BGRS\SB-2018

The 11th International Conference

Bioinformatics of Genome Regulation and Structure\ Systems Biology

BGRS\SB-2018
NOVOSIBIRSK, RUSSIA
20–25 AUGUST, 2018

CONF.BIONET.NSC.RU/BGRSSB2018

Institute of Cytology and Genetics, Siberian Branch of Russian Academy of Sciences
Novosibirsk State University

BIOINFORMATICS OF GENOME REGULATION AND STRUCTURE\SYSTEMS BIOLOGY (BGRS\SB-2018)

The Eleventh International Conference

Abstracts

20–25 August, 2018
Novosibirsk, Russia

Novosibirsk
ICG SB RAS
2018

УДК 575
В60

Bioinformatics of Genome Regulation and Structure\Systems Biology (BGRS\SB-2018) :
The Eleventh International Conference (20–25 Aug. 2018, Novosibirsk, Russia); Abstracts /
Institute of Cytology and Genetics, Siberian Branch of Russian Academy of Sciences;
Novosibirsk State University. – Novosibirsk: ICG SB RAS, 2018. – 267 pp. –
ISBN 978-5-91291-036-4.

ISBN 978-5-91291-036-4

International program committee

- Nikolai Kolchanov, Institute of Cytology and Genetics of SB RAS, Novosibirsk, Russia (Conference Chair)
- Ralf Hofestädt, University of Bielefeld, Germany
- Mikhail Fedoruk, Novosibirsk State University, Novosibirsk, Russia
- Lubomir Aftanas, State Scientific-Research Institute of Physiology & Basic Medicine, Novosibirsk, Russia
- Ivo Grosse, Halle-Wittenberg University, Halle, Germany
- Olga Lavrik, Institute of Chemical Biology and Fundamental Medicine of SB RAS, Novosibirsk, Russia
- Dmitry Afonnikov, Institute of Cytology and Genetics of SB RAS, Novosibirsk, Russia
- Andrey Lisitsa, IBMC, Moscow, Russia
- Mikhail Voevoda, Institute of Internal Medicine and Preventive Medicine – ICG SB RAS, Russia
- Vladimir Ivanisenko, Institute of Cytology and Genetics of SB RAS, Novosibirsk, Russia
- Sergey Kabanikhin, Institute of Computational Mathematics and Mathematical Geophysics of SB RAS, Novosibirsk, Russia
- Olga Krebs, Heidelberg Institute for Theoretical Studies, Germany
- Inna Lavrik, Otto von Guericke University, Magdeburg, Germany
- Yuri Matushkin, Institute of Cytology and Genetics of SB RAS, Novosibirsk, Russia
- Evgeny Nikolaev, Institute of Energy Problems of Chemical Physics of RAS, Moscow, Russia
- Nikolay Podkolodny, Institute of Cytology and Genetics of SB RAS, Novosibirsk, Russia
- Mikhail Moshkin, Institute of Cytology and Genetics of SB RAS, Novosibirsk, Russia
- Leonid Moroz, University of Florida, Gainesville, USA
- Sergey Peltek, Institute of Cytology and Genetics of SB RAS, Novosibirsk, Russia
- Sergey Lashin, Institute of Cytology and Genetics of SB RAS, Novosibirsk, Russia
- Evgeny Rogaev, University of Massachusetts, USA
- Alexey Rodionov, Institute of Computational Mathematics and Mathematical Geophysics of SB RAS, Novosibirsk, Russia
- Elena Salina, Institute of Cytology and Genetics of SB RAS, Novosibirsk, Russia
- Natalya Kolosova, Institute of Cytology and Genetics of SB RAS, Novosibirsk, Russia
- Alexey Kochetov, Institute of Cytology and Genetics of SB RAS, Novosibirsk, Russia
- Yuri Orlov, Institute of Cytology and Genetics of SB RAS, Novosibirsk, Russia
- Vladimir Poroikov, Institute of Biomedical Chemistry, Moscow, Russia
- Maria Samsonova, St. Petersburg State Polytechnic University, St. Petersburg, Russia
- Denis Larkin, The Royal Veterinary College, London, United Kingdom
- Ancha Baranova, George Mason University, Fairfax, USA

Local organizing committee

- Svetlana Zubova, ICG SB RAS (Chairperson)
- Olga Petrovskaya, ICG SB RAS
- Tatyana Karamysheva, ICG SB RAS
- Erlan Tokpanov, ICG SB RAS
- Alina Morkovina, ICG SB RAS
- Larisa Eliseeva, ICG SB RAS
- Ksenia Strygina, ICG SB RAS
- Andrey Kharkevich, ICG SB RAS
- Nadezhda Glebova, ICG SB RAS
- Tatyana Chalkova, ICG SB RAS
- Ilya Akberdin, ICG SB RAS
- Aleksandr Ovsienko, NSU
- Anastasia Klimenok, NSU
- Grigoriy Khazankin, NSU Information Technology Department, IEEE R8 Russia Siberia Section

Contacts

Russian Federation
630090 Novosibirsk, Lavrentyeva, 10

Tel: +7(383) 363-49-80
Fax: +7(383) 333-12-78

URL - ICG SB RAS: www.bionet.nsc.ru
URL - BGRS\SB-2018: http://conf.bionet.nsc.ru/bgrssb2018/en/

Organizing committee: bgrs2018@icg.sbras.ru

Organizers

- Institute of Cytology and Genetics SB RAS (ICG SB RAS)
- Novosibirsk State University (NSU)
- The Federal Agency for Scientific Organizations (FASO Russia)
- Siberian Branch of the Russian Academy of Sciences (SB RAS)
- The Vavilov Society of Geneticists and Breeders
- Federal State Scientific Budgetary Institution "Scientific Research Institute of Physiology & Basic Medicine"(FSSBI "SRIPBM")
- Chair of Information Biology NSF NSU
- Limited liability company "Scientific service"
- Vavilov Journal of Genetics and Breeding

Sponsors

Russian Foundation for Basic Research Grant No. 18-04-20047

Federal Agency for Scientific Organizations FASO Russia

Ministry of Education and Science of the Russian Federation (Minobrnauki of Russia) Grant No. ДНИТ 28.12487.2018/12.1

GOLD SPONSORS

HUAWEI

intel Software

SILVER SPONSORS

MP Biomedicals

Albiogen

Roche Diagnostics Rus Ltd.

Skygen, Ltd

Bioline, Llc

Khimexpert Ltd.

DIA-M, Ltd

BASIC SPONSORS

geneXplain GmbH

IOS Press

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

Director: Corr. Member of the RAS Alexey V. Kochetov
Science Head: Professor, Full Member of the RAS Nikolay A. Kolchanov
Academic Secretary: Candidate of Sciences (Biology) Galina V. Orlova
Phone: +7(383) 363-49-85 ext. 1336; e-mail: OrlovaGV@icg.sbras.ru, gorlova@bionet.nsc.ru

The Institute was founded in 1957, among the first institutions of the Siberian Branch of the Russian Academy of Sciences. Presently, ICG SB RAS is an interdisciplinary biological center, which ranks among the leading biological institutions in Russia. The second step of the restructuring of the Federal Research Center Institute of Cytology and Genetics was completed in May 2017. Presently, ICG includes three branches:

- Siberian Research Institute for Plant Industry and Breeding,
- Institute of Clinical and Experimental Lymphology, and
- Institution of Internal and Preventive Medicine.

Tasks of ICG SB RAS: Solution of top-priority problems in the development of the Russian science and technology sector in the fields of plant genetics and breeding, animal genetics and breeding, human genetics, biotechnology, and fundamental medicine by applying methods of molecular genetics, cell biology, and computational biology.

Strategic goal: Integrated studies in plant genetics and breeding, animal genetics and breeding, human genetics, fundamental medicine, and biotechnology by applying methods of molecular genetics, cell biology, and computational biology from the generation of academic knowledge to the solution of top priority problems set by Russian agricultural, biotechnological, biomedical, and pharmaceutical industries.

Staff: As on January 1, 2018, ICG included 88 scientific units, which employed 1397 members, of them 492 researchers, 2 RAS Advisors, 8 Full Members of the RAS, 4 Corresponding Members of the RAS, 98 Doctors of Science, and 305 Candidates of Science.

Postgraduate education and residency training: As on January 1, 2018, ICG trains 87 postgraduates and 18 interns.

Publications: The Institute intensely contributes to Russian and foreign journals and ranks among acknowledged leaders in Russian biology. In 2017, 344 papers were enrolled in the international online service Web of Science (WoS). The overall number of publications in peer-reviewed journals exceeds 450. In 2017 papers of the researchers of the Institute received over 4719 WoS citations.

Auxiliaries: Core facility "Center for Genetic Resources of Laboratory Animals", which includes the unique research unit "SPF vivarium", and seven shared access centers www.bionet.nsc.ru/uslugi.

The Federal Research Center Institute of Cytology and Genetics is looking to cooperate with scientific and commercial enterprises.

Address: Prosp. Akad. Lavrentiev 10, Novosibirsk, 630090 Russia
Phone/fax: +7(383) 363-49-80 / +7(383) 333-12-78
URL: www.bionet.nsc.ru
e-mail: icg-adm@bionet.nsc.ru

*NSU IN FACTS & FIGURES

7000 STUDENTS

1400 FOREIGN STUDENTS from 57 countries

6 departments

3 institutes

2500 UNIVERSITY STAFF MEMBERS

880 ASSOCIATE PROFESSORS

570 FULL PROFESSORS with doctoral degrees

73 MEMBERS OF THE RAS

MORE THAN **80** COUNTRIES REPRESENTED BY MOOCs STUDENTS

141 PARTNER UNIVERSITIES in 24 countries

JOINT PROGRAMS WITH COMPANIES:

Baker Huges, OSCiAl, Intel, Yandex, 2GIS, Academpark, Biotechnopark, Sberbank Technologies

35 PARTNER INSTITUTES in Academgorodok

139 LEADING RESEARCHERS

H>20

RESEARCH GROUPS ARE INVOLVED IN NSU **38** INTERNATIONAL COLLABORATIONS in the field of elementary particle physics and life sciences

140 LABORATORIES

6 MIRROR LABORATORIES

5-100 LEADER OF TOP 5-100 PROJECT –
the project oriented at making leading Russian universities competitive on world market of education services

TOP-100 IN PHYSICS
QS University Rankings

TOP-500 ACADEMIC RANKING OF WORLD UNIVERSITIES
3rd position among Russian universities

20th POSITION
QS UNIVERSITY RANKINGS BRICS

2nd POSITION
QS UNIVERSITY RANKINGS EMERGING EUROPE AND CENTRAL ASIA

63th POSITION
QS UNIVERSITY RANKINGS NATURAL SCIENCES

190th POSITION
GOOGLE SCHOLAR CITATIONS

291st POSITION
QS UNIVERSITY RANKINGS
among 20 000 world universities

N*

Contents

Genomics, transcriptomics and bioinformatics

Proteomics and bioinformatics

Systems computational biology

Bioinformatics and systems biology of plants

Animal genetics

Systems biology of aging

GENOMICS, TRANSCRIPTOMICS AND BIOINFORMATICS

DOI 10.18699/BGRSSB-2018-001

Analysis of repetitive DNA sequences in genomes of *Porodaedalea niemelaei*, *P. chrysoloma* and *Armillaria borealis*

A. Aksyonova[1]*, Yu. Putintseva[1], N. Oreshkova[1, 2], I. Pavlov[1, 2], Y. Litovka[3], K. Krutovsky[1, 4, 5, 6]

[1] *Siberian Federal University, Krasnoyarsk, Russia*
[2] *V.N. Sukachev Institute of Forest SB RAS, Krasnoyarsk, Russia*
[3] *M.F. Reshetnov Siberian State Aerospace University, Department of Silviculture, Krasnoyarsk, Russia*
[4] *Georg-August University of Göttingen, Department of Forest Genetics and Forest Tree Breeding, Göttingen, Germany*
[5] *Vavilov Institute of General Genetics RAS, Moscow, Russia*
[6] *Texas A&M University, College Station, Texas, USA*
** e-mail: aksien96@mail.ru*

Key words: *Armillaria*, fungi, genomes, *Porodaedalea*, repetitive DNA, transposable elements, transposons

Motivation and Aim: We are interested in the evolution of the fungi genes and genomes. Repetitive DNA sequences may play significant role in determining the structure and evolution of genes and genomes. Therefore, it is very important to identify and study them in the fungi genomes, such as *Porodaedalea* and *Armillaria* that are among the most important fungi in boreal and temperate forests.
Methods and Algorithms: The *P. niemelaei* genome was sequenced and annotated using the PacBio technology by the US DoE Joint Genome Institute (JGI) in collaboration with Department of Forest Genetics and Forest Tree Breeding, Georg-August University of Göttingen, Germany and Laboratory of Forest Genomics, Siberian Federal University (SFU), Russia. The length of the full genomic assembly was 53 Mbp. The genome assembly of *P. chrysoloma* was obtained from the JGI fungal database (http://genome.jgi.doe.gov/programs/fungi/index.jsf). The length of the full genomic assembly was 44 Mbp. The *A. borealis* genome was sequenced and annotated in the Laboratory of Forest Genomics, SFU using Illumina HiSeq 2000. The length of the full genomic assembly was 66 million Mbp.
Results: In total, 161 highly repetitive elements (REs) including 127 unknown REs were found in the *P. niemelaei* genome, 122 REs including 94 unknown REs were found in the *P. chrysoloma* genome, and 886 REs including 835 unknown REs – in the *A. borealis* genome.
Conclusion: Known as well as mostly unknown repetitive elements were identified in the *Porodaedalea* and *Armillaria* genomes. Their study and comparison with other genomes of basidiomycetes will help us better understand the functional and evolutionary role of REs and to reveal their evolutionary relationships.
Acknowledgements: This study was funded by a research grant No. 14.Y26.31.0004 from the Government of the Russian Federation.

DOI 10.18699/BGRSSB-2018-002

Genome *de novo* sequencing, assembly and functional annotation of pathogenic fungi *Armillaria borealis*

V. Akulova[1]*, V. Sharov[1], Yu. Putintseva[1], N. Oreshkova[1, 2], S. Feranchuk[1, 3, 4], D. Kuzmin[1], I. Pavlov[1, 2], K. Krutovsky[1, 5, 6, 7]
[1] *Siberian Federal University, Krasnoyarsk, Russia*
[2] *V.N. Sukachev Institute of Forest SB RAS, Krasnoyarsk, Russia*
[3] *Irkutsk National Research Technical University, Irkutsk, Russia*
[4] *Limnological Institute SB RAS, Irkutsk, Russia*
[5] *Georg-August University of Göttingen, Göttingen, Germany*
[6] *Vavilov Institute of General Genetics RAS, Moscow, Russia*
[7] *Texas A&M University, College Station, Texas, USA*
** e-mail: vfedotova@sfu-kras.ru*

Key words: genome assembly, functional annotation, fungal pathogenicity

Motivation and Aim: The forest decline is observed almost everywhere as a result of negative anthropogenic and climatic effects, often aggravated by pests, fungi and other phytopathogens. The pollution, increased average annual temperature, decreased precipitation, more frequent droughts and other climatic extremes can weaken trees and make fungi much more destructive. The forest conservation has become a very serious problem, since the scale of tree death caused by the phytopathogenic fungi is enormous. *Armillaria borealis* (Marxm. & Korhonen) is a fungi from the Physalacriaceae family (Basidiomycota) widely distributed in Siberia and the Far East and is also causing the root rot disease that weakens and often kills woody plants. Our goal was to *de novo* sequence, assemble and characterize the genome of *Armillaria borealis* and to generate data that can be used to identify the fungi virulence factors.
Methods and Algorithms: The fungi material was collected from active mycelia of *A. borealis* taken from the *Abies sibirica* trees died in 2015. DNA was sequenced using the 250-bp insert paired-end libraries on the Illumina MiSeq platform at the Laboratory of Forest Genomics of the Siberian Federal University. A *de novo* genome assembly was performed using the SPAdes genome assembler. Protein coding regions were identified in the genome using Exonerate. The EVidenceModeler and Augustus software were used to predict genes using gene models. Finally, the functional annotation was done using predictions as well as protein and transcript alignments and assignments based on PFAM, InterPro and GO ontology.
Results: The *A. borealis* genome assembly contained ~79 Mbp and was comparable with 60 and 84 Mbp for the *A. ostoyae* and *A. gallica* genomes, respectively. The N50 for contigs equaled 15659 bp. Functional annotation revealed 6703 protein coding genes, which was also comparable with 7797 and 8261 in *A. ostoyae* and *A. gallica*, respectively, and provided important data for further comparative analysis.
Conclusion: We are currently reconstructing metabolic pathways of *Armillaria* core genes and pathogenicity. This genome study provides much needed knowledge regarding the woody plant fungal pathogenicity, and useful insights towards identifying specific genes associated with pathogenesis and other metabolic functions.
Acknowledgements: This study was supported by research grant No. 14.Y26.31.0004 from the Government of the Russian Federation.

DOI 10.18699/BGRSSB-2018-003

Proteomic analysis of affinity captured LINE-1 macromolecular complexes

I. Altukhov[1]*, M.S. Taylor[2], K.R. Molloy[3], P. Mita[4], H. Jiang[5], E.M. Adney[4, 6], A. Wudzinska[6], S. Badri[7], D. Ischenko[1], G. Eng[2], K.H. Burns[5, 8], D. Fenyö[4], B.T. Chait[3], D. Alexeev[9], M.P. Rout[5], J.D. Boeke[4], J. LaCava[4, 5]

[1] *Moscow Institute of Physics and Technology, Dolgoprudny, Russia*
[2] *Department of Pathology, Massachusetts General Hospital, Boston, United States*
[3] *Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, United States*
[4] *Department of Biochemistry and Molecular Pharmacology, Institute for Systems Genetics, NYU Langone Health, New York, United States*
[5] *Laboratory of Cellular and Structural Biology, The Rockefeller University, New York, United States*
[6] *McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, United States*
[7] *Department of Pathology, NYU Langone Health, New York, United States*
[8] *Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, United States*
[9] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: ilya.altukhov@gmail.com*

Key words: protein-protein interactions, protein complex, mass-spectrometry, LINE-1, retrotransposon, data analysis

Motivation and Aim: Long Interspersed Nuclear Element-1 (LINE-1, L1) is a mobile genetic element active in human genomes. L1-encoded ORF1 and ORF2 proteins bind L1 RNAs, forming ribonucleoproteins (RNPs). These RNPs interact with diverse host proteins, some repressive and others required for the L1 lifecycle. In a never-ending battle, our cells have been fighting to keep LINE-1 and its ancestors from replicating, and so evolved various defense mechanisms. Yet, LINE-1 has learned to circumvent these barriers, and continues to replicate and cause disease. Our understanding of these defenses and of how LINE-1 evades them is limited.

Methods and Algorithms: Using differential affinity purifications, quantitative mass spectrometry, and next generation RNA sequencing, we have characterized the proteins and nucleic acids associated with distinctive, enzymatically active L1 macromolecular complexes.

Results: We described a cytoplasmic intermediate that we hypothesize to be the canonical ORF1p/ORF2p/L1-RNA-containing RNP, and we described a nuclear population containing ORF2p, but lacking ORF1p, which likely contains host factors participating in target-primed reverse transcription. We additionally explored proteomes associated with catalytically-inactivated ORF2p point mutants and monitored the rates of protein exchange from L1 macromolecules in vitro. Taken together, our data support the existence of a variety of putative L1-related protein complexes.

Conclusion: Custom computer code written in the R programming language was used in the analysis of mass spectrometry and RNA sequencing data; it has been published at https://github.com/elifesciences-publications/altukhov-line-1.

DOI 10.18699/BGRSSB-2018-004

Comparative transcriptomics of the effects of prionization and inactivation of the Swi1 protein in *Saccharomyces cerevisiae*

K.S. Antonets[1, 2]*, S.F. Kliver[1, 2], D.E. Polev[1], A.R. Shuvalova[1], E.A. Andreeva[1, 3], S.G. Inge-Vechtomov[1, 3], A.A. Nizhnikov[1, 2]

[1] *Laboratory for Proteomics of Supra-Organismal Systems, All-Russia Research Institute for Agricultural Microbiology, Pushkin, St. Petersburg, Russia*

[2] *Department of Genetics and Biotechnology, St. Petersburg State University, St. Petersburg, Russia*

[3] *Vavilov Institute of General Genetics RAS, St. Petersburg, Russia*

** e-mail: kirantonez@gmail.com*

Key words: prion, amyloid, yeast, SWI1, transcriptome

Motivation and Aims: Prions are infectious, self-perpetuating conformational states of proteins. Most part of known prions were found in yeast *Saccharomyces cerevisiae.* Usually, prion state is associated with the formation of protein fibrils (amyloids) and considered to be equal to functional inactivation of the protein. Swi1 is a component of the key chromatin remodeling complex SWI/SNF of yeast and was found to form prion [*SWI*$^+$]. Inactivation of Swi1 affects many different processes in yeast cells and has the nonsense-suppression phenotype (growth on the media without adenine) in strains with the mutant variants of the *SUP35* gene. The goal of this study was to compare the effects of prionization and deletional inactivation of the Swi1 protein on the transcription of different genes.

Methods: The next-generation RNA sequencing (RNA-seq) of the yeast transcriptomes of the [*SWI*$^+$], [*swi*$^-$], and *swi1*Δ strains was performed using Illumina HiSeq 2500 platform. The expression levels of several genes were analyzed using quantitative real-time PCR.

Results: Using RNA-seq we compared transcriptome-wide effects of prionization and deletional inactivation of Swi1 and found significant differences. In particular, about 20 yeast genes that are downregulated in the *swi1*Δ strain, are upregulated in the [*SWI*$^+$] strain. In addition, we found that nonsense-suppression phenotype had also different mechanisms in the [*SWI*$^+$] and *swi1*Δ strains [1]. The deletion of *SWI1* leads to increased expression of the *ade1-14* mutant allele, while in the [*SWI*$^+$] strains nonsense-suppression is caused by downregulation of the *SUP45* gene encoding eRF1 release factor.

Conclusion: Prionization of Swi1 protein and deletion of *SWI1* have different effects on transcription of yeast genes and, in some cases, the consequences of prion formation are similar to "gain-of-function" mutation.

Acknowledgements: This study was supported by the Grant of the President of the Russian Federation (MK-3240.2017.4 to AAN) and the Russian Foundation for Basic Research (17-04-00816 to SGIV and 16-34-60153 to AAN). Sequencing of the transcriptomes was carried out in the Research Resource Centre "Biobank" of Research Park of SPbSU.

References

1. Antonets K.S. et al. (2017) Distinct mechanisms of phenotypic effects of inactivation and prionization of Swi1 protein in *Saccharomyces cerevisiae.* Biochemistry (Mosk.). 82(10):1147-1157.

DOI 10.18699/BGRSSB-2018-005

Bioinformatics approach for prediction of metabolic capabilities for synthesis of essential vitamins and amino acids in human gut microbiome

G.A. Ashniev*, A.A. Arzamasov, S.N. Iablokov, D.A. Rodionov
Institute for Information Transmission Problems RAS, Moscow, Russia
** e-mail: escobar.morente@gmail.com*

Key words: amino acids, vitamins, metagenomics, human, gut, bacteria, biosynthesis, metabolomics

Introduction: Bacteria are abundantly presented in almost every part of the world, living in sundry communities with complex interactions and various phenotype features. As soon as cooperation and rivalry within microbial communities is well known as central part of their stability, continuity and longevity, there is lack of knowledge about the general principles of vitamins amino acids metabolism, particularly biosynthesis. Understanding of basic features of amino acid formation is key for inner-community regulation and its influence on host organism.
Results: We used the metabolic subsystem approach implemented in SEED genomic platform to reconstruct essential biosynthetic pathways in bacteria that inhabit human gastrointestinal tract. We analyzed metabolic pathways for synthesis of 9 vitamins and 20 amino acids in 2228 human gut bacteria with sequenced genomes and inferred their vitamin and amino acid phenotypes (prototrophy or auxotrophy). We also analyzed genomic distributions of known uptake transporters for amino acids and vitamins. The predicted pathways allowed us to classify the studied organisms with respect in both to their biosynthetic and transport capabilities. The obtained metabolic phenotypes in reference genomes were applied to human stool samples from different published metagenomics. A unique approach for taxonomy-based mapping allowed us to compute the cumulative amino acid and vitamin biosynthetic phenotypes of microbial samples from a large number of metagenomics studies.
Conclusion: The studied bacteria showed high level of conservation of amino acid and vitamin biosynthesis phenotypes on the taxonomic level of species. Incomplete biosynthesis pathways for some intermediates suggest certain deficiencies could be alternatively supplemented by their metabolic precursors. Vitamin auxotrophic phenotypes are much more prevalent in the human gut microbiota, whereas the larger number of studied bacteria is capable of *de novo* synthesis of all 20 amino acids. Histidine and tryptophan have the most variable phenotypes. The comparison of metabolic profiles of gut microbiota between various population groups revealed applicability of this approach to classify microbiota and predict the functional consequences of dietary changes in composition of gut microbiota.
Acknowledgments: This research was supported by the Russian Science Foundation (grant No. 14-14-00289).

DOI 10.18699/BGRSSB-2018-006

-Omics approaches to help decipher molecular control of root biotic interactions in the model legume *Medicago truncatula*

C. Ben[1]*, M. Toueni[1], M. Rickauer[1], MIRMED consortium (PIs : M. Crespi[2], L. Gentzbitte[1], J. Gouzy[3], A. Niebel[3], C. Roux[4]), L. Gentzbitte[1]

[1] *EcoLab, Université de Toulouse, CNRS, Toulouse INP, UPS, Toulouse, France*

[2] *IPS2, Orsay, France*

[3] *LIPM, Toulouse, France*

[4] *LRSV, Toulouse, France*

** e-mail: ben@ensat.fr*

Key words: transcriptomics, RNAseq, microRNA, differentially expressed genes (DEGs), gene co-expression network (WGCNA), Gene ontology (GO) enrichment, symbiosis, root diseases

Motivation and Aim: Legume roots undergo various abiotic and biotic constraints. A '-omics' high-throughput analysis was conducted to shed light on the molecular mechanisms regulating root plasticity in response to biotic stimuli in the legume model *Medicago truncatula*.

Methods and Algorithms: Next-generation sequencing technologies (NGS) were used to assess microRNA and gene expression in roots challenged with various symbiotic and pathogenic interactions. Differentially expressed miRNAs and genes were detected using the DESeq and edgeR packages of the R/Bioconductor statistical language. miRNA targets were predicted *in silico* with miRanda 3.3a customized to consider criteria available for plant miRNA targets. To identify relevant regulatory pathways involved in biotic interactions and symbiotic signals, gene co-expression network (using WGCNA R package) and Gene ontology (GO) enrichment analyses were conducted.

Results: MicroRNA regulation of the root response to different pathogenic (*Verticillium alfalfae* and *Ralstonia solanacearum*) and symbiotic (*Sinorhizobium meliloti*, *Rhizophagus irregularis* and Nod and Myc-LCO treatments) interactions were compared (MIRMED project, [1]). Specific or, conversely, common miRNAs to bacterial and fungal pathogens or symbionts have been detected. Connections in symbiotic signaling by Nod and Myc-LCO factors have also been highlighted. Regulatory network analyzes revealed co-expressed miRNAs modules that each regulate distinct sets of targets involved in different cellular functions. Molecular control of the set-up of quantitative resistance to verticillium wilt in *M. truncatula* was analyzed in more detail by transcriptomic analysis comparing the genes expressed in the roots of a resistant (A17) and susceptible (F83005.5) line at a key stage of the infective process (up to 24 h after inoculation) [2]. The results suggest that resistance in the A17 line of *M. truncatula* may be due to innate immunity combining preformed defense mechanisms and others triggered by PAMP (Pathogen Associated Molecular Patterns) and involves only a small number of induced genes. Weighted Gene-coexpression network analysis revealed five major regulatory modules in the resistant line. One module, that is significantly associated with the response to inoculation in A17, contains the majority of differentially expressed genes, and genes associated with PAMP perception and hormonal signaling as well as transcription factors. *In silico* analysis has shown that many of these genes also respond to other telluric pathogens in *M. truncatula*, suggesting a common core of transcriptional response to root pathogens.

Conclusion: The combination of '-omics' studies have revealed regulatory factors (miRNA and protein-encoding genes) potentially playing key roles in the control of pathogenic and symbiotic interactions. These candidates, which may be promising breeding targets, are undergoing functional validation by mutant analysis and genetic transformation.

Acknowledgements: This work was supported by a grant from the MIRMED project (Genoscope, CNRS). M. Toueni was supported by a doctoral grant from Mayotte department.

References

1. Formey D. et al. (2014) Diversity, conservation and plasticity of the miRNAome from *Medicago truncatula* roots under symbiotic and pathogenic interactions. Genome Biology. 15:457.
2. Toueni M. et al. (2016) Quantitative resistance to *Verticillium* wilt in *Medicago truncatula* involves eradication of the fungus from roots and is associated with transcriptional responses related to innate immunity. Front. Plant Sci. http://dx.doi.org/10.3389/fpls.2016.01431

DOI 10.18699/BGRSSB-2018-007

Identification of targets genes for miRNAs of the pathogenic fungus *Fusarium oxysporum* in a *de novo* transcriptome assembly of the Siberian larch (*Larix sibirica* Ledeb.)

V. Biryukov[1]*, N. Oreshkova[1, 2], Yu. Putintseva[1], K. Krutovsky[1, 3, 4, 5]
[1] *Siberian Federal University, Krasnoyarsk, Russia*
[2] *V.N. Sukachev Institute of Forest SB RAS, Krasnoyarsk, Russia*
[3] *Georg-August University of Göttingen, Göttingen, Germany*
[4] *Vavilov Institute of General Genetics RAS, Moscow, Russia*
[5] *Texas A&M University, College Station, Texas, USA*
** e-mail: vladislav.v.biriukov@gmail.com*

Key words: *Fusarium oxysporum*, fungi pathogenicity, miRNA, Siberian larch, targets genes, transcriptome

Motivation and Aim: MicroRNAs (miRNAs) play important regulatory roles in animals and plants and can regulate metabolic pathways related to ontogenesis, cell differentiation and response to environmental factors, such as stress [1]. In addition, there is evidence that miRNAs play a particular role in cross-species communication due to miRNA transmission [2]. We studied whether miRNAs of a pathogenic fungus *Fusarium oxysporum* could potentially regulate host's (Siberian larch) biotic stress response pathways through targets genes in the host genome.
Methods and Algorithms: We used the Siberian larch (*Larix sibirica*) transcriptome that was originally *de novo* sequenced and assembled in the Laboratory of Forest Genomics at Siberian Federal University. The sequence data of *L. sibirica* was obtained using the Illumina HiSeq 2000 sequencer. *De novo* assembly was done using the Trinity assembler. To predict potential targets in the Siberian larch transcriptome for previously published miRNAs of *F. oxysporum* a specially developed plant small RNA target analysis server (psRNATarget) was used [3, 4].
Results: 576 target sequences were detected in total transcriptome of Siberian larch. Annotation of the targets showed that the most matches are associated with leucine-rich and pentatricopeptide repeat domains. Targets with glutathione S-transferase, RNA-recognition and ribonuclease H activities were also found.
Conclusion: Based on our own and published data we suggest that the detected targets can be involved in complex plant defense responses to pathogens. Thereby, miRNAs produced by fungi can weaken the defense mechanisms of the host tree. These results may be also of particular interest in further fungi pathogenicity studies.
Acknowledgements: This study was funded by a research grant No. 14.Y26.31.0004 from the Government of Russian Federation.

References

1. Bartel D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 116(2):281-297.
2. Liang H., Zen K., Zhang J., Zhang C.Y., Chen X. (2013) New roles for microRNAs in cross-species communication. RNA Biol. 10(3):367-370.
3. Chen R., Jiang N., Jiang Q., Sun X., Wang Y., Zhang H., Hu Z. (2014) Exploring microRNA-like small RNAs in the filamentous fungus *Fusarium oxysporum.* PLoS One. 9(8):e104956.
4. Dai X., Zhao P.X. (2011) psRNATarget: a plant small RNA target analysis server. Nucleic Acids Res. 39(Suppl. Iss.2):W155-W159.

DOI 10.18699/BGRSSB-2018-008

Genome rearrangements in bacterial genomes

O.O. Bochkareva[1]*, M.S. Gelfand[1, 2, 3, 4]
[1] *Institute of Information Transmission Problem RAS, Moscow, Russia*
[2] *Lomonosov State University, Moscow, Russia*
[3] *Higher School of Economics, Moscow, Russia*
[4] *Skolkovo Institute of Science and Technology, Moscow, Russia*
** e-mail: olga.bochkaryova@gmail.com*

Key words: bacterial evolution, genome rearrangements, phylogeny reconstruction

Motivation: Gene order in prokaryotes is relatively poorly conserved making it a convenient tool for the analysis of the species and strain evolution, when changes in protein, and even gene sequences do not provide sufficient resolution [1]. In addition, genome rearrangements are less sensitive to homologous recombination and hence allow for an alternative approach to construction of phylogenetic trees, as even a small number of genome rearrangements may resolve topological ambiguities in a phylogenetic tree [2].
Results: For several bacterial species we reconstructed phylogenetic trees based on sequences similarities and gene order. Our results show that rearrangement rates differ dramatically in different bacterial species, that is likely to be related to the adaptation to different ecological niches. For newly formed pathogens such as *Y. pestis* and *B. mallei* with a particularly high rate of rearrangements we revealed the correlation between mutations rates and inversions rates. Analysis of contradictions between the obtained evolutionary trees yielded numerous parallel rearrangements. Numerous gene losses likely have been caused by a high rate of intragenomic recombination between limited number of repeated elements such as transposases and 16S-23S rRNA clusters, that also creates the plasticity of gene order in chromosomes. We revealed cases when inversions found in separated branches in the phylogenetic trees may result in phase variation. Also analysis of syntenies positions allowed us to reveal gene cassettes that are spreading horizontally such as iron uptake genes cluster in *B. cepacia* group.
Conclusion: Our data indicate that an integrated analysis of sequence-based and inversion-based trees enhances the resolution of phylogenetic reconstruction. At that, inversions may resolve branches with low bootstrap support; on the other hand, sequence analysis may distinguish between parallel inversions and single inversion propagated by homologous recombination. The reconstructed patterns of rearrangements indicate strong avoidance of large inter-replichore inversions and translocations that is likely to be caused by selection against transfer of large fractions of genes between the leading and the lagging strands.
Acknowledgements: Supported by the Russian Science Foundation under grant 14-50-00150.

References

1. Wolf Y.I., Rogozin I.B., Kondrashov A.S., Koonin E.V. (2001). Genome alignment, evolution of prokaryotic genome organization, and prediction of gene function using genomic context. Genome Res. 11(3):356-72.
2. Darling A.E., Miklós I., Ragan M.A. (2008). Dynamics of genome rearrangement in bacterial populations. PLoS Genet. 4:e1000128.

DOI 10.18699/BGRSSB-2018-009

How to analyze transcriptome of each of 800 000 cells, using automatized library prep for Illumina from 10× Genomics (technology overview)

K. Bolaeva
SkyGen LLC
e-mail: kermenbolaeva@gmail.com

Key words: transcriptome of single cells, UMI, unique molecular barcodes, gel bead technology, RNA-seq on Illumina platform, microfluidics

Motivation and Aim: Nowadays plenty of new approaches for library prep for Illumina sequencing are developed every day. As well, scientists are more and more interested, what is going on in a different types of tissue or in a tumor microenvironment. So, in this technology review we are to describe a protocol from 10× Genomics automated library preparation, which enables to mark cDNA from 80000 single cells using microfluidics, gel beads coated with millions of oligos and unique molecular barcoding technology (UMI). Currently scientists know that in a tissue there are different types of cells, each of them has it's own function. Great example for that is a brain tissue. It was shown, that 10× Genomics approach got it possible to analyze more than 1 million of cells of a rat brain, and divide all cells into different types based on their expression profile. As well, in a tumor tissue we also have healthy cells, tumor cells and it could be interesting to see and quantify ratio of them. Another amazing example of the technology application is investigation of blood cell, and especially of peripheral blood mononuclear cells (PBMC). PBMC cells were divided by different types of gene expression, and therefore different *Methods and Algorithms*: One of the main advantages of the technology discussed here is, that due to automatization using Chromium Controller from 10× Genomics, we can mark cDNA from 800000 cells in 7 minutes, allowing to distinguish transcriptome data from every singe cell finally. It is possible to analyze thousands of single cells in every run. 10× technology suite performs millions of parallel reactions to enable gene expression profiling at scale with single cell resolution. It is crucial to have suspension of alive eukaryotic cells for the experiment. Then suspension of cells, thousands gel beads coated with millions of oligos (oligos from each gel bead are unique, as well as oligos within one gel bead are with UMI), and oil are put into a microfluidic chip of Chromium Controller machine. Cells and beads under pressure are migrating in a microfluidic channels, where they are mixed with oil so, that finally re get drops in the oil. In each drop we expect to have one gel bead and one cell. Cell is to lysates within the drop, and mRNA molecules are attaching to poly(T) part of oligos immobilized on the surface of gel beads. After that, gel beads are also dissolved. After reverse transcription reaction, molecules of cDNA has on their ends UMI, 10× Genomics barcode, and Illumina P5 and P7 primers. Ready cDNA can then be estimated, and then used for the Illumina library prep. Finally it was described that around 20000 reads can be obtained per cell.

References

1. Wang J. et al. (2018) Gene expression distribution deconvolution in single-cell RNA sequencing. Proc Natl Acad Sci USA.
2. Lambrechts D. et al. (2018) Phenotype molding of stromal cells in the lung tumor microenvironment. Nat Med.
3. Griffiths J.A. et al. (2018) Detection and removal of barcode swapping in single-cell RNA-seq data. Nat Commun.

DOI 10.18699/BGRSSB-2018-010

Minor variation in the 3' downstream region of *eGFP* reporter gene substantially increases level of its expression in mouse and human cells

L.V. Boldyreva[1]*, L.A. Yarinich[1, 2], E.N. Kozhevnikova[1], A.V. Pindyurin[1, 2]
[1] *Institute of Molecular and Cellular Biology SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: asd@mcb.nsc.ru*

Key words: transcription termination, mRNA processing, pre-mRNA cleavage, *eGFP* reporter expression

Motivation and Aim: The process of transcription termination appears to be complex and provides one of the levels of gene regulation [1]. We explored the influence of transcription termination defect on the level of *eGFP* reporter mRNA in transiently transfected mouse, human and *Drosophila* cell cultures.
Methods and Algorithms: We generated "wild type" and mutant double-reporter plasmids encoding mCherry (inner reference) and eGFP (reporter) fluorescent proteins. In the mutant constructs, a point deletion was introduced shortly (32 nt) downstream of the AAUAAA polyadenylation signal (PAS) of the *eGFP* reporter. Using RT-qPCR and FACS analyses, we determined the reporter mRNA and protein levels in transiently transfected human HEK293T, mouse C57BL/3T3-LCD, and *Drosophila* Kc167 cultured cells. We obtained the C57BL/3T3-LCD cell line by immortalization of primary mouse embryonic fibroblasts isolated from C57BL/6J embryos according to [2]. 3′-RACE method [3] was used to identify the 3′ end(s) of the mature *eGFP* mRNA molecules in HEK293T cells.
Results: We found that the one-nucleotide deletion introduced shortly downstream of the PAS causes up to 4-fold increase in both mRNA and protein production in transiently transfected human and mouse, but not in *Drosophila* cells. This deletion of a single C 32 nt downstream of the PAS leads to twice more frequent cleavage of pre-mRNA molecules within 14 nt downstream of the PAS (85 % of cases vs 43 % in the control), whereas the wild-type pre-mRNA molecules are typically cleaved more distally: at 31 nt downstream of the PAS and a number of other sites.
Conclusion: Even a small change in the region immediately downstream of 3′ UTR, which is not present in the mature mRNA, can substantially affect the expression level of the upstream gene. The results indicate the great regulatory potential of the 3′ downstream region, which also may be applied in biotechnology.
Acknowledgements: Supported by the Russian Science Foundation grant No. 16-14-10288. We thank S.E. Kulemzin and E.S. Reshetnikova from the Laboratory of immunogenetics of IMCB SB RAS for help in FACS data acquiring.

References

1. Neve J. et al. (2017) Cleavage and polyadenylation: Ending the message expands gene regulation. RNA Biology. 14(7):865-890.
2. Xu J. (2005) Preparation, culture, and immortalization of mouse embryonic fibroblasts. Curr. Protoc. Mol. Biol. Ch.28:Unit 28.1.
3. Scotto-Lavino E. et al. (2006) 3′ end cDNA amplification using classic RACE. Nat. Protoc. 1(6):2742-2745.

DOI 10.18699/BGRSSB-2018-228

Profiling neuronal epigenomes in a mouse model of alcohol use disorder

C.R. Bridges, O. Ponomareva, C.T. Tulisiak, T. Larina, R.A. Harris, I. Ponomarev*
Waggoner Center for Alcohol and Addiction Research, The University of Texas, Austin, Texas, USA
** e-mail: ponomarev@utexas.edu*

Key words: alcoholism, neuroscience, epigenomics, INTACT, ATAC-Seq

Motivation and Aim: Alcohol use disorder (AUD) is characterized by widespread changes in gene expression in humans and animal models. Mounting evidence points to a central role of chromatin (epigenomic) modifications in controlling gene expression and behavior in AUD, but there is a critical gap in our knowledge regarding AUD-associated epigenomic changes in specific neuronal populations in the addiction neurocircuits. Here, we tested two recently developed methods: INTACT (Isolation of Nuclei TAgged in specific Cell Types) and ATAC-Seq (Assay for Transposase-Accessible Chromatin using Sequencing) to obtain epigenomic profiles from limited number of cells [1]. INTACT is based on the Cre-loxP system in mice to express a tagged nuclear membrane protein, allowing affinity purification of tagged nuclei from genetically defined cell populations. ATAC-Seq provides a comprehensive description of the epigenomic state in the cell, including chromatin accessibility (open chromatin), nucleosome positioning and occupancy of DNA binding proteins. The aim of this study was to examine the effects of alcohol on the epigenome of glutamatergic neurons in the prefrontal cortex, a key structure with a role in the preoccupation/anticipation stage of AUD.
Methods and Algorithms: Alcohol dependence was induced by ethanol vapor inhalation [2] in genetically modified mice expressing green fluorescent protein (GFP) in *Camk2a*-positive cells. Twenty four hours after ethanol was removed brains were dissected for molecular profiling. We used INTACT to isolate nuclei from glutamatergic cortical neurons and ATAC-Seq to obtain genome-wide distributions of open chromatin marks.
Results: We achieved >95 % specificity and up to 70 % yield of GFP+ neurons. We tested ATAC-Seq using various numbers of nuclei and were able to obtain reliable epigenomic profiles from as few as 5,000 nuclei. Statistical and bioinformatics analysis are underway to identify genes and gene networks affected by alcohol.
Conclusion: Identifying alcohol-induced, cell type-specific molecular changes is critical to our understanding of the roles of individual neuronal populations in alcohol actions, and a combination of INTACT and ATAC-Seq will be instrumental in studying neuronal epigenomes in a brain region – and neural circuit – specific manner.
Acknowledgements: Supported by NIAAA grants AA024586 and Bruce/Jones Graduate Fellowship in Addiction Biology.

References

1. Mo A. et al. (2015) Epigenomic signatures of neuronal diversity in the mammalian brain. Neuron. 86:1369-1384.
2. Lopez M.F., Becker H.C. (2005) Effect of pattern and number of chronic ethanol exposures on subsequent voluntary ethanol intake in C57BL/6J mice. Psychopharmacol. 181:688-696.

DOI 10.18699/BGRSSB-2018-011

Study of bioresources of salt lakes of Novosibirsk oblast

A. Bryanskaya[1]*, Yu. Uvarova[1], A. Rozanov[1], E. Lazareva[2], O. Taran[3], T. Ivanisenko[1], S. Peltek[1]

[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Institute of Geology and Mineralogy SB RAS, Novosibirsk, Russia*
[3] *Boreskov Institute of Catalysis SB RAS, Novosibirsk, Russia*
** e-mail: alla@bionet.nsc.ru*

Key words: methagenomic, biodiversity of microorganisms, salt lakes, microbial communities

Motivation and Aim: In the Kulunda and Baraba Steppes there are many brackish and saline lakes. These lakes are unstable water systems, characterized by significant fluctuations in the water level so the microorganism communities developing in them have various adaptations that allow them to withstand significant fluctuations in environmental factors. The aim of this study was to investigate the fundamentals of biogeochemistry and microbiology of extreme environments using small saline drainless lakes of the Novosibirsk oblast.
Methods and Algorithms: As a model object we took the Solenoye Lake (Novosibirsk oblast). We studied the composition of the benthic microbial community using parallel amplicon sequencing of the V3 variable region of the 16S rRNA gene, as well as geochemical parameters (pH, Eh, cationic/anionic composition, gas components). We described and compared different layers of the bottom sediments and found relations between chemical composition and microbial content of the bottom sediments.
Results: As a result, it has been established that geochemical data on the study of the structure of the microbial community are well correlated. The upper 30–35 cm of the sediment by the ratio of the main groups of microorganisms are somewhat similar, and below, where the change occurs and the main geochemical and mineralogical indices of the substance, the composition of the community changes significantly. Throughout the sediment column, Actinobacteria and/or Proteobacteria predominated. In the layers of the columns the number of prokaryotic sequences found was from 96 to 414. The proportion of archaeal sequences in the bottom sediments was 5.5–6.6 %, in water and microbial mathematics it was 2.0 and 1.2 %. In the lake water 50 prokaryotic sequences were identified, in the microbial mate – 81, in the columns of bottom sediments – 390. In water and microbial lake mate, Cyanobacteria and Proteobacteria dominated. It is established that the ecosystem of the Salt Lake is characterized by: (1) variations in geochemical parameters; (2) developing mass of microbial community with prevalence of cyanobacteria; (3) the presence of powerful layers of bottom sediments with a high taxonomic diversity of prokaryotes; (4) intensive development of the Artemia crab; (5) tracking changes in biological and geochemical characteristics in dynamics and extrapolating the data obtained to other lake ecosystems in the Novosibirsk region, characterized by a similar geochemical composition of natural solutions [1, 2].
Acknowledgements: Supported by the RFBR (17-44-540815), by the Government of the Novosibirsk Region and by Budget Project of the IC&G SB RAS (VI.58.1.8).

References

1. Lazareva E.V. et al. (2012) Research of elements distribution between components of salt lake system by SR'XRF technique. Poverkhnost = J. Surface Investigation: X-Ray, Synchrotron and Neutron Techniques. 12:70-80. (in Russian)
2. Bryanskaya A.V. et al. (2016). BMC Microbiology. 16(1):S4.3.

