RUSSIAN ACADEMY OF SCIENCES, SIBERIAN BRANCH
INSTITUTE OF CYTOLOGY AND GENETICS

THE 6th INTERNATIONAL YOUNG SCIENTISTS SCHOOL "BIOINFORMATICS AND SYSTEMS BIOLOGY"

Abstracts

BGRS\SB-2014
Novosibirsk, Russia
June 24–27, 2014

INTERNATIONAL PROGRAM COMMITTEE

Prof. Nikolay A. Kolchanov, Academician RAS, Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia (Chairman)
Prof. Eugeny I. Rogaev, N.I. Vavilov Institute of General Genetics, Moscow; University of Massachusetts Medical School, USA;
Prof. Eugeny V. Berezikov, European Research Institute for the Biology of Ageing, University Medical Center Groningen, Groningen, The Netherlands
Prof. Maria G. Samsonova, St.Petersburg State Polytechnic University, St.Petersburg, Russia
Prof. Alexander M. Samsonov, A.F. Ioffe Physico-Technical Institute RAS, St.Petersburg State Polytechnic University, St.Petersburg, Russia
Prof. Oleg L. Serov, Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
Prof. Nikolay N. Dygalo, corresponding member of the RAS, Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
Prof. Suren M. Zakian, Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
Prof. Limsoon Wong, National University of Singapore, Singapore
Prof. Ming Chen, Zhejiang University, Hangzhou, China

LOCAL ORGANIZING COMMITTEE

Ilya Akberdin, PhD
Dmitry Afonnikov, PhD
Nadejda Glebova
Tatyana Karamysheva, PhD
Galina Kiseleva, PhD
Andrey Kharkevich
Anna Onchukova
Olga Popik
Alex Semenichev
Yuriy Orlov, PhD
Erlan Tokpanov
Svetlana Zubova

CONTACTS

Organizing committee: bgrs2014@bionet.nsc.ru
Siberian Branch of the Russian Academy of Sciences
RUSSIAN FEDERATION
630090 Novosibirsk, Lavrentyeva, 10
Tel: +7(383) 363-49-80
Fax: +7(383) 333-12-78
URL - ICG SB RAS: www.bionet.nsc.ru
URL - BGRS\SB-2014: http://conf.nsc.ru/BGRSSB2014

ISBN 978-5-91291-019-7

Organizers

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

Chair of Information Biology, Novosibirsk State University (NSF NSU)

Siberian Branch of the Russian Academy of Sciences

Company «Scientific service» Ltd.

The Vavilov Society of Geneticists and Breeders

The program "Genomics, proteomics and bioinformatics" of the SB RAS

Center of neurobiology and neurogenetics, IC&G SB RAS

Sponsors

Russian Foundation for Basic Research

SILVER SPONSORS

Helicon

Limited Liability Company Eppendorf Russia

Genotek

Bio-Rad

Intel

InterLabService

BRONZE SPONSORS

Roshe diagnostics

CONTENTS

ROTATIONAL DYNAMICS OF BASES IN THE GENE CODING INTERFERON ALPHA 17 (IFNA17)

Krasnobaeva L.A.[1, 2]*, Yakushevich L.V.[3]
[1]*Siberian State Medical University, Tomsk, Russia;*
[2]*Tomsk State University, Tomsk, Russia;*
[3]*Institute of Cell Biophysics RAS, Pushchino, Moscow region, Russia*
e-mail: kla1983@mail.ru
**Corresponding author*

Key words: interferon alpha 17 (IFNA17), rotational oscillations of bases, DNA

Motivation and Aim: The group of proteins – interferons – plays an important role in medicine. Due to them cells become immune to viruses. Therefore the studying of the genes coding these proteins is an interesting and actual task. In the present work, rotational oscillations of nitrogenous bases in the sequence of the gene coding interferon alpha 17 (IFNA17), are investigated.

Methods and Algorithms: As a mathematical model simulating oscillations of the bases, we use a system of two coupled nonlinear partial differential equations that take into account the effects of dissipation, the action of external fields and the dependence of the coefficients on the sequence of bases. We apply the methods of the theory of oscillations to to solve the equations in linear approach and to construct dispersion curves. To consider the system of the equations in the general (nonlinear) case, the approximation of the average field, which allows to reduce the problem of two coupled equations to the problem of a single equation simulating oscillations of the bases of one of two polynucleotide chains in the average field, induced by the second polynucleotide chain is was used. This equation has been solved by two methods: the method of the concentration [1], and the energy method [2].

Results: In the linear approach, the solutions of the model system of differential equations have been obtained, and the dispersive curves determining the dependence of the frequency of the plane waves (ω) on the wave vector (q) have been constructed. In the nonlinear case, the solution in the form of was obtained, and the main characteristics of the kink: the energy density (ρ), the total energy (E), the rest energy (E_0), the rest mass (m_0) and the size (d), were calculated. With the help of energy method the kink velocity (υ), the path (S) that kink passed, and the lifetime (τ) have been obtained.

References:

1. L.V. Yakushevich, L.V. Krasnobaeva. (2008) A new approach to studies of non-linear dynamics of kinks activated in inhomogeneous polynucleotide chains, *International Journal of Nonlinear Mechanics*, **43:** №10, 1074–1081.
2. D.W. McLaughlin, A.C. Scott. (1978) Perturbation analysis of fluxon dynamics, *Phys. Rev. A*, **18:** 1652-1660.

miRNAs BINDING TO mRNAs OF RICE MYB GENES

Pinsky I.V.*, Ivashchenko A.T.
Al-Farabi Kazakh National University, Almaty, Kazakhstan
e-mail: ilya.pinskyi@mail.ru
**Corresponding author*

Key words: plant, MYB transcription factors, miRNAs, mRNAs, CDS

Motivation and Aim: Transcription factors (TF) play a key role in gene expression regulation. 2025 transcription factors of *Oryza sativa L. ssp. indica* are included in 63 families. The family of MYB TF genes plays a key role in the plant growth, development and stress reactions. Expression of transcription factor genes is controlled by miRNAs binding with mRNAs of genes and blocking translation. So it was important to study which miRNAs bind to mRNAs of *O. sativa* MYB genes, which genes are targets for each miRNA family and calculate free energy of miRNA-mRNA interaction.

Methods and Algorithms: The binding sites of osa-miRNA in mRNA of 136 gene of MYB family was revealed using the program MirTarget, which defines the location of the miRNA with mRNA binding sites, calculates the free energy hybridization (ΔG, kJ/mole) and the ratio $\Delta G/\Delta G_m$ (%), where ΔG_m equal to the free energy of miRNA binding with fully complementary nucleotide sequence. The binding sites of miRNA with mRNA selected by $\Delta G/\Delta G_m$ equal or above to 85%.

Results: At present 661 miRNAs and 136 MYB TF genes are known in the genome of rice. It was found that only 173 miRNAs bind to mRNAs of *O. sativa* MYB genes with value of score ranging from 85% to 96% and only 86 genes are controlled by miRNAs. Osa-miR2102-5p has the largest number of target genes. It bound to mRNAs of 23 genes with value of score from 86 to 91%. Osa-miR5075 bound to mRNAs of 11 genes with value of score equal 86-93%. In the genome of rice some miRNAs bound to mRNAs of one or several genes. Six miRNAs of osa-miR159a-f family bound to mRNAs of OsIBCD015177, OsIBCD018450, OsIBCD021789 genes with value of $\Delta G/\Delta G_m$ equal 87-96%. All ten miRNAs of osa-miR439a-j family bound to mRNAs of only one gene OsIBCD007874 with the same value of $\Delta G/\Delta G_m$ equal 85%. Six miRNAs of osa-miR164a-f family bound to mRNAs of OsIBCD014655, OsIBCD001147, OsIBCD000203 and OsIBCD028679 genes with values of $\Delta G/\Delta G_m$ equal 87-96%. Osa-miR169a-e bound to mRNAs of five genes (OsIBCD015070, OsIBCD022880, OsIBCD003079, OsIBCD032138 и OsIBCD008009) with score equal 85-87%. Osa-miR159 bound to mRNAs of three genes (OsIBCD015177, OsIBCD018450, OsIBCD021789) with value of score ranging 87-96%. Osa-miR1846d-3p bound to mRNAs of OsIBCD018922, OsIBCD023878, OsIBCD026910 and OsIBCD035359 genes. All rice miRNA binding sites are located in the protein-coding part of MYB TF family mRNAs.

Conclusion: As a result of study of *O. sativa* MYB TF gene family there was found that: 1) some target genes have more than one miRNA binding site, beginnings of which are located successively through one-three nucleotides that increases probability of translational process control by miRNA binding; 2) not all MYB family genes are targets for miRNA; 3) nucleotides of miRNA binding sites encode oligopeptides of MYB proteins in different reading frames. Results obtained show that osa-miRNAs can regulate expression the most of MYB genes and influence the plant productivity and resistance.

SHARED BIOINFORMATICS DATABASE WITHIN UniPro UGENE

Protsyuk I.V.[1, 2]*
[1]*Novosibirsk State University, Novosibirsk, Russia;*
[2]*Center of Information Technologies "UniPro", Novosibirsk, Russia*
e-mail: iprotsyuk@unipro.ru
**Corresponding author*

Key words: Bioinformatics, Custom Shared Databases, Open Source, Cross Platform, Data Management

Motivation and Aim: Unipro UGENE [1] is an open-source bioinformatics toolkit that integrates popular tools along with original instruments for molecular biologists within a unified user interface. Nowadays most bioinformatics desktop applications, including UGENE, make use only of local user files when processing different types of data. Such an approach causes inconvenience to scientists working cooperatively and relying on the same data. The most obvious issues are the need to make multiple copies of certain resources for every workplace and do synchronization in further. There are tools that provide the needed capabilities but they are proprietary and quite expensive.

Methods and Algorithms: It was decided to implement the client-server architecture where the server is represented by a remote database server and the client is an instance of the UGENE tool run on a local host. Thereby certain data models were developed for all the types of biological data supported by the UGENE platform. One of the most important goals was to minimize the count of queries to the server. Major efforts were made to avoid downloading of the entire data, requested by a user as well. Loading on demand was realized instead.

Results: Initial tests showed that the system is capable of storing hundreds of thousands of objects (annotated sequences, multiple alignments, etc.) with shared read/write access to them. Large objects are supported as well. For example, the assembly of Denisovan genome (~10 GB in BAM format) was successfully imported to the database and clients could navigate through its short reads.

Conclusion: Thus, UGENE may provide a unified access to shared data for users located, for example, in the same lab or an institution. This work itself presents a basis for further development of a distributed computational system. Namely, workflow manager, a part of the UGENE platform, coupled with a shared database may operate remotely and do jobs for clients aiming to process the shared data. Therefore, this opens an opportunity for using easy-to-install computational clusters requiring the UGENE suite only.

Availability: UGENE binaries are freely available for MS Windows, Linux and Mac OS X at {{http://ugene.unipro.ru/download.html}}. UGENE code is licensed under the GPLv2.

References:

1. Konstantin Okonechnikov et al. (2012), Unipro UGENE: a unified bioinformatics toolkit, Bioinformatics, 28: 1166-1167.

ON A PARALLEL ALGORITHM FOR MORPHOGENE DIFFUSION-REACTION PROCESSES SIMULATION ON A 2D CELL ENSEMBLE

Troeglazova T.S.[1]*, Baishibaev D.Ja.[1], Penenko A.V.[2], Nikolaev S.V.[3], Zubairova U.S.[3]
[1]*Novosibirsk State University, Novosibirsk, Russia;*
[2]*Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia;*
[3]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
e-mail: tatyanameda@gmail.com
**Corresponding author*

Key words: morphogenesis, diffusion-reaction model, parallel computations, discrete-analytical scheme

Motivation and Aim: Diffusion-reaction PDE systems can be used to model morphogene transport and dynamics over a plant tissue. The aim of the work is to create parallel implementation of a diffusion-reaction solver capable for high performance simulations. High performance is needed for inverse problems solution algorithms that involve multiple direct and adjoint problem evaluations.

Methods and Algorithms: A tissue is modeled with 2D cell ensemble. A system is split with respect to different processes i.e. to the reaction and diffusion parts. To do model decomposition the additive-averaged splitting scheme is used. Reaction part is implemented with explicit discrete-analytical scheme [1] which guarantees positive morphogene concentrations.

