

UDC 577.218

The start of systems biology in Ukraine

**M. Yu. Obolenskaya, B. T. Tokovenko, A. V. Kuklin, A. A. Frolova,
R. R. Rodriguez, V. A. Dotsenko, O. O. Dragushchenko**

Institute of Molecular Biology and Genetics, NAS of Ukraine
150, Akademika Zabolotnogo Str., Kyiv, Ukraine, 03680
m.obolenska@gmail.com

The first laboratory of Systems Biology in Ukraine (IMBIG NASU) represents a track record of its scientific results. They include the pioneered development of a web-based tool for genome-wide surveys of eukaryotic promoters for the presence of transcription factors binding sites (COTRASIF); the deciphered mechanisms of the fine-tuned and balanced response of primary hepatocytes to interferon alpha levels recorded after partial hepatectomy; the elaboration of a novel method of gene regulatory network inference compatible with GRID environment and the development of a stoichiometric model of folate-related one carbon unit metabolism in human placenta and its application for the characteristics of the system's behavior as a whole at different human pathologies.

Keywords: systems biology, gene regulatory network, metabolic modeling, microarray.

«... we need to overcome the idea, so prevalent in both academic and bureaucratic circles, that the only work worth taking seriously is highly detailed research in a speciality. We need to celebrate the equally vital contribution of those who dare to take what I call «a crude look at the whole».»

Murray Gell-Mann, Nobel Laureate in Physics, 1994.

The beginning of XXI century was marked by the emergence of the new branch in biological sciences – Systems Biology that is defined as the quantitative analysis of the dynamic interactions between the components of a biological system and aims at understanding the behaviour of the system as a whole, as opposed to the behaviour of its individual constituents [1]. Systems Biology recognizes that: the overall systemic features differ from the collective features of its constituents; each system has subsystems being itself a subsystem of the system of higher order; the system has a hierarchic structure; it is a dynamic formation, which changes during its functioning and evolution. Systems Biology uses the high-throughput technologies of molecular biology, mathematics, physics, informatics and engineering. It

handles the massive amounts of data; its measurements must be as quantitative as possible – for the ultimate integration and modeling this information is a quantitative process; it combines discovery-driven and hypothesis-driven approaches to systems studies; the modeling is the ultimate objective of systems biology that helps to understand the mechanistic underpinnings of particular biological systems phenomena that emerge from the integrated operation of the information components (emergent properties) [2]. Systems Biology promotes the close cooperation of scientists from different areas of science and technology. The promises of Systems Biology are great – from creation of virtual organs to virtual man. It will affect many areas of the biomedical sciences from drug design to medical treatment going from the statistically average for typical patient to the individual with his/her specific geno- and phenotype.

Institute of Molecular Biology and Genetics, NAS of Ukraine, 2014

Table 1
Statistics of COTRASIF visits by countries

Countries	Total number of tasks	Percent of tasks
United States	1132	23.80
Ukraine	879	18.48
Germany	361	7.59
United Kingdom	289	6.08
China	217	4.56
France	146	3.07
Canada	144	3.03
Norway	128	2.69
Italy	126	2.65
India	102	2.14
Other	1233	25.9
Total	4,757	100 %

In 2009 the first laboratory of Systems Biology in Ukraine (Prof. Maria Obolenskaya) was organized in the Institute of Molecular Biology and Genetics of National Academy of Sciences of Ukraine. Herein we present the initial results obtained by our small group.

COTRASIF: conservation-aided transcription factor binding site finder. The transcriptional regulation of gene expression relies on the effects of the transcription factors (TFs) bound to specific regulatory elements, transcription factor binding sites (TFBS) in the promoters of the genes, activating or repressing the corresponding gene. While experimental identification of TFBS within single gene promoters is common, with no prior information this process is both effort- and time consuming. Computational prediction of TFBS marks the potential targets for further experimental verification, and presents self-sufficient data on the gene regulation patterns, associated with each specific TF.

We have developed an easy to use web-based tool for genome-wide surveys of eukaryotic promoters for the presence of TFBS [3]. The most common way of representing conserved sequences (such as TFBS) is to use consensus strings (built using IUPAC – International Union of Pure and Applied Chemistry – nomenclature) but these strings contain only a small portion of the information, available from the set of initial sequen-

ces. Instead, position frequency matrix (PFM) consists of each of four possible nucleotide counts per each position of the identified binding site. Position weight matrix (PWM) represents the complete nucleotide occurrence probabilities for each position using position weight matrices and information content evaluation of each matrix position [4–7]. It also allows the quantification of the similarity between the weight matrix and a potential TFBS detected in the target sequence [6, 8–10]. However, even the matrix representation is prone to information loss: it is perfect for recording single nucleotide frequencies, but higher order groupings (like pairs or triplets of nucleotides) are not accounted for. Joining hidden Markov models (HMMs) to PFM and PWM helps to preserve this kind of information as well and yields the method of storing full TFBS characteristics.

