/v2.9.x/docs/developerGuide/README.md>
- Wood, D.E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol* 20, 257 (2019). <https://doi.org/10.1186/s13059-019-1891-0>, <https://github.com/DerrickWood/kraken2/wiki>
- Roth, A.C., Gonnet, G.H. & Dessimoz, C. Algorithm of OMA for large-scale orthology inference. *BMC Bioinformatics* 9, 518 (2008). <https://doi.org/10.1186/1471-2105-9-518>
- Justin B Miller, Brandon D Pickett, Perry G Ridge, JustOrthologs: a fast, accurate and user-friendly ortholog identification algorithm, *Bioinformatics*, Volume 35, Issue 4, February 2019, Pages 546–552, <https://doi.org/10.1093/bioinformatics/bty669>

- Guillaume Fertin, Falk Hüffner, Christian Komusiewicz, Matching algorithms for assigning orthologs after genome duplication events, Computational Biology and Chemistry, Volume 74, June 2018, Pages 379-390, <https://doi.org/10.1016/j.compbiolchem.2018.03.015>