**RUSSIAN ACADEMY OF SCIENCES**
**SIBERIAN BRANCH**

**INSTITUTE OF CYTOLOGY AND GENETICS**

SCIENTIFIC PROGRAM

And

ABSTRACTS

Of

# Young scientists school "Bioinformatics and systems biology"

**30 June – 2 July 2012**

**Novosibirsk**
**2012**

## SCIENTIFIC PROGRAM

# Young scientists school "Bioinformatics and systems biology"

Novosibirsk, Russia, 30 June – 2 July 2012
Institute of cytology and genetics SB RAS

## 30 June, Saturday

**8.00-9.30** **Registration** (Conference hall, 3rd floor)

**9.30-10.00** **Welcome speech**
*acad. Nikolai Kolchanov* (Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia)
***Review of the previous young scientists schools***
*Dr. Victoria Mironova* (Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia)

**10.00-11:30** **Lecture**
*Prof. Eric Mjolsness* (University of California, Irvine, USA)
***Quasi-equilibrium and non-equilibrium modeling of biochemical reaction networks***

**11.30-13.00** **Lecture**
*Prof. Roman Efremov* (Institute of Bioorganic Chemistry RAS, Moscow, Russia)
***Numerical experiment in molecular biophysics of proteins and membranes: tasks and solutions for proteomics***

***13.00-14.30*** ***Lunch***

**14.30-16.00** **Lecture**
*Prof. Evgeny Rogaev* (Mental Health Research Center RAMS, Moscow, Russia)
***New trends in medical genomics***

**16.00-17.00** **Lecture**
*Prof. Robert Giegerich* (Bielefeld University, Bielefeld, Germany)
***Dynamic Programming in Bellman's GAP***

**17.00-17.30 Young scientists' poster session**

**17.30-18.30 Parallel practical trainings**

1. **Computer analysis of high-throughput sequencing and chromatin-immunoprecipitation data (ChIP-seq)**
   *Dr. Yuri Orlov* (Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia)
   Conference hall (3rd floor)

2. **Multi-trait analysis of linked QTLs**
   *Prof. Yefim Ronin (*Institute of Evolution, University of Haifa, Israel)
   Room № 1343 (3rd floor)

3. **Enjoy Dynamic Programming in Bellman's GAP**
   *Prof. Robert Giegerich* (Bielefeld University, Bielefeld, Germany)
   Computer class (Basement)

**19.00 Excursion "Novosibirsk in the evening"**

## 1 July, Sunday

**9.30-10.30 Lecture**

*Dr. Maksim Zakhartsev* (University of Stuttgart; University of Heidelberg, Germany)

***Fundamentals of flux analysis***

**10.30-11.30 Lecture**

*Dr. Olga Krebs* (Heidelberg Institute for Theoretical Studies, Germany)

***Data integration in systems biology: SySMO DB on the focus***

***11.30-11.45- Coffee-break***

**11.45 -12.45 Lecture**

*Fernando Izquierdo* (HITS gGmbH, Germany)

***Parallel computations applied to phylogenetics***

***12.45- 14.30 Lunch***

**14.30-15.30 Parallel practical trainings**

1. **Data integration in systems biology: SEEK, JERM and RightFiled tools**
   *Dr. Olga Krebs* (Heidelberg Institute for Theoretical Studies, Germany)
   Room № 1343 (3rd floor)
2. **Computational proteomics for beginners**
   *Evgenij Tiys (*Institute of Cytology and Genetics, Novosibirsk, Russia)
   Library hall (2nd floor)
3. **Parallel computations applied to phylogenetics**
   *Fernando Izquierdo (HITS gGmbH, Germany)*
   Computer class (Basement)

**15.30 – 16.30 Young scientist's poster session**

**16.30-18.30 Young scientists' session "Software and algorithms development"**

*Ilya Malakhin* "Unsupervised algorithm based on wavelet analysis for extraction of information about hippocampal neuronal activity characteristics from the experimental data"

*Ilya Kiselev* "BioUML: modular modeling of complex biological systems"

*Evgenij Pozdnyak* "Tools for the describing scientific workflows"

*Zakhar Mustafin* "Haploid evolutionary constructor: parallelization and high performance simulations of prokaryotic communities' evolution"

*Fedor Kazantsev* "High Performance Computing with MGSmodeller"

**19:30 – Football game**

## 2 July, Monday

**9.00-10.30 Young scientists' session "Plant systems biology"**

*Vera Gorelova* "Study of *Arabidopsis thaliana* ornithine aminotransferase gene expression control"

*Ekaterina Sergeeva* "Towards an analysis of the structure of the short arm of 5B chromosome of the bread wheat *Triticum aestivum* L."