Taking into account the length of matrices used for TFBS prediction (usually less than 15 nucleotides), the search may yield numerous false-positive results, occurring by chance. To reduce their number, and avoid the loss of sensitivity additional processing of results – gene orthology information [11, 12], was applied.

Therefore, a new web-based tool, called COTRASIF, allows conducting genome wide searches of the user-specified TFBS or sequence sets, and provides an option to filter the list of results based on evolutionary conservation [13]. COTRASIF allows using either PFM, or a set of sequences. When a set of sequences is provided, the hybrid HMM-PWM search method is used instead of the PWM-only search method. COTRASIF contains its own, regularly updated database of eukaryotic promoters (automatically imported from Ensembl genome annotations), integrated JASPAR CORE database of high-quality TFBS matrices and evolutionary conservation filter. Therefore COTRASIF is a fully integrated solution which is freely available at <http://biomed.org.ua/COTRASIF/>. There is an extensive online help available for the tool.

COTRASIF application. COTRASIF has become a popular tool all over the world that is clearly demonstrated by the Table 1.

Using COTRASIF we identified and analyzed the rat genes of the potential primary response to interferon alpha (IFN α). The dominating IFN α transduction pathway is realized via ISGF3 transcription factor (interfe-

ron-stimulated gene factor 3), which binds specifically to ISRE (interferon-stimulated response element) [14]. Thus, searching for ISRE has revealed 162 genes with putative conserved ISRE among 17725 protein-coding genes of *Rattus norvegicus*. Using web-based tool FatiGO [15] we have predicted previously unknown potential targets of IFN α referring to the complement system and components of synapse. The experimental verification of this prediction is currently in progress.

Gene expression profiling. Gene expression profiling is the measurement of expression of thousands of genes at once, to create a global picture of transcriptome and ultimate system level information for inference of gene regulatory network. There are two principal approaches used for the gene expression profiling – based on hybridization of RNA sample with previously identified target genes attached to a solid surface (microarray) and sequence based technique (*e. g.* RNA-seq).

We have analyzed the gene expression profile in primary rat hepatocytes isolated from intact liver and treated during 3 h and 6 h with IFN α (250 U/ml) [16]. The early up-regulated expression, specific antiviral activity and activated intracellular IFN α signaling were recently discovered in regenerating rat liver [17–20] against a background of different other manifestations of the innate immune response [21–23]. To approach the problem we mimicked *in vitro* the quantitative and temporal local levels of IFN α after partial hepatectomy (PH) and delineate the response of primary hepatocytes to such treatment.

The gene expression profile was analyzed with GeneChip® Rat Genome 230 2.0 Array (Affymetrix). It provides comprehensive coverage of the transcribed rat genome on a single array, comprised of more than 31000 probe sets, analyzing over 30000 transcripts and variants from over 28000 well-substantiated rat genes.

The analysis of gene expression profiles after 3 h- and 6 h-incubation of hepatocytes with IFN α revealed 24 and 128 differentially expressed up-regulated genes, a smaller number than that induced by longer treatment with higher doses of IFN α , but only partly overlapping with them. In total, they differ in that the response is smaller in magnitude, that there is a simultaneous expression of activators and inhibitors of transcription and intracellular signaling, co-regulators of transcription factors are involved, as well as several specific modifi-

ers of proteins. As such the response is cell and situation specific [24, 25] and temporally regulated.

The sequence of how the differentially expressed genes enter the process is of principal biological importance. At the early stage the cells exploit the strategic advantage of the energy saving principle by using available resources and protein changes to get the fastest response: autophagy to eliminate unneeded cellular materials; reversible ISGylation of preexisting proteins exceeding other types of protein modification; partial inhibition of the energy consuming translation; co-activation of preexisting nuclear receptors to provide faster changes in gene expression in comparison with other transcription factors acting as target proteins of slower responding signaling cascades; switching the mitochondrial dNTP synthesis to the salvage pathway and, lastly, the «insistent appeal» to other hepatic cells for their involvement in the process.

The further development of the response is associated with a more intensive nuclear control of gene expression via up-regulated STAT1, STAT2 and IRF9 against the background of activated preexisting Jak/STAT, PI3K/Akt and p38/MAPK pathways; enlargement of the list of differentially expressed genes and increasing the response of previously expressed genes (in most cases) and involving restraining factors. The simultaneous expression of proinflammatory and restraining factors seems to be a distinguishing feature of the hepatocyte response to a specified IFN α impact. The propagation of IFN α signaling via activated transmission from MAVS (mitochondrial antiviral signalling protein) to TBK1 (TANK-binding kinase 1) may be opposed by inhibition of RIG1 (Retinoic Acid Inducible Gene 1 Protein) and its pathway; T cell activation by CXC ligands and MHC complexes by T cell inhibition via negative co-receptor PD-1; initiator caspase 8 by the most potent cellular inhibitor of apoptosis, XIAP [16].

