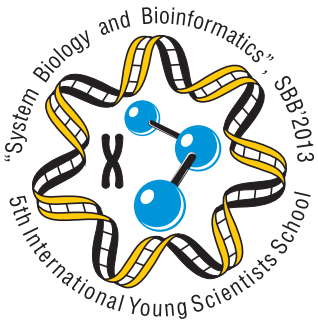
RUSSIAN ACADEMY OF SCIENCES
SIBERIAN BRANCH

INSTITUTE OF CYTOLOGY AND GENETICS

5th International Young Scientists School

"SYSTEMS BIOLOGY & BIOINFORMATICS"



Program
&
Abstracts

Novosibirsk, Russia
June 23–26, 2013

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. Konstantin V. Anokhin corresponding member of the RAS and RAMS, NRC "Kurchatov Institute"; P.K.Anokhin Institute of Normal Physiology of Russian Academy of Medical Sciences

Prof. Sergei L. Kiselev N.I. Vavilov Institute of General Genetics, Moscow, Russia

Prof. Tatiana V. Chernigovskaya St.Petersburg State University, NRC "Kurchatov Institute", Moscow, Russia

Prof. Boris M. Velichkovskiy corresponding member of the RAS, NRC "Kurchatov Institute"; Institute of Psychology, Unit of Engineering Psychology and Cognitive Ergonomics, Dresden University of Technology, Germany

Prof. Elza K. Khusnutdinova Institute of Biochemistry and Genetics, Ufa Research Center RAS, Ufa, Russia

Prof. Maria G.Samsonova St.Petersburg State Polytechnic University, St.Petersburg, Russia

Prof. Evgenii E. Vityaev Sobolev Institute of Mathematics, Novosibirsk, 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. Alexander S. Ratushnyak Design Technological Institute of Digital Techniques 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. Sergei V. Netesov corresponding member of the RAS, Novosibirsk State University, the State Research Center of Virology and Biotechnology "Vector", Koltsovo, Novosibirsk, Russia

Prof. Limsoon Wong National University of Singapore, Singapore

Prof. Ming Chen Zhejiang University, Hangzhou, China

Prof. Haihua Bai Director of Human Genetics & Medical Genetics Research Institute, Inner Mongolia University for the Nationalities, China

Prof. Vyacheslav I. Molodin academician RAS, Institute of Archeology and Ethnography SB RAS, Novosibirsk, Russia

Prof. Anatoly P. Derevyanko academician RAS, Institute of Archeology and Ethnography SB RAS, Novosibirsk, Russia

LOCAL ORGANIZING COMMITTEE

Ilya Akberdin, PhD
Dmitry Afonnikov, PhD
Nadejda Glebova,
Tatyana Karamysheva, PhD
Galina Kiseleva, PhD
Andrey Kharkevich,
Sergey Lashin, PhD
Anna Onchukova,
Alex Semenichev
Yuriy Orlov, PhD
Erlan Tokpanov,
Svetlana Zubova (chair)

CONTENTS

BINDING OF miR396 FAMILY WITH mRNA OF GROWTH-REGULATING FACTORS IN RICE AND MAIZE

Bari A.*, Pinsky I., Sagaidak A.
Al-Farabi Kazakh National University, Almaty, Kazakhstan
e-mail: assyl.bari@gmail.com
** Corresponding author*

Key words: *microRNA, binding site, growth-regulating factors, plant*

Motivation and Aim: microRNAs (miRNAs) are involved in plant growth and development by regulating post-transcriptional gene expression. miR396 family directly regulates growth processes in plants via targeting growth-regulating factor (*GRF*) genes family. It is important to identify the interaction characteristics of miR396 family with mRNA of *GRF* genes in rice and maize.

Methods and Algorithms: Gene nucleotide sequences of *Oryza sativa* and *Zea mays* were obtained from GenBank (http://www.ncbi.nlm.nih.gov). miRNAs nucleotide sequences were received from miRBase (http://www.mirbase.org). The free energy (ΔG) of hybridization of miRNA and mRNA, the position of potential binding sites, and the interaction schemes were calculated by using the RNAHybrid 2.1 software (http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/). The E-RNAhybrid software (http://sites.google.com/site/malaheenee/software/) was used to compute the ΔG/ΔGm value and p-value. The ΔG/ΔGm value was used as comparative criterion of the miRNA and mRNA interaction force.

Results: We found that among 661 miRNAs of *O. sativa* only osa-miR396a-i are shown to have strong binding sites with mRNAs of nine *GRF* genes. These miRNAs bound with mRNAs of *GRF* genes with various degree of prediction reliability and distributed into five groups: miR396a,b; miR396c; miR396d,e; miR396f; and miR396g,h,i. The ΔG/ΔGm value for miR396 binding sites in mRNAs of *Os02g0701300, Os06g0116200, Os02g0776900, Os03g0729500, Os02g0678800, Os03g0674700, Os11g0551900, Os04g0600900* and *Os12g0484900* genes ranged from 75.9% to 100% of the maximum free energy, which indicates a strong interaction of these miRNAs with the mRNA of the *GRF* gene family. osa-miR396a-i binding sites in mRNA of these *GRF* genes are located in the protein-coding sequence, are highly homologous and encode the same RSRKHVE heptapeptide. We have also investigated miRNAs and *GRF* genes as their targets in maize. Our results reveal that among 321 miRNAs of *Z. mays* only zma-miR396 family bind to the mRNA of 36 *GRF* genes with various degree of prediction reliability. By their different binding ability zma-miR396 family may be distribute into five groups: miR396a,b; miR396c; miR396d,e; miR396f; miR396g,h. zma-miR396 binding sites in mRNA of *GRF* genes in *Z. mays* are located in the protein-coding sequence, are highly homologous and encode RSRKHVE heptapeptide. In all cases, the free energy of zma-miR396a-h binding with mRNAs of 36 *GRF* family members in maize varied from 78.8% to 100% of the maximum free energy, which shows a high rate of interaction. All paralogs of *GRF* genes in *Z. mays* bind all members of miR396 family.

Conclusion: Our results show that miR396 binding sites are located in the protein-coding sequence of the mRNA, and they are highly conserved in rice and maize. We suggest the expression of growth-regulating factors, that affect the productivity of rice and maize, is under strong miR396 family control.

MicroRNAs BINDING SITES IN mRNAs OF SOME ONCOGENES

Berillo O.*, Moldagalieva A., Baidildinova G.
National Nanotechnology Laboratory, al-Farabi Kazakh National University, Almaty, Kazakhstan
e-mail: devolia18@mail.ru
* *Corresponding author*

Key words: *binding site, cancer, miRNA, mRNA, oncogene*

Motivation and Aim: MicroRNAs (miRNAs) are small noncoding RNAs. They bind with mRNA of target genes and suppress their translation. MiRNAs participate in different cell processes: apoptosis, cell cycle, differentiation, *etc.* [1] Recently it was showed that miRNA binding sites are located not only in 3'UTRs, but in the 5'UTRs and CDSs of mRNAs [2]. The expression of many oncogenes increases in during of tumourigenesis [3] and some of them were revealed as important participants of malignant transformation [4]. MiRNAs are potential noninvasive biomarkers, because they are stable molecules in body fluids (serum, plasma, *etc.*) and there are their concentrations changes in depend on cancer [5]. MiRNA dysregulation in control of oncogene expression may be a cause of cancer development (breast cancer, lung cancer, *etc.*). Data on the locations of miRNA binding sites could improve understanding of the interactions between miRNAs and mRNAs of oncogenes.

Methods and Algorithms: The nucleotide sequences of mRNAs were downloaded from Genbank. MiRNA nucleotide sequences were obtained from the miRBase. Schemes of binding sites were found, and free energies of miRNA interactions (ΔG) was calculated using the RNAHybrid 2.1 program. The $\Delta G/\Delta G_m$ ratio, significance (p), and mRNA regions were found using the E-RNAhybrid script. The $\Delta G/\Delta G_m$ value (%), where ΔG_m equals the miRNA binding energy with a perfectly complementary nucleotide sequence, was calculated.

Results: The binding sites of 2037 miRNAs were investigated within 474 mRNAs of oncogenes. These miRNAs accomplish a tumour-suppressor function, because they may decrease expression of oncogenes. MiRNA binding sites with $\Delta G/\Delta G_m$ values equaled greater than 90% were revealed in 3'UTRs (55.37%), CDSs (23.73%), and 5'UTRs (20.90%) of mRNAs. It was predicted, that 177 binding sites were formed between 83 miRNAs and 128 mRNAs: 45 binding sites were formed by 26 intergenic miRNAs with 43 mRNAs; 136 binding sites were formed by 58 intragenic miRNAs (ing-miRNAs) with 96 mRNAs. MiR-1285-3p encoded in intron of *KRIT1* gene and in intergenic region. Ing-miRNAs are coded in introns, exons or untranslated regions (5'UTR, 3'UTR) of 52 genes (host genes). There are connections between host genes and target-genes via these ing-miRNAs. Changes in host gene expression lead to changes in ing-miRNA expression, which influence to the mRNA translations of target-genes. The expression of 96 target-genes depend on expression of 52 host genes, which proteins are not oncogenes. Some genes had several miRNA binding sites. For example, *RABL5* mRNA had miR-4452, miR-5096 (2 sites), miR-548d-5p, and miR-5585-3p binding sites. It was predicted that single miRNA as miR-3656 regulate translation of mRNAs of *ATF1, CEBPB, FASN, GRINA*, and *HRAS* genes.

Conclusion and Availability: As a result, it was predicted, that 128 mRNAs of oncogenes had 117 miRNA binding sites with $\Delta G/\Delta G_m$ value equaled more than 90%. Revealed data about host genes, miRNAs and their binding sites with target mRNAs are important for using in biomedicine.

References:

1. J. Krol, W.J. Krzyzosiak. (2004) Structural aspects of microRNA biogenesis, *IUBMB Life,* **56(2):** 95-100.
2. X. Zhou et al. (2009) Abundant conserved microRNA target sites in the 5'-untranslated region and coding sequence, *Genetica,* **137(2):** *159-164.*
3. F.H. Shen et al. (2013) The G-quadruplex ligand, SYUIQ-FM05, targets proto-oncogene c-kit transcription and induces apoptosis in K562 cells, *Pharm Biol,* **51(4):** 447-454.
4. N.O. Chimge, B. Frenkel. (2013) The *RUNX* family in breast cancer: relationships with estrogen signaling, *Oncogene,* **32(17):** 2121-2130.
5. M.A. Cortez et al. (2012) Circulating microRNAs as noninvasive biomarkers in breast cancer, *Recent Results Cancer Res.,* **195:**151-61.