*Victoria Shtratnikova* "High-throughput sequencing of microRNA of roots and nodules of *Pisum sativum* L."

*Belenikin Maxim* "Sequencing and de novo transcriptome assemby of *Stellaria media* (L.) Vill."

***10.30-10.45 Coffee-break***

**10.45-11.45 Lecture**

*Dr. Daniil Naumoff, (*Institute of Microbiology, Moscow, Russia)
***Genome projects and gene annotation***

**11.45 -13.00 Young scientists' session "Computational medicine"**

*Nikolai Barducov "*Inbreeding and differently directed dynamics of ISSR-PCR and IRAP-PCR markers polymorphism in musk oxen populations*"*

*Petr Ponomarenko* "Experimental examining of prognoses *in silico* TBP binding to TATA box with SNP associated with human diseases"

*Ulyana Potapova* "NS2B/NS3 protease: analysis of allosteric effects of mutations associated with the pathogenicity of tick-borne encephalitis virus"

*Bogumil Kaczkowski* "Integrative analyses reveal novel strategies in HPV11, -16 and -45 early infection"

***13.00- 14.30 Lunch***

**14.30-15.30 Young scientists' session "Computational proteomics"**

*Anastasia Bakulina* "Molecular modeling of cytosolic part of a2-subunit of mouse V-ATPase"

*Pavel Davidovich* "E3 Ligase and p53 family proteins interaction modeling"

*Timofey Ivanisenko "*Reconstruction of the associative genetic networks based on integration of automated text-mining methods and protein-ligand interactions prediction*"*

*Sergey Pintus* "Computational evaluation of impact of amino acid substitution p.W172C on structure and function of GAP-junction protein connexin 26 and its association with hearing impairment"

***15.30-15.45 Coffee-break***

**15.45-17.00 Young scientists' session "Mathematical biology"**

*Katyakova Vera* "About cycles of functioning of model of gene network regulatory loops"

*Ulyana Zubairova* "Activation of CLV3 gene expression in model of the stem cell niche structure regulation in the shoot apical meristem"

*Elena Mamontova* "Spatially distributed modeling of bacterial communities with haploid evolutionary constructor"

*Temlyakova Evgenia* "Clustering of *E.coli* promoter electrostatic profiles"

**17.00-18.00 Parallel practical trainings**

1. **Introduction to bioinformatics: search and analysis of genomic sequences**
   *Dr. Dmitry Afonnikov (*Institute of Cytology and Genetics, Novosibirsk, Russia)
   Room № 1343 (3rd floor)
2. **Practical exercises in flux analysis**
   *Dr. Maksim Zakhartsev* (University of Stuttgart; University of Heidelberg, Germany), Computer class (Basement)

**19.00 Closing ceremony**

# Young scientists' poster session

## *Conference hall*

30 June 17.00-17.30
1 July 15.30 – 16.30

1. *Andrey Akinshin* "COMPUTER ANALYSIS OF PHASE PORTRAITS IN GENE NETWORKS MODELS
2. *Anton Bogomolov* "FLUORESCENCE IN SITU HYBRIDIZATION WITH CHROMOSOME-DERIVED DNA PROBES ON *OPISTHORCHIS FELINEUS* AND *METORCHIS XANTHOSOMUS* CHROMOSOMES WITHOUT SUPPRESSION OF REPETITIVE DNA SEQUENCES"
3. *Eugene Bushmelev* "POPULATION STUDY OF THE VARIATION IN TRIPLET DISTRIBUTIONS OBSERVED ALONGSIDE A CHROMOSOME, FOR YEAST SPECIES"
4. *Vasilina Chernova* "ANALYSIS OF TRANSCRIPTIONAL AND POSTTRANSCRIPTIONAL REGULATION OF AUXIN CARRIER *AtPIN1"*
5. *Mikhail Domrachev* "NEURAL NETWORK MODELS OF THE PROCESS OF PERCEPTION OF TEXTUAL INFORMATION"
6. *Maria Elkina* "IRAP-PCR MARKERS AND MICRONUCLEI TEST IN THE CHARACTERIZATION OF GENETIC STRUCTURE OF THE KALMYK SHEEP AND TYPES OF THE EDILBAY SHEEP"
7. *Timur Erkenov* "GENOME SCANNING OF HORSE BREEDS BY USING OF ISSR-PCR MARKERS"
8. *Timofey Ermak* "KINET – A NEW WEB DATABASE ON KINETICS DATA AND PARAMETERS FOR *E. COLI*"
9. *Zhamiga Imangalieva* "A DISCRETE DYNAMICAL SYSTEM ON A DOUBLE CIRCULANT WITH AN ADDITIVE FUNCTION OF THE VERTICES"
10. *Mikhail Kabakov* "A COMPUTER SYSTEM FOR KINETIC ANALYSIS OF GENE NETWORKS"
11. *Artem Kasianov* "*DE NOVO* SEQUENCING, ASSEMBLY AND CHARACTERIZATION OF TRANSCRIPTOME IN TETRAPLOID PLANT *CAPSELLA BURSA-PASTORIS*"
12. *Bogumil Kaczkowski* "SOMATIC COPY-NUMBER ALTERATION CAN HELP PREDICT THE TISSUE ORIGIN OF CANCERS OF UNKNOWN PRIMARY"