Results: The algorithm has been realized for model of structuring the stem cell niche in shoot apical meristem of Arabidopsis Thaliana [2]. A discrete-analytical scheme performance has been tested in the application to morphogenesis.

Conclusion: Parallel structure of the algorithm allows high performance simulations of morphogenesis processes.

Availability: The program is available on request from the authors.

References:

1. V.V. Penenko et al. (2014) Variational Approach and Euler's Integrating Factors for Environmental Studies, accepted in CAMWA, 2014.
2. S.V. Nikolaev et al. (2013) Model of Structuring the Stem Cell Niche in Shoot Apical Meristem of Arabidopsis thaliana, Doklady Biological Sciences, 452:336–338.

DYNAMIC INSTABILITIES OF MICROTUBULES

Barlukova A.[1]*, Honoré S.[2], Hubert F.[1], Petit M.[2]
[1]Aix-Marseille University, I2M CNRS UMR 7353, Marseille, France;
[2]Aix-Marseille University, Inserm, CRO2 UMR 911, Marseille, France
e-mail: ayuna.barlukova@etu.univ-amu.fr
**Corresponding author*

Key words: cancer, microtubules, dynamic instability, MT targeting drugs

Motivation and Aim: Microtubules (MTs) are long tube polymers of tubulin, found throughout the cytoplasm. MTs are very important in all crucial cellular processes in cancer progression such as cell division and migration. The aim of the work is to develop new mathematical models that account for the effects of MT targeting agents on MT instabilities.

Methods and Algorithms: We propose a new deterministic mathematical model inspired by the work of P. Hinow et al. [1] to simulate the behavior of a MT population. The model couples transport equations with ordinary differential equations (ODE) with nonlocal terms endowed with suitable boundary conditions for both catastrophe and rescue. The mathematical model is built from biological observations obtained by the pharmacologist of our interdisciplinary research group [2]. Numerical results are obtained in MATLAB by using upwind scheme with adaptive time step for the partial differential equations and the explicit Euler method for the ODE.

Results: We obtain graphs for time evolution of the average total length of MTs in polymerization state and their caps and average length of MTs in depolymerization state (similar to data obtained by kymograph); concentrations of free GTP and GDP tubulin, total quantities of tubulin incorporated in MTs in polymerization and depolymerization states; time evolution of distribution of MTs in polymerization state.

Conclusion: Computational simulations describe diverse concepts of behavior of MT populations with and without impact of drugs. New model allows us to analyze the pharmacological action of different anti-microtubule drugs, including that influence on MT "aging", on MT instabilities. Numerical results are in a good agreement with biological observations.

Acknowledgements: The program is funded thanks to the support of the A*MIDEX project (n° ANR-11-IDEX-0001-02) funded by the «Investissements d'Avenir» French Government program, managed by the French National Research Agency (ANR).

References:

1. P. Hinow, V. Rezania, J.A. Tuzszynski. Continuous model for microtubule dynamics with catastrophe, rescue, and nucleation processes. Phys Rev E Stat Nonlin Soft Matter Phys. 2009 Sep; 80(3 Pt 1): 031904.
2. R. Mohana, E.A. Katrukhaa, H. Doodhia, I. Smalb, E. Meijeringb, L.C. Kapiteina, M.O. Steinmetzc, A. Akhmanova. End-binding proteins sensitize microtubules to the action of microtubule-targeting agents. Proc. Natl. Acad. Sci. U.S.A. V. 110, 1091-6490, 2013.

ANALYSIS OF GENETIC SUBDIVISION OF TICK-BORNE ENCEPHALITIS VIRUS WITHIN EURASIAN AREAL

Borisenko A.Yu.[1, 2], Bukin Yu.S.[3], Dzhioev Yu.P.[1, 2], Zlobin V.I.[1]
[1]Research Institute of Biomedical Technology, Irkutsk State Medical University, Irkutsk, Russia;
[2]FSBI "Scientific Center of family health problems and human reproduction" SB RAMS, Irkutsk, Russia;
[3]Limnological Institute, Irkutsk, Russia

Tick-borne encephalitis virus proposes a great danger to human health. In the Russian Federation there are hundreds of people suffer from this disease. Like all other viruses, tick-borne encephalitis virus continually rapid transformation of the genome. This leads to the emergence of new strains of virus of the known methods of protection. It is known that all the transformations in the genomes appear and are fixed at the population level. The research of tick-borne encephalitis virus at the population level will allow as predicting the dynamics of accumulation of variables in its genome. This will prearrange prevention and treatment of viral infections.

Investigation of the population structure of the virus was carried out on sones transcribed sequences of the viral envelope protein E. This protein is involved in virus penetration into the cell and acts when antigen after immune response of human. Currently Genbank database accumulated about 375 decrypted sequence of this protein. This number sequence is sufficient for population genetic analysis.

For population analysis of the all data array has been broken by belonging of sequences to certain regions on Eurasian habitat virus. Total allocated three regions: Far East and East Asia; The European part of Russian, Siberia and Mongolia; Eastern and Western Europe. In each of the selected population was estimated genetic diversity using different variations of the parameters Theta. Were determined flows gens between populations with the help of Fst criterion. For each population was built distribution of pairwise genetic distances and appreciate the history of population size changes in the past. To determine the population parameters we used software packages: Arlequin311 [1]; DNASP [2]; BEAST [3].

Another important aspect of the study was to compare the polymorphism within the virus population with polymorphism insider genotype of virus. Total allocated 4 genotype of tick-borne encephalitis virus, characterized by different features of the gene and clinical manifestations of the disease. As a result, within the sequences belonging to the same genotype their different geographical regions was identified less than genetic differentiation between regions.

References:

1. Excoffier, L.G. Laval, S. Schneider (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50.
2. P. Librado, J. Rozas, (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25:** 1451-1452.
3. A.J. Drummond, M.A. Suchard, Xie Dong, A. Rambaut. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29, 1969-1973.

COMPUTER SIMULATION OF SELF-ORGANIZATION IN THE BACTERIAL MinCDE SYSTEM

Vitvitsky A.A.
Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia
e-mail: vitvit@ssd.sscc.ru

Key words: computer simulation, MinCDE system, cellular automata, self-organization

Motivation and Aim: Bacterial cell division begins with formation of a ring-like structure, called the Z-ring, on the cell membrane. A proper position of Z-ring in the midcell is controlled by the certain self-organization mechanisms. A bright example of these mechanisms is a MinCDE protein complex. Currently, the processes leading to a self-organization in this MinCDE system are not quite clear, but are intensively studied [1, 2]. However, exact theoretical description of these processes is a hard task because of the difficulty to obtain information about spatiotemporal dynamics of individual particles from bulk biochemical assays. Hence, computer simulation plays an important role in this problem and helps to conform (or to disprove) the proposed hypotheses.

Methods and Algorithms: Cellular Automata (CA) was chosen as a simulation tool. CA is a discrete mathematical model consisting of a set of finite state automata called cells [3]. CA models have the following advantages: ability to simulate complex nonlinear processes in active environments, simple rules of simulating and the natural fine-grained parallelism.

Results: CA-model of self-organization MinCDE protein complex based on theoretical description from [2] has been developed. The dependence of protein concentration in time and space, obtained as a result of computer simulation, revealed similarity with that, obtained by the experiments *in vitro.* In addition, the visualization of computational experiments showed propagating protein waves similar to those that emerge *in vitro*.

Conclusion: The evidence in favor of the hypothesis from [2], which claims that self-organization in the MinCDE system arises from an interplay of two opposing mechanisms: cooperative binding of MinD to the membrane, and accelerated MinD detachment due to persistent MinE rebinding have been obtained.

References:

1. M. Loose, E. Fischer-Friedrich, C. Herold, K. Kruse, P. Schwille (2011). Min protein patterns emerge from rapid rebinding and membrane interaction of MinE. Nat. Struct. Mol. Biol. 18, 577583. doi:10.1038/nsmb.2037.
2. V. Ivanov, K. Mizuuchi (2010). Multiple modes of interconverting dynamic pattern formation by bacterial cell division proteins. Proc Natl Acad Sci USA 107: 80718078.4
3. O. Bandman Implementation of Large-Scale Cellular Automata Models on Multi-Core Computers and Clusters // High Performance Computing and Simulation (HPCS), 2013 Int. Conf. IEEE Conference Publications, 2013. pp.304–310.

1D MODELLING OF DIFFERENT TIME REGIMES OF ENHANCED EXTERNAL COUNTERPULSATION

Gamilov T.M.*, Simakov S.S.
Moscow Institute of Physics and Technology, Dolgoprudny, Russia
e-mail: gamilov@crec.mipt.ru
**Corresponding author*

Key words: cardiovascular system, enhanced external counterpulsation, EECP, hemodynamic, vessels network, cardiovascular modelling, elastic tube

Motivation and Aim: Enhanced external counterpulsation (EECP)is a non-invasive procedure that involves surrounding patient's legs with inflatable cuffs that are pressurized and depressurized out-of-phase with systole. It is used for the treatment of patients with different heart diseases [1,2]. Despite some success [2], the mechanisms by which EECP improves cardiac function remain unclear. This leads to problems in defining some EECP parametrs for each particular case. The aim of this work is to perform numerical simulations of different EECP regimes and estimate their impact on hemodynamics in coronary region.

Methods and Algorithms: The model viscous incompressible fluid flow through the network of elastic tubes described in [3] was used. This model was validated by a number of methods, including shock wave formation study [4], and enhanced with autoregulation and venous valves models. Heart cycle was considered to be constant and equals 1 s.

Results: Different time regimes of EECP were simulated. Regimes are denoted by the time inside a heart cycle when the pressurization starts: 0.25 s, 0.3 s, 0.35 s, 0.4 s. The increase in average blood pressure in a terminal coronary artery for each regime is 29.2 ± 0.2 %. The emptying effectiveness of leg vessels is different for each regime: 13 %, 16 %, 19 % and 22 % respectively.

Conclusion: Results show that the impact of different EECP regimes on blood circulation in coronary vessels is approximately the same as long as pressurization is out-of-phase with systole. Regimes with a "late start" provided better emptying effectiveness thus , providing higher venous return. We conclude that later start is more effective in terms of assisting in heart work.

Acknowledgements: This work was supported by grant RSCF 14-21-00107.

References:

1. E.T. Ozawa, K.E. Bottom, X. Xiao, R.D. Kamm. (2001) Numerical simulation of enhanced external counterpulsation, *Ann Biomed Eng* **29:** 284-297. W.E. Lawson et al. (1992) Efficacy of enhanced external counterpulsation in the treatment of angina pectoris, *Am J Cardiol* **70**: 859-862.
2. S.S. Simakov, A.S. Kholodov. (2009) Computational study of oxygen concentration in human blood under low frequency disturbances, *Math Mod Comp Sim* **1(2)**: 283-295.
3. S. Canic, E.H. Kim. (2003) Mathematical analysis of the quasilinear effects in a hyperbolic model blood flow through compliant axi-symmetric vessels, *Math Meth Appl Sci* **26**: 1161-1186.

ImageJ ADDON FOR 2D ELECTROPHORESIS GEL ANALYSIS

Gurkov A.N.*, Kondratyeva E.M., Bedulina D.S.
Institute of Biology at Irkutsk State University, Irkutsk, Russia
e-mail: a.n.gurkov@gmail.com
**Corresponding author*

Key words: ImageJ, 2D electrophoresis, 2D gel analysis software

Motivation and Aim: Two-dimensional gel electrophoresis of proteins is widely used in modern proteomics. There is a number of commercial software for analysis of obtained 2D gels, however there is no free software, which provides similar functionality for full and automated analysis. Existing free software like RegStatGel [1] and Pinnacle [2] perform only certain steps of analysis, but not all of them. Whereas ImageJ is a perfect base for this kind of image analysis: it is free, open source and cross-platform, it is very popular and provides powerful extensibility. A workflow for 2D gel analysis with ImageJ have been published [3], but the algorithm contains some mistakes and no macro or plugin have been published to automate the workflow. The aim of this work was to develop an addon for ImageJ, which allows to fulfil full analysis of 2D electrophoresis gels as automatically as possible.

Methods and Algorithms: The addon represents the media that brings together under ImageJ or Fiji several ImageJ plugins, new implementation of developed earlier algorithm and new algorithms to perform following steps of 2D gel analysis: image alignment (plugins «Extract SIFT Correspondences», «Extract MOPS Correspondences» and «bUnwarpJ»), protein spots identification (implementation of foreign algorithm [4], original algorithm or plugin «Watershed Algorithm») and spots quantification (original algorithm).