EXPERIMENTAL AND THEORETICAL STUDY OF WATER AND ELECTROLYTE BALANCE OF RENAL COLLECTING DUCT PRINCIPAL CELLS

Ilyaskin A.V.*
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
e-mail: ilyaskin@bionet.nsc.ru
**Corresponding author*

Key words: *collecting duct cell, biophysical model, ion and water membrane permeability*

Motivation and Aim: Kidney collecting duct principal cells play the key role in regulated tubular water and sodium reabsorption and potassium secretion. The aim of the current work was the investigation of ion and water transport processes across plasma membrane of principal cells of rat OMCD.

Methods and Algorithms: The results of investigation of transport processes through these cells are distorted due to the invasive nature of the experimental protocols used. In recent years fluorescent indicators proved to be an efficient approach to study intact cells. A wide spectrum of fluorescent probes allows one to measure cell volume, intracellular concentration of potassium, sodium and chloride ions and thus could be used to investigate transmembrane ion and water fluxes. One shortcoming of the fluorescent methods is that sophisticated analytical tool is needed to interpret the data and to obtain quantitative estimates of cell physiological characteristics. We used appropriate fluorescent indicators to measure the dynamics of cell volume in hypotonic or Cl-free medium and dynamics of intracellular concentration of sodium ions in low-sodium medium with sub-second time resolution. Experimental data were analyzed by the biophysical model of membrane transport.

Results: The combination of experimental approaches and modeling allowed us to obtain quantitative estimates of plasma membrane permeabilities for water, sodium, potassium and chloride ions. Also fluorescent dyes were applied for identification of membrane permeabilities of the same cell type isolated from rats subjected to the chronic salt loading. It was shown that increased NaCl intake caused more than 3 fold decrease of both potassium and sodium permeabilities.

Conclusion: To conclude, the current study reveals that the fluorescent indicators in combination with the appropriate mathematical model could be used to study transmembrane water and ion transport trough intact epithelial cells in dynamically changing extracellular environment.

Acknowledgements: This work was supported by RFBR grants #12-04-00369-a, 12-04-00370-a and 12-04-31298.

DIFFERENTIAL EXPRESSION ANALYSIS OF microRNAS IN THE ROOT OF *M. truncatula* WHEN SYMBIOSIS WITH *S. meliloti* AND *G. intraradices*

Danfeng Jin[1, 2], Yuhua Zhao[2], Ming Chen*[1]
[1] *Department of Bioinformatics, College of Life Sciences, Zhejiang University, Hangzhou 310058, China;*
[2] *Institute of Microbiology, Zhejiang University, Hangzhou 310058, China*
e-mail: mchen@zju.edu.cn
* *Corresponding author*

Key words: *microRNA, Medicago truncatula, symbiosis, root*

Motivation and Aim: Legumes and many nonleguminous plants enter symbiosis with microbes, forming special nodule or arbuscular mycorrhizas (AM), to convert atmospheric nitrogen to ammonia and utilize it as a nitrogen source. Nitrogen-fixing bacteria, such as *Rhizobium leguminosarum*, *Sinorhizobium meliloti*, and *Bradyrhizobium japonicum*, can form endosymbiotic nodules association with roots of legumes. Some fungi, named as arbuscular mycorrhizal fungi (AM fungi) penetrate the cortical cells of the roots of vascular plants and form AM.

The symbiotic interactions are regulated by some small RNAs (sRNAs) such as microRNA (miRNA). miR166, miR169, miR396, miR482 and miR1512 in plants have been found to play important roles in root nodule development and nodule numbers control. miR399 also play a role in the regulation of cellular response to local stress during AM symbiosis.

Our research is to analyze the differential expression of miRNAs in the root of *Medicago truncatula* when it form nodules and AM with different microbes.

Methods: Small RNA sequencing was data from Gene Expression Omnibus (GEO). Tools and software such as Python, fastx, Bowtie2 and BLASTN were used in our work.

Results: We used the data to analyze the differential expression of miRNAs when *M.truncatula* symbiosis with *S. meliloti* (in the form of nodule) and *Glomus intraradices* (forming AM), respectively. In both condition, miR1509, miR2118 and miR2597 showed higher abundances than the mature miRNAs annotated in miRBase. miR393 had a low abundance in nodule, while its mature sequence was missing in AM.

REGULATION OF MITOCHONDRIAL BIOGENESIS IN HUMAN EMBRYONIC STEM CELLS

Kao L.P. *[1], Ovchinnikov D.A.[1], Prescott M.[2], Tra T.[3], Turner J.P.[1], Cooper-White J.J.[1], Wolvetang E.J.*[1]

Institutions: [1]Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, St Lucia, Australia, [2] Department of Biochemistry & Molecular Biology, and [3] ARC Centre of Excellence in Structural and Functional Genomics, Monash University, Clayton Campus, Victoria 3800.
e-mail: kaolp@yahoo.com.tw, and e.wolvetang@uq.edu.au

** Corresponding authors*

Key words: *Mitochondrial Biogenesis, Mitochondrial Rho Minus Cells, Mitophagy, Human Embryonic Stem cells, Human Foreskin Fibroblast cells, Transcription Factors*

Motivation and Aim: The mitochondrion plays an important role in human life for example in mitochondrial biogenesis, and in mitophagy. The number of mitochondria is very sensitive to the energy requirement of cell types, and to environmental stress. Environmental stress can involve altered efficiency and/or fidelity of mtDNA synthesis or damage from overproduction of reactive oxygen species as a result of mitochondrial dysfunction. The number and quality of mitochondria in a cell are determined by the balance between mitochondrial biogenesis, and mitochondrial turnover. Mitochondrial turnover, mitophagy, regulates the quantity and quality of mitochondria. The number of mitochondria can be associated with disorders of the twenty-first century such as fertilisation dysfunction, and Mitochondrial Depleted Syndrome (MDS).

Methods and Results: Firstly, we established a mitochondrial minus cell line in order to compare the metabolic, mitochondrial and nuclear responses to impaired mitochondrial protein synthesis, using chloramphenicol, or DNA replication, using EtBr, and assessed these parameters both during treatment and after a two week recovery period. Our data show that both during treatment and during the recovery phase EtBr and CAP impaired mitochondrial function in primary human fibroblasts to differing extents. This was accompanied by differential expression of both mtDNA-encoded genes and nuclear transcription factors that control mitochondrial biogenesis.

Next, we used human embryonic stem cells (hESCs) as a tool to understand mitochondrial biogenesis during development, and to examine whether how damage to mitochondria causes mitophagy in these cells. Mitophagy is a specific form of autophagy whereby the mitochondrion destined for degradation is sequestered within a vesicle with a membrane marked by Light Chain protein III (LC3). The autophagosome subsequently fuses with lysosomes to engulf the damaged mitochondria. We show that mitophagy can be monitored in hESCs by measuring co-localisation of Mitotracker red and LC3-GFP labelled vesicles. Using this assay we show increased levels of mitophagy in hESCs treated with rapamycin (an mTOR inhibitor), 500ng/ml EtBr (a mitochondrial DNA damaging agent), and CCCP (mitochondrial membrane depolarizer).

We further discovered that differentiated hESCs show more differentiated mitochondria with mature transverse cristae, a greater number of mitochondria accompanied by elevations in expression levels of mitochondrial biogenesis-related genes, mitochondrial membrane potential and mitochondrial ATP synthesis, and increased mitophagy events.

Conclusions: Our results suggest that mitophagy is involved in the selective turnover and quality control of mitochondria in hESC during or prior to differentiation. A better understanding of the molecular mechanisms that control mitopahgy in hESCs will allow this process to be manipulated. An understanding of mitochondrial quality in long term cultured hESC or hiPS cells will benefit their use in future regenerative medicine approaches and is the future focus of our research.

MOLECULAR FUNCTIONAL SYSTEMS OF COGNITIVE PROCESSES NEURON

Ratushnyak A.S.*, Proskura A.L., Zapara T.A.
Design Technological Institute of Digital Techniques, SB RAS, Novosibirsk, Russia
e-mail: ratushniak.alex@gmail.com
** Corresponding authors*

Key words: *neuron, cognitive processes, learning, memory, protein networks, functional system*

Motivation and Aim: The actual direction of research and development in the field of neuroscience is the study of brain functioning principles, creating, on this basis, the new generations of information devices, biotechnology interfaces, biosensors, hardware and software aids for diagnostic, prevention and correction of neuropathology, target personalized medicine and screening of pharmacologically advanced compounds. The shortage of fundamental progress in these studies is first of all conditioned by the lack of conceptual understanding of the principles and mechanisms of functioning of the biological prototypes of such systems. The analysis and modeling of cognitive systems requires the development of the structure functional model of the basic element - the neuron. Owing to the latest investigations it became clear that the complication of molecular cellular organization underlies the evolutionary development of a brain [1; 2; 3]. The complexity of cell organization allows us to consider it as a multilevel functional system. The concept of functional systems was first proposed by Anokhin [4]. The functional systems emerge for integrative activity and maintenance of adaptive parameters within a physiological range. The functional systems are governed by the set of motivations and genetic programs adapted by learning to the conditions of environment. The functional systems within a neuron are the molecular networks of compartments performing a specific cellular function. The interaction of these systems leads to the creation of new integrative properties that are not the part of the individual components, the so-called emergent-systemic properties and functions. These newly emerged properties allow the nerve cells to function as the complex of molecular information systems that underlie cognitive processes.

A large amount of data about the intracellular processes involved in learning and memory leads to the fact that the reconstitution of a neuron complete structure scheme is not feasible now. The perception (by investigator) of the complex processes that occur even in separate neuronal compartments without special tools is a difficult, if at all possible, task. The problem lies in the organization easy access to the acquired knowledge and visualization – the representation of these complex processes in the form of structural schemes, maps (entity-relationship diagram). One of the tools that can facilitate the analysis of complex systems with many interacting elements is the technology, which combines the capability to store data in a database with visualize the data in the form of diagrams including further using for developing the simulation models.