13. *Alexandra Klimenko* "HAPLOID EVOLUTIONARY CONSTRUCTOR: A GRAPHICAL USER INTERFACE FOR SIMULATING BACTERIAL COMMUNITIES EVOLUTION"

14. *Anna Koval* "STRUCTURAL PARAMETERS OF TRYPTOPHAN ENVIRONMENT IN *PHOTOBACTERIUM LEIOGNATHI* LUCIFERASE"

15. *Denis Korobko* "GENETICS AND DISEASE PROGRESSION OF FAMILIAL MULTIPLE SCLEROSIS IN NOVOSIBIRSK REGION OF RUSSIA"

16. *Elena Kulish* "ASSOCIATION OF ITGB3 AND GNB3 VARIANTS WITH THE DEVELOPMENT OF VASCULAR COMPLICATIONS IN PATIENTS WITH ACUTE CORONARY SYNDROME"

17. *Ilya Malakhin* "THEORETICALLY-EXPERIMENTAL RESEARCH OF VESICLE TRAFFICKING MECHANISMS IN THE SYNAPTIC PLASTICITY PROCESS"

18. *Azhar Nazhmidenova* "THE DISCRETE DYNAMIC SYSTEM ON A DOUBLE CIRCULANT WITH DIFFERENT FUNCTIONS AT THE VERTICES"

19. *Ekaterina Novoselova* "ROLE OF AUXIN DOSE-DEPENDENT CONTROL IN SPECIFICATION OF ROOT VASCULAR CELLS"

20. *Nikolas Rapin "*CLASSIFICATION OF PURIFIED BONE MARROW POPULATIONS SORTED VIA MULTICOLOR FLOW CYTOMETRY, APPLICATIONS IN ACCUTE MYELOID LEUKEMIA"

21. *Maria Savina* "MATHEMATICAL MODEL OF AUXIN RESPONSIVE REPORTER DR5 ACTIVITY IN PLANT CELL"

22. *Alexander Semenychev* "DISTRIBUTED RESTFUL-WEB-SERVICES FOR THE RECONSTRUCTION AND ANALYSIS OF GENE NETWORKS"

23. *Alexey Sokolov* "THE UNDERLYING MECHANISMS OF REPROGRAMMING OF HUMAN UMBILICAL VEIN ENDOTHELIAL CELLS (HUVEC)"

24. *Yuri Vaskin* "DATA MINING TOOL FOR ANALYSIS OF REGULATORY REGIONS OF GENES: INTEGRATION OF ExpertDiscovery AND UGENE"

# Lectures

***Lecture 1.***

**Quasi-equilibrium and non-equilibrium modeling of biochemical reaction networks**

Prof. Eric Mjolsness (Department of Information and Computer Science; Department of Mathematics University of California, Irvine, USA)

Basic tools of mathematical physics, including equilibrium and nonequilibrium statistical mechanics, sey **new application in computational systems biology** generally and biochemical reaction networks in particular. For example, transcriptional regulation networks and allosteric enzymes can often be modeled with quasi-equilibrium or simple nonequilibrium methods. General reaction networks can be decomposed into dynamics given by equilibrium free energies, thermodynamic forces on cycles, and transition state free energies. Relate technical tools will be introduced as well.

**Lecture 2.**

**Numerical experiment in molecular biophysics of proteins and membranes: tasks and solutions for proteomics**

Prof. Roman Efremov (head of laboratory of biomolecular modeling, Institute of bioorganic chemistry RAS, Moscow, Russia)

Nowadays computational modeling has the same impact in research of molecular-genetics events as traditional experimental methods. In numerical experiment a unique information about the structure, dynamics and the mechanisms of biomolecules and complexes can be often received. From the level of single small molecules which was possible to analyzed earlier, modern In silico methods allow analyze in details more complex molecular substances: proteins and biomembraines. The physical models of such objects will be discussed on the lecture. Also the modern computational methods which are in use for a large amount of tasks in molecular biophysics will be introduced.