Results and Conclusion: The worked out addon can substitute existing commercial software and provide the improved and highly automated workflow for 2D-PAGE analysis for free.

Availability: Debugging of the addon is finishing now. The addon is available on request from the authors, but soon it will be available through the web-pages of ImageJ and Fiji.

References:

1. Li F., Seillier-Moiseiwitsch F. (2010) Differential analysis of 2D gel images, Methods in enzymology, 487: 595-609.
2. Morris J.S. et al. (2009) Evaluating the performance of new approaches to spot quantification and differential expression in 2-dimensional gel electrophoresis studies, Journal of proteome research, 9, 1: 595-604.
3. Natale M. et al. (2011) Image Analysis Workflow for 2-D Electrophoresis Gels Based on ImageJ, Proteomics Insights, 4: 37-49.
4. Kostopoulou E. et al. (2014) An Effective Approach for Detection and Segmentation of Protein Spots on 2-D Gel Images, Biomedical and Health Informatics, 18, 1: 67-76.

EVALUATION OF VIRE2-COMPLEXES BY MOLECULAR DYNAMIC METHODS

Gusev Yu.S., Chumakov M.I.*
Institute of Biochemistry and Physiology of Plants and Microorganisms RAS, Saratov, Russia
e-mail: chumakov@ibppm.sgu.ru
**Corresponding author*

Key words: Agrobacterium, VirE2-dependent pores; ssDNA transfer

Motivation and Aim: VirE2 virulence protein of the soil bacteria *Agrobacterium tumefaciens* provides the Ti-plasmid (T-DNA) fragment transfer into the genomes of a wide range of plant, and some animal species as a ssDNA-protein complex [1]. The aim of this work was to study the structure and function of the VirE2 protein complexes for VirE2-mediated ssDNA transfer.

Methods and Algorithms: For the VirE2 complex formation the GRAMM-X program was used. Mobility of VirE2 protein complexes was estimated by the normal modes method using the ElNemo program. The modeling of VirE2 complexes by the molecular dynamic methods was carried out at the MDWeb server.

Results: The two-, four-, and six- VirE2 protein subunits complexes were reconstructed from the VirE2 protein X-ray diffraction model, using GRAMM-X program [2]. The stable state for the VirE2-VirE1 complex with the modeling time up to 500 ps by the molecular dynamics method was demonstrated. The oscillating motions of the VirE2-VirE1 proteins, and two-, four- VirE2 proteins subunits complexes were for the first time estimated by the normal modes method. In the VirE2-VirE1 protein model global movements of domains are found, and the possible gating mechanism into the two-VirE2-proteins-complex is discussed.

Conclusion: In the computation model of VirE2 complexes formed by two and four individual proteins the pore formation with channel diameters of 1.2–1.6 and 1.4–4.6 nm, respectively, was predicted. It is shown by molecular dynamic methods that the model of VirE2-VirE1 proteins reaches an equilibrium, stable state at the modeling time up to 500 ps. By the normal mode method is shown the possible gating mechanism in the complex consisting from two VirE2 protein subunits.

Availability: GRAMM-X {{http://vakser.bioinformatics.ku.edu}}, ElNemo {{http://www.igs.cnrs-mrs.fr/elnemo}}, MDWeb {{http://mmb.irbbarcelona.org/ MDWeb/index.php}}.

Acknowledgements: This work was supported in part by the RFBR grant (14-04-31206).

References:

1. M.I. Chumakov (2013) Protein apparatus for horizontal transfer of Agrobacterial T-DNA to eukaryotic cells (review), Biochemistry (Moscow), 78: 1321-1332.
2. I.V. Volokhina, Yu.S. Gusev, S.I. Mazilov, M.I. Chumakov. (2012) VirE2-dependent pores for ssDNA transfer across artificial and cell membranes, *J. Bioinform. & Comput. Biol.,* 10: 1241009.

COMPUTER DATA ANALYSIS OF GENOME SEQUENCING BY TECHNOLOGY ChIP-seq AND Hi-C

Kulakova E.V.[1], Podkolodnaya O.A.[2], Serov O.L.[2], Orlov Y.L.[1, 2]*
[1]*Novosibirsk State University, Novosibirsk, Russia;*
[2]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
e-mail: orlov@bionet.nsc.ru
**Corresponding author*

Key words: ChIP-seq, ChIA-PET, Hi-C, chromosome contacts

Motivation and Aim: Currently there is an active development of high-tech DNA sequencing, including using Chromatin Immuno-Precipitation (ChIP-seq, ChIA-PET), which leads both to the rapid accumulation of a large number of genome-wide information and qualitatively new tasks of bioinformatics, requiring the development of new computer tools. The aim of this work was to develop a statistical data of the computer program of genes and chromatin domains, analysis of experimental data ChIP-seq, Hi-C in the mouse genome (stem cells) and in the human genome [1], [2].

Methods and Algorithms: We use data on the location spatial domains in the human genome and mouse genome[3] and unpublished data Hi-C that is received at the Institute of Cytology and Genetics SB RAS on the methodology presented in the Battulin's work [4]. As standard genomic annotations we used the UCSC Genome Browser - integrated environment for displaying any requested part of the genome at any size, along with dozens of standardized annotations (http://genome.ucsc.edu).

Results: Identified genes that lies on the spatial boundaries of chromosomal domains. Gene Ontology categories were identified with the help of Internet resources of DAVID (http://david.abcc.ncifcrf.gov/). Coexpression communications were analyzed using internet resource String-db (http://string-db.org/).

Conclusion: Genes on domain boundaries represent approximately 2.3% of the total number of genes. During the analysis of gene ontology categories was observed that the most of the genes that lie on the domain boundaries are responsible for allocating the phosphoprotein. The most important gene ontology categories based on the significance factor that is connected with the plasmatic membrane. Which corresponded to 1/4 of these genes.

References:

1. G. Li et al., (2012), Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation, *Cell* V.148. – N.1-2. – P.84-98.
2. M.J. Fullwood et al., (2009) An oestrogen-receptor-alpha-bound human chromatin interactome, *Nature*. V. 462. – N. 7269. – P. 58-64.
3. J.R. Dixon et al. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions, *Nature,* V. 485. – N. 7398. – P. 376-380.
4. N.R. Battulin, et al., (2012) 3C-based methods for 3D genome organization analysis*, Vavilov journal of genetics and breeding*, V. 16. – № 4/2. – C. 872-878.

RECONSTRUCTION OF THE MOUSE BRAIN VASCULAR NET ACCORDING TO THE DATA OF HIGH-FIELD MRI-SCANNER

Maltseva S.V.[1, 2]*, Akulov A.E.[3], Derevtsov E.Yu.[1, 2], Cherevko A.A.[4, 2], Chupakhin A.P.[4, 2], Khe A.K.[4, 2]

[1]*Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia;*
[2]*Novosibirsk State University, Novosibirsk, Russia;*
[3]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[4]*Lavrentyev Institute of Hydrodynamics SB RAS, Novosibirsk, Russia*
e-mail: sv_maltseva@mail.ru
**Corresponding author*

Key words: MRI, brain vasculature modelling, approximation, noise reduce, segmentation

Motivation and Aim: the problem is to reconstruct mouse brain vasculature [1] by given tomographic data; the recovery of the vascular net should be one-connected. This problem is very important not only for fundamental sciences but also for application once (medicine, biology).

Methods and Algorithms: the reconstruction is realized by automatic segmentation according to the mathematically preprocessed data. Initial data are some tomographic packets with the different spatial slice orientation. Preprocessed data are obtained from data packets using some additional technics allowing to correct data intensity, to make vessels walls more contrast and to decrease noise.

Results: the algorithm proposed allows to reconstruct the vascular net more efficient as compared with standard tomographic reconstruction.

Conclusion: detailed and one-connected geometrical model of the vascular net allows to realize haemodynamic CFD-modelling [2] and to discover correlations between structural features of the net and physiological characteristics [3] of the animal units.

Availability: not available.

Acknowledgements: RFBR 14-01-00036, RFBR 14-04-31829-mol_a, Interdisciplinary integration project SB RAS 60 and 61, SB RAS IP № 44, Department of Mechanics RAS № 2.13.4.

References:

1. S.W. Oh et al. (2014) A mesoscale connectome of the mouse brain, Nature, 508, 207-214.
2. C.G. Caro et al. (1979) The mechanics of the circulation, Cambridge University Press.
3. A.D. Nozdrachev, E.L. Polyakov. (2001) Anatomy of the rat, St. Petersburg, Lan'.

BOINC-BASED DESKTOP GRID INFRASTRUCTURE FOR VIRTUAL DRUG SCREENING♣

Nikitina N.N.[1, 2]*, Ivashko E.E.[2], Gupta Y.[1], Ludwig R.[1], Möller S.[1]
[1]*University of Lübeck, Department of Dermatology, Ratzeburger Allee 160, 23538 Lübeck, Germany;*
[2]*Institute of Applied Mathematical Research, Karelian Center of the RAS, Petrozavodsk, Russia*
e-mail: nikitina@krc.karelia.ru
**Corresponding author*

Key words: virtual drug screening, protein-ligand docking, desktop grid, volunteer computing, BOINC

Motivation and Aim: Virtual drug screening aims at specifying and ranking a set of drug candidates according to their predicted binding properties against target proteins. The considerable computational effort typically prevents an adoption of the technology in small research environments. This work demonstrates how *in silico* screening can be facilitated by volunteer computing concepts to complement local research questions in preclinical wet-lab environments.

Methods and Algorithms: Large sets of compounds are freely available from the ZINC database {{http://zinc.docking.org}} and are docked using AutoDock Vina software {{http://autodock.scripps.edu}}. Computational load was distributed with the BOINC middleware {{http://boinc.berkeley.edu}}.

Results: The setup was implemented within the Lübeck Department of Dermatology with research emphasis on autoimmune blistering diseases. Over 40 desktop grid clients, locally and remotely as provided by friends and families of the researchers, have completed over four million computational jobs in two months. The compute facilities allowed to *in silico* confirm the docking scores for multiple structural variants of the receptor. Among many thousands of compounds with higher affinity than the known natural ligand are 20 of known drugs. The priorisation of the many will depend on the results of the *in vitro* validation of those already FDA-approved compounds. Further iterations between wet- and dry-labs are expected.

Conclusion: This work proves the readiness of the virtual screening technology in preclinical environments. Additional work is required to allow for IT skilled biochemists to adopt the technology for themselves and overcome hurdles e.g. for working with hundreds of thousands of files in single directories.

Availability: The package *boinc-server-autodock,* distributed with the experimental section of Debian Linux, provides scripts for an automated setup of complete desktop grid-ready drug screening project.

Acknowledgements: The authors thank Stefano Forli and his colleagues at The Scripps Research Institute, Gianfranco Costamagna and Antony Travis of Debian Med, Minke Informatics, and all our colleagues contributing to our BOINC docking projects.

♣ The work is supported by a DAAD fellowship to N.N. Nikitina and grant RFBR 13-07-00008, "Research and development of mathematical models and algorithms for creation of adaptive BOINC-based Desktop Grids".

AUXIN-INDUCED CHANGES OF TRANSCRIPTOME IN *ARABIDOPSIS THALIANA* ROOTS

Novikova D.D.*, Filipenko E.A., Klimova N.V., Vasiliev G.V., Omelyanchuk N.A., Kochetov A.B., Mironova V.V.
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
e-mail: novikovadd@bionet.nsc.ru
**Corresponding author*

Key words: roots, *A. thaliana*, auxin-induced transcription, AVG-induced transcription, RNA Seq

Motivation and Aim: Phytohormone auxin is the main regulator of plant growth and development. In roots auxin controls many processes: stem nishe maintenance, lateral root development, vascular cell differentiation and others. Auxin-induced transcription of genes is regulated by family of auxin response transcription factors, ARFs and their corepressors Aux/IAA [1]. In arabidopsis genome there are 23 ARF genes and 25 Aux/IAA. Despite of great importance of auxin response in plant development still little is known on ARF target genes and only few of them were experimentally proved. It is why auxin signaling pathways requires modern experimental and bioinformatical methods to be studied in details.