Methods and Algorithms: The methods of theoretical and experimental analysis were used for systematization of literary and experimental data and develop of the conceptual structure-functional circuit of neuron. To reconstruct the protein networks underlying the functioning of one of the elements of nerve cells - a synapse, we used the technology GeneNet [5]. In the diagram the network entities and relations between them are provided with links to the published literature data, as well as to the database Swiss-Prot, EMBL, PubMed, MGI, GeneCard, TRRD.

Results: As a result of analysis it were developed the general conceptual hypothetical

schemes of a neuron as a cognitive functional system. The main structure-functional blocks of the neuron it is a set of molecular systems performing motivational subprograms and providing the main task of the cell - the maintenance of homeostasis. The blocks of base motivations, are equipped with internal molecular memory elements and algorithms of parameter stabilization. It were separated blocks of conversion and input of information, blocks performing the comparison of external signals with recorded, blocks of decision-making and blocks of executive elements. Molecular mechanisms and principles of functioning the blocks of conversion and input of information in a synapse were examined in more detail.

Dendritic spines are the postsynaptic component of a synapse. It protrusions less than 1 μm^3 and originates from an axial dendrite. Spines are the structure of the excitatory synapse majority of a mammal brain. The network of dendritic spine proteins in the CA1 hippocampal region of rodent have been reconstructed (http://wwwmgs.bionet.nsc.ru/mgs/gnw/genenet/viewer/Early long-term potenti-ation.html). This network reflects the elements of a real network, which provides increasing the amplitude of a neuronal response to an intense and short-term release of neurotransmitter. This pattern of glutamate receptor (NMDA, AMPA type) activation leads to transition of a spine to the level, corresponding to more efficient synaptic transmission. The network reflects the ability of nanosized compartments of a neuron to self-development (the transition to the new level of efficient synaptic transmission) using only own resources. However, for maintaining the new state of a spine the resources from other compartments of a neuron are needed. Maintaining the new level of transfer is accompanied by replacement of AMPA1/2 receptors on AMPA2/3 subtype. The molecular network, mediating these processes is available at: http://wwwmgs.bionet.nsc.ru/mgs/gnw/genenet/viewer/AMPA.html.

Conclusion: Analysis of a cell as a multilevel functional system component that interacts in a wide range of time and space allows us to conceptualize the processes of the origination of integrative properties of nerve cells associated with the concepts of cognition. The protein-protein network reflects the molecular functional system of a spine in a state of «readiness» perception and action for change of synaptic connectivity using its own resources. Based on the conceptual model of a neuron that takes into account the basic properties of simulation systems development it may be revealed the functional capabilities of such systems in their biologic prototypes, thus creating the opportunity for medicine to correct the pathologic states involving the cognitive impairment. These findings require further theoretical and experimental analysis and the development of additional conceptual models of biologic information systems.

*Acknowledgements:*The work was supported by RAS base fundamental research project 35.1.5, Integration project presidium SB RAS 136, RFBR grant 12-01-00639 a.

References:

1. L.Lin et al. (2010) Evolution of alternative splicing in primate brain transcriptomes, Hum. Mol. Genet. 19: 2958-2973.
2. C.P.Andrew et al. (2012) Activity-mediated AMPA receptor remodeling, driven by alternative splicing in the ligand-binding domain, Neuron, 76: 503–510.
3. N.L.Barbosa-Morais et al. (2012) The evolutionary landscape of alternative splicing in vertebrate species. Science, 338:1587-1593.
4. P.K.Anokhin, (1974) Systems analysis of the integrative activity of the neuron. Pavlov J. Biol. Sci. 19: 43–101.
5. F.A.Kolpakov, E.A.Ananko (1999) Interactive data input into the GeneNet database. Bioinformatics 15: 713–714.

RECOGNITION OF LIVER PROMOTERS WITH EXPERT DISCOVERY AND UGENE INTEGRATED SYSTEM

Vaskin Y.Y.*[1], Khomicheva I.V. [2], Vityaev E.E.[2]
[1] *Novosibirsk State University, "UniPro" Company, Novosibirsk, Russia;*
[2] *Istitute of Mathematics SB RAS, Novosisbrsk, Russia*
e-mail: vaskin90@gmail.com
* *Corresponding author*

Key words: *Complex signals, relation data mining, bioinformatics, UGENE, recognition, gene regulatory regions*

Motivation and Aim: Rapid development in the field of sequencing technology caused accumulation of huge amount of data. To analyses this data effectively, we need to use modern bioinformatics tools. The ExpertDiscovery system is designed to analyze one of the most important parts of eukaryotic genome – gene regulatory regions. The task of extraction of the hierarchical structure of the regulatory regions is an actual and poorly studied problem.

The first goal of the project is development of the scope of the ExpertDiscovery system for more convenient, fast and efficient recognition of the complex signals. Due to this development, experts will have a tool, which is in the context of interacting modules of UGENE. The second goal of the project is a using and presentation of the system on real data.

Methods and Algorithms: The ExpertDiscovery system which is integrated into UGENE package performs data mining of sequences using semantic probabilistic inference based on a training set. Rules, which are discovered by the system, are complex signals that reflect the hierarchical structure of the region. Then we can recognize signals on any sequence, taking into account different statistics which are displayed as a few numerical parameters and graphs. The key feature of the system is a possibility of combining of results, which are obtained with other methods, to extract more complex rules.

Results: A convenient tool for analysis of regulatory regions has been integrated into UGENE package, providing a good possibility of automation of expert's work. All the operations are performed within the environment of one tool which speeds up productivity of work.

Conclusion: An original approach to data integration of different methods and knowledge of experts was developed. During the study, significant rules, which cannot be extracted automatically with other similar methods, have been found.

Availability: www.ugene.unipro.ru. Free open-source software. GPLv2 license.

AN EVALUATION OF THE SOFTWARE TOOLS FOR PLANTS microRNA DEEP-SEQUENCING DATA ANALYSIS

Yue Wang, Ming Chen*
Department of Bioinformatics, College of Life Sciences, Zhejiang University, Hangzhou 310058, China.
e-mail: mchen@zju.edu.cn
** Corresponding authors: Ming Chen*

Key words: *deep-sequencing data; software tools; miRNA prediction*

MiRNA is a class of small RNA, which plays important roles in gene regulation. Even though both plants and animals miRNA provide essential regulation of some cellular processes and are excised from endogenous hairpin pre-miRNAs, the biogenesis of miRNA among the two kingdoms is different.

As the next-generation sequencing (NGS) techniques develop, some software tools emerge to detect miRNA in deep-sequencing data, which are based on the biogenesis of mature miRNA or the homology of the same family members among different organisms. At the beginning, several tools were developed to detect animal miRNA. Then some groups update their softwares to be used both in animals and plants, such as miRanalyzer-2. And some groups develop some special tools to analyze plants miRNA deep-sequencing data, such as miRDeep-P, mirdeepFinder.

Although deep-sequencing is popular for detecting miRNA, the data should be viewed carefully. Last year, two groups respectively evaluated some tools in animals. An overall evaluation of these tools in plants is lacking, although there have been several tools for plant miRNA deep-sequencing data analysis. Here, we evaluate six softerware tools in plants by using both simulated data and experimental data. We hope to provide useful information for researchers to facilitate their selection of tools for miRNA analysis in plant kingdom.

ANALYSIS OF PHASE PORTRAITS IN SOME GENE NETWORKS MODELS

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

Key words: *gene networks models, dynamical systems, equilibrium points, periodic trajectories, numerical modeling, Andronov-Hopf bifurcation, first Lyapunov coefficient*

Motivation and Aim: learning the dynamic behavior of some gene networks models, developing a special software for building the interactive visualizations of corresponding phase portraits, searching equilibrium points, modelling of behavior, various calculations.

Methods and Algorithms: chemical kinetic method.

Results: the analysis of special gene network models (symmetric system of differential equations with Hill's function, see [1,2]) were performed (include description of equilibrium points, Andronov-Hopf bifurcation cycles, stability question with first Lyapunov coefficient for delay-argument equation, see [3]). The results were verified by numerical experiments with developed software.

Fig. 1. The oscillation in 9-D model.

Fig. 2. The chaos in 5-D model.

Availability: You can download demo-version of the program from http://ppanalyzer.ru/.

Acknowledgements: the author expresses a great appreciation to V.P.Golubyatnikov and V.A.Likhoshvai for the valuable advices. The work was supported by the RFBR grant 12-01-00074, by the grant SP-561.2012.5 of the President of Russian Federation.

References:

1. V.A.Likhoshvai, V.P.Golubyatnikov et al. (2008) Theory of gene networks, In: *System computerized biology*, (N.A.Kolchanov, S.S.Goncharov), Novosibirsk, SB RAS, 397–480.
2. A.A.Akinshin, V.P.Golubyatnikov (2012) Geometric characterization of cycles in some symmetric dynamical systems, *Novosibirsk State University Bulletin, Mathematics Series*, **12(2):** 3–12.
3. A.V.Ion (2004) On the Bautin bifurcation for systems of delay differential equations, *Acta Univ. Apulensis*, **8:** 235–246.

ELECTROCARDIOGRAM RECOGNITION WITH WAVELET ANALYSIS

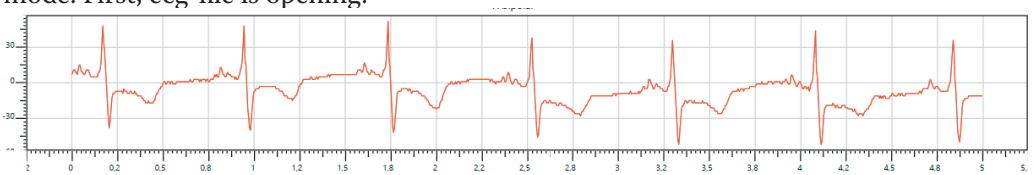
Bogdanov M.R.*, Zakharov A.V.
Bashkir State Pedagogical University, Ufa, Russia
e-mail: bogdanov_marat@mail.ru
** Corresponding author*

Key words: *Electrocardiogram, THEW, telemetric and Holter ECG warehouse Initiative, wavelet, diagnosis of heart disease*

Motivation and Aim: Cardiovascular diseases are the leading reason of disability and premature death in the population. Doctors traditionally make decision on patient diagnosis based on the waveform of electrocardiogram. We would like to use wavelet analysis for ecg-signals recognition. We designed a software prototype, which recognize ecg-signals very well, and we are going to develop an ecg diagnostic system as telemedicine service.