**Lecture 3.**

**New trends in medical genomics**

Prof. Evgeny Rogaev (head of laboratory of brain molecular genetics in Mental Health Research Center RAMS, Moscow, Russia; head of laboratory in University of Massachusetts Medical School)

**Lecture 4.**

**Dynamic Programming in Bellman's GAP**

Prof. Robert Giegerich (Bielefeld University, Germany)

Applications of the dynamic programming paradigm are ubiquitous computer science, and especially so in bioinformatics. The discipline of algebraic dynamic programming liberates programmers from cumbersome coding and debugging,

as algorithms can be described on a declarative level of abstraction. Teaching DP algorithms becomes more rewarding, as crucial ideas are better exhibited in a more abstract representation. The algebraic discipline underlies a good number of bioinformatics tools, such as RNAshapes, RNAlishapes, RNAhybrid, PknotsRG, KnotInFrame, Locomotif, and pKiss, (see bibiserv.cebitec.uni-bielefeld.de), some of which are used quite widely.

Quite in contrast to this success, the proliferation of algebraic dynamic programming as a method remains marginal. It has been hindered by the fact that its original implementation was based on a Haskell-derived syntax and provided only marginal compiler optimizations.

***Lecture 5.***

**Computer analysis of high-throughput sequencing and chromatin-immunoprecipitation data (ChIP-seq)**

Dr. Yuriy L Orlov (Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia)

Advances in high throughput sequencing technologies coupled with chromatin-immunoprecipitation (ChIP-seq) allow identification of transcription factor (TF) binding sites in genome scale. Such genome wide TF binding maps in human include Oct4, Sox2 and Nanog in stem cells and cancer related transcription factors oestrogen receptor ERα, FOXA1. TF binds to only a small fraction of sequence motifs or eligible binding sites in the genome. Key problem of gene expression regulation analysis is detection of functional binding sites responsible for gene activation. We discuss here interplay between TF binding and chromatin landscape revealed by ChIP-sequencing together with statistical issues.

***Lecture 6.***

**Fundamentals of flux analysis**

Dr. Maksim Zakhartsev (Institute of Biochemical Engineering, University of Stuttgart; Institute of Pharmacy and Molecular Biotechnology, University of Heidelberg, Germany)

Analysis of intracellular fluxes in microbial cells is one of the major research tools in modern Metabolic Engineering and Systems Biology. In this lecture we will review modern analytical techniques to perform experiments and measure metabolic fluxes as well as consider fundamentals of unstructured and structured metabolic models and methods for their mathematical analysis. We are going to use MATLAB and some specialized software (for example InSilico Discovery) to perform an analysis of experimental data.

***Lecture 7.***

**Data integration in systems biology: SySMO DB on the focus**

Dr. Olga Krebs (Heidelberg Institute for Theoretical Studies, Germany)

*Lecture 8.*

**Parallel computations applied to phylogenetics**

Fernando Izquierdo (HITS gGmbH, Germany)

Due to the rapid pace of molecular data accumulation and the slower increase in CPU speeds, parallel computing has established as a demanded standard technique in Bioinformatics. We will introduce some of the main paradigms in parallel computer architectures and show how these can be applied to the field of computational phylogenetics to enable the analysis of large datasets.

*Lecture 9.*

**Genome projects and gene annotation**

Dr. Daniil Naumoff, *(*Institute of Microbiology, Moscow, Russia)

Several dozen million of protein sequences are available now. The majority of them have been obtained during genome projects. As a rule, these proteins were not characterized experimentally therefore it is necessary to predict their functions by bioinformatics approaches. Progress in genome sequencing as well as the gene annotations will be discussed during the lecture. Special attention is given to protein misannotations in the modern databases.

# Young scientists school
# ABSTRACTS

# COMPUTER ANALYSIS OF PHASE PORTRAITS IN GENE NETWORKS MODELS

A.A. Akinshin
*Altai State Technical University, Barnaul, Russia*
*e-mail: andrey.akinshin@gmail.com*

**Key words:** *gene networks models, nonlinear dynamical systems, stationary points, periodic trajectories, numerical modeling)*

*Motivation and Aim:* developing a computer program that will build the graphical representation of the phase portraits in gene networks models (2d- and 3d- projections on different vectors), search stationary points, calculate eigenvalues and eigenvectors, simulate the trajectory.