Thus, the realization of protein activities encoded by differentially expressed genes would affect all regulatory levels of gene expression and provide a time-dependent, strictly balanced and energy sparing response. We hypothesize that elevated expression of IFN α at the beginning of regeneration process may induce the similar processes.

There are several facts either confirming the involvement of several genes that *in vitro* respond to IFN α or

Table 2

Characteristics of regenerating liver related to the *in vitro* hepatocytes response to IFN [16]

Characteristics		
Compromised liver restoration with <i>KO</i> genes	Up-regulated individual mRNAs*	Up-regulated protein content and enzyme activity [†]
<i>Er</i> , <i>Ppar</i> , <i>Rar/Rxr</i>	<i>Ifn</i> ^{h,t} , <i>Stat1</i> ^h , <i>Stat2</i> ^h , <i>Atf2</i> ^h , <i>Atf4</i> ^h	Accumulation of autophagosomes
<i>Stat5</i> , <i>Nf- B</i>	<i>Mx1</i> ^h , <i>Mx2</i> ^h	Lysosome proteolysis
<i>Dab</i> , <i>Casp8</i> , SCF(-TrCP) complex	<i>p28</i> ^{GANK^t}	Caspase 8
	<i>Uba7</i> ^t , <i>Trim25</i> ^t	
<i>B2m</i>	<i>Eif2</i> <i>k2</i> ^h	Cytosolic 5'-nucleotidase
	<i>Casp8</i> ^t	3' 5'-exonuclease

*«h» and «t» designate hepatocytes and liver tissue which RNA and protein were used for the study.

arguing in favor of their potential participation in liver restoration (Table 2). The formal resemblance between the expression of a limited number of genes in our experiment and the processes activated during the transition period after PH is not sufficient to consider IFN as a candidate regulator at this stage of liver restoration, but it offers a challenge for a more in-depth study and gives a potential clue to elucidating its place in the context of liver restoration.

Inferring gene-regulatory network. Essentially all biological functions of a living cell are carried out through the interplay between many protein-coding and protein-noncoding genes. Identifying these gene networks and their functions is a main challenge in systems biology, which using the high-throughput biotechnologies enables the collection of large-scale genomics data on gene expressions, genome-wide location of transcription factor binding sites, protein–protein interactions, genetic variations, and many other types of data. These data provide valuable system level information on different aspects of the complex biological processes and make it possible to infer the underlying networks.

Various models have been proposed in literature to represent and simulate the behavior of Gene Regulatory Networks (GRNs). Boolean, Bayesian Networks, Differential Equations, Information Theory, supervised learning methods are some of the prominent ones. Boolean networks use a binary variable (gene is «on» or «off») to represent the state of a gene activity and a directed graph to represent genes connections. The state of a particular gene in time $t + 1$ can be described as

Boolean function of other genes in time t . Bayesian models make use of Bayes rules and consider gene expressions as random variables. In Bayesian models each node is associated with a probability function that takes as input a particular set of values for the node's parent variables and gives the probability of the variable represented by the node. Difference and differential equations allow more detailed descriptions of network dynamics, by explicitly modeling the concentration changes of molecules over time. Information theory models correlate two genes by means of a mutual information correlation coefficient and a threshold.

Each of the models has their strength and weakness and for each of them it is possible to apply different methods which may be realized by different algorithms. Despite substantial research on this topic, it remains a great challenge to elucidate the complete network due to the complexity of the transcription processes and the noisy nature of high-throughput raw data [26–28].

One of the significant problems the scientist encounters is a NP-hard (Non-deterministic Polynomial-time hard) character of the GRN inference that is the time required to create GRN using any currently known algorithm increases very quickly as a size of the data grows [29]. To overcome the problem the modern reconstruction algorithms use heuristics that reduce the field of search at expense of the accuracy of reconstruction and its biological relevance.

We proposed a novel integrated approach to infer GRN – an ensembl-method that combines the networks obtained by other methods, selects their best traits, and

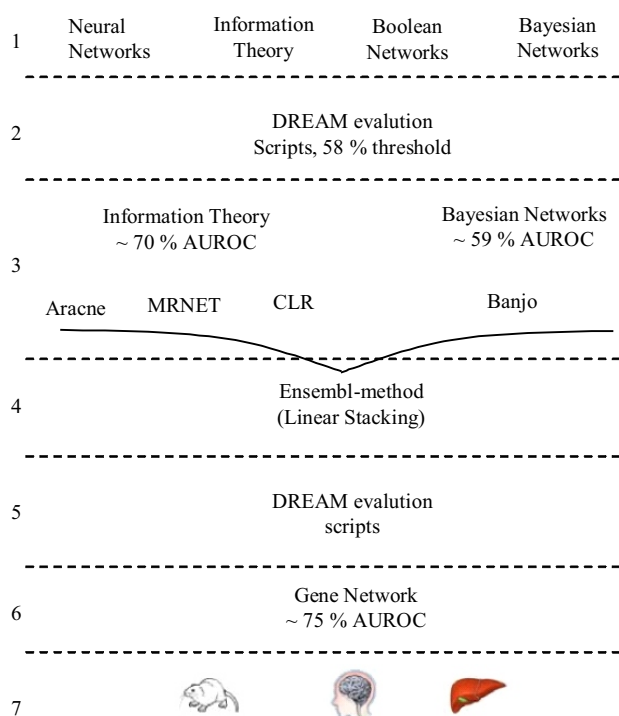


Fig. 1. The scheme of step-by-step evaluation of ensembl-method

uses Grid environment that enables the inference in reasonable time with minimal heuristics (if needed) due to high level of parallelization of calculations [30]. The six sequential steps testifying the applicability of specific combination of the models with methods and algorithms for the tasks parallelization in GRID environment are represented in the Fig. 1. The GRN reconstruction technique was validated and evaluated by the «golden standard» (The Dialogue for Reverse Engineering Assessments and Methods (DREAM) program; <http://www.the-dream-project.org/>) and synthetic networks. The DREAM provides the expression data and corresponding GRNs and encourages researchers to develop new efficient computation methods and evaluate them by comparison with the results of the «golden standard». The synthetic GRNs of different size were self-made on the basis of DREAM-4 InSilico Challenge (<http://gnw.sourceforge.net/dreamchallenge.html#dream4challenge>) by the GeneNetWeaver (<http://gnw.sourceforge.net>) tool [30].

The conventional AUROC (the Area under the Receiver operating Characteristics curve) and the AUPR (the Area Under Precision/Recall curve) evaluation measu-

res have been used to compare the obtained networks with the above mentioned standards*. The best methods (Aracne, Banjo, CLR, MrNet) were chosen by these criteria and compared with the ensembl-method that has combined these methods using their best features. The characteristics of ensembl-method appeared substantially better than those of each specified methods but its obvious advantage is that it may be used in GRID environment. The pilot inference of GRN in rat hepatocytes, treated with different pharmacological and xenobiotic compounds (7700 genes, 5288 animal studies; gene expression Omnibus, GEO; <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE8858>), has taken 15 days of parallel computations on GRID while serial computations on average modern computer would take up to 580 days (unpublished). So, the ensembl-method combined with GRID technologies has definite advantage and is admissible for inferring GRNs. Our virtual organization Sysbio is registered in Ukrainian National grid.

A booming of computational biology, which sheerly was a theoretician's fantasy twenty years ago, has become a reality. Various GRNs are successfully reconstructed from different species, tissues and cells [31–33]. But the challenging problems in this area still remain. Just to name a few, the number of variables (genes) in a system under investigation is considerably more than the number of measurements (observations) of those variables that does not correspond to the requirements of matrix formation; the absence of reference GRN for human and other model mammalian organisms. There are several propositions to overcome these problems [34, 35] but the elaboration of new methods and optimization of the existing ones are greatly encouraged.

Modeling and simulation of folate-mediated one-carbon metabolism in human placenta. Folate-mediated one-carbon metabolism (FOCM) in human placenta is one of the areas of our scientific interests [37–39]. The FOCM is represented by a network of interconnected metabolic pathways necessary for the synthesis of purine nucleotides, thymidylate, the remethylation of homocysteine (Hcy) to methionine (Met), prevalent

*AUROC is evaluated using the ratio between the true positive (TP) rate and the false positive (FP) rate as a performance measure. The AUPR is calculated using the ratio between precision (i. e. $TP/(TP + FP)$) and recall (i. e. $TP/(TP + FN)$) with FN denominating the number of false negative data.

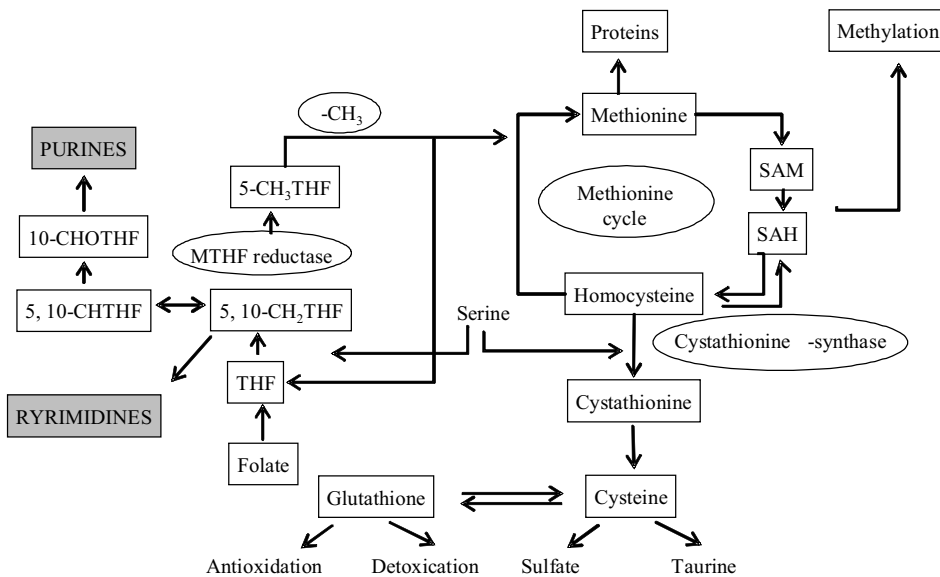


Fig. 2. The general scheme of folate-mediated one-carbon unit metabolism: 5-CH₃THF – 5-methyltetrahydrofolate; 5,10-CH₂THF – 5,10-methylenetetrahydrofolate; THF – tetrahydrofolate; 5,10-CHTHF – 5,10-methenyltetrahydrofolate; 10-CHOTHF – 10-formyltetrahydrofolate; SAM – S-adenosylmethionine; SAH – S-adenosylhomocysteine

majority of methylation reactions and the synthesis of glutathione (Fig. 2).

Thus, FOCM is involved in the fundamental molecular functions and biological processes connected with them – proliferation, transcription, translation, maintenance of cellular redox state and detoxication [40, 41]. According to the specificity of its biochemical reactions the FOCM network is highly sensitive to nutritional status of amino acids and several vitamins (folate and vitamins B12, B6, and B2), that are cofactors in numerous FOCM reactions, and to numerous penetrant gene variants as the majority of its enzymes are highly polymorphic [42]. At the same time it has to be robust and steady as it is involved in fundamental biological processes. The regulation of this balance is an intriguing question. Numerous pathological states with unknown etiologies (cancer, cardiovascular diseases *etc.*) are associated with perturbations in this network, particularly during pregnancy, affecting the development of a fetus, the health of the child post partum and maternal health [43, 44].

Although a considerable research has elucidated biochemical details of FOCM, most studies have focused primarily on single reactions or pathways in isolation, failing to capture the overall functioning of the system. Even large-scale international trials would not be able to experimentally follow the FOCM variability in its endless combinations. Mathematical modeling has proven to be a powerful tool for filling such gap.

The mathematical representations of cellular metabolism have many facets, ranging from topological and stoichiometric descriptions to kinetic models of metabolic pathways. We have applied the stoichiometric one as it goes beyond merely topological arguments, takes the specific relative quantities of reactants into account, the reversibility/irreversibility of reactions, and allows putting constraints on the functional capabilities of metabolic networks. At the same time it does not require the knowledge of kinetic characteristics of enzymes, concentrations of metabolites *etc.* which are difficult to obtain but are necessary for kinetic modeling. The stoichiometric analysis largely relies on the steady state assumption and does not allow accounting for allosteric regulation.

We have addressed several questions – the impact of elevated concentration of Hcy, down-regulated activities of methyleneTHF reductase (MTHFR; EC 1.5.1.20), cystathionine synthase (CBS; EC 4.2.1.22), methionine synthase (MS; EC 2.1.1.13), Met nutritional deficiency and several combinations of these factors in the FOCM network in human placenta [45, 46]. Elevated concentration of Hcy is a well known marker of perturbations in the placental FOCM [37]; three enzymes are polymorphic, their mutated forms possess lower catalytic activities than the wild forms and their prevalence at Caucasians is substantial [42]. MTHFR catalyses the reduction of 5,10-methyleneTHF to 5-methylTHF which donates methyl group to methionine

cycle; CBS is responsible for irreversible Hcy elimination and MS – for the synthesis of Met; methionine deficiency is a frequent event among vegetarians, vegans and vulnerable social groups.

We have mined the literature for placental specific gene expression of FOCM enzymes, for the first time modeled placental specific FOCM and simulated the situations close to real-life ones. The two wellknown variants of stoichiometric analysis, namely elementary flux modes (EFM) and flux balance analysis (FBA) were applied [45].

One common characteristic of the mentioned approaches is that they both compute the complete set of EFMs given the network stoichiometry. EFM is a minimal operational unit at steady state in metabolic networks satisfying the thermodynamic constraint regarding the reversibility of each reaction [46]. In other words, EFM is a minimal biochemical pathway that, at steady state, catalyses a set of reactions between input and output metabolites. Each metabolic network is characterized by unique set of EFMs.

We arranged all biochemical reactions in descending order according to their frequency in the set of EFMs. Serine hydroxymethyltransferase (SHMT; EC 2.1.2.1) and serine (Ser) input are the most frequent while the reactions of taurine synthesis are the least frequent ones pointing to the importance of the former as main suppliers of one carbon units in FOCM and partial independence of the latter from folate metabolism. This conclusion coincides with experimental data [40, 41].

The simulation of the abovementioned situations has revealed that they affect the majority of reactions in the network in quite specific way [45, 46]. For example, the twofold increase in Hcy concentration is associated with the down-regulation in the fluxes supporting the synthesis of nucleic acids precursors and up-regulation of cysteine production providing the elimination of Hcy excess that correlates with the experimental data obtained on the placental explants cultivated with Hcy [39]. The C677T heterozygosity of MTHFR stimulates the pronounced accumulation of Hcy, down-regulation of Hcy remethylation and increasing requirement in 5-MTHF. In this situation the folate therapy is justified while in the case of MS deficiency the rapidly increasing concentration of Hcy, down-regulated reactions of methionine cycle are associated with the inhibited flux

through MTHFR. In this case the folates prescription will not improve the situation as the flux through MTHFR is inhibited. In context with these results the homozygous MS knockout mice embryos survive through implantation but die soon thereafter. Nutritional supplementation during pregnancy was unable to rescue embryos that were completely deficient in methionine synthase [47].

As can be seen from above the simulation of close-to-real-life situations in FOCM model induces on the one hand the changes similar to those experimentally obtained that underscore the validity of the model. On the other hand the results of analysis draw attention to several previously unknown interrelations in FOCM that were not experimentally explored so far and have to be taken into account. The more detailed analysis of FOCM at the patients is required for the justified therapy.

Looking ahead. Therefore, the Systems Biology has officially started in Ukraine and we have got nearly four years experience and received our first results. During the nearest time we intend to experimentally verify the predictions made by COTRASIF, gene expression profiling, FOCM simulations and also master the reconstruction and analysis of GRNs.

Our general considerations refer to the perspectives of Systems Biology development in Ukraine. Understanding the function of complex biological systems is one of the greatest challenges facing the modern science. The scientific community in the world is going along this way. Ukraine has all possibilities to move in the same direction. We have highly qualified specialists in different areas of fundamental science who might master and spread the new knowledge in scientific community. The courses on Systems Biology have to be included in the programs of high schools. The special grants have to be established for the development of Systems Biology. Systems Biology may start from the pure computational biology using a wealth of available data and software and widespread partnership with our foreign colleagues but the investments in facilities for high-throughput technologies are indispensable. In conclusion – a journey of a thousand miles begins with a single step. The first steps are already made in the Institute of Molecular Biology and Genetics of NAS of Ukraine. Let it be a start of great promise.

Започаткування системної біології в Україні

М. Ю. Оболенська, Б. Т. Токовенко, А. В. Куклін, А. О. Фролова,
Р. Р. Родригес, В. А. Доценко, О. О. Драгущенко

Резюме

Представлено наукові здобутки першої в Україні (ІМБіГ НАН України) лабораторії системної біології. Вони включають створення веб-інструменту для всегеномного пошуку сайтів зв'язування транскрипційних факторів у промоторах еукаріотних генів; розкриття тонко збалансованого механізму відповіді первинних гепатоцитів на дію інтерферону альфа в дозі, зареєстрованій на першому етапі регенерації печінки після часткової гепатектомії; розробку нового методу інференції мережі генної регуляції для його використання в середовищі ГРІД і створення стехіометричної моделі фолатного циклу у плаценті людини для характеристики поведінки системи за різних патологічних станів.

Ключові слова: системна біологія, генні регуляторні мережі, метаболічне моделювання, мікроаррей.

Старт системной биологии в Украине

М. Ю. Оболенская, Б. Т. Токовенко, А. В. Куклин, А. А. Фролова,
Р. Р. Родригес, В. А. Доценко, А. А. Драгущенко

Резюме

Представлены научные наработки первой в Украине (ИМБіГ НАН Украины) лаборатории системной биологии. Они включают создание веб-инструмента для всегеномного поиска сайтов связывания транскрипционных факторов в промоторах эукариотических генов; раскрытие тонко сбалансированного механизма ответа первичных гепатоцитов на действие интерферона альфа в дозе, регистрируемой в печени на раннем этапе после частичной гепатектомии; разработку нового метода инференции сетей генной регуляции для его использования в среде ГРИД и создание модели фолатного цикла в плаценте человека для характеристики поведения системы при разных патологических состояниях.

Ключевые слова: системная биология, генные регуляторные сети, метаболическое моделирование, микроаррей.