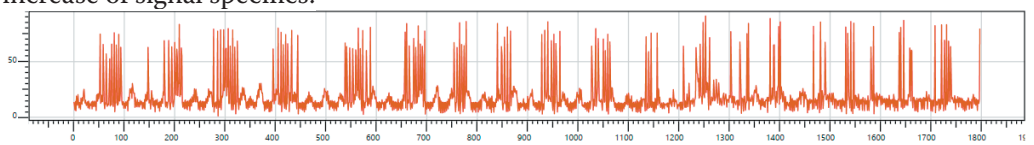
Methods and Algorithms: Wavelet is a function, given by numerical array, which can characterize a signal. We think traditional diagnostic methods have some disadvantages. In particular, modern electrocardiographs generate relatively high-frequented signal, having a lot of information, and a doctor operates on a smooth curve obtained by the recorder. Thus, in our view part of the valuable information is lost. Wavelet analysis, in our opinion, makes better use of information hidden in the signal; it can identify the signal features that are not available with other methods. We used ecg-files in THEW form (three leads; 200 Hz, 24 hour).

The software for ecg recognition include wavelet designing mode and ecg-recognition mode. First, ecg-file is opening.



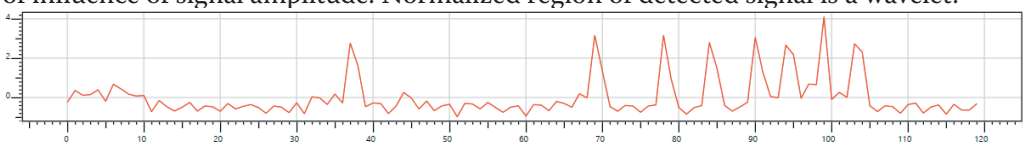
The electrocardiogram of coronary artery disease patient

Then we have to do detecting of ecg-signal to eliminate of signal redundancy and increase of signal specifics.



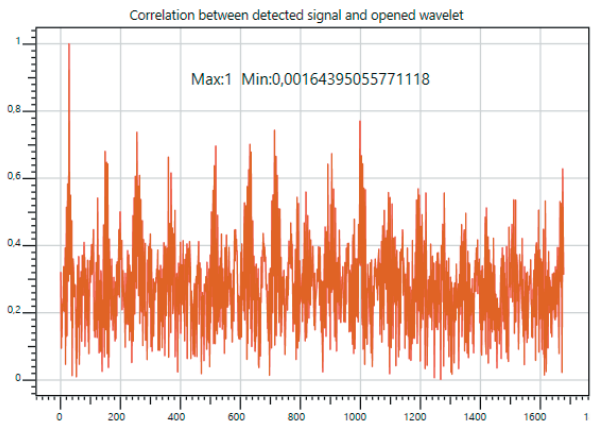
The detected signal of coronary artery disease patient

Then we have to find of specific region of detected signal and normalize it to eliminate of influence of signal amplitude. Normalized region of detected signal is a wavelet.



The wavelet of coronary artery disease patient's electrocardiogram

Under recognition, the detection of signal is making with the same detecting frame, when creating wavelet. Then we have to choice a recognition step equal to wavelet size. We have to determine the correlation coefficient between the segment of testing signal and known wavelet within step of recognition. Finding the coefficients of correlation of the selected step, we shift the frame of recognition to a single element and resume processes. If the correlation coefficient is close to one, we can make a decision on the success of recognition.



The correlation function between ecg-signal and wavelet of CAD patient

Results: We designed a software for ecg-recognition with wavelet analysis. Our wavelets show high specificity to healthy heart, coronary artery disease and acute myocardial infarct.

Conclusion: For accurate medical diagnosis, we should include into consideration, all leads information. This time we are improving the software.

Availability: The proposed approach can be used as a telemedicine service.

References:

1. http://thew-project.org/database.html

LEARNING OF LOCOMOTION AND CHEMOTAXIS IN 3D MODEL OF THE *C. elegans* NEMATODE

Demin A.V.*[1], Vityaev E.E.[2]
[1] *A.P. Ershov Institute of Informatics Systems SB RAS, Novosibirsk, Russia;*
[2] *Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia*
e-mail: alexandredemin@yandex.ru
* *Corresponding author*

Key words: *adaptive behavior, control system, knowledge discovery*

Motivation and Aim: The C. elegans nematode is currently the only organism, for which the structure of its nervous system is fully known, and the connectome is obtained in the first approximation. However, we still have not understood the working principles of such a simple nervous system. In this paper, we propose a learning control system, which models the work of neural circuits controlling locomotion and chemotaxis of the C. elegans nematode.

Methods and Algorithms: For our experiments, we used a realistic 3D-simulator of the nematode with graphic interface [1]. The model of the worm's body inside the simulator is built according to the anatomic scheme (http://www.wormatlas.org), providing a sufficiently accurate approximation of the nematode's muscular system. A model of learning logical neurons has been proposed to create learning neural control circuits. The process of training neurons is based on the semantic probabilistic inference algorithm [2].

Results: Based on the proposed model, we have developed a neural control circuit for locomotion and a control circuit for chemotaxis. Using the 3D-simulator of the nematode, we have conducted a series of successful experiments in training the proposed model. The control system can stably learn an effective way of movement in 100 working cycles on the average, and identify an optimal chemotaxis strategy in 1000 cycles on the average. At the same time, we observe a considerable visual likeness between the behavior of the model and the behavior of a real nematode.

Conclusion: The results of experiments have shown that the movement function and associated orientation mechanisms of a nematode can be obtained by way of learning only in interaction with the environment. Practically, the results show that the proposed model can be successfully used to control complex objects.

Acknowledgements This research was supported by the Russian Science Foundation grant #11-07-0388-a, Integration project #15/10 of the RAS and Integration project #136 of the SB RAS

References:

1. A. Palyanov et al. (2012) Towards a virtual C. elegans: A framework for simulation and visualization of the neuromuscular system in a 3D physical environment, *Silico Biology*, **11 (3):** 137-147.
2. E.E.Viyaev, A.V.Demin (2011) Recursive subgoals discovery based on the Functional Systems Theory, *Biologically Inspired Cognitive Architectures 2011*: 425-430.

KiNET – A NEW WEB DATABASE ON KINETICS DATA AND PARAMETERS FOR *E. coli*

Ermak T.*, Akberdin I., Timonov V., Khlebodarova T. M., Likhoshvai V.A.
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
e-mail: timofei.ermak@live.com
** Corresponding author*

Key words: *kinetics data, parameters, E.coli, web database, biochemical reactions processes*

Motivation and Aim: Mathematical modeling of biochemical reactions has many applications in such modern scientific field as systems biology. For example, development of *in silico* cell or whole cell model (theoretical platform for performance of computational experiments) demands development of robust and verified mathematical models describing whole complex gene network of cell's functioning. In turn, it requires accumulation and analysis of published information and data in web-warehouses on operation of the cell as an integrated system.

Previously, the KiNET database was developed in the ICG SB RAS [1]. The KiNET is a database of kinetics data and parameters of biochemical processes for *E.coli.* The database accumulates unique published data collected in the Institute over seven years. However, the KiNET was not accessible as a web resource for wide range of scientists.

Methods and Algorithms: Web application was developed on Java platform using Vaadin (http://vaadin.com/) and Spring (http://www.springsource.org/) web frameworks. Software for data comparison was developed on .NET platform using C# and Entity Framework ORM-technology.

Results: To solve the problem with availability to the Kinet we have developed the web application that has user-friendly interface and easy access to unique data. In addition, to understand how our data is unique we have compared the KiNET data with information collected in another well-known database Sabio-RK (http://sabio.villa-bosch.de/). We have developed software, which imports data from Sabio-RK by using its web-services and compares data by substances that take part in the same reactions and by experimental conditions. As a result, we have obtained that only 2% of processes in two databases are identical.

Conclusion: We have presented unique kinetics data of biochemical reactions (conditions of experiments, values of kinetics parameters, concentrations of key metabolites, cell enzymes and other) as the web-source. The KiNET will be helpful in carrying out of large-scale *in silico* studies. The comparative analysis of the developed database with Sabio-RK has also shown that the KiNET's data is unique.

Availability: http://kinet.biomodelsgroup.ru/

Acknowledgements: This study was partially supported by the Scientific school № 5278.2012.4, RFBR grant (13-01-00344-a) and Program of the RAS "Molecular and cell biology" (6.6)

References:

1. Khlebodarova T.M *et al.*, Information sources for modeling *Escherichia coli* metabolic pathways// Proc. of the 5th BGRS International Conference, Vol. 2, pp. 19-24.

GENERATION OF WEB SERVICES - THE UNIFICATION OF ACCESS TO BIOINFORMATICS RESOURCES

Komyshev E.G.*[1, 2], Genaev M.A.[1], Afonnikov D.A.[1]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[2] *Novosibirsk State University, Novosibirsk, Russia*
e-mail: komyshev@bionet.nsc.ru
* *Corresponding author*

Key words: *generation, web-services, unification of access, pipelines*

Motivation and Aim: The analysis of biological data in bioinformatics usually consists of several steps performed by different programs subsequently. During the analysis progress, the output of one calculation module serves as an input of the other module, etc. The overall procedure could be organized as a workflow [1, 2]. For example, the calculation of the phylogenetic tree for protein family requires protein sequence extraction from databases, multiple sequence alignment, phylogeny estimation. Each step of process requires a certain input data. These data may be located in different stores, or may be the result of any of the software module. Thereby, there is the problem of producing a simple, standardized access to such heterogeneous bioinformatics resources. One possible solution to this problem is the implementation of web-services based on the resources that will provide programming interface for various bioinformatics resources. We develop a system which implements the automatic generation of web-services for such resources. These services can be used in various systems, such as Galaxy [1], Taverna [2], so the creation of web-services is an urgent task.

Methods and Algorithms: Generation of web-services is produced automatically by java application. Generation system gets the information about the resource for which the service is performed, for example, from meta-descriptions in an XML file. As a result is generated war-file that represents REST or SOAP / WSDL web-service for resource.

Results: At the moment the system generation of web-services is part of the BioinfoWF[3] system. The system of web-services generation includes the following elements: generation of web-services for computing modules described in a special meta-language; generation of web-services for pipelines of the BioinfoWF system; web-based interface to generate a descriptions of computational modules, descriptions of BioinfoWF pipelines and web-services. The developed system was used to generate interface for web-service to access the PeffDB database [http://pixie.bionet.nsc.ru/peff].

Conclusion: The system developed can facilitate collection of data needed for specialists in the bioinformatics area. Moreover using web-services simplify the task of its integration with third-party systems.

Availability: http://bioinfowf.bionet.nsc.ru

References:

1. J. Goecks, A. Nekrutenko, J. Taylor, Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences, Genome Biology (2010).
2. Oinn T., Addis M., Ferris J., Marvin D., Senger M., Greenwood M., Carver T., Glover K., Pocock M.R., Wipat A., Li P. (2004) Taverna: a tool for the composition and enactment of bioinformatics workflows, Bioinformatics, 20:3045-3054.
3. M.A. Genaev, K.V. Gunbin, D.A. Afonnikov. BioinfoWF: workflow management system for bioinformatics analysis.

NEURONAL LOSS IN SENESCENCE-ACCELERATED OXYS RATS HIPPOCAMPUS: HISTOLOGICAL XAMINATION

Maksimova K.Y.
Siberian State Medical University, Tomsk, Russia
e-mail: kseniya.maksimova.88@mail.ru

Key words: *aging, hippocampus, neuronal loss, OXYS rats*

Motivation and Aim. One of the major challenges in neurodegenerative research is modeling systemic aging. To date there is no perfect rodent model able to fully mimic human neurodegenerative diseases and particularly Alzheimer's disease (AD). The widely recognized genetic model of aging and aging-associated diseases is the SAMP8 mouse strain. But the mounting evidence suggests the utility of the senescence-accelerated OXYS rats as a rat model of AD-like pathology, and consequently this model is expected to provide more insight into the mechanisms of AD pathogenesis. The manifestations of accelerated brain aging in OXYS rats are behavioral alterations and learning and memory deficits that develop during the period from 4 to 12 weeks of age against the background of the first signs of neurodegeneration detected by magnetic resonance imaging. With age, such changes are amplified against pathological accumulation of beta-amyloid protein which results in the amyloid plaques, hyperphosphorylation of tau protein - markers of AD - in the cerebral cortex and hippocampus. However, not previously been evaluated the assessment of age-related neuronal loss in the hippocampus of OXYS rats. It is known that the degenerative changes in the neurons of CA1, CA3 pyramidal layer fields of the hippocampus and the dentate gyrus are caused by disturbance of the afferent projection pathways in the hippocampus originating from other brain regions. The aim of this work was the analysis of structural changes with age in the hippocampus of OXYS rats. Wistar rats were used as a control.

Methods. After dislocation, brains were dissected, fixed with 10% formalin, embedded in paraffin, and serially cut in coronal sections (5μm thickness). These sections were stained with Cresyl violet and examined with a photomicroscope (Carl Zeiss Axiostar plus, Germany). The total number of hippocampal pyramidal cells in the CA1, CA3 fields, and the dentate gyrus was estimated in 14-day-old, 5- and 15-month-old OXYS and Wistar rats (n=6) on the 5 slices of each brain sections. The number of neurons with chromatolysis, hyperchromatic with darkly stained cytoplasm and shrunken neurons were calculated as degenerative neurons.

Results. The histological examination have not revealed in the total number of hippocampal neurons in CA1, CA3 fields and the dentate gyrus of OXYS and Wistar rats during early postnatal development (age 14 day). This number increased by age 5 months and decreased with age in the hippocampus of both rat strains. The neurodegenerative processes in OXYS rats developed already at age 5 months compared with those in age-matched Wistar rats. There were significantly bigger the neurons with chromatolysis, hyperchromatic and shrunken neurons in the CA1, CA3 fields of the hippocampus and the dentate gyrus of 5- and 15-months-old OXYS rats, compared with those in age-matched Wistar rats. Also in OXYS rats, there were more neurons with nuclear pyknosis, which indicates decline of reserve capacity of neurons and is a sign of hypoxia.

Conclusion. Thus, the features of age-dependent neuronal loss in the hippocampus of OXYS rats identified by histological methods serves as proof of the processes that contribute to accelerated aging of the brain. This results support the utility of the OXYS strain as a rat model of neurodegenerative pathology, and consequently this model is expected to provide more insight into the mechanisms of brain aging.

MOLECULAR DYNAMICS SIMULATION ANALYSIS OF INFLUENCE OF HIGH TEMPERATURE AND HIGH PRESSURE ON THE NIP7 PROTEINS FROM THE HYPERTHERMOPHILIC ARCHAEA

Medvedev K.E.*1, Afonnikov D.A.[1, 2], Vorobjev Y.N.[3]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[2] *Novosibirsk State University, Novosibirsk, Russia;*
[3] *Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia*
e-mail: kirill-medvedev@yandex.ru
* *Corresponding author*

Key words: *Nip7 protein, high pressure, molecular dynamic simulation hyperthermophiles, water-protein interaction*

Motivation and Aim: High temperature and hydrostatic pressure are severe for living cells. However some organisms can survive and proliferate at deep-sea hydrothermal vents with high temperatures (above 100°C) and pressures (hundreds times greater than atmospheric). The molecular mechanisms of the adaptation of these organisms to extreme environments remain unclear. In this work we apply comparative analysis of the molecular dynamics simulation of Nip7 proteins from the hyperthermophilic archaea P. abyssi (deep sea habitat) and P. furiosus (shallow water habitat) species.

Methods and Algorithms: The structure of P.furiosus protein was obtained by homology modeling using Nip7 from P.abyssi as template (PDBID 2p38). MD simulations and structure analysis were performed using GROMACS [1].

Results: Obtained data demonstrated that the RNA-binding domain of the Nip7 protein is more flexible than N-terminal one. This flexibility may provide C-terminal domain functionality: non-specific binding of the poly-U/poly-AU RNA sequences. N-terminal domain demonstrates stable hydrophobic core and flexible loop regions. Additionally we found that segment of the Nip7 sequence 45-55 (helices α2-α3) demonstrate high structure fluctuation with increasing temperature in proteins from both species. Interestingly, the sequence of this segment is highly conserved in archaeal Nip7 proteins. We suggested that hydrophobic patch formed by this segment may be involved in protein-protein interactions.

Conclusion: In general, these data demonstrate the importance of water-protein interactions for the protein stability under high pressure and temperature.

Acknowledgements: This work was supported by RFBR grant No. 11-04-01771, RAS program Biosphere Origin and Evolution, RAS program A.II.6.8, SB RAS Integration projects 39 and 130.

References:

1. D.Van der Spoel et al. (2005) GROMACS: Fast, Flexible and Free. *J. Comp. Chem.* **26**:1701-1718.
2. P. P. Coltri et al. (2007). Structural insights into the interaction of the Nip7 PUA domain with polyuridine RNA. *Biochemistry*, **46**:14177-14187.

HOW MULTIPLE NFAT5 SITES MAY IMPACT TO THE NFAT5-INDUCIBLE GENE EXPRESSION: AN ILL-POSED INVERSE PROBLEM SOLUTION

Ponomarenko P.M.*[1], Ponomarenko M.P. [1], Kolchanov N.A.[1, 2, 3], Eisenhaber B.[4], Eisenhaber F.[4, 5, 6]

[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[2] *Novosibirsk State University, Novosibirsk, Russia;*
[3] *National Research Centre "Kurchatov Institute", Moscow, Russia;*
[4] *Bioinformatics Institute; A*STA R; Singapore;*
[5] *Department of Biological Sciences; National University of Singapore;*
[6] *School of Computer Engineering; Nanyang Technological University; Singapore*
e-mail: pon@bionet.nsc.ru
* *Corresponding author*

Key words: *NFAT5, stress response, hypertonicity, osmolarity, ill-posed inverse problem*

Motivation and Aim. The transcription factor NFAT5 (Nuclear Factor of Activated T-cells 5; synonyms: TonEBP and OREBP as Tonicity Enhancer Binding Protein and Osmotic Response Element-Binding Protein, respectively) is a master-gene for the osmotic stress response. It is, therefore, critical for many developmental and regeneration processes such as wound healing and it is known that certain drugs interfere with this pathway [1]. Multiple response elements (RE) within the regulatory gene regions are characteristic of the glucocorticoid, steroid, thyroid, auxin (AuxRE), and many other hormones as well as of the heat shock, hypertonic, osmotic, and many other stresses. There is few data on the molecular mechanisms underlying how multiple REs contribute to gene expression whereas a bulk of the data on the only single RE is available.

Methods and Algorithms. In our previous work [2], we have introduced an approximate solution of an ill-posed inverse problem, of dissecting the expression, φ, of a given gene originating from different variants of multiple REs and to decode the sequence-activity relationships of each of them:

$$\varphi = \varphi_0 \left(1 + \sum_{1 \le k \le N} P_k \sum_{\substack{1 \le n_1 \le N-k \\ n_1 < n_2 \le N-1 \\ \cdots \\ n_{k-1} < n_k \le N}} \varphi_{n_1 n_2 \ldots n_{k-1} n_k} \right) \quad (1)$$

here: ϕ_0, a basal expression level; P_k, a probability of k from amount N REs, $\Sigma_{1 \le k \le N} P_k \equiv 1$; N>1; $\phi_{\zeta\xi\ldots\varsigma}$ is an impact of a given set of ζ-th, ξ-th, … and ς-th REs into the induced expression level. As for N REs Eq. (1) contains 2^N variables, the optimal size of independent experimental data for its estimation by the multiple linear regression is 2^{2N}, i.e. 16, 64, 256, … at N∈ {2, 3, 4, …} that is unavailable. But two extreme cases, additive $\{P_1=1; P_{k>1}=0\}$ and multiplicative $\{P_{k<N}=0; P_N=1\}$, are solvable by the standard tools STATISTICA, the last case by Sandwich theorem. By this way we found the insignificant additive and significant multiplicative impacts of the multiple AuxREs into the expression level magnitudes of three auxin-responsive genes in plants, and a repression of these genes without auxin [2] that is impossible to do otherwise.