DOI 10.18699/BGRSSB-2018-012

Candidate SNP markers of social dominance, which may affect the affinity of the TATA-binding protein for human gene promoters

I.V. Chadaeva[1, 2], D.A. Rasskazov[1], E.B. Sharypova[1], L.K. Savinkova[1], P.M. Ponomarenko[3], M.P. Ponomarenko[1, 2]*

[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

[2] *Novosibirsk State University, Novosibirsk, Russia*

[3] *University of La Verne, La Verne, California, USA*

** e-mail: pon@bionet.nsc.ru*

Key words: TBP, TATA-box, promoter, TBP-promoter affinity, gene, SNP, change in gene expression, significance, SNP marker, dominant, subordinate

Motivation and Aim: The purpose of this work is a prognosis based on the reference human genome of candidate SNP markers of predisposition to social dominance and its absence (which determines the subordinate (subordinate) rank of the individual). To achieve it, we proposed a hypothesis about the possibility of such a prediction on the basis of taking into account physiological markers of animal aggressiveness.

Methods and Algorithms: To achieve it, we proposed a hypothesis about the possibility of such a prediction on the basis of taking into account physiological markers of animal aggressiveness. Within the framework of this hypothesis, we applied the Web service [1] to the analysis of human genes homologous to animal genes for a wide range of protein marker functions associated with aggressive behavior: hormones, biosynthetic enzymes and neurotransmitter receptors, transcriptional and neurotropic factors.

Results: As a result, we found 92 markers of predisposition to domination and submission in people, which can be actual information when planning experimental computer studies of human social behavior.

Conclusion: We uncovered a number of SNPs that can affect the affinity of the TBP to human gene promoters. These genes are homologous to the ones found in animal genes associated with dominant and subordinate behavior [2]. We propose that these candidate SNP markers may be linked to social dominance in humans. As a computational prediction, our result requires further experimental verification.

References

1. Ponomarenko M. et al. (2015). How to use SNP_TATA_Comparator to find a significant change in gene expression caused by the regulatory SNP of this gene's promoter via a change in affinity of the TATA-binding protein for this promoter. Biomed. Res. Int. 2015:359835. DOI 10.1155/2015/359835.
2. Chadaeva I. et al. (2017) Candidate SNP markers of social dominance, which may affect the affinity of the tata-binding protein for human gene promoters. Russ. J. Genet.: Appl. Res. 7(5):523-537.

DOI 10.18699/BGRSSB-2018-013

Adaptation and selective constraint throughout the *Drosophila melanogaster* life-cycle

M. Coronado-Zamora[1], I. Salvador-Martínez[2, 3], D. Castellano[4], I. Salazar-Ciudad[1], A. Barbadilla[1]*

[1] *Genomics, Bioinformatics and Evolution. Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Spain*

[2] *EvoDevo Helsinki Community, Centre of Excellence in Experimental and Computational Developmental Biology, Institute of Biotechnology, University of Helsinki, Helsinki, Finland*

[3] *Department of Genetics, Evolution and Environment, University College London, London, UK*

[4] *Bioinformatics Research Center, Aarhus University, Aarhus C, Denmark*

** e-mail: antonio.barbadilla@uab.es*

Key words: adaptation, selective constraint, population genomics, transcriptomics, evo-devo

Motivation and Aim: Numerous studies have tried to test whether embryonic development for morphological, gene expression and genomic metrics follows the von Baer [1] or the hourglass [2] hypotheses of embryonic developmental evolution. Here we integrate population genomics data with the developmental transcriptome to measure both adaptation and selective constraint over the embryonic development and life cycle of the species *Drosophila melanogaster*.

Methods: We gather gene expression data for each temporal stage from the modENCODE database. We use three different methods combining genome polymorphism data in *D. melanogaster* and divergence data between *D. melanogaster* and *D. yakuba*. In addition, for each developmental stage, we estimate gene features such as codon usage bias, intron length, expression bias, number of exons and number of transcripts.

Results: (i) Genes expressed in the pupa and adult male stages exhibit the highest levels of adaptive substitutions. (ii) Genes expressed in mid and late embryonic development stages show the highest sequence conservation and the most complex structure: they are larger, with more exons, more transcripts and longer introns. This gene structure complexity may account for the observed conservation. (iii) Earliest stages of embryonic development are the most divergent, but in contrast with the adaptive explanation of pupa and adult stages, this increased divergence seems to be due to the diminished efficiency of natural selection on maternal genes and the simpler gene structure in these stages.

Conclusion: Considered over the whole life-cycle, *D. melanogaster* seems to fit the hour-glass model of embryonic development evolution.

Acknowledgements: Supported by the Spanish Ministerio de Economía, Industria y Competitividad [Project MINECO CGL2017-89160-P] to AB, the Finnish Academy to IS-C and the Generalitat de Catalunya grant 2017SGR 1379.

References

1. von Baer K.E. (1828). Über Entwicklungsgeschichte der Thiere. Beobachtung und Reflexion. Königsberg, Bei den Gebrüdern Borntträger.
2. Medawar P.B. (1954). The significance of inductive relationships in the development of vertebrates. J. Embryol. Exp. Morphol. 2:172-174.

DOI 10.18699/BGRSSB-2018-014

Flax (*Linum usitatissimum* L.) response to non-optimal acidity and zinc deficiency

A. Dmitriev[1]*, G. Krasnov[1], A. Zyablitsin[1], T. Rozhmina[1, 2], R. Novakovskiy[1], P. Kezimana[1], A. Kudryavtseva[1], N. Melnikova[1]
[1] *Engelhardt Institute of Molecular Biology RAS, Moscow, Russia*
[2] *All-Russian Research Institute for Flax, Torzhok, Russia*
** e-mail: Alex_245@mail.ru*

Key words: flax, *Linum usitatissimum,* zinc deficiency, stress response, high-throughput sequencing

Motivation and Aim: Flax (*Linum usitatissimum* L.) is grown for fiber and seed production. Unfavorable environments, such as nutrient deficiency and non-optimal soil acidity, decrease quantity and quality of yield. Cultivation of resistant to stress varieties can significantly reduce the crop losses. Understanding the mechanisms of flax response to the stresses and identification of resistance gene candidates will help in breeding of improved cultivars. In the present work, response of flax plants to increased pH level and zinc (Zn) deficiency was studied.
Methods and Algorithms: About 200 plants of flax cultivars, Norlin (resistant) and Mogilevsky (sensitive), were grown under control conditions, Zn deficiency (1000-fold reduced Zn content), increased pH level (7.5), and both Zn deficiency and pH 7.5. Total RNA was isolated from root tips. Illumina TruSeq Stranded Total RNA Sample Prep Kit was used for library preparation. The cDNA libraries were sequenced on Illumina NextSeq500 with 80-nucleotide read length. Transcriptome assembly, annotation, and expression analysis was performed using Trinity, Trinotate, TransDecoder, Bowtie2, and RSEM as described earlier [1].
Results: High-throughput sequencing of 16 cDNA libraries of flax cultivars, grown under control conditions, increased pH level, Zn deficiency, and both stresses simultaneously, was performed, and about 40 million reads were obtained for each experiment type. At pH 7.5, upregulation of stress-related genes and downregulation of genes, associated with the biogenesis of the cell wall, were revealed in flax plants. Under Zn deficiency, expression alterations were identified for genes involved in transmembrane transport and photosynthesis. Besides, distinct expression changes were revealed for flax genotypes with diverse resistance to studied stresses.
Conclusion: We identified genes with expression alterations in flax under non-optimal acidity and Zn deficiency. These genes are involved in diverse processes, including cell wall biogenesis, transmembrane transport, and photosynthesis, and could play an important role in flax response to the studied stresses. Moreover, genes with distinct expression changes between resistant and sensitive genotypes could determine the mechanisms of flax resistance to non-optimal acidity and Zn deficiency.
Acknowledgements: This work was financially supported by the Russian Science Foundation, grant 16-16-00114.

References

1. Dmitriev A.A. et al. (2016) Glutathione S-transferases and UDP-glycosyltransferases are involved in response to aluminum stress in flax. Front. Plant Sci. 7:1920. DOI 10.3389/fpls.2016.01920.

DOI 10.18699/BGRSSB-2018-015

The CpG island methylator phenotype (CIMP) in colorectal cancer is associated with energy metabolism alterations

M. Fedorova[1]*, E. Lukyanova[1], A. Snezhkina[1], G. Krasnov[1], S. Kharitonov[1], A. Dmitriev[1], N. Melnikova[1], E. Pudova[1], M. Kiseleva[2], M. Chernichenko[2], D. Sidorov[2], A. Kudryavtseva[1]

[1] *Engelhardt Institute of Molecular Biology RAS, Moscow, Russia*

[2] *National Medical Research Radiological Center, Ministry of Health of the Russian Federation, Moscow, Russia*

** e-mail: fedorowams@yandex.ru*

Key words: Colorectal cancer, glycolysis, differential gene expression, qPCR, bioinformatics analysis

Motivation and Aim: The accurate cause of dense aberrant DNA methylation in CIMP tumors remains incompletely clear [1]. However, several factors that may be associated with this process were found, including metabolic alterations such as glycolysis activation [2]. Therefore mRNA expression for genes involved in energy metabolism in CIMP+ and non-CIMP colorectal tumors (CRC) was compared.

Methods and Algorithms: Methylation profiling and gene expression data from The Cancer Genome Atlas were analyzed using R package DESeq2. mRNA level of the glycolytic genes was estimated for Russian population with qPCR (20 CIMP+ and 20 non-CIMP CRC samples).

Results: Bioinformatics analysis revealed increased expression level (1.5–3 folds) of many glycolytic genes in CIMP+ tumors compared to non-CIMP. Increased expression level was demonstrated for genes *ENO2* (3-fold), *PFKP*, *HK*3 and *PKM* (2-fold). Expression of genes involved in the Krebs cycle in CIMP+ tumors was slightly altered. Although decreased expression of the *OGDHL* (8-fold) was demonstrated. Expression of *PKLR* gene, a participant of gluconeogenesis, was 20-fold decreased. These results were verified with qPCR analysis. Frequent increase of *ENO2* mRNA level in CIMP+ CRC samples as well as 7 and 4-fold decrease of *OGDHL* and *PKLR* genes were evaluated [3].

Conclusion: Association between CIMP+ phenotype and activation of glycolytic genes was demonstrated in CRC samples. Both characteristics commonly correlated with aggressiveness of tumors and unfavorable prognosis [4].

Acknowledgements: This work was funded by the Russian Science Foundation, grant 14-15-01083. qPCR was performed using the equipment of EIMB RAS “Genome” center (http://www.eimb.ru/rus/ckp/ccu_genome_c.php).

References

1. Kudryavtseva A.V. et al. (2016) Important molecular genetic markers of colorectal cancer. Oncotarget. 7(33):53959-53983.
2. Krasnov G.S. et al. (2015) Evaluation of hexokinase gene expression in colorectal cancer using bioinformatics tools. Biophysics. 60(6):870-875.
3. Fedorova M.S. et al. (2015) Downregulation of OGDHL expression is associated with promoter hypermethylation in colorectal cancer. 49(4):678-88.
4. Snezhkina A.V. et al. (2016) Differential expression of alternatively spliced transcripts related to energy metabolism in colorectal cancer. BMC Genomics. 17(14):1011.

DOI 10.18699/BGRSSB-2018-016

Functional analysis of SNVs affecting splicing in congenital aniridia

A. Filatova[1]*, T. Vasilyeva[1], M. Skoblov[1, 2], A. Voskresenskaya[3], A. Marakhonov[1, 2], R. Zinchenko[1, 4]

[1] *Research Center for Medical Genetics, Moscow, Russia*

[2] *Moscow Institute of Physics and Technology* (*State University*)*, Dolgoprudny, Russia*

[3] *Cheboksary Branch of S. Fyodorov Eye Microsurgery Federal State Institution, Cheboksary, Russia*

[4] *Pirogov Russian National Research Medical University, Moscow, Russia*

** e-mail: maacc@yandex.ru*

Key words: congenital aniridia, PAX6, minigene assay, intronic SNVs, splicing, exon skipping, exon elongation

Motivation and Aim: A congenital aniridia (OMIM 106200) is a rare autosomal dominant panocular disorder caused by mutations in the *PAX6* gene or chromosome 11p13 rearrangements. Haploinsufficiency due to LoF mutations in *PAX6* gene is thought to be the main mechanism of congenital aniridia ethiopathogenesis. Mutations affecting splicing are known to be a cause of Mendelian disorders, but their pathogenic status cannot always be provided bioinformatically. This is a reason why some aniridia associated *PAX6* nucleotide variants remain outside the mutations database. Minigene splicing assay *in vitro* allowed us to analyze functional consequences of six different intronic sequence variants, one missense and one synonymous substitution in *PAX6* gene found in patients with congenital aniridia.

Methods and Algorithms: DNA samples for analysis were obtained from patients with congenital aniridia (110 patients from 84 unrelated families). The search for mutations in the *PAX6* gene was carried out by Sanger sequencing, MLPA and analysis of heterozygosity loss (LOH) in proband. Human Splicing Finder and IntSplice on-line tools were used to predict the effect of identified SNVs on PAX6 pre-mRNA splicing. To confirm the effect of SNVs on splicing experimentally we used *in vitro* minigene assay.

Results: Molecular analysis of a large cohort of aniridia patients from Russia conducted earlier in RCMG revealed a significant proportion of *PAX6* mutations affecting splicing (17.3 %). We focused on 8 SNVs located out of canonical splicing site dinucleotides: 6 deep-intronic and 2 exonic. These variants were classified as variant of unknown significance (VUS), benign or likely pathogenic according to ACMG recommendations. On-line prediction tools analysis revealed that they could influence PAX6 pre-mRNA splicing. *In vitro* minigene assay showed that all investigated sequence variants except one affected splicing. The variants resulted in open reading frame shifting, premature termination codon formation followed by aberrant mRNA degradation by nonsense-mediated decay. Thus, investigated SNVs produce a null allele and haploinsufficiency of the PAX6 function. So investigated mutations were reclassified as loss of function.

Conclusion: Using functional *in vitro* analysis we confirmed the pathogenicity of 7 *PAX6* mutations affecting splicing. Our results emphasized the necessity of such an analysis and advanced search for *PAX6* mutations.

Acknowledgements: Supported by the RFBR grant (17-04-00475).

DOI 10.18699/BGRSSB-2018-017

Predicting pathological promoter-enhancer rewiring in chromosomal rearrangements

V. Fishman[1, 2]*, P. Belokopytova[1, 2], P. Salnikov[1, 2], D. Fishman[2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: minja@bionet.nsc.ru*

Key words: genome architecture, chromosomal rearrangements, promoters, enhancers, Hi-C

Motivation and Aim: in mammals transcription of genes is precisely regulated by multiple molecular mechanisms. A crucial part of regulatory process is specific pairing of promoters and enhancers. Despite recent progress in genome-wide identification of enhancer sequences, it is not clear what determine specificity of 3-dimentional interactions between regulatory elements. Recently, several computational algorithms were employed to predict spatial interactions based on epigenetic signatures of promoters, enhancers and surrounding genomic regions. Here, we aimed to extend these algorithms to explain or predict alterations of 3-dimentional organization of chromatin and accompanying changes of gene expression, induced by chromosomal rearrangements.
Methods and Algorithms: We modify recently published machine-learning algorithm [1] to predict promoter-enchanter interactions in mouse and human genomes. To train algorithm and estimate importance of predictors, we used known promoter-enhancer interactions from [2]. Next, we used the trained model to predict gain or loss of interactions in genome altered by chromosomal rearrangement. We utilized publically available mouse (IMPC, Jaxon Lab) and human (OMIM, ClinVar) databases to suggest phenotypical consequences associated with promoter-enhancer rewiring caused by chromosomal rearrangement. Finally, we used CRISPR/Cas9 tools to establish experimental system for validation of model prediction.
Results and Conclusion: We developed a machine learning algorithm to predict loss or acquirement of promoter-enhancer interactions caused by chromosomal rearrangements. We estimated impact of different biological features on prediction accuracy and suggested the optimal set of features. We used the algorithm to predict promoter-enhancer rewiring caused by defined chromosomal rearrangements. Finally, we showed that CRISPR/Cas9 system can be utilized to validate prediction of developed algorithm.
Acknowledgements: This study was supported by the RSF (grant 17-74-10143).

References

1. Whalen S. et al. (2016) Enhancer–promoter interactions are encoded by complex genomic signatures on looping chromatin. Nat. Genet. 48(5):488-496.
2. Shen Y. et al. (2012) A map of the cis-regulatory sequences in the mouse genome. Nature. 488(7409): 116-120.

DOI 10.18699/BGRSSB-2018-018

Entropic hourglass patterns of animal and plant development and the emergence of biodiversity

A. Gabel[1], H.-G. Drost[2], M. Quint[1], I. Grosse[1, 3]*
[1] *Institute of Computer Science at Martin Luther University Halle-Wittenberg, Germany German Centre for Integrative Biodiversity Research* (*iDiv*) *Halle-Jena-Leipzig, Germany*
[2] *University of Cambridge, Cambridge, UK*
[3] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: grosse@informatik.uni-halle.de*

Key words: embryogenesis, phylotranscriptomic hourglass

One surprising observation going back to pioneering works of Karl Ernst von Baer in 1828 and Ernst Haeckel in 1866 is that embryos of different animal species express on average evolutionarily young genes at the beginning of embryogenesis, evolutionarily old genes in mid-embryogenesis, and again evolutionarily young genes at the end of embryogenesis.

Focusing our attention on plants, which represent the second major kingdom in the tree of life that evolved embryogenesis, we have found that this phylotranscriptomic hourglass pattern also exists in plant embryogenesis, which is surprising as multicellularity and embryogenesis evolved independently in animals and plants. Moreover, we have found that phylotranscriptomic hourglass patterns also exist in the two main transitions of post-embryonic plant development, germination and floral transition, suggesting the convergent evolution of phylotranscriptomic hourglass patterns in animal and plant development.

The origin of these phylotranscriptomic hourglass patterns has remained concealed, but here we find that not only the mean age of expressed genes changes in an hourglass-like manner, but the whole age distribution of expressed genes changes. When studying the entropy of these age distributions as functions of time, we find hourglass patterns that surprisingly are orders of magnitude more significant than the original phylotranscriptomic hourglass patterns of the mean, which might indicate that the entropic hourglass patterns are more fundamental than, and possibly even the origin of, the original phylotranscriptomic hourglass patterns of animal and plant development.

DOI 10.18699/BGRSSB-2018-019

The analysis of functional activity of extracellular matrix genes in carotid atherosclerotic plaques

I.A. Goncharova[1]*, A.V. Markov[1], A.A. Sleptsov[1], Y.A. Koroleva[1], D.V. Sharysh[3], A.A. Zarubin[3], N.R. Valiahmetov[1], E.F. Muslimova[2], D.S. Panfilov[2], B.N. Kozlov[2], S.A. Afanasiev[2], V.P. Puzyrev[1, 3], M.S. Nasarenko[1, 3]

[1] *Research Institute of Medical Genetics, Tomsk, Russia*

[2] *Research Institute of cardiology, Tomsk, Russia*

[3] *Siberian State Medical University, Tomsk, Russia*

** e-mail: irina-goncharova@medgenetics.ru*

Key words: extracellular matrix genes, carotid artery plaque

Motivation and Aim: arterial sclerotic disease is the reason of cardiovascular diseases progression and the development of severe acute events, which are directly connected with stability of atherosclerotic plaque. Metabolism of extracellular matrix is essential in forming a thick fibrous layer and keeping plaque whole. Our previous researches showed the association between polymorphic variants of genes *ADAMDEC1*, *MMP3*, *ITGA4*, *ITGB5*, *TIMP2*, which protein products are involved in extracellular matrix metabolism, and the variety of diseases, including myocardial infraction [1, 2]. The aim of this research is to estimate the functional activity of those genes, essential in extracellular matrix metabolism, in the case of carotid artery sclerotic disease.

Methods and Algorithms: Samples of carotid artery plaque (CAP), which were taken from three patients after carotid endarterectomy, were used to perform full genome gene expression analysis. Samples of intact internal mammary arteries (IMA), taken from other patients, were used as control. Gene expression analysis was performed using microchip HumanHT-12 BeadChip (Illumina). Bioinformatics analysis of materials was performed using lumi, limma packages in R software environment (Bioconductor). Genes with difference of the expression level between sample groups $|FC| \geq 2$ and $p_{FDR} < 0.05$, were considered to be differently expressed. Functional annotation of differentially expressed genes was performed in the Web-based GEne SeT AnaLysis Toolkit.

Results: 469 differently expressed transcripts, belonging to 445 genes, were discovered in carotid artery plaque, comparing to intact arteries. Most of overexpressed genes in carotid artery plaque are those, which protein products are involved in extracellular matrix organization (GO:0030198; *CD44*, *COL1A2*, *COL3A1*, *COL5A2*, *FMOD*, *HAPLN1*, *ITGA11*, *ITGAV*, *SPARC*, *SPP1*, *SULF1*, *TIMP1*; $p_{FDR} = 1.44 \times 10^{-07}$). Genes *ADAMDEC1*, *ITGB5* и *TIMP2* are also overexpressed in arteries, damaged by atherosclerosis ($p_{FDR} = 0.018$; $p_{FDR} = 0.011$; $p_{FDR} = 0.006$).

Conclusion: Consequently, the functional activity of genes, encoding proteins, are involved in extracellular matrix organization, was increased in carotid artery plaque at the late stages of pathological process.

Acknowledgements: The study was partially supported by the RFBR (16-04-00840).

References

1. Goncharova I.A. et al. (2016) Genes for fibrogenesis in the determination of susceptibility to myocardial infarction. Mol. Biology. 50(1):81-90.
2. Tarasenko N.V. et al. (2017) Association of genes of different functional classes with type 1 diabetes. Russ. J. Genet. 53(8):923-929.

DOI 10.18699/BGRSSB-2018-020

GWAS-MAP: a platform for storage and analysis of the results of thousands of genome-wide association scans

D.D. Gorev[1, 2, 3]*, T.I. Shashkova[1, 2, 3], E. Pakhomov[2], A. Torgasheva[4], L. Klaric[5, 6], A. Severinov[1, 3], S. Sharapov[2, 4], D.G. Alexeev[1, 2], Y.S. Aulchenko[2, 3, 4]

[1] *LLC 'Knomics', Moscow, Russia*

[2] *Novosibirsk State University, Novosibirsk, Russia*

[3] *Moscow Institute of Physics and Technology, Moscow, Russia*

[4] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

[5] *Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK*

[6] *Genos Glycoscience Laboratory, Zagreb, Croatia*

** e-mail: gorev.d@phystech.edu*

Key words: genome-wide association study, SNP, database, web-service

Motivation and Aim: Hundreds of genome-wide association studies (GWAS) of human traits are performed each year, and are, together with results from tens of thousands of previously reported GWAS, freely available. These results are published in the form of summary statistics (where for each SNP the allelic frequencies, estimates of the coefficients of regression and their standard errors are typically reported). This information can be used for multiple purposes – from research in fundamental biology and genetics, to biomarker and therapeutic intervention of targets discovery. At the same time, while the amount of information accumulated by the scientific community is very large, the use of this valuable information is unfortunately restricted by lack of reporting guidelines and facilities that would allow for quality control (QC), long-time storage, and analysis of such data. This situation forces researchers to spend a lot of time and efforts on data collection, data preprocessing to accommodate different analytical tools, and QC. In this work, we designed a platform for storage, QC, and analysis of GWAS summary statistics.

Results: The original data harmonisation algorithm was developed to effectively store and quickly access GWAS data. For data storage and manipulations on our platform we use two database management systems, ClickHouse, to store harmonized GWAS results and PostgreSQL for meta-data storage. Clickhouse provides us with rapid-access storage accessible via powerful and flexible SQL interface. The platform implements several GWAS QC algorithms. It also embeds several methods often used for analysis of GWAS summary statistics, such as LDsr and MRbase libraries that facilitate genetic correlations and mendelian randomization analyses, respectively, and our own implementation of the summary data-based mendelian randomization and heterogeneity in dependent instruments (SMR-HEIDI) testing that allows for analysis of pleiotropy. As a proof of concept, 429 GWASs, totalling to ~3 billions of SNP-trait associations, were uploaded to the platform. On average, the selection of all SNPs in a 500 kbp range from all GWAS in the database to a Python Dataframe takes ~17 seconds, while selection of a specific GWAS ~34 seconds.

Conclusion: We have developed a platform for storage, quality control and analysis of summary GWAS data. The platform is capable of storage and high throughput retrieval and analysis of results of thousands of GWAS.

Availability: GWAS-MAP will soon be available as a web-service platform.

DOI 10.18699/BGRSSB-2018-021

Mining of non-teleost fish transcriptomes and genomes uncovers the evolutionary history of immunoglobulin light chains

S. Guselnikov[1, 2]*, K. Baranov[1], N. Chikaev[1], A. Najakshin[1], S. Kulemzin[1], A. Makunin[1], V. Trifonov[1, 2], A. Taranin[1, 2]

[1] *Institute of Molecular and Cellular Biology SB RAS, Novosibirsk, Russia*

[2] *Novosibirsk State University, Novosibirsk, Russia*

** e-mail: sguselnikov@mcb.nsc.ru*

Key words: IgL isotype, CDR length, *Acipenseriformes*, *Polyperiformes*, Holostei, Teleostei, jawed vertebrates, phylogenetic analysis

Motivation and Aim: Subdivision of mammalian immunoglobulin light (IgL) chains into kappa and lambda types has for long time remained obscure from both the functional and evolutionary perspective. The aim of this study was to fill the important gaps in the evolutionary history of light chains by examining the structure and diversity of IgL genes in non-teleost ray-finned fish.

Methods and Algorithms: Standard experimental (RACE PCR cDNA cloning, Sanger sequencing) and bioinformatic (BLASTN, TBLASTN, BLASTP, MUSCLE sequence alignment, NJ/ME/ML phylogenetic analysis with bootstraps) methods were used.

Results: We performed bioinformatic analysis of recent transcriptomic and/or genomic resources for four *Acipenseriformes* species: sterlet (*Acipenser ruthenus*), Siberian sturgeon (*Acipenser baerii*), Chinese sturgeon (*Acipenser sinensis*), and American paddlefish (*Polyodon spathula*). The results obtained were used for the detailed experimental characterization of the IgL genes in sterlet. This species was shown to possess three loci of genes for IgL kappa-like chains, a single Ig lambda-like VJC cluster as well as one sigma-like V and one sigma-like C gene. The data obtained on *Acipenseriformes* were extended by the bioinformatic identification of IgL genes in a holostean spotted gar (*Lepisosteus oculatus*) and in two polypterid species, saddled bichir (*Polypterus endlicheri*) and ropefish (*Erpetoichthys calabaricus*). The inclusion of IgL sequences from non-teleost ray-finned fish into phylogenetic analysis for the first time showed subdivision of IgL chains into five ancient types. The teleostean IgL "lambda" chains turned out to be a kappa and lambda chain paralog that emerged before the radiation of ray-finned fish but have been lost in *Acipenseriformes*. We designate this type lambda-2. Each vertebrate lineage has its own combination of IgL types (from one to four) with the kappa and lambda chains being the most broadly spread. Sequence comparisons showed that, in contrast to sigma, sigma-2 and lambda-2 chains, the kappa and lambda are highly variable in the length of either CDR1 (kappa) or both CDR1 and CDR2 (lambda).

Conclusion: The IgL chains are subdivided into five ancient isotypes differentially evolved in various lineages of jawed vertebrates. High variability of the CDR length in the kappa and lambda chains may be responsible for their preferential evolutionary retention by providing more flexibility in association with IgH V domains.

Acknowledgements: Supported by the RFBR grant No. 17-04-01519 (cDNA cloning) and by the RSF grant No. 18-44-04007 (*Acipenseriformes* genomes and transcriptomes analysis).

DOI 10.18699/BGRSSB-2018-022

Bayesian modelling of gene network alterations during blood cells differentiation and cancerogenesis

A.A. Igolkina*, M.G. Samsonova
Peter the Great St. Petersburg Polytechnic University, Mathematical Biology and Bioinformatics Laboratory, St. Petersburg, Russia
** e-mail: igolkinaanna11@gmail.com*

Key words: gene network, Structural Equation Modeling, Bayesian inference, branching process, Wiener process, Leukemia

Blood is a heterogeneous tissue with the known tree of differentiation. While the process of differentiation from hematopoietic stem cells through different progenitor states to mature states (monocytes, lymphocytes, neutrophils, etc.) is continuous we hypothesise that changes in genetic networks along the tree also follow the continuous-states process. We developed the model of one gene network alterations along the tree based on the following assumptions. (1) In each inner node and outer leaf of the tree, gene network satisfies the Structural Equation Model (SEM); (2) the change of gene covariance matrix together with coefficients of gene-gene interactions follow the continuous-states, time-homogeneous Markov Process, specifically the Wiener Process. We used gene expression data within the leaves (microarray Human Map dataset) and optimised all parameters of both SEM model and the Wiener Process drawing the Bayesian Inference. We worked with RAS signalling network as it involves in Leukemia development. We predicted the states of this gene network in inner nodes and, using parameters of the Wiener Process, predicted the point on the tree where the cancer cells (T-cells and B-cells) have its own branch. The knowledge of this point can potentially help in leukaemia treatment.
We consider, the developed methodology can be easily applied to other cell development and also phylogenetic studies.

DOI 10.18699/BGRSSB-2018-023

PyMPFA: python pipeline for massively parallel functional assays used for characterization of DNA regulatory elements

A. Ivankin[1]*, A. Pindyurin[1, 2]
[1] *Institute of Molecular and Cellular Biology SB RAS, Novosibirsk, Russia*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
** e-mail: anton.ivankin@gmail.com*

Key words: massively parallel functional dissection assay, DNA regulatory elements, python

Motivation and Aim: The spatio-temporal regulation of gene expression is a complex process that determines the destiny and function of various cells and tissues. The regulation of this process consists of many stages and to a large extent is orchestrated by diverse DNA regulatory elements (promoters, terminators, enhancers, etc.) and epigenetic factors associated with them. Recently, a number of massively parallel functional assays (MPFAs) aimed to identify and dissect DNA regulatory elements were developed. Most of these assays make use of short DNA sequences, which are most frequently referred to as "tags" or "barcodes", to track the expression activity (and in some assays genomic localization) of individual gene reporter constructs. Typically, as assay readout, barcodes are PCR-amplified from cDNA and DNA samples obtained from the studied cells and subsequently identified and counted by using the high-throughput sequencing.
Methods and Algorithms: Python language (v. 2.7.6) was used to implement the pyMPFA algorithm designed to extract barcodes and associated with them variable DNA sequences (hereafter, "features", which are mutant variants or genomic locations) from the high-throughput sequencing reads and subsequently identify and report the most reliable barcode-feature combinations. The follow-up analysis of the data was carried out in R language (v. 3.4.4).
Results: We have developed a flexible pyMPFA pipeline, which is applicable to high-throughput sequencing reads of different structure. As a pilot test, we applied this pipeline to datasets generated by MPFAs aimed to characterize (1) the influence of minor DNA sequence variations in the 3′ downstream region of the reporter gene on its episomal expression in cultured human cells and (2) the influence of local chromatin environment of the reporter gene activity in cultured drosophila cells. As a result, we found that the running time of the pyMPFA pipeline strongly depends on the number of unique barcodes present in a dataset. There is an exponential dependence between the number of processed unique barcodes and the time of the pipeline run. For example, reads with about 200 thousand unique barcodes can be processed by the pyMPFA pipeline (on CPU Intel Core-i7 3770K, 32Gb DDR3) in single-threaded mode in 4–5 hours.
Conclusion: Despite the flexibility of the pipeline, it still operates slowly when processing large datasets. The processivity of the pyMPFA pipeline will be improved in its future releases. Source code of the pipeline is available on github: https://github.com/wiw/pyMPFA/tree/trip_0.3/pyMPFA.
Acknowledgements: Supported by the RSF (16-14-10288).

DOI 10.18699/BGRSSB-2018-024

HPC clusters and Big Data storage for data analysis in scientific research – Huawei experience

A. Iyudin
Huawei Technologies
e-mail: iyudin.alexey@huawei.com

Key words: high-performance computing, large-scale computing

High-performance computing (HPC for short) is a computer cluster system that connects multiple computer systems using various interconnection technologies and utilizes the integrated computing capability of all connected systems to process large-scale computing tasks. That's why HPC is also known as HPC cluster.

Huawei is dedicated to providing efficient, easy-to-manage, energy-saving, and flexible HPC solutions. Huawei hardware is a cornerstone of the Huawei HPC solution.

Huawei provides a wide range of hardware devices that can be used in the HPC system, including servers, storage devices, and network devices. Huawei servers have bettered 365 world records in SPEC tests, and the number of records is still increasing. Huawei OceanStor 9000 provides up to 5,000,000 operations per second (OPS) and 200 GB/s throughput, and features high reliability and optimal performance.

The Huawei HPC solution provides board-level to system-level energy saving measures, intuitive, real-time monitoring, and dynamic energy saving technologies, reducing power consumption by up to 40 %.

Huawei has delivered nearly a hundred HPC projects, customers all over the five continents, in education, scientific research, energy, animation, meteorology and other fields to play a role.

DOI 10.18699/BGRSSB-2018-025

Motif analysis of regulatory SNPs reveals Krüppel-like transcription factors as putative tumor suppressors in colorectal carcinoma

O. Kel-Margoulis[1]*, A. Kel[1, 2], E. Wingender[1, 3]
[1] *geneXplain GmbH, Wolfenbuettel, Germany*
[2] *Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia*
[3] *Institute of Bioinformatics, University Medical Center Göttingen, Georg August University, Göttingen, Germany*
** e-mail: olga.kel-margoulis@genexplain.com*

Key words: regulatory SNP, transcription factors, signal transduction, colorectal cancer, CRC

Motivation and Aim: Genome-wide associated studies (GWAS) reveal multiple SNPs correlated with predisposition to various cancers. Majority of such SNPs are found in non-coding regions of genome, which makes their interpretation a special task. It becomes evident that such SNPs mark regions of genomes that posses important regulatory functions. It is important to understand the role of regulatory SNPs in molecular mechanisms of cancer initiation and progression. In the current work we study SNPs that are revealed by eQTL while analyzing massive transcriptomics (RNA-seq) and genotyping data of a cohort of 103 patients with colorectal carcinoma (CRC) [1]. Many of them are also shown to be correlated with the change of expression of neighbor genes. We applied a machine learning techniques combined with graph search algorithms in order to reconstruct regulatory circuits involving genes located near regulatory SNPs.
Methods and Algorithms: In the current work we applied a novel approach, which we refer to as "walking" pathways. This approach is further development of "upstream analysis" [2]. First, we applied an improved version of Composite Module Analyst (CMA) [3] to reveal specific combinations of transcription factor binding sites (TFBS) in the genomic regions around regulatory SNPs. Next, we performed a graph search in signal transduction network for common regulatory nodes upstream of transcription factors that were revealed at the previous step. At this step we identified positive feedback loops, which direct the network search towards potential master-regulators of a self-inducing pathological state of the system.
Results: We found that binding sites for KLF transcription factors are significantly enriched around regulatory SNPs. We also found that KLF transcription factor binding sites are co-localized with binding sites for E2F1, TP53 and other factors forming composite elements in the regions around regulatory SNPs. Analysis of gene expression data in tumor samples revealed significant down-regulation of three KLF family members. Finally, the upstream analysis with positive feedback loops (walking pathways) identified several potential master-regulators. Among them we identified several cyclin/cdk complexes, which are significantly up-regulated in the tumor samples. The reconstructed signal transduction circuit involves cell cycle regulators, Cdk1, Cdc2, Plk1, p107 as well as transcription factors c-Myc, TGIF, PEA3 whose expression is highly unregulated in tumor samples.
Conclusion: These findings, together with the fact of significant down-regulation of three KLF family members, allow suggesting GKLF as a putative tumor suppressor. Down-regulation of GKLF in many CRC patients may lead to activation of oncogenic regulatory circuits. These results shed a light on the important role of regulatory SNPs in cancer genomes.
Acknowledgements: Supported by the Targeted Program "Research and development on priority directions of science and technology in Russia, 2014–2021", Contract No. 14.604.21.0101, unique identifier of the applied scientific project: RFMEFI60414X0101, and by the grant of the EU FP7 program: "SYSCOL".

References

1. Ongen H. et al. (2014) Putative cis-regulatory drivers in colorectal cancer. Nature. 512(7512):87-90.
2. Koschmann J. et al. (2015) "Upstream Analysis": An integrated promoter-pathway analysis approach to causal interpretation of microarray data. Microarrays. 4:270-286. DOI 10.3390/microarrays4020270.
3. Waleev T. et al. (2006) Composite Module Analyst: identification of transcription factor binding site combinations using genetic algorithm. Nucleic Acids Res. 34(Web Server iss.):W541-5.

DOI 10.18699/BGRSSB-2018-026

Design of genus-specific primer panel for detection and identification of viral DNA in environmental samples using next-generation sequencing

K. Khafizov[1, 2]*, A.A. Ayginin[1, 2], E.V. Pimkina[1], A.D. Matsvai[1, 2], A.S. Speranskaya[1, 3], M.V. Safonova[1], E.A. Blinova[1], I.V. Artyushin[1, 3], V.G. Dedkov[1], G.A. Shipulin[1]

[1] *Central Research Institute of Epidemiology, Moscow, Russia*

[2] *Moscow Institute of Physics and Technology, Dolgoprudny, Russia*

[3] *Lomonosov Moscow State University, Moscow, Russia*

** e-mail: khafizov@cmd.su*

Key words: virome analysis, pathogens detection, NGS

Motivation and Aim: The advances in the next generation sequencing (NGS) technologies have significantly increased our ability to detect new viral pathogens and systematically determine the spectrum of viruses that persist in various biological samples [1]. Such studies led to the discovery of new viral pathogens, as well as to establishing the associations of viromes with many diseases [2]. However, unlike the metagenomics studies using *16S* rRNA for bacteria detection, it is impossible to create universal oligonucleotides to target all known and novel viruses due to the viral genomes diversity and their variability, whereas whole-genome sequencing is still expensive and relatively low-sensitive for such purposes.

Methods and Algorithms: In this study, we designed a genus-specific oligonucleotide panel for targeting enrichment of viral nucleic acids in different samples and demonstrated possibility of its application for virus detection in samples collected from migratory birds. The role of migratory birds in the circulation and spreading of a number of viral pathogens is well known [3]. Thereby the analysis of bird's seasonal migration routes and their viromes allows predicting the future directions and timing of pathogens spreading, and could help to prevent infectious outbreaks.

Results: Our panel has been tested using a number of collected bird samples and has demonstrated superior efficiency in pathogen detection and identification. The reliable bioinformatics pipeline for the rapid classification of the sequences was crucial to success of these efforts. Existing approaches for designing oligonucleotides are usually oriented to developing primers or DNA-probes for detection of particular viral species or genera [4] using PCR, but not families or higher taxonomic orders. Moreover, it is nearly impossible to design multiplex primer panel using mentioned algorithms. Thus, we developed computational pipeline for designing oligonucleotides that cover high number of known viruses belonging to different taxonomic orders and also their novel variants. In this work we also developed an NGS-based data analysis module, and demonstrated the functionality of our tool both for detecting novel viruses and for analyzing the virome diversity.

Conclusion: We introduced a method for designing oligonucleotide panels for targeted viral nucleic acids enrichment, where the main idea is to use a minimal number of oligonucleotides to cover the maximal number of diverse viral taxa, and this can be performed in one PCR reaction. We applied this approach to design genus-specific primer pairs for target nucleic acids enrichment of zoonotic viruses and evaluated it using a number of samples from migratory birds. The superior increase in the viral genome coverage has been shown.

Acknowledgements: Supported by the RSF grant (17-74-20096).

References

1. Yeh Y.-T. et al. (2016) Tunable and label-free virus enrichment for ultrasensitive virus detection using carbon nanotube arrays. Sci. Advances. 2(10):e1601026.
2. Briese T. et al. (2015) Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. mBio. 6(5):e01491-15.
3. Verhagen J.H. et al. (2014) Migratory birds reinforce local circulation of avian influenza viruses. PLoS One. 9(11): e112366.
4. Gall A. et al. (2008) Universal primer set for amplification and sequencing of HA_0 cleavage sites of all influenza A viruses. J. Clim. Microbiol. 46(8):2561-2567.

DOI 10.18699/BGRSSB-2018-027

Genomic organization of serratiochelin cluster in the environmental and clinical strains of *Serratia marcescens*

I. Khilyas[1]*, M. Sharipova[1], L. Bogomolnaya[1, 2]
[1] *Institute of Fundamental Medicine and Biology KFU, Kazan, Russia*
[2] *Texas A&M University Health Science Center, Bryan, Texas, USA*
** e-mail: irina.khilyas@gmail.com*

Key words: siderophores, serratiochelin, *Serratia*

Motivation and Aim: Bacterial siderophores are small secondary metabolites with a strong ferric iron-binding capacity. Availability of iron in the mammal hosts during infection is extremely limited for support of bacterial survival. The non-ribosomal peptide synthetase (NRPS) – mediated combinatorial biosynthesis of siderophores is induced by the iron starvation and functions to replenish iron supply in the bacterial cells [1]. Thus, siderophores play an essential role as a virulence factor of pathogenic bacteria. A goal of this study was to identify the serratiochelin gene cluster in the genomes of the environmental and clinical strains of *S. marcescens,* SM6 and SR 41-8000.
Methods and Algorithms: *S. marcescens* genomes were analyzed for the presence of siderophores-encoding gene clusters using AntiSmash software and using RAST platform for gene annotation [2].
Results: Genome analysis of *S. marcescens* strains Sm6 and SR41-8000 showed the presence of NRPS modules SM6_200-203 and SR41_3114-3117, respectively. Closely related bacteria *Serratia* sp. V4 was recently showed to synthetize siderophores serratiochelines A, B and C [3]. The BLAST analysis of amino acid sequences of identified NRPS *S. marcescens* SM6 and SR 41-8000 against serratiochelin NRPS *Serratia* sp. V4 demonstrated a high homology. However, a comparison of *Serratia* sp. V4 serratiochelin cluster organization with clusters, identified in *S. marcescens* SM6 and SR 41-8000 strains revealed significant differences.
Conclusion: Both strains *S. marcescens* SM6 and SR 41-8000 encode NRPS needed for serratiochelin production. The organization of serratiochelin gene clusters in environmental and clinical strains of *S. marcescens* differs from *Serrati a* sp. V4.
Acknowledgements: This work was supported by the Russian Science Foundation project 16-14-10200 and performed in accordance with the Russian Government Program of Competitive Growth of Kazan Federal University.

References

1. Crosa J.H., Walsh C.T. (2002) Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. Microbiol. Mol. Biol. Rev. 66(2):223-249.
2. Weber T. et al. (2015) AntiSMASH 3.0 – a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res. doi 10.1093/nar/gkv437.
3. Seyedsayamdost M.R. et al. (2012) Mixing and matching siderophore clusters: structure and biosynthesis of serratiochelins from Serratia sp. V4. J. Am. Chem. Soc. 134(33):13550-13553. DOI 10.1021/ja304941d.

DOI 10.18699/BGRSSB-2018-028

Osteogenic-related gene expression by human adipose-derived mesenchymal stem cells

I. Khlusov[1, 2], L. Litvinova[2]*, K. Yurova[2], V. Shupletsova[2], O. Khaziakhmatova[2], E. Melashchenko[2], L. Pokrovskaya[1]
[1] *Tomsk State University, Tomsk, Russia*
[2] *Immanuel Kant Baltic Federal University, Kaliningrad, Russia*
** e-mail: larisalitvinova@yandex.ru*

Key words: *in vitro*, quantitative polymerase chain reaction, osteoblastic differentiation

Motivation and Aim: Mesenchymal stem cells (MSCs) are multipotent cells, which among other cell lineages, give rise to adipocytes and osteoblasts. The differentiation of adipose-derived MSCs (AMSCs) into osteoblasts occurs through the cross talk between complex signaling pathways including those derived from bone morphogenic proteins (BMPs), fibroblastic growth factors (FGF), osterix (OSX), and runt-related transcription factor 2 (Runx2). Transcription factors Runx2 and OSX are the main molecular switches and determinants of MSC osteoblastogenesis [1]. Therefore, an expression of osteogenic genes has been studied for 14-day culture of human AMSCs (hAMSCs).
Methods and Algorithms: hAMMSCs were isolated from lipoaspirate (permission no. 7 of December 9, 2015; Local Ethics Committee, Immanuel Kant Baltic Federal University) as described in [2]. Cells were cultured with standard nutrient medium (90 % DMEM/F12 (1:1), 10 % fetal-calf serum, 50 mg/mL gentamicin, and 280 mg/mL L-glutamine without osteogenic additions) for 14 days with medium exchange every 3–4 days. The expression of osteogenic-related genes (*RUNX2*, *BMP2*, *BMP6*, *FGF10*, and *ALPL*) was analyzed *via* multiplex quantitative polymerase chain reaction (qPCR). mRNA was extracted from cells using the ExtractRNAkit (Evrogen, Russia) according to the manufacturer's instructions. Reverse transcription of RNA samples was performed using the MMLV RT kit (Evrogen, Russia). qPCRmixHS reagents (Evrogen, Russia), specific TaqMan probes and primers (Beagle, Russia) were used for qPCR with the help of CFX96 amplifier (Bio-Rad, USA). *RPLPO* served as reference gene.
Results: The expression of pro-osteogenic genes was determined in 14 days of hAMMSCs culturing. Testing of corresponding transcription proteins in particular Runx2 and OSX is planning.
Conclusion: Isolated hAMSCs are useful cells for *in vitro* genetic investigation of osteoinductive properties of scaffolds proposed for bone tissue engineering.
Acknowledgements: Supported by the Federal Target Program of the Ministry of Education and Science of the Russian Federation (Agreement No. 14.575.21.0164, ID number RFMEFI57517X0164).

References

1. Muruganandan S. et al. (2008) Adipocyte differentiation of bone marrow-derived mesenchymal stem cells: Cross talk with the osteoblastogenic program. Cell. Mol. Life Sci. 66(2):236-253.
2. Zuk P.A. et al. (2011) Multilineage cells from human adipose tissue: implications for cell-based therapies. Tissue Engineering. 7(2):211-228.