*Methods and Algorithms:* gradient descent, Runge–Kutta methods, QR decomposition.

*Results:* the program has successfully developed and widely used for various numerical experiments with natural and hypothetical gene networks models.

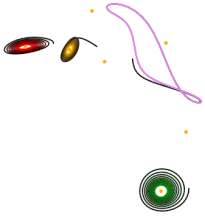

Fig. 1. 3-D model with 7 stationary points.

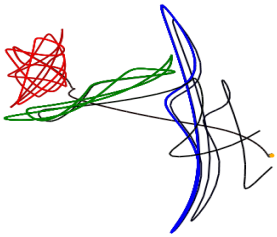

Fig. 2. Three cycles in 15-D symmetric model.

*Conclusion:* the program identifies new regularity of the functioning in some gene networks models. The results are used to describe a hypothetical and natural gene networks (e.g., the gene network determining development of drosophila melanogaster mechanoreceptors).

*Availability:* The program currently is in testing mode and is unavailable over the Internet. To get the trial version should write to the author by email.

*Acknowledgements:* the author makes a great appreciation V.P. Golubyatnikov for the valuable advice on the mathematical background.

*References:*

1. V.A.Likhoshvai, V.P.Golubyatnikov et al. (2008) Theory of gene networks, In: *System computerized biology*, (N.A.Kolchanov, S.S.Goncharov), Novosibirsk, SB RAS, 397-480.
2. A.A.Akinshin, V.P.Golubyatnikov (2011) Some mathematical and computational problems of bioinformatics, *National scientific-technical conference of students and young scientists "Science and the youth"*, Barnaul, 6-9.
3. A.A.Akinshin, V.P.Golubyatnikov (2011) Mathematical and computer modeling of periodic modes of gene networks, *International Scientific and Practical Conference of Students and Young Scientists "Modern Techniques and Technology"*, Tomsk, 283-284.

# INTEGRATIVE ANALYSES REVEAL NOVEL STRATEGIES IN HPV11, -16 AND -45 EARLY INFECTION

Bogumil Kaczkowski[a]*, Maria Rossing[b], Ditte Andersen[c], Anita Dreher[c], Melissa A. Visser[c], Ole Winther[a,d], Finn Cilius Nielsen[b], Bodil Norrild[c]

*a) The Bioinformatics Centre, Department of Biology and Biomedical Research and Innovation Centre, Copenhagen University, Ole Maaloes Vej 5, 2200 Copenhagen, Denmark*

*b) Department of Clinical Biochemistry, Copenhagen University Hospital, Blegdamsvej 5, 2100 Copenhagen, Denmark*

*c) Institute of Cellular and Molecular Medicine, DNA Tumor Virus Laboratory, University of Copenhagen, Panum Institute, Blegdamsvej 3, 2200 Copenhagen, Denmark*

*d) DTU Informatics, Technical University of Denmark, 2800 Lyngby, Denmark*

The interaction between human papillomavirus (HPV) and host cells is not well understood. We investigate the early stage of HPV infections by global expression profiling in a cell model, in which HaCaT cells were transfected with HPV11, HPV16 or HPV45 genomes. We report the differential expression of genes not previously implicated in HPV biology, such as the PSG family and ANKRD1, and of genes implicated in the biology of other viruses, e.g. MX1, IFI44 and DDX60. Carcinogenesis-related genes, e.g. ABL2, MGLL and CYR61, were upregulated by high-risk HPV16 and -45. The integrative analysis revealed the suppression of DNA repair by HPV11 and -16, and downregulation of cytoskeleton genes by all HPV types. Various signalling pathways were affected by the HPVs: IL-2 by HPV11; JAK-STAT by HPV16; and TGF-β, NOTCH and tyrosine kinase signalling by HPV45. This study uncovered novel strategies employed by HPV to establish infection and promote uncontrolled growth.

# NEURAL NETWORK MODELS OF THE PROCESS OF PERCEPTION OF TEXTUAL INFORMATION

M.A. Domrachev
*Kemerovo State University, Kemerovo, Russia*
*e-mail: m.domrachev.scientist@gmail.com*

**Key words:** *Markov chain, Edelman assortment, neural network, textual information.*

*Motivation and Aim.* The aim was to study and modeling of the process of perception of new textual information. The materials for the study of texts were advertising brochures. One of the objectives of the study was the selection of a new type of perception of textual information.