REFERENCES

- Dollery C, Kitney R. Systems Biology: a vision for engineering and medicine. Report of the Academy of Medical Sciences and Royal Academy of Engineering Working Group on Systems Biology. London, 2007; 65 p.
- Hood L, Rowen L, Galas DJ, Aitchison JD. Systems biology at the Institute for Systems Biology. *Brief Funct Genomic Proteomic*. 2008; 7(4):239–48.
- Tokovenko B, Golda R, Protas O, Obolenskaya M, El'skaya A. COTRASIF: conservation-aided transcription-factor-binding site finder. *Nucleic Acids Res*. 2009; 37(7):e49.
- Sandelin A, Alkema W, Engstrom P, Wasserman WW, Lenhard B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res*. 2004; 32, Database issue D91–D94.
- Lenhard B, Wasserman WW. TFBS: Computational framework for transcription factor binding site analysis. *Bioinformatics*. 2002; 18(8):1135–6.
- Schones DE, Sumazin P, Zhang MQ. Similarity of position frequency matrices for transcription factor binding sites. *Bioinformatics*. 2005; 21(3):307–13.
- Stormo GD. DNA binding sites: representation and discovery. *Bioinformatics*. 2000; 16(1):16–23.
- Hertz GZ, Stormo GD. Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics*. 1999; 15(7–8):563–77.
- Hertz GZ, Hartzell GW 3rd, Stormo GD. Identification of consensus patterns in unaligned DNA sequences known to be functionally related. *Comput Appl Biosci*. 1990; 6(2):81–92.
- Stormo GD, Hartzell GW 3rd. Identifying protein-binding sites from unaligned DNA fragments. *Proc Natl Acad Sci USA*. 1989; 86(4):1183–7.
- Wasserman WW, Palumbo M, Thompson W, Fickett JW, Lawrence CE. Human-mouse genome comparisons to locate regulatory sites. *Nat Genet*. 2000; 26(2):225–8.
- Lenhard B, Sandelin A, Mendoza L, Engstrom P, Jareborg N, Wasserman WW. Identification of conserved regulatory elements by comparative genome analysis. *J Biol*. 2003; 2(2):13.
- Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M, Bayerlein M, Werner T, MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics*. 2005; 21(13):2933–42.
- Lon F, Liu H, Hahn C, Sumazin P, Zhang MQ, Zilberstein A. Genome-wide prediction and analysis of function-specific transcription factor binding sites. *In Silico Biol*. 2004; 4(4):395–410.
- Draguschenko OO, Tokovenko BT, Obolenskaya MYu. Initial analysis of the results of the genome-wide search for the genes of response to interferon alpha. *Ukr Biokhim Zh*. 2010; 82(1):82–9.
- Kuklin A, Tokovenko B, Makogon N, Oczko-Wojciechowska M, Jarzab B, Obolenskaya M. Hepatocytes response to interferon alpha levels recorded after liver restoration. *J Interferon Cytokine Res*. 2013; Oct 9. [Epub ahead of print].
- Perepelyuk MM, Fedorchenko DB, Rybalko SL, Obolenskaya MYu. Interferon expression in the rat liver after partial hepatectomy. *Biopolym Cell*. 2006; 22(4):276–82.
- Perepelyuk MM, Kuklin AV, Shcherba IaV, Tokovenko BT, Makogon NV, Gogler A, Szala S, Obolenskaya MYu. Interferon and protein kinase R during rat liver restoration after partial hepatectomy. *Biopolym Cell*. 2009; 25(2):145–9.
- Han H, Zhang L, Xu C. Effects of the pathogen infection-related genes on rat liver regeneration following 2/3 hepatectomy. *Life Sci J*. 2007; 4(2):50–5.
- Chen X, Xu C, Zhang F, Ma J. Microarray approach reveals the relevance of interferon signaling pathways with rat liver restoration post 2/3 hepatectomy at cellular level. *J Interferon Cytokine Res*. 2010; 30(7):525–39.
- Markiewski MM, DeAngelis RA, Lambris JD. Liver inflammation and regeneration: two distinct biological phenomena or parallel pathophysiologic processes? *Mol Immunol*. 2006; 43(1–2):45–56.
- Riehle KJ, Dan YY, Campbell JS, Fausto N. New concepts in liver regeneration. *J Gastroenterol Hepatol*. 2011; (Suppl. 1):203–12.
- Hosoya S, Ikejima K, Takeda K, Arai K, Ishikawa S, Yamagata H, Aoyama T, Kon K, Yamashina S, Watanabe S. Innate immune responses involving natural killer and natural killer T cells promote liver regeneration after partial hepatectomy in mice. *Am J Physiol Gastrointest Liver Physiol*. 2013; 304(3):G293–299.
- de Veer MJ, Holko M, Frevel M, Walker E, Der S, Paranjape JM, Silverman RH, Williams BR. Functional classification of interferon-stimulated genes identified using microarrays. *J Leukoc Biol*. 2001; 69(6):912–20.
- Zocco MA, Carloni E, Pescatori M, Saulnier N, Lupascu A, Nisita EC, Novi M, Candelli M, Cimica V, Mihm S, Gasbarrini G, Ramadori G, Gasbarrini A. Characterization of gene expression