DOI 10.18699/BGRSSB-2018-029

Siamese neural networks for metagenomics binning

B. Kirillov
Skolkovo Institute of Science and Technology, Moscow, Russia
e-mail: Bogdan.Kirillov@skoltech.ru

Key words: deep learning, few-shot learning, metagenomics, binning

Motivation and Aim: A metagenomic sample usually contains DNA from a lot of different organisms. Each organism has its own genome but only small part of it can be found in pool of sequenced reads and it is hard to identify the source species for each of the resulting reads [1]. This operation is called "binning" and it is one of the major challenges in current metagenomics. It can be done using a reference database or without prior knowledge of taxonomy, in this study I concentrate on the latter case. Most taxonomy-independent binning algorithms rely either only on low-level k-mer features [2] or on additional data, such as DNA methylation [3] or coverage profiles [4]. Ability of deep neural networks to extract high-level features from the sequence may provide an improvement. The research's end goal is to create and evaluate a deep neural network-based solution for binning.
Methods and Algorithms: I use a special neural network architecture, so-called Siamese Networks [5]. The model learns from data how to answer the question "Did that pair of reads come from the same species?" In such a setup the binning problem becomes a binary classification task. To train and test the model I use simulated data and freely accessible dataset of human gut microbiome [6].
Results: Currently, I am in process of hyperparameter selection and optimization. The model shows good generalization ability, also there are indicators that the model is able to identify reads from previously unseen species correctly.
Conclusion: This study provides a novel method of taxonomy-independent metagenomics binning using Deep Learning methods. The trained model, training and read simulation scripts will be available at GitHub.
Acknowledgements: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Wooley J.C., Godzik A., Friedberg I. (2010) A primer on metagenomics. PLoS Comput. Biol. 6(2):e1000667.
2. Girotto S., Pizzi C., Comin M. (2016) MetaProb: accurate metagenomic reads binning based on probabilistic sequence signatures. Bioinformatics. 32(17):i567-i575.
3. Beaulaurier J. et al. (2018). Metagenomic binning and association of plasmids with bacterial host genomes using DNA methylation. Nat. Biotechnol. 36(1):61-69.
4. Lin H.H., Liao Y.C. (2016) Accurate binning of metagenomic contigs via automated clustering sequences using information of genomic signatures and marker genes. Sci. Rep. 6:24175.
5. Koch G., Zemel R., Salakhutdinov R. (2015) Siamese neural networks for one-shot image recognition. In: ICML Deep Learning Workshop. Vol. 2.
6. Nielsen H.B. et al. (2014) Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. Nat. Biotechnol. 32(8):822.

DOI 10.18699/BGRSSB-2018-030

Genetic regulation of immunoglobulin G *N*-glycosylation

L. Klaric[1, 2]*, Y. Tsepilov[3, 13], M. Mangino[4], T. Esko[5, 6, 7], P. Salo[8], S.J. McGurnaghan[1], F. Vuckovic[2], H.M. Colhoun[1, 9], M. Dunlop[10], M. Perola[8], K. Fischer[5], O. Polasek[11], J. Wilson[1, 12], T. Spector[4], Y.S. Aulchenko[3, 13], C. Hayward[1], G. Lauc[2, 14]

[1] *Human Genetics Unit, Medical Research Council, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK*
[2] *Genos Glycoscience Research Laboratory, Zagreb, Croatia*
[3] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[4] *Department of Twin Research and Genetic Epidemiology, King's College London, London, UK*
[5] *Estonian Genome Center, University of Tartu, Tartu, Estonia*
[6] *Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, USA*
[7] *Division of Endocrinology, Boston Children's Hospital, Cambridge, USA*
[8] *Genomics and Biomarkers Unit, Department of Health, National Institute for Health and Welfare* (*THL*)*, Helsinki, Finland*
[9] *Department of Public Health, NHS Fife, Kirkcaldy, UK*
[10] *Colon Cancer Genetics Group, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK*
[11] *Department of Public Health, Faculty of Medicine, University of Split, Split, Croatia*
[12] *Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK*
[13] *Novosibirsk State University, Novosibirsk, Russia*
[14] *Faculty of Pharmacy and Biochemistry, University of Zagreb, Zagreb, Croatia*
***** *e-mail: Lucija.Klaric@ed.ac.uk*

Keywords: immunoglobulin G, glycosylation, genome-wide association studies, pleiotropy

Motivation and Aim: Glycans are complex carbohydrates attached to the surface of the protein. With glycosylation being amongst the most abundant post-translational modification, glycans are expected to have an important role in many physiological processes and most diseases. Although the main enzymes of the glycosylation pathway are known, little is understood about how this template-independent process is regulated to result in a faithful synthesis of a specific glycoform or how it is related with genetic regulation of complex traits and disease. To address these questions, we performed genome-wide association analyses (GWAS) of IgG *N*-glycosylation, followed by extensive *in-silico* functional follow-up.

Methods and Algorithms: We performed GWAS of 77 IgG *N*-glycosylation traits in eight European populations (discovery $N = 8090$, replication $N = 2368$) using HapMap2 imputed genotypes. We prioritised candidate genes according to pleiotropy with gene expression, coding region variants and enrichment in relevant gene-sets. We assessed pleiotropy with complex traits and diseases by comparing regional associations using Mendelian Randomisation based analysis.

Results: We found 27 loci significantly associated ($p \leq 2.4 \times 10^{-9}$) with IgG N-glycosylation. Twelve of these loci replicate findings from Lauc et al. [1] and Shen et al. [2], while 15 are novel. For 9 genes we found evidence of a non-synonymous amino acid change, 4 were pleiotropic with expression in B-cells and 11 with expression in peripheral blood. The remaining genes were prioritised based on the enrichment in antibody synthesis related pathways and cells. Based on the similarity of glycome-wide association estimates we proposed how these genes are connected in the functional network regulating main glycosylation enzymes. In six IgG N-glycosylation loci we found evidence of pleiotropy with autoimmune and inflammatory diseases. In one locus we also showed that IgG N-glycosylation is pleiotropic with both expression of ORMDL3, GSDMB and IKZF3 genes in B-cells and peripheral blood and risk for inflammatory bowel disease, rheumatoid arthritis, cirrhosis, asthma and allergy.

Conclusions: With this study we expanded the network of genes involved in glycosylation of immunoglobulin G providing us with further insights how these molecules could be involved in complex human diseases.

Acknowledgments: This research was supported by FP7 grants MIMOmics (contract No. 305280), HighGlycan (No. 278535), IntegraLife (No. 315997), HTP-GlycoMet (contract No. 3 24400), PainOmics (contract No. 602736) and by Medical Research Council UK.

References

1. Lauc G., Huffman J.E., Pučić M., Zgaga L., Adamczyk B., Mužinić A., … Rudan I. (2013) PLoS Genetics. 9(1). DOI 10.1371/journal.pgen.1003225.
2. Shen X., Klarić L., Sharapov S.Z., Mangino M., Ning Z., Wu D., … Aulchenko Y.S. (2017) Nat. Commun. 1-10.

DOI 10.18699/BGRSSB-2018-031

Evolution of mitochondrial genomes in three closely-related *Armillaria* species

A. Kolesnikova[1], Yu. Putintseva[1], S. Jain[1], N. Oreshkova[1, 2], I. Pavlov[1, 2], V. Sharov[1], D. Kuzmin[1], S. Makolov[1], K. Krutovsky[1, 3, 4, 5]*

[1] *Siberian Federal University, Krasnoyarsk, Russia*
[2] *V.N. Sukachev Institute of Forest SB RAS, Krasnoyarsk, Russia*
[3] *Georg-August University of Göttingen, Göttingen, Germany*
[4] *N.I. Vavilov Institute of General Genetics RAS, Moscow, Russia*
[5] *Texas A&M University, College Station, USA*
** e-mail: konstantin.krutovsky@forst.uni-goettingen.de*

Key words: mitochondrial genome, mitochondrial evolution, homing endonuclease genes, plasmid DNA

Motivation and Aim: Comparative analysis of mitochondrial genomes of closely related organisms allows obtaining new information about their phylogenetic and evolutionary relationships. Plant and fungi mitochondrial genomes are especially interesting due to their large mtDNA size.

Methods and Algorithms: In this study we analyzed mitochondrial genomes of *Armillaria borealis* (116443 bp) and *A. sinapina* (103563 bp) sequenced and annotated in Laboratory of Forest Genomics of the Siberian Federal University. The complete sequence of the *A. solidipes* (122167 bp) mitochondrial genome was retrieved for annotation and comparative analysis as a single scaffold from the JGI Genome portal. Genome rearrangements were detected using the MAUVE 2.0 program. Duplicated sequences were identified by local BLASTn searches of mtDNAs against themselves with a cut-off e-value of 10^{-3}. Intronic nucleotide sequences of three species were compared with each other and NCBI GenBank database using BLAST.

Results: Despite conserved gene content, there were significant genome rearrangements in a region between *rps3* and *atp9* genes. We found gene duplications in all species. There was an 87 bp long truncated duplication of *atp9* located on the minus strand together within *rnl* in *A. solipides* and *A. sinapina*. We also found a 42 bp long duplication of *atp9* in *A. borealis*, which was located next to *atp9* after mobile element of the LAGLIDADG family.

Conclusion: Our study revealed active intronic and mobile genetic elements acquisition during evolution. Some introns have homology with rather distant species. We consider that horizontal DNA transfer has played a significant role in size variation and genome structure of the *Armillaria* mitochondrial genomes.

Acknowledgements: This study was funded by the Research Grant No. 14.Y26.31.0004 from the Government of the Russian Federation.

DOI 10.18699/BGRSSB-2018-032

Population genetic history the red-necked stint

F. Kondrashov
Institute of Science and Technology Austria (*IST Austria*)*, Klosterneuburg, Austria*
e-mail: fkondrashov@gmail.com

Key words: evolution, population genetics, polymorphisms, diversity, red-necked stint

Motivation and Aim: The nature of genetic changes contributing to species extinction and impeding population recovery remains poorly understood. How did living organisms become the way we know them today? This is the fundamental question that our laboratory is preoccupied with. We are less concerned with understanding the organisms themselves, our main focus is on how they evolved, that is how they changed over time. To address these fundamental issues we employ a diversity of modern tools. We apply mathematical modeling, we use available bioinformatic data and we perform our own experiments with a diversity of different model and non-model organisms.
Methods and Algorithms: We study the population genetic history of the critically endangered spoon-billed sandpiper and its sister species, the red-necked stint, which is of least concern.
Results: We found that while the red-necked stint population was relatively constant across 500,000 years, the spoon-billed sandpiper population peaked 15,000–25,000 years ago during the last glacial maximum, when suitable breeding habitat was likely abundant, and has been declining since. The increase of the population prior to the ongoing decline led to accumulation of recessive deleterious polymorphisms, imposing a risk of inbreeding depression on the spoon-billed sandpiper population.
Conclusion: Thus, demographic changes leading to gain of deleterious genetic diversity pose an additional risk to species survival and recovery by increasing the cost of inbreeding. Specifically, species that had greater habitat availability during the last glacial maximum may be especially prone to this effect.
Acknowledgements: The participation at BGRS-2018 multiconference has been supported by Russian Ministry of Science project 28.12487.2018/12.1.

DOI 10.18699/BGRSSB-2018-033

Molecular evolution analysis of the antioxidant system and heme metabolism in helminths

D. Konstantinov[1, 2]*, M. Lvova[2], A. Doroshkov[2]
[1] *Novosibirsk State University, Novosibirsk, Russia*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
** e-mail: konstantinov@bionet.nsc.ru*

Key words: heme, helminths, evolution, antioxidant system

Motivation and Aim: Often endoparasite feed on blood and harmful to vertebral (schistosomiasis, opisthorchiasis, etc.). Common adaptations to parasitism are the reduction of morphology and the change in biochemistry. For the normal functioning of the parasite-hematophagus the most important are the heme metabolism system and the antioxidant system. The choice of these two systems is due to the fact that with this type of food, a large number of blood components enter the parasite organism and it needs to utilize toxic excess iron and heme derivatives. Also the activity of immune cells (they produce reactive oxygene species), and chemical conditions (pH, etc.). These enzymatic systems of many parasites may be targets for therapy. The automatic annotation is complicated due to the large evolutionary distances between model organisms (such as *C. elegance* and *S. mediterranea*) and the most important human hematophagous helminths (schistosomes, opisthores, etc.). In our report, we present a comprehensive study of the molecular evolution of genes encoding enzymes of biosynthesis and heme degradation, as well as antioxidant systems of hematophagous helminths.
Methods and Algorithms: Using the text mining technology we find targets of drags in a model organism (such as *S. mansoni*, *O. viverrini*, *H. robusta*, *N. americanus*). For each target was performed extraction of sets of homological sequences presets from databases was carried out using the reciprocal BLAST search. For each sample set was conducted a domain compositions and protein secondary structure comparative analysis. During molecular phylogeny reconstruction orthologues and paralogues relationships, as well as the functional correspondence between proteins were identified.
Results: Complete set of heme biosynthesis genes identified in Flatworms, this is probably related to the fact that part of the lifecycle of the species conducts in the external environment. At the same time a number of differences in protein structure were revealed. Differences in the set of genes responsible for the synthesis of components of antioxidant systems, as well as genes of enzymes responsible for heme degradation, have been revealed, indicating significant differences in the organization of these systems between free-living and parasitic organisms. The obtained data allows us to define targets for further development of modern methods of control of helminths.
Conclusion: A number of differences in protein structure were revealed, for example, the difference in the domain composition of proteins catalyzing the reaction of the formation of 5-aminolevulinic acid from glycine and succinyl-CoA suggests a different protein localization in the cell.

DOI 10.18699/BGRSSB-2018-034

A novel approach to identify highly connected and differentially expressed gene subnetworks in metastasizing endometrial cancer

K. Kusonmano[1, 2, 3]*, M.K. Halle[2, 4], E. Wik[5, 6], E.A. Hoivik[2, 4], C. Krakstad[2, 4], K.K. Mauland[2, 4], I.L. Tangen[2, 4], A. Berg[2, 4], H.M.J. Werner[2, 4], J. Trovik[2, 4], A.M. Øyan[4, 7], K.-H. Kalland[4, 7], I. Jonassen[1, 8], H.B. Salvesen[2, 4], K. Petersen[1]

[1] *Computational Biology Unit, Department of Informatics, University of Bergen, Bergen, Norway*
[2] *Department of Obstetrics and Gynecology, Haukeland University Hospital, Bergen, Norway*
[3] *Bioinformatics and Systems Biology Program, School of Bioresources and Technology, King Mongkut's University of Technology Thonburi, Bangkhuntien, Bangkok, Thailand*
[4] *Center for Cancer Biomarkers, Department of Clinical Science, University of Bergen, Bergen, Norway*
[5] *Center for Cancer Biomarkers, Department of Clinical Medicine, University of Bergen, Bergen, Norway*
[6] *Department of Pathology, The Gade Institute, Haukeland University Hospital, Bergen, Norway*
[7] *Department of Microbiology, Haukeland University Hospital, Bergen, Norway*
[8] *Center for Cancer Biomarkers, Department of Informatics, University of Bergen, Bergen, Norway*
** e-mail: kanthida.kus@mail.kmutt.ac.th*

Key words: differential expression analysis, gene set analysis, biological network analysis, differentially expressed subnetworks, metastasis, endometrial cancer

Motivation and Aim: Differential expression analyses is a common approach to study molecular changes between two phenotypes of interest, and gene set and network analysis complementary identifying subset of associated genes for functional study and interpretation. The standard application of gene set analysis evaluates enrichment of differentially expressed genes in publicly available gene sets, while the network analysis identifies likely functional gene modules with their interactions.

Methods and Algorithms: We propose a novel approach combining gene set and network analyses to identify phenotypically relevant highly connected and differentially expressed gene subnetworks for better understanding biological mechanisms involved. The method allows data integration between available biologic knowledge of *a priori* gene-gene relations and assessment of internal correlations of expression data. The proposed pipeline includes permutation tests to assess the significance of the identified subnetworks being differentially expressed.

Results: The pipeline was applied to study genes differentially expressed between primary tumors and metastatic lesions in endometrial carcinomas, and made freely available as an R implementation. The identified top ranked subnetworks are associated significantly with disease aggressiveness as well as patient survival, and include genes related to cell proliferation and epithelial-mesenchymal transition.

Conclusions: We have developed a workflow for differential expression analysis providing gene-gene relations inside detected subgroup of related to the studied phenotype. We integrated of known protein-protein interactions and gene expression correlations of aggressive primary tumors and metastatic lesions. The identified subnetworks links to biological processes relevant for the metastasizing process, and also significantly associates with patient survival. We suggest further functional studies of their biological relevance for development of metastatic disease in endometrial cancer.

DOI 10.18699/BGRSSB-2018-035

The overlapped motifs co-occurrence in ChIP-seq data

V.G. Levitsky[1, 2]*, D.Y. Oshchepkov[1], E.V. Zemlyanskaya[1, 2], V.V. Mironova[1, 2], E.V. Ignatieva[1, 2], O.A. Podkolodnaya[1], T.I. Merkulova[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: levitsky@bionet.nsc.ru*

Key words: composite element, chromatin immunoprecipitation, transcription factor binding site prediction

Motivation and Aim: The cooperative binding of transcription factors (TFs) is the common mechanism of their functioning [1]. Recently developed collections of whole-genome datasets for ChIP-seq peaks [2] and derived motifs for TF binding models [3] require development of adequate tools for prediction of potential composite elements (CEs) consisting of anchor/partner motifs separated by relatively short spacer (not more than some tens base pairs). Existing bioinformatics approaches for prediction of motif co-occurrence in ChIP-seq datasets can't treat the motif overlapping, i. e. only motifs separated by a spacer of zero/positive length were considered as a potential CE (e. g. [4]).
Methods and Algorithms: We propose a new algorithm that can infer motif co-occurrence in ChIP-seq data without limitation for overlap. First, we compute the frequencies of anchor/partner motifs in a peak. Second, for each motif hit we count the number of overlapped motifs of the same type. These two measures help to generate the permuted sequences for a peak. We apply the Fisher's exact test to estimate the enrichment of the CE content in peaks in comparison with that in permuted data. Additionally, we use the Tomtom tool [5] to filter out possible overpredictions related to a significant match between the anchor and partner motifs.
Results: We analyzed more than hundred ChIP-seq datasets for about fifty TF types of mammals and plants and found that the majority (~95 %) of the overrepresented co-occurred motif pairs are overlapped. Our results are in a good accordance with earlier analysis of motif co-occurrence in specific cell lines [6] and the applicaion of *in vitro* SELEX modelling for cooperative TF binding [7].
Conclusion: We found that motifs overlap is widespread in ChIP-seq data. The application of our novel tool will substantially contribute to their careful annotation.
Acknowledgements: The work was supported by RFBR 18-04-01130 and ICG SB RAS budget project 0324-2018-0017.

References

1. Morgunova E., Taipale J. (2017) Structural perspective of cooperative transcription factor binding. Curr. Opin. Struct. Biol. 47:1-8.
2. Kulakovskiy I.V. et al. (2018) HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. Nucleic Acids Res. 46(D1):D252-D259.
3. Yevshin I. et al. (2017) GTRD: a database of transcription factor binding sites identified by ChIP-seq experiments. Nucleic Acids Res. 45(D1):D61-D67.
4. Whitington T. et al. (2011) Inferring transcription factor complexes from ChIP-seq data. Nucleic Acids Res. 39:e98.
5. Gupta S. et al. (2007) Quantifying similarity between motifs. Genome Biology. 8(2):R24.
6. Jankowski A. et al. (2013) Comprehensive prediction in 78 human cell lines reveals rigidity and compactness of transcription factor dimers. Genome Res. 23(8):1307-1318.
7. Jolma A. et al. (2015) DNA-dependent formation of transcription factor pairs alters their binding specificity. Nature. 527(7578):384-388.

DOI 10.18699/BGRSSB-2018-036

Mutation load in carotid paragangliomas

E. Lukyanova[1], A. Snezhkina[1]*, D. Kalinin[2], M. Fedorova[1], E. Pudova[1], Z. Guvatova[1], S. Kharitonov[1], N. Melnikova[1], G. Krasnov[1], A. Dmitriev[1], A. Kudryavtseva[1, 3]

[1] *Engelhardt Institute of Molecular Biology RAS, Moscow, Russia*

[2] *Vishnevsky Institute of Surgery, Ministry of Health of the Russian Federation, Moscow, Russia*

[3] *National Medical Research Radiological Center, Ministry of Health of the Russian Federation, Moscow, Russia*

** e-mail: leftger@rambler.ru*

Key words: carotid paragangliomas, exome, mutation load, high-throughput sequencing

Motivation and Aim: Carotid paragangliomas are tumors of head and neck that have neuroendocrine origin and arise at carotid bifurcation. High vascularity and location of the tumors make them a surgical challenge. Carotid paragangliomas are belonged to tumors with uncertain potential for malignancy; in 5–15 % of cases, they become to be malignant. Treatment options for patients with carotid paragangliomas (particularly with metastasis) are limited. The development the effective management strategy of these tumors based on early diagnosis and new therapeutic approaches is important. In this work, we performed the analysis of the mutation load (ML) in carotid paragangliomas. It was shown that in many tumors ML is associated with sensitive to immunotherapy, which is very promising for tumor treatment [1].

Methods and Algorithms: In the study, we used the collection of blood and tissues (tumors and lymph nodes) derived from 12 patients with carotid paragangliomas at Vishnevsky Institute of Surgery, Ministry of Health of the Russian Federation. The isolation of DNA from the samples was performed using FFPET DNA Isolation Kit (Roche, Switzerland). DNA from blood was isolated with MagNA Pure Compact Nucleic Acid Isolation Kit I (Roche) on MagNA Pure Compact Instrument (Roche). Exome library preparation was performed using TruSeq Exome Kit (Illumina, USA) according to the manufacturer's protocol. Sequencing of exome libraries was carried out on NextSeq 500 System (Illumina) with 2x75 bp paired-end reads at EIMB RAS "Genome" center (http://www.eimb.ru/rus/ckp/ccu_genome_c.php).

Results: The average ML was estimated for examined samples as a number of mutations per megabase of coding regions and was compared with the data obtained [3]. The values were approximately equal (~7/Mb).

Conclusion: Additional analysis of exome data from tumors with paired blood and lymph nodes tissues allowed accurately estimating ML in carotid paragangliomas.

Acknowledgements: This work was financially supported by the Russian Foundation for Basic Research (grant 16-04-01521a), ICGEB project CRP/RUS15-01, and Presidential Scholarship (SP-4247.2018.4).

References

1. Alexandrov L.B. et al. (2013) Signatures of mutational processes in human cancer. Nature. 500(7463): 415-421.
2. Snezhkina A.V. et al. (2018) Exome analysis of carotid body tumor. BMC Medical Genomics. 11(Suppl 1):17.

DOI 10.18699/BGRSSB-2018-037

Multiple omics ageing clocks

E. Macdonald-Dunlop[1]*, P.K. Joshi[1], J.F. Wilson[1, 2]
[1] *Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Teviot Place, Edinburgh, UK*
[2] *MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh, UK*
** e-mail: e.macdonald.dunlop@ed.ac.uk*

Key words: ageing, omics, penalized regression, biological age, healthy ageing, multi-omics

Biological age, a measure of deterioration and ageing that is distinct from chronological age has been found to predict disease and mortality [1]. The first published measure of biological age was Hannum's epigenetic clock. Hannum's clock took the form of an elastic net regression model predicting age, built from whole blood CpG methylation data, the ratio of this predicted age to chronological age was used to measure apparent methylomic ageing rate (AMAR) [2]. Hannum's work was extended by Horvath using the same methodology, but a larger number of: CpG markers, individuals and tissue types [3].

Since the publication of these landmark epigenetic clocks, ageing clocks have been built using telomere length [4], facial morphology [5], metabolomics [1], glycomics [6] and proteomics [7]. Each of these ageing clocks have shown that chronological age compared to biological age can be used as an indicator of health outlook. These models have the potential to inform health and lifestyle advice in order to improve individuals' health [7]. Here we test replication of published ageing clocks using approximately 1,000 individuals from the cross-sectional population cohort ORCADES, that are highly annotated with 877 phenotypes spanning glycomics, lipidomics, metabolomics and proteomics.

References

1. Fischer K. et al. (2014) Biomarker profiling by nuclear magnetic resonance spectroscopy for the prediction of all-cause mortality: An observational study of 17,345 persons. Plos Medicine. 11.
2. Hannum G. et al. (2013) Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol. Cell. 49:359-367.
3. Horvath S. (2013) DNA methylation age of human tissues and cell types. Genome Biology. 14.
4. Zhang W.G. et al. (2014) Select aging biomarkers based on telomere length and chronological age to build a biological age equation. Age. 36:1201-1211.
5. Chen W.Y. et al. (2015) Three-dimensional human facial morphologies as robust aging markers. Cell Research. 25:574-587.
6. Kristic J. et al. (2014) Glycans are a novel biomarker of chronological and biological ages. J. Gerontol. Ser. a-Biol. Sci. Med. Sci. 69:779-789.
7. Enroth S., Enroth S.B., Johansson A., Gyllensten U. (2015) Protein profiling reveals consequences of lifestyle choices on predicted biological aging. Sci. Rep. 5.

DOI 10.18699/BGRSSB-2018-038

Selfish elements drive mitochondrial and nuclear genome size in opposite directions

A.A. Mikhaylova[1]*, A.G. Mikhaylova[1], K. Ushakova[1], D. Knorre[2], I. Mazunin[1], A. Reymond[3], K. Gunbin[1, 4], K. Popadin[1, 3]

[1] *School of Life Science, Immanuel Kant Federal Baltic University, Kaliningrad, Russia*
[2] *Belozersky Institute of Physico-Chemical Biology, Moscow State University, Russia*
[3] *Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland*
[4] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
** e-mail: mihailovaalina777@yandex.ru*

Key words: comparative genomics, mitochondrial DNA, repeats, mammals

Motivation and Aim: Nuclear genome size (nucDNA) is higher in long-lived chordata species, where selection against expansion of selfish elements is relaxed due to: (i) lower effective population size (Ne), (ii) higher cell volume, (iii) extended cell cycle period, etc. Mitochondrial genome (mtDNA) has one more, intracellular level of selection, which might change evolutionary forces maintaining the genome size.
Methods and Algorithms: To understand genetic-ecological correlations better we analyzed 840 complete mtDNAs of Mammalia. For each genome we derived dozens of genetic traits (genome length; GC content; abundance, skewness and densities of direct, symmetrical, inverted, complementary and tandem repeats; GC/AT skew etc) and correlated them with several ecological and physiological traits (longevity, body mass, metabolic rate, fecundity etc) using comparative methods like PIC and PGLS.
Results: Analysing 840 complete mitochondrial genomes of mammals species we observed that: (i) genome length variation is explained mainly by variation in control region, which is driven by the abundance of tandem repeats; (ii) tandem repeats are more common in short-lived species; (iii) tandem repeats are AT rich. Putting together all these observations we postulate that mtDNAs of short-lived species are longer and AT-richer as compared to mtDNAs of long-lived species. We explain these observations by two non-mutually exclusive hypotheses: (i) short-lived species have an evolutionary force to maintain fast-replicating mtDNA, i.e. to be more AT rich and to have tandem repeats which can promote fast replication time, simulating or strengthening the origin of replication site [1]; (ii) long-lived species have an evolutionary force to maintain stable mtDNA to decrease somatic mutation rate (high GC content; no repeats which introduce genomic instability) [2]. Since selfish elements propagate more effectively in mtNDA of short-lived species and in nucDNA of long-lived species, we expect and we do observe a negative correlation between the mtDNA size and nucDNA size, emphasizing that opposite forces affects evolution of mtDNA and nucDNA.
Acknowledgements: The study was supported by the 5 Top 100 Russian Academic Excellence Project at the Immanuel Kant Baltic Federal University.

References

1. Ma H., O'Farrell P.H. (2016) Selfish drive can trump function when animal mitochondrial genomes compete. Nature Genetics. 48(7), 798.
2. Lehmann G., Segal E., Muradian K.K., Fraifeld V.E. (2008). Do mitochondrial DNA and metabolic rate complement each other in determination of the mammalian maximum longevity? Rejuvenation Research. 11(2):409-417.

DOI 10.18699/BGRSSB-2018-039

Expression of *Fusarium oxysporum* genes upon infection of *Linum usitatissimum* plants

R. Novakovskiy[1]*, G. Krasnov[1], T. Rozhmina[1, 2], P. Kezimana[1, 3], N. Melnikova[1], A. Dmitriev[1]

[1] *Engelhardt Institute of Molecular Biology RAS, Moscow, Russia*
[2] *All-Russian Research Institute for Flax, Torzhok, Russia*
[3] *Peoples' Friendship University of Russia* (*RUDN University*)*, Moscow, Russia*
* *e-mail: 0legovich46@mail.ru*

Key words: *Fusarium oxysporum*, flax, gene expression, high-throughput sequencing, plant pathogen

Motivation and Aim: *Fusarium oxysporum* sp. *lini*, which leads to fusarium wilt, is one of the most harmful flax (*Linum usitatissimum* L.) pathogens. Flax cultivars are varied in their resistance to fusarium wilt [1]. Specific expression alterations were revealed for a number of genes in resistant flax genotypes compared to susceptible ones under inoculation of *F. oxysporum* [2]. However, the question of whether there is any difference in expression of the fungus genes, when the pathogen infects genotypes with diverse resistance, remains open. In the present study, we performed such investigation.

Methods and Algorithms: Seedlings of resistant (Dakota and #3896) and susceptible (AP5 and TOST) to *F. oxysporum* flax cultivars were inoculated with the pathogen isolate #39 from the collection of the All-Russian Research Institute for Flax. RNA was isolated from plant roots in duplicate for each variety. TruSeq Stranded Total RNA Sample Prep Kit (Illumina) was used for cDNA library preparation and high-throughput sequencing on NextSeq500 sequencer (Illumina) using 80-nucleotide pair-end reads was performed. The number of reads for each *F. oxysporum* transcript was assessed, and expression analysis was performed using the edgeR.

Results: Eight cDNA libraries, which were obtained from roots of inoculated with *F. oxysporum* flax cultivars with diverse resistance, were sequenced, and about 25 million pair-end reads were obtained for each cultivar. About 30–40 % of reads were mapped on genome of *F. oxysporum*. The similarity of gene expression profiles was higher within groups of varieties with similar resistance to the pathogen. *F. oxysporum* genes with the greatest differences in expression level upon inoculation of resistant and susceptible flax genotypes were identified.

Conclusion: The study of *F. oxysporum* expression changes when the fungus infects flax genotypes with diverse resistance is important for understanding of plant-pathogen interactions. Data on genes with distinct expression when pathogen infects resistant and susceptible cultivars bring new insights into mechanisms of fungal infection.

Acknowledgements: This work was financially supported by the Russian President Grant MK-5828.2018.4.

References

1. Diederichsen A. et al. (2008) Variation patterns within 153 flax (*Linum usitatissimum* L.) genebank accessions based on evaluation for resistance to fusarium wilt, anthracnose and pasmo. Plant Genetic Resources: Characterization and Utilization. 6(1):22-32.
2. Dmitriev A.A. et al. (2017) Differential gene expression in response to *Fusarium oxysporum* infection in resistant and susceptible genotypes of flax (*Linum usitatissimum* L.). BMC Plant Biology. 17(Suppl. 2):253.

DOI 10.18699/BGRSSB-2018-040

Evolutionary analysis of 3-dimensional chromatin structure

M. Nuriddinov[1]*, V. Fishman[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: nuriddinov@bionet.nsc.ru*

Key words: Hi-C, chromatin architecture, spatial chromosome organization, genome evolution

Motivation and Aim: Development of the genome-wide chromatin conformation capture techniques such as the Hi-C method allowed discovering of the new layer of spatial chromatin organization, formed by the closely interacting chromosomal regions, in other words, topologically associated domains (TADs). The TAD architecture appears to play important roles in gene regulation, transcription and replication [1]. Furthermore, chromatin domain structure reveals high conservation across both cell types within one organism and evolutionary lineages [1, 2]. Nevertheless, there have been a number of Hi-C data of various species, this are not used to investigate evolutionary changes of chromatin architecture in details. To bridge the gap, we have developed bioinformatic tool allowing cross evolution comparison of chromatin structure.
Methods and Algorithms: We developed a python-based software that uses standard output from HiC-Pro pipeline [3] to compute frequencies of 3-dimentional chromatin contacts in each specie. The software estimates variance of contact frequencies by binomial distribution and normalizes contact counts by vanilla coverage. To compare contact frequencies between species, our algorithm solves several problems:
– finding orthologues contacts of studied species;
– accounting for species-to-species changes of linear distance between contacting loci;
– computing expected contact frequency for one specie based on observed contact frequency for the other specie;
– estimating significance of difference or similarity between observed and expected contact frequencies.
Results: We developed the software that compares Hi-C data of two species and discriminating differential and conserved chromatin interactions. The software can visualize obtained results in a heatmap-like format. As a proof-of-principle, we compared chromatin architecture in mouse and human and showed conservation of spatial contacts around conserved TAD boundaries.
Conclusion: Described bioinformatic tool allows evolutionary comparison of 3-dimentional chromatin structure. This gives new approaches to process chromosome conformation data, exploration of evolution hot spots and identification of regulatory elements. This could give fresh insights to genome evolution and molecular mechanisms forming chromosomal architecture.
Acknowledgements: This study was supported by the RSF (grant 17-74-10143)

References

1. Dixon J.R. et al. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 485:376-380.
2. Battulin N. et al. (2015) Comparison of the 3D organization of sperm and fibroblast genomes using the Hi-C approach. Genome Biol. 16:77.
3. Servant N. et al. (2015) HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol. 16:259.

DOI 10.18699/BGRSSB-2018-041

Two-step emulsion PCR to prevent formation of the chimeric molecules

E. Omelina[1]*, A. Ivankin[1], A. Leshchenko[1, 2], L. Yarinich[1], M. Lebedev[1, 2], A. Pindyurin[1]

[1] *Institute of Molecular and Cellular Biology SB RAS, Novosibirsk, Russia*

[2] *Novosibirsk State University, Novosibirsk, Russia*

** e-mail: omelina@mcb.nsc.ru*

Key words: emulsion PCR, chimeric molecules, next-generation sequencing

Motivation and Aim: Application of the conventional PCR method for the amplification of the random nucleotide sequence libraries often causes a formation of the undesired chimeric molecules. Invention of the water-in-oil emulsion PCR (ePCR) approach allowed to reduce the probability of the chimeric molecule formation compared to the conventional PCR. However, in the non-optimized conditions even ePCR causes the formation of the chimeric products.

Methods and Algorithms: We developed a step-by-step protocol for the ePCR consisting of two consequent rounds. We found that both an initial amount of the DNA template and number of amplification cycles play a critical role in the formation of the chimeric molecules. We suggest to use only 10^6 DNA molecule copies for the first round of ePCR.

Results: We analyzed a formation of the chimeric products during amplification of heterogeneous plasmid library in different conditions. To assess the percent of the formed chimers we used Illumina MiSeq platform. A proportion of chimeric molecules under optimal conditions was lower than 0.25 %. We suppose that this two-step ePCR approach may be useful for the preparation of heterogeneous sequence libraries for the next-generation sequencing and other issues which demand avoiding of the chimeric molecule formation.

Conclusion: ePCR approach allows to separate the DNA template molecules from each other using the water-in-oil emulsion. Our data suggest that ePCR approach is suitable for the preparation of random DNA libraries for the next-generation sequencing. However it requires additional adjustment to reduce a formation of chimeric molecules as much as possible.

Acknowledgements: Supported by the RSF (16-14-10288).

DOI 10.18699/BGRSSB-2018-042

Tandem repeats in mammalian genomes

D. Ostromyshenskii[1]*, O. Podgornaya[1, 2]
[1] *Institute of Cytology RAS, Saint Petersburg, Russia*
[2] *Saint Petersburg State University, Saint Petersburg, Russia*
** e-mail: necroforus@gmail.com*

Key words: tandem repeat, satellite DNA, heterochromatin

Motivation and Aim: Heterochromatin definitely plays an important role in chromatine 3D structure, but the detailed mechanisms not determined. Tandemly repeated sequences are the DNA class, which is absent in prokaryote but appear in eukaryotes. Tandem repeats (TR, or satDNA – satellite DNA) is the mostly fast evolving genome component. It is impossible to determine TR role and functions in the genome housekeeping without their classification and annotation. We are going to classify TR in the mammalian genomes with similar degree of assembly available in databases and to check the pattern of TR evolution *in silico.*

Methods and Algorithms: The modified pipeline of the one used previously in our Lab [1] applied to the several mammalian genomes. For TR searching we used mammalian genomes assembly from www.ncbi.nlm.nih.gov/assembly with N50 metrics > 40 Kb. Totally we searched TR in assembled genomes of 57 species of 11 mammalian orders. For each genomes we accounted consensus sequence of each TR family, monomers length, GC-content and variability within the TR arrays. For *Mus musculus*, *Cricetulus griseus*, *Mesocricetus auratus* and *Sus scrofa* we confirmed in silico prediction by molecular biology methods such as FISH and PCR.

Results: We found full sets of TR in 57 mammalian genomes assembly. Comparison of sets of TR in the different genomes showed that despite different primary sequences of TR of different genomes, the distribution of TR within each genome according to GC-content, monomer length and variability within the TR arrays is similar. Almost in all genomes, the family of major TR has been found. In different groups of mammals, it can be formed as AT-rich TR (for example, rodents and primates) or GC-rich TR (for example, carnivorans and artiodactyls), but always it is characterized by organization in high order repeats, long monomers, centromeric-pericentromeric localization. The second family of TRs by representation in genome is often formed by TR with a relatively short monomer, more complex arrays and has chromosome-specific variants (for example, the HS1-4 human satellite or the mouse TRPC-21A-MM family).

Conclusion: By comparison of the TR set from the genomes we show that the TR fields distribution is similar in the genomes. Such a set reflects the “bar-code”, which predetermine the hierarchy of chromosome domains associations. We expect that heterochromatic “bar-code” made up of the TR could be the base for the hypothetical General Morphogenetic program attributed to the heterochromatin.

Acknowledgements: This work was supported by the granting program of Russian Foundation for Basic Research (No. 18-34-00238) and the granting program ‘Molecular and cell biology’ of the Russian Academy of Sciences (No. 01.2.01457147).

References

1. Komissarov A.S., Gavrilova E.V., Demin S. et al. (2011) Tandemly repeated DNA families in the mouse genome. BMC Genomics. 12(1):531-552.

DOI 10.18699/BGRSSB-2018-043

Electrostatic up-element and promoter strength

A. Osypov[1, 2]*, G. Krutinin[3], E. Krutinina[3], P. Beskaravayny[3], S. Kamzolova[3]
[1] *Institute of Higher Nervous Activity and Neurophysiology RAS, Moscow, Russia*
[2] *Institute of Theoretical and Experimental Biophysics RAS, Pushchino Moscow Region, Russia*
[3] *Institute of Cell Biophysics RAS, Pushchino Moscow Region, Russia*
** e-mail: aosypov@gmail.com*

Key words: DNA electrostatics, transcription regulation, genome evolution

Motivation and Aim: It is known that not only the consensus sequence text is essential for RNA polymerase-promoter recognition and regulation, but physical properties, especially electrostatics, play important role particularly at the early stages of this process. One of the elements that may play a crucial role in the promoter strength regulation is a so-called "up-element", which interacts with the alpha-subunit of RNAP and thus facilitates its binding to the promoter. There is no text consensus in the "up-element" (though high AT content is often attributed) and functionality of this region is defined by its physical properties. We have shown earlier, that electrostatics is responsible for its functioning during the global transcription switch under the T4 bacteriophage infection and that strong T4, early T7-like, phage *Lambda* and *E. coli* ribosomal promoters with pronounced up-element have high levels of the electrostatic potential within it.
Methods and Algorithms: DEPPDB and its tools [1, 2] were used to carry out the analysis. Mutant promoters' sequences and strength are taken from [3, 4].
Results: In the strong E. coli ribosomal rrnB P1 promoter and its up-element mutants the promoter strength depends upon the size of the electrostatic up-element so that the bigger the element – the stronger the promoter is. However, if the element is too big the strength decreases slightly due to possible trapping of the polymerase that hinders the transcription initiation or elongation.
Conclusion: Electrostatics may play important role in the transcription regulation in the first step of initial promoter recognition and the second step of the transcription initiation due to interactions between the RNA-polymerase alpha subunit and the promoter up-element that has electrostatic nature.
Acknowledgements: Supported by the RFBR (16-04-01865 and 18-34-00942).

References

1. Osypov A.A. et al. (2010) DEPPDB – DNA Electrostatic Potential Properties Database: Electrostatic properties of genome DNA. JBCB. 8(3):413-425.
2. Osypov A.A. et al. (2012) DEPPDB – DNA Electrostatic Potential Properties Database: Electrostatic properties of genome DNA elements. JBCB. 10(2):1241004.
3. Estrem S.T. et al. (1998) Identification of an UP element consensus sequence for bacterial promoters. Proc. Natl. Acad. Sci. USA. 95(17):9761-9766.
4. Estrem S.T. et al. (1999) Bacterial promoter architecture: subsite structure of UP elements and interactions with the carboxy-terminal domain of the RNA polymerase alpha subunit. Genes Dev. 13:2134-2147.

DOI 10.18699/BGRSSB-2018-044

Selecting codons not for proteins: codons and amino acids biases around proteins binding DNA sites are due to their electrostatics

A. Osypov[1, 2]*, G. Krutinin[3], E. Krutinina[3], P. Beskaravayny[3], S. Kamzolova[3]
[1] *Institute of Higher Nervous Activity and Neurophysiology RAS, Moscow, Russia*
[2] *Institute of Theoretical and Experimental Biophysics RAS, Pushchino, Moscow Region, Russia*
[3] *Institute of Cell Biophysics RAS, Pushchino, Moscow Region, Russia*
** e-mail: aosypov@gmail.com*

Key words: DNA electrostatics, transcription regulation, genome evolution

Motivation and Aim: DNA is a highly charged molecule and its electrostatic and other physical properties define its interactions with different proteins, especially those regulating transcription in Prokaryotes. Electrostatic potential (EP) is distributed non¬uniformly along DNA and correlates with GC content, depending on the sequence arrangement and its context [1]. Binding sites of transcription factors of different protein families in different taxa are located in wide areas of high electrostatic potential, multiple times of the protein size; EP distribution on transcription factors protein surface reflects that of their binding sites [2]. Promoters in average have high value of EP profile [1]. Hundreds of transcription factors binding sites lie in protein coding areas. Some protein coding areas also host promoters for the genes mainly located on the opposite strand.
Methods and Algorithms: DEPPDB and its tools [1, 2] were used to carry out the analysis.
Results: We found that in such diverse bacteria as *E. coli*, *B. subtilis* and *Corynebacterium glutamicum* (and other species) there are codons and even amino acids biases around promoters and transcription factors binding sites spanning for a hundred of codons and that these biases are due to physical properties of considered codons providing proper electrostatic attractors for transcription regulating proteins.
Conclusion: The data obtained demand serious rethinking of the concept of molecular evolution. The apparent non-synonymy of synonymous substitutions may lead to different wrong estimations of the sequences fate, including misuse of the molecular clock and mistaken evaluations of specific mutations biomedical importance. The amino acids bias leads to even more important shift in the conception of natural selection in proteins. It substantiates the view on the DNA as not only the text of the first step in realizing the Central Dogma, but a complex organ of heredity that fulfills different and sometimes contradictory demands.
Acknowledgements: Supported by the RFBR (16-04-01865).

References

1. Osypov A.A. et al. (2010) DEPPDB – DNA Electrostatic Potential Properties Database: Electrostatic properties of genome DNA. JBCB. 8(3):413-425.
2. Osypov A.A. et al (2012) DEPPDB – DNA Electrostatic Potential Properties Database: Electrostatic properties of genome DNA elements. JBCB. 10(2):1241004.

DOI 10.18699/BGRSSB-2018-045

Transcriptomic comparative analysis of hippocampal tissue and primary cultures after hyaluronidase treatment

V. Pershin[1, 2]*, A. Balashova[2], E. Guryev[1, 2], L. Kurbatov[3], M. Gainullin[1, 4], I. Mukhina[1, 2]

[1] *Privolzhsky Research Medical University, Nizhny Novgorod, Russia*
[2] *Lobachevsky State University, Nizhny Novgorod, Russia*
[3] *Orekhovich Institute of Biomedical Chemistry, Moscow, Russia*
[4] *Oslo University Hospital, Oslo, Norway*
** e-mail: bp1995@yandex.ru*

Key words: Extracellular matrix, epilepsy, hippocampus, gene expression, transcriptomics

Motivation and Aim: It is well-known that traumatic injury or enzymatic digestion of brain extracellular matrix can cause seizure-like activity in primary hippocampal culture [1]. However, underlying molecular mechanism and similarity between *in vitro* and *in vivo* models of epilepsy is poorly investigated. The aim of this study was to investigate changes of gene expression and perform comparative transcriptomic analysis while modeling of epileptogenesis.

Methods and Algorithms: C57BL/6J mice were used to prepare hippocampal cell culture and intra-hippocampal hyaluronidase injection. Hyaluronic acid, which is the basis of extracellular matrix, has been removed by 75U/ml hyaluronidase. Transcriptomic analysis was performed by mouse full-genome 2-colour Agilent Microarrays. Normalization, processing and detection of differentially-expressed genes were performed by online free software – Babelomics 5.0 [2]. Functional annotation clustering was made by David bioinformatics resource [3].

Results: Hyaluronan digestion resulted in significant changes of expression of 3 gene clusters in neuronal hippocampal cultures – synaptic gene cluster, ribosomal gene cluster and mitochondrial gene cluster, while in hippocampus hyaluronidase injection resulted in changes of such functional gene clusters: synaptic gene cluster, neurogenesis gene cluster, actin-binding gene cluster, circadian rhythm cluster and ubiquitinylation gene cluster. When comparing control groups of *in vivo* and in vitro samples there were found 3 significant gene clusters: calcium ion-binding cluster, cluster of genes expressing EGF-domains and cortical cytoskeleton gene cluster.

Conclusion: The results obtained in this study suggest that hyaluronic acid digestion mainly affect synaptic gene cluster that indicate to similarity of *in vivo* and in vitro models. A wide variety of gene clusters from isolated hippocampus point to high complexity of biological processes in mice brain. Transcriptomic comparative analysis revealed strong difference between cell culture and hippocampus that should be considered when interpreting in vitro data.

Acknowledgements: Supported by the RFBR (16-04-01763).

References

1. Vedunova M.V. et al. (2013) Seizure-like activity in hyaluronidase-treated dissociated hippocampal cultures. Front. Cellular Neurosci. 7(149):81-89. DOI 10.3389/fncel.2013.00149.
2. Alonso R. et al. (2015) Babelomics 5.0: functional interpretation for new generations of genomic data. Nucleic Acids Res. 43(1):117-121.
3. Huang D.W. et al. (2009) Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nat. Protoc. 4(1):44-57.