Methods and materials: Plants *A. thaliana L*. Columbia (Col-0) ecotype were grown on S MS at 23 C under long-day conditions (16-hr-light/8-hr-dark cycles). Three-days-old plants were incubated in S MS for 6 hours with 1mM IAA (1), 40 mM AVG (2). Total RNA from seedling roots purified with RiboMinus Plant Kit (Invitrogen) and treated with Solid Total RNA-Seq Kit (Invitrogen) was sequenced on Solid 5500. Bioinformatic analyses of RNA Seq data was performed in collaboration with Ivo Grosse group (Martin Luther University, Germany). Differential expression of top genes was verified by qRT-PCR.

Results and Conclusion: It's the first time, when auxin-induced transcriptome of *A. thaliana* was studied with RNA Seq. Differentially expressed genes were found in auxin and AVG treated samples compared with the control ones. Among these genes, known auxin targets were found (IAA3, IAA5, LAX3, YDK1, GH3.5 et al.) as well as new ones. We compared RNA Seq results with results of microarray experiment GDS1515 [2]. Functional annotation of genes, which expression increased in auxin treatment and decreased in AVG treatment, was performed. After analyses of coexpression of the most interesting genes with their putative regulators from ARF family and Aux/IAA, gene networks of ARF-Aux/IAA-"target" interactions were composed forming predictions to be tested further in experiments.

The work is supported by RFBR (11-04-01254-a, 12-04-33112, 5278.2012.4) and Dynasty foundation.

References:

1. Hagen G., Guilfoyle T. (2002) Plant molecular biology. T. 49. №. 3-4. C. 373-385.
2. Vanneste S. et al. (2005) Plant Cell. 17(11):3035-50.

BIOINFORMATICS DETECTION OF POTENTIAL RECOMBINATION SITES IN TICK-BORNE ENCEPHALITIS VIRUS

Paramonov A.I.[1]*, Dzhioev Yu.P.[1, 2], Kozlova I.V.[1]
[1]*FSBI "Scientific Center of family health problems and human reproduction" SB RAMS,* Irkutsk, Russia;
[2]*Research Institute of Biomedical Technology, Irkutsk State Medical University, Irkutsk, Russia*
e-mail: Paramonov_a.i@mail.ru
**Corresponding author*

Key words: tick-borne encephalitis virus, recombination sites, bioinformatics detection

Motivation and Aim: Tick-borne encephalitis virus (TBEV) causes severe neurological infection in humans, in some cases with paralysis, paresis and high fatality [1]. Recombination is noted in all DNA and many RNA viruses, including a number of members of the genus Flavivirus [2]. However, the existence of recombination in TBEV is still disputed due to conflicting data from different authors [3]. The aim of this work was the discovery of potential sites of recombination in genomic sequences of TBEV strains using highly sensitive software-based methods to detect recombination events.

Methods and Algorithms: Material is presented by coding part of the genome of 55 TBEV strains from the GenBank database. Determining the position of the recombination sites is performed using software methods implemented in the software package RDP 3.34: RDP, GENECONW, BootScan, Chimaera, 3Seq, SiScan, Maxchi [4]. Phylogenetic hypothesis of presence of recombinations was obtained using the program Splits Tree 4.1 [5]. Statistical test was carried out using the method Phi Test for Recombinations [6] of software system Splits Tree.

Results: We identified 21 strain which contain potential recombination sites. 5 unique recombination sites were recorded. The first is located at the intersection of M and E genes and identified in strains – 205 and Glubinnoe. The second is located in the gene of the structure protein E in the European strain AS33. Third - at the junction of NS4a and NS4b genes and is found in many strains of Far Eastern TBEV genotype, the fourth is also located at the intersection of NS4a and NS4b genes, but is much longer and is found only in the strain Joutseno, fifth - in the Chinese strains Senzhang, MDJ-01, MDJ-02, MDJ-03 in the gene NS5. Phylogenetic networks constructed by Neighbor-net show partitions, indicating the possibility of reticular evolution.

Conclusion: We identified 21 strain, which genomes contain potential recombination sites belonging to the five unique sites. Region inside genes NS4a and NS4b forms a "hot spot" of recombination. Recombination sites are tied to the two genotypes - Far East and European, they are only recorded in the peripheral regions of TBEV area - European and Far Eastern. In the strains of the Siberian region, occupying the center of TBEV area, all software methods we used show no reliable recombination sites.

Availability: available on request from the authors.

References:

1. Thiel H.-J., Collett M.S., Gould E.A. et.al. Family Flaviviridae. In Virus Taxonomy: Classification and Nomenclature // Eighth Report of the International Committee on the Taxonomy of Viruses / Edited by C. M. Fauquet et.al. - Amsterdam: Elsevier.- 2005.-P. 981–998
2. Bertrand Y., Töpel M., Elväng A. et.al. First dating of a recombination event in mammalian tick-borne flaviviruses // PLoS One. – 2012.-V.7.-P. 1-12.
3. Twiddy S.S., Holmes E.C. The extent of homologous recombination in members of the genus Flavivirus // J. of General Virology.-2003.-V. 84.- P. 429–440.
4. Martin D.P., Lemey P., Lott M., Moulton V., Posada D., Lefeuvre P. (2010). RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26, 2462-2463.
5. Bruen T.C., Philippe H., Bryant D. A quick and robust statistical test to detect the presence of recombination //Genetics.- 2006.-V.172.-P. 2665–2681.
6. Huson D.H., Scornavacca C.A survey of combinatorial methods for phylogenetic networks //Genome Biol.Evol. -2011.- V.3.- P. 23-35.

MATHEMATICAL MODELING OF PEPTIDOGLYCAN PRECURSOR BIOSYNTHESIS IN THE CYTOPLASM OF *ESCHERICHIA COLI* CELL

Perfilyeva O.A.[1]*, Likhoshvai V.A.[1, 2], Khlebodarova T.M.[1]
[1]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[2]*Novosibirsk State University, Novosibirsk, Russia*
e-mail: perf@bionet.nsc.ru
**Corresponding author*

Key words: peptidoglycan biosynthesis, cell growth

Motivation and Aim: Determination of bacterial cell growth strategy is of fundamental interest and plays a key role in solving the problem of coordinating cell growth and replication of the genome [1]. Inability to determine cell growth law with experimental methods [1, 2] addresses the problem to theoretical studies. Mathematical modeling of peptidoglycan synthesis dynamics, which is the main component of the cell wall and defines cell growth, is one of the possibilities to solve the problem. Developing peptidoglycan biosynthesis mathematical model of the cytoplasmic stage in E.coli is the first step in determining the growth strategy.

Methods and Algorithms: The modeling process was performed using chemical kinetic approach, as well as King-Altman method. Genetic algorithms on the basis of pulsed nonequilibrium kinetics data and characteristics of enzyme reactions, such as K_m (Michaelis constant) and V_{max}, were used to estimate the elementary reaction model parameters.

Results: An elementary reaction models database of nucleotide peptidoglycan precursor synthesis catalyzed by enzymes MurA, MurB, MurC, MurD, MurE and MurF was developed. Based on it equilibrium and nonequilibrium peptidoglycan biosynthesis models of cytoplasmic stage, which used the concentrations of the enzymes as additional variable, were constructed. The value of the concentrations was estimated according to the accumulation of peptidoglycan precursor during the cell cycle. The numerical analysis of negative regulation loop influence on activity of enzymes MurA, MurC, MurD, and MurE by intermediate products of nucleotide peptidoglycan precursor synthesis was made.

Conclusion: A model displays a stable kinetic of intermediate peptidoglycan precursor products synthesis in the cytoplasm of the cell and can be used to study the cell growth laws.

Acknowledgement: The work was supported by the Program of Siberian Branch of the Russian Academy of Sciences (Integration project № 80).

References:

1. Likhoshvai V.A., Khlebodarova T.M. (2014) Mathematical modeling of bacterial cell cycle: The problem of coordinating genome replication with cell growth. J. Bioinform. Comput. Biol. V.12. № 3. 1450009.
2. Cooper S. (2006) Distinguishing between linear and exponential cell growth during the division cycle: Single-cell studies, cell-culture studies, and the object of cell-cycle research. Theor. Biol. Med. Model. V. 3. P. 10.

SPECIES SPECIFICITY OF ATP-DEPENDENT EFFLUX IN THE LIVER FLUKE *OPISTHORCHIS FELINEUS*

Pirozhkova D.S.*, Tsyganov M.A.
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
e-mail: pirozhkova82@mail.ru
**Corresponding author*

Key words: helminth, ABC-transporters, P-glycoprotein, drug resistance

Motivation and Aim: ATP-binding cassette (ABC) superfamily proteins are the main components of excretion/secretion system of cells and some of them are related to drug resistance. These transmembrane proteins are found in all taxons from prokaryotes to humans. B1-subfamily proteins (Multi Drug Resistance– MDR, P-glycoprotein – Pgp) work as efflux pump and are worth special attention because their involvement to MDR in many different species including helminthes. *Opisthorchis felineus* is a liver fluke, an epidemiologically important parasite of humans and other mammals. It is widespread on the territory of Russia, Kazakhstan, and Eastern European countries. The aim of present work was to investigate ABC transporters B1 subfamily in *O. felineus*.

Materials and methods: NCBI tools (Blast, ORFfinder, CD-search), Sanger sequencing, 3'-RACE, real-time PCR and droplet digital PCR.

Results: Four ABCB1 proteins were found in *O. felineus.* These proteins, called Pgp 1-4, consist of 2 halves, each with 6 transmembrane sections and a cytosolic ATP-binding domain. Results of phylogenetic study confirmed their belonging to ABCB1 (Pgp) subfamily. Three of four Pgp had close homologs in *C. sinensis,* but Pgp4 was unique. Coding region of Pgp 1-3 were completely reconstructed. cDNA of Pgp4 including 3'UTR was sequenced. We also determined exon-intron structure of all Pgp genes.

Pgp gene expression through the life cycle and under treatment with praziquantel, albendazole and carbendazim was studied. It was found that Pgp4 was the prevailing ABCB1 in adult worms: its expression was almost 20 times higher than other ones. Pgp4 expression ratio in adult worms was increased after treatment with carbendazim and praziquantel (1.5-fold and 3-fold increase respectively). We have also shown that hemoglobin causes10-fold increase of Pgp4 gene expression in juvenile worms.

Conclusion: ATP-dependent cell excretion/secretion system with conservative organization was found in *O. felineus*. Pgp4 is species specific and probably crucial protein of ABCB1 subfamily. These results suggest that Pgp4 can be a potential molecular target for overcoming drug resistance in *O. felineus.*

COMPUTER ANALYSIS OF HUMAN SNP CONTAINING SITES BY METHODS OF TEXT COMPLEXITY ESTIMATIONS

Safronova N.S.[1]*, Orlov Y.L.[1, 2]
[1]*Novosibirsk State University, Novosibirsk, Russia;*
[2]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
e-mail: taschasafronova@bionet.nsc.ru
**Corresponding author*

Key words: SNP, mutations, human genome, text complexity, nucleotide poly-track

Motivation and Aim: Analysis of genomic sequences issues the challenge to search for the regions with low text complexity, which could be functionally important in genomes. It could affect structural characteristics of DNA sequences and decrease accessibility of such genome regions for high-throughput sequencing. The methods of nucleotide sequence complexity estimation [1] have wide area of application; also they can be effectively used for DNA analysis on the whole genome, short genome regions and for analysis of SNP in special. SNPs (single nucleotide polymorphisms) are very important in studying human diseases. This work continues the work started earlier [2]. It's devoted to the using the developed program complex with new module in SNP research in human chromosome. For the analysis there were taken long regions [-100;+100] nt on the flanks around the mutation position and short regions [-10;+10] nt on the flanks from UCSC Genome Browser. Also we used available SNP data from rat and mouse genomes (http://www.ncbi.nlm.nih.gov/SNP/).

Methods and Algorithms: Firstly the short regions were studied, and for them there was admitted a certain type of changes of the complexity near the mutation, so secondly the long regions were studied to find possible random changes of complexity. We studied the complexity of these genome sequences by different measures with sliding window. We used Lempel-Ziv complexity, linguistic complexity, Shannon's entropy and monomers frequency measures. Then we studied the nucleotide composition, analyzed the frequency of oligonucleotides occurrence by lengths 5-7 and built the profiles of occurrence for the most frequent oligonucleotides.

