**Program: Systems biology**

**Admission: year 2019**

**Published: January, 2020**

|  |  |  |
| --- | --- | --- |
| **Topic with description** | **Contact person** | **Contact email** |
| **1) TITLE:** **Computational mapping of nucleosome positions across variety of organisms**  **ABSTRACT.** Nucleosome positioning DNA sequence patterns (NPS) - usually distributions of particular dinucleotides or other sequence elements in nucleosomal DNA - at least partially determine chromatin structure and arrangements of nucleosomes that in turn affect gene expression. Statistically, NPS are defined as oscillations of the dinucleotide periodicity with about 10 base pairs (bp) which reflects the double helix period. Recently few distinctive patterns in nucleosomal sequences were observed that can be termed as packing and regulatory referring to distinctive modes of chromatin function [1]. Our working hypothesis for the future studies is that packing patterns tend to be preferred by evolutionary lower organisms and regulatory ñ by higher organisms.  Given vast amount of publicly available nucleosome maps in various organisms [2,3] it is possible to perform a computational mapping of nucleosomes at the loci of gene promoters in various organisms using already available dnpatterntools v1.0 software (https://github.com/erinijapranckeviciene/dnpatterntools).  Although such mapping was attempted previously, the results were inconclusive because of a small number of available nucleosomal DNA seqences at that time [4].  Results of a new computational experiment using data from NucMap [4] and NucPosDB[3] would provide more iformation on pattern preferences across organisms and subsequently better prediction of nucleosome's position at a specific loci. This would eventually lead to the better tools of computational mapping of nucleosome positions.  **REFERENCE**  1. Nucleosome positioning sequence patterns as packing or regulatory. Erinija Pranckeviciene, Sergey Hosid, Nathan Liang, Ilya Ioshikhes. bioRxiv 755272;doi:https://doi.org/10.1101/755272  [Accepted for publication in PLOS Computational Biology]  2. Teif V.B. (2016). Nucleosome positioning: resources and tools online.  Briefings in Bioinformatics 17, 745-757. [https://generegulation.org/nucleosome-positioning-database/]  3. Yongbing Zhao, Jinyue Wang, Fang Liang, Yanxia Liu, Qi Wang, Hao Zhang, Meiye Jiang, Zhewen Zhang, Wenming Zhao, Yiming Bao, Zhang Zhang, Jiayan Wu, Yan W Asmann, Rujiao Li, Jingfa Xiao, NucMap: a database of genome-wide nucleosome positioning map across species, Nucleic Acids Research, Volume 47, Issue D1, 08 January 2019, Pages D163ñD169, https://doi.org/10.1093/nar/gky980  [https://bigd.big.ac.cn/nucmap/]  4. Liang N, Pranckevicience E and Ioshikhes I. Matching nucleosome patterns to the complexity of the organism  [version 1; not peer reviewed]. F1000Research 2019, 8:127 (poster) (https://doi.org/10.7490/f1000research.1116423.1) | Erinija Pranckevičienė | erinija.pranckeviciene@mf.vu.lt |
| **2) TITLE:** **Molecular shape similarity in virtual screening for drug discovery**  **ABSTRACT.** Virtual Screening can drastically accelerate drug discovery processes. Molecular shape similarity is essential in virtual screening for drug discovery. Shape similarity is used to compare in detail the shape of a query molecule against a large database of potential drug compounds. In order to evaluate shape similarity accurately the molecules should be optimally adjusted. In this work, optimization problems for molecular shape similarity are investigated aiming at fast solution and acceptable accuracy. | Julius Žilinskas | julius.zilinskas@mif.vu.lt |
| **3) TITLE:** **Comprehensive Pathology Analytics of Renal Tissue Based on Segmentation of Multi-stained Tissue Sections**  **ABSTRACT.** The aim of this study is to develop a microscopy image analytics platform for kidney pathology quantification. At the core of the platform is the segmentation of the kidney morphology from tissue sections subjected to different immunohistochemical stainings, each stain optimizing the delineation of a different kidney morphology. Whole slide kidney microscopy images with annotations of different tissue components are available for both the machine/deep learning segmentation and the development of a comprehensive set of quantitative kidney pathology indicators. We expect the use of Python for the machine learning part and R as the statistical back-end of the analytics pipeline, and we look forward to a motivated and focused applicant to join our digital pathology team. | Arvydas Laurinavičius | arvydas.laurinavicius@vpc.lt |