*Methods and Algorithms.* As a mathematical model of the perception of textual information, we used neural networks, but the choice of particular groups of neurons was presented as a discrete Markov random process with discrete time. Thus, the process of perception was presented as a succession of different states of the system. The dominant group of formal neurons exists in each state that allows them to differentiate.

*Results.* As a result of the research we have identified new types of perception of textual information, created a model of this process, as well as a special method of forecasting.

*Conclusion.* In the future we plan to apply these models to other material.

# STUDY OF *ARABIDOPSIS THALIANA* ORNITHINE AMINOTRANSFERASE GENE EXPRESSION CONTROL

Gorelova V.V.*, Gerasimova S.V., Kochetov A.V.
*Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia*
*e-mail: gorelova@bionet.nsc.ru*
*Corresponding author**

*Motivation and Aim:* Mitochondria localized enzyme delta-ornithine aminotransferase (OAT) converts ornithine into pyrrolin-5-carboxylate (P5C) with following glutamate production. Due to P5C accumulation as a product of the *trans*-amination reaction, OAT has been assumed to be involved in cytosolic proline accumulation via ornithin pathway that is an alternative to the main glutamate pathway. Since proline is a compatible solute which concentration increases during osmotic stress, OAT has been assumed to take part in regulation of cellular response to osmotic stress. In our laboratory it has been shown earlier that tobacco plants over-expressing OAT exhibited better stress tolerance and avoided growth inhibition. However, this tolerance was not accompanied by proline accumulation. This fact allows us to propose that OAT-related stress-tolerance is mediated through other molecular mechanisms. The objective of the research is to study an expression control of *A. thaliana* OAT gene in order to clarify the role of this gene in plant stress response mechanisms.

*Methods:* To study the expression control of OAT we have generated a set of genetic constructions carrying deletion variants of the *A.thaliana* OAT promoter region, including 5` UTR region. The constructions harbour the reporter gene of beta-glucurinidase (GUS) driven by the truncated promoter variants. *Nicotiana tabacum* and *Arabidopsis thaliana* transgenic plants carrying these constructions were obtained. We have identified basal OAT expression pattern and traced promoter activity during germination and under various conditions. To study the role of OAT in stress response we have screened promoter in order to predict *cis*-elements responsible for transcription control. The screening has been implemented using PLACE database and program ACTIVITY [1] for auxin responsive elements prediction. The most plausible predictions have been checked experimentally.

*Results:* OAT promoter was found to be active in plant meristems (apical bud, lateral buds, neogenic roots, shoot-root junction). The promoter activity was specifically induced in root tips in response to application of auxin and inorganic forms of nitrogen ($NH_4$+, $NO_3$-) into MS medium, that allows us to speculate about its role in nitrogen assimilation. We suppose that OAT is involved in assimilation of inorganic forms of nitrogen that are converted into glutamine prior to be transported to the upper parts of plants. It is very tempting to presume that the OAT promoter activity in response to inorganic nitrogen application is regulated by auxin. There are some premises for this hypothesis: 1) the pattern of OAT expression visualized with the aid of the reporter gene strikingly resembles the auxin distribution pattern in plants, 2) promoter activity is inducible by auxin application, there are AuxRE sites predicted in the promoter region, 3) regulation of some other genes involved in nitrogen assimilation is mediated by auxin both directly and indirectly.

*References:*

1. Ponomarenko et al. (2001) ACTIVITY: a database on DNA/RNA sites activity adapted to apply sequence-activity relationships from one system to another, *Nucleic Acid Research*, **29:1**: 284-287.

# ABOUT CYCLES OF FUNCTIONING OF MODEL OF GENE NETWORK REGULATORY LOOPS

V.A. Katyakova*, A.A. Evdokimov
*Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia*
*e-mail: arevtak@gmail.com*
**Corresponding author*

**Key words:** *gene network, discrete model, regulatory loop, functional graph, working vector, threshold function, cycles and fixed points, dynamics of states.*

*Motivation and Aim:* Gene networks have an important role in the living systems functioning. The aim of this work was to investigate gene networks in terms of the discrete models of the regulatory loops functioning.

*Methods and Algorithms:* The regulatory loop is determined by specifying of a circulant digraph $G_{n,k}(V, D)$ with the vertex set $V=\{v_1, \dots, v_n\}$ and the arc set $D=\{v_i v_{j(mod\ n)} \mid 1\le i\le n,\ 1\le j\le k\}$. Each vertex $v_i$ of graph is determined by full number weight $0 \le x_i \le p-1$ which characterize concentration of products of genes. Changes of these concentrations in time are described by action of additive mapping $A$ with threshold functions in graph vertices.