Results: Studying the complexity showed that in the mutation region there was a special character of complexity observed: before the mutation the complexity profile got visibly lower, then followed the hop and then the complexity fell again. So the graph of the complexity in this region looked like pit-peak-pit. The composition analysis showed that regions of mutation are saturated by A and T nucleotides. Further frequency analysis showed that there's a significant saturation with poly-A and poly-T tracks.

Conclusion: Neighboring-nucleotide effects on mutation rate can be revealed by text complexity estimations. Human and mouse genome SNP containing sites have similar complexity profiles.

References:

1. Y.L. Orlov, V.N. Potapov (2004) Complexity: an internet resource for analysis of DNA sequence complexity. Nucleic Acids Res. 32(Web Server issue):W628-33.
2. P. Putta et al. (2011) Relatively conserved common short sequences in transcription factor binding sites and miRNA, *Vavilov journal of genetics and breeding,* **15** (4): 750-756.

APPLICATION OF NONLINEAR TIME SERIES ANALYSIS FOR HEMODYNAMIC MODEL VALIDATION ON THE BASE OF PHOTOPLETHYSMOGRAM SIGNAL

Sviridova N.V.[1, 2]*, Sakai K.[1]
[1]*Tokyo University of Agriculture and Technology, Tokyo, Japan;*
[2]*Computing Center, Far-Eastern Branch, Russian Academy of Science, Khabarovsk, Russia*
e-mail: nina.v.svr@gmail.com
**Corresponding author*

Key words: hemodynamics, cardiovascular system, photoplethysmography, nonlinear time series analysis

Motivation and Aim: Estimation of the cardiovascular system (CVS) performance has been an object of many recent studies, due to promising applications for human health care and monitoring. Mathematical modelling of CVS can provide us with valuable information for estimating blood circulation changes and trends in CVS under different conditions. However utilizing of all advantages of modelling in hemodynamics requires careful model validation, which often becomes an issue. Nowadays medical devices allows us to measure various biological signals produced by CVS that provide significant information about actual processes in CVS.

Photoplethysmography (PPG) is a simple and low-cost optical technique that can be used to detect blood volume changes in the microvascular bed of tissue. Although the origins of the components of PPG signal are not fully understood, it's generally accepted that PPG can provide valuable information about the CVS. Utilizing of PPG for model validation would be quite promising; however usage of PPG signal for model validation is not so straightforward, since measured signal differs from the parameters included in the most of conventional mathematical model. To overcome mentioned problem, in this study we have suggested to use nonlinear time series analysis (NTSA) methods as a tool for model and PPG signal's features comparison.

Methods and Algorithms: Applying set of NTSA methods such as largest Lyapunov exponent, nonlinear deterministic prediction, Wayland test translation error in this study we have indirectly compared numerically calculated output from system of equations describing hemodynamics in the vessel with PPG signal obtained from healthy young subjects in controlled environment.

Results and Conclusion: Obtained results of comparison between actual signal and model output allowed us to conclude that NTSA methods can be efficiently used to verify model's adequacy to actual processes in CVS.

GRAPH DATABASE FOR MOLECULAR BIOLOGY: ADVANTAGES OF THE GRAPH REPRESENTATION OF DATA

Sergeev A.V.[1]*, Temlyakova E.A.[2], Sorokin A.A.[2]
[1]*Institute of Mathematical Problems of Biology RAS, Pushchino, Russia;*
[2]*Institute of Cell Biophysics RAS, Pushchino, Russia*
e-mail: antonsergeevphd@gmail.com
**Corresponding author*

Key words: graph database, data integration, database structure, and molecular biology data representation

Motivation and Aim: There are a great number of knowledge-specific databases that store essential and detailed information about a particular organism, types of reactions, molecular mechanisms and objects, etc. Most of that data stored in tabular form either in local files or in relational databases. On the contrary, many aspects of modern biology and medicine are naturally mapped to a graph or network structures. Another important aspect is a data structure mismatch when two databases represent the same biological object differently. This makes a fast and reliable data integration system an urgently needed.

Methods and Algorithms: We suggest using graph representation of molecular biology data taken from major authoritative well-known databases (UniProt, GenBank, RegulonDB, etc). All the data are stored as a one colored attributed graph with molecular objects as nodes with predetermined types of relations and object-specific properties. Data from different sources could be stored in different nodes, so let end user (not the database designer) to decide which data source have to be used. The highly flexible and easily scaled graph-orientated storage system Neo4j [1] serves as a backend and query system for the database.

Results: The prototype of graph database contained information about *Escherichia coli* genome, proteome, metabolome, transcriptional and translational regulation, etc is constructed. The key design issues and decisions are described and their consequences are analyzed.

Conclusion: Graph representation of molecular biology data provides new opportunities for queering database, information retrieval and analyzing patterns in data.

Availability: The first version will be available soon.

Acknowledgements: This work was supported by RFBR grant 14-04-31793 mol_a.

References:

1. Neo4j web-site (Neo Technology): http://www.neo4j.org/

DE NOVO TRANSCRIPTOME ASSEMBLY OF THE STARFISH *ASTERIAS RUBENS*

Sadritdinova A.F.*, Snezhkina A.V., Uroshlev L.A., Dmitriev A.A., Kudryavtseva A.V.
Engelhardt Institute of Molecular Biology RAS, Moscow, Russia
e-mail: asiya.bun@gmail.com
**Corresponding author*

Key words: High-throughput sequencing, Transcriptome, Heat shock, Echinodermata, *Asterias rubens*

Background: The starfish *Asterias rubens L.* is one of the most abundant echinoderm species in the White, Barents, North and Baltic Seas. This species is an important component of marine ecosystems and a model object for certain biological studies, in particular those requiring quantitative estimation of gene expression. We have performed a preliminary analysis of gene expression in tube feet of *A. rubens* using high-throughput sequencing under normal and heat shock conditions.

Methods and Algorithms: Tissues of tube feet of normal adults and underyearlings and of adults after heat shock were used. High-throughput transcriptome sequencing of four samples from tube feet of *A. rubens* was carried out on Illumina GAIIx sequencer (Illumina, USA). The first (1) and the second samples (2) were obtained from adult starfish under normal conditions; the third (3) sample – after a 30-min at 25°C; the fourth (4) – after 30 min at 27°C. Paired-end reads 75 nucleotides long were used. About 50 million reads were obtained for each sample. The reads were assembled into contigs using the Velvet/Oases software. Analysis of differential expression was carried out with the edgeR software package.

Results: We generated a cDNA library containing expressed sequence tags (ESTs) from starfish *A. rubens*. A total of 61273037 (GC-content – 40.8%) and 73538249 (GC-content – 42.0%) 75 nt paired-end reads were obtained from high throughput sequencing for normal conditions (the first and the second sample) 74525446 (GC-content – 40.4%), 73810609 (GC-content – 40.2%) and 15202358 (GC-content – 63.1%) for different stress conditions (the third, the fourth and the fifth sample respectively). For assembly we estimated length of k-mer via kmergenie software. We obtained optimal k-mer lengths equal to 19, 17, 21, 25 and 21. The total number of transcripts was 102169 (N50 – 536) for sample 1, 131085 (N50 – 223) for sample 2, 88712 (N50 – 925) for sample 3, and 77810 (N50 – 1243) for sample 4. The max length of transcripts was 22697 bp for sample 1, 8564 bp for sample 2, 18067 for sample 3, and 24412 for sample 4. Thus, we obtained 1200-1500 genes approximately. Transcriptome of echinoderm *Strongylocentrotus purpuratus* was used as a reference.

Conclusions: The assembly allowed us to identify several candidate reference genes for the transcriptome analysis under heat shock as well as additional candidates for differentially expressed genes in *A. rubens*. In future, we have a plan to test the candidate reference and differentially expressed genes using quantitative polymerase chain reaction in real-time (qRT-PCR). In addition, the identified transcripts of the starfish enrich the limited knowledge base of echinoderm's molecular mechanisms of adaptation.

Acknowledgements: The high-throughput sequencing was performed at the EIMB RAS "Genome" center.

ANALYSIS OF BACTERIA AND ARCHAEA GENOMES AVAILABLE IN GENBANK DATABASE BY "ELOE" PROGRAM

Sokolov V.S.*, Matushkin Yu.G.
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;
Novosibirsk State University, Novosibirsk, Russia
e-mail: sokovlad1@bionet.nsc.ru
**Corresponding author*

Key words: translation elongation efficiency, codon composition, mRNA secondary structures, Bacteria, Archaea

Motivation and Aim: Various organisms are known to optimize primary structure of genes in different ways to increase its translation efficiency [1]. Some organisms optimize codon composition, others – count and/or stability of mRNA secondary structures. There are also organisms which optimize both of this nucleotide sequence features. The main aim of this study is to find out how prokaryotes are grouping into five types with particular evolutional optimization of gene nucleotide sequences.

Methods and Algorithms: EloE program performs the analysis of genes nucleotide sequences. The key parameters of the analysis are gene codon composition and presence of perfect local inverted repeats in mRNA [1]. Gene sequences were obtained from the GenBank database (ftp://ftp.ncbi.nih.gov/genomes/Bacteria/) in gbk file format. The program calculates five types of EEI index and determines which type work most effectively in each organism.

Results: Genomes of 2747 prokaryotes (2582 bacteria, 165 archaea) were analyzed. Consequently it was shown, that distributions of Bacteria and Archaea differ. The majority of Bacteria (45%) belong to first type of EEI index. It means that organisms in this group optimized genes codon composition during evolution. On the other hand, the majority of Archaea (59%) belong to fourth type of EEI. These organisms optimized genes codon composition as well as count of secondary structures in mRNA.

Conclusion: As a result genome-wide analysis of Prokaryote was carried out. It was shown, that distributions of Bacteria and Archaea differ.

Availability: http://www.bionet.nsc.ru/razrabotki/prikladnyie-razrabotki/programmyi-dlya-evm.html.

Acknowledgements: This work was supported in part by the program of the Presidium of Russian Academy of Sciences "Origin and evolution of the biosphere" (no. 15).

References:

1. V.A. Likhoshvai, Y.G. Matushkin. Differentiation of single-cell organisms according to elongation stages crucial for gene expression efficacy // FEBS. 2002.

COMPUTER ANALYSIS OF HUMAN GENE EXPRESSION DATA USING BioGPS DATABASE OF MICROARRAY AFFYMETRIX U133

Spitsina A.M.[1, 2]*, Orlov Y.L.[1]
[1]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[2]*Novosibirsk State University, Mechanics and Mathematics Department*
e-mail: anastasia.spitsina@gmail.com
**Corresponding author*

Key words: gene expression, microarray, gene network, human genome

Motivation and Aim: Analysis of gene expression using microarrays is associated with many problems in bioinformatics – e.g. for cancer research. There are dozens of microarray types, but we interested in the most common method - is the technology of synthesis of short oligonucleotide probes on a surface of the microchip developed by the Affymetrix, which allow to simultaneously measure the expression of thousands of genes. Microchips GeneChip U133 series are widely used in clinical practice, large amount of experimental data was collected (BioGPS (biogps.org), GEO NCBI databases). Initial design of oligonucleotide probes microchip may not correspond to a target transcript (gene) and contain a number of technical problems associated with genomic annotation of samples. It leads to inconsistent results, so for working with data expression it is required to develop a specific software package that allow to filter noisy expression signals on the microchip and make it easier to work with large amount of data, and aim of work is to create a set of programs for multipurpose computer research on expression of genes on the microchips.

Methods and Algorithms: To meet issues of human genome analysis of expression data was developed set of computer programs to work with existing databases. Tools in C++ for processing Affymetrix U133 microarray data, including algorithms for estimating coefficients of correlation and filtration for samples of genes, are implemented to perform the different tasks, e.g. to identify characteristics of genes which actively expressed in human brain tissues and analyze characteristics of the expression of pairs of transcripts which co-localized in the genome including cis-antisense transcripts.

Results: Program for analyzing correlation was developed, which is also useful for analyzing human brain gene expression and gene networks (e.g. were obtained correlation matrix for networks of cholesterol level regulation and circadian rhytm regulation), for research microchip quality [2], and analyzing tissue specificity (group of gene with high expression in brain cells was identified using BioGPS DB [1]), including 3-dimensional chromosomal contacts.