| **4) TITLE:** **Patterns of adaptive positive selection in  human populations from high-density SNP data**  **ABSTRACT.** The analysis of geographically specific regions and the characterization of fine-scale patterns of genetic diversity may facilitate a much better understanding of the microevolutionary processes affecting local human populations.  Geographically specific microevolutionary processes can be inferred when exploring local patterns of population structure and adaptation within the global and historical genetic context established from large general population and ancient publicly available reference panels.  The Lithuanian population is partially isolated with genetic distinctiveness within the European context and with preserved ancient genetic composition. A small genetic differentiation among the Lithuanian ethnolinguistic groups exist (Urnikyte et al., 2019).  Aim of the practice: to analyse signatures of adaptive positive selection in high-density genotyping data previously generated in two main ethnolinguistic population groups from Lithuania.  The student will perform the main following tasks:  Task 1. Parsing and transforming data generated in six Lithuanian populations with the Illumina Infinium whole-genome SNP array to appropriate formats for subsequent analyses.  Task 2. Compiling, parsing and transforming external reference population datasets including the CEU (Utah residents with ancestry from northern and western Europe), FIN (Finnish in Finland), TSI (Toscani in Italy) and GBR (British in England and Scotland) populations as wells as additional reference populations from Africa, East Asia and South Asia available from the 1000 Genomes Project (http://www.internationalgenome.org/about/).  Task 3. Performing population structure analysis of six Lithuanian ethnolinguistic groups by PCA using different regional European and worldwide global contexts.  Task 4. Calculating kinship and inbreeding coefficients in the Lithuanian population.  Task 5. Computing selection tests based on population differentiation (XP-EHH statistic, FST ) and site frequency spectrum (Tajima’s D) from SNP genotyping phased/unphased data and using different pairwise population comparisons when appropriate to identify and distinguish signals of positive selection specific of Lithuanians from adaptations shared with other Europeans populations.  Task 6. Identifying top significant candidate genomic windows for positive selection, as wells as pathways enriched for signals of positive selection in the Lithuanian population.  Task 7. Detecting potential functional variation linked to the signatures of selection identified using different functional element annotations of the human genome and diverse in silico prediction methods (ANNOVAR,  RefSeqGene, dbSNP147 and CADD).  The student will acquire basic knowledge and skills on the analysis of genetic structure on human populations as well as on the detection of natural selection from high-density genome-wide SNP data. population genetics classic software such as Plink, shapeit, selscan, popgenome and annovar among others.  Requirements: Bioinformatic skills. | Alina Urnikytė | alina.urnikyte@mf.vu.lt |
| **5) TITLE:** **Application of Deep Neural Networks for Eye Fundus Image Analysis**  **ABSTRACT:** Aim of the study is to develop workflow for the eye fundus images analysis with the view to detect various anatomical changes. Early detection of the changes can indicate and even prevent such diseases as diabetes, glaucoma, cataract and etc. The images of known structure changes will be available as well as the consultations with the skilled ophthalmologist. Proposed solution has to be built on deep neural network application and tested. Students selecting the theme will face the Python as a primary programming language that will utilize the Keras and Tensorflow frameworks as those mostly used tools for deep neural network application with the strong community support. The student should have strong motivation to learn new skills all other knowledge will be acquired during the master study. | Povilas Treigys | povilas.treigys@mif.vu.lt |