Define the function $f$ compared to the vertices of $G_{n,k}$ for the arbitrary values of $n$, $k$ and $p$ by the following way $f(x_i, \dots, x_{i+k}) = x_i'$.

The gene network functioning is characterized by changes of the substance concentrations, i.e. changes of n-vectors corresponding to the values of $f$ in $n$ graph vertices at every moment of time, let's call this mapping $A$. Changes in state of whole network is determined by mapping $A$.

*Results:* Our description of actions of mapping $A$ on n-vectors $x^i = (0^{i-1}, p-1, 0^{n-i})$ in dependence from $k$ was more detailed than was published previously. Also we described of actions of mapping A on other vectors of special types. We described some cycles in functioning graph and estimated length of path from vectors of special type to cycles.

*Conclusion:* Obtained results adds new information to descriptions of cycles and reverse states of additive mapping of gene networks.

*References:*

1. A.A. Evdokimov, E.O. Kutumova. (2010) Discrete model of gene networks regulatory loops with threshold function, The International Conference on Bioinformatics of Genome Regulation and Structure\Systems Biology, Novosibirsk, Russia, Abstracts: 155.

# STRUCTURAL PARAMETERS OF TRYPTOPHAN ENVIRONMENT IN *PHOTOBACTERIUM LEIOGNATHI* LUCIFERASE

A.A. Koval*, E.V. Nemtseva
*Siberian Federal University, Krasnoyarsk, Russia*
*e-mail: anya.a.koval@gmail.com*
**Corresponding author*

**Key words:** *protein structure, protein fluorescence, mutations, geometry optimization*

*Motivation and Aim:* The aim of the study is to examine the main characteristics of tryptophan environment in bacterial luciferase to simplify the fluorescence analysis of this protein. The tryptophan fluorescence is one of the most widely used spectral properties for the study protein structure and conformation [1]. To define the contribution of each of seven tryptophan residues to the protein fluorescence a computational simulation was done.

*Methods and Algorithms:* As the crystal structure of the *P. l.* bacterial luciferase has not been solved yet, the structure of *Vibrio harveyi* (*V. h.*) bacterial luciferase (PDB entry: 3FGC), highly homologous enzyme, was used. Residues located at the distance of 7.0A from indole ring atoms of tryptophan residues were determined for *V. h.* luciferase. Differing amino acids for *P. l.* luciferase were mutated according to the alignment. Then an optimization of obtained structure was made using molecular mechanics and semi empirical optimization algorithms: GROMOS force field implementation of Swiss-PdbViewer and PM6 method implemented in MOPAC 2009, respectively. For residues located at the distance of 7.0A from indole ring atoms of tryptophan residues a geometry optimization was done, while the remaining residues kept fixed. The structural characteristics of new tryptophan environment were analyzed.

*Results:* The structural analysis of the environment of seven tryptophan residues was done. We defined the polar protein atoms (O, N and S) located at the distance of 7.0A from indole ring atoms of tryptophan residues. These parameters were used to find out what residues are quenched by the environment interactions and what tryptophan residues participate in the enzyme fluorescence. Packing density in the sphere for every residue was also determined.

*Conclusion:* Assignment of tryptophan residues of *P. l.* bacterial luciferase to the five spectral-structural classes [1] is made relying on computational study of highly homologous *V. h.* bacterial luciferase.

*Acknowledgements:* This work was supported by the Federal agency of science and innovations (contract No 02.740.11.0766), the Program of the Government of Russian Federation "Measures to attract leading scientists to Russian educational institutions" (grant No 11.G34.31.058)

*References:*

1. 1. E.A.Burstein. (1983) The intrinsic luminescence of proteins is a method for studies of the fast structural dynamics, *Molecular Biology,* **17:** 455-467.

# HIGH-THROUGHPUT SEQUENCING OF microRNA OF ROOTS AND NODULES OF PISUM SATIVUM L.

V.Yu. Shtratnikova[1]*, Y.A. Pekov[1], V. Zhukov[2]
*1CIT BAC RAS, Moscow, Russia*
*2Laboratory of Genetics of Plant-Microbe Interactions, All-Russia Research Institute for Agricultural Microbiology, St-Petersburg-Pushkin, Russia.*
*e-mail: vtosha@yandex.ru*
**Corresponding author*

**Key words:** *next-generation sequencing, microRNA*

Root tips and nodules of line SGE of Pisum sativum L. were used for RNA isolation. Small RNAs were isolated with Qiagen miRNeasy kit. MicroRNA library (insertion fragment 18-32 b.) was prepared with TruSeq Small RNA sample preparation kit (Illumina, USA) by instruction from manufacturer. RNA-adapters were ligated to 3' and 5' ends. It is necessary for 1) reverse transcription, 2) fragment amplification, 3) attaching to flowcell for sequencing. The libraries were amplified on flowcell for cluster formation to intensify fluorescent signal. Sequencing by synthesis was carry out on GAIIx. Read length was 36 bases.