Complexity of the working with obtained experimental data – is large size of databases, which included about forty thousand lines and hundred columns. Created tool contain several options to simplify work with databases, which size is large. It allow to integrate data from one database to another and use obtained database for further processing or analysis. Program had understandability text interface, and for its future development uses clearly structures and functions.

Conclusion: Among samples microchip Affymetrix U133, that are represented in the BioGPS database, were identified some with high expression, collected samples of genes for which the expression is higher in brain structures, created a tool that allows to create a chart for correlations of expression of gene pairs. Structural features of genes with high expression (number of exons, the length of the transcript, link with alternative splicing) were detected. It is planning to integrate this tool into a package for statistical data processing expression of genes that is developed in ICG SB RAS.

Availability: https://github.com/amspitsina/gene-expr, http://about.me/anastasiaspitsina

References:

1. I.V. Medvedeva, O.V. Vishnevsky, N.S. Safronova, O.S. Kozhevnikova, M.A. Genaev, A.V. Kochetov, D.A. Afonnikov, Y.L. Orlov (2013) Computer analysis of data on gene expression in brain cells obtained by microarray tests and high-throughput sequencing // Vavilov Journal of Genetics and Breeding, V.17, № 4/1, p. 629-638.
2. Orlov Yu.L. et al (2007) Quality assessment of the Affymetrix U133A&B probesets by target sequence mapping and expression data analysis. In Silico Biol. 7(3):241-60.

THE SYSTEM OF 3D VISUALISATION FOR SOFTWARE PACKAGE "HAPLOID EVOLUTIONARY CONSTRUCTOR 3D"

Chekantsev A.D.
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;
Novosibirsk State University, Novosibirsk, Russia
e-mail: lashin@bionet.nsc.ru

Motivation and Aim: Modeling of the evolution of bacterial communities is an important task of modern bioinformatics. To resolve this problem, in the ICG SB RAS a technique for modeling and software tool "Haploid Evolutionary Constructor" (HEC) were developed. The HEC is designed to simulate the coexistence of populations of trophically related unicellular haploid organisms, particularly bacteria [1]. HEC allows the user to simulate the interactions between a population and the environment, including a variety of evolutionary population processes such as horizontal gene transfer between bacteria of different species, loss of genes, and mutations.

Since the HEC is a software product that automates research, and the most part of information people perceive visually, so it is very important to present information in clear view. Results obtained by using the HEC should be clear and understandable to the end user, which promotes efficient data analysis. A computer model of prokaryotic communities has many customizable settings, so it is important to provide interactive work of the HEC by developed management tools to achieve greater flexibility in the software package.

The aim of this work is to create a new 3D GUI version of HEC for visual imaging of necessary parameters, as well as creating the necessary control for setting the initial parameters of the modeling process and setting the various options of the simulations.

Methods and Algorithms: The HEC implements multilayer modeling approach, where each "layer" represents a single computer model of biological objects and processes related to a particular level of biological organization. In developing of the interface, we used the technology OpenGL and free cross-platform development environment Qt Creator. For charts and graphics creation, we used free library QCustomPlot.

Results: Designed interface allows user to consider a predetermined volume and dispersed bacterial populations as an interactive 3D model. One can observe all the parameters and variables of the model with plots of dynamics of population growth, substrate concentrations, genetic spectrum of populations and graph of trophic relationships between different types of bacteria, both for overall environment and for its specific subcell.

Conclusion: The interactive interface allows user to quickly obtain information about any element of the investigated model and presents it in a user-friendly way.

References:

1. Lashin S.A., Matushkin Yu.G. Haploid evolutionary constructor: new features and further challenges // In Silico. Biol. 2012. V.11. N. 3. P. 125-135.

PIN-TRANSPORTERS IN THE ROOT MERISTEM OF *ARABIDOPSIS THALIANA* L.: IMAGE ANALYSIS OF EXPRESSION PATTERNS

Chernova V.V.[1]*, Doroshkov A.V.[1], Omelyanchuk N.A.[1], Mironova V.V.[1, 2]
[1]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[2]*Novosibirsk State University, Novosibirsk, Russia*
e-mail: chernovavasilina@yandex.ru
**Corresponding author*

Key words: auxin, *Arabidopsis thaliana*, image analysis, microscopy

Motivation and Aim: Plant hormone auxin is the main regulator of plant growth and development. The PIN-FORMED (PIN) family proteins provide auxin polar transport, forming tissue gradients and localized concentration maxima of this hormone, instructive for differentiation of various cell types. Auxin regulates PIN protein expression at different levels, including transcription, protein stability, intracellular localization and transport. In this study we investigated the dose-dependent effect of auxin on PIN protein expression in the root *A. thaliana*.

Methods and Algorithms: We used in the experiment four following A. thaliana reporter lines: PIN1::PIN1-GFP, PIN2::PIN2-GFP, PIN4::PIN4-GFP and PIN7::PIN7-GFP. Transgenic plants were grown under standard conditions on solid nutrient medium S MS without auxin as a control and containing different auxin concentrations (NAA 0.0002, 0.002, 0.02, 1 mM / l). Microscopic analysis of reporter activity in the root cells of 3 days old seedlings was carried out after 6, 18 and 29 hours on a confocal microscope. The intensity of the luminescence of the reporter protein GFP and the length of its expression domain in micro-images were analyzed by ImageJ. Quantitative data obtained by image processing were further processed in Excel. To assess the significance of differences between the treated samples and the control we used Student's t-test.

Results: Analysis of changes in expression profiles of PIN1, PIN2, PIN4 and PIN7 proteins using GFP reporter 6, 18 and 29 hours after auxin treatment with different doses revealed that these changes and their dynamics are specific for each PIN protein under investigation and auxin dose dependent.

Conclusion: Auxin changes the expression profiles of PIN1, PIN2, PIN4 and PIN7 proteins in the root. The dynamics and direction of these changes depend on the auxin dose and the time passed aftter treatment. Our investigation allowed to establish the range auxin concentrations specific for each from PIN1, PIN2, PIN4 and PIN7 proteins at which these proteins more effectively perform their functions.

Acknowledgements: Microscopy was performed in the Shared Facility Center for Microscopic Analysis of Biological Objects SB RAS. The work is partially supported by the Dynasty Foundation grant for young biologists, RAS program A.II.6, Integration project SB RAS 80 and RFBR grants 11-04-01254-a, 12-04-33112 and 5278.2012.4.

METHOD TO PREDICT THE PERCENTAGE OF CELL TYPES IN HUMAN BLOOD

Igolkina A.A.*, Samsonova M.G.
St. Petersburg State Polytechnical University, Russia
e-mail: igolkinaanna1@gmail.com
**Corresponding author*

Key words: blood, heterogeneous tissue, cell type percentage, expression deconvolution

Motivation and Aim: Blood is the most investigated heterogeneous tissue. It contains a variety of cell types, of which the major types are Lymphocytes, Monocytes, Granulocytes, Erythrocytes, Megakaryocyte (Lymphocytes and Granulocytes are complex cell groups in turn). Gene expression data from blood genomics studies is widely used in medical diagnosis. Most of these studies are based on the analysis of total peripheral blood mononuclear cells (PBMCs). PBMCs are composed of over a dozen cell types, the proportion of which varies in blood samples from individual people. This variability significantly influences genome-wide gene expression data. The heterogeneity of blood distorts the data, however, it is often discarded due to the lack of data on the composition of the samples. The application of experimental methods to separate or quantify constituents from each sample is time-consuming and does not solve the problem. Therefore, an attractive alternative is to accurately deconvolve gene expression data. Here we develop a method to predict the percentage of cell types in a blood sample from whole genome gene expression data.

Methods and Algorithms: We used a data set of 628 patients with known gene expression levels and percentages of 5 cell types in blood samples. The samples were divided into training set, testing set and set for prediction. We built and tested various predictive models based on PCA, linear regression (with and without prior knowledge of cell type specific signatures obtained from pure cell types [1-2]) and SVM with different kernel types. To select a gene subset which provides the best prediction of cell proportions we use random sampling, filter methods(correlation, mutual info), wrapped methods (RFE), as well as some prior knowledge in the form of sets of marker genes [3]. It is noteworthy that both feature selection and predictive methods were constructed for each individual cell type independently. To estimate the performance of different approaches Pearson correlation coefficient between estimated and true cell type proportion in data was calculated.

Results: We achieved the best estimation of cell type proportions using filter or wrapped feature selection method and SVM with Gaussian kernel. This approach significantly improved the Pearson correlation between true and estimated cell type proportions to approximately 0,8-0.95, that is high enough for further studies.

Conclusion: We have developed a method that can accurately predict the percentage of cell types from whole genome gene expression data in human blood samples. Our approach can be used to predict the percentage of cell types in other tissues.

Availability: The MATLAB script is available on request from the author.

References:

1. A.R. Abbas et al. (2009) Deconvolution of blood microarray data identifies cellular activation patterns in systemic lupus erythematosus. *PloS one* 4.7.
2. T. Gong et al. (2011) Optimal deconvolution of transcriptional profiling data using quadratic programming with application to complex clinical blood samples. *PloS one* 6.11.
3. R. Gaujoux, C. Seoighe (2013) An introduction to gene expression deconvolution and the CellMix package. *Bioinformatics.* 29 (17): 2211-2212.

COMPARISON OF TWO METHODS FOR DETECTION OF EXCEPTIONAL WORDS IN GENOMES

Rusinov I.S.,[1] Ershova A.S.,[2,3,4,*] Karyagina A.S.,[2,3,4] Spirin S.A.,[1,2,5] Alexeevski A.V. [1,2,5]
[1]*Lomonosov Moscow State University, Faculty of Bioengineering and Bioinformatics, Moscow, Russia;*
[2]*Lomonosov Moscow State University, Belozersky Institute of Physical and Chemical Biology, Moscow, Russia;*
[3]*Gamaleya Institute of Epidemiology and Microbiology, Moscow, Russia;*
[4]*Institute of Agricultural Biotechnology RAS, Moscow, Russia;*
[5]*Scientific Research Institute for System Studies RAS (NIISI RAS), Moscow, Russia;*
**Corresponding author. isrusinov@gmail.com*

Key words: restriction sites, word compositional bias, site occurrence

Motivation and Aim: Several methods are widely used for detection of short sequences (words) in genomes that are under evolutionary pressure. Two most common methods are maximal order Markov model (Mmax) based one (see S.Schbath [1]) and S. Karlin's method [2]. The latter takes into account observed frequencies of all subwords of the word, including discontinuous, for expected frequency estimation while the former – only subwords obtained by deleting one letter from either 5', 3' or both word ends. We compared these two methods in terms of detection of restriction sites avoided in a genome.

*Methods and Algorithms:*1) Genome sequences and restriction sites. We used 2141 complete prokaryotic genomes. We chose recognition sites of restriction-modification systems as target short words for methods comparison. Avoidance of many restriction sites in prokaryotic genomes was shown [3]. Among all restriction sites only degenerate ones, like CCNNGG where N is any nucleotide, were selected for a comparison of methods precision.

2) Estimation of site avoidance. We estimated the bias in sites representation with the observed to expected number ratio, called "contrast". If the ratio is less 1 then there are less sites in a genome sequence than expected, and vice versa. Mmax based method and Karlin's one were used for computation of expected site number.

3) Simulation of evolutionary press. We used 100000 bp random Bernoulli sequence as a base for selection simulation. Certain words were randomly mutated with a given probability. The both contrasts were computed for all six or less bp words in the resulted sequences.

Results: Contrasts of recognition sites of restriction-modification systems in prokaryotic genomes were computed with both methods. We found that two lists of 4% of the most underrepresented sites differ by 40 %. Thus, the method used has significant impact on the obtained results.

If degenerate restriction site (like CCNNGG) is under pressure to reduce number of its occurrences, then we expect to observe this site more avoided then its non-degenerate variants (CCATGG, CCTGGG, and so on), especially non-palindromic ones. Hence, the method seems to be better if it matches this rule more often. We found that Karlin's contrast showed better results than Mmax based one: 87% vs. 38% of cases matching the assumption (only non-palindromic variants were taken into account).

We known at least one experimentally confirmed case approving this result. Gelfand and Koonin [3] predicted specificity of MjaIV restriction-modification system as

GTYRAC based on site underrepresentation in the genome of *Methanocaldococcus jannaschii*. Mmax based method was used. MjaIV system was experimentally characterized [4] and its specificity in REBASE is GTNNAC. Notably, Karlin's contrast reveals underrepresentation of GTNNAC but not GTYRAC and non-degenerate variants of GTNNAC.