DOI 10.18699/BGRSSB-2018-046

Identification of 3 unknown bacterial strains, characterization of copper oxidase genes

M. Plotnikova
Voronezh State University, Department of Genetics, Cytology and Bioengineering, Voronezh, Russia
e-mail: j.mariia@rambler.ru

Key words: copper oxidase, genome sequence, 16R RNA, identification of strains

Motivation and Aim: The purpose of this research is identification of unknown strains of microorganisms and characterization of Copper Oxidase Genes in their genomes.
Methods and Algorithms: Objects of research are unknown strains from RCM (Russian Collection of Microorganisms, Pushchino). Bacterial cells were provided by the staff of the RCM department (http://www.vkm.ru) of the Scryabin Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences. DNA strains were isolated using the Genomic DNA Purification Kit (Thermo Fisher Scientific) by the manufacturer's procedure with modifications: treatment with lysozyme cells before lysis solution was applied. The fragment of the 16s rRNA gene was amplified using specific primers 27F (AGA GTT TGA TCC TGG CTC AG) and 1492R (ACG GYT ACC TTG TTA CGA CTT). Polymerase chain reaction included denaturation of the matrix in the first cycle at 95 °C for 5 min and 35 cycles of amplification in the following regime: 95 °C – 30 s, 55 °C – 30 s, 72 °C – 1 min. PCR products were analyzed in a 1.3 % agarose gel with ethidium bromide (0.0002 %). PCR products of about 1500 bp were purified from the gel using the Cleanup Mini kit (Eurogen, Russia) and transferred to Eurogen to determine the nucleotide sequence.
Results: Using the methods of molecular biology, we identified 3 unknown strains of microorganisms from RCM. After analysis of the nucleotide sequence, it was found that the strains studied belong to the following species: strain 1 – Cellulosimicrobium cellulans (99 %); strain 2 – Cellulosimicrobium funkei (98 %).
Conclusion: The strains of microorganisms from the RCM were identified. Nowadays, we research the expression of Copper Oxidase Genes in these genomes (Cellulosimicrobium cellulans, Cellulosimicrobium funkei).
Acknowledgements: Supported by the Voronezh State University and Scryabin Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences.

DOI 10.18699/BGRSSB-2018-047

Potential biomarkers associated with poor prognosis for locally advanced prostate cancer without lymphatic dissemination

E. Pudova[1]*, E. Lukyanova[1], K. Nyushko[2], S. Kharitonov[1], A. Snezhkina[1], M. Fedorova[1], A. Kaprin[2], B. Alekseev[2], A. Kudryavtseva[1, 2]

[1] *Engelhardt Institute of Molecular Biology RAS, Moscow, Russia*

[2] *National Medical Research Radiological Center, Ministry of Health of the Russian Federation, Moscow, Russia*

** e-mail: pudova_elena@inbox.ru*

Key words: prostate cancer, prognostic biomarkers, TCGA

Motivation and Aim: Prostate cancer (PC) is a major cause of cancer-related deaths in men worldwide. Locally advanced prostate cancer (LAPC) is characterized by invasion of the prostatic capsule without evidence of nodular or distant metastatic spread. Patients within this clinical category have different risks of recurrence [1, 2]. Our study is aimed at identifying novel prognostic biomarkers for LAPC, which will lead to optimization of treatment and development of appropriate clinical recommendations.

Methods and Algorithms: We performed bioinformatics analysis of The Cancer Genome Atlas (TCGA) project RNA-Seq data using the computational resources of EIMB RAS "Genome" center (http://www.eimb.ru/rus/ckp/ccu_genome_c.php); 130 samples (only derived from Caucasian patients) of LAPC without lymphatic dissemination were divided into two groups depending on the cancer prognosis (favorable or unfavorable). In the analysis, the genes previously identified by published data as potential prognostic biomarkers were considered.

Results: Additional eight genes, that are characterized by differential expression, were found. Increased expression of *TWIST1*, *TUBB3*, and *CHAT* and decreased expression of *CYP1B1*, *IGSF1*, *EDN3*, *MSMB*, and *SERPINA3* were detected in the group of patients with poor prognosis in comparison to one with favorable prognosis. These genes are involved in the key processes of carcinogenesis, such as angiogenesis, proliferation and migration of tumor cells (*CYP1B1*, *EDN3*, *TWIST1*, *TUBB3*, and *CHAT*), as well as with disturbance of regulation of signaling cascades (*IGSF1*, *SERPINA3*, and *MSMB*).

Conclusion: Thus, we identified a number of genes as potential prognostic biomarkers associated with poor prognosis of LAPC without lymphatic dissemination.

Acknowledgements: This work was funded by the Program of fundamental research for state academies for 2013–2020 years (No. 1201363819) and by RFBR according to the research project No. 17-29-06083.

References

1. Nyushko K.M. (2014) Surgical treatment of patients with prostate cancer with a high risk of progression. Oncourology. 2:38-45.
2. Alekseev B.Ya. (2012) Combined treatment of prostate cancer: clinical observations. Oncourology. 4:77-87.

DOI 10.18699/BGRSSB-2018-229

Genomes of three conifer species: *Larix sibirica*, *Pinus sibirica* and *Pinus sylvestris*

Yu. Putintseva[1]*, V. Sharov[1], D. Kuzmin[1], N. Oreshkova[1, 2], S. Feranchuk[3, 4], V. Biryukov[1], S. Novikova[1], K. Miroshnikova[1], S. Makolov[1], M. Sadovsky[1], K. Krutovsky[1, 5, 6, 7]

[1] *Siberian Federal University, Krasnoyarsk, Russia*
[2] *V.N. Sukachev Institute of Forest SB RAS, Krasnoyarsk, Russia*
[3] *National Research Technical University, Irkutsk, Russia*
[4] *Limnological Institute, SB RAS, Irkutsk, Russia*
[5] *George-August University of Göttingen, Göttingen, Germany*
[6] *Vavilov Institute of General Genetics RAS, Moscow, Russia*
[7] *Texas A&M University, College Station, USA*
** e-mail: yaputintseva@mail.ru*

Key words: conifers, *de novo* assembly, genome, Siberian larch, Siberian pine, Scots pine

Motivation and Aim: The enormous size of conifer genomes makes the task of sequencing them extremely difficult because of both amount of sequencing data needed for good coverage and huge computer resources needed to process them. However, the reference genomes of conifers are much needed for further studies of evolutionary, biochemical and physiological processes in these organisms. Only a few nuclear genomes of conifers have been sequenced and published so far. We recently sequenced and assembled nuclear genomes in three conifer species – *Larix sibirica*, *Pinus sibirica* and *P. sylvestris*.
Methods and Algorithms: Sequencing was performed using the Illumina HiSeq2000 and MiSeq platforms. We developed an original stepwise method of *de novo* assembly by parts (sets), which allowed us to bypass the limitations of modern assemblers associated with a huge amount of data being processed. We performed the scaffolding with program BESST and scaffolding with RNA reads using the RaScaf program. The gap-filling program Sealer was used also to improve the assembly. RepeatModeler and RepeatMasker were used to identify repeats. Genome annotation was performed using the MAKER2 pipeline.
Results: The Siberian larch genome assembly contained 12.34 Gbp with N50 of 6,443 bp and average GC of 35.41 %. The Siberian pine draft assembly contained 13.56 Gbp with N50 of 6,920 bp and average GC of 36.6 %. The Scots pine draft assembly contained 14.79 mln scaffolds with a total length of 7,8 Gbp, N50 of 654 bp.
Conclusion: Using the new stepwise *de novo* assembling method, the genome of Siberian larch, *Larix sibirica* Ledeb. (12.34 Gbp) was for the first time completely assembled *de novo*. It is the first genome assembly for any larch species. We also present draft genome assemblies for two pine species – *P. sibirica* and *P. sylvestris*.
Acknowledgements: This work was funded by a research grant no. 14.Y26.31.0004 from the Government of the Russian Federation.

DOI 10.18699/BGRSSB-2018-048

Graph-oriented database for analysis of prokaryotic communities -*omics* data

A. Ryasik[1]*, T. Ermak[1, 2], M. Orlov[1], E. Zykova[1], A. Sorokin[1]
[1] *Institute of Cell Biophysics RAS, Pushchino, Russia*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
** e-mail: ryasik.aa@gmail.com*

Key words: metabolic reconstruction, metabolic model assembly, -*omics* data integration, metagenomics

Motivation and Aim: We have constructed a graph database containing information about proteome, metabolome, transcription and translation regulation in various prokaryotes, that is suitable as for analysis of physiology of microbial communities, as for analysis of strain and isolates differences within species. All the data were taken from well-known external databases (Genbank, Uniprot, RAST, etc.). Advantages of data storage in graph structures allowed us to construct metabolic models of the organisms and strains based on the similarity with reference genomes. The aim of the research is the automatic assembly of the metabolic network flux models for bacteria by their protein similarity to reference genome, preparation of metabolic networks for the communities and comparative analysis of these models that can be applied for the distinction of the strains and prediction of physiological abilities of the communities.
Methods and Algorithms: We recently developed the BioGraph database, graph-oriented storage for information about prokaryotic organisms. We collected various -*omics* data (genomic, proteomic, taxonomic) and integrated it by graph representation with predefined types of nodes and edges. This approach was used to create strict rules for integration of -*omics* data of various origin. The implemented network of structural and semantic similarity relationships between proteins, genes, organisms, and reactions enables constructing models for close strains and species using the reference model as a template. Furthermore, such approach makes it possible to extract the existing model from the database as an SBML file. The developed tool was applied for analysis of different bacterial societies such as human microbiome. Also, we build a system of asynchronous queries based on actors for complicated comparison between the large number of nodes with multiple conditions, which could not be completed with online queries. Also, asynchronous queries allow fetching the data from remote sources such as UniProt to keep the BioGraph up to date.
Results: We developed a graph database that contains different -*omics* data; developed tools that allow to assemble metabolic flux models in automatic mode from different -*omics* data such as genomic annotation or proteins and genes similarity; tools for complex graph queries and database updates.
Availability: Source code of the software is available on Github: https://github.com/arc7an/scalaBiomeDB

DOI 10.18699/BGRSSB-2018-049

Genomic reconstruction of histidine metabolism and regulation in human microbiome

N.V. Sernova, I.A. Zharov, G.A. Ashniev, S.A. Leyn, D.A. Rodionov*
Institute for Information Transmission Problems RAS, Moscow, Russia
** e-mail: rodionovdmitry.delmar@gmail.com*

Key words:

Introduction: Histidine is one of the essential amino acids in mammals. Prokaryotes, fungi, and plants can synthesize it using a common biosynthesis pathway. Histidine biosynthetic pathway in Bacteria consists of ten biochemical steps. In *Escherichia coli* and other gamma-proteobacteria, the histidine biosynthesis genes are organized into the *his* operon, which is regulated via the leader peptide-dependent transcription attenuation mechanism. However, the histidine-dependent transcriptional regulation is not described in Gram-positive bacteria. The human microbiota represents a complex assemblage of microbial species and a significantly larger set of gene functions that are organized into complex metabolic and transcriptional networks. Bacteria from the human microbiome outnumber the number of human cells within a person by an order of magnitude and are represented by a large number of diverse taxa. Mapping gene functions and biological networks in the human microbiome is critically important for our understanding of mutualistic relationships including both microbe-microbe and host-microbe interactions and its relation to diseases such as obesity, diabetes, Crohn's disease.
Results: Genome-scale mapping and reconstruction of metabolic pathways and transcriptional regulatory networks in taxonomically diverse microbes is one of the critical tasks of microbial genomics. Here we present a study of distribution of histidine biosynthetic and transport genes across the reference set of 2228 bacterial genomes from 11 phyla out of sequenced consortia of Human Gut Microbiome. By using the comparative genomics approach we have reconstructed the biochemical pathways and transcriptional regulons for histidine metabolism in HMP genomes. The majority of histidine prototrophs belong to the *Actinobacteria*, *Proteobacteria* and *Bacterodetes* phyla that contain 90 %, 96 % and 76 % of prototrophic genomes, respectively. We have performed the comparative genomics reconstruction of regulons for a novel histidine-specific transcriptional regulator HisR that is homologous to the TrpR family of tryptophan-sensing repressors. Orthologs of HisR were found in Firmicutes, Xanthomonadales and some Alphaproteobacteria. In all genomes possessing HisR orthologs, we have identified DNA binding motifs of HisR, palindromes with consensus YACTTTANYNNRNTAAAGTR, and reconstructed the respective HisR regulons that mostly include the histidine biosynthetic operons and putative histidine transporters The regulon analysis allowed us to identify novel types of transporters that are potentially involved in histidine uptake.
Conclusions: The comparative genomics-based inference of histidine metabolism, uptake and regulation networks allows improve a metabolic modelling and provide a reference dataset for interpretation of transcriptomics data for human microbiota.
Acknowledgments: This research was supported by the Russian Science Foundation (grant No. 14-14-00289).

DOI 10.18699/BGRSSB-2018-050

Mining large database of genome-wide associations to identify biomarkers and intervention targets

T.I. Shashkova[1, 2, 3]*, A. Torgasheva[4], D.D. Gorev[1, 2, 3], E. Pakhomov[2], L. Klaric[5, 6], A.V. Severinov[1, 3], S. Sharapov[2, 4], Y.A. Tsepilov[2, 4], D.G. Alexeev[1, 2], J.F. Wilson[5], P. Joshi[5], Y.S. Aulchenko[2, 4]

[1] *LLC 'Knomics', Moscow, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Moscow Institute of Physics and Technology, Moscow, Russia*
[4] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[5] *University of Edinburgh, Edinburgh, UK*
[6] *Genos Glycoscience Laboratory, Zagreb, Croatia*
** e-mail: shashkova@phystech.edu*

Key words: genome-wide association studies, coronary artery disease, biomarker, intervention target

Motivation and Aim: Results from tens of thousands of GWAS that have been performed over the last decade are publicly available. The results are typically presented in the form of genome-wide summary statistics (for each SNP the allelic frequencies, estimates of the coefficients of regression and their standard errors are usually reported). This information can be used for multiple purposes – from research in fundamental biology and genetics, to biomarker and target discovery for therapeutic intervention. The aim of this work is to demonstrate that mining large databases of GWAS results allows identification of biomarkers and intervention targets. We focussed on coronary artery disease (CAD), one of the most economically and socially significant and one of the most studied complex common human diseases.

Methods and Algorithms: We developed a system, 'GWAS-MAP', that allows for a platform for storage, quality control and analysis of GWAS summary statistics. Our platform embeds LDsr and MRbase libraries, facilitating genetic correlations and mendelian randomization (MR) analyses, respectively. The analysis of pleiotropy is possible via our own implementation of summary-level mendelian randomization (SMR)/heterogeneity in dependent instruments (HEIDI) testing. We populated our database with about 220 GWASes. These included GWASes of lipid levels, GWASes for 128 metabolites, and 82 proteins from OLINK panel. We have also used eQTL data from a range of sources.

Results: Observed genetic correlations were consistent with previous studies. We selected 51 loci that were associated with CAD at genome-wide significant level in published GWASes. To define loci profiles in "omics" space we considered SMR results for CAD in metabolomic and proteomic space, which allowed us to cluster loci in several biologically meaningful groups. To understand the biological bases of the locus action, and to potentially provide a drug target, within each locus, we have prioritised the genes using the SMR/HEIDI test between CAD and gene expression. Our results confirmed existing knowledge of CAD mechanisms and suggested several CAD biomarkers and intervention targets.

Conclusion: The analysis of CAD using GWAS-MAP allowed for identification of biomarkers and potential intervention targets. Some of these biomarkers and targets are already well known and are used in clinical practice, validating the approach, whilst some are new – showing our approach to biomarker and target discovery is effective.

DOI 10.18699/BGRSSB-2018-051

Identification of an ABC-type multidrug efflux pump MacAB genes in the genome of *Serratia marcescens* SM6

T.V. Shirshikova[1]*, D.S. Pudova[1], D.A. Kabanov[1], A.O. Tikhonova[1], M.R. Sharipova[1], L.M. Bogomolnaya[1, 2]

[1] *Kazan (Volga region) Federal University, Kazan, Russia*

[2] *Texas A&M University Health Science Center, Microbial Pathogenesis and Immunology, Bryan, USA*

** e-mail: tatyana-shirshikova@yandex.ru*

Key words: *Serratia marcescens*, efflux pump, antimicrobial resistance

Motivation and Aim: Bacterial multidrug resistance is a global problem worldwide. *Serratia marcescens* is an opportunistic pathogen with increasing clinical importance. Infections caused by *S. marcescens* are often difficult to treat due to multiple drug resistance. One of the mechanisms of drug resistance is based on the activation of efflux pumps. A number of efflux pumps that belong to several classes were identified in reference genome *S. marcescens* Db11 [1], including a macrolide-specific ABC-type efflux pump MacAB. The aim of this study was to identify *macAB*-like genes in the genome of laboratory strain *S. marcescens* SM6.

Methods and Algorithms: Analysis of *macAB*-like genes was done in RAST (http://rast.nmpdr.org/rast.cgi).

Results: Analysis of *S. marcescens* SM6 genome showed that this strain harbors three independent gene clusters encoding for MacAB-like efflux pumps. Interestingly, *macAB*-like genes in genome locus SM6_1728-1729 share approximately 70 % of homology to similar genes in *E. coli* and *Salmonella enterica* serovar Choleraesuis str. SC-B67 and have a surrounding similar to that in those bacteria. Two additional *macAB*-like loci *S. arcescens* SM6 genome, SM6_875-876 and SM6_1583-1584 are absent from *E. coli* and salmonellae genomes but could be found in other genomes of *Serratia* sp. Function of these additional gene clusters in *S. marcescens* drug resistance is currently unknown.

Conclusion: Therefore, in contrast to a single macrolide-specific MacAB efflux pump present in *E. coli* genome, distantly related bacterium *S. marcescens* possesses three gene clusters encoding for homologues of this efflux pump. Detailed characterization of the impact of these gene clusters on resistance to antimicrobials is needed.

Acknowledgements: This work was supported by the Russian Foundation for Basic Research project 18-34-00458, the Russian Science Foundation project 16-14-10200 and performed in accordance with the Russian Government Program of Competitive Growth of Kazan Federal University.

References

1. Mardanova et al. (2014) Efflux systems in *Serratia marcescens*. Microbiology. 82(6):668-679.

DOI 10.18699/BGRSSB-2018-052

Dynamically regulated miRNA-mRNA network in lymph node metastasis of prostate cancer patients

M. Shkurnikov*, E. Knyazev, D. Mikhailenko, A. Zotikov, B. Alekseev
National Medical Research Radiology Center, Obninsk, Russia
** e-mail: mshkurnikov@imscs.msu.ru*

Key words: prostate cancer, metastasis, microarrays

Motivation and Aim: The lymph node metastasis in prostate cancer is associated with poor prognosis of overall survival of patients. Recent evidences suggest that metastatic activity of tumors could be regulated by extracellular circulating microRNAs which are released by various cells of the body. In this study we have analyzed the microRNA profiles in blood plasma of prostate cancer patients as well as the expression of microRNA-targeted genes in prostate cancer tumors with and without metastases to lymph nodes.
Methods and Algorithms: Blood plasmas from overall 101 prostate cancer patients (89 – pN_0M_0 stage, and 12 – pN_1M_0 stage) have been collected in a period between 2015–2016, and profiled for extracellular microRNAs using GeneChip miRNA 4.0 Arrays (Affymetrix, USA). The microRNAs that showed at least 2-fold (p-value < 0.05) change in expression between the analyzed groups of patients were selected for the subsequent analysis. Next, the list of mRNA genes targeted by at least two of the pre-selected circulating microRNAs simultaneously, was compiled. The expression data of these genes in tumors of prostate cancer patients with (N = 75) and without (N = 321) lymph node metastases were obtained from the TCGA database. All experiments have been conducted in accordance with the principles of the Declaration of Helsinki of World Medical Association.
Results: The analysis of microRNA profiles in blood plasmas of prostate cancer patients with lymph node metastases has revealed a significant decrease in the levels of nine circulating microRNAs including hsa-miR-92a-3p, hsa-miR-16-5p, hsa-miR-451a, hsa-miR-93-5p, hsa-let-7c-5p, hsa-miR-320a, hsa-miR-17-5p, hsa-miR-106a-5p, and hsa-miR-25-3p. We hypothesize that these microRNAs could be secreted by cells from tumor microenvironment, or other distant tissues (including the immune system components) and suppress tumor metastasis. Furthermore, elevated expression of six mRNA-coding genes (including *BIRC5*, *CDC20*, *MELK*, *RRM2*, *UBE2C*, *TUBB3)*, which were among the potential targets for the pre-selected nine miRNAs, was observed in tumor tissues from patients with lymph node metastases. Finally, each of these mRNA genes was an independent predictor of cancer metastasis with AUC > 0.66 and adjusted p-value was less than 0.05.
Conclusion: Our data demonstrate the role of nine circulating microRNAs in suppressing the lymph nodes metastases during prostate cancer, presumably by inhibiting the expression of six target genes in tumor tissues. These findings could be applied for the development of new approaches for stage differentiation and treatment of prostate cancer.
Acknowledgements: Supported by the MES of Russia (14.616.21.0092, RFMEFI61618X0092).

DOI 10.18699/BGRSSB-2018-053

Metastatic prostate cancer cells in lymph nodes perturb nodal laminin expression

M. Shkurnikov*, D. Maltseva
National Medical Research Radiology Center, Obninsk, Russia
** e-mail: mshkurnikov@imscs.msu.ru*

Key words: prostate cancer, metastasis, microarrays, laminins

Motivation and Aim: Metastasis of the most epithelial tumors occurs primarily through the lymphatic system. It is supposed that the presence of tumor cells in the lymph nodes (LNs) is not only a marker, but also an important mediator of metastatic disease. Normal stromal infrastructure of LN is a crucial for orchestrating immune cell homeostasis and adaptive immunity. Extracellular matrix (ECM) proteins play the most important role in LN parenchyma organization. Here, we compared the expression profile of genes encoding ECM components in LNs with and without histologically confirmed metastatic cells from a patient with prostate cancer.
Methods and Algorithms: Six LNs with and without histologically confirmed metastatic cells were obtained surgically from a patient with prostate cancer. A comparative transcriptomic analysis was performed using the microarray technology. The differential expression of the most relevant ECM protein genes was confirmed by qPCR.
Results: Decreased expression levels of genes encoding laminin 411 (LM-411) chains (LAMA4 – 3.4 times, LAMB1 – 2.3 times, LAMC1 – 2.2 times) was revealed in LNs with metastatic cells. Lymphocytes penetrate the blood capillary wall only at that places where the basement membrane (BM) contains LM-411. Thus, the reduction of LM-411 expression may disrupt the transport of naïve lymphocytes in LNs. LM-411 is also one of the two laminins comprising the BM of the conduit system, a special system that provides closer contact of LN resident dendritic cells (DCs) with lymph and antigens therein. The reduction of LM-411 expression may prevent contact of DCs with antigens, lymphocyte migration through LN space to find the antigen-bearing DCs and further activation and maturation of naïve T cells. A reduced mRNA level of other important BM component collagen type IV (1.7 and 1.6-fold lower expression of α1 and α2 subunits, respectively) was also shown in LNs with metastatic cells. Such decrease of the ECM protein expressions is apparently compensated by upregulation of the others, e. g. α2 laminin chain (1.6 times), fibronectin (2.5 times), collagen type I α1 and α2 subunits (1.4 times), collagen type VI α3 subunit (1.6 times), collagen type XIV α1 subunit (1.7 times). An increased collagen expression was earlier detected in LNs with metastatic cells in breast cancer.
Conclusion: The presence of metastatic cells in LNs may perturb their normal infrastructure and function. This might be one of the mechanisms used by tumor to suppress immune system.
Acknowledgements: Supported by the MES of Russia (14.616.21.0092, RFMEFI61618X0092).

DOI 10.18699/BGRSSB-2018-054

Artificial intelligence in the problems of analysis and interpretation of omics human data

A. Shlikht*, N. Kramorenko
Far Eastern Federal University, Vladivostok, Russia
** e-mail: schliht@mail.ru*

Key words: genomics, proteomics, common digital space, bioinformatics, artificial intelligence

Motivation and Aim: For efficient storage, analysis and interpretation of increasing omics data (BigData), it is necessary to convert it to SmallData using special types, formats, data structures [1]. The use of artificial intelligence technologies, databases, knowledge bases, expert systems are crucial here. The purpose of the work was to create an intelligent bioinformation system with effective structures and formats of presentation of omics data and knowledge of human.
Methods and Algorithms: The basis of the developed intelligent bioinformation system is the strategy of a common digital space (CDS). CDS is formed on the basis of highly structured, unified, complete and consistent registers, classifiers and coded dictionaries using standard structures and data types in the format of databases and knowledge bases. A characteristic feature of CDS is the presence of an autonomous information core, which focuses encoded highly structured data with a minimum of textual information, which allowed to move from BigData to Smalldata. The information core contains code parts of registers, classifiers and coded dictionaries (on the principle of machine – code, man – word). The creation of an autonomous information core has solved the problems of remote interactive access to world portals. This presentation of data in the core provides end-to-end information on all omics structures, many times increases the speed and validity of the analysis and interpretation of data. For filling of a CDS created by intelligent agents, which carry out conversion multi-format semi-structured data world portals, as well as the results of sequencing and mass spectrometry in a format the core CDS. This approach allowed moving from script procedural programming languages of 3rd generation (R, Phython) to the declarative languages of databases and knowledge bases of 4–5 generation, which frees the user from writing complex procedures for data extraction, analysis and interpretation. This presentation of data and knowledge makes it possible to build complex end-to-end logical chains in the tasks of genom-centered diagnosis and precision medicine.
Results: We developed an intelligent system of analysis and interpretation omics data, comprising: a stand-alone repository of highly structured data; a specific browser with a search system for navigating the space omics data and knowledge; subsystem knowledge extraction; subsystem access to world portals for updating data.
Conclusion: The created bioinformation system is an effective tool for biologists, physicians, geneticists, informatics, students in their scientific, practical and educational activities.

References

1. Shlikht A.G., Kramorenko N.V. (2016) Bioinformatic expert system of analysis and interpretation of omics sequence of the human genome. MM-HPC-BBB-2016:94.

DOI 10.18699/BGRSSB-2018-055

Determining the pathogenicity of genetic variants affecting splicing in Mendelian disorders

M. Skoblov[1, 2]*, A. Marakhonov[1, 2], Y. Vyakhireva[1], P. Sparber[1], V. Freire[1], A. Filatova[1]
[1] *Research Center for Medical Genetics, Moscow, Russia*
[2] *Moscow Institute of Physics and Technology* (*State University*)*, Dolgoprudny, Russia*
** e-mail: mskoblov@gmail.com*

Key words: Mendelian disorders, splicing alterations, minigene assay

Motivation and Aim: Despite the success of whole-exome sequencing (WES) in the diagnosis of Mendelian disorders, ~50–75 % of the patients still do not receive a genetic diagnosis [1]. One of the reasons for this is the limitation of bioinformatic approaches to detect pathogenic variants of the nucleotide sequence, especially in noncoding regions. It is now known that mutations affecting splicing can cause the Mendelian disorder. But because of the complexity of splicing regulation, it is not always possible to predict accurately the effect of nucleotide variants on splicing events and RNA structure. In this work, we focused on functional analysis of genomic variants affecting splicing in a variety of Mendelian disorders.
Methods and Algorithms: To determine the effect of mutations we used two experimental approaches: (1) RT-PCR from available patient's samples and (2) *in vitro* minigene assay. For different cases, we performed one or both methods. Human Splicing Finder and IntSplice on-line tools were used to predict the effect of different nucleotide variants on pre-mRNA splicing.
Results: We analyzed >27 previously uncharacterized genetic variants in >12 genes, associated with different Mendelian disorders. These variants are located in both exons and introns and mostly were classified as variant of unknown significance (VUS). We determined the effect of these variants on mRNA structure; it allowed us to classify most of them as pathogenic and to make assumption of the mechanisms involved in the molecular pathogenesis of diseases (e. g. RNA degradation by NMD, disruption of functional domain of protein). Additionally, we compared our experimental data with prediction tools for splicing events and revealed that it is not always possible to predict accurately the effect of mutation on splicing.
Conclusion: Although it is now known that mutations affecting splicing can cause the Mendelian diseases, however their contribution may be underrepresented due to limitation of diagnostic procedures. To prove the pathogenicity of these mutations, additional functional analysis is often required.

References

1. Kremer L.S. et al. (2017) Genetic diagnosis of Mendelian disorders via RNA sequencing. Nat. Commun. 8:15824.

DOI 10.18699/BGRSSB-2018-056

MYH3 is a novel gene associated with carotid paragangliomas

A. Snezhkina[1]*, E. Lukyanova[1], D. Kalinin[2], E. Pudova[1], M. Fedorova[1], S. Kharitonov[1], Z. Guvatova[1], I. Abramov[1], G. Krasnov[1], N. Melnikova[1], A. Dmitriev[1], A. Kudryavtseva[1, 3]

[1] *Engelhardt Institute of Molecular Biology RAS, Moscow, Russia*

[2] *Vishnevsky Institute of Surgery, Ministry of Health of the Russian Federation, Moscow, Russia*

[3] *National Medical Research Radiological Center, Ministry of Health of the Russian Federation, Moscow, Russia*

** e-mail: leftger@rambler.ru*

Key words: carotid paragangliomas, exome, mutations, MYH3, high-throughput sequencing

Motivation and Aim: Carotid paragangliomas are relatively rare tumors arising from the paraganglia of carotid body at the bifurcation of carotid arteries. Approximately 40 % of paragangliomas are hereditary, and currently over 30 different genes have been associated with the disease. However, molecular genetic mechanisms underlying carcinogenesis of carotid paragangliomas are still not fully understood.

Methods and Algorithms: We have collected 52 carotid paragangliomas derived from Vishnevsky Institute of Surgery, Ministry of Health of the Russian Federation. The isolation of DNA and RNA from the samples was performed. Specialized Illumina kits (USA) were used for exome and transcriptome library preparation. Exome sequencing was performed in 2x75 bp paired-end model using NextSeq 500 System (Illumina) at EIMB RAS "Genome" center (http://www.eimb.ru/rus/ckp/ccu_genome_c.php). Sequencing of transcriptome libraries was carried out using HiSeq 4000 (Illumina) with 2x100 bp paired-end reads. Bioinformatic analysis of the data was performed using XSEQ software that allows computing the influence of somatic mutations on gene expression profile [1].

Results: Using XSEQ software, we analyzed the exome and transcriptome data from carotid paragangliomas. We found that mutations in *MYH3* gene impact its expression and, therefore, can have potential phenotypic effect. Different mutations were observed in *MYH3* gene including three probably pathogenic ones (NM_002470.3: p.Ile264Thr/c.791T>C (rs763347751, chr17: 10550688), NM_002470.3: p.Ala1752Thr/c.5254G>A (rs34393601, chr17: 10534960), and NM_002470.3: p.Ala1604Thr/c.4810G>A (rs201488879, chr17: 10535939), according to Sift and PolyPhen 2 prediction tools. These mutations can change the structure of the protein and, therefore, affect its function.

Conclusion: A novel gene, *MYH3*, which can be involved in carcinogenesis of carotid paragangliomas, was identified.

Acknowledgements: This work was funded by the Russian Science Foundation, grant 17-75-20105.

References

1. Ding J. et al. (2015) Systematic analysis of somatic mutations impacting gene expression in 12 tumour types. Nat. Commun. 6:8554.

DOI 10.18699/BGRSSB-2018-057

Hsp as a long-term buffer of the genome-wide mutation burden

A. Sokol[1]* , K. Ushakova[1], A.A. Mikhaylova[1], A.G. Mikhaylova[1], D. Knorre[3], I. Mazunin[1], A. Reymond[5], K. Gunbin[1, 4], K. Popadin[1, 5]**

[1] *The School of Life Sciences, Immanuel Kant Baltic Federal University, Kaliningrad, Russia*

[2] *Institute of Biodiversity Animal Health and Comparative Medicine, University of Glasgow, Glasgow, UK*

[3] *The A.N. Belozersky Institute of Physico-Chemical Biology, MSU, Moscow, Russia*

[4] *The Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

[5] *Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland*

** anastasia3sokol@yandex.ru, ** konstantinpopadin@gmail.com*

Key words: hsp90, gene expression level, effective population size

Motivation and Aim: At short evolutionary times, the expression level of hsp90 works as a buffer, i. e. strongly influences the manifestation of slightly deleterious nonsynonymous variants (SDNVs) in populations of model organisms (*Drosophila*, *Arabidopsis*, etc.) and human. The increased expression level of hsp90 compensates the burden of SDNVs by correct folding of numerous proteins. Oppositely, the decreased expression level of hsp90 can not maintain correct folding of these proteins, uncovering previously hidden burden of SDNVs. We hypothesized that the expression level of hsp90 may play an important role also at long evolutionary periods compensating genome-wide burden of slightly deleterious variants. Taking into account that the number of fixed slightly deleterious variants is higher in species with low effective population size, we expect that the expression level of hsp90 in these species is increased to partially compensate the high burden of SDNVs.

Methods and Algorithms: In this project using comparative-species approach we test our hypothesis on three different levels: (i) comparison of the expression level of hsp90 between different species and tissues, (ii) investigation of the patterns of molecular evolution of hsp90 between species and (iii) analysis of the number of copies of the hsp90 genes in different species.

Results: We derived hsp90 data-base, containing the expression levels, patterns of molecular evolution and copy numbers of hsp90 genes in all eukaryotes.

Conclusion: The database will allow us to address our questions and test the hypothesis.

Acknowledgements: The study was supported by the 5 Top 100 Russian Academic Excellence Project at the Immanuel Kant Baltic Federal University.

DOI 10.18699/BGRSSB-2018-058

Analysis of biosynthetic gene clusters of *Rhodococcus* sp. S10

A. Sorokina[1]*, L. Shafigullina[1], A. Elistratova[1], M. Belenikin[2], M. Sharipova[1], L. Bogomolnaya[1, 3], I. Khilyas[1]

[1] *Institute of Fundamental Medicine and Biology KFU, Kazan, Russia*

[2] *Department of Molecular and Biological Physics, Moscow Institute of Physics and Technology* (*State University*), *Dolgoprudny, Moscow Region, Russia*

[3] *Texas A&M University Health Science Center, Bryan, Texas, USA*

** e-mail: alvita.94@yandex.ru*

Key words: siderophores, secondary metabolites, NRPS, *Rhodococcus*

Motivation and Aim: Bacteria produce a broad spectrum of biologically active natural compounds, including peptides, synthetized by using large multifunctional nonribosomal peptide synthetases (NRPSs) or polyketide synthases (PKSs). Among microbial peptides the compounds with a metal chelating activity draw a particular interest. Siderophores are secreted low molecular weight compounds, which can chelate Fe (III) with an extremely high affinity [1]. Bacteria from the genus *Rhodococcus* have been shown to produce a wide range of secondary metabolites [2]. Nowadays, the only known siderophores produced by members of the genus *Rhodococcus* include heterobactin A and rhodobactin [2]. The aim of this study was to identify additional biosynthetic gene clusters in the genome of *Rhodococcus* sp. S10.

Methods and Algorithms: *Rhodococcus* sp. S10 genome was analyzed for secondary metabolite and siderophore biosynthetic gene clusters using antiSmash software; RAST software was used for gene annotation [3].

Results: One hundred two biosynthetic gene clusters were predicted to be present in the genome of *Rhodococcus* sp. S10 by the antiSMASH software. Among those, two putative PKSs and ten putative NRPSs gene clusters were identified. Two NRPS clusters have a high sequence homology to known siderophores. Thus, cluster 99 has 100 % similarity to heterobactin gene cluster of *R. qingshengii* BKS 20-40, *Rhodococcus* sp. ADH and *R. erythropolis* SK121; while cluster 56 has 57 % similarity to albachelin gene cluster of *R. qingshengii* BKS 20-40 and *R. erythropolis* CCM2595. Five NRPS gene clusters of *Rhodococcus* sp. S10 did not show any homology to any known bacterial NRPS clusters.

Conclusion: Analysis of *Rhodococcus* sp. S10 genome allowed us to identify a number of putative biosynthetic gene clusters. High homology of several *Rhodococcus* sp. S10 gene clusters to genes involved in siderophores synthesis encourages the search for new metabolites with metal chelating activity, which might promote *Rhodococcus* sp. S10 growth and adaptation to the extreme environments.

Acknowledgements: This work was performed within the Program of Competitive Growth of Kazan Federal University and supported by RFBR (grant no. 16-34-60200) and the scholarship of the President of the Russian Federation for young scientists and graduate students.

References

1. Crosa J.H., Walsh C.T. (2002) Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. Microbiol. Mol. Biol. Rev. 66(2):223-249.
2. Carrano C.J. et al. (2001) Heterobactins: A new class of siderophores from *Rhodococcus erythropolis* IGTS8 containing both hydroxamate and catecholate donor groups. BioMetals. 14:119-125.
3. Weber T. et al. (2015) antiSMASH 3.0 – a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res. DOI 10.1093/nar/gkv437.

DOI 10.18699/BGRSSB-2018-059

Evaluation of MinION nanopore platform for HIV whole coding regions sequencing

A.S. Speranskaya*, A.E. Lopatukhin, K. Khafizov, A.A. Ayginin, E.V. Korneenko, D.E. Kireev, G.A. Shipulin
Central Research Institute of Epidemiology, Moscow, Russia
** e-mail: hanna.s.939@gmail.com*

Key words: sequencing technology, genomics, biomedicine, HIV, genome assembly

Motivation and Aim: Human immunodeficiency virus has a significant impact on economies worldwide due to high morbidity rates. The search for the new effective approaches for both rapid and non-expensive identification of drug resistance mutations is needed in order to reduce the HIV-induced mortality. Complete sequencing of the HIV genomes now is possible by various high-throughput second generation sequencing (HTS) methods. Both Illumina and Ion Torrent semiconductor platforms are widely used for the analysis of sequence diversity, recombinants determination and estimation of genetic distance between the HIV-1 quasi-species. Second generation HTS platforms has efficiency, accuracy and ability to detect the minor populations at the level as low as 1 %. However, the usage of these approaches is labour-consuming and therefore the analysis of clinical samples might be significantly delayed. Recently, the long read HTS sequencing of a third generation has become available from Oxford Nanopore (nanopore sequencing technology). The MinION is a low-cost portable device, which requires short time for the sample preparation and potentially can generate up to 10–20 Gb of DNA sequencing data per run. Oxford Nanopore sequencing technology is often described as highly error-prone (~8–40 % error rate). Our study aimed at estimating of the accuracy of the whole HIV coding region identification of HIV viral standards using the MinION sequencer.
Materials and Methods: For experiments we used the pNL4-3 (NIH AIDS Reagent, aidsreagent.org, USA), which are vectors containing the genome-wide replicatively and infectiously active HIV DNA. The four amplified fragments of the HIV genome (of 2,500–3,500 bp length each) were used for library preparation. MinION non-barcoded libraries were prepared using the Oxford Nanopore Genomic DNA Sequencing protocol provided by the manufacturer. The Rapid 1D Sequencing Kit was used to determine if the using of 1D reads would be adequate for HIV single nucleotide analysis. The canu software was used to assemble the reads into contigs. BWA program was then employed to map the contigs to the reference genome.
Results: The MinION produces data in real time so the analysis can be done during the experiment step by step. We estimated what number of reads should be produced for the best results. Finally, we made 700,000 more reads to achieve extra-large covering of pNL4-3 reference sequence (which is 8,883 bp long). The complete genome-wide HIV consensus sequences were assembled using MinION sequencing data. The platforms accuracies were estimated by comparing of the consensuses to the pNL4-3 reference sequence. We found that 98.95 % of the nucleotides analyzed positions were correctly identified when all obtained reads were used for assembling. Another 1.05 % were single nucleotide insertions (0.21 %), deletions (0.54 %) or ambiguous positions (0.29 %).
Conclusion: An efficient approach to obtain HIV genome sequencing data using the portable sequencer MinION was demonstrated.

DOI 10.18699/BGRSSB-2018-060

BRCA mutation and castration-resistant prostate cancer, association with the AKT/m-TOR signaling cascade

L.V. Spirina[1, 2]*, A.K. Gorbunov[1], E.A. Usynin[1], E.M. Slonimskaya[1, 2], I.V. Kondakova[1]
[1] *Cancer Research Institute, TNIMC, Tomsk, Russia*
[2] *Siberian State Medical University, Tomsk, Russia*
** e-mail: spirinalvl@mail.ru*

Key words: castration-resistant prostate cancer, BRCA, AKT/m-TOR

Motivation and Aim: Mutations of BRCA genes are an independent poor prognostic factor in the development of prostate cancer [1]. It is believed that its effect on the prognosis is higher than the degree of differentiation, stage, level of the prostatic specific antigen. It is known that the mutations BRCA1/2 are most often associated with the AKT/m-TOR signaling cascade hyperactivation [2–4]. The purpose was to study the AKT/m-TOR pathway components in castration-resistant prostate cancer patient, depending on the presence of the BRCA mutations.
Methods and Algorithms: 40 patients with prostate cancer, 15 patients with castration-resistant prostate cancer and 20 patients with benign hyperplasia are enrolled in the investigation. The expression of AKT, c-Raf, GSK-3, PDK1, and m-TOR, 70-S64, E-BP1 was determined by real-time PCR. The BRCA 1/2 mutation was determined in allele-specific PCR in real time.
Results: Activation of the AKT/m-TOR signaling cascade was detected in prostate cancers. The high levels of AKT and m-TOR expression were revealed. The increase in the level of phosphatase PTEN was found in benign hyperplasia and cancer tissues [5]. The level of mRNA 4E-BP1 was decreased in castration-resistant prostate cancer patients. At the next stage of the study, the incidence of inherited BRCA 1/2 mutations were studied in patients with castration-resistant cancer. The *BRCA1-5382insC* mutation was detected in 3 patients (20 %), *BRCA1-4153delA* – in 5 patients (33 %), *BRCA1-185delAG* – in 2 patients (13 %), *BRCA1-T300G* – in 2 patients (13 %) and *BRCA2-6174del* – in 4 patients (27 %). BRCA1-deficiency activates the AKT oncogenic pathway, one of the most common alterations associated with human malignancy. Mutation of *BRCA1* gene increases the phosphorylation and the kinase activity of AKT. The decreased AKT expression in cancers was found in patients with *BRCA1-5382insC* mutation. Mutation of *BRCA1-4153delA* increased expression of 70S, m-TOR, in the presence of *BRCA1-T300G* – increased PTEN. The inherited *BRCA2-6174del* mutation was correlated with the increased expression of AKT.
Conclusion: Therefore, the development of PCa is accompanied by activation of this signaling cascade, even more pronounced in the presence of mutations *BRCA2-6174del*, *BRCA1-4153delA*, *BRCA1-T300G*. It should be noted that the frequency of occurrence of these mutations varies from 13 to 33 %.

References

1. Katsnelson L.B. et al. (2011) Contribution of mechanical factors to arrhythmogenesis in calcium overloaded cardiomyocytes: Model predictions and experiments. Prog. Biophys. Mol. Biol. 107(1):81-89.
2. Lashin S.A., Matushkin Yu.G. (2012) Haploid evolutionary constructor: new features and further challenges. In Silico. Biol. 11(3):125-135.
3. Modena A. et al. (2016) Investigating BRCA mutations: A breakthrough in precision medicine of castration-resistant prostate cancer. Target Oncol. 11(5):569-577.
4. Xiang T. et al. (2016) Targeting the Akt/mTOR pathway in Brca1-deficient cancers. Oncogene. 30(21): 2443-50. DOI 10.1038/onc.2010.603. Epub 2011 Jan 17.
5. Spirina L.V. et al. (2018) Transcription factor brn-3α and prostate cancer, relation with hormone receptors and AKT/m-TOR signaling pathway activation. Sib. J. Oncology. 17(1):26-31. (in Russ.) DOI 10.21294/1814-4861-2018-17-1-26-31.

DOI 10.18699/BGRSSB-2018-061

The interplay of Piwi and heterochromatin proteins in transposable element silencing in the germline of *Drosophila melanogaster*

A. Stolyarenko[1]*, T. Leinsoo[1], Y. Abramov[1], P. Mazin[2], M. Klenov[1]
[1] *Institute of Molecular Genetics RAS, Moscow, Russia*
[2] *Center for Data-Intensive Biomedicine and Biotechnology, Skolkovo Institute of Science and Technology, Moscow, Russia*
** e-mail: stol@img.ras.ru*

Key words: Piwi, Argonaute, HP1a, heterochromatin, transcriptional silencing, RNAseq

Motivation and Aim: Transposable element (TE) silencing is critical for germline genome stability and fertility of animals, including *Drosophila melanogaster*. The piRNA-binding protein Piwi of the Argonaute family was shown to induce transcriptional repression of TEs in *Drosophila* ovaries [1]. However, the interplay of Piwi and the main heterochromatin protein HP1a and other heterochromatin proteins in this process is still not fully understood, and Piwi is usually considered to team with HP1a to carry out TE repression [2].
Methods and Algorithms: We took advantage of a genetic approach of combining in one genotype germline knockdowns (KD) of Piwi and HP1a and comparing them with single Piwi and HP1a KDs in the same genetic background via RNA-seq. We divided TEs responding to the double KD according to their response to Piwi and HP1a KDs into four groups.
Results: A group of TEs are equally controlled by Piwi or HP1a which points to the participation of these two proteins in a single pathway to silence the elements of this group. Another group of TEs are controlled by Piwi alone, but not HP1a. Another group of elements are derepressed only upon double KD. This implies independent double control by systems based on Piwi and HP1a in combination. Interestingly, according to our previously published ChIP-seq data, the chromatin of these TEs is especially highly enriched in the repressive H3K9me3 mark, which is indicative of their localization in heterochromatin. Finally, there is a group of Piwi-independent elements controlled by HP1a alone. Interestingly, we found that these elements coincided with the ones known to be derepressed upon mutation of endosiRNA-binding protein Ago2 [3]. We propose that the repression of these elements by Ago2 is assisted by HP1a in the germline.
Conclusion: Our results indicate that Piwi-induced silencing and Piwi-independent repression by heterochromatic proteins may represent two distinct systems that silence transcription of TEs in the ovarian germline. These two systems can act autonomously or perhaps in association with each other to repress an element.
Acknowledgements: Supported by the RFBR (16-04-01524).

References

1. Klenov M. et al. (2007) Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the *Drosophila melanogaster* germline. Nucleic Acids Res. 35(16):5430-5438.
2. Sienski G. et al. (2015) Silencio/CG9754 connects the Piwi–piRNA complex to the cellular heterochromatin machinery. Genes Dev. 29(21):2258-2271.
3. Czech B. et al. (2008) An endogenous small interfering RNA pathway in Drosophila. Nature. 453(7196): 798-802.