- profile in rat Kupffer cells stimulated with IFN-alpha or IFN-gamma. *Dig Liver Dis.* 2006; **38**(8):563–77.
26. Cerulo L., Elkan C., Ceccarelli M. Learning gene regulatory networks from only positive and unlabeled data. *BMC Bioinformatics* 2010; **11**:228.
 27. Gupta R., Stinccone A., Antczak P., Durant S., Bicknell R., Bikfalvi A., Falciani F. A computational framework for gene regulatory network inference that combines multiple methods and datasets. *BMC Syst Biol.* 2011; **5**:52.
 28. Cai X, Bazerque JA, Giannakis GB. Inference of gene regulatory networks with sparse structural equation models exploiting genetic perturbations. *PLoS Comput Biol.* 2013; **9**(5): e1003068.
 29. Lee WP, Tzou WS. Computational methods for discovering gene networks from expression data. *Brief Bioinform.* 2009; **10**(4): 408–23.
 30. Frolova AO. Overview of methods of reverse engineering of gene regulatory networks: Boolean and Bayesian networks. *Bio-polym Cell.* 2012; **28**(3):163–170.
 31. Peter IS, Faure E, Davidson EH. Predictive computation of genomic logic processing functions in embryonic development. *Proc Natl Acad Sci USA.* 2012; **109**(41):16434–42.
 32. Marbach D, Roy S, Ay F, Meyer PE, Candeias R, Kahveci T, Bristow CA, Kellis M. Predictive regulatory models in *Drosophila melanogaster* by integrative inference of transcriptional networks. *Genome Res.* 2012; **22**(7):1334–49.
 33. Bonn S, Zinzen RP, Girardot C, Gustafson EH, Perez-Gonzalez A, Delhomme N, Ghavi-Helm Y, Wilczynski B, Riddell A, Furlong EE. Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat Genet.* 2012; **44**(2):148–56.
 34. Hurley D, Araki H, Tamada Y, Dunmore B, Sanders D, Humphreys S, Affara M, Imoto S, Yasuda K, Tomiyasu Y, Tashiro K, Savoie C, Cho V, Smith S, Kuhara S, Miyano S, Charnock-Jones DS, Crampin EJ, Print CG. Gene network inference and visualization tools for biologists: application to new human transcriptome datasets. *Nucleic Acids Res.* 2012; **40**(6):2377–98.
 35. Reece-Hoyes JS, Diallo A, Lajoie B, Kent A, Shrestha S, Kadrepapa S, Pesyna C, Dekker J, Myers CL, Walkout AJ. Enhanced yeast one-hybrid assays for high-throughput genecentered regulatory network mapping. *Nat Methods.* 2011; **8**(12):1059–64.
 36. Clarke BL. Stability of complex reaction networks. *Advances in Chemical Physics.* Eds S. A. Rice, I. Prigogine. New Yorks: John Wiley, 1980:1–215.
 37. Martseniuk OP, Romanets' KL, Obolens'ka Mlu, Huppertz B. Effect of homocysteine on the structure and functions of human placenta trophoblast. *Ukr Biokhim Zh.* 2009; **81**(5):40–9.
 38. Obolenskaya MYu, Rodrigues RR, Martsenyuk OP. Folate-related processes in human placenta: gene expression, aminothiols, proliferation and apoptosis. *Ukr Biokhim Zh.* 2011; **83**(1):5–17.
 39. Mislanova C, Martsenyuk O, Huppertz B, Obolenskaya M. Placental markers of folate-related metabolism in preeclampsia. *Reproduction.* 2011; **142**(3):467–76.
 40. Scotti M, Stella L, Shearer EJ, Stover PJ. Modeling cellular compartmentation in one-carbon metabolism. Wiley Interdiscip. *Rev Syst Biol Med.* 2013; **5**(3):343–65.
 41. Tibbetts AS, Appling DR. Compartmentalization of Mammalian folate-mediated one-carbon metabolism. *Annu Rev Nutr.* 2010; **30**:57–81.
 42. Stover PJ. Polymorphisms in 1-carbon metabolism, epigenetics and folate-related pathologies. *J Nutrigenet Nutrigenomics.* 2011; **4**(5):293–305.
 43. Forges T, Monnier-Barbarino P, Alberto JM, Gueant-Rodriguez RM, Daval JL, Gueant JL. Impact of folate and homocysteine metabolism on human reproductive health. *Hum Reprod Update.* 2007; **13**(3):225–38.
 44. Shenoy V, Kanasaki K, Kalluri R. Pre-eclampsia: connecting angiogenic and metabolic pathways. *Trends Endocrinol Metab.* 2010; **21**(9):529–36.
 45. Rodrihes RR, Lushchuk IS, Obolens'ka Mlu. Stoichiometric model of folate-dependent metabolism of one-carbon units in human placenta. *Ukr Biokhim Zh.* 2012; **84**(4):20–31.
 46. Schuster S, Hilgetag C. On elementary flux modes in biochemical reaction systems at steady state. *J Biol Syst.* 1994; **2**(2):165–82.
 47. Swanson DA, Liu ML, Baker PJ, Garrett L, Stitzel M, Wu J, Harris M, Banerjee R, Shane B, Brody LC. Targeted disruption of the methionine synthase gene in mice. *Mol Cell Biol.* 2001; **21**(4):1058–65.

Received 15.07.13