In addition, we evaluated the methods dependence on subword frequency biases with simulated selection against certain word in random bernoulli sequence and tracing effects on representation of other short words in the sequence with both methods. In case of degenerative word under selection results are just the same that were observed for the degenerative restriction sites in the real genome sequences. We also found, that the Mmax based ratio always disperse from 1 worse than the Karlin's one. This effect could be explained with much more information about sequence use in case of Karlin's method.

Conclusion: We found the difference in occurrences estimate between two methods. We showed that Karlin's method is more precise, probably due to use of all site subwords frequencies for the representation evaluation.

References:

1. Schbath S, Prum B, de Turckheim E. Exceptional motifs in different Markov chain models for a statistical analysis of DNA sequences. J Comput Biol. 1995;2(3):417-437.
2. Karlin S, Cardon LR. Computational DNA sequence analysis. Annu Rev Microbiol. 1994;48:619-654.
3. Gelfand M, Koonin E. Avoidance of palindromic words in bacterial and archaeal genomes: a close connection with restriction enzymes. Nucleic Acids Res. 1997;25(12):2430-2439.
4. Zheng Y, Posfai J, Morgan RD, Vincze T, Roberts RJ. Using shotgun sequence data to find active restriction enzyme genes. Nucleic Acids Res. 2009;37(1):e1.

STATISTICAL ANALYSIS OF GENE EXPRESSION DATA BY RANK CORRELATION COEFFICIENTS

Spitsina A.M.[1,2]*, Kulakova E.V.[1,2], Safronova N.S.[1,2], Orlova N.G.[1,3]*
[1]*Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[2]*Novosibirsk State University, Novosibirsk, Russia;*
[3]*Siberian University of Consumer Cooperation, Novosibirsk, Russia*
e-mail: anastasia.spitsina@gmail.com; orlovanina2@mail.ru
**Corresponding author*

Key words: statistics, rank correlation, gene expression, microarray, human genome

Motivation and Aim: Computer analysis of gene expression on microarrays demands development of statistical tools. Co-expression of genes might be caused by noise signals (such as wrong annotation of nucleotide probes) and biological co-regulation. One of the measures of joint gene expression is correlation coefficient between two vectors of expression levels (set of experimental samples, such as patients groups, etc). Sources of data for correlation analysis are microarrays, in particular widely used Affymetrix GeneChip U133 series. It was shown that location of genes probes in cis-antisense orientation in genome gives false correlations [1]. To estimate statistical significance of co-expression coefficients on new microarrays data one needs computer tools [2]. Aim of the work is programming of statistical module for rank correlation analysis based on Kendall tau and Spearman rank correlation coefficients.

Methods and Algorithms: The set of computer programs was developed for gene expression data analysis. Tools in C++ for processing Affymetrix U133 microarray data, including algorithms for processing and filtering of expression data by probe sets were developed.

Results: Program for analyzing correlation allows calculate linear Pearson correlation coefficient, as well as Kendall tau and Spearman rank correlation coefficients. Statistical significance of the coefficient dependent on sample size is estimated based on pre-defined statistical table data. Based on BioGPS database (biogps.org) containing tissue-specific expression data we calculated correlation matrix for specific gene groups (from gene networks) and for random gene samples. It was shown that for random genes groups the correlation coefficients distribution follows Gaussian distribution. Both rank correlation coefficients have more contact distribution shape.

Conclusion: Among Affymetrix U133 data samples represented in the BioGPS database, we identified sets of genes correlated across 80 tissues and cell types in human. The computer tool developed allows calculation of correlation coefficients distribution on large data sets, up to 40000 probes. We are going to integrate this tool into JACOBI-4 package for statistical data processing of gene expression data.

Acknowledgements: The work is supported by RFBR (14-04-01906).

References:

1. Y.L. Orlov et al. (2007) Quality assessment of the Affymetrix U133A&B probesets by target sequence mapping and expression data analysis // In Silico Biol. 7(3):241-60.
2. Y.L. Orlov, V.M. Efimov, N.G. Orlova (2011) Statistical estimates of transposable element expression in the human genome based on clinical microarray data on expression // Vavilov journal of genetics and selections. 15(2): 327-339. (In Russian).

MODELING OF CELL DYNAMICS IN THE ROOT APICAL MERISTEM WITH DYNAMICAL GRAMMAR

Lavrekha V.V.[1*], Omelyanchuk N.A.[1], Mironova V.V. [1,2]
[1]*Institute of Cytology and Genetics SB RAS, 10 Lavrentyeva, Novosibirsk, Russia*
[2] *Novosibirsk State University, Pirogova 2, Novosibirsk, Russia*
[*]*Corresponding author: vvl@bionet.nsc.ru*

Key words: *RAM, cell dynamics, Dynamical Grammar, auxin, cytokinin.*

Background. Root apical meristem (RAM) is the plant stem cell niche which provides for the formation and continuous development of the root. In the root tip along the central axis there are several zones with principally different cell dynamics: columella, quiescent center, proliferation domain, transition domain and elongation zone. It is known that in the root tip, auxin and cytokinin distributions are nonuniform with maxima and gradients. The distal maxima of auxin and cytokinin in the root tip maintains the stem cell niche. Here we propose that patterning of the plant hormones can be sufficient condition for the zones emergence and maintenance. This hypothesis was tested by mathematical modeling with Dynamical grammar [1].

Material and Methods. Previously [2] we developed a model to study auxin patterning in the stem cell niche. In the model cell divisions were regulated by auxin and hypothetical morphogen [2.] The model was created using package Plenum [3] in Mathematica. The cell dynamics was modeled by Dynamical Grammar rules that specify a weak (breakable) spring potential function between neighboring cells[1]. Here we extended the model with the rules for cytokinin effects on cell dynamics. We substituted the *Division Factor* by cytokinin and auxin and considered their effects on cell cycle.

Results and Conclusions. In the resulting model we observed simultaneous formation of (1) auxin distribution; (2) cytokinin distribution; (3) zonation of the root along the central axis with respect to cell growth/division. We found the model parameters under which the proliferation domain was stable in development. We assume that this model will allow us to study an effect of auxin and cytokinin on cell dynamics in the plant.

Acknowledgments. The work is supported by the Dynasty Foundation grant for young biologists and RSF 14-14-00734.

References:

1. E. Mjolsness. (2010) Towards Measurable Types for Dynamical Process Modeling Languages.Electr. Notes Theor. Comput. Sci. 265: 123-144
2. V.V. Mironova et al. (2010) A plausible mechanism for auxin patterning along the developing root. BMC Systems Biology, 4:98.
3. http://computableplant.ics.uci.edu/theses/guy/Plenum.html

ADVANCED ACHIEVEMENTS OF ILLUMINA NEXT-GENERATION SEQUENCING

Shapovalov I.
InterLabService, Moscow

Key words: cloud computing, bioinformatics, illumina, data analysis, next-generation sequencing

Scince 1970 when the first DNA sequencing method was invented by Ray Wu determination of nucleotide sequence evolved into the ultra-high output process capable of providing the whole genome sequences within a day. High throughput and affordability of the next-generation sequencing (NGS) results in huge amount of sequence data produced every day.

The very first Illumina sequencer, Genome Analyzer, presented in 2006 was capable of sequencing human genome in one run. And it was very complex device that required trained personal with high skills in operating the sequencer. In 2014 Illumina presents NextSeq500, tabletop whole-genome sequencer that opens the new era of very user-friendly NGS systems. It takes several minutes to start the run and requires no special skills to operate.

Nowadays getting NGS data becomes a simple and routine procedure, so other sides of NGS are prone to be bottleneck of the workflow. According to the latest surveys the main difficulty that faces the researcher is data analysis.

Huge amount of data that is generated during the run requires computational analysis using variety of the bioinformatics tools. Bioinformatic tools are usually created and used by those who have background in mathematics and informatics, so the programs aren't easy to operate for researchers in other fields. Another limitation for bioinformatic analysis is a need for high-throughput servers.

Another way to do bioinformatics is to apply cloud computing to the process. The development of the Internet gave start to cloud services that produce computational resources that are available from every computer with global network access. To get rid the bottleneck of data analysis from NGS process Illumina created **BaseSpace** (http://basespace.illumina.com/), the cloud bioinformatic service that provide experiment management features and is intended to store and process data obtained on Illumina sequencers.

BaseSpace provides the tools for the whole experiment management process. One can create list of samples, combine samples in pools for sequencing, form samplesheets for the sequencer and define run parameters. Samplesheets and run protocol are transferred to the device via internet, so that there is no need to enter this information manually on the sequencer.

Data from sequencer is transferred to the **BaseSpace** during the run and is stored on the encrypted Amazon storage service. All data is downloadable, so anyone who have the access to the appropriate **BaseSpace** account or was given the access to the data can download standard FASTQ files to analyze data offline.

The data model is based on two structures, *Runs* and *Projects*. *Runs* include samples that were processed in the single run, and *Project*s combine samples and data that refer to the same experiment. Samples that were sequenced in different *Runs* can be combined in *Projects* for simultaneous data analysis.

Simplicity of data analysis is achieved by using over 25 **BaseSpace** *Applications*. These user-friendly web-applications based on the latest and wide-used bioinformatic algorithms. *Applications* cover the variety of data analysis tasks like raw data processing,

variants identification, statistics and summary generation. Separate *Applications* represent Illumina *Workflows* that perform standard data analysis equivalent to the MiSeq Reporter analysis.

All the Applications are easy-to-use one-click bioinformatic tools with downloadable results of the analysis. Integrated sharing system allow quick and easy sharing of the initial raw data and all of the analysis results.

All Illumina and most of the non-Illumina applications are free of charge, others can be accessed through internal **BaseSpace** payment system.

Illumina **BaseSpace** is the service that integrates the processes of sequence data obtaining and consequent bioinformatic analysis with NGS raw data being seamlessly transferred to web-based bioinformatics center. Simple analysis from raw data to the final summary report with **BaseSpace** *Applications* require no sophisticated high-throughput servers, so one can perform analysis from every computer with internet access.

KINETIC SIMULATION OF MITOCHONDRIAL SHUTTLES

Korla K.
School of Life Sciences, University of Hyderabad, India 500046
e-mail: kskorla@gmail.com

Key words: *ETC, NADH, Krebs cycle, malate-aspartate shuttle, citrate-pyruvate shuttle*

Motivation and Aim: Shuttles are used to carry metabolites across biological membrane to maintain appropriate concentrations on both sides. Besides metabolites, shuttles also facilitate transfer of redox equivalents, across the mitochondrial membrane, as it is impermeable to NADH/NAD couple. The three shuttles reported here work in close association with Krebs cycle to maintain metabolite concentrations and supply redox equivalents to the electron transport chain in mitochondria. In the present work, we aim to make a kinetic model of three shuttles, malate-aspartate, citrate-pyruvate and glycerol phosphate, and integrate it with our earlier model of Krebs cycle and oxidative phosphorylation (Oxphos)[1]. Such a model will enable us to study the system in a kinetic perspective and the effect of several components in the kinetics of the system.

Methods and Algorithms: The set of reactions were simulated using Octave, a high-level, open source interpreted language. It provides the numerical solutions of linear and non-linear problems and performs other numerical experiments[2]. Ordinary differential equation solver was used for simulation calculations and graphical representations (GNUPLOT). Scripts were written individually for the three shuttles and were integrated with the scripts for Krebs cycle and Oxphos.

Results: Simulation curves of shuttles individually and the combined reaction set, shuttles combined with Krebs cycle and Oxphos, were obtained. These curves relate the concentration of various components in the same time frame and can help us to view the impact of a single component on all the other components of the system as a single simulation result.

Conclusion: Kinetic model of shuttles with Krebs cycle and Oxphos have been successfully made. Such a model can be effectively used to study the system and the effects of perturbation in the present system.

Acknowledgements: he author gratefully acknowledges the receipt of Senior Research fellowship from University Grants Commission, India. Thanks are also due to V Lakshmipathi, M Durga Prasad and C K Mitra for their valuable guidance.