DOI 10.18699/BGRSSB-2018-062

Comparative transcriptomics of the moss *Physcomitrella patens* under biotic stress

N. Sukhikh[1], E. Egorova[1], A. Predeus[2], S. Vinogradova[1]*
[1] *Institute of Bioengineering, Research Centre of Biotechnology RAS, Moscow, Russia*
[2] *Bioinformatics Institute, Saint Petersburg, Russia*
** e-mail: svetlana.vinogradova@biengi.ac.ru*

Key words: *Physcomitrella patens*, moss, gene expression

Motivation and Aim: The moss *Physcomitrella patens* is a non-vascular, multicellular land plant. It used for studying cell biology and evolution processes. In contrast to the polyploid flowering plants, mosses have prevalent the haploid gametophytic stage in their life cycle. *P. patens* is an interesting model organism for exploring the mechanisms of adaptation in plants due to its evolutionary stage between green algae and flowering plants. Here we present research project related to complex analysis of high-throughput sequencing data from our study and publicly available transcriptomics data. This project provides information on the regulatory mechanisms that developed during land plant evolution.
Methods and Algorithms: Moss gametophores were cultivated on Knop's medium and inoculated with suspensions of *Pseudomonas syringae*, *Pseudomonas viridiflava* and *Xanthomonas arboricola* (OD 0.4). A five days post inoculation gametophores were grinded, mRNA was extracted, and cDNA libraries were constructed. Sequencing of cDNA libraries was performed on SOLiD platform. Sequence reads were mapped to the *P. patens* genome and transcriptome V3.3 using Bowtie and Tophat. For quantifying gene and isoform abundances RSEM was used. Additionally we have compiled a gene set database for pathway analysis in *P. patens*. For this purpose we have collected previously published transcriptome datasets of *P. patens* from Gene Expression Omnibus (GEO, www.ncbi.nlm.nih.gov/geo/) and ArrayExpress (www.ebi.ac.uk/arrayexpress/). Gene IDs from various sources were converted to Phytozome gene symbols. STAR, kallisto and RSEM were used for data analysis. Data visualization was performed with IGV Tools and Phantasus.
Results and Conclusion: About 84 M of 50 bp reads were obtained for each sample and nearly 68 % of reads were aligned to transcriptome. As a result, it was shown 33055 differentially expressed transcripts. The majority of genes up-regulated by *P. syringae* and *P. viridiflava* encode endocytosis, while down-regulated genes encode nitrogen metabolism, histidine metabolism and cyanoamino acid metabolism. As to *X. arboricola*, the majority of up-regulated genes encode ribosome biogenesis, RNA transport, photosynthesis, and flavonoid biosynthesis, whereas down-regulated genes encode alpha-Linolenic acid metabolism, glutathione metabolism, and nitrogen metabolism. 10 RNA-Seq datasets taken from GEO and ArrayExpress. KEGG database was used for analysis of signaling networks. Transcriptomic profiling of moss samples under different biotic and abiotic conditions has provided an opportunity for investigation of gene expression and molecular pathways.
Acknowledgements: This study was supported by RFBR, research project no. 18-34-00822 and was performed using the experimental climate control facility U-73547.

DOI 10.18699/BGRSSB-2018-230

Workflows for classification of NGS metagenomic data

D. Sukhomlinov[1, 2]*, Y. Algaer[2], A. Tiunov[2], E. Pushkova[2], O. Golosova[2]
[1] *Novosibirsk State University, Novosibirsk, Russia*
[2] *Unipro Center of Information Technologies, Novosibirsk, Russia*
** e-mail: dsukhomlinov@unipro.ru*

Key words: bioinformatics, NGS, data analysis, metagenomics, classification, workflow, graphical user interface

Motivation and Aim: Next Generation Sequencing (NGS) technologies, invented a few decades ago, have opened new opportunities for scientists. In particular, the technologies are used for identification of microorganisms in metagenomic clinical and environmental samples. Although, there is an extensive set of computational tools for classification of the sequencing data, the tools are commonly command-line and require additional configuration. Also, usage of multiple tools, that may be useful for improving quality of classification, requires a lot of routines from a scientist in case of stand-alone tools. We are motivated to provide convenient and user-friendly graphical user interface for some of these tools to make a work process more effective and simple. We believe that a well-organized process of the classification will improve viral and bacterial pathogen detection and discovery as well as disease control and prevention.
Methods and Algorithms: Unipro UGENE [1] is a desktop multiplatform software package that integrates dozens of widely used bioinformatics tools. The Workflow Designer component of the software allows one to use graphical interface to create and run workflows, composed of different tools. In addition, it is intended to store and investigate the results and re-run workflows on different datasets. The described infrastructure has served as the basis for a new framework for whole-genome NGS data classification, which includes the following popular tools: CLARK [2] (CLAssifier based on Reduced K-mers), supplied with NCBI RefSeq viral and bacterial database; Kraken [3], supplied with MiniKraken database; DIAMOND [4], a sequence aligner, similar to NCBI BLAST, supplied with UniRef50 and UniRef90 databases and WEVOTE [5] (WEighted VOting Taxonomic idEntification), used to ensemble classification data, produced by other tools.
Results: The NGS data classification framework was integrated into the open source Unipro UGENE software and contains sample workflows. The first serial workflow sequentially runs Kraken, CLARK and DIAMOND tools, filtering out NGS reads after each step, and reports classification information, produced by the tools. The second parallel workflow runs these tools and ensembles the output data using WEVOTE. The third workflow classifies NGS scaffolds, assembled de novo by SPAdes [6]. Advanced users can also use individual blocks of these workflows to create a new one, suitable for their purposes. The results are available on Linux and macOS platforms.
Conclusion: The new framework is intended for optimization of a scientist's work in the area of whole-genome NGS data classification, easy to set up for a biologist and require no additional configuration.
Availability: http://ugene.net/download.html
Acknowledgements: The project was supported by the VIROGENESIS consortium [7].

References:
1. Okonechnikov K., Golosova O., Fursov M. (2012) the UGENE team, Unipro. Bioinformatics. 28:1166-1167. DOI 10.1093/bioinformatics/bts091
2. Ounit R. et al. BMC Genomics (2015) 16:236. DOI 10.1186/s12864-015-1419-2
3. Wood D.E., Salzberg S.L. (2014) Kraken: ultrafast metagenomic sequence classification using exact alignments. Gen Biol. 15:R46.
4. Buchfink B., Xie C., Huson D. (2015) Nature Methods. 12:59-60.
5. Metwally A., Dai Y., Finn P.W., Perkins D.L. (2016) WEVOTE: Weighted Voting Taxonomic Identification Method of Microbial Sequences. PloS ONE.
6. Bankevich A. et al. (2012) SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. Journal Computational Biol.
7. VIROGENESIS project web site: http://www.virogenesis.eu/

DOI 10.18699/BGRSSB-2018-063

The genomic basis of human lifespan

P.R.H.J. Timmers[1], N. Mounier[2, 3], K. Läll[4, 5], K. Fischer[4], Z. Ning[6], X. Feng[7], A. Bretherick[8], D.W. Clark[1], X. Shen[1, 6], T. Esko[4], Z. Kutalik[2, 3], J.F. Wilson[1, 8], P.K. Joshi[1, 2]*

[1] *Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Teviot Place, Edinburgh, UK*

[2] *Institute of Social and Preventive Medicine, University Hospital of Lausanne, Lausanne, Switzerland*

[3] *Swiss Institute of Bioinformatics, Lausanne, Switzerland*

[4] *Estonian Genome Center, University of Tartu, Tartu, Estonia*

[5] *Institute of Mathematics and Statistics, University of Tartu, Tartu, Estonia*

[6] *Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden*

[7] *State Key Laboratory of Biocontrol, Guangdong Provincial Key Laboratory of Plant Resources, Key Laboratory of Biodiversity Dynamics and Conservation of Guangdong Higher Education Institutes, School of Life Sciences, Sun Yat-sen University, Guangzhou, China*

[8] *MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh, UK*

* *e-mail: peter.joshi@ed.ac.uk*

Key words: lifespan, longevity, genetics, genomics

Motivation and Aim: Whilst of great interest to us all, investigation into the genomic basis of longevity has been hampered by limited sample sizes. As a result, until very recently, only 4 genome-wide significant loci had been discovered and replicated, limiting the inferences that be made about its genetic basis [1].

Results: Here using an independent replication cohort, we examine 20 published [2] but unreplicated genome-wide significant loci for longevity, validating associations at or near CDKN2B-AS1, ATXN2/BRAP, FURIN/FES, FOXO3A, 5q33.3/EBF1, ZW10, PSORS1C3, 13q21.31, and provide evidence against previous findings near CLU, CHRNA4, PROX2, and d3-GHR. In a GWAS using all data combined, totalling over 1m lifespans, we next find 15 further loci at genome-wide significance. Of the life lengthening loci significant in our analyses, we find many protect against cardiovascular, glycaemic and neurological diseases (CGND), but not cancer. Using our GWAS, we then create polygenic scores for survival in independent sub-cohorts, and are able to partition populations, using DNA alone, into deciles of expectation of life, with a difference in excess of five years from top to bottom decile.

Conclusion: It seems that natural selection has been more effective in purging common variants affecting lifespan through cancer, but in our modern obesogenic and long lived environment has yet to as effectively purge variants affecting CGND. Materially accurate predictions of lifespan can now be made from DNA.

Acknowledgements: We thank the UK Biobank Resource, approved under application 8304; we acknowledge funding from the UK Medical Research Council Human Genetics Unit. We thank Tom Haller of the University of Tartu, for tailoring RegScan [3] so we could use it with compressed files. We would also like to thank the researchers, funders and participants of the LifeGen consortium.

References

1. Joshi P.K. et al. (2017) Genome-wide meta-analysis associates HLA-DQA1/DRB1 and LPA and lifestyle factors with human longevity. Nat. Commun. 8:910.
2. Pilling L.C. et al. (2017) Human longevity: 25 genetic loci associated in 389,166 UK biobank participants. Aging (Albany NY). 9:2504-2520.
3. Haller T., Kals M., Esko T., Magi R., Fischer K. (2015) RegScan: a GWAS tool for quick estimation of allele effects on continuous traits and their combinations. Brief Bioinform. 16:39-44.

DOI 10.18699/BGRSSB-2018-064

The use of a horizontally scalable infrastructure in the search for genetic similarity in biodiversity

A. Tskhai[1, 2]*, S. Murzintsev[3]
[1] *Institute for Water and Environmental Sciences SB RAS, Barnaul, Russia*
[2] *Altai State Technical University named after I.I. Polzunov, Barnaul, Russia*
[3] *Altai State University, Barnaul, Russia*
** e-mail: taa1956@mail.ru*

Key words: similarity of genomes, large data, nonrelational databases, search algorithms for repetitions

Motivation and Aim: The exploration of the structural-functional organization of biodiversity continues to be a main direction, developing at the intersection of biology and information technologies [1].
Methods and Algorithms: In this scientific research the problem of rapid detection of genetic similarity is considered in the analysis of databases (DB) of genomes of individuals from ecosystems of different levels. The distributed non-relational DB MongoDB and the Winnowing data processing algorithm are used as the basis for creating the information system. Using a non-relational database to identify genetic similarity, a variant of representing the prints of the structural variations of the genomes in the form of "key-value" was proposed, a program implementation of the developed model was carried out, and computational experiments were carried out, which confirmed the possibility of using the proposed method of genetic similarity search, for example, in a personified analysis of deviations in the gene level.
Results: The results obtained in the scientific research allow to find confirmation of the hypothesis of the applicability of distributed non-relational databases for comparison and searching for deviations in the development of living beings on the basis of a personified analysis of their genomes. The database of the Kyoto encyclopedia of genes and genes KEGG GENOME (Sequenced genomes of various living organisms) was selected as a source of elements of the genome database of organisms, including decoded representations of about five thousand living beings (http://www.genome.jp/kegg/catalog/org_list.html). The genome size of one living creature reaches 1 GB in compressed form. The total volume of genomes of living beings in this information resource, therefore, is approximately five TB. Due to the large amount of data that occurs during the processing of the original information, the transition from relational databases to non-relational databases has been carried out, both to a more flexible and scalable database.
Conclusion: The development of scientific research is considered from several sides. The first direction is the solution of the problem of determining the moment at which it is necessary to add a node to the cluster of computers with increasing the number of deviations considered and increasing the number of genomes in the database of organisms. The second is the practical filling and further formation of the database with as many real genomes as possible. The application of the results obtained in interdisciplinary studies would allow us to speak about the development of the proposed direction of research of genomic disorders. This direction is focused on obtaining an assessment of the probability of genetic abnormalities at the stage of recognition of the potentially unfavorable development of the situation. In the case of gaining access to the production databases of genomes of humans, animals and plants and conducting joint research with genetic specialists, all of the above looks real.

References

1. Noisy V.C., Shokin Yu.I., Kolchanov N.A., Fedotov A.M. (2006) Biodiversity and ecosystem dynamics: information technology and modeling. Novosibirsk: SB RAS Publ. House, 2006.

DOI 10.18699/BGRSSB-2018-065

Genomics in analysis of *Chlamydia psittaci* host-adaptation

O. Voronina*, M. Kunda, E. Aksenova, N. Sharapova, N. Ryzhova, N. Bondareva, N. Zigangirova
N.F. Gamaleya National Research Center for Epidemiology and Microbiology, Moscow, Russia
** e-mail: olv550@gmail.com*

Key words: *Chlamydia psittaci*, polymorphic membrane proteins, plasticity zone, type III secretion system

Motivation and Aim: *Chlamydia psittaci* is an economically relevant pathogen in livestock and pet animals, where it causes psittacosis/ornithosis, and also a human pathogen causing atypical pneumonia after zoonotic transmission. Conjunctivitis, lung diseases, enteritis, animals aborted fetus are associated with this pathogenic bacteria. Earlier we proved that three *Chlamydia* strains, isolated from synovial fluid of patients with reactive arthritis were *C. psittaci*. Genome sequencing of these strains (Accession numbers CP024451, CP024453, CP024455, ST (sequence type) 24) revealed that they are closely related. The main zones of variability in very conservative *Chlamydia* genomes are plasticity zone (PZ) and zones of polymorphic membrane proteins (pmp). These zones of the synovial strains had the highest similarity with 01DC11 strain isolated from pig conjunctiva (ST24). Extending these strains investigation we analyzed type III secretion system (T3SS) as main pathogenicity factor participating in host-microbe interactions.
Methods and Algorithms: RAST v.2.0 server was used for genome annotation; TMHMM Server v.2.0 – for prediction of protein localization; Virulence Factor DataBase and EffectiveDB (EffectiveT3) – for revealing the genes associated with T3SS and its effectors; COILS server – for prediction of helix coiled-coil domains.
Results: Analysis of the synovial strains genomes revealed 62 genes coding intact proteins and effectors of T3SS that is able to inject proteins directly into cytoplasm of the eukaryotic cells. Only few groups of these genes were organized in the operons, in contrast with Betaproteobacteria having one big T3SS gene cluster. Among ST24 strains of the animal origin we found only 3 genes which were differed from synovial T3SS genes by 1–2 mismatches. Human ST24 strain from bronchial lavage had 3–4 mismatches in the other 3 genes. In contrast, ST35 strain isolated from throat of humans with an influenza-like infection demonstrated differences in 79 % of T3SS genes. The most number of changes, 94 mismatches and 2 gaps, were in the gene of putative cytotoxin. The genes coding proteins of the export apparatus were predominant among the genes with 100 % similarity. There were only two genes of the effectors in this group. Moreover among 28 pmp genes of ST35 strain there was only one with 100 % similarity with pmp genes of the synovial strains.
Conclusion: *C. psittaci* ST24 strains are wide spread, polyhostal, adapted to the wide host range, including birds, livestock animals and human, and are able to infect the cell of the different tissue, causing the large list of the diseases. PZ, pmp, T3SS genes of ST24 strains are significantly differ from the genes of the strains with the rare ST. These data allow to elaborate common medicine for eradication of *C. psittaci* ST24 strains.
Acknowledgements: Supported by the State Order to the N.F. Gamaleya National Research Center for Epidemiology and Microbiology.

DOI 10.18699/BGRSSB-2018-066

MiRNA seed shifting: origin and functional implications

P. Vorozheykin[1]*, N. Yazikov[1], I. Titov[1, 2]
[1] *Novosibirsk State University, Novosibirsk, Russia*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
** e-mail: pavel.vorozheykin@gmail.com*

Key words: miRNA, seed shifting, targets, miRNA origin, isomiRs, miRNA biogenesis

Motivation and Aim: miRNAs constitute an extensive class of small RNAs. Mostly the miRNAs suppress the undesirable transcripts and genetic materials. Nucleotides 2–8 of the miRNA (so-called "seed" region) play a crucial role in the target recognition and binding; due to the imperfect binding rules one miRNA can regulate expression of many genes and thus control many biological functions. With the growth of deep-sequencing data, a diversity of miRNA location, length and sequence have been discovered (so-called isomiRs).

Results: To make out the isomiR role in the miRNA life cycle we consider the causes of the miRNA variants (evolutionary changes, processing mechanisms [1] and annotation errors [2]) and their functional implications. During the evolution the accumulated mutations in the pri-/pre-/miRNA change the miRNA boundaries (including "seed" location), sequence and secondary structure of the precursor. Other reasons of miRNA end heterogeneity are an inaccuracy of processing by the RNase proteins or by the splicing, their predisposition to the precursor sequence and structure and possible interconnection between Dicer cleavage errors and the length of the pre-miRNA hanging ends [1, 3]. The considered changes can lead to the new biological functions of miRNA and diverse levels of miRNA expression.

Conclusion: The results obtained in the work can be useful to predict new miRNAs and their functions based on the similarity to the existing ones. They allow to take a fresh look at the evolution of the silencing and can help to reconsider the role of the miRNA seed and the rest of the miRNA sequence and its precursor.

References

1. Titov I.I., Vorozheykin P.S. (2018). Comparing miRNA structure of mirtrons and non-mirtrons. BMC Genomics. 19(3):114.
2. Fromm B., Billipp T., Peck L.E. et al. (2015). A uniform system for the annotation of vertebrate microRNA genes and the evolution of the human microRNAome. Annu. Rev. Genet. 49:213-242.
3. Starega-Roslan J., Witkos T.M., Galka-Marciniak P., Krzyzosiak W.J. (2015). Sequence features of Drosha and Dicer cleavage sites affect the complexity of isomiRs. Int. J. Mol. Sci. 16(4):8110-8127.

DOI 10.18699/BGRSSB-2018-067

Prediction of gene expression level by using ChIP-Seq-derived data from the GTRD database and transcription start sites identified by the Fantom5 project

I. Yevshin[1], Yu. Kondrakhin[1, 2], R. Sharipov[1, 3]*, F. Kolpakov[1, 2]
[1] *BIOSOFT.RU, Ltd, Novosibirsk, Russia*
[2] *Institute of Computational Technologies SB RAS, Novosibirsk, Russia*
[3] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: shrus79@gmail.com*

Key words: ChIP-Seq, transcription factors, binding regions, GTRD database, transcription start sites, gene expression levels

Motivation and Aim: Since the beginning of the current millennium, ChIP-Seq has become the most powerful experimental technique for the genome-wide study of interactions between transcription factors (TFs) and DNA. For our analysis, we have used 4232 datasets of human transcription factor binding regions (TFBRs) identified by ChIP-Seq and collected in a GTRD database [1]. These datasets represented TFBRs of 694 distinct TFs. We also have exploited the expression profiles measured in 1829 main human primary cell types and tissues [2]. For each cell type or tissue, these expression levels were measured for collection of 209911 transcription start sites (TSSs) identified in the frames of the FANTOM5 project [2].

Methods and Algorithms: For each considered cell line or tissue, we developed a prediction model that consisted of classification and regression submodels. The goal of the classification submodel was to discriminate between expressed and non-expressed genes while the aim of the regression submodel was to predict the real-valued expression levels. For construction of the classification submodel, we used such machine learning approaches as random forest, support vector machine, perceptron, and Fisher's discriminant analysis. To construct regression submodel we used random forest, support vector machine, least squares regression, and regression on principal components. The features of regression submodels were defined as presence/absence of TFBRs in promoter regions (–5000, –1000), (–1000, –500), (–500, –200), (–200, –100), (–100, –1), (–1, +1), (+1, +100), and (+100, +500).

Results: We constructed the set of accurate models for prediction of expression levels in distinct cell lines and tissues. In particular, in case of HepG2 cell line, the accuracy of the prediction model was estimated by high correlation (0.866) between observed and predicted expression levels as well as by high value of the explained variance (74.9 %). In this case, the most significant feature for prediction was the occurrence of TFBRs within the following promoter regions: HNF-4a in (–1, +1), TAF-1 in (+100, +500), TAF-1 in (–100, –1), TAF-1 in (+1, +100), HNF-4a in (+100, +500), ZNF274 in (–5000, –1000), and TR4 in (–1, +1).

References

1. Yevshin I., Sharipov R., Valeev T. et al. (2017) GTRD: a database of transcription factor binding sites identified by ChIP-seq experiments. Nucleic Acids Res. 45:D61-D67.
2. The FANTOM Consortium and The RIKEN PMI and CLST. (2014) A promoter-level mammalian expression atlas. Nature. 507:462-471.

DOI 10.18699/BGRSSB-2018-068

Exome-wide survey of the Siberian Caucasian population

A.A. Yurchenko[1], N.S. Yudin[1, 2], M.I. Voevoda[1, 3]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Institute of Internal and Preventive Medicine – Branch of Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
** e-mail: andreyurch@gmail.com*

Key words: exome sequencing, population structure, associations, Siberia, Russia

Motivation and Aim: Population structure is a very important factor in medical genetic association studies which can compromise modern genomic methods not being properly accounted for. In this study, we identified exome genetic variants for 39 individuals from Novosibirsk and compared them with the previously published genome-wide data and exomes from the 1000 Genomes Project to understand the extent of the population stratification and compared allele frequencies in our sample and European dataset for medically and pharmacogenetically important variants.
Methods and Algorithms: SNVs and indels were identified using GATK pipeline according to the GATK Best Practices workflow with the sensitivity filter equal to 99.9 in the 39 samples. The variants were used for the PCA with the European and previously published Russian populations, ADMIXTURE analysis with the European dataset from 1000 Genomes Project and pairwise Fst estimation. We tested medically (ClinVar) and pharmacogenetically (PharmGKB) relevant variants for the differences in allele frequencies between the populations with PLINK software.
Results: A total of 136276 SNVs and 14464 indels were identified in the target regions of the Agilent SureSelect V5. The PCA demonstrated an intermediate emplacement of the Novosibirsk population between the Finnish and other European populations confirming the full congruity of the exome samples and previously published microarray data of the Novosibirsk population and Siberian Starovers. The results of the ADMIXTURE analysis showed a higher Finnish component in the Novosibirsk population than in other European groups and clearly distinguished the Novosibirsk population from the others at K = 4. We identified a highly significant albeit low (Fst = 0.005–0.009) level of the genetic differentiation between the Novosibirsk and other European Non-Finnish (ENF) populations. Among the 452 pharmocogenetically and 210 medically important variants we found 3 and 7 variants respectively which showed significant allele frequency differences between the Novosibirsk and the ENF population after the multiple testing correction. The most significant differences in allele frequencies were attributed to such genes as FCGR3B, TYR, OCA2, FABP6 and SLC4A1.
Conclusion: The Caucasian Novosibirsk population is quite homogeneous and significantly differentiated from other European populations from 1000 Genomes Project demonstrating a higher Finnish component and genetic congruence with the previously published Russian dataset including partially isolated Siberian Starovers.
Acknowledgements: This study was supported by the grant from the Russian Science Foundation (project no. 16-15-00127).

DOI 10.18699/BGRSSB-2018-069

Search for a hidden structure in genomic data based on a compressing autoencoder

A. Zarubin[1, 2]*, A. Markov[1], V. Stepanov[1], V. Kharkov[1]
[1] *Tomsk National Research Medical Center RAS, Research Institute of Medical Genetics, Tomsk, Russia*
[2] *Siberian State Medical University, Tomsk, Russia*
** e-mail: aleksei.zarubin@medgenetics.ru*

Key words: compressing autoencoder, genomic data, neural networks

Motivation and Aim: Recently, the problem of genomic data processing is becoming more urgent. The main difficulty lies in the "curse of dimension" – because the number of polymorphisms determined is usually several orders of magnitude greater than the number of genotyped samples. Usually the principal component analysis (PCA) is used to solve this problem. PCA also allows analyzing the hidden data structure by finding new variables. But it has some drawbacks, especially in analyzing complex interactions. Therefore, the application of a compressing autoencoder for genomic data as an approach that can approximate nonlinear interactions might be promising.
Methods and Algorithms: In order to evaluate how well the autoencoder manages to find the hidden structure in the data, we used a dataset on genotypes of 894 people from 28 populations from Russia and neighboring countries [1]. After filtering out the missing values, the remaining 113 749 polymorphic variants were used as a training sample. Artificial neural networks of different architectures were modeled in the R software environment using the Keras library [2].
Results: In the selection and evaluation of hyperparameters and the architecture of the neural network, the linear activation function for the output layer of the encoder and exponential linear unit for all the fully-connected and convolutional layers were most successful. The most effective algorithm of optimization was the Adam algorithm. As the final test model, we selected a 7 layered fully-connected perceptron with a total of 117 646 423 parameters and two linear outputs from the encoder. The training was gradual in 20 iterations with batch size 20. Seven populations are separated into mono groups quite well and quickly, but due to some non-linearity of axes, it is necessary to reduce the speed of training, or gradually exclude from training the samples, which have already clearly separated into a separate cluster and repeat the learning process in a smaller sample. The remaining populations are more difficult to differentiate, although they form a number of clusters. An increase in the number of output neurons from the encoder to 3 makes it possible to isolate up to 11 populations. In turn, using PCA, only up to 6 populations can be clearly identified.
Conclusion: Compressing autoencoder shows a higher efficiency comparing to PCA for searching the hidden structure in the genomic data and lowering the dimension. A reliable differentiation of populations requires the determination of many more hyperparameters the most effective use of linear and piecewise linear activation functions, Adam as an optimization algorithm and a reduced learning rate.

References

1. Triska P. et al. (2017) Between Lake Baikal and the Baltic Sea: genomic history of the Gateway to Europe. BMC Genetics. 18(1):110.
2. Chollet F.(2015) Keras. https://github.com/fchollet/keras

DOI 10.18699/BGRSSB-2018-070

Improvement of a test-system for detecting inherited and non-inherited genetic changes in living cells

A. Zhuk[1, 2]*, E. Stepchenkova[1, 2], S. Inge-Vechtomov[1, 2]
[1] *Saint-Petersburg State University, Saint-Petersburg, Russia*
[2] *Saint-Petersburg Branch of Vavilov Institute of General Genetics, Saint-Petersburg, Russia*
** e-mail: ania.zhuk@gmail.com*

Key words: DNA damages, mutations, mutation clusters, genetic safety, alpha-test, genetic toxicology, toxicogenomics

Motivation and Aim: To date, many different tests for evaluating mutagenic activity of chemical, physical and biological factors have been developed and widely used. New dataset of genetic sequence information and new technologies for genetic variation or functional gene expression analysis offer new opportunities for further improvement of classical methods used for evaluating potential environmental toxicants. Here we present an approach that enabling us to increase the efficiency of the alpha-test, an original method which was developed at the Department of Genetics and Biotechnology of the Saint-Petersburg State University.
Methods and Algorithms: Yeast *Saccharomyces cerevisiae* where used as a model organism. Alpha-test, forward and reverse mutagenesis assays and whole-genome sequencing where used to estimate genome instability. Additionally, we used a wide range of molecular genetic methods (gene cloning, yeast and bacterial cells transformation, polymerase chain reaction, primer and plasmid design methods, chromosome loss induction method), fluorescence microscopy methods, and flow cytometry.
Results: Using the alpha-test we evaluated the spectrum and frequency of inherited and non-inherited changes of genetic material, arising during induction of primary DNA lesions of various types induced by reference mutagens. Using yeast model we investigated distribution of mutations in the genome and the formation of homozygous mutation clusters in diploid yeast cells [1]. Combination of whole-genome sequencing with classical yeast genetics allowed us to investigate initial stage of mutation formation and possible mechanisms of phenotypic changes in non-dividing cells.
Conclusion: An integrated approach that combines methods of classical and molecular genetics, modern molecular genetics, cytological approaches and genome sequencing reveal opportunity for further development of a highly sensitive test system for genetic toxicology, that allows to detect genotoxic factors in eukaryotic cells without using a large number of selective media and manual labor.
Acknowledgements: The reported study was funded by RFBR according to the research project No. 18-34-00130.

References

1. Lada A.G. et al. (2017) Recombination is responsible for the increased recovery of drug-resistant mutants with hypermutated genomes in resting yeast diploids expressing APOBEC deaminases. Front. Genet. 8:A.202.

DOI 10.18699/BGRSSB-2018-071

DNA sequence features that may establish H3K27ac mark

A. Zubritskiy[1]*, Y.A. Medvedeva[1, 2, 3]
[1] *Institute of Bioengineering, Research Centre of Biotechnology, RAS, Moscow, Russia*
[2] *Department of System Biology and Bioinformatics, Vavilov institute of General Genetics, RAS, Moscow, Russia*
[3] *Department of Biological and Medical Physics, Moscow Institute of Physics and Technology, Dolgoprudny, Russia*
** e-mail: a.zubrit@gmail.com*

Key words: histone modification, H3K27ac, CpG dinucleotide

Motivation and Aim: Histone posttranslation modification (PTM) is a key mechanism of epigenetic regulation. Histone PTMs are highly variable, but in some genomic loci they may be highly similar between many cell lines. It has been reported that the attraction of polycomb repressive complex 2 (PRC2) and consequent trimethylation of Lys27 of histone H3 (H3K27me3) is correlated with the local density of CpG dinucleotides [1]. Other sequence patterns, such as transcription factor (TF) binding sites (TFBS), also affect the presence of histone modifications: the ENCODE project demonstrated a specific histone modification profile around binding sites of many TFs [2].
Another histone PTM, acetylated Lys27 of histone H3 (H3K27ac) is mutually exclusive with H3K27me3, but it is still unclear if H3K27ac shares any sequence specific pattern. In this work we performed a direct experiment to test whether a specific genomic sequence is capable of recovering H3K27ac.
Methods and Algorithms: We generated stably transfected Caki1 cell line that contain selection cassette flanked by LoxP sites in the beta-globin locus. This cassette can be exchanged via *Cre*-mediated recombination to any other DNA sequence of interest flanked by LoxP sites.
Results: We performed insertion of 10 DNA sequences into beta-globin locus. This step could help us to unify the epigenetic background and remove any preceding histone PTMs from DNA fragments. 3 of inserted sequences are GC-rich promotor regions with high number of CpG dinucleotides, while other DNA sequences have less than two CpG's per fragment and CG content in range between 27 and 68 %.
Surprisingly, none of the GC- and CpG rich promoter regions, that were acetylated in their original genomic loci recovered H3K27ac after relocation to a beta globin locus, while two extremely GC-rich but CpG poor sequences gained H3K27ac in the beta globin locus, while in their native location they had no H3K27ac.
Conclusion: The establishment of H3K27ac may not depend on a CpG content of the sequence per se, as opposed to its antagonistic mark H3K27me3, but might depend on an elevated GC content.
Acknowledgements: Supported by the grant from Russian Science Foundation (15-14-30002).

References

1. Jermann P. et al. (2014) Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. Proc Natl Acad Sci USA. 111(33):E3415-3421.
2. Hoffman M.M. et al. (2013) Integrative annotation of chromatin elements from ENCODE data. Nucleic Acids Res. 41(2):827-841.

PROTEOMICS AND BIOINFORMATICS

DOI 10.18699/BGRSSB-2018-072

Investigation of proteome of Baikal endemic amphipod *Eulimnogammarus cyaneus* (Dybowsky, 1874) using LC-MS

D. Bedulina[1]*, S. Kalkhof[2], A.Gurkov[1], P. Drozdova[1], M. Timofeyev[1], T. Luckenbach[3]
[1] *Irkutsk State University, Irkutsk, Russia*
[2] *The University of Applied Sciences, Coburg, Germany*
[3] *Helmholtz Centre for Environmental Research, Leipzig, Germany*
** e-mail: daria.bedulina@gmail.com*

Key words: proteogenomics, amphipoda, Baikal, eco-proteomics, LC-MS proteomics

Motivation and Aim: Lake Baikal is unique ancient ecosystem, currently threatened by the anthropogenic impact and climate change [1]. Thus, developing of the monitoring bioassays for this lake if of the great importance. Aim of the current study was to investigate proteome of a potential model biotest species for Lake Baikal – amphipod *Eulimnogammarus cyaneus* (Dybowsky, 1874), using a proteogenomic approach.
Methods and Algorithms: We used 1D-PAGE of total protein lysate of *E. cyaneus*, followed by gel fragmentation, in-gel trypsinolysis and nano-HPLC/nano-ESI-MS/MS using LTQ Orbitrap XL™ (Thermo Fisher Scientific). Transcriptome assembly of the respective species was used to generate a specific database for protein identification, which was carried out with MaxQuant software [2].
Results: MS-spectra were assigned to 1028 protein groups. Among them the first five major groups were belonging to the following families (identified by InterProScan Ids): tubulins, ubiquitin, heat shock protein 70 kDa, actin and hemocyanin. Used approach allowed us to validate approximately 9 % of the transcriptome. Label free quantification using MaxQuant revealed 16 differently expressed ($p < 0.05$, Bonferroni correction) protein groups in males and females of the studied species.
Conclusion: Thus, for Baikal endemic amphipod *E. cyaneus* experimentally-validated proteomic database was established and a sexual dimorphism of proteomes was investigated. This database will allow to carry out a broad spectrum of eco-physiological and eco-toxicological researches with amphipods and to reveal new specific biomarkers of anthropogenic impact to the unique ecosystem of Lake Baikal.
Acknowledgements: Supported by the RFBR (14-01-00885, 14-01-31134), and Grant of the President of the Russian Federation (МК-6804.2018.4).

References

1. Timoshkin O.A., Samsonov, D.P., Yamamuro, M., Moore, M.V., Belykh, O.I., Malnik, V.V., ... Fedorova G.A. (2016) Rapid ecological change in the coastal zone of Lake Baikal (East Siberia): Is the site of the world's greatest freshwater biodiversity in danger? J. Great Lakes Research. 42(3):487-497.
2. Cox J., Mann M. (2008) MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. Nature Biotechnol. 26(12):1367.

DOI 10.18699/BGRSSB-2018-073

Novel amyloid-forming protein in *Escherichia coli*

M. Belousov[1]*, S. Bondarev[1], A. Kosolapova[1, 2], K. Antonets[1, 2], M. Belousova[2], A. Sulatskaya[3], M. Sulatsky[3], G. Zhouravleva[1], I. Kuznetsova[3], K. Turoverov[3, 4], A. Nizhnikov[1, 2]

[1] *Saint Petersburg University, St. Petersburg, Russia*
[2] *All-Russia Research Institute for Agricultural Microbiology, Pushkin, St. Petersburg, Russia*
[3] *Institute of Cytology RAS, St. Petersburg, Russia*
[4] *Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia*
** e-mail: belousovmix@gmail.com*

Key words: secreted metalloprotease, functional amyloids, bacteria, *Escherichia coli*

Motivation and Aim: Amyloids are protein fibrils with a characteristic spatial structure. In previous studies, using a method for the proteomic screening and identification of amyloids (PSIA) [1], we identified 61 proteins of *Escherichia coli* that formed detergent-resistant aggregates *in vivo* and without overproduction [2]. Among these proteins, YghJ was the most enriched with bioinformatically predicted amyloidogenic regions. YghJ is a lipoprotein and important virulence factor of *E. coli* containing the zinc metalloprotease M60-like domain ($YghJ_M$) that is involved in the pathogenesis of the enterotoxigenic strains via mucin degradation in the intestine.

Methods and Algorithms: To analyze amyloid properties of the $YghJ_M$ fibrils *in vitro*, transmission electron, confocal and polarization microscopies of the fibrils were used. For *in vivo* analysis, curli-dependent amyloid generator system (C-DAG), which provides export of the target protein to the surface of bacterial cells, was employed.

Results: We detected detergent-resistant aggregates of $YghJ_M$ by SDS-PAGE and SDD-AGE and confirmed that these aggregates are resistant to α-chymotrypsin protease treatment. We analyzed the fibrillary morphology of the obtained $YghJ_M$ aggregates using transmission electron microscopy. Next, we demonstrated that the $YghJ_M$ aggregates bind Thioflavin-T amyloid-specific dye and exhibit CD-spectra typical for protein aggregates rich in β-sheets. The Congo red stained $YghJ_M$ fibrils demonstrated apple-green birefringence which is considered to be the "gold standard" for verification of the amyloid structure. Finally, we showed that $YghJ_M$ forms amyloid fibrils on the surface of the *Escherichia coli* cells using C-DAG [3].

Conclusion: We demonstrated that $YghJ_M$ forms amyloid fibrils *in vitro* and *in vivo*. Our data on the amyloid properties of the YghJ protein highlight the role of this protein in the pathogenesis and suggest possible mechanism for its functioning.

Acknowledgements: Supported by the Grants of the President of the Russian Federation (MK-3240.2017.4, MK-512.2017.4) and by RFBR (16-34-60153, 17-54-150002).

References

1. Nizhnikov A.A. et al. (2016) Interaction of prions causes heritable traits in *Saccharomyces cerevisiae*. PLoS Genetics. 12(12). e1006504.
2. Antonets K.S. et al. (2016) Proteomic analysis of *Escherichia coli* protein fractions resistant to solubilization by ionic detergents. Biochemistry (Moscow). 81(1):34-46.
3. Belousov M.V. et al. (2018) M60-like metalloprotease domain of the *Escherichia coli* YghJ protein forms amyloid fibrils. PLoS ONE. 13(1):e0191317.

DOI 10.18699/BGRSSB-2018-074

Proteomic analysis of extremely stable soluble high molecular mass multi-protein complex of human placenta

E. Burkova[1]*, G. Nevinsky[1, 2]
[1] *Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: burkova_evgeniya_e@mail.ru*

Key words: placenta, protein complex, catalytic activities

Motivation and Aim: Many biological function of placenta are performed not just a set of individual proteins, but also different oligomeric structures and complexes. Herewith activities of complexes may considerably differ from activities of individual proteins. The research of protein complexes will allow to characterize some molecular mechanisms of functioning of placenta more fully. We recently obtained from human placenta an extremely stable multi-protein complex (SPC, ~1000 kDa) [1]. This complex possesses DNase and catalase activities [1, 2]. The aim of the present work was to investigate a protein composition and catalytic activities of the extremely stable high molecular mass multi-protein complexes (SPC) from placenta of healthy mother.
Methods and Algorithms: We identified the proteins of SPCs by MALDI mass MS and MS/MS spectrometry using proteins tryptic hydrolyzates after proteins separation by 1D- and 2D-electrophoresis. Also we analyzed catalytic activities of this very stable complex.
Results: It was shown that SPCs contain twelve proteins and their different isoforms: hemoglobin, alkaline phosphatase, cytoplasmic actin, human serum albumin, chorionic somatomammotropin hormone, heart shock protein beta-1, peroxiredoxin-1, 78 kDa glucose-regulated protein, protein disulfide isomerase A3, serotransferrin, annexin A5, and IgGs. It was shown the SPCs possess RNase, peroxidase (H_2O_2-dependent) and oxidoreductase (H_2O_2-independent) activities. An addition, investigation of cytotoxic effect on human cancerous cell lines has shown that the SPCs reveal high cytotoxicity.
Conclusion: Progress in the study of placental protein complexes can promote understanding of their biological functions.
Acknowledgements: Supported by Russian State funded budget project (VI.62.1.5, 0309-2016-0003), Grant of RFBR No. 16-04-00609.

References

1. Burkova E.E., Dmitrenok P.S., Sedykh S.E., Buneva V.N., Soboleva S.E., Nevinsky G.A. (2014) Extremely stable soluble high molecular mass multi-protein complex with DNase activity in human placental tissue. PLoS One. 9(11):e111234.
2. Burkova E.E., Nevinsky G.A. (2016) Very stable high molecular mass multi-protein complex from human placenta. FEBS J. 283(Suppl. S1):211.

DOI 10.18699/BGRSSB-2018-075

The search of blood-based biomarkers for schizophrenia by proteomics methods

E. Dmitrieva[1]*, L. Smirnova[1], A. Seregin[1], A. Letova[2], E. Kornetova[1], A. Semke[1], V. Zgoda[3], S. Ivanova[1]
[1] *Mental Health Research Institute, Tomsk, Russia*
[2] *Siberian State Medical University, Tomsk, Russia*
[3] *Institute of Biomedical Chemistry, Moscow, Russia*
** e-mail: Lena-312tom@yandex.ru*

Key words: biomarkers, schizophrenia, proteomics

Motivation and Aim: Schizophrenia is a complex neuropsychiatric disorder whose symptoms lead to significant disability throughout life, causing a massive personal and social burden. Understanding pathogenesis of schizophrenia requires studies of not only gene expression and DNA variations, but also studies, of the abundance and modifications of various proteins, and their distribution at anatomical, cellular, and subcellular levels. Proteomic studies of proteins of this disease may contribute to the understanding of the molecular mechanisms of schizophrenia. They may also indicate pathological changes in brain cell's in patients with schizophrenia.
Methods and Algorithms: For the research we used the serum of 20 healthy and 30 patients with schizophrenia. Patients were treated in clinics of Mental Health Research Institute, Tomsk, Russia. Diagnostics was carried out in accordance with the current classification ICD-10. Preparation of samples included: purification from serum major proteins by affinity chromatography, separation of proteins by one-dimensional electrophoresis, in-gel tryptic hydrolysis of the separated proteins. LC-MS/MS analysis of the resulting peptides using an ion trap XCT Ultra mass spectrometer (Agilent Technologies). Identification of proteins was carried out using Mascot software Ver. 2.1 («Matrix Science», USA). Statistical analysis was performed using Fisher's exact test with Yates' correction using the program STATISTICA 10.0.
Results: The proteins found in our study are involved mostly in biological processes, such as regulation of nucleic acid metabolism, immune response and also unknown processes. We found Centromere-associated protein E and Bromodomain adjacent to zinc finger domain protein 2B their DNA binding protein function and transcription factor, respectively. Complement factor H-related protein 2 is involved in the regulation of complement activation. Dermisidine protein is involved in the antimicrobial humoral response, but one of the functions is related to the survival of neurons and displays of phosphatase activity.
Conclusion: Identified proteins can be included in a sensitive and specific biomarker panel for diagnosis of schizophrenia, both for diagnosis and for subsequent response to treatment.
Acknowledgements: Mass spectrometric analysis was carried out using the equipment of 524 "Human Proteome" Core Facility of the Institute of Biomedical Chemistry (IBMC) Moscow which 525 is supported by Ministry of Education and Science of the Russian Federation (unique project ID 526 RFMEFI62117X0017) Support by Grant of RSF (18-15-00053).

DOI 10.18699/BGRSSB-2018-231

Molecular characterization of aquaglyceroporine: a novel mutation in *LMAQP1* from *LEISHMANIA MAJOR* (MRHO/IR/75/ER)

G. Eslami*, M. Ghavami
Research Center for Food Hygiene and Safety, Shahid Sadoughi University of Medical Sciences, Yazd, Iran
** e-mail: eslami_g2000@yahoo.com*

Key words: Aquaporin 1, Leishmania, Molecular Dynamics Simulation, Antimony

Motivation and Aims: The first line treatment for cutaneous leishmaniasis is pentavalent antimony such as sodium stibogluconate (pentostam) and meglumine antimonite (glucantime). One of the most important way to uptake the drug is by a transmembrane protein, called aquaglyceroporin encoded by *Aquaglyceroprotein1* (*LmAQP1*). So far, there is no report on *LmAQP1* from *L. major* (MRHO/IR/75/ER), therefore, in this study, molecular characterization of *LmAQP1* was reported.
Methods and Algorithms: *L. major* (MRHO/IR/75/ER) promastigotes were cultured, and then DNA extraction and RNA extraction were done and followed by cDNA synthesis. Amplicons resulted from PCR and RT-PCR using specific primers were purified and sequenced. Molecular characterization was done by bioinformatically softwares such as BLAST, *ClustalW2*, and RMSD.
Results: Amplicons resulted from PCR and RT-PCR showed equal size in length. BLASTn analysis showed a point nucleotide change in *LmAQP1* gene that encoded 282-amino-acid long protein with a mutation at position 154 including replacement of alanine by threonine. The observed mutation in the interested gene was assessed using the above mentioned software. The mentioned gene was submitted at GenBank, NCBI with accession number of KU514052.
Conclusion: The functional prediction of the protein encoded from *LmAQP1* showed that the mentioned mutation could not affect the three dimension structure, but it may modify the drug uptake potential of this important channel. Based on from *LmAQP1* role, it seems to be an appropriate candidate for drug development. According to search through internet, this is the first report of *LmAQP1* from *L. major* (MRHO/IR/75/ER).
Availability: The mentioned strain is available in esearch Center for Food Hygiene and Safety, Shahid Sadoughi University of Medical Sciences, Yazd, Iran.

References

1. Marquis N. et al. (2005) Modulation in aquaglyceroporin AQP1 gene transcript levels in drug-resistant Leishmania. Mol. Microbiol. 57:1690-1699.
2. Figarella K. et al. (2007) Biochemical characterization of *Leishmania major* aquaglyceroporin LmAQP1: possible role in volume regulation and osmotaxis. Mol. Microbiol. 65:1006-1017.
3. Mukhopadhyay R. et al. (2011) The role of alanine 163 in solute permeability of *Leishmania major* aquaglyceroporin LmAQP1, Mol. Biochem. Parasitol. 175:83-90.
4. Monte-Neto R. et al. (2015) Intrachromosomal amplification, locus deletion and point mutation in the aquaglyceroporin AQP1 gene in antimony resistant Leishmania (Viannia) guyanensis. PLoS Negl Trop Dis. 9:e0003476.

DOI 10.18699/BGRSSB-2018-076

Unraveling of gene expression control in genome-reduced bacteria. The rally goes on…

V. Govorun
Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical Biological Agency, Moscow, Russia
e-mail: Vgovorun@yandex.ru

Key words: *Mycoplasma gallisepticum*, gene expression regulation, RNA-seq, ribosomal profiling, cell cycle control

Mycoplasmas of class Mollicutes represent extremely reduced bacteria yet capable of self-replication without the aid of eukaryotic cell. Thus they are good models to study the basic principles of the organization of living cell. In the present work we used a combination of high-throughput technologies to elucidate the transcription control network in a model organism *Mycoplasma gallisepticum*. Using RNA-seq we identified transcription units and performed identification and activity quantitation of promoters of *M. gallisepticum*. We used comparative genomics and promoters' identification across different Mollicutes species as well as a set of perturbation models to identify putative regulatory sequences. We used ribosomal profiling and proteomics to study the transfer of genetic information to the protein level. We identified that the significant amount of transcription regulation in *M. gallisepticum* is achieved via weak determinants of the core promoter. In addition *M. gallisepticum* features conditional terminators, which may undergo read-through under the stress and contribute to transcription profile. Ribosomal profiling demonstrated that while under exponential growth translation is near equal to transcription under stress ribosome demonstrates selectivity towards particular transcripts. Analysis of cross-species promoters' conservation demonstrated that transcriptional control is largely species specific, while the most conserved transcriptional regulators control cell cycle.
Acknowledgements: This study supported by project Russian Science Foundation No. 14-24-00159.