Results. In this work we analyzed six independent datasets on the multiple NFAT5 binding sites both wild-typed and subjected to genetic manipulations within promoters of six genes, namely: *HSP70-2* (heat shock protein 70-2) in human [3] and in mouse [4], *AR*

(aldose reductase) in human [5] and in rat [6], *SMIT* (sodium/myo-inositol cotransporter) in human [7], *AQP2* (aquaporin) in mouse [8]. As an example, upon 19 probes of four TonREs - A, B, C, D, - of murine *HSP70-2* promoter, x, upstream *Luc* (luciferase) gene of pGL2 plasmid by which the murine IMCD cells were transfected [4], we found that both additive$^{(a)}$ and multiplicative$^{(m)}$ impacts of the TonREs into Luc-activity, φ, are significant: $\{\varphi_0^{(a)}=1.0, \varphi_A^{(a)}=5.2, \varphi_B^{(a)}=3.6, \varphi_C^{(a)}=-6.3, \varphi_D^{(a)}=2.7\}$ and $\{\varphi_0^{(m)}=0.5, \varphi_A^{(m)}=7.3, \varphi_B^{(m)}=3.8, \varphi_C^{(m)}=0.2, \varphi_D^{(m)}=2.3\}$. There are the significant correlations between $\varphi^{(a)}$ and $\varphi^{(m)}$ values (r=0.98, $\alpha<0.025$), between reconstructions $\varphi(\varphi^{(a)}(x))$ and $\varphi(\varphi^{(m)}(x))$ based on each of them (r=0.96, $\alpha<0.05$), and, also (r=0.98, $\alpha<0.025$), between the magnitudes φ and their joint reconstruction $\varphi(x)$ using Eq.(2), $\delta(true)=1$ and $\delta(false)=0$, as:

$$\varphi(x)=\frac{2}{3}\left(1+\frac{1}{2}\sum_{\xi\in\{A,B,C,D\}}\varphi_\xi^{(a)}\delta(\xi\in x)+\frac{1}{2}\prod_{\xi\in\{A,B,C,D\}}\left(\varphi_\xi^{(m)}\right)^{\delta(\xi\in x)}\right).$$

Both negative $\varphi_C^{(a)}=-6.3$ additive and fractional $\varphi_C^{(m)}=0.2$ multiplicative impacts mean the repressive affect of TonRE-C to murine *HSP70-2* that was consciously lost earlier [4] because it is immeasurable experimentally against three rest TonREs activating this gene.

Conclusion. Finally, we confirmed the repression of *HSP70-2* using an independently derived human dataset [3]. Upon the majority of the investigated NAT5 clusters [3-8] we found two sorts of NFAT5-sites that make opposing impacts, namely: one NFAT5 site represses the gene, whereas the all others activate the same gene. This dual NFAT5 role may be an important feature of the expression pattern of the NFAT5-dependent genes.

Acknowledgements. This work was partially supported by the grants RFBR 11-04-01888, 13-04-00359, SciSh-5278.2012.4, Project 8740 from the Russian Ministry of Education & Science, and RAS Programs "Molecular and cellular biology".

References

1. L.K.Schenk et al. (2010) Cyclosporin-A induced toxicity in rat renal collecting duct cells: interference with enhanced hypertonicity induced apoptosis Cell Physiol Biochem, 26:887-900.
2. V.V.Mironova et al. (2013) How multiple auxin responsive elements may interact in plant promoters: a reverse problem solution. J Bioinform Comput Biol., 11:1340011.
3. J.I.Heo et al. (2006) The role of tonicity responsive enhancer sites in the transcriptional regulation of human hsp70-2 in response to hypertonic stress. Exp Mol Med., 38:295-301.
4. S.K.Woo et al. (2002) TonEBP/NFAT5 stimulates transcription of HSP70 in response to hypertonicity. Mol Cell Biol., 22: 5753-5760.
5. B.C. Ko et al. (1997) Identification and characterization of multiple osmotic response sequences in the human aldose reductase gene. J Biol Chem 272: 16431-16437.
6. K. Aida et al. (1999) Induction of rat aldose reductase gene transcription is mediated through the cis-element, osmotic response element (ORE): increased synthesis and/or activation by phosphorylation of ORE-binding protein is a key step. Endocrinology 140: 609-617.
7. J.S. Rim et al. (1998) Transcription of the sodium/myo-inositol cotransporter gene is regulated by multiple tonicity-responsive enhancers spread over 50 kilobase pairs in the 5'-flanking region. J Biol Chem 273: 20615-20621.
8. S.Z. Li et al. (2007) Calcineurin-NFATc signaling pathway regulates AQP2 expression in response to calcium signals and osmotic stress. Am J Physiol Cell Physiol 292: C1606-C1616.

SYSTEMS BIOLOGY ANALYSIS FOR PROMOTER ACTIVITIES OF GENETIC CONSTRUCTIONS FROM DELETION AND LINKER-SCANNING EXPERIMENTS

Savina M.S.*[1, 2], Mironova V.V.[1], Akberdin I.R.[1], Omelyanchuk N.A.[1], Likhoshvai V.A.[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia;*
[2] *Novosibirsk State University, Novosibirsk, Russia*
e-mail: captains-daughter@yandex.ru
* *Corresponding author*

Key words: *mathematical modeling, deletion analysis, linker-scanning mutagenesis, AuxREs*

Motivation and aims: Reporter constructions under promoters carrying deletions and linker-scanning mutations are widely used to study the input of regulatory sites into promoter activity. Systems biology offers new approaches for the analysis of reporter activity in these experiments. Data on expression of various mutant promoters being formalized in mathematical models may clarify the mechanisms for interaction of regulatory sites in response to external stimuli, as well as estimate their possible damage caused by genetic manipulation. Genes with multiple transcription factor binding sites of the same type play an important role in regulation of an organism development. For genes with this kind of promoters we developed a systems biology approach to study the experimental data on deletion analysis and linker-scanning mutagenesis.

Methods and Algorithms: In the approach we combine the methods of regression analysis and mathematical modeling with generalized Hill functions [1]. Regression analysis was carried out applying the method of least squares in the STATISTICA package. For parameter estimation we used gradient descent method in *Mathematica*. Euclidean distance from calculations to experimental data is served as a function of minimization.

Results: Our approach includes the following: (1) formalization of experimental data on the proposed protocol; (2) development of the linear additive and multiplicative models for regulation of transcription initiation from the target promoter, selection of the most appropriate model by multiple linear regression analysis; (3) extension mathematical model obtained on the step (2) using generalized Hill function method, selection of the models which successfully reproduce the experimental data. This approach was employed to study the effect of auxin-responsive elements (AuxRE) on the efficiency of functioning of the following gene promoters: 1) *Glycine max GH*3 [2]; 2) *Pisum sativum IAA*4/5 [3], 3) *Arabidopsis thaliana DR*5 [4]. Our approach allowed estimating the extent of the damage AuxRE-sites as a result of genetic manipulation for all three genes [2-4].

Conclusion: On the basis of the developed approach, we obtained the following results: (1) the auxin induction level of auxin-responsive promoters depends nonlinearly on the number of AuxRE sites; (2) the mixed additive-multiplicative models which successfully reproduce the experimental data were found; (3) it was shown that promoters of the auxin-responsive genes are in the repressive state before auxin treatment. Our approach may be applied to expression analysis of the other genes, activity of which is regulated by binding of transcription factors to multiple sites of the same type.

Acknowledgements: This work was partially supported by The Dynasty Foundation grant for young biologists, RFBR grants 11-04-01254-a, 11-04-01888-a, 11-04-01748-a and 12-04-33112, SS-5278.2012.4, Integration SB RAS programs 80.

References:

1. Likhoshvai V, Ratushny A. (2007), *JBCB* **5***: 521–531.*
2. Liu ZB et al. (1994), *Plant Cell* **6**: 645-657.
3. Ballas N et al. (1995), *ProcNatlAcadSci USA* **92**: 3483-3487.
4. Ulmasov T et al. (1997) *The Plant Cell* **9**: 1963-1971.

RECONSTRUCTION OF SQUAMOUS CELL CARCINOMA "ASSOCIOME" BASED ON ANALYSIS OF DIFFERENTIAL GENE EXPRESSION ACCORDING TO RNASEQ DATA AND INFORMATION STORED IN DATABASES

Saik O.V.*, Demenkov P.S., Ivanisenko V.A.
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
e-mail: saik@bionet.nsc.ru
** Corresponding author*

Key words: *"associome", squamous cell carcinoma, RNASeq, databases*

Motivation and Aim: Despite the advances of modern medicine, the rate of mortality from diseases is still high. Reconstruction of disease associations network (disease "associome") could lead to implementing of integrated approach in diagnostic, treatment and prevention of diseases especially based on RNASeq data and information from databases.

Methods and Algorithms: automatic analysis of information stored in databases, text-mining and data-mining approaches implemented in ANDSystem software [1], analysis of differential gene expression according to RNAseq data by EgdeR [2].

Results: Based on the analysis of differential gene expression according to RNASeq data and information from databases a network of squamous cell carcinoma associations at the level of 16 polymorphisms, 13 cellular localizations, 146 miRNAs, 71 transcription factors, 1734 enzymes and 81 metabolites was reconstructed. A profile of 67 metabolites - potential biomarkers of squamous cell lung cancer was proposed. Of these 32 metabolites are referred to in the literature as associated with cancer, 35 are potential biomarkers. It was shown that gene expression is increased in lung squamous cell carcinoma compared to control.

Conclusion: At present a huge amount of information concerning the relationships of genes, microRNAs, proteins and metabolites with diseases is accumulated in different databases. One way of the rational and efficient use of such information in biology and medicine is the identification and study of disease-associated modules in molecular genetic networks. This approach provides a deeper and more comprehensive understanding of the pathological processes and makes it possible to develop new tools for diagnosis, treatment and prevention of diseases. Use of the power of high-throughput sequencing (including RNASeq) for studing of diseases at the genome expression level allows to expand the molecular-genetic networks associated with the diseases, reconstructed based on the information from databases. Analysis of differential gene expression according to RNASeq data and information from databases were used for the reconstruction of squamous cell carcinoma association network ("associome"), containing 16 polymorphisms, 13 cellular localizations, 146 miRNAs, 71 transcription factors, 1734 enzymes, 81 metabolites.

References:

1. P.S. Demenkov et al. (2011-2012) ANDVisio: a new tool for graphic visualization and analysis of literature mined associative gene networks in the ANDSystem, *In Silico Biol,* **11(3-4)**:149-61.
2. M.D. Robinson et al. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data, *Bioinformatics*, **26(1)**:139-40.

COMPUTER PROGRAMS FOR COMPLEXITY ESTIMATION AND OLIGONUCLEOTIDE ANALYSIS OF REGULATORY DNA

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

Key words: *bioinformatics,computer analysis, transcription factor binding sites, complete genomes*

Motivation and Aim: The analysis of gene regulatory nucleotide sequences is an actual problem of modern bioinformatics. Especially the problem of transcription factor binding sites analysis in genome scale is important. Universal statistical measures that bioinformatics prefer to use for symbol texts are Shannon entropy [1] and text complexity [2]. Also scientists are interested in common frequency content for short gene regulatory regions and for long genome fragments. The aim of this work was analyzing the frequency content and studying complexity of the sequences depending on the oligonucleotide repeats using novel computer tools.

Methods and Algorithms: We used "Complexity" software developed at the Institute of Cytology and Genetics SB RAS earlier [3]. The new program module was written on C++. The program allows counting of common frequencies, frequency profiles and Shannon's entropy using approaches suggested in [2, 3]. Transcription factor binding site sequences were downloaded from the JASPAR database (http://jaspar.genereg.net/) as well as samples from TRRD database. The miRNA sequences were downloaded from the miRNA database (http://www.mirbase.org/ftp.shtml). Next we counted number of hits of each oligonucleotides in the TF dataset (the size of oligonucleotides was 2, 3, .., 8). Numeration algorithm was taken from [4]. The sequences were assigned numerical values (lexically ordered) for computational convenience (by coding in degrees of 4; we set 1 for A, 2 for T or U, 3 for G and 4 for C). For example, AAAAAA = 0, AAAAAT = 1, etc. For our research we used total frequencies for their own oligonucleotides and their complementary. Also we introduced a new measure of complexity which shows the diversity of changes of current symbol (nucleotide). This measure is close to linguistic complexity estimates.

Results: The analysis of the DNA and miRNA sequences from freely available databases was done for oligonucleotides of lengths of 2 to 9. The test sequences with high frequencies were selected and analyzed separately comparing with transcription factor binding sites motifs. We found common patterns for TFBS and miRNA and compared them to known binding motifs using STAMP web-tool (http://www.benoslab.pitt.edu/stamp) following approach suggested in [4]. Earlier we analyzed oligonucleotide frequencies and GC content in bacterial genomes and found correlations of such features with environment niches of the organisms studied [5]. To continue this work using new computer tools we constructed profiles of oligonucleotides' frequencies and Lempel-Ziv complexity in common. The peaks of frequencies of high-frequency nucleotides were marked in regions with low complexity for the tested samples. The correlation between introduced new complexity measure and Shannon entropy was calculated for short sequences. The correlation coefficient of 0.5 in a rounded average was obtained for sequences containing transcription factor binding sites.

Conclusion: The oligonucleotides with high frequencies were found for gene regulatory regions which are specific for some transcription factors. Their specialty can mean the fact that they may play the role of signals in multi-stage regulation of gene expression. It is interesting to note the dependence of complexity profile changes and sequences flanking to tandem repeats in genomes (for example, interspersed repeats in rye genome).

Availability: The program is available by request to authors.

Acknowledgements: Authors are grateful to V.G. Levitsky for science discussion and provided materials. The work was financially supported by Russian Ministry of Education and Science (project 8740).

References:

1. C. E. Shannon (1948) A mathematical theory of communication, *Bell System Technical Journal*, vol. 27, pp. 379-423 and 623-656, July and October, 1948.
2. V.D. Gusev, L.A. Nemytikova and N.A. Chuzhanova (1999) On the complexity measures of genetic sequences, *Bioinformatics* 15(12):994-999.
3. Yu.L. Orlov, V.N. Potapov (2004) Complexity: Internet-resource for analysis of DNA sequence complexity, *Nucleic Acids Res.,* Web-issue 2004, 32:628-633.
4. P. Putta, Y.L. Orlov, N.L. Podkolodnyy, C.K. Mitra (2011) Relatively conserved common short sequences in transcription factor binding sites and miRNA, *Vavilov journal of genetics and selection*, 15(4):750-756 (In Russian).
5. N.S. Safronova *et al.* (2012) Oligonucleotide frequencies and GC content of bacterial genomes are related to the environment evolution. In: *Proceedings of the Eighth International Conference on Bioinformatics of Genome Regulation and Structure \ System biology (BGRS\SB'2012),* 25-29 June 2012, ICG Press, Novosibirsk. P.276.

GENE EXPRESSION AND mRNA SECONDARY STRUCTURES IN MYCOPLASMA STRAINS

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

Key words: *codon composition, mRNA secondary structure, translation efficiency, mycoplasma*

Motivation and Aim: In a modern biology, study of genes expression efficiency and its optimization is timely and significant problem both for theory and for practice. From the fundamental point of view, a solution of this problem is important for theoretical estimation of expression level of genes, which experimental data is not available yet. The results of this research can also be helpful in practical biology, for instance in planning genetic engineering experiments or increasing producer (genetically modified organism) productivity. The main aim of this work is the analysis of translation elongation features in *Mycoplasma*.

Methods and Algorithms: We have implemented the algorithm for research of completely sequenced genomes. For each gene in a genome, the algorithm calculates parameters associated with mRNA codon composition and stability of its secondary structures. Consequently, so-called elongation efficiency index (EEI) is assigned to each gene [1, 2].

Results: We have analyzed genomes of 42 various *Mycoplasma* strains. It was found, that genes expression level in almost all organisms depends on number of mRNA secondary structures and doesn't depend on codon composition. Also three strains which significantly differed from other were revealed (*Mycoplasma haemocanis Illinois, Mycoplasma haemofelis Ohio2 and Mycoplasma haemofelis Langford1*). Genes of these organisms contain fewer numbers of local inverted repeats, which can form secondary structures in mRNA, than other *Mycoplasma* strains. In addition, we have constructed and analyzed the average profiles of mRNA secondary structures for all available *Mycoplasma* genomes.

Conclusion: Based on our data and information available in literature about *Mycoplasma* phylogeny we suppose, that particular *Mycoplasma* strains evolved in the line of reducing number of mRNA secondary structures for increasing genes expression level. It can be consequence of parasitic life style of these organisms.

Availability: The special software will be available on the server of ICG soon.

References:

1. N.V.Vladimirov, V.A.Likhoshvai, Y.G.Matushkin. (2007) Correlation of codon biases and potential secondary structures with mRNA translation efficiency in unicellular organisms, *Molecular Biology*, V. 41. № 5. P. 843-850.
2. V.A.Likhoshvai, Y.G.Matushkin. (2002) Differentiation of single-cell organisms according to elongation stages crucial for gene expression efficacy, *FEBS*.

INHIBITION OF PROTEIN SYNTHESIS DISSOCIATES DIFFERENT COGNITIVE DOMAINS IN A MOUSE MODEL OF POSTTRAUMATIC STRESS DISORDER

Toropova K.A.*[1], Anokhin K.V. [1, 2]
[1] *National Research Center "Kurchatov Institute", Moscow, Russia;*
[2] *P.K. Anokhin Institute of Normal Physiology, Moscow, Russia*
e-mail: xen.alexander@gmail.com
* *Corresponding author*

Key words: *memory; protein synthesis; stress; PTSD*

Motivation and Aim: Posttraumatic stress disorder (PTSD) belongs to one of the most frequent psychiatric disorders. However the basic neurobiological mechanisms of PTSD are still poorly understood. One of the current hypotheses is that PTSD exploits the same mechanisms of protein synthesis dependent neuronal plasticity that are implicated in consolidation of the normal long-term memory. The purpose of this study was to evaluate the role of protein synthesis in two cognitive outcomes of traumatic experience in a mouse model of PTSD: behavioral sensitization and traumatic memory.

Methods: In our PTSD model, we used an electric footshock as the aversive encounter. Seven days after footshock application mice were re-exposed to the shock context and their freezing response was recorded as a measure of conditioned (associative) fear memory. Twenty four hours later the freezing response to a neutral tone in a novel environment was scored as a measure of sensitized fear reaction. Twenty four hours after sensitized fear test, mice were tested in the elevated plus maze and open field to access trauma-induced alterations in anxiety-like behavior.

Results: Administration of protein synthesis inhibitor cycloheximide (100 mg/kg, i.p.) 30 min prior to a single footshock application (2 s, 1.0 mA) led to associative memory impairment, exaggerated sensitized fear amelioration and decrease of anxiety-like behavior compared to saline-injected controls. Cycloheximide injection before exposure to the intense aversive stimulus (three consecutive footshocks 10 s, 1.5 mA each) prevented behavioral sensitization and decreased anxiety level in the shocked mice, but had no amnestic effect.

Conclusion: Our results suggest that the development of PTSD is associated with protein synthesis dependent process which is similar to memory consolidation. Unimpaired traumatic memory in mice that received protein synthesis inhibitor injection before the intense stressful event implies distinct neural mechanisms for the formation of classical associative memory and the process of sensitization in the PTSD.

EXAMINATION OF THE CARDIOVASCULAR SYSTEM AND CELL CYCLE GENES INTERACTION

Yazeva A.S., Gorbunova V.Y., Vorobyova E.V.
Bashkir State Pedagogical University named after M. Akmullah, Ufa, Russian Federation
e-mail: annayazeva@mail.ru
** Corresponding author*

Key words: *renin-angiotensin system, the blood coagulation system, ACE (I/D)geneMTHFR(C677T) gene, FV(G1691A)gene*

*Motivation and objectives:*The study of genetic polymorphism, the products of which provide various links of the same chain or close conjugated metabolic chains, that is, within the same gene chain is of a great importance.The purpose of this study is to analyze the interaction between allelic states of genes of the cardiovascular system and cell cycle genes in oncopathology.

*Research methods and algorithms:*We used standard molecular genetic techniques: DNA isolation by phenol-chloroform extraction, PCR, RFLP analysis.The criterion $\chi 2$ (P) for 2x2 contingency tables with Yates's correction for continuity, the coefficient D' calculated in the 2LD program were used for thestatistical data processing. We applied the gene-gene simulation method of nonparametric program MDR - Multifactor-Dimensionality Reduction for the analysis of interactions between genes.

*Results.*The *ACE (I/D*) gene showed a significant increase in the frequency of allele *D(**χI**= 10,85 P=0,0018) and the frequency of the genotype *D/*Din a group of cancer patients (**χI**= 4,4327; P=0,0355) compared with the healthy individuals. According to the*FV(G1691A)*genes a proven prevalence of genotype *A/*G (χI= 5,69; P=0,0176) in cancer patients, that is, the presence of diseases will be of epigenetic nature. No significantdifferences in the *MTHFR(C677T*) gene were found. The intergene interaction analysis showed mutual negative effects of *ACE(I/D)* gene and *FV(G1691A)*, *MTHFR(C677T*) and *p53* (arg/pro), *ACE(I/D)*and *P53(arg/pro)*, which is associated with hypertension, hypoxemia, ischemia.

*Conclusion.*The methods of mathematical and computer modeling are increasingly being applied in practice to study the dynamic behavior of biological systems. This approach to the study may provide information as to the role and function of one or another gene in a complex genetic circuit that forms the physical function of the human body.

TOWARDS A COMPREHENSIVE microRNA INTERACTOME NETWORK

Chen M.*
College of Life Sciences, Zhejiang University, 310058 Hangzhou, China
e-mail: mchen@zju.edu.cn
** Corresponding author*

Key words: *microRNA, networks, bioinformatics*

Motivation and Aim: MicroRNAs (miRNAs), a class of 19-24 nucleotide long non-coding RNAs derived from hairpin precursors, have been widely shown to modulate various critical biological processes, including differentiation, apoptosis, proliferation, the immune response, and the maintenance of cell and tissue identity.

Current achievements in the plant miRNAs research area are inspiring. Cloning and functional elucidation of miRNA genes have greatly advanced our understanding of these small molecules. However, the ultimate goal for these related research projects is to draw a comprehensive view of miRNA-mediated gene regulatory networks in plants. Numerous bioinformatics tools competent for network analysis have been available. However, the bottleneck of plant miRNA-mediated network construction is to obtain sufficient reliable data to start this work.

Methods and Algorithms: In our previous work, we introduced a comprehensive workflow to obtain the desired data, which serve as the cornerstones for network construction [1]. For the upstream analyses of miRNA genes, the sequence features of miRNA promoters should be characterized. Regulatory relationships between transcription factors and miRNA genes need to be investigated. For the downstream part, we emphasized that the high-throughput degradome sequencing data were especially useful for genuine miRNA—target pair identification. Functional characterization of the miRNA targets is essential to provide deep biological insights into certain miRNA-mediated regulatory pathways. For miRNAs themselves, studies on their organ- or tissue-specific expression patterns and the mechanism of self-regulation of the miRNA genes were also investigated [2-4].

Results: In order to discover novel miRNAs or other small RNAs (sRNAs), and their target genes we established a new framework with adjustable parameters of current available algorithms for reverse identification of miRNA—target pairs in plants [5].

Moreover, it was reported that plant miRNAs could regulate the gene expression not only in plant, but also in the animals, it is a great development for researches on relationship between plant and animal miRNAs. Based on our expertise on plant sRNAs, genome sequencing and bioinformatics, we plan to conduct a comprehensive study of the plant sRNAs in animals, in particular the miRNAs that can lead to cross kingdom regulation of genes. We are motivated to *in silico* identify and characterize those miRNAs, and try to construct miRNA regulatory networks across species. Firstly, based on the current databases and NGS sequencing data, we are going to find out miRNAs in several edible plants such as rice, maize, wheat and some vegetables. Then we perform bioinformatics analysis to identify miRNA and its target sequences in the human, mouse, or rat genome. *Conclusion:* Genomic and functional analysis of these microRNAs and their targets can be performed to model a regulatory network of plant miRNA and animal RNA or genes. In addition, several miRNAs have been found to have links with some types of cancer or other diseases. We expect to find a new way to resist these disease by introduce plant

microRNAs. Interesting problem is to study interaction of miRNA with transcription factor in embryonic stem cells [6].

Availability: Tools and databases are available at web-site of Bioinformatics Department, Hangzhou: http://www.cls.zju.edu.cn/binfo/index.htm

Acknowledgements: Author is grateful to Chinese Ministry of Science and Technology (grant No.CR16-11) and Russian Ministry of education and sciences (project 8740).

References:

1. Yijun M, et al. (2011) Towards microRNA-mediated gene regulatory networks in plants, *Briefings in Bioinformatics*, 12(6): 645-659..
2. Yijun M, et al. (2009) Genome wide survey of rice microRNAs and microRNA-target pairs in the root of a novel auxin-resistant mutant, *Planta*, 230: 883-898.
3. Yijun M, et al. (2011) Construction of MicroRNA- and MicroRNA*-mediated Regulatory Networks in Plants, *RNA Biology*, 8(6): 1124-1148..
4. Yijun M, et al. (2012) Expression-based functional investigation of the organ-specific microRNAs in Arabidopsis, *PLoS One*, 7(11): e50870.
5. Chaogang S, et al. (2012) A reverse framework for the identification of microRNA—target pairs in plants, *Briefings in Bioinformatics*, doi: 10.1093/bib/bbs040.
6. Orlov Y, et al. (2012) Computer and statistical analysis of transcription factor binding and chromatin modifications by ChIP-seq data in embryonic stem cell, *J Integr Bioinform.*, **9**(2):211.

THE FUNCTIONAL ROLE OF GENES INVOLVED IN LIPID METABOLISM IN THE COURSE OF METABOLIC SYNDROME

Mulyukova R.V.*, Vorobyova E.V., Gorbunova V.U.
Bashkir State Pedagogical University named after M.Akmullah, Ufa, Russian Federation
e-mail: newidimc@mail.ru

Key words: *metabolic syndrome, obesity, lipid metabolism genes*

Motivation and objectives: Currently, the metabolic syndrome (MetS) is one of the major and socially important problems of medicine and genetics. The interest in the metabolic syndrome has been steadily growing; this is due to its widespread prevalence (up to 30% of the population). *The purpose of the present study* is a complex analysis of the impact of a number of genes (LEP, LEPR, PPARG, LPL) on the metabolic syndrome.

Research methods and algorithms: Standard genetic and molecular and statistical (analysis of gene interactions) methods were used.

Results: The analysis of gene and gene-gene interactions established the four-factor DNA loci interaction model to the genotype combinations that may lead to the dramatic increase in the total cholesterol levels (above 5.2 mmol/L) were assigned 17 various combinations.

37 various genotype combinations that determine the total cholesterol level as normal (up to 5.2 mmol / L) were found.

The GMDR analysis revealed statistically significant two-factor and three-factor models of interaction between loci, which are associated with high levels of triglycerides (TG) in the blood.

When analyzing the two-factor model of DNA loci interaction, the combinations of genotypes that may lead to the increase in the TG level have not been identified.

The 3 combinations of genotypes that determine the TG level as normal were identified.

When analyzing the three-factor model of DNA loci interactions to the combinations of genotypes that may lead to the increase in the TG level have been identified 2 combinations.

The 7 various combinations of genotypes that determine the TG level as normal were identified.

Conclusion.

The methods of mathematical and computer modeling are increasingly being applied in practice to study the dynamic behavior of biological systems. This approach to the study may provide information as to the role and function of one or another gene in a complex genetic circuit that forms the physical function of the human body.

STUDYING LONG QT SYNDROME USING PATIENT-SPECIFIC INDUCED PLURIPOTENT STEM CELLS

Dementyeva E.V.*, Grigor'eva E.V., Vyalkova A.V., Medvedev S.P., Shevchenko A.I., Zakian S.M.

Insitute of Cytology and Genetics SB RAS, Novosibirsk, Russia;
State Research Institute of Circulation Pathology, Novosibirsk, Russia
e-mail: dementyeva@bionet.nsc.ru

* *Corresponding author*

Key words: *long QT syndrome, induced pluripotent stem cells, cardiomyocytes*

Motivation and Aim: Long QT syndrome is a disease detected as a prolongation of the QT interval on electrocardiogram. Long QT syndrome is characterized by an increased risk of ventricular tachycardia, which can cause syncope, ventricular fibrillation, and sudden death. There are two forms of long QT syndrome – acquired and congenital. The congenital form is due to mutations in the genes involved in ionic channel functioning. To date it has been described more than 700 mutations in 13 genes. Despite the progress, which has been made in the understanding of long QT syndrome mechanisms and its treatment, a good model to study the disease has not been developed so far. One of the most perspective trends in this field is using induced pluripotent stem cells. It is based on derivation of patient-specific induced pluripotent stem cells carrying mutations causing long QT syndrome and their differentiation into cardiomyocytes. This approach allows obtaining cardiomyocytes from long QT syndrome patients at any time and in amounts sufficient for studying long QT syndrome molecular mechanisms and drug screening, thereby opening prospects for so-called personalized medicine.

Methods and Results: In this work, we sequenced the exons of two long QT syndrome genes (*KCNQ1* and *KCNH2*) in several patients. Two mutations were found in the *KCNQ1* fifth and seventh exons. A fibroblast culture was derived from the patient with the mutation in the *KCNQ1* fifth exon and was transduced with lentiviruses carrying the *Oct4*, *Sox2*, *Klf4*, and *c-Myc* cDNA to generate induced pluripotent stem cells. The cells obtained were characterized in detail. They demonstrated the same pattern of gene expression and methylation as seen in pluripotent cells and differentiated into derivatives of three germ layers. Moreover, the patient-specific induced pluripotent stem cells were able to form cardiac derivatives during spontaneous differentiation.

Conclusions: Patient-specific induced pluripotent stem cells carrying a mutation causing long QT syndrome were generated. These cells can be further differentiated into cardiomyocytes to model long QT syndrome in vitro.

Научное издание

Материалы международной школы молодых ученых «Биоинформатика и системная биология» (Новосибирск, 2013)

на английском языке

Program and abstracts of International young scientists school “Bioinformatics and Systems biology” (Novosibirsk, 2013)

Abstracts have been printed without
editing as received from the authors

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

Дизайн и компьютерная верстка: А.В. Харкевич

Подписано к печати 17. 06. 2013 г.
Формат бумаги 70 ×108 1/16. Печ. л. 3,1. Уч.-изд. л. 3,1
Тираж 80.

Отпечатано в Институте цитологии и генетики СО РАН
630090, Новосибирск, пр. акад. М.А. Лаврентьева, 10