References:

1. K.Korla, C.K.Mitra. (2014) Modelling the Krebs cycle and oxidative phosphorylation, Journal of Biomolecular Structure and Dynamics, 32(2), 242-256.
2. GNU Octave: http://www.gnu.org/software/octave/

COMPUTATIONAL AND FUNCTIONAL ANALYSIS OF AUXIN RESPONSE ELEMENTS TGTCTC DIMERS IN *ARABIDOPSIS THALIANA* L. GENES PROMOTERS

Wiebe D.S.
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
e-mail: wiebe@bionet.nsc.ru

Key words: Arabidopsis, auxin, ARF

Motivation and Aim: Plant hormones called auxins regulate various growth and developmental processes including apical domination, roots formation and formation of new organs. The primary response to auxin is regulated on the level of transcription via cis-regulation elements AuxRE, which are binding sites of the ARF family of transcription factors. Previously, the most common method of AuxRE elements identification in plant genes promoters was consensus TGTCTC. However in Ulmasov et al. (1997), revealed the core TGTC as the most conservative part of binding site since the mutations in other hexamer positions did not strongly affect the affinity. It is known that transcription factors from ARF family can? function as dimers. Hence AuxRE may be observed in clusters [1]. Two AuxRE may be located in different strands and orientations (direct, inverted and everted repeats). In this work we analyzed these types of repeats in the promoters of *Arabidopsis thaliana L*. genes. We compared the sublists of ID's of the genes that contain repeats with the gene ID's from microarray experiments data. Also for all genes that have repeats in their promoters the functional annotation was carried out.

Methods and Algorithms: Promoter sequences of *Arabidopsis thaliana L*. genes were taken from the TAIR database [2]. The promoter sequences of different length: 1000 bp + 5UTR and 2000 bp + 5UTR were used. For the analysis of the expression of genes that contain repeats we used the data from 12 microarray experiments with different concentrations of auxin and different time of treatment from GEO database [3]. Analysis was carried out for the repeats of TGTC motif with spacing from 2 to 30 bp Functional annotation was carried out using the DAVID tool from the web resource of National Institute of Allergy and Infectious Diseases [4].

Results: Sublists of genes which contained repeats of TGTC motif with different spacers in their promoters were compared with sublists of gene ID's from the data of microarray experiments on auxin treatment, as a result the genes that contain specific repeats were identified, these genes were revealed as up-regulated in several microarray experiments with different conditions of auxin treatment. For functional annotation we also composed sublists of genes with repeats in their promoters with different spacing and orientation. As a result of the annotation we identified many sublists of genes with specific repeats, which were enriched in terms of functional annotation related to the response to auxin, response to other hormones which relate to processes of plant growth and development, for instance, ethylene.

Conclusion: We have identified an interesting intersection of experimental data and AuxRE repeats recognition data and strengthened these results with functional annotation.

References

1. Ulmasov, T., Hagen, G. and Guilfoyle, T.J. (1997) «ARF1, a transcription factor that binds auxin response elements» *Science 276:* p. 1865–1868.
2. [Web resource] // Carnegie Institution for Science Dept. of Plant Biology. URL: http://www.arabidopsis.org
3. [Web resource] // National Center for Biotechnology Information. URL: http://www.ncbi.nlm.nih.gov/geo
4. [Web resource] // National Institute of Allergy and Infectious Diseases (NIAID), NIH. URL: http://david.abcc.ncifcrf.gov

Author index

T

U

V

Y

Z

Компания Хеликон – крупнейший поставщик в России самого современного оборудования и реактивов для широкого спектра научных исследований, молекулярной диагностики, криминалистики. Клиентами Компании Хеликон являются исследователи в области молекулярной и клеточной биологии, протеомике, молекулярной диагностике, медико-генетической экспертизе и идентификации личности, а также ветеринарии и селекции в сельском хозяйстве.

Мы являемся дистрибьютором ведущих мировых брендов:

Life Technologies,
Thermo Fisher Scientific,
BioRad,
Eppendorf,
Affymetrix,
Fluidigm,
SSI и др.

Почему клиенты работают с нами:

- Мы динамичная, быстро развивающаяся компания.
- Мы являемся не просто дистрибьютором ведущих мировых производителей, но и их методическим и сервисным центром в России.
- Мы всегда на шаг впереди использования и внедрения новых методов и технологий в молекулярной биологии.
- Мы тесно работаем с ведущими биотехнологическими компаниями и известными институтами.
- Мы постоянно проводим практические школы по NGS секвенированию, цитометрии, микроскопии.
- Мы работаем по всей России, имея представительства в 7 регионах.
- Мы обеспечиваем обучение персонала и постоянную методическую поддержку, технический сервис, своевременную доставку оборудования и расходных материалов.

119991 Москва, Ленинские горы, МГУ, д. 1, строение 40,
НИИ Физико-химической биологии им. А.Н.Белозерского, Лабораторный корпус «А»
Телефоны: +7 (499) 705-50-50, +7 (499) 769-51-60 Факс: +7 (495) 930-00-84
mail@helicon.ru, helicon.ru

Представительство в Сибирском регионе:
г. Новосибирск, ул. Инженерная, д. 28
+7 (383) 207-84-85, novosibirsk@helicon.ru

Представительство в Северо-Западном регионе: г. Санкт-Петербург, ул. Ушинского, д. 5, корп. 2, лит. А, пом. 1-Н.
+7 (812) 244-85-52, spb@helicon.ru

Представительство в Южном регионе:
г. Ростов-на-Дону,

Представительство в Приволжском регионе:
г. Казань, ул. Петербургская,

Компания Genotek оказывает услуги по высокопроизводительному секвенированию и генотипированию различных организмов с 2010 г. и является ведущим поставщиком на рынке генетических услуг.

Основателями компании являются выпускники биологического и механико-математического факультетов МГУ им. М.В. Ломоносова. В Genotek работают выпускники ведущих вузов России: МГУ, МИФИ, РГМУ им. Пирогова, МГТУ им. Н. Э. Баумана и др. Многие из нас имеют опыт работы в ведущих международных лабораториях.

Наши клиенты - ведущие научно-иследовательские институты и фармкомпании России и стран СНГ, такие как Институт общей генетики им. Вавилова, Институт цитологии и генетики РАН, Московская медицинская академия им. И. М. Сеченова и др.

Благодаря доступности передовых платформ для секвенирования, Genotek предлагает широкий спектр возможностей для Ваших исследований. А значительный объем заказов, получаемых компанией, позволяет нам оптимально использовать производственные мощности оборудования, чтобы всегда предлагать Вам самые низкие цены.

Подробнее об услугах: www.genotek.ru/R&D

Phone: +7(495) 767 85 95 или +7 (985) 411 64 07

Email: info@genotek.ru или ok@genotek.ru

ООО «Био-Рад Лаборатории»
117105, Москва, Варшавское шоссе, д.9 стр. 1Б,
Офисный комплекс Loft - квартал "ДМ - 1867"
ТЕЛ: (495) 721-1404
ФАКС: (495) 721-1412
post@bio-rad.com
www.bio-rad.com

Компания **Bio-Rad Laboratories, Inc USA** (Био-Рад, США) является одним из мировых лидеров производства оборудования и реагентов для научных исследований. В рамках взаимодействия с научными, медицинскими, биотехнологическими и образовательными организациями Био-Рад предлагает современные технологии, оборудование и реагенты.

Геномные технологии *(генная экспрессия и генная модуляция)*

- **Амплификация** *(уникальный спектр приборов)*
- **Цифровой капельный ПЦР третьего поколения**
- **Гельэлектрофорез** *(горизонтальный и вертикальный форматы)*
- **Experion™** *(автоматизированная система капиллярного электрофореза нуклеиновых кислот и белков)*
- **Системы визуализации** *(колориметрия, флюоресценция, хемилюминесценция, хемифлюоресценция, радиоизотопные метки)*
- **Перенос генов** *(электропорация, болистика, химическая трансфекция)*

Протеомные технологии *(структурная и функциональная протеомика)*

- **BioLogic DuoFlow™** *(модульная гибкая система для биохроматографии)*
- Широкий спектр колонок и носителей
- **NGC ™** *(автоматизированная хроматографическая система очистки рекомбинантных белков)*
- **Rotofor™** (аналитический и препаративный изоэлектрофокус)
- Оборудование для анализа и процессинга 2-D протеомных карт
- **Bio-Plex MagPix** *(мультиплексный количественный анализ биомолекул, панели для определения цитокинового профиля, белков сигнальной трансдукции, реагенты для создания собственных уникальных наборов)*
- **ProteOn XPR36™ Protein Interaction Array System** (*матричный интерактивный анализ биомолекулярных взаимодействий методом поверхностного плазмодинного резонанса*)
- **S3 сортер клеток**

Обзор инструментов
http://software.intel.com/en-us/intel-sdp-home/

Пакеты XE Suites

Intel® Parallel Studio XE
http://software.intel.com/en-us/articles/intel-parallel-studio-xe/

Intel® Cluster Studio XE
http://software.intel.com/en-us/articles/intel-cluster-studio-xe/

Статьи и материалы для самостоятельного изучения

Технические вебинары по программным инструментам
http://software.intel.com/en-us/articles/intel-software-tools-technical-webinar-series

Документация
http://software.intel.com/en-us/articles/intel-software-product-tutorials/#studioxe

Сертификационная программа
http://www.intuit.ru/academies/companies/972/info

Учебная лаборатория
http://software.intel.com/en-us/articles/intel-learning-lab/

Intel Premier Support
http://software.intel.com/en-us/articles/performance-tools-for-software-developers-intel-premier-support/

ООО «Рош Диагностика Рус»

Оборудование и реагенты для молекулярно-биологических, геномных и клеточных исследований

Компания предлагает оборудование и реагенты для лабораторий, работающих в области молекулярной и клеточной биологии, а также молекулярной диагностики:

- системы для секвенирования нового поколения 454
- системы для ПЦР в реальном времени LightCycler
- системы экстракции нуклеиновых кислот MagNAPure
- ДНК-чипы для обогащения образцов перед секвенированием нового поколения NimbleGen Sequence Capture
- клеточные анализаторы Innovatis
- высококачественные наборы реагентов

Компания предлагает покупателю не просто свою продукцию, а единую систему, включающую технический сервис, обучение персонала, постоянную методическую поддержку, своевременную доставку реактивов и расходных материалов.

ООО «Рош Диагностика Рус»
Россия, 115114, Москва
Летниковская ул., д.2/2
Бизнес-центр «Вивальди Плаза»
Тел.: +7-495-229-69-99
Факс: +7-495-229-62-64
ras.russia@roche.com
www.roche-applied-science.ru
www.roche-applied-science.com

Россия, 630132, Новосибирск
Ул.Красноярская, д.35, офис 103
Бизнес-центр «Гринвич»
Тел./Факс: +7-383-335-88-04/03
e-mail: olga.yatsenko@roche.com

ИнтерЛабСервис

передовые технологии
молекулярной диагностики

Общество с ограниченной ответственностью «ИнтерЛабСервис»

Страна: Россия
Индекс: 115035
Адрес фактический: г. Москва,
ул. Садовническая, д. 20/13, стр. 2
Телефон: (495) 664-28-84,
Факс: (495) 664-28-89
E-mail: info@interlabservice.ru
Сайт www.interlabservice.ru

Основные направления деятельности компании:

Оснащение клинико-диагностических, научно-исследовательских и экспертных лабораторий современным оборудованием для ПЦР-диагностики и молекулярно-биологических исследований, в том числе для выявления и характеристики патогенов, медицинской генетики, полногеномного секвенирования, изготовления и анализа биочипов.

Поставки широкого спектра наборов реагентов и высококачественных расходных материалов.

Разработка комплексных решений для автоматизации ПЦР-диагностики.

Обучение, сервисное обслуживание и клиентская поддержка.

Научное издание

Тезисы 6-ой Международной Школы молодых ученых «Системная биология и биоинформатика», SBB'2014

на английском языке

Abstracts of the 6th International Young Scientists School "Bioinformatics and Systems Biology"

Printed without editing

Подготовлено к печати
в редакционно-издательском отделе
ИЦиГ СО РАН
630090, Новосибирск, пр. акад. М.А. Лаврентьева, 10

Дизайн и компьютерная верстка: А.В. Харкевич

Подписано к печати 17. 06. 2014 г.
Формат бумаги 70 × 108 1/16. Печ. л. 4,2. Уч.-изд. л. 4,9
Тираж 80

Отпечатано на полиграфической базе ИЦиГ СО РАН