DOI 10.18699/BGRSSB-2018-077

Protein structural domain prediction via machine learning approach

D. Iakovlev*, A. Kobchenko, E. Semina
Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia
** e-mail: yakovlevd@niboch.nsc.ru*

Key words: protein domain, protein visualization, machine learning, clustering

Motivation and Aim: Amount of solved protein structures in databases such as PDB is growing incredibly fast, making manual investigations in this field more and more challenging. One of a basic and, usually, manual steps of protein analysis is a structural domains annotation. A concept sometimes taken as a rough working definition of a structural domain is that, if excised, the domain should remain folded as a stable structure [1]. Therefore, residues in protein domains are distributed more densely than averagely in protein and would be detectable by clustering algorithms. Despite there are many tools for protein structure analysis and visualization, no one of them can automatically split protein into domains using only structural information (e. g. a PDB file). Using protein architecture database such as CATH or SCOP2 is also difficult if we deal with proteins that do not have annotated homologues in there. Some methods for automatic detection of protein domains have been already developed earlier [1] but we improved them using modern algorithms and computational approaches.
Methods and Algorithms: Training sample of structural domains boundaries was obtained from CATH database [2]. Boundaries were tested to not cross secondary structure features (α-helices and β-sheets), yielding a set of approximately 130000 marked up polypeptide chains. To predict structure domains we used clusterization of C_α atoms of polypeptide chain by BIRCH [3] algorithm with some *ad hoc* modifications: handling of helix/sheet integrity and amino acid hydrophobicity. To improve clustering precision neural network was used for optimization of hyperparameters: BIRCH branching factor and threshold and post-clustering division factor. Quality of domain prediction was estimated by scoring function which matches each predicted domain with known ones and counts a proportion of correctly predicted amino acid residues.
Results: A new machine learning-based method for protein structural domain prediction was implemented. Predictive power of the clustering model was proved by cross-validation technique.
Availability: A web service for protein domain prediction and visualization is available at protein-clustering.ru
Acknowledgements: We express our gratitude to Yuri Vyatkin for formulation of the problem and curation during our investigation.

References

1. Taylor W.R. (1999) Protein structural domain identification. Protein Engineering. 12(3):203-216.
2. Sillitoe I. et al. (2015) CATH: comprehensive structural and functional annotations for genome sequences. Nucleic Acids Res. 43(D1):D376-D381.
3. Zhang T., Ramakrishnan R., Livny M. (1996) BIRCH: an efficient data clustering method for very large databases. Proceedings of the 1996 ACM SIGMOD international conference on Management of data - SIGMOD ‘96:103-114.

DOI 10.18699/BGRSSB-2018-078

SRM-based approach for β-lactamases profiling

E. Ilgisonis[1]*, M. Rubtsova[2], G. Presnova[2]
[1] *Institute of Biomedical Chemistry, Moscow, Russia*
[2] *Lomonosov Moscow State University, Moscow, Russia*
** e-mail: ilgisonis.ev@gmail.com*

Key words: β-lactamase, antibiotics, mass-spectrometry, proteomics

Motivation: In recent years, multidrug resistant strains are actively spreading, which are resistant to several or even all antibiotics. Resistance to antibiotics is a big botnet in the selection of personalized therapy. In this regard, there is a problem of identifying bacteria resistable to β-lactam antibiotics. The main mechanism of this type of resistance is the production of beta-lactamase enzymes. They are extremely different in structure and substrate specificity. In this work several approaches have been developed for the MS profiling of beta-lactamases to determine the family and type, depending on which antibiotic therapy is selected.
Materials and methods: From the UniProt and literary sources amino acid sequences belonging to mutant forms of beta-lactamases (1) of different types (2) were loaded. Using Skyline virtual trypsinolysis was performed and using BLAST-search proteotypic peptides (specific for each form) were determined for subsequent mass spectrometric analysis.
Results: Proteorypic peptides for several TEM, SHV, CTX-M families were chosen. Methods for their detection using SRM mass-spectromical technology were developed. It was shown that this peptides can be used for mass-spectrometric identification basing on more than 50 experiments in PRIDE. At the same time, it was not possible to select specific peptides that distinguish mutant forms TEM-1 from TEM-84.
Conclusions: Bioinformatic analysis of mutant forms of beta-lactamases allowed to determine specific peptides, taking into account the changes in the nucleotide sequence for the subsequent creation of a set of peptides and their mass spectrometric identification in the samples. An experimental verification of the proposed approach presupposes a directional mass spectrometric analysis of beta-lactamases in the periplasmic fraction of producer strains to identify types and key, clinically relevant mutations.
Acknowledgements: Ek.I. acknowledge the Leading Scientific School of Prof. Andrey Lisitsa (No. NSh6313.2018.4).

DOI 10.18699/BGRSSB-2018-079

Human blood proteins after long duration space flights

D.N. Kashirina[1]*, A.J. Percy[2], L.Kh. Pastushkova[1], C.H. Borchers[2], E.N. Nikolaev[3], I.M. Larina[1]

[1] *Institute for Biomedical Problems – Russian Federation State Scientific Research Center RAS, Moscow, Russia*

[2] *Genome British Columbia Proteomics Centre, University of Victoria, Victoria, BC, Canada*

[3] *Skolkovo Institute of Science and Technology, Skoltech, Moscow region, Russia*

** e-mail:daryakudryavtseva@mail.ru*

Key words: blood plasma, cosmonauts, mass spectrometry

Motivation and Aim: The conditions of space flight have a significant effect on the physiological processes in the human body, yet the molecular mechanisms driving physiological changes remain unknown.

Methods and Algorithms: Blood samples of 18 Russian cosmonauts who had conducted long-duration missions to the International Space Station were collected 30 days before launch (L-30), on the first (R+1) and seventh (R+7) days after landing. A panel of 125 proteins in the blood plasma was quantitated by a well-established and highly-regarded targeted mass spectrometry approach involving multiple reaction monitoring (MRM) in conjunction with stable isotope-labeled standards at the University of Victoria - Genome BC Proteomics Centre.

Results: In our data set, 125 plasma proteins were detected and quantitated in fmol/µl. Concentrations of most proteins were reduced at R+1 with a gradual return of their concentrations to the background level by R+7, except apolipoprotein A-II and serotransferrin whose concentrations remained reduced. This concentration dynamics partly reflected the dilution of the plasma during the restoration of the plasma circulation volume at R+1 and the gradual return of the plasma composition to the background values by R+7. Therefore, proteins whose concentration although insignificantly, but increased on the first day after the landing were of interest. It was determined 30 such proteins, including complement system proteins, acute phase proteins, proteases and their inhibitors. The only one protein whose concentration significantly increased at R+1 was S100A9 protein, that broadly regulates vascular inflammation and contributes to the biological response to vascular injury. This protein could serve as a marker of activation of proinflammatory reactions or damage of endothelial cells due to landing stress.

Conclusion: A decrease in the level of certain blood' proteins, without these levels being restored to their pre-flight levels 7 days after the flight, could clearly be interpreted as the impact of space flight on the human body. The increase in proteins involved in the regulation of inflammatory reactions and the immune response may adversely affect the functioning of the endothelium.

Acknowledgements: The work was supported by RFBR grant No. 18-34-00524. We express special thanks to Bruker Daltonics for financial and technical support.

DOI 10.18699/BGRSSB-2018-080

Synergy of shotgun-MS and 2DE for analysis of proteome heterogeneity

O. Kiseleva*, E. Poverennaya, S. Naryzhny, V. Zgoda, A. Lisitsa
Institute of Biomedical Chemistry, Moscow, Russia
** e-mail: olly.kiseleva@gmail.com*

Key words: proteomics, proteoforms, proteome heterogeneity, shotgun MS, 2DE

Motivation and Aim: Proteome heterogeneity is an unavoidable factor that complicates proteomic studies. Necessity for investigation of such events as alternative splicing, single amino acid polymorphisms and posttranslational modifications at protein level is determined by significant influence of these events on expression and functional properties of proteins. Careful analysis of proteome, which takes into account aberrant proteoforms, provides a basis for understanding the machinery of complex biochemical systems.
Methods and Algorithms: Basing on the results of RNASeq analysis of HepG2 cell line, we created a transcriptomespecific library, containing 52 thousand protein sequences, encoded by 12 thousand genes. Such library allows to focus on individual variations and helps to avoid uncontrolled extension of search area by populational data, thus reducing FDR. One of the most popular methods of proteomic analysis – shotgun mass spectrometry – is characterized by low (ca. 20 %) coverage of protein sequence. Short peptides, detected with shotgun-MS, often do not allow to distinguish highly homologous proteoforms. To empower shotgun approach, we added two dimensional electrophoresis (2DE). After 2DE fractioning we cut the gel in 96 cells and analyzed every cell by MS. Fractioning of protein mixture before MS allows to enrich the results with coordinates of proteoform on the gel (p*I* and MW).
Results: 2DE profiling with further MS analysis allowed to discover over 2358 proteoforms (1658 of which were canonical, 224 – splice-forms, and 115 – with amino acid polymorphisms) encoded by 1904 genes. Without 2DE we identified only 925 proteoforms because of sample heterogeneity and lack of knowledge of physicalchemical parameters of proteoforms, used for specification of certain proteoform.
Conclusion: Effective tandem of 2DE/MS allows us to forecast modifications, which can change physicalchemical parameters (and the location of protein spot on the gel, consequently). Obtained results consist not only of evaluation of proteoforms implemented at the protein level, but also of improvement of experimental approaches to cell proteotyping.
Acknowledgements: We acknowledge the Leading Scientific School of Prof. Andrey Lisitsa (grant the Russian Federation of President NSh6313.2018.4).

DOI 10.18699/BGRSSB-2018-081

The study of protein composition of *Triaenophorus* sp. at different stages of the life cycle and in different body segments

A. Kochneva[1]*, E. Borvinskaya[1, 2], D. Bedulina[2]
[1] *Institute of Biology of the Karelian Research Centre of the Russian Academy of Sciences, Petrozavodsk, Russia*
[2] *Research Institute of Biology of Irkutsk State University, Irkutsk, Russia*
** e-mail:kochnevaalbina@gmail.com*

Key words: Cestoda, *Triaenophorus* sp., proteomics, stages of the life cycle, different body segments

Motivation and Aim: Members of class Cestoda are unique organisms, which are obligate parasites with a complex life cycle that includes the life stages at very different habitats. Thus, cestodes can be considered as an appropriate model for the study of molecular adaptations. It is known that the tapeworms body structure is simplified, but a pronounced functional differentiation exists along the anteroposterior axis of the body and of the helminth strobila. The aim of this research was to investigate a protein polymorphism in different body segments (scolex, immature and mature segments of the strobila) and life stages of two species of Cestoda – *Triaenophorus nodulosus* (Pallas, 1781) and *T. crassus* (Forel, 1868).
Methods and Algorithms: The 2D-DIGE (two-dimensional difference gel electrophoresis) was used to study the electrophoretic mobility of proteins. A comparative analysis of gel images was carried out using the SameSpots software (TotalLab, England). Protein identification was performed with mass spectrometry (Agilent ESI-Q-TOF 6538 UHD (Agilent Technologies)) combined with HPLC (Agilent 1260 (Agilent Technologies)).
Results: Comparative analysis of 2D-DIGE electrophoregrams of plerocercoid and adult worms of *T. crassus* revealed two protein spots, significantly different in intensities. In *T. nodulosus*, significant differences were found for three protein spots. Significant differences in spot intensities were found in 2D-DIGE protein patterns of scolex, immature and mature segments of *Triaenophorus* sp. Two of these spots were identified as tropomyosin. Intensities of three spots were found significantly different on 2D-DIGE gels of plerocercoids of *T. nodulosus* and *T. crassus*. Protein patterns of adult worms differed in two spot intensities.
Conclusion: Thus, for the first time the electrophoretic protein mobility of two fish parasites species were described. This data will be useful in future research to develop species-specific antibody for helminthosis diagnostics in fishery.
Acknowledgements: Supported by the Russian Foundation for Basic Research (No. 17-04-01700_A).

DOI 10.18699/BGRSSB-2018-082

Screening of *Salmonella enteritidis* and *Bacillus thuringiensis* proteomes for potentially amyloidogenic proteins

A.O. Kosolapova[1, 2], K.S. Antonets[1, 2]*, A.A. Nizhnikov[1, 2]
[1] *Laboratory for Proteomics of Supra-Organismal Systems, All-Russia Research Institute for Agricultural Microbiology, Pushkin, St. Petersburg, Russia*
[2] *Dept. of Genetics and Biotechnology, St. Petersburg State University, St. Petersburg, Russia*
** e-mail: kirantonez@gmail.com*

Key words: amyloid, pathogen, bacteria

Motivation and aims: *Salmonella enteritidis* and *Bacillus thuringiensis* are pathogens in different groups of organisms. *S. enteritidis* causes intestinal infections of rodents and *B. thuringiensis* is insecticidal pathogen used to control pests. Recently it was shown that bacterial factors of virulence might form amyloids and it might play an important role in infection [1]. The aim of this study was to find potentially amyloidogenic proteins in the proteomes of *S. enteritidis* and *B. thuringiensis*.
Methods: To find amyloidogenic regions in proteins we used Waltz tool [2], which was developed to predict amyloid-forming regions; Aggrescan [3], which finds aggregation-prone regions; and SARP algorithm [4], which was designed to find compositionally biased regions. All protein sequences and their annotations were downloaded from Uniprot database (uniprot.org).
Results: We found that about 30 % of all proteins of *S. enteritidis* and *B. thuringiensis,* harbored regions predicted with Waltz, up to 1 % of proteins were enriched with amyloidogenic amino acids asparagine and glutamine (QN-rich) and more than 80 % of proteins contained regions found with Aggrescan. Proteins with Waltz-predicted regions are mostly membrane-associated and participate in transport or biosynthesis. The group of flagellar proteins were significantly enriched with QN-rich proteins, found with SARP. The flagellum proteins participate in host-pathogen interactions. All proteins attributed to the process of pathogenesis by GeneOntlogy database, contained regions found with Aggrescan.
Conclusion: We have found that various factors of virulence of the phylogenetically distant pathogenic bacteria *Salmonella enteritidis* and *Bacillus thuringiensis* are rich in amyloidogenic regions. Thus, amyloid formation by the factors of virulence might represent evolutionary conservative molecular mechanism underlying pathogenesis of different bacterial infections.
Acknowledgements: The analysis of amyloidogenic regions with Waltz and SARP programs was supported by the Grant of the President of the Russian Federation (MK-3240.2017.4 to AAN). Prediction of aggregation-prone regions in *B. thuringiensis* and *S. enteritidis* within Aggrescan was supported by Budget Project (0664-2015-0018).

References

1. Belousov M.V. et al. (2018) M60-like metalloprotease domain of Escherichia coli YghJ protein forms amyloid fibrils. PLoS One. 13(1):e0191317.
2. Maurer-Stroh S. et al. (2010) Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. Nat. Methods. 7:237-242.
3. Conchillo-Sole O. et al. (2007) AGGRESCAN: a server for prediction and evaluation of "hot spots" of aggregation in polypeptides. BMC Bioinformatics. 8:65.
4. Antonets K.S., Nizhnikov A.A. (2013) SARP: a novel algorithm to assess compositional biases in protein sequences. Evol. Bioinform. 9:263-73.

DOI 10.18699/BGRSSB-2018-083

Residue-residue contacts in modeling protein structure

M. Kotulska[1]*, P.P. Wozniak[1], B. Konopka[1], M. Kurczynska[1], G. Vriend[2], A. Sokolowska[1], J. Wojciechowski[1]

[1] *Wroclaw University of Science and Technology, Faculty of Fundamental Problems of Technology, Department of Biomedical Engineering, Wroclaw, Poland*

[2] *Radboud University Medical Centre, Centre for Molecular and Biomolecular Informatics, Nijmegen, The Netherlands*

** e-mail: malgorzata.kotulska@pwr.edu.pl*

Key words: correlated mutations, contact maps, mirror chirality

Motivation and Aim: Experimental study of the molecular structure is expensive, time consuming and not always feasible. The difficulty with experiments results in a great disproportion between the numbers of known structures and protein sequences. Bioinformatical modeling of a protein predicts its molecular structure. Various computational methods use different approaches and generate different structural models of the same protein. Also models from the same method can differ. Choosing the best model is a difficult task, especially that accuracy of each method is determined based on its average efficiency related to available experimental structures. It does not necessarily correspond to the situation of modeling an individual unknown protein. Information of contact distribution supports the modeling as it brings additional different information.

Methods and Algorithms: We applied Direct Contact Analysis (DCA) methods, which use correlated mutations of amino acids. Obtained contact maps were first used to support prediction of molecular structure. The studies regarded characteristics of the results and obstacles appearing in modeling of protein structures based on full and incomplete contact maps. The maps resulting from modeling are incomplete mainly because of insufficient accuracy of current methods. We proposed several methods to improve it, based on machine learning and structural filters, as well as a method to forecast efficiency of contacts prediction for an individual protein.

Results: The studies showed that for correct protein reconstruction 30 % of random contacts is sufficient, however it depends on the protein class. Density of contacts and their positions relative to the interior of the molecule are also important in proper reconstruction [1]. Contact maps do not provide all the necessary information though, allowing for models of mirror chirality difficult to filter out. The knowledge of even incomplete maps allows selecting misfolded from properly folded proteins [2]. We designed a regression model that forecasts the accuracy of contact prediction for individual proteins with an average error of 7 percentage points [3]. Contact prediction can be further raised by including machine learning methods and structural filters.

Conclusion: The results obtained in our studies show potential of contact maps in modeling protein structures.

References

1. Konopka B.M. et al. (2014) Automated procedure for contact-map-based protein structure reconstruction. Journal Membrane Biol. 247(5):409-420.
2. Wozniak P.P. et al. (2017) Forecasting residue-residue contact prediction accuracy. Bioinformatics. 33(21):3405-3414.
3. Wozniak P.P. et al. (2017) Correlated mutations select misfolded from properly folded proteins. Bioinformatics. 33(10):1497-1504.

DOI 10.18699/BGRSSB-2018-084

Integrated experimental and computational pipeline for proteome-wide *in tissue* crosslinking analysis

K.A.T. Makepeace[1], Y. Mohammed[1, 2], E.L. Rudashevskaya[3], R.D. Brown[1], E.V. Petrotchenko[1]*, A. Sickmann[3], C.H. Borchers[1, 4]

[1] *University of Victoria – Genome British Columbia Protein Centre, University of Victoria, Vancouver Island Technology Park, Victoria, Canada*

[2] *Center for Proteomics and Metabolomics, Leiden University, Leiden, Netherlands*

[3] *Leibniz Institut für Analytische Wissenschaften – e.V., Dortmund, Germany*

[4] *Dept of Biochemistry and Microbiology, University of Victoria, Victoria, Canada*

** e-mail: jenyavp@gmail.com*

Key words: proteome, protein structure, crosslinking analysis, peptides

Proteome-wide crosslinking analysis has the potential for the structural characterization of proteins in their natural cellular and tissue environments. The technique is able to capture weakly interacting and transient complexes, and can provide structural information on protein-protein interaction interfaces and conformations. The analysis is performed by chemical crosslinking proteins *in situ*, enzymatically digesting the crosslinked proteins, and identifying the crosslinked peptides using MS. Although the process is fairly straightforward, for proteome-wide applications the approach is dramatically complicated by the combinatorial nature of crosslinked peptides, requiring the identification of the interacting proteins based on a single peptide from each protein in a crosslinked peptide pair. Here, we describe the complete experimental and computational workflow using the isotopically-coded affinity-enrichable CID-cleavable crosslinker CBDPS.

DOI 10.18699/BGRSSB-2018-085

Statistical analysis of macromolecular B values

R. Masmaliyeva[1]*, G.N. Murshudov[2]
[1] *Institute of Molecular Biology and Biotechnologies ANAS, Baku, Azerbaijan*
[2] *MRC Laboratory of Molecular Biology, Cambridge, UK*
** e-mail: rmasmaliyeva@gmail.com*

Key words: bayes statistics, validation, atom displacement parameters, refinement

Motivation and Aim: The parameters of macromolecular atomic models consist of positional, occupancy and temperature factors – B values for each atom. B values reflect the combined information about uncertainties of atoms, their mobility as well the resolution and noise in the experimental scattering data. There have been number of research articles directed towards the validation of positional parameters of atoms [1]. However, there are only few research papers covering the validation of atomic B values. Analysis of B value distributions as well as their differences for neighbouring atoms can shed light into the validity of atomic models and even may show locations of misinterpretation of the experimental data. Since B values are proportional to the variances of positional parameters, their distribution should obey the distribution of variances. In Bayesian statistics it is common to use the inverse gamma distribution as the conjugate prior for the variance parameters of normal distribution. Therefore we expect that the probability distribution of B values can be described as a shifted inverse gamma distribution. We add shifts to the inverse gamma distribution to account for possible mishandling of the experimental data by under- or over-blurring.
Methods and Algorithms: The inverse gamma distribution is a continuous univariate distribution. Parameters are: shape parameter (α) and scale parameter (β) and for shifted inverse gamma distribution there is an additional parameter – B_0 that may be equal to the minimum B value. We used Maximum likelihood estimation to evaluate the parameters of the inverse gamma distribution for each macromolecule from the PDB. For optimization of the likelihood function we used the modified Newton-Raphson method with Fisher information matrix.
Results: Our analyses confirmed that for most of the PDB [2] entries the inverse gamma distribution describes well the distribution of temperature factors. In this work analysis of differences between neighbouring B values in the PDB entries will also be presented.
Conclusion: Estimating distributions of temperature factors of protein molecules can be used as a validation tool for B values. We also plan to use the derived probability distribution of B values for designing new restraints for refinement thus increase reliability of derived parameters. Although for parameter estimation we use high-resolution macromolecular structures we expect that these restraints will be useful for medium and low-resolution structure refinement.
Acknowledgements: Supported by the Presidium of Azerbaijan National Academy of Sciences, grant of decree No. 5/9 dated on 15.03.2017.

References

1. Read R.J., Adams P.D., Arendall W.B. et al. (2011) A New Generation of Crystallographic Validation Tools for the Protein Data Bank. Structure. 19:1395-1412.
2. Berman H.M., Battistuz T., Bhat T.N. et al. (2002) The Protein Data Bank. Acta Crystallogr. 58:899-907.

DOI 10.18699/BGRSSB-2018-086

The secretome of *Serratia marcescens* SM6 under oxidative stress conditions

L.E. Matrosova[1]*, I.V. Khilyas[1], L.M. Bogomolnaya[1, 2]
[1] *Institute of Fundamental Medicine and Biology KFU, Kazan, Russia*
[2] *Texas A&M University Health Science Center, Bryan, Texas, USA*
** e-mail: M.Lilia.Evg@yandex.ru*

Key words: efflux system, MacAB, secretome, *Serratia*

Motivation and Aim: Multidrug efflux pumps representing the membrane protein complexes play a significant role in a drug resistance and natural physiology of pathogenic bacteria such as *S. marcescens* [1]. Four different classes of efflux pumps were identified in the genome of *S. marcescens*. Clear understanding of the natural functions of efflux pumps is required for the development of new generation of antimicrobials to combat bacterial drug resistance. The objective of this work was to study the antioxidant effect of extracellular metabolites produced by environmental strain *S. marcescens* SM6 in response of oxidative stress and secreted in efflux pump MacAB-dependent fashion.
Methods and Algorithms: All experiments were performed in M9 medium at 37 °C in the presence/absence of hydrogen peroxide. Cell-free growth media used for cultivation of *S. marcescens* wild type SM6 or Δ*macAB* strains were passed through Supelco Discovery DSC-18 solid phase extraction C18 cartridges. Purification and fraction collection of metabolites was performed on an Acclaim® PolarAdvantage II (PA2) C18 reverse-phase column using UltiMate 3000 UHPLC system (Thermo Scientific, Dionex, USA).
Results: Cultivation of Δ*macAB S. marcescens* mutant strain in the presence of a hydrogen peroxide led to a drastic drop in bacterial viability. Interestingly, extracellular metabolites produced by *S. marcescens* SM6 wild type under oxidative stress conditions had a protective effect on the growth of *S. marcescens* Δ*macAB* in the presence of peroxide. HPLC analysis detected several unique peaks with retention time of 12.360, 12.757, 19.237, 23.317 и 24.797 min, respectively, that were present only in the extracts from conditioned media used for growth of *S. marcescens* wild type but not the Δ*macAB* mutant strain. Addition of these fractions to the peroxide-containing media rescued growth of Δ*macAB* mutant strain.
Conclusion: Results of this study show that hydrogen peroxide affected the composition of secreted metabolites of *S.marcescens* SM6. Particularly, HPLC analysis determined the presence of several peaks in the samples of the conditioned media used for growth of wild type strain SM6 under stress conditions.
Acknowledgements: This work was supported by the Russian Science Foundation project 16-14-10200 and performed in accordance with the Russian Government Program of Competitive Growth of Kazan Federal University.

References

1. Mahlen S.D. (2011) Serratia infections: from military experiments to current practice. Clin Microbiol Rev. 24(4):755-791.
2. Nishino K. (2005) Bacterial multidrug exporters: insights into acquisition of multidrug resistance. Science [online publication]. 2005. http://www.sciencemag.org/fea$ture/data/prizes/ge/2004/nishino.dlt

DOI 10.18699/BGRSSB-2018-087

Development of quantitative MRM assays for the measurement of 3,000 proteins across 20 mouse tissues

S.A. Michaud[1], A.L. Palmer[1], N.J.T. Sinclair[1], H. Pětrošová[1], I. Feldmann[2], Y. Mohammed[1, 3], A. Sickmann[2], C.H. Borchers[1, 4, 5, 6]

[1] *University of Victoria – Genome British Columbia Proteomics Centre, Victoria, Canada*
[2] *Leibniz-Institut für Analytische Wissenschaften – e.V., Dortmund, Germany North Rhine-Westphalia*
[3] *Center for Proteomics and Metabolomics, Leiden University Medical Center, Leiden, Netherlands South Holland*
[4] *University of Victoria, Department of Biochemistry and Microbiology, Petch Building Room 207, 3800 Finnerty Rd., Victoria, BC, V8P 5C2, Canada*
[5] *Gerald Bronfman Department of Oncology, Jewish General Hospital, McGill University, 3755 Cote-Ste-Catherine Road, Montreal, QC H3T 1E2, Canada*
[6] *Jewish General Hospital Proteomics Laboratory, McGill University, Lady Davis Institute, 3755 Chemin de la Côte-Sainte-Catherine, Montréal, QC H3T 1E2, Canada*
** e-mail: christoph@proteincentre.com*

Key words: proteomics, protein expression mouse, mass spectrometry, assay

Detailed characterization of protein expression in mouse tissues is challenging to perform due to the lack of available tools for rapid and robust quantitation. To simplify this process, we are developing highly multiplexed panels of assays to quantify 3,000 unique proteins across 20 mouse tissues by MRM mass spectrometry. Our method requires minimal sample pre-processing and uses stable isotope-labeled standard (SIS) peptides for precise and sensitive quantitation. Assay development involves determination of the LLOQ, linear range, and assay variability for each peptide. This rigorous characterization ensures the quality of each assay. Ultimately these assays will provide the first steps towards large scale, multi-tissue quantitation, and will allow researchers to gain an improved understanding of complex biological processes and diseases.

DOI 10.18699/BGRSSB-2018-088

Effect of salt bridges rupture on the activity and thermostability of bovine chymosin

E.S. Novoselova[1]*, A.P. Rudometov[2], A.V. Kriger[3], V.V. Elchaninov[3]
[1] *Novosibirsk State University, Novosibirsk, Russia*
[2] *State Research Center of Virology and Biotechnology "Vector" Rospotrebnadzor, Koltsovo, Russia*
[3] *Federal Altai Scientific Centre of Agro-BioTechnologies, Barnaul, Russia*
** e-mail: batmaiden@gmail.com*

Key words: salt bridges, chymosin, thermostability, molecular dynamics

Motivation and Aim: It is known that salt bridges are one of the factors influencing the thermostability of bovine chymosin [1]. Salt bridges result from the interaction of positively and negatively charged side radicals of amino acid residues on the surface of the protein. This work aimed at creating the functional mutants of chymosin with reduced thermostability. Chymosin is an enzyme that catalyzes the hydrolytic cleavage of the peptide bond between phenylalanine and methionine of the main protein of kappa-casein milk with the formation of a milk clot.

Methods and Algorithms: To determine salt bridges, molecular dynamics in the GROMACS 5 program at various temperatures – 300, 350, and 400 K was used. The trajectories of 5 ns duration were obtained for each temperature. The program VMD 1.9.3 was used for processing the trajectories. Mutant variants of chymosin were obtained using site-directed mutagenesis. The initial and mutant variants of chymosin were obtained using the *E. coli* pET21a vector system. The activity of the enzyme was determined by measuring the milk-coagulating activity on a standardized dry milk substrate. Thermostability was determined by the change in enzyme activity with increasing temperature.

Results: 10 salt bridges with persistence of no less than 50 % were detected in chymosin at least one of the investigated temperatures. Four of them were selected for the destruction effect to be investigated experimentally.

For the two mutant variants E198K and E363Q, the milk-binding activity was too weak to be further investigated. For the other two mutant variants D156V and D209N, the milk-binding activity was lower than that of the wild-type, but it was noticeable. Thermostability of these two variants proved to be higher than that of the wild-type chymosin.

Conclusion: The mutant variants of chymosin with disrupted salt bridges were investigated. Such point mutations turned out to significantly affect the milk-swilling activity. Surprisingly, an increase in thermal stability was obtained in two cases instead of the expected reduction in thermal stability. These results demonstrate how great could be the effect of a single mutation on the enzyme activity and it can encourage the further research.

Acknowledgements: The reported study was funded by RFBR according to the research project No. 17-44-220540 Supported by Russian Ministry of Science and Education under 5-100 Excellence Programme.

References

1. Karshikoff et al. (2015) Rigidity versus flexibility: the dilemma of understanding protein thermal stability. FEBS J. 282(20):3899-917.

DOI 10.18699/BGRSSB-2018-089

A number of blood biochemical parameters and endothelium-associated urine proteins of healthy people at head down bed rest

L.Kh. Pastushkova[1]*, D.N. Kashirina[1], A.G. Brzhozovsky[1], A.M. Nosovsky[1], A.S. Kononikhin[1, 2], V.A. Ivanisenko[3], E.S. Tiys[3], E.N. Nikolaev[2], I.M. Larina[1]

[1] *Institute for Biomedical Problems – Russian Federation State Scientic Research Center, RAS, Moscow, Russia*

[2] *Institute of Energetic Problems of Chemical Physics RAS, Moscow, Russia*

[3] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

** e-mail: lpastushkova@mail.ru*

Key words: urine, chromatography-mass spectrometry, head-down bed rest, cardiovascular system, endothelial dysfunction

Motivation and Aim: Head down bed rest (HDBR), a strict bed rest with negative tilt of the head-end of the bed, triggers adaptive mechanisms of cardiovascular, endocrine, central and peripheral nervous systems. With regard to cardiovascular system, HDBR induces hypovolemia and cardiovascular deconditioning with alteration in vascular functions of various body regions, that is accompanied by endothelial dysfunction. Carbohydrate metabolism is also undergoing changes: insulin secretion increases and glucose tolerance is impaired. It is known that endothelial dysfunction is the first step in the development of atherosclerosis associated with insulin resistance syndrome. Therefore, we aimed to reveal the relationships between endothelium-associated proteins and the biochemical variables related to carbohydrate metabolism and its regulation.

Results: Mass spectrometry-based proteomics was employed to analyze urine samples from 8 healthy volunteers who remained at the bed rest study for 21 days, with an angle of inclination relative to the longitudinal axis of the body horizontal position – 6°. ANDSystem software which builds associative networks was used to identify urinary proteins functionally related to the endothelium. We identified 7 endothelium-related biological processes, directly linked to 13 urine proteins. Analysis of correlations with biochemical variables revealed a positive correlation between fasting blood glucose and the next urine protens: albumin, CD44 antigen, endothelial protein C receptor, mucin-1, osteopontin, receptor tyrosine kinase. As well, we found a positive correlation between HOMA-insulin resistance index and urine protens: endothelial protein C receptor and syndecan-4.

Conclusion: These results might suggest the involvement of above-mentioned proteins in glucose metabolism and their effects on carbohydrate metabolism. Thus, proteomics methods allowed obtaining new data about changes induced by HDBR, and suggesting the possible mechanisms of these changes.

Acknowledgements: The work was supported by RFBR grant No. 18-34-00524. We express special thanks to Bruker Daltonics for financial and technical support.

DOI 10.18699/BGRSSB-2018-090

Nonthermal impact of terahertz (THz) radiation on living systems

S.E. Peltek[1]*, I.A. Meshcheryakova[1], T.N. Goryachkovskaya[1], E.V. Kiseleva[1], A.S. Rozanov[1], E.V. Demidova[1], K.V. Starostin[1], A.V. Bryanskaya[1], S.V. Sergeeva[1], P.S. Loshenova, D.Y. Oshchepkov[1], E.A. Demidov[1], G.L. Dianov[1], M.A. Logarkova[3], S.L. Kiselev[3], M. Timofeeva[2], N.A. Vinokurov[2], V.M. Popik[2], M.A. Scheglov[2]

[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

[2] *Budker Institute of Nuclear Physics SB RAS, Novosibirsk, Russia*

[3] *Vavilov Institute of General Genetics RAS, Moscow, Russia*

** e-mail: peltek@bionet.nsc.ru*

Key words: terahertz, *E. coli*, biosensor, proteomics, SILAC

An analysis was made of the reaction of mesophilic and extremophilic microorganisms to the non-thermal impact of terahertz radiation (THz) of the Novosibirsk free-electron laser. Transcriptome analysis of irradiated and control cells revealed the activation of genes controlling of adhesive structure of the cell wall, transport through the cell membrane and the metabolism of organic substrates. Additional, by electron microscopic method shows, shown that irradiated *E. coli* cells with THz increases the unevenness of the positioning of the pilus on the surface of the bacteria, and also causes the pilus to stick together and form their 2–3-layer bundles.

In this study we also performed a differential proteomic analysis of the total microorganisms soluble protein fraction after exposure to THz radiation. Based on results of the analysis of the proteome of irradiated and control cells the biosensor *E. coli*/glnA-gfp which react to impact to THz was created. Using this and other biosensors it has been shown that systems sensitive to oxidative stress and the presence of metal ions react to the action of THz and the system sensitive to the presence of antibiotics do not respond.

The results of proteomic analysis of the response to the influence of THz on extremophilic bacteria Halorubrum indicate a change in the expression of genes that control the regulation of the components of the cell wall.

Analysis of the reaction of human cells to the nonthermal exposure of terahertz radiation showed that in the genetic apparatus of human embryonic stem cells, mutations also do not occur. Changes in the transcription activity of genes are primarily due to the genetic apparatus of the mitochondria.

The SILAC method shows that the expression of mitochondrial proteins involved in the synthesis of ATP and the activity of the gene network controlling the synthesis of the myelin protein increases in the human THz-exposed fibroblasts.

DOI 10.18699/BGRSSB-2018-091

The size of the Human Proteome: how many protein species are detectable today?

E. Ponomarenko*, A. Kopylov, V. Zgoda, E. Poverennaya, E. Ilgisonis, A. Lisitsa, S. Naryzhny, E. Petrenko, S. Radko, A. Archakov
Institute of Biomedical Chemistry, Moscow, Russia
** e-mail: 2463731@gmail.com*

Key words: Human Proteome Project, proteoforms, proteome size, human chromosome 18

Motivation and Aim: The size of the human proteome, which is determined by the number of protein species («*width*») and the number of copies of an individual proteoform in a biosample («*depth*»), is still unknown [1, 2]. Given the limitations of sensitivity of analytical methods and the absence of amplification reaction in proteomics, it remains a challenge to define the part of the proteome that can be observed experimentally.
Methods and Algorithms: Here, meta-analysis of neXtProt knowledge base is proposed for theoretical prediction of the number of different proteoforms arising from alternative splicing, single amino acid polymorphisms and post-translational modifications. Experimental part was performed using targeted MS for chromosome 18 encoded proteins, along with the estimation of copy numbers in plasma, liver, and HepG2 cell line. The proposed approaches for estimation of proteome *width* and *depth* were validated using UPS1 and UPS2 protein calibration standards provided by Sigma Aldrich and consisting of 48 proteins present in same (UPS1) and different (UPS2) concentration.
Results: A range of 0.55-7.14 millions of protein species (proteoforms) in the human body was estimated using different methods of calculation based on the average number of variations per gene from neXtProt. In particular, 275 protein-coding genes predicted for human chromosome 18 could potentially encode about 8 to 18 thousand of proteoforms. In total, proteins were detected and measured for only approximately 30 % of the predicted protein-coding genes in selected types of biomaterials. When using UPS1 and UPS2 standards we found that shotgun LC-MS/MS analysis allows to identify only a half of proteins presented in the sample in a pure solution. In comparison, there was just four proteins we were not able to detect by targeted MS method. The number of undetected proteins increases when we add a matrix, such as human blood plasma or E-coli protein.
Conclusion: Taking chromosome 18 as an example, the size of the human proteome was predicted based on NeXtProt data. We found that biological matrix significantly affects the list of detected proteins and obviously affects proteome *width* and *depth*: MS-signals from presenting proteins may be lost or new MS-signals may appear resulting in false-positive results.
Acknowledgements: El. P., Ek. P. and Ek. I. acknowledge the Leading Scientific School of Prof. Andrey Lisitsa (No. NSh6313.2018.4).

References

1. Ponomarenko E., Poverennaya E., Ilgisonis E. et al. (2016) The Size of the Human Proteome: The Width and Depth. Int J Anal Chem. 2016:7436849.
2. Aebersold R., Agar J.N., Amster I.J. et al. (2018) How many human proteoforms are there? Nat Chem Biol. Feb 14;14(3):206-214.

DOI 10.18699/BGRSSB-2018-092

The good, the bad, the aberrant: the role of prevailing splice-forms in proteomic studies

E. Poverennaya*, O. Kiseleva, M. Pyatnitskiy, E. Ilgisonis, E. Ponomarenko, A. Lisitsa
Institute of Biomedical Chemistry, Moscow, Russia
** e-mail: k.poverennaya@gmail.com*

Key words: Human Proteome Project, proteoforms, alternative splicing, splice-forms, RNA-seq, mass-spectrometry, human chromosome 18

Motivation and Aim: Thousands of genes expressed in a particular cell determine the functionality of the cell. Each step in the flow from DNA via RNA and finally to protein supplies the cell with a potential control point for self-regulation of its functions. This control is implemented by alternation of the amount and type of proteoforms the cell generates by "fine adjustment" of different aberrations. One of the fundamental resources of such aberrations is alternative splicing. The difference between splice-forms is defined by a combination of exons. While the majority of proteomic studies are devoted to examination of canonical sequences, we focused on more sophisticated task of analysis of aberrant forms. It is non trivial because of high homology between the canonical and splice-forms: the selection of proteotypic peptide which uniquely characterizes the splice-form and does not map on the canonic form is often challenging.
Methods and Algorithms: RNAseq data of normal liver tissue, lung, bladder, kidney, stomach and esophagus, publicly available in SRA database, were analyzed to identify the prevailing form (canonical or aberrant) by level of expression for each gene. The transcriptomic data were processed by PPLine software. Then transcripts were translated into amino acid sequences, which were *in silico* cleaved by trypsin. The frequency of detection of proteotypic peptides in proteomic experiments was calculated according to the GPMdb.
Results: According to the results of the transcriptomic analysis of six tissues, splice-forms prevail over canonical forms in more than 2.7 thousand cases. Notably, that in 1.6 thousand cases gene products were presented only by splice-forms. We analyzed the prevalence of splice-forms on the example of human chromosome 18, key chromosome of the Russian part of Chromosome-Centric Human Proteome Project. For 52 of 275 genes of chr18 we revealed, that 54 splice-forms prevail over canonical forms in 6 examined types of tissues: 24 forms were found in all tissues, two were detected only in esophagus, one – in stomach, bladder, and lung each, while in kidney tissue there was no specific splice-form.
According to the UniProt database, there are ca. 550 thousand of proteotypic peptides in human proteome: 50 % of them correspond both to canonical and splice-forms, 40 % (223 thousand) are specific only for canonical form, and 4 % are specific exclusively for splice-forms.
Conclusion: Knowledge about prevailing form is crucial on every step of a proteomic study – from experiment design to data interpretation.
Acknowledgements: El. P., Ol. K,. Ek. P. and Ek. I. acknowledge the Leading Scientific School of Prof. Andrey Lisitsa (No. NSh6313.2018.4).

DOI 10.18699/BGRSSB-2018-093

Milk exosomes: isolation, proteins, and nucleic acids

L. Purvinsh[1, 2]*, S. Sedykh[1, 2], G. Nevinsky[2]
[1] *Novosibirsk State University, Novosibirsk, Russia*
[2] *Institute of Molecular Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia*
** e-mail: ladapurvinsh13@gmail.com*

Key words: exosomes, proteins, microRNA

Motivation and Aim: Exosomes are extracellular membrane vesicles with the diameter of 40-100 nm. They are secreted by cells and detected in various biological fluids. Over the past decade, the role of exosomes in many physiological and pathological processes in the body has become clear. The presence of proteins, peptides, DNA, mRNA and, especially, microRNA in their composition is shown, and the role of nucleic acids contained in exosomes in the development of diseases is clarified. The collected data of the exosome composition and functioning serve as the basis for the development of new non-invasive methods for diagnosing various diseases and means of targeted drug delivery in the body. The study of exosomes from milk is very important in the perspective of the development of new approaches to the exosome isolation from complex biological fluids containing a large number of proteins.
Methods and Algorithms: We have developed an approach that combines standard protocol of exosome isolation (ultrafiltration, ultracentrifugation) with additional gel-filtration. According to transmission electron microscopy, this technique allows obtaining homogeneous vesicles. Immunohistochemical staining using antibodies to tetraspanins – the main surface proteins of exosomes was confirmed that the obtained vesicles are exosomes. In addition, the same antibodies were used to create columns for affinity selection of exosomes.
Results: Analysis of protein composition in highly purified milk exosomes showed the presence of a small number of proteins. This indicates that some milk proteins, which were previously described as exosomal, are not part of the exosomes, but co-isolate with them [1]. For the analysis of nucleic acids, we obtained RNA samples at different stages of the exosome isolation, carried out reverse transcription and PCR in real time using primers to several main microRNAs of milk exosomes. Quantitative amplification of this microRNA can serve as an effective high-performance method for determining the content of exosomes in preparations.
Conclusion: The results obtained by us indicate that with the right technology of isolation, milk can serve as a promising source of exosomes necessary for their study and development of new diagnostic approaches and therapeutic drug delivery systems.

References

1. Sedykh S.E., Purvinish L.V., Monogarov A.S., Burkova E.E. et al. (2017) Purified horse milk exosomes contain an unpredictable small number of major proteins. Biochimie Open. 4:61-72.

DOI 10.18699/BGRSSB-2018-094

In silico prenylation predictions for human proteins as a novel class of naturally occurring post-translational peptides

J. Ravich[1], A. Randhawa[1], A. Bragin[2], B. Eisenhaber[3], F. Eisenhaber[3], E. Marakasova[1], A. Baranova[1, 4]

[1] *Center for the Study of Chronic Metabolic and Rare Diseases, School of Systems Biology, George Mason University, Manassas, USA*

[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

[3] *Bioinformatics Institute, Agency for Science, Technology and Research, Singapore*

[4] *Research Centre for Medical Genetics, Moscow, Russia*

** e-mail: jwravich@gmail.com*

Key words: prenylation, post-translational modification, human proteins

Prenylation is a form of post-translational modification in which isoprenoid lipids are covalently bonded to the C-terminus of newly synthesized proteins [1, 2]. This process enhances intracellular protein localization and alters activity of numerous proteins. It is believed that prenylation of proteins assists in their incorporation into human cell membranes [2, 3]. However, even though prenylation has been experimentally identified in bacteria [3], no experimental evidence supports prenylation of human proteins.

The research objective was to collect data of modified human proteins that supports the existence of prenylation in human proteins. Modified human proteins (775) were gathered, based on laboratory evidence, from the EROP-Moscow Database (http://erop.inbi.ras.ru/query-erop.php?submit=%C2%A0%C2%A0Query+EROP), and all of the available unmodified human proteins were collected from the NCBI Homo sapiens mRNA Protein Database (https://www.ncbi.nlm.nih.gov/refseq/). Homo sapiens proteases (6,425) from the CutDB Proteolytic Event Database (http://cutdb.burnham.org/) were applied to the NCBI data in order to generate in silico cut protein segments. Unique Python scripts were created and used to perform in silico bioinformatic experimentation and analysis, resulting in the synthesis of 69,706 protein segments from the NCBI data, in addition to the 775 already collected proteins from the EROP data.

These results were then submitted to PrePS (http://mendel.imp.ac.at/PrePS/) as a batch job to computationally identify human protein prenylation targets. Consequently, 35 peptides were positively predicted as human protein prenylation targets. The proteases, which are involved in peptide modificaton for prenylation, include cathepsin, bacterial collegenase, retropepsin (human T-cell leukemia virus), and eupitrilysin. Of these 4 proteases, bacterial collegenase, retropepsin (human T-cell leukemia virus), and eupitrilysin were identified as physiological proteases due to repetitive occurrence in the formation of prenylated peptides.

References

1. Casey P.J. (1992) Biochemistry of protein prenylation. Journal Lipid Research 33:1731-40.
2. Benetka W., Koranda M., Eisenhaber F. (2006) Protein Prenylation: An (Almost) Comprehensive Overview on Discovery History, Enzymology, and Significance in Physiology and Disease. Monatsh. Chem. 137:241-81.
3. Marakasova E.S., Akhmatova N.K., Amaya M., Eisenhaber B., Eisenhaber F. et al. (2013) Prenylation: from bacteria to eukaryotes. Mol Biol. 47:622-33.

DOI 10.18699/BGRSSB-2018-095

Dynamics of *S. cerevisiae* proteomic and transcriptomic response to changes in aeration conditions

A. Rozanov*, I. Meshcheryakova, S. Bannikova, D. Oshchepkov, G. Vasiliev, S. Peltek
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: sibiryak.n@gmail.com*

Key words: metabolic analysis, S.cerevisiae, RNAseq, proteomics

Motivation and Aim: The microorganisms widely used in biotechnologies to produce various substances, from simple primary metabolites such as alcohol to complex organic molecules and proteins with a specific set of modifications. Yeasts are the most widely used producers of various substances. Traditionally, yeast has been used for cultivation in anaerobic environments to produce alcohol. Currently, they are widely used in aerobic processes. Modern technological processes are carried out in large volumes and with high-density cultivation. Microorganisms are strained in such conditions because of the difficulties in mass transfer. The fluctuation of the oxygen concentration is the most significant. In the work, we study the dynamics of the transcriptom and proteome response S. cerevisiae to stop to stop aeration.
Methods and Algorithms: Scheme of the experiment: yeast was cultivated in a 5l Biostat b-dcu II Sartorius fermenter. Strain FY1679 was used. Medium: YNB + 2 % glucose, + required amino acids, 2.5l. Conditions: temperature 30 °C, mixing speed 800 rpm, air 2 liters per minute. The culture was grown until an optical density of OD_{600} = 0.8. In that time takin samples. The air supply was off and the nitrogen supply was on. Samples for analysis were taken at 5, 10, 20, 40 and 80 minutes. Further RNAseq and differentiated proteomic analysis of the samples were performed.
Results: According to proteome analysis, half of all visualized proteins were observed in concentration changes. a concentration of 3 % (15 of 500) proteins decreasing after the 10^{th} or 5^{th} minute after oxygen off, and 30 % (151 of 500) of proteins – increased ($p < 0.05$). The main changes in the protein content occur between the 10^{th} and 20^{th} minutes, which is undoubtedly a reflection of the metabolic rearrangements occurring at this time in culture. At the moment, 162 proteins have been identified. Также в настоящий момент идет обработка данных RNAseq. The results of bioinformatic processing and analysis results will be presented in the report.
Acknowledgements: The reported study was funded by ICG SB RAS No. 0324-2018-0017.