The reads were converted to fastq by CASAVA software. Initial number of the reads was 37 202 109 for the root tips and 15 858 836 for the nodules. Adapters was removed from the reads with fastq-mcf sortware. After that each read of length less than 18 b. was also removed. As the result 32 904 411 reads remained for the roots, 11 445 068 reads for the nodules. We apply a quality filter (fastq_quality_filter software from GALAXY toolkit, more than 80% of bases of read should have quality more than 30) and got 31 566 793 reads for the root tips and 10 949 097 for the nodules of high quality reads. The reads were mapped to pea rRNAs and tRNAs from EMBL and the mapped reads was discarded. Number of the rest reads: 24 529 726 for root tips, 10 159 253 for nodules. We discarded all the reads which coincided more than 95% with cd-hit-est software to get a list of unique reads. We get 2 641 248 unique reads for the root tips and 2 915 224 for the nodules.

There is no pea whole genome sequenced by the date, so for next treatment we use next databases: known sequence of Pisum genes and Pisum transcriptome (EST) from GenBank; mature microRNAs and hairpins as pre-mRNAs of Medicago truncatula (nearest species with sequenced genome) - miRBase.

| Base | Number of the mapped reads | |
|---|---|---|
| | Root tips | Nodules |
| Pisum genes | 378 054 | 250 536 |
| Pisum transcriptome | 118 612 | 84 055 |
| Medicago hairpins | 215 | 655 |
| Medicago mature RNA | 100 | 271 |
| Reads mapped to Medicago mature RNA coincided with reads mapped to Medicago hairpins | 9 | 101 |

The pea nodules contain significantly more microRNAs coincided with microRNAs of Medicago truncatula than the root tips. Searching for new microRNAs of pea is carried at the present time.

The other abstracts of Young scientists school are published in the Abstracts of The Eighth International Conference on Bioinformatics of Genome Regulation and Structure\Systems Biology (BGRS\SB'12).

**Below is citation of these abstracts:**

## PROGRAM COMMITTEE

**Nikolay Kolchanov** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia (Chairman)

**Dmitry Afonnikov** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

**Roman Efremov** Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Moscow, Russia

**Robert Giegerich** Bielefeld University, Germany

**Olga Krebs** Heidelberg Institute for Theoretical Studies, Germany

**Eric Mjolsness** University of California, Irvine, USA

**Yuriy Orlov** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

**Evgeny Rogaev** Head of laboratory of brain molecular genetics in Mental Helth Research Center RAMS, Moscow; Head of laboratory in University of Massachusetts Medical School

**Maksim Zakhartsev** Institute of Biochemical Engineering, University of Stuttgart; Institute of Pharmacy and Molecular Biotechnology, University of Heidelberg, Germany

## LOCAL ORGANIZING COMMITTEE

**Victoria Mironova** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia (Chairperson)

**Svetlana Zubova** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

**Ilya Akberdin** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

**Fedor Kazantsev** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

**Evgeny Tiys** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

**Nikolai Alemasov** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

**Erlan Tokpanov** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

**Nadezhda Glebova** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

**Tatyana Karamysheva** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

**Andrey Kharkevich** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

**Galina Kiseleva** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

**Anna Onchukova** Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

# Organizers


Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences

Siberian Branch of the Russian Academy of Sciences


The Vavilov Society of Geneticists and Breeders


Department of Systems Biology


Chair of Information Biology

Shared Facility Center «Bioinformatics» SB RAS


PBSoft Ltd.

# Sponsors

Russian Foundation for Basic Research

Skolkovo Innovation Centre

## GOLD SPONSOR

Life Technologies

## SILVER SPONSORS

Bruker Corporation

Agency “Khimexpert”

Hewlett Packard

“OPTEC”, LLC

OOO «Miass Factory of Medical Equipment» & ZAO «Aceptic Medical Systems»

## BRONZE SPONSORS

Intel

Bio-Rad

Roshe diagnostics

Тираж 120. Заказ 51.

Отпечатано на полиграфической базе ИЦиГ СО РАН
630090, Новосибирск, пр. акад. Лаврентьева, 10