DOI 10.18699/BGRSSB-2018-096

Study of pathogenic features of stress-related disorders by proteomics methods

A. Seregin[1], L. Loginova[1], A. Letova[2]*, E. Dmitrieva[1], N. Karageorgiy[3], V. Rudnitckiy[1], L. Smirnova[1]

[1] *Mental Health Research Institute, Tomsk National Research Medical Center of the RAS, Tomsk, Russia*

[2] *Siberian State Medical University, Tomsk, Russia*

[3] *Tomsk State Pedagogical University, Tomsk, Russia*

** e-mail: anastasia.a.letova@gmail.com*

Key words: biomarkers, biological psychiatry, adjustment disorder, protein markers

Motivation and Aim: Stress-related disorders are the group of mental illnesses, combined conditions induced by acute or chronic stress. One of these disorders is the adjustment disorder – the pathological condition of subjective distress and emotional disorder, which make difficulties for social activity in adaptation period to significant changes in life. In the base of the pathogenesis of adjustment, disorder lays a classical theory of stress, but many aspects of its still cannot be explained. Proteomics methods could promote the solution of this problem and answer the question about laboratory diagnostic of some mental disorders.

Methods and Algorithms: The base of study is the comparative proteomic analysis of blood serum 10 healthy donors and 10 patients with adjustment disorder.Firstly, the serum was purified from six major proteins by affinity chromatography. Thereupon received protein mix was separated by 1D electrophoresis. After the sample processing, including trypsinolysis and extraction of the peptides, HPLC/mass-spectrometry LTQ Velos (Thermo Scientific) identified the proteins. The validity of results was checked by the non-parametric Fisher exact test with the Yates correction.

Results: According to the results of mass-spectrometry, the next proteins were obtained: Inter-alpha-trypsin inhibitor complex component III (93402 Da), Glucocorticoid receptor AF-1 specific elongation factor (46240 Da), Secretory actin-binding protein (16562 Da) and Hypothetical protein CAE 93899 (14466 Da). Three proteins are participants in modulating the immune response. According to the literature analysis, the secretory actin-binding protein is probably the main acting link in the function of the immune system [1, 2]. The function of the hypothetical protein is not established yet; the remaining proteins realize the adaptive response of the organism to stress.

Conclusion: Presented proteins glucocorticoid receptor AF-1 specific elongation factor and actin-binding protein could pretend to the role of adjustment disorder biomarkers, reflecting key features of pathogenesis this illness.

References

1. Adair J. (2009) Inter-α-trypsin Inhibitor Promotes Bronchial Epithelial Repair after Injury through Vitronectin Binding. J Biol Chem. 284(25):16922-16930.
2. Choi-Miura N.H. (2001) Quantitative measurement of the novel human plasma protein, IHRP, by sandwich ELISA. Biological Pharmaceutical Bulletin. 24(3):214-217.

DOI 10.18699/BGRSSB-2018-097

The analysis of blood proteins by ESI-mass spectrometry for endogenous mental disorders

A.A. Seregin[1]*, L.P. Smirnova[1], E.M. Dmitrieva[1], A.A. Letova[2], L.V. Loginova[1], G.G. Simutkin[1], S.A. Ivanova[1], V. Zgoda[3]

[1] *Mental Health Research Institute, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia*

[2] *Siberian State Medical University, Tomsk, Russia*

[3] *Institute of Biomedical Chemistry (IBMC)*

** e-mail:apocalips1991@mail.ru*

Key words: Bipolar disorder, protein markers, proteomics, blood serum

Motivation and Aim: Bipolar affective disorder (BD) refers to endogenous disorders and is similar to schizophrenia; it develops without the influence of external factors. However, their biological differences are not revealed. Nowadays proteomic studies of the blood composition in patients with BD via ESI-mass spectrometry are missing.

Methods and Algorithms: In this work, serum of 8 patients with BD and 5 healthy individuals were analyzed. The samples were subjected to affinity chromatography, 1D SDS page and in-gel trypsinolysis of the proteins. The mass-spectrometric analysis was carried out via mass spectrometer-Thermo Scientific LTQ Velos. Identification of proteins was carried out via Matrix Science resources. Statistical significance was assessed by Fisher exact test with the Yates correction, $p < 0.05$.

Results: As a result of the analysis, from 150 to 500 proteins were identified in every sample. Proteins, presenting in BD, are mostly involved in the processes of cell growth and maintenance, regulation of nucleic acid metabolism, followed by the immune response, protein metabolism and unknown processes. Including detected: Plectin – with a detection rate of healthy 20 %, in patients – 50 %; Vimentin – detection rate in patients – 75 %, of healthy is not found; Glutamate ionotropic receptor NMDA type subunit 1 (NMDAR) – frequency of detection in patients – 25 %, not detected in the control; Unconventional myosin-Va – detection rate in healthy 20 %, in patients – 75 %; SH3 and multiple ankyrin repeat domains protein 1 – frequency of detection in patients – 37.5 %, not detected in the control group.

Conclusion: SH3 and multiple ankyrin repeat domains protein 1, as a cytoskeleton protein, represents a great interest because it connects postsynaptic membrane receptors including metabotropic glutamate and NMDA receptors with the actin cytoskeleton. That several point mutations are revealed in SHANK gene encoding SH3 and multiple ankyrin repeat domains protein 1 presumably contributing to the development of autism spectrum disorders. Unconventional myosin-Va and other cytoskeletal proteins influence the regulation of NMDAR and are in the blood due to the violation of the permeability of blood-brain barrier.

Acknowledgements: Mass spectrometric analysis was carried out using the equipment of "Human Proteome" Core Facility of the Institute of Biomedical Chemistry (IBMC) Moscow which is supported by Ministry of Education and Science of the Russian Federation (unique project ID RFMEFI62117X0017) Support by Grant of RSF No. 18-15-00053.

DOI 10.18699/BGRSSB-2018-098

Cytotoxic proteins of human placenta

S. Skripkin[1]*, E. Burkova[2], G. Nevinsky[1, 2]
[1] *Novosibirsk State University, Novosibirsk, Russia*
[2] *Institute of Chemical Biology and Fundamental Medicine, Novosibirsk, Russia*
** e-mail: skripkin.stan@gmail.com*

Key words: human placenta, proteins of the placenta, protein complexes, inhibition of cell proliferation, apoptosis

Motivation and Aim: Numerous biological functions of the placenta are carried out by various oligomeric structures and complexes. The activity of protein complexes can significantly differ from the activity of individual proteins. Previously, a highly stable soluble protein complex of the human placenta was isolated, with a molecular mass reaching 1000 kDa. It was found that the complex includes human serum albumin, transferrin, IgG, annexin and other proteins. This complex exhibited DNA-hydrolyzing and catalase activity [1, 2]. The purpose of this work is to study proteins and protein complexes of the human placenta, to study their composition and properties.
Methods and Algorithms: The placentas of ten healthy women was used in the work after normal physiological parturition. The proteins of the placenta extract were separated by gel filtration on a column with Sepharose 4B sorbent. The protein fractions were examined using an MTT test and flow cytofluorometry with treatment with Annexin B and propidium iodide. SDS electrophoresis was used to evaluate the molecular masses of proteins.
Results: After gel filtration of protein extracts in all cases two protein peaks were observed. The MTT test of the obtained fractions showed that some fractions of the second peak of the gel filtration of the human placenta extract inhibit the proliferation of cells of both the cancer line MCF-7 and normal WI-38. Flow cytofluorometry with treatment with Annexin V and propidium iodide revealed that cytotoxically active fractions of the human placenta cause cell death by the mechanism of apoptosis. Using SDS electrophoresis, molecular weights of proteins of cytotoxic fractions of human placenta have been determined. In the main in the factions are proteins with molecular masses of 64.7, 51.2, 24.8 and 14.8 kDa, minor – 179.4, 44 and 33.1 kDa.
Conclusion: The study of proteins and protein complexes of the human placenta will allow a deeper understanding of the functional characteristics of certain proteins, as well as the mechanisms of the functioning of the placenta.
Acknowledgements: The work is supported by the project No. VI.62.1.5. The program of FSI SB RAS VI.62.1 (0309-2018-0003) and the RFBR grant No. 16-04-00609.

References

1. Burkova E.E., Dmitrenok P.S., Sedykh S.E., Buneva V.N., Soboleva S.E., Nevinsky G.A. (2014) Extremely Stable Soluble High Molecular Mass Multi-Protein Complex with DNase Activity in Human Placental Tissue. PLOS One. DOI 10.1371/journal.pone.0111234.
2. Burkova E.E., Nevinsky G.A. (2016) Very stable high molecular mass multi-protein complex from human placenta. FEBS. 283(1):211.

DOI 10.18699/BGRSSB-2018-099

Comparative proteomic analysis form patients with schizophrenia and bipolar disorder

L.P. Smirnova[1]*, E.M. Dmitrieva[1], A.A. Letova[2], A.A. Seregin[1], G.G. Simutkin[1], A.V. Semke[1], S.A. Ivanova[1]

[1] *Mental Health Research Institute, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia*

[2] *Siberian State Medical University, Tomsk, Russia*

** e-mail: lpsmirnova@yandex.ru*

Key words: schizophrenia, вipolar disorder, protein markers, proteomics, blood serum

Motivation and Aim: Schizophrenia and bipolar disorder (BD) are the most important mental disorders for social life. They represent a heterogeneous group of endogeneous mental disorders with unclarified ethiology and pathophysiological mechanisms at present. Since diagnostics of mental disorders is based only on clinical symptoms, there is a necessity in development of additional methods of biochemical/paraclinical diagnostics. The search for blood based biomarkers which may be used for diagnostics and prognosis of therapy efficacy are very important presently. Purpose of study is revealing significant differences in serum proteomes in schizophrenia, BD and matched healthy controls.

Methods and Algorithms: The sample preparation included affinity removing of six major proteins, separation by 1D electrophoresis, in-gel tryptic hydrolysis, and LC-MS/MS peptide analysis using LTQ Orbitrap Velos mass spectrometer.

Results: When comparing proteome profiles, different unique protein sets were revealed (absent in other groups): 22 proteins typical for schizophrenia, and 20 – for BD. Protein set in schizophrenia was mostly associated with nucleic acid and protein metbolism, immune response, cell communication, and cell growth and maintenance. Protein set in BD was mostly associated with cell growth and maintenance, nucleic acid metabolism regulation, immune response, protein metabolism, transport and cell communication. Concentrations of ankyrin repeat domain-containing protein 12 (ANKRD12), coagulation factor XIII, and cadherin 5 in serum samples were determined by ELISA. Significant difference between three groups was revealed in ANKRD12 concentration ($p = 0.02$), with maximum elevation of ANKRD12 concentration (mediana level) in schizophrenia followed by BD. Cadherin 5 concentration differed significantly ($p = 0.035$) between schizophrenic patients with prevailing positive symptoms and those with prevailig negative symptoms.

Conclusion: Our results are presumably useful for discovering the new pathways involved in endogeneous psychotic disorders.

Acknowledgements: Mass spectrometric analysis was carried out using the equipment of “Human Proteome” Core Facility of the Institute of Biomedical Chemistry (IBMC) Moscow which is supported by Ministry of Education and Science of the Russian Federation (unique project ID RFMEFI62117X0017). Support by Grant of RSF No. 18-15-00053.

DOI 10.18699/BGRSSB-2018-100

Molecular design of a new class of inhibitors for ion channel of influenza protein

Y.N. Vorobjev, O.S. Fedorova
Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia
Novosibirsk State University, Novosibirsk, Russia
** e-mail: ynvorob@niboch.nsc.ru*

Key words: M2 protein, ionization, drug binding, histidine ionization

Motivation: A design of novel anti-influenza drugs is a task of great importance due to a capability of influenza viruses to infect fast a large human population by occasional cross of inter-species barriers and to rapid mutate.
Methods: Transport of H^+ ion through ion channel of protein M2 of cell membrane can be blocked by drug molecule bound inside of ion channel. A new class of molecular blockers is suggested. Binding mode and binding energies are calculated for a set of novel molecular structures constructed on the base of diazobicyclooctane.
Results: A set of molecular structures as a derivatives of leading compound is suggested. Binding modes and energies of binding are calculated by method of hierarchical blind docking [1]. A molecular structure of an optimal molecular blocker is suggested.
Acknowledgements: The work is supported by grant RFFI No. 18-04-00005 and by Russian State funded project program (No. VI.57.1.2, 0309-2018-0001).

References

1. Vorobjev Y.N. (2010) J Comput Chem. 31:1080-1092.

DOI 10.18699/BGRSSB-2018-101

Top-down venomics: a *de novo* sequencing approach

K. Vyatkina[1, 2, 3, 4]*, D. Petras[5]
[1] *Saint Petersburg State University, Saint Petersburg, Russia*
[2] *Saint Petersburg Academic University of the Russian Academy of Sciences, Saint Petersburg, Russia*
[3] *ITMO University, Saint Petersburg, Russia*
[4] *Saint Petersburg State Electrotechnical University "LETI", Saint Petersburg, Russia*
[5] *University of California, San Diego, CA, USA*
** e-mail: k.vyatkina@spbu.ru*

Key words: proteomics, top-down mass spectrometry, snake venomics, toxins, *de novo* sequencing

Motivation and Aim: Understanding of the composition of snake venoms can significantly contribute to drug discovery research, and is important for the development of effective antidotes. However, often the lack of comprehensive genome or transcriptome databases, as well as alternative splicing events and PTMs substantially complicate their analysis [1, 2]. The aim of this research is to develop an algorithmic framework for identifying new proteoforms of known toxins from top-down mass spectrometry through *de novo* sequencing. Efficiency of the proposed approach is illustrated for the venoms of green and black mamba.

Methods and Algorithms: The venom samples of green and black mamba were reduced with TCEP and analyzed by LC-MS/MS using an Agilent 1260 HPLC system coupled to a Thermo Orbirtap LTQ XL mass spectrometer, as described in [2]. The obtained top-down MS/MS spectra were processed with a modified version of the Twister *de novo* sequencing algorithm [3] to derive a number of highly accurate sequence tags of length 3. Those 3-tags were applied to match their underlying spectra against the 157 mamba toxin sequences available in the NCBI database, with a goal of identifying novel proteoforms of the known venom proteins. For each suggested identification, the spectra witnessing for it were annotated against the putative proteoform sequence.

Results: For the green mamba and black mamba venom samples, putative novel proteoforms were proposed for 17 and 27 toxins from the NCBI database, respectively. Upon a more thorough examination, appearance of 6 and 8 out of those, respectively, was attributed to misinterpretation of common PTMs. Among the remaining suggestions, 8 and 10, respectively, were confidently confirmed by the annotated mass spectra. A decision on the remaining hypotheses could presumably be made from native mass spectrometry of the venom samples.

Conclusion: The proposed technique for analyzing toxin families based on *de novo* sequencing of top-down tandem mass spectra has proven to be a handy tool for studying snake venoms.

Acknowledgements: The first author was supported by the RFBR (grant No. 16-54-21006).

References

1. Petras D. (2015) Venom proteomics of Indonesian king cobra, Ophiophagus hannah: Integrating top-down and bottom-up approaches. J. Proteome Res. 14(6):2539-2556.
2. Petras D. et al. (2016) Top-down venomics of the East African green mamba, Dedndroaspis angusticeps, and the black mamba, Dendroaspis polylepis, highlight the complexity of their toxin arsenals. J. Proteomics. 146:148-164.
3. Vyatkina K. et al. (2015) De novo sequencing of peptides from top-down tandem mass spectra. J. Proteome Res. 14(11):4450-4462.

SYSTEMS COMPUTATIONAL BIOLOGY

DOI 10.18699/BGRSSB-2018-102

Methane utilization in *Methylomicrobium alcaliphilum* 20Z^R: new routes of C_1-metabolism

I.R. Akberdin[1, 2, 3]*, M. Thompson[1], R. Hamilton[1], D. Collins[1], M.G. Kalyuzhnaya[1, 4]
[1] *Department of Biology, San Diego State University, San Diego, USA*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[3] *Novosibirsk State University, Novosibirsk, Russia*
[4] *Viral Information Institute, San Diego State University, San Diego, USA*
** e-mail: akberdin@bionet.nsc.ru*

Key words: methane utilization, *Methylomicrobium alcaliphilum* 20Z^R, genome-scale model

Motivation and Aim: Biological methane utilization is one of the main sink of the greenhouse gas in nature and it represents an attractive platform for production of fuels and value-added chemicals. A number of methanotrophic bacteria have been isolated, but only a few of them, including *Methylomicrobium alcaliphilum* 20Z^R (20Z^R), have been established as promising and efficient catalysts for industrial applications. Robust growth, complete and expert-annotated genomic information, and a large set of genetic tools make 20Z^R one of the best systems for understanding and exploiting biological methane oxidation.

Methods and Algorithms: We applied a set of systems-level approaches for comprehensive investigation of the methane utilization network in 20Z^R. A genome-scale metabolic model of 20Z^R was constructed and refined using enzymatic, metabolic, proteomics, and gene expression data collected under different environmental conditions.

Results: The model simulations demonstrated that a significant portion of consumed methane, a quarter of consumed carbon, is used for non-growth-associated energy maintenance. Developed genome-scale model has highlighted the dynamic behavior of methane oxidation machinery and indicated the necessity of an additional constraint on the O_2 consumption rate to correctly reproduce experimentally observed parameters (growth rate and corresponding yields).

Conclusion: The flux balance analysis of the model combined with global, non-targeted metabolomic profiling and enzymatic assays highlighted the importance of the substitution of ATP-linked steps with PPi-dependent reactions and supported the presence of a carbon shunt from acetyl-CoA to the pentose-phosphate pathway and highly branched TCA cycle [1].

Acknowledgements: The study was financially supported by DOE under FOA DE-FOA-0001085 and by NSF-CBET award 1605031.

References

1. Akberdin I.R. et al. (2018) Methane utilization in Methylomicrobium alcaliphilum 20ZR: a systems approach. Scientific reports. 8(1):2512.

DOI 10.18699/BGRSSB-2018-103

CRISPR-Cas regulation: a systems biology approach

M.J. Djordjevic[1]*, A. Rodic[1], B. Blagojevic[2], M.R. Djordjevic[2], K. Severinov[3]
[1] *Faculty of Biology, Institute of Physiology and Biochemistry, University of Belgrade, Serbia*
[2] *Institute of Physics Belgrade, University of Belgrade, Serbia*
[3] *Skolkovo Institute of Science and Technology, Moscow, Russia*
** e-mail: dmarko@bio.bg.ac.rs*

Key words: computational systems biology, CRISPR-Cas, small RNAs, gene expression regulation

Motivation and Aim: Bacterial immune systems (CRISPR-Cas and restriction-modification systems) defend bacterial cells against invasion by viruses or plasmids. CRISPR array is transcribed as a long transcript (pre-crRNA), which is processed by Cas proteins to small interfering RNAs (crRNAs). CRISPR-Cas is typically silent under normal conditions, and one of the main questions in understanding CRISPR-Cas functioning is how this normally silent system is induced. Previous work (by us and others) shows that the system is regulated both at transcriptional and post-transcriptional (pre-crRNA processing) level.
Approach and Methods: To understand mechanism of CRISPR-Cas induction, we use a systems biology approach, combining computational modeling with bioichemical experiments, and first available *in-vivo* measurements of molecule dynamics in bacterial immune systems (done by Konstantin Severinov lab, Skoltech). Moreover, we exploit that CRISPR-Cas and more rudimentary restriction-modification (R-M) systems, likely exhibit similar constraints in their dynamical response, so that modeling better characterized R-M systems can aid understanding CRISPR-Cas. Computational modeling of CRISPR-Cas and R-M system induction/establishment is based on thermodynamical modeling of transcription regulation, and on modeling dynamics of the relevant molecular species (RNA and proteins). Wild-type systems are *in-silico* perturbed to assess how key regulatory features contribute to their dynamics.
Results: We show that computational modeling can reasonably explain both *in-vitro* measurements of transcription regulation, and *in-vivo* measurements of protein dynamics in R-M systems. We furthermore explore R-M systems with different architectures, and find that their (otherwise diverse) regulation can be explained in terms of few simple design principles. For CRISPR-Cas we show that both of the key system features (high cooperativity in transcription regulation and fast pre-crRNA processing) are responsible for a fast (switch-like) transition of the system from "OFF" to "ON" state, which we also obtain for R-M systems, and which can be explained in terms of efficiently protecting the host cell. Additionally, fast pre-crRNA processing also leads to a delay in crRNA generation, similarly to the delay in restriction enzyme synthesis during R-M establishment, and which may be related with evading auto-immune response. Finally, we find that the cooperative transcription regulation qualitatively leads to a cross-over to the regime where at higher pre-crRNA processing rates crRNA generation approaches the limit of (an infinitely abrupt) system induction.
Conclusion: We propose that few simple design principles may be behind regulation of mechanistically diverse bacterial immune systems, which is likely related with a highly efficient host cell protection, while at the same time evading autoimmunity.
Acknowledgement: Support by MESTD of Serbia OI173052 and SNSF IZ73Z0_152297.

DOI 10.18699/BGRSSB-2018-104

Patterns and models of flowering of some Campanulaceae Juss. species

E. Fomin[1]*, T. Fomina[2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Central Siberian Botanical Garden SB RAS, Novosibirsk, Russia*
** e-mail: fomin@bionet.nsc.ru*

Key words: flowering models, Campanulaceae

Motivation and Aim: Flowering is the most important event in the seasonal cycle of plant development, the realization of which largely determines reproductive success. This process can be studied at different levels of organization – organ, individual, population, phytocenotic, each of which is characterized by its own set of tasks and research methods. There are a number of prognostic models of phenological development, including flowering, for various taxonomic and biomorphological groups of plants [1, 2]. However, most of the existing models describe the flowering process at the population level, while insufficient attention is paid to modeling individual flowering. The purpose of this work is to study the flowering patterns of a number of species of the Campanulaceae family and to construct a computer model suitable for analysis and prediction of their decorativeness.

Methods and Algorithms: The subjects of the study were representatives of the family Campanulaceae Juss. and a large-bellied large-billed Platycodon grandiflorus (Jacq.) A. DC. from the collection of ornamental plants of the natural flora of the Central Siberian Botanical Garden of the SB RAS (Novosibirsk). The investigated species are characterized by a summer-brightening phenorhythmotype, with summer flowering dates for Campanula (late June to mid-July) and late-summer for Platycodon grandiflorus (second half of July). Phenological observations were carried out on shoots of each species during the entire flowering period. Observation data were statistically processed, repeatability in shoot structure and stable relationships between the position of flowers on them and the time of their disclosure were revealed.

Results: A stochastic structural dynamic model of flowering of shoots was constructed and implemented in the form of a computer program. It allows to predict the structure of the shoot and the phase of development of its generative organs for any date from the beginning of flowering. Three different flowering patterns were found: unimodular (C. bononiensis), bimodular (C. sarmatica) and a bimodular one with a strongly protracted phase of the beginning of flowering (P. grandiflorus). The reasons for the formation of these patterns were analyzed.

References

1. Normand F., Habib R., Chadoeuf J. (2002) A stochastic flowering model describing an asynchronically flowering set of trees. Annals Botany. 90:405-415.
2. Erwin J. (2007) Factors affecting flowering in ornamental plants. In: Flower breeding and genetics (issues, challenges and opportunities for the 21st century). Ed. N.O. Anderson. Minnesota. 7-48.

DOI 10.18699/BGRSSB-2018-105

Petri-net-Framework: modeling and simulation of biological networks based on Petri-nets

R. Hofestädt*, C. Brinkrolf
AG Bioinformatics and Medical Informatics, University Bielefeld, Technical Faculty, Germany
** e-mail: ralf.hofestaedt@uni-bielefeld.de*

Key words: Petri-Net, simulation, new tool

Motivation and Aim: Based on a short introduction of Petri-nets and an overview of existing simulation shells this presentation will focus to a new Petri-net simulation shell based on the OpenModelica software tool. A user interface will be presented, which allows the access to the Petri net library (PNlib) ofOpenModelica. The PNlib-Shell provides a powerful simulation environment for the simulation of biological networks. Based on this new shell Petri net models can be easily created, simulated and analyzed. In addition the new system includes basic features to check and evaluate the model and to analyze simulation results generated by the simulation back-end.

References

1. Hofestädt R. (2017) Advantages of Petri-Net modeling and simulation for biological networks. J. Bioscience, Biochemstry Bioinformatics. 7(4):221-229. DOI 10.17706/ijbbb.2017.7.4.221-229.

DOI 10.18699/BGRSSB-2018-106

A catalog of human genes and a gene network controlling feeding behavior and body weight

E.V. Ignatieva
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
Novosibirsk State University, Novosibirsk, Russia
e-mail: eignat@bionet.nsc.ru

Key words: feeding behaviour, body mass index, database, PPI network, obesity

Motivation and Aim: Obesity is a heritable disease with only a few safe and long-term effective therapies and intervention strategies. In efforts to understand the genetic basis of obesity and to collect data on potential targets for new therapies, we (1) created a catalog of human genes regulating feeding behavior (FB) or body weight (BW); (2) constructed protein-protein interaction (PPI) network involving genes/proteins from the catalog; (3) prioritized genes/proteins according to their degree in the PPI network.
Methods and Algorithms: The set of genes controlling FB or BW were collected from publications: (1) 578 human genes from the compendium of genes involved in regulation of FB or BW described in [1]; (2) 134 GWAS genes associated with elevated body mass index (BMI) revealed in [2]; (3) additional 190 GWAS genes associated with elevated BMI revealed in [3] at a less stringent significance level (between $5*10^{-8}$ and 10^{-5}).
Results: We created a catalog comprising 853 human genes controlling FB and BW. According to the type of evidence, genes were classified into three functional categories: (1) GWAS genes from [1] and [2] (Sublist A); (2) other genes from [1] (Sublist B); (3) genes revealed in [3] at a less stringent significance level of 10^{-5} (Sublist C). Using STRING, GeneMANIA and Cytoscape, we constructed PPI network formed by genes/proteins from the catalog and ranked all genes according to the number of neighbors in the network. The top genes/proteins were: (1) *MAPK3*, *PPARG*, and *IRS1* in Sublist A; (2) *ESR1*, *CREBBP*, and *STAT3* in Sublist B; (3) *YWHAZ*, *NCK1*, and *MTOR* in Sublist C. We revealed that six genes/proteins from the Sublist C (*NCK1*, *ADCY5*, *PAM*, *NCAM1*, *BACE2*, *MRAS*) had PPIs with genes/proteins associated with monogenic non-syndromic obesity.
Conclusion: By systematic review and curation of multiple lines of evidence, we created a comprehensive catalog of genes regulating FB and BW. Based on PPI network we found hubs among genes revealed by GWAS. We propose to keep in mind these genes as potential candidates for investigating the genetic factors predisposing to elevated body weight.
Acknowledgements: The work was supported by the Integrational Project of SB RAS No. 0324-2018-0021.

References

1. Ignatieva E.V. et al. (2016) A compendium of human genes regulating feeding behavior and body weight, its functional characterization and identification of GWAS genes involved in brain-specific PPI network. BMC Genetics. 17(Suppl 3):15.
2. Turcot V. et al. (2018) Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. Nat Genet. 50(1):26-41.
3. Locke A.E. et al. (2015) Genetic studies of body mass index yield new insights for obesity biology. Nature. 518:197-206.

DOI 10.18699/BGRSSB-2018-107

Adaptive strategies of motile bacteria in dynamic aquatic ecosystems. A simulation study

A.I. Klimenko*, Yu.G. Matushkin, S.A. Lashin
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
Novosibirsk State University, Novosibirsk, Russia
** e-mail: klimenko@bionet.nsc.ru*

Key words: marine bacteria, chemotaxis, motility, energy budgets, adaptive strategies, individual-based modelling

Motivation and Aim: There are two different strategies living organisms resort to in their struggle for existence facing the environmental changes – either adaptation to existing conditions, e. g. by adjusting their metabolism to alternative energy source, or migration following the optimal environmental conditions invading new biotopes where they compete with the species of a local community. Both strategies are relevant for microbial ecosystems and both the mechanisms that drive the emergence of species following different adaptation strategies and the conditions that favor the domination of certain variants are of a fundamental scientific interest. Current estimates show that energy expenditure of various marine bacteria ranges from 2 to 50 % of their total energy budgets [1] with different species varying in motility. However, the causes underlying the sustainability of such a diversity in energy expenditure range remain obscure as its relation to metabolic efficacy does. According to these questions, we have investigated which adaptive strategies emerge during the evolution of traits associated with migration.
Methods and Algorithms: We have used the Haploid Evolutionary Constructor 3D (HEC 3D) [2] software complex to build the models of coevolution of metabolic and migratory traits in populations of microorganisms. The HEC 3D allows creating multilayer models of evolution of trophically interlinked populations of motile microorganisms inhabiting spatially structured environments.
Results: Investigating the models of microbial populations characterized by the polymorphism in gene that controls migratory energy costs, we have found out that given homogeneous initial distribution of cells there are two different adaptive strategies: either decreasing energy costs of migration and performing regular adaptive migrations towards nutrient rich biotopes or reluctance in spending energy for migration along with a respective increase in energy costs of motility.
Conclusion: It has been demonstrated that the nutrient variable conditions and homogeneous initial distribution of cells cause the disruptive selection by the trait of migratory energy costs. We suggest such a mechanism to be underlying the corresponding processes that are observed in nature.
Acknowledgements: The study was supported by the budget project 0324-2018-0017.

References:

1. Stocker R., Seymour J.R. (2012). Ecology and physics of bacterial chemotaxis in the ocean. Microbiol. Molecular Biology Reviews. 76(4):792-812.
2. Klimenko A.I. et al. (2015) Modeling evolution of spatially distributed bacterial communities: a simulation with the haploid evolutionary constructor. BMC Evol. Biol. 15:S3.

DOI 10.18699/BGRSSB-2018-108

Infrastructure systems biology europe (ISBE): emergence of innovative systems biology servicing

A. Kolodkin[1, 2, 3, 4, 5]*, L. Alberghina[6], J.L. Snoep[7], H.V. Westerhoff[1, 4, 5, 8]
[1] *Infrastructure for Systems Biology Europe* (*ISBE*)
[2] *Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Luxembourg*
[3] *Federal Centre Institute of Cytology and Genetics of Siberian Division of the Russian Academy of Sciences, Russia*
[4] *Molecular Cell Physiology, VU University Amsterdam, The Netherlands*
[5] *Synthetic Systems Biology, SILS, University of Amsterdam, The Netherlands*
[6] *SysBio Centre of Systems Biology, University of Milano-Bicocca, Italy*
[7] *Department of Biochemistry, University of Stellenbosch, South Arica*
[8] *Manchester Centre for Integrative Systems Biology, UK*
* *e-mail: a.n.kolodkin@vu.nl*

Key words: systems biology, networks, modelling, service

The Infrastructure Systems Biology Europe (ISBE) provides stewardship and help with biological and medical data, their acquisition, their analysis and their understanding. It consists of 5 interconnected infrastructure pillars: 1) The Make Me My Model (M4) pillar consists of a software infrastructure that helps customers to make their various types of data (genome sequence, transcriptome, proteome, metabolome, physiological, kinetic, etc.) predictive and understood via modelling (www.isbe.nl). 2) The Do Me an Experiment pillar is a distributed hardware-plus-service infrastructure that performs systems-biology quality assays as a service (M5; Make Me My Mass Spectra Measurements, enzyme kinetics, metabolomics, and epigenetics (http://www.sysbio.it/isbe/)). 3) The Live Model Repository (LMR) of ISBE is a software infrastructure of interconnectable, systems-biology-quality kinetic models through JWS Online (https://jjj.bio.vu.nl/). 4) The Data and Model Stewardship of ISBE called FAIRDOM (http://fair-dom.org). FAIRDOM platform supports Systems Biology to make Data (models, data, SOPs, samples, workflows) FAIR and platform-exchangeable. 5) Help Me to Model (HMTM) provides training to customers wishing to make models themselves, in online or workshop tutorials (www.isb.nl).

The Netherlands branch of ISBE (www.isbe.nl) focuses on M4, LMR and HMTM. It provides phenomenological modelling (top-down) as well as bottom-up mechanism-based modelling, followed by model analyses that help understand the system under study, e. g. predicts the effects of therapeutic or biotechnological interventions useful for model-driven experimental design and bioengineering, or for therapeutic practice.

ISBE.NL is currently providing CORBEL clients with online modelling services using the FAIRDOM HUB website (https://fairdomhub.org/) and coordinating with other European Research Infrastructures, such as EuroBioImaging and Elixir. ISBE.NL profits from the expertise of distinctive systems biologists in a wordwide collaborative network extending from the Netherlands, Luxembourg, Italy, Spain, UK, Germany, and Slovenia, to Russia, and shows the way for innovative network-accommodating biology, biotechnology and medicine.

DOI 10.18699/BGRSSB-2018-109

The "transformer" model of ROS management for PD and cancer: Network diversification as the source of precision in blue-print modelling

A. Kolodkin[1, 2, 3, 4, 5]*, A.M. Colangelo[6], L. Alberghina[6], R.P. Sharma[7], V. Kumar[7], D. Jennen[8], J.J. Briede[8], H.V. Westerhoff[1, 4, 5, 9]

[1] *Infrastructure for Systems Biology Europe* (*ISBE*)
[2] *Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Luxembourg*
[3] *Federal Centre Institute of Cytology and Genetics of Siberian Division of the Russian Academy of Sciences, Russia*
[4] *Molecular Cell Physiology, VU University Amsterdam, The Netherlands*
[5] *Synthetic Systems Biology, SILS, University of Amsterdam, The Netherlands*
[6] *SysBio Centre of Systems Biology, University of Milano-Bicocca, Italy*
[7] *Center of Environmental Food and Toxicological Technology* (*TecnATox*)*, Departament d'Enginyeria Química, Universitat Rovira i Virgili, Spain*
[8] *Department of Toxicogenomics, Maastricht University, The Netherlands*
[9] *Manchester Centre for Integrative Systems Biology, UK*
* *e-mail: a.n.kolodkin@vu.nl*

Key words: systems biology, dynamic modelling, blueprint modelling, cancer, Parkinson's disease

Biomolecular and functional similarity of various cell types allows constructing a universal blueprint model of the cell. Most processes and qualitative descriptions of biomolecular interactions should be similar. They would differ in kinetic parameter values and expression levels. Re-parameterization should allow switching the model from one cell type, e. g. hepatocyte to another, e. g. dopaminergic neuron. This resembles transformers – children toys that can change shape from humanoid, to vehicle, to weapon, with just a few simple turns.

We have built a blueprint model of ROS management in several instantiations, e. g. hepatocyte and neuron-like cells. Then we used two independently obtained data sets showing the response of those cells to oxidative stress. Fitting the model to one data set, should somewhat enhance the confidence to the model describing another data set. This is then an example of how blueprint modelling of the cell may become more and more precise by taking into account various instantiations and data sets.

This blueprint concept goes hand-in hand with implications for building a universal model for different diseases. Striking examples are Parkinson's disease and cancer, which look like opposing diseases: in cancer the cell lives, in PD the cell dies. Our ROS model explains both cases by demonstrating how an overexpression of just one protein (DJ1), relevant for some cancers, makes cell "immortal" under oxidative stress, while downregulation of the same protein, which is relevant for some cases of PD, accelerates cell death under oxidative stress.

DOI 10.18699/BGRSSB-2018-110

Development of nonlinear regression models of flowering time control by climatic factors in soybean and chickpea

K. Kozlov[1]*, L. Novikova[1, 2], I. Seferova[2], S. Nuzhdin[1, 3], M. Samsonova[1]
[1] *Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia*
[2] *Federal Research Center the N. I. Vavilov All-Russian Institute of Plant Genetic Resources, St. Petersburg, Russia*
[3] *University of Southern California, Los Angeles, CA, USA*
** e-mail: kozlov_kn@spbstu.ru*

Key words: soybean, chickpea, flowering time, nonlinear regression, grammatical evolution

Motivation and Aim: Climate impact on agriculture is getting stronger due to combined effects of rising average temperatures, reduced water supply in dry regions and more frequent extreme weather conditions. Agriculture must intensify, become more sustainable, and possess greater resilience to pests and climate. Crucial to this effort are predictive models that connect agricultural traits to climatic factors.
Methods and Algorithms: The analytic representation of non-linear dependence of a agronomic traits on climatic factors is build using a formal approach called "Grammatical Evolution". This technique constructs N functions from "words" of length L according to the rules of a defined grammar. In this grammar a word represents expression that may encode predictor (X) or operation on expressions. The model is further build as a linear combination of N functions with LASSO approach [1]. The set of predictors is determined by minimization of approximation error with Differential evolution (DEEP). The model quality was assessed with determination coefficient R^2. Each model is characterized by the vector of fixed length that contains predictor indices and coefficients, thus the approach allows us to quantitatively compare the model structures.
Results: The method was applied to predict the flowering time in two datasets, namely 379 samples of 9 soybean accessions of different origin phenotyped at Pushkin VIR station in 1999–2013 [2] ($R^2 = 0.60$) and chickpea VIR landraces from Turkey and Ephiopia, phenotyped in Syria ($R^2 = 0.47$). The model vectors build for Turkey and Ephiopia chickpea varieties were uncorrelated (Pearson $r = 0.13$, $p = 0.4 > 0.05$) and their difference was statistically significant according to Mann-Whitney test ($p = 0.04 < 0.05$).
Conclusion: The flowering time models developed in this work are more accurate in comparison with earlier models [2] but possess more complex structure and contain hyperbolic functions. The results obtained with chickpea dataset showed the ability to quantitative compare varieties using statistical testing of difference in the model structures.
Acknowledgements: Supported by the Federal Targeted Program (Agreement No. 14.575.21.0136). Calculations were performed in Supercomputer Center of Peter the Great St. Petersburg Polytechnic University.

References

1. Kozlov et al. (2018) A mathematical model of the impact of climatic factors on soybean development. Biophysics. 63(1):175-176.
2. Seferova I., Novikova L. (2015) Climatic factors that impact development of fast ripening accessions of soybean in Nothwest Russia. Applied Botany, Genetics and Selection. 176:88-97.

DOI 10.18699/BGRSSB-2018-111

FAIRDOMHub: a repository and collaboration environment for sharing systems biology research

O. Krebs[1]*, K. Wolstencroft[2], J.L. Snoep[3], N.J. Stanford[4], F. Bacall[4], M. Golebiewski[1], S. Owen[4], A.R. Williams[4], W. Mueller[1], C. Goble[4]
[1] *Heidelberg Institute for Theoretical Studies, Heidelberg, Germany*
[2] *Leiden Institute of Advanced Computer Science,Netherlands*
[3] *University of Stellenbosch, Stellenbosch , South Africa*
[4] *School of ComputerScience, University of Manchester, UK*
** e-mail: olga.krebs@h-its.org*

Key words: systems biology, experimental data management, databases, mathematical modeling

The systems biology approach has an iterative cycle of experimental and modeling analyses. Experimental results inform mathematical model design and refinement, and modeling simulations direct further laboratory experiments. Data are highly heterogeneous and the relationships between multiple different data sets and mathematical models must be clearly maintained. The interlinking of the experimental data, standard operating procedures (SOPs) and models is essential for interpreting and understanding results.
The FAIRDOMHub is a repository for publishing FAIR (Findable, Accessible, Interoperable and Reusable) Data, Operating procedures and Models (https://fairdomhub.org/) for the Systems Biology community. It is a web-accessible repository for storing and sharing systems biology research assets. There are several well established databases for mathematical models or types of experimental data (e. g. omics data and kinetics), but FAIRDOMHub combines data and models and provides services that enable the integration, interlinking and publishing of experimental and modeling results in the context of the overall systems biology experiment, from a project perspective. It enables researchers to organize, share and publish data, models and protocols, interlink them in the context of the systems biology investigations that produced them, and to interrogate them via API inter-faces. By using the FAIRDOMHub, researchers can achieve more effective exchange with geographically distributed collaborators during projects, ensure results are sustained and preserved and generate re-producible publications that adhere to the FAIR guiding principles of data stewardship.
The FAIRDOMHub (https://fairdomhub.org/) has been developed as a joint action between ERA-Net ERASysAPP (https://www.erasysapp.eu/), an EU-wide consortium of applied systems biology, the European Research Infrastructure and Infrastructure for Systems Biology in Europe (ISBE) (http://project.isbe.eu/).

DOI 10.18699/BGRSSB-2018-112

Mathematical modeling of formation and supporting of the structure of the root apical meristem *Arabidopsis thaliana* L.

V.V. Lavrekha[1, 2], V.V. Mironova[1, 2]*
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: victoria.v.mironova@gmail.com*

Key words: mathematical modeling, auxin, cytokinin, root apical meristem, cell cycle, stem cell niche

Motivation and Aim: The stereotypic root structure, formed of concentric cell layers (the epidermis, cortex, endodermis, and pericycle, which encircle the central vascular system), is generated by the stem cell niche in the root apical meristem (RAM). The apparent simplicity of the root structure derives from an intricate pattern of developmental trajectories inside the RAM. E. g. in Arabidopsis, the diarch symmetry of vascular tissues in the root, including the pericycle, disturbs the radial root structure of the RAM. Plant hormones auxin and cytokinin play major roles in both cell division and differentiation. From another hand, their distributions are known to be nonuniform in the RAM with maxima and gradients. Here we infer the minimal regulatory mechanism for longitudinal patterning of the RAM within the gradients of auxin and cytokinin.
Methods and Algorithms: We hypothesized a plausible mechanism for regulation of the cell cycle in the RAM by auxin and cytokinin. We implemented the mechanism in 1D mathematical model. The hybrid model created using Dynamic grammar in the Plenum package for the Mathematica system. It simulates cell growth and division under control of the proposed circuit, wherein cell growth, active and passive transport of auxin and cytokinin between the cells, and their degradation are described by continuous functions. Transitions between the cell cycle phases are discrete events, which are described by stochastic rules.
Results: In the numerical calculation of the model, we observed the self-organization the proliferation domain between the maximum of auxin in the quiescent center (QC) and the maximum of cytokinin on its shootward edge. The model predicted the difference in length of proliferation activity for different vascular lineages. The model predictions have been verified experimentally.
Conclusion: Auxin and cytokinin regulate the size of the proliferation domain of RAM through concentration-dependent control of the mitotic cycle.
Acknowledgements: The work was supported by Russian Federation President Grant for young scientists (MK-1297.2017.4) and the Russian Ministry of Science and Education under 5-100 Excellence Programme.

DOI 10.18699/BGRSSB-2018-113

On evolutionary analysis of gene networks by the Orthoscape software

Z.S. Mustafin[1]*, D.A. Afonnikov[1, 2], Yu.G. Matushkin[1, 2], S.A. Lashin[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: mustafinzs@bionet.nsc.ru*

Key words: Cytoscape plugin; ortholog; paralog; metabolic pathway; gene regulatory network; evolution; phylostratigraphy; evolution

Motivation and Aim: There is a huge amount of networks of different types in biology. There is also a number of software intended to visualization and analysis of such networks. One of the key software frameworks to work with biological networks is Cytoscape (http://cytoscape.org/). Recently we had presented Orthoscape, a Cytoscape application for evolutionary analysis and visualization of gene networks and gene sets [1]. Such an analysis, which includes phylostratigraphic analysis and Darwinian selection analysis, may shed light on origin and evolution of complex traits determined by gene networks
Methods and Algorithms: We used KEGG (http://www.kegg.jp/) as the source of biological data. We used CyKEGGParser plugin [2] to get networks from the KEGG Pathway and original KEGG API to get lists or homologs with identity and SW Score values, protein domains, nucleotide and amino acid sequences and taxonomic information. K_a/K_s was calculated by PAML [3] using pairwise sequence comparisons for the taxa under analysis. Sequences were aligned using the Needleman–Wunsch algorithm realization adapted from (http://zhanglab.ccmb.med.umich.edu/NW-align/). The gene divergence value based on every gene-ortholog pair. K_a/K_s result, is allows us to discriminate the diversifying and the stabilizing selection.
Results: 81 human diseases networks from the KEGG Pathway were analyzed. The most evolutionary young genes influenced into immune system group of diseases. They have also shown the highest evolutionary rate (the highest K_a/K_s value) in comparison with the Primates. The oldest genes have been found within the substance dependencies group. There are small amount of genes at all and the most part of them influenced in basic biological processes like breath regulation.
Availability: http://apps.cytoscape.org/apps/orthoscape; https://github.com/ZakharM/Orthoscape
Acknowledgements: The study is supported by the integration project 0324-2018-0021.

References

1. Mustafin Z.S., Lashin S.A., Matushkin Y.G., Gunbin K.V., Afonnikov D.A. (2017) Orthoscape:a cytoscape application for grouping and visualization KEGG based gene networks by taxonomy and homology principles. BMC Bioinformatics. 18(S1):1-9.
2. Nersisyan L., Samsonyan R., Arakelyan A. (2014) CyKEGGParser: tailoring KEGG pathways to fit into systems biology analysis workflows. F1000Research. 145:1-11.
3. Yang Z. (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24(8):1586-1591.

DOI 10.18699/BGRSSB-2018-114

Circadian rhythm of biological processes in mouse liver and kidney: analysis of RNA-seq and ribosome profiling data

N.L. Podkolodnyy[1, 2]*, N.N. Tverdokhleb[1, 3], O.A. Podkolodnaya[1]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia*
[3] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: pnl@bionet.nsc.ru*

Key words: circadian rhythm, translation, GO enrichment analysis, tissue specificity, biological processes, phase characteristics

Motivation and Aim: A large-scale analysis of the circadian dynamics of transcriptomes and translatomes makes it possible to understand the role of transcriptional and posttranscriptional regulation of circadian clocks of the organism better and to reveal their relationship with the daily dynamics of biological processes. The purpose of this work is to study the tissue-specific features of the circadian phase characteristics of biological processes based on the GO functional analysis of genes with a pronounced circadian dynamics of transcription and translation in the mouse liver and kidney.
Methods and Algorithms: We used experimental data on the gene expression at the mRNA level and ribosome profiling (GSE67305 and GSE81283) in mouse liver and kidney. Identification of genes demonstrating the circadian dynamic is carried out by a modified method based on JTK_CYCLE. Identification of circadian patterns of gene expression (mRNA level and ribosome profiling) is carried out by the methods of correlation, cluster analysis and the principal component analysis. An analysis of the enrichment of groups of genes by the terms of GO was carried out using the bioinformatics resource DAVID.
Results: We have identified genes that demonstrate pronounced circadian dynamics of transcription and translation. For the twelve time points (ZT0-ZT22) we have identified for each of the tissues groups of genes that were in a phase with an increased level of expression. For these groups of genes the GO enrichment analysis was carried out.
We have identified processes, the rhythmicity of which is characteristic for both the liver and the kidneys. Some of these processes showed similar circadian phase characteristics. At the same time, processes with significantly different temporal phase patterns in these two organs were identified. Also revealed processes with strict tissue-specific rhythmic translation. The circadian patterns of the mRNA level and ribosome profiling of the liver and kidney genes of the mouse were revealed. We analyzed the distribution of gene expression patterns across these 12 groups of genes and shown, in particular, that genes with a pronounced the impulse pattern of expression (the gene is expressed shortly during the day (no more than 2-4 hours)) most often have a maximum of expression at ZT16.
Conclusion: Our approach allows us to analyze the tissue-specific phase characteristics of biological processes, and the results emphasize the need to take into account the phase circadian characteristics when comparing the features of their course of these processes in various organs.
Acknowledgements: Supported by Presidium of SD RAS (0324-2018-0021) and by RF Government (0324-2018-0017).

DOI 10.18699/BGRSSB-2018-115

3D multiscale hybrid modeling and simulation of vascular tumour growth including spatio-temporal distribution of central metabolism

M. Reuss*, A. Lapin
Stuttgart Research Center Systems Biology, University Stuttgart Germany
** e-mail: reuss@srcsb.uni-stuttgart.de*

Key words: modeling, metabolism, cancer, tumour, cellular automaton, 3D multiscale model

Multiscale modeling and simulation in systems medicine is an emerging methodology and discipline to tackle the enormous challenges posed by complex diseases. Cancer modeling is an outstanding and important examples for solving problems which have important features at multiscale of time and space. The presentation aims at the application of a 3D hybrid multiscale hybrid discrete-continuum model to simulate angiogenesis and vascular tumour growth [1].The models is based on a cellular automaton approach and couples intracellular processes, active cell movement, cell-cell interaction, extracellular diffusion, an a dynamically evolving vascular network. For use of FDFG-PET data further extension of the model regarding the glucose balance are required. Coupling a dynamic model of the central metabolism in Hepatoma cells developed in our lab [2] which has been experimentally verified by quantitative measurements of metabolites and dynamic C13 flux analysis permits for the first time to consider a spatio-temporal distribution of intracellular central metabolism within a tumour.

References

1. Perfahl H., Byrne H., Chen T., Estrella V., Alarcon T., Lapin A., Gatenby R.A., Gillies R.J., Lloyd M.C., Maini P.K., Reuss M., Owen M.R. (2011) Multiscale modelling of vascular tumour growth in 3D: the role of domain size and boundary conditions. PLoS One. 6:e14790.
2. Maier K., Hofmann U., Reuss M., Mauch K. (2010) Dynamics and control of the central carbon metabolism in hepatoma cells. BMC Systems Biology. 4:54.

DOI 10.18699/BGRSSB-2018-116

Comparison of high- and low-resolution MS data for direct tissue profiling on a way from laboratory to clinic

A. Sorokin[1]*, V. Shurhkhay[1, 3], K. Bocharov[1, 2], I. Popov[1, 2, 4], E. Zhvansky[1], A. Potapov[3], E. Nikolaev[1, 2, 5]

[1] *Moscow Institute of Physics and Technology (State University), Dolgoprudny, Russia*
[2] *Institute of Energy Problems of Chemical Physics of RAS, Moscow, Russia*
[3] *N.N. Burdenko Scientific Research Neurosurgery Institute, Moscow, Russia*
[4] *Emanuel Institute of Biochemical Physics of Russian Academy of Sciences, Moscow, Russia*
[5] *Skolkovo Institute of Science and Technology, Skolkovo, Moscow region, Russia*
** e-mail: lptolik@gmail.com*

Key words: high-resolution mass spectrometry, phospholipids, biomarkers, brain tumor, database

Motivation and Aim: It is known that the lipid metabolism reprogramming is one of the new hallmarks of cancer. Over the last decade, great progress was made in the understanding of the role of lipid metabolism in the progression of cancer. High-resolution mass spectrometers are widely used in research facilities but applying mass spectrometry to clinical use requires reproducing the results on the low-resolution instruments, which commonly used in routine applications. For the effective automatic classifications, the spectra of one specific type of sample have to represent the same array of peaks characterizing this type of samples. Nevertheless, in practice high- and low-resolution spectra obtained even by similar instruments from one sample are not as similar as one could expect. In this work, the feature selection approach to determine features that are common for the different resolution spectra demonstrated.
Methods and Algorithms: All experiments performed on Thermo LTQ Orbitrap XL instrument. Mass spectrometry data obtained with the novel direct-spray-from-tissue approach ion source in the negative mode as described in [1]. For registration of high-resolution spectra (resolution 56,000 at 800 Th), we used Orbitrap analyzer. The low-resolution spectra (resolution 1,000 at 800 Th) we obtained using LTQ analyzer of the same instrument. All biological samples collected from dissected brain tumors during neurosurgical operations in the N.N. Burdenko Scientific Research Neurosurgery Institute. Mass-spectrometry data preprocessing and feature selection was performed in R environment by custom script available on request from authors.
Results: Both types of spectra from the same tumor fragment were processed separately through the data analysis pipeline for denoising, aggregation, normalization, peak picking and peak alignment. Two datasets were then aligned to each other and analyzed. The low-resolution spectra contain about one-third of the peaks, detectable in high-resolution spectra; however, all major peaks found in both types of spectra. The vast majority of features distinct in tumor and unmodified brain were found in both types of spectra. More than, we have created mapping schema, which allows using classifier trained on high-resolution spectra with low-resolution spectra.
Conclusion: It is shown that low-resolution spectra preserve distinctive features of brain tumor samples and could be used in sample classification.
Acknowledgements: Supported by the RSF (16-15-10431).

References

1. Kononikhin A., Zhvansky E., Shurkhay V., Popov I., Bormotov D., Kostyukevich Y., Karchugina S., Indeykina M., Bugrova A., Starodubtseva N., Potapov A., Nikolaev E. (2015) A novel direct spray-from-tissue ionization method for mass spectrometric analysis of human brain tumors. Analytical and Bioanalytical Chemistry. 407(25):7797-7805.

DOI 10.18699/BGRSSB-2018-117

Evolutionary computations and modular organization of the gene regulatory regions

A. Spirov[1]*, E. Myasnikova[2]
[1] *I.M. Sechenov Institute of Evolutionary Physiology and Biochemistry RAS, St. Petersburg, Russia*
[2] *Peter the Great St. Petersburg Polytechnical University, St. Petersburg, Russia*
** e-mail: alexander.spirov@gmail.com*

Key words: Genetic Algorithms, Royal Road function, binding site, promoter, in silico gene design, synthetic biology

Motivation and Aim: Modular organization and functioning of gene sequences, RNA and polypeptides attracts much attention during recent decades (e.g., reviewed in [1]). Modularity in gene regulatory regions (promoters and enhancers) is crucial for our understanding of the gene functioning and evolution (e.g. [2]). In the vast area of Evolutionary Computations, inspired by the ideas and concepts from evolutionary biology, it was paid a special attention to the theoretical foundations for the evolutionary search efficacy. At first place these were Schema theorem and Building block hypothesis (by J. Holland [3]), that laid the foundation for this area in genetic algorithms (GA). On the way to further develop the theory, the Royal Road problem and Royal Road functions (RRFs) were introduced [4] and comprehensively studied [5]. Here we are considering several case-studies of the modular gene regulatory regions which could be treated as RRFs implementations in directed molecular evolution (SELEX, etc.).
Methods and Algorithms: Analytical tools from statistical mechanics, dynamical systems theory, and mathematical population genetics gave possibilities to develop a detailed and quantitative description of the search dynamics for the RRF class of problems [van Nimwegen with co-authors]. The approach bridges evolutionary computations from benchmark cases, such as RRF, which are well-understood theoretically, to biological cases, which can serve as a basis for more efficient directed molecular evolution in the test tube.
Results: By introducing GA crossover operators that perform well on RRFs, we are developing computational techniques to deal with the real design problems for bacterial and yeast promoters. In particular, we are introducing GA crossover operators that work like retroviral or sexual PCR recombination. The case examples are the bacterial promoters in comparison with the yeast promoters. We found that our algorithms are capable to achieve the polynomial efficacy of the evolutionary search. The common, standard GA algorithms are the the exponential-time algorithms for the problems.
Conclusion: Computational theory from GA can contribute to both understanding how real gene structures have evolved and to speeding up laboratory work on directed evolution of promoters and other gene regulatory elements.
Acknowledgements: Supported by the Russian Science Foundation (grant 17-18-01536).

References

1. Lorenz D.M. et al. (2011) The emergence of modularity in biological systems. Physics Life Rev. 8:129-160.
2. Dynan W.S. (1989) Modularity in Promoters and Enhancers. Cell. 58:1-4.
3. Holland J.H. (1975) Adaptation in Natural and Artificial systems, University of Michigan Press, Ann Arbor, MI. 183 p.
4. Mitchell M. et al. (1992) The royal road for genetic algorithms: Fitness landscapes and GA performance. In F.J. Varela and P. Bourgine, eds. Towards a practice of autonomous systems. 245-254.
5. Jansen T., Wegener I. (2005) Real royal road functions – where crossover provably is essential. Discrete Applied Mathematics. 149:111-125.

DOI 10.18699/BGRSSB-2018-118

Pulse wave velocity measurement in the human radial artery

B.G. Vainer
Rzhanov Institute of Semiconductor Physics SB RAS, Novosibirsk, Russia
Novosibirsk State University, Novosibirsk, Russia
e-mail: BGV@isp.nsc.ru

Key words: pulse wave velocity, radial artery, humans, computer-controlled brachial cuff inflation/deflation, midget microphones, time lag between acoustic signals, influence of the brachial artery cross-clamping

Motivation and Aim: The evaluation of pulse wave velocity (PWV) provides information about the elastic properties of arterial system and some other physiological attributes of the organism. At present, the "gold standard" in the PWV studies is a carotid-femoral measurement [1]. A brachial-ankle configuration is also popular and usable [2]. Nevertheless, as it is noted in [3], the PWV estimation remains a challenge for the engineers and clinicians. In this contribution, a simple and robust novel approach helping to solve the problem in question is presented. In addition, some valuable results obtained in humans using the described new method are demonstrated.
Methods and Algorithms: Two originally framed midget microphones were attached to the wrist and upper-limb bend of elbow and served as the pulse wave-induced pressure transducers. Biopac MP 100 system was used to receive and digitize the microphone signals with 200 Hz temporal discretization. An original computer-controlled brachial cuff inflation system was used for both the acoustic signal enhancement and the arterial pressure automated measurement. The latter was realized during inflation and deflation in the same diagnostic cycle. The method was tested in humans of different ages.
Results: The measured pulse pressure waveforms were quite similar in appearance that increased the accuracy in the determination of the time lag between acoustic signals. The problem of the pulse wave path length definition was fallen away because the distance between the signal sources (two microphones) was certainly the length of the forearm that could be measured easily. It has been found that the radial artery (RA) PWV may be appreciably influenced by the proximal artery transient cross-clamping. The PWV measurement results combined with the diagnostic manifestations synchronously observed using infrared thermography [4] are also presented.
Conclusion: A simple and feasible novel method aimed at the RA PWV measurement has been presented. It is suitable to the scientific research and clinical routines. The RA PWV may be influenced by a proximal artery transient cross-clamping or occlusion.
Acknowledgements: Supported by the RFBR (grant No. 18-08-00956).

References

1. Asmar R., Benetos A., Topouchian J., Laurent P., Pannier B., Brisac A.-M., Target R., Levy B.I. (1995) Assessment of arterial distensibility by automatic pulse wave velocity measurement. Validation and clinical application studies. Hypertension. 26:485-490.
2. Tanaka H., Munakata M., Kawano Y., Ohishi M., Shoji T., Sugawara J., Tomiyama H., Yamashina A., Yasuda H., Sawayama T., Ozawa T. (2009) Comparison between carotid-femoral and brachial-ankle pulse wave velocity as measures of arterial stiffness. Journal Hypertension. 27:2022-2027.
3. Pereira T., Correia C., Cardoso J. (2015) Novel methods for pulse wave velocity measurement. Journal Medical Biological Engineering. 35:555-565.
4. Vainer B.G. (2018) A novel high-resolution method for the respiration rate and breathing waveforms remote monitoring. Annals Biomedical Engineering. https://doi.org/10.1007/s10439-018-2018-6.

DOI 10.18699/BGRSSB-2018-119

Selection versus adaptation: network diversification and the origins of life, ageing and cancer

H.V. Westerhoff
Synthetic Systems Biology and Nuclear Organization, University of Amsterdam
Molecular Cell Physiology, Vrije Universiteit Amsterdam
Systems Biology, The University of Manchester
Infrastructure Systems Biology Europe. NL
e-mail: h.v.westerhoff@uva.nl

Key words: networks, functional heterogeneity, drug resistance, flux balance analysis, stochastic modelling

Networking enables functions that are otherwise thermodynamically impossible, such as the synthesis of ATP, proteins and DNA. We shall here highlight a lesser known function of networking, i.e. diversification. Network diversification followed by selection, sprouted the tree of Life, but that very tree hides a forest of diversity. Early Life on this planet may have benefitted from diversification of the redox network around acetogenesis. Flux Balance Analysis (FBA) of the genome-wide metabolic network of *Cl. Ljungdahlii* reveals carbon fixation at various ATP/acetate stoicheiometries. This may have enabled early organisms to survive the erratic environmental conditions by shifting gears. The flexible ATP yield enabled by the Warburg effect may help do so for tumor cells.

Our organs are subject to a drizzle of somatic mutations, which leads to cell diversification with age. We shall review an FBA methodology that simulates this and then demonstrate that this may initially enhance the metabolic versatility of organs such as T-cells and liver, in the absence of adaptation. More somatic mutations would cause loss of function and ageing. And, subpopulations of asocial cells would develop into tumors with Warburg and the new WarburQ (i. e. glutamine dependent) phenotypes.

Likewise, noise such as required by the third law of thermodynamics should diversify cell populations. FISH and deep sequencing experiments show even stronger noise than this, which should thereby be subject to regulation, e. g. through transcription bursting. We shall show that even though such noise varies with time, it may be selectable and may lead to drug resistance of tumor cell populations, either because of nonlinearities or because of the 'Waddington' genetic landscape, a remnant of developmental biology.

DOI 10.18699/BGRSSB-2018-120

Reconstruction of whole-genome metabolic model of Atlantic salmon *Salmo salar* (SALARECON)

M. Zakhartsev*, J.O. Vik, F. Grammes
Norwegian University of Life Sciences, Ås, Norway
** e-mail: maksim.zakhartsev@nmbu.no*

Key words: constrained based stoichiometric model, genome-scale metabolic model, diets, osmoregulation, farmed salmon, optimization of growth conditions

Introduction: Atlantic salmon *Salmo salar* is Norway's main livestock and the biggest export commodity after the oil. Currently, the plant feedstock is the main feed for the farming salmon, consequently, it influences both growth rate and quality of the fish biomass. The whole-genome (WG) metabolic model of *S. salar* (SALARECON) will allow multifactorial optimization of the fish growth conditions by means of a diet for the application in salmon farming. Special interest will be paid to the optimization of fat yield and growth limitations due to dietary amino acids. The reconstruction of WG metabolic model requires integration of available genomic, biochemical and physiological information.

Methods: The genome and metabolic information are available at NCBI and KEGG. The SALARECON is the integrated stoichiometric model that describes steady state growth of the biomass. The reconstruction workflow includes: (1) Gene-Protein-Reaction associations (GPRs) using in-house SAPP system; (2) network setup and network topological analyses; (3) Flux Balance Analysis for optimization of the network using Insilico Discovery package. The model allows integration of high-throughput experimental omics-data (transcriptomics, proteomics, metabolomics).

Results and Discussion: The model integrates all important biochemical reactions/pathways that lead to polymerization of major biomass constituents: proteins, carbohydrates, fats/lipids, polynucleotides. The modeling methodology is based on energy-centric approach because energy- and redox-balances in different subcellular compartments play a central role in coupling and harmonizing activities of different metabolic modules and pathways. The energy costs for the ion balance is one of the main contributor to the maintenance costs. The model will be optimized for two osmotic scenarios: fresh and sea waters.

Currently, the SALARECON integrates 311 transformer steps (performed by products of 931 genes), 255 balanced compounds belonging to 56 pathways which are allocated in 3 compartments. For the network reconstruction, the in-house developed SAPP system is used to predicts GRP associations as well as subcellular compartment localizations of the corresponding biochemical reactions.

The model aims to optimize a whole range of possible diets of farmed salmon. Exact formulation of the biomass composition (model's output) defines the input fluxes through the network of reactions, which are compared with experimentally measured dietary requirements. The model is validated using variety of experimental omics-data.

DOI 10.18699/BGRSSB-2018-121

Deciphering cell's robustness by a multi-scale framework integrating cell cycle and metabolism in yeast

L. van der Zee, T.D.G.A. Mondeel, H. Westerhoff, M. Barberis*
Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, The Netherlands
** e-mail: M.Barberis@uva.nl*

Key words: Boolean modeling, constraint-based modeling cell cycle, metabolism, budding yeast

Motivation and Aim: Cell cycle and metabolism are coupled networks. For example, cell growth and division require synthesis of macromolecules which is dependent on metabolic cues. Conversely, metabolites involved in nucleotide and protein synthesis are fluctuating periodically as a function of cell cycle progression. Although computational models of these networks are being developed for some time, to date no effort has been made to integrate these two systems in any organism. We aim to investigate cell cycle robustness by generating the first multi-scale model that integrates cell cycle with metabolism, and investigating their bidirectional regulation. Connections among these two biochemical networks have been recently elucidated in budding yeast. However, high-throughput and manually curated studies point at many more physical interaction, which relevance for precise cell cycle timing remains unknown.
Methods and Algorithms: A framework is presented that integrates a Boolean cell cycle model with a constraint-based model of metabolism, incorporating mechanistic and high-throughput interactions. Directionality and effect are incorporated for the mechanistic interactions. Conversely, as this information is unknown for the high-throughput interactions, an informed optimization algorithm has been developed to generate models that can incorporate it iteratively.
Results: To verify the results of the informed optimization algorithm against metabolomic data, changes in flux through a number of metabolic pathways are compared to metabolic pathway enrichment time-series. The multi-scale model predicts expected changes in a number of pathways, ranging from amino-acid to pentose phosphate to lipid metabolism. Many model variants that differ in number and directionality of interactions robustly predict the effect of definite cell cycle-metabolism pairs. Furthermore, the integrative model shows a temporal export of acetate, pyruvate and alanine, reminiscent of yeast metabolic oscillations.
Conclusion: Altogether, our multi-scale framework is able to integrate computer models of biological networks with high-throughput data, to capture the functional connectivity among their elements that ultimately results in systems robustness.
Acknowledgements: Supported by the SILS Starting Grant of the University of Amsterdam, UvA and by the UvA-Systems Biology Research Priority Area Grant.

References

1. Barberis M., Todd R.G., van der Zee L. (2017) Advances and challenges in logical modeling of cell cycle regulation: perspective for multi-scale, integrative yeast cell models. FEMS Yeast Research. 17(1).

BIOINFORMATICS AND SYSTEMS BIOLOGY OF PLANTS

DOI 10.18699/BGRSSB-2018-122

Using the mathematical modeling approach to access the behavior of plant antioxidant system under abiotic stress conditions

A. Bobrovskikh[1, 2]*, A. Doroshkov[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: avb@bionet.nsc.ru*

Key words: antioxidant system, modelling, ROS, stress response

Motivation and Aim: Reactive oxygen species (ROS) is one of the key factors that damages living systems. ROS production during stress dramatically increases, which causes membrane damage and can lead to death of the living system. The plants have an antioxidant system (AOS) that effectively fights against the ROS. Biochemistry of this system is well-known. AOS includes a number of antioxidant enzymes (ascorbate-glutathione cycle, catalase, superoxide dismutase, catalase) and antioxidant species (ascorbate, glutathione). However, the regulation of this system insufficiently studied. There are many kinetic data on antioxidants species and expression data. Integration of this data and dynamic modeling is a perspective approach to describing the complex response of the system to external disturbances, such as abiotic stress.
Methods and Algorithms: Literature sources were used to obtain kinetic data. Expression experiments was extracted from GEO databases. We used the COPASI software [1] to build a dynamic model of the antioxidant system.
Results: We present a new model of AOS in a plant cell. In addition to chloroplasts, models of peroxisome and mitochondrial components of the AOS were created. Based on these data, stationary states are calculated in normal and stressed conditions (salt stress, water-deficient).
Conclusion: Our study includes a complex description of the dynamics of the plant's AOS in response to stress, taking into account the division into compartments. The kinetic mechanisms and kinetic parameters of enzymes of the antioxidant system, concentration of enzymes in different cell compartments and generation rates of free radicals are based on the literature data and databases.
Acknowledgements: We thank Dr. Alexey Kolodkin for consultations that improved the model quality.

References

1. Hoops S. et al. (2006) COPASI—a complex pathway simulator. Bioinformatics. 22(24):3067-3074.

DOI 10.18699/BGRSSB-2018-123

Discordant evolution of cellular genomes in peas (*Pisum* L.) as evidenced from complete sequences of plastid genomes and partial sequences of mitochondrial genomes

V.S. Bogdanova[1], A.V. Mglinets[1], N.V. Shatskaya[1], O.E. Kosterin[1, 2]*, V.I. Solovyev[1, 2], G.V. Vasiliev[1]

[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

[2] *Novosibirsk State University, Novosibirsk, Russia*

** e-mail: kosterin@bionet.nsc.ru*

Key words: pea crop wild relatives, *Pisum*, plastid genome, mitochondrial genome, discordant evolution

Motivation and Aim: The genus *Pisum* L. (Fabaceae) includes an important crop, the pea (*Pisum sativum* L.) and its wild relatives, the genetic diversity of which can be used for crop improvement. At the same time wild peas are a good model for studying microevolutionary phenomena. Earlier we revealed among them a principal concordant divergence of genes from the three cellular genomes [1], later data appeared suggesting a more complicated evolutionary history [2]. To clarify this we carried out an analysis of organellar genomes rather than single markers.

Methods and Algorithms: The plastid genome and a mitochondrial genome fragment were sequenced using the Ion Torrent platform in 16 wild and 3 cultivated accessions representing pea genetic diversity. The genomes were assembled with MIRA4 program, visualized with the Tablet software and checked manually. Phylogenetic analysis was carried out with Bayesian MCMC using of BEAST 1.8.4 software

Results: Phylogenetic reconstructions based on the plastid genome, a mitochondrial genome fragment, and a nuclear gene for the histone H1 subtype 5, updated with phenotypic information, evidenced for at least four events of introgression of organellar genomes through hybridisation of diverged evolutionary lines in the course of the genus evolution. One of the cases concerned introgression of plastid genome and three other cases introgression of mitochondrial genome from a distinct evolutionary lineage, as judged from the phylogenetic pattern revealed from the nuclear gene *His5*. A plastid genome intermediate between the consensuses of the common pea *P. sativum* L. and its distant wild relative, *P. fulvum* Sibth and Smith was revealed, and accession Pe 013 of wild pea from Tokat Province of Turkey was nominated as a 'missing link' between the mentioned species.

Conclusion: Discordant evolution of the three cellular genomes because of introgression may be widespread in plant microevolution and should not be underestimated in reconstruction of phyletic relationships.

Acknowledgements: Supported by the the project No. 0324-2018-0018.

References

1. Kosterin O.E., Bogdanova V.S. (2008) Relationship of wild and cultivated forms of *Pisum* L. as inferred from an analysis of three markers, of the plastid, mitochondrial and nuclear genomes. Genet Res Crop Evol. 55(5):735-755.
2. Zaytseva O.O. et al. (2017) Refinement of the collection of wild peas (*Pisum* L.) and search for the area of pea domestication with a deletion in the plastidic psbA-trnH spacer. Genet Res Crop Evol. 64(6):1417-1430.

DOI 10.18699/BGRSSB-2018-124

Sequencing and assembly of mitochondrial genomes in three conifer species *Larix sibirica*, *Pinus sibirica* and *Pinus sylvestris*

E. Bondar[1]*, A. Kirichenko[1], V. Sharov[1], Yu. Putintseva[1], N. Oreshkova[1, 2], S. Feranchuk[3, 4], Yu. Konstantinov[5], V. Shmakov[5], V. Belkov[5], D. Kuzmin[1], S. Sadovsky[1], K. Krutovsky[1, 6, 7, 8]

[1] *Siberian Federal University, Krasnoyarsk, Russia*
[2] *V.N. Sukachev Institute of Forest SB RAS, Krasnoyarsk, Russia*
[3] *National Research Technical University, Irkutsk, Russia*
[4] *Limnological Institute, SB RAS, Irkutsk, Russia*
[5] *Siberian Institute of Plant Physiology and Biochemistry SB RAS*
[6] *George-August University of Göttingen, Göttingen, Germany*
[7] *Vavilov Institute of General Genetics RAS, Moscow, Russia*
[8] *Texas A&M University, College Station, USA*
** e-mail: bone-post@ya.ru*

Key words: conifers, *de novo* assembly, mitochondrial genome, Siberian larch, Siberian pine, Scots pine

Motivation and Aim: Conifers are ancient group of plants of great economic value and ecological importance represented by more than 600 species including Siberian larch (*Larix sibirica*), Siberian pine (*Pinus sibirica*) and Scots pine (*Pinus sylvestris*) that are keystone species in Siberian boreal forest. Due to the huge sizes both nuclear and mitochondrial genomes are still understudied in conifers. Only a few nuclear and mitochondrial genomes in conifers have been sequenced and published so far. We recently sequenced and assembled nuclear and mitochondrial genomes in two Siberian conifers – *L. sibirica* and *P. sibirica* and in *P. sylvestris*.

Methods and Algorithms: Sequencing was performed using Illumina HiSeq2000 and DNA extracted from purified mitochondria isolated from needles. CLC Assembly Cell was used to assemble paired-end reads into contigs, and the gap-filling program Sealer was used to build scaffolds. To select mitochondrial sequences from nuclear and plastid sequences, assembled contigs were mapped to the nucleotide database of complete and partial plant mitochondrial genomes. To further verify putative mitochondrial contigs in *L. sibirica*, *k*-means clustering in R was used, and the obtained assembly was compared to the whole genome sequencing data. RepeatModeler and RepeatMasker were used to identify repeats.

Results: The Siberian pine mitochondrial genome assembly contained 3.22 Mbp with N50 of 598,637 bp and average GC of 41.78 %. The Scots pine assembly contained 2.34 Mbp with N50 of 289,815 bp and average GC of 42.31 %. The Siberian larch assembly contained 53 scaffolds with a total length of 11.37 Mbp, N50 of 578,009 bp and the largest scaffold of 1,423,047 bp.

Conclusion: The Siberian and Scots pine mitochondrial genome lengths were similar to those earlier studied in other conifers – spruces *Picea glauca* and *P. abies*, while the mitochondrial genome of Siberian larch was much larger and similar to an annual flowering plant *Silene conica* (11.3 Mbp). It does not contain many genes, and repeated and low complexity regions cover only 14.46 % of the sequence. The verification and detailed annotation of the assembled mitochondrial genomes are in progress.

Acknowledgements: This work was supported by research grants No. 16-04-01400 from the Russian Foundation for Basic Research and No. 14.Y26.31.0004 from the Government of the Russian Federation.

DOI 10.18699/BGRSSB-2018-125

Assembly and annotation of genomes of some species from the apomictic genus *Boechera* and evolutionary analysis of apomixis-associated genes

V. Brukhin[1, 2]*, S. Kliver[1], M. Rayko[1], A. Komissarov[1], E. Bakin[1], C. Rushworth[3], T. Mitchell-Olds[4], U. Grossniklaus[5]

[1] *Dobzhansky Center for Genome Bioinformatics, St. Petersburg State University, St. Petersburg, Russia*

[2] *Department of Plant Embryology & Reproductive Biology, Komarov Botanical Institute Russian Academy of Sciences, St. Petersburg, Russia*

[3] *University and Jepson Herbaria, University of California, Berkeley, USA*

[4] *Department of Biology, Duke University, Durham NC 27708-0338 NC, USA*

[5] *Department of Plant & Microbial Biology Zurich-Basel Plant Science Center, University of Zurich, Zurich, Switzerland*

** e-mail: vbrukhin@gmail.com*

Key words: *Boechera*, Brassicaceae, genome, assembly, annotation, apomixis

Apomixis is asexual way of plant reproduction through seeds, which could be found in more than 400 plant species representing almost 40 families. It is believed that apomixis evolved independently in several taxa from sexual ancesters. Apomixis could be considered as a developmental variation of sexual reproduction in which some steps are lost, reduced, deregulated, desynchronized or changed. Thus, apomictic and sexual reproduction are closely related and they share many regulatory components. Molecular and genetic basis underlying apomixis and amphymixis (sexual reproduction) regulation still remains poorly understood. The ability to produce maternal clones and therefore to fix useful traits in the further generations of various crop plants could streamline agricultural breeding strategies based on the genetic aspects of apomixis. The potential of apomixis as a next generation technology for plant breeding attracts huge interest to elucidate molecular and genetic mechanisms of its regulation. Closely related to the model plant *Arabidopsis thaliana*, the genus *Boechera* is known to contain both sexual and apomictic species or accessions. *Boechera retrofracta* is a diploid sexually reproducing species and is thought to be an ancestral parent species of apomictic species. In the presentation will be reported the *de novo* assembly of the *B. retrofracta* genome using short Illumina and Roche reads from 1 paired-end and 3 mate pair libraries. The distribution of 23-mers from the paired end library has indicated a low level of heterozygosity and the presence of detectable duplications and triplications. The genome size was estimated to be equal 227 Mb. N50 of the assembled scaffolds was 2.3 Mb. 27048 protein-coding genes were predicted using a hybrid approach that combines homology-based and *de novo* methods. Also repeats, tRNA and rRNA genes were annotated. Finally, genes of *B. retrofracta* and 6 other Brassicaceae species were used for phylogenetic tree reconstruction. Also, we explored the histidine exonuclease *APOLLO* locus, related to apomixis in *Boechera*, and proposed model of its evolution through the series of duplications. An assembled genome of *B. retrofracta* will help in the challenging assembly of the highly heterozygous genomes of hybrid apomictic species such as *B. divaricarpa*. The *B. retrofracta* genome will also provide a basis to decipher the hybridogenesis events that led to the formation of apomictic *Boechera* accessions.

DOI 10.18699/BGRSSB-2018-126

Genome-scale metabolic model predicts carbon flux partitioning towards starch biosynthesis in storage root of cassava

P. Chiewchankaset[1], W. Siriwat[2], S. Kalapanulak[2, 3], T. Saithong[2, 3]*
[1] *Division of Biotechnology, School of Bioresources and Technology, King Mongkut's University of Technology Thonburi* (*Bang Khun Thian*)*, Bangkok, Thailand*
[2] *Systems Biology and Bioinformatics Research Group, Pilot Plant Development and Training Institute, King Mongkut's University of Technology Thonburi* (*Bang Khun Thian*)*, Bangkok, Thailand*
[3] *Bioinformatics and Systems Biology Program, School of Bioresources and Technology, King Mongkut's University of Technology Thonburi* (*Bang Khun Thian*)*, Bangkok, Thailand*
** e-mail: treenut.sai@kmutt.ac.th*

Key words: genome-scale metabolic model (GMM), carbon flux partitioning, cassava

Motivation and Aim: Cassava is one of the top rank stable crops feeding at least 800 billion people each year. The great amount of starch containing in underground roots have made cassava being a promising carbohydrate source for world population. In contrast to its fascination, how starch is highly produced and accumulated in cassava roots is still uncovered, and become a challenge research question.
Methods and Algorithms: To disclose the mystery of cassava metabolism, the genome-scale metabolic model (GMM) was developed to disentangle the complex metabolic processes of carbon assimilation in cassava roots. The constraint-based model of carbon assimilation pathway was constructed based upon flux balance analysis technique.
Results: The model of carbon assimilation covered 393 metabolites related to 352 biochemical reactions, and 116 the transport reactions required for mobilizing the metabolites between cytoplasm, plastid, and mitochondria subcellular compartments. The model was used to relate the intracellular carbon flux partition to the observed yield of storage roots grown in the field under rainfed condition. Optimized to root growth rate, the model successfully simulated carbon flux towards cellular biomass production that reflected the metabolic phenotype underlying the growing biomass as well as starch accumulation in storage roots of cassava. The model was literature validated and was supported by model sensitivity analysis.
Conclusion: The GMM of cassava storage roots, herein, is the first and crucial step for inquiring more insightful questions of carbon metabolism that yields superior starch production than other starchy crops.
Acknowledgements: Supported by partially supported by King Mongkut's University of Technology Thonburi (KMUTT) under the Higher Education Research Promotion and National Research University Project of Thailand, Office of the Higher Education Commission, Ministry of Education, Thailand, and by NRCT and NSTDA (P-17-52192).

DOI 10.18699/BGRSSB-2018-127

Genetic regulation of wheat plant development and architecture

O.B. Dobrovolskaya[1, 2]*, A.E. Dresvyannikova[1, 2], K.I. Popova[1],
M.P. Ponomarenko[1], Yu.L. Orlov[1, 2], A.A. Krasnikov[3], P. Martinek[4]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Central Siberian Botanical Garden, SB RAS, Novosibirsk, Russia*
[4] *Agrotest Fyto, s.r.o., Kroměříž, Czech Republic*
** e-mail: oxanad@bionet.nsc.ru*

Key words: inflorescence development, plant architecture, meristems, cereal, wheat

Motivation and Aim: Wheat is one of the most important food crop in the world; the yield of grains from this crop is largely dependent from plant and, especially, inflorescence architecture. The inflorescence of wheat is a spike with the main axis (spike rachis) carrying lateral sessile spikelets that are directly attached to the rachis and a terminal spikelet. The study of the genetic factors that determine the structural features of the spikelet, a reduced branch bearing the reproductive organs (the florets), is important to understand the mechanisms underlying plant developmental processes and has obvious practical importance. In wheat, difference in the number of fertile florets per a spikelet is genetically determined by the level of ploidy and interspecific variability, but the genes that determine this trait are currently little studied. The aim of our research was to identify genes that control the development of the bread wheat inflorescence and regulate meristem identity and determinacy.
Methods and Algorithms: A set of classical and modern approaches of genetics and developmental biology, including novel methods of plant genome analysis, such as high-throughput genotyping (GBS), as well as light and electron microscopy methods and modern bioinformatic approaches have been used to characterize unique experimental models. The models included wheat accessions with abnormalities in spike development.
Results: It was found that several genes located on chromosomes 5AL, 2AL, and 2AS control the formation of fan-shaped «flabellum» spikelet of bread wheat *T. aestivum*. Results of SEM analysis showed that the formation of the "flabellum" spikelet was associated with the peculiarities of spikelet development. The features of interaction of plant developmental genes and regulation of gene expression, especially those related to determining identity and determinacy of inflorescence meristems have been studied.
Conclusion: Wheat forms with altered plant/spike morphology represent an important genetic tool for research on the development of the wheat spike and for identification of genes that control meristem activities. Further studies on different non-standard morphotypes and wheat lines with altered spike morphology will allow researchers to identify new genes that control meristem identity and determinacy, to elucidate the interaction between the genes, and to understand how these genes, acting in concert, regulate the development of the wheat spike.
Acknowledgements: Supported by the RFBR (No. 18-04-00483) and the IC&G Budgetary Project No. 0324-2016-0001.

DOI 10.18699/BGRSSB-2018-128

Design of genetic sensors to ethylene based on bioinformatics analysis of whole-genome data

V. Dolgikh
Novosibirsk State University, Novosibirsk, Russia
e-mail: volleydollmc1@gmail.com

Key words: ethylene, ETHYLENE-INSENSITIVE3 (EIN3), ChIP-seq, RNA-seq, *Arabidopsis thaliana*

Motivation and Aim: Ethylene is a gaseous hormone that regulates a wide range of physiological processes involved in plant growth, development and stress responses. Ethylene signaling proceeds via a linear pathway leading to the stabilization of EIN3 transcription factor – the major transcriptional regulator of ethylene response. It activates the expression of a large pool of genes upon binding to the specific sites in their promoters. However, in many cases EIN3 binding to gene promoter is not enough to trigger transcriptional response, which unveils the complexity of EIN3-dependent regulation [1]. Ethylene sensitive *Arabidopsis thaliana* reporter lines are important tools to localize the domains of ethylene biosynthesis in plants. The available genetic sensors used to date utilize several copies of EIN3 binding sites from direct EIN3 targets EDF1 or ERF1 inserted upstream of a minimal promoter followed by a reporter gene [2, 3]. However, it is quite possible that the reporter expression pattern in these lines does not strictly match the domain of ethylene biosynthesis due to the mentioned above complexity of EIN3-promoted regulation. Here we perform the bioinformatics analysis of whole genome data on EIN3 binding coupled with the analysis of ethylene-induced transcriptomes in *A. thalliana* to design a set of genetic sensors to ethylene, which strictly define the domains ethylene biosynthesis.

Methods and Algorithms: We used publicly available ChIP-seq and RNA-seq data from [1] for bioinformatics analysis. Regulatory elements were identified de novo with Homer tool [4]. The significance of motif enrichment in ethylene responsive genes was determined with Fisher's exact test.

Results: We distinguished several EIN3 regulatory patterns and coupled them with ethylene-induced gene expression. We used these predicted regulatory patterns to design a set of fluorescent genetic sensors to ethylene based on ethylene-induced GFP protein expression. The following experimental verification of these biosensors will expand our understanding of the issues that affect hormonal regulation in plants.

Acknowledgements: Supported by Russian Ministry of Science and Education under 5-100 Excellence Programme.

References

1. Chang K.N. et al. (2013) Temporal transcriptional response to ethylene gas drives growth hormone cross-regulation in Arabidopsis. eLife. 2:e00675.
2. Stepanova A.N. et al. (2007) Multilevel interactions between Ethylene and Auxin in Arabidopsis roots. Plant Cell. 19(7):2169-2185.
3. Solano R. (1998) Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. Genes and Development. 12(23):3703-3714.
4. Heinz S. et al. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 38:576-589.

DOI 10.18699/BGRSSB-2018-129

Molecular evolution analysis of genetic network components related to plant trichome development

A.V. Doroshkov[1, 2]*, D.K. Konstantinov[1], K.V. Gunbin[1, 2], D.A. Afonnikov[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: AD@bionet.nsc.ru*

Key words: molecular evolution, trichome, MBW complex, gene regulatory network

Motivation and Aim: Trichomes are involved in many significant functions such as the transpiration, thermoregulation and protection from insect attacks. On the other hand specialized cell formation is an fruitful model system for analyzing the molecular mechanisms of plant cell differentiation, including cell fate choices, cell cycle control, and cell morphogenesis. However, the evolution of specialized epidermal cell formation genetic network remains unclear.
Methods and Algorithms: We reconstructed the network of interactions between known leaf pubescence genes using A. thaliana as a model organism using Cytoscape and Pathway Studio software. For each network node (gene) the extraction of sets of homological protein sequences was carried out using the DELTA-BLAST. Then multiple sequence alignment was conducted with MAFFT algorithm. The PhyML and IQTree maximum likelihood algorithms was used to reconstruct the phylogenetic trees of protein families and bootstrap resampling technique was used for testing the topology. In order to find phylogenetic tree branches at which accelerated evolution occurred, we reconstruct protein ancestral sequence state spectra in each inner tree nodes using Bayesian approach implemented in PhyloBayes. After that for each protein family we reconstructed relative rate matrix of amino acid substitutions by modelestimator software and compared it with the inferred frequencies of amino acid substitutions on each tree branch that estimated by the ancestral sequence state spectra comparison on these branches. Additionally, for each residue of each ancestral proteins we predicted its structural properties using RaptorX Property pipeline. As a result of these procedures, for each internal branch of protein family phylogenetic tree we derived several statistical values describing acceleration of sequence and structure evolution.
Results: The main stages of gene network evolution has been traced down to the evolutionary time of appearance of its components. It was found that the LCA of dicotyledonus plants in a comparison with *A. thaliana* have a reduced set of R2R3-MYB factors (7 versus 13), and at the LCA of all flowering plants also have a reduced set of R3-MYB inhibitors (1 versus 7). The reduction of the network component sets resulted in at least 4 protein families: ttg1, egl3, glabra, cpc. The above described analysis of protein evolution in each inner branches of that protein families demonstrated that in contrast to glabra protein family accelerated protein evolution events associated with cpc, egl and ttg1 concentrated on orthologs formation events related with divergence between flower plants and gymnosperms.
Conclusion: Thus, we found that the main adaptive evolution events of the gene network involved in plant trichome development related with divergence between flower plants and gymnosperms.

DOI 10.18699/BGRSSB-2018-130

Genetic mechanisms of resistance to golden potato cyst nematode *Globodera rostochiensis* in *Solanum phureja*

A.A. Egorova[1, 2]*, N.A. Shmakov[1], S.V. Gerasimova[1], G.V. Vasilyev[1], N.V. Shatskaya[1], K.V. Strygina[1], D.A. Afonnikov[1], A.V. Kochetov[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk StateUniversity, Novosibirsk, Russia*
** e-mail:egorova@bionet.nsc.ru*

Key words: potato, RNA-seq, *Solanum phureja*, resistance, *Globodera rostochiensis*

Motivation and Aim: Golden potato cyst nematode (GPCN) is an important pathogen of potatoes, tomatoes, and other plants in the family *Solanaceae*. GPCN juveniles damages the roots of susceptible plants. Nowadays GPCN is found worldwide and it appears to be the one of the most deleterious pathogens for potato. Protection against GPCN is complicated because the eggs of nematodes can remain viable in the soil for more than 30 years. In addition, most chemical nematicides are not efficient, thus the protection is mostly obtained through the introduction of the resistance genes.
In this study we analyzed the resistant cultivar of diploid potato *Solanum phureja* to reveal new resistance genes through comparison of root transcriptomes of resistant and susceptible genotypes.
Methods and Algorithms: For RNA-seq, total RNA was extracted from root samples collected in time points 0, 24 and 72 hours after inoculation with GPCN. Sample preparation was carried out by colleagues from Vavilov Institute of Plant Genetic Resources and All Russian Research Institute for Plant Protection (Saint Petersburg, Russia) according to [1]. Sequencing was performed on Illumina NextSeq 500 platform. FastQC and Prinseq tools were used to assess sequences quality and filter the libraries. STAR and TopHat were used to map the filtered libraries to the reference genome. Search for the differentially expressed genes was performed using Cufflinks pipeline and EdgeR package for R. Lists of differentially expressed genes (DEGs) were further analyzed with Biomart and the databases AgriGO, KEGG, and PlantCyc. De novo transcriptome assembling was carried out with Trinity software. Prediction of NBS-LRR genes was based on their typical domain structure.
Results: Analysis of *S. phureja* transcriptomic data revealed differential expression of a number of genes. According to the literature data, the most probable candidate gene(s) for resistance against GPCN may belong to the NBS-LRR family. These receptors recognize specific molecular patterns and can induce programmed cell death at the site of a pathogen invasion. Thus, we focused our analysis on this type of R-genes. Candidate resistance genes expressed in *S. phureja* resistant cultivar were predicted for further analysis of segregating populations.
Conclusion: Candidate pathogen major resistance genes were predicted by transcriptome analysis of two *S. phureja* genotypes contrasting in GPCN resistance.
Acknowledgements: Supported by grant from RSF No. 16-16-04073.

References

1. Kochetov A.V. et al. (2017) Differential expression of NBS-LRR-encoding genes in the root transcriptomes of two Solanum phureja genotypes with contrasting resistance to Globodera rostochiensis. BMC Plant Biol. 17(Suppl 2):251.

DOI 10.18699/BGRSSB-2018-131

Genetic modification of cereals and increasing size of caryopsis

I. Erokhin
NBK LLC, Moscow, Russia
e-mail: i.erokhin@inbox.ru

Key words: caryopsis, cereals, ontogenesis, junk DNA , Genome Tree Theory

In accordance with the Genome Tree Theory [1], the genome of a multicellular organism is structured (in terms of graphs theory) in the form of an oriented binary tree, whose tops are represented by similar logical elements "genome tree cycle step", and arcs are the transfers of control between the "steps". Each cell of the multicellular organism has its step in its Genome Tree. It is controlled by the step of the mother cell, initiates performance of the "cell program" which determines its development and division, and returns control upon its completion and transfers management to the steps of daughter cells. The cell program is a network (multi-level embedded Dijkstra loop), where regulatory genes, as well as structural ones, are used in a certain order. The step of the Cell Program may be represented by a couple of non-coding regulatory genes with identical complex promotors, and the Genome Tree step – by three non-coding regulatory genes. Separate steps of the Genome Tree may also be represented by protein-coding regulatory genes.

Transcription of the non-coding regulatory gene is supposedly initiated with the help of the nucleoprotein formed with general transcription factors of Pol III polymerase that contains a short RNA transcribed from another non-coding regulatory gene. One of the sites (ligand) of this short RNA should be complimentary to some site (acceptor) of promotor of the initiated regulatory gene. Another site (emitter) of the initiated regulatory gene encodes the new ligand in the transcribed short RNA. This new ligand may help to initiate another regulatory non-coding gene. Thus, control is transmitted between regulatory genes of Cell Programs and Genome Tree.

The size and shape of caryopsis of cereals are determined by the relevant branch of the Genome Tree, dominant allele of triploid genome of endosperm cells. In particular, its size (number of cells that form an endosperm) depends on the number of steps that form this branch of the Genome Tree. Adding another step to this branch leads to an increase in the number of cells and, consequently, the size of caryopsis. The number of added cells depends on the position of the additional step in the Genome Tree branch. The closer is the added step to the root (in terms of graphs theory) of the branch responsible for endosperm formation, the more new cells are added. Adding a step in extreme tops of the Genome Tree leads to an insignificant increase in the number of new cells. The endosperm properties depend on the activated Cell Programs. To insert an additional step between the two available (previous and subsequent) steps requires adding to the genome three extra non-coding regulatory genes with identical complex promotors (or modifying the three inactive genes that exist). An additional step in other branches of the Genome Tree will lead to increase in sizes of the relevant organs or the whole plant. Such accidental mutations of regulatory genes are responsible for a variety of sizes and shapes of Angiospermae and Gramíneae, among others.

References

1. Erokhin I. (2016) Genome tree theory. International Conference on the Mathematical Modeling and High-Performance Computing in Biomedicine and Biotechnology, August 29–September 2, Novosibirsk, Russia.

DOI 10.18699/BGRSSB-2018-132

Peculiarities of genome structure of flowering plants

I. Erokhin
NBK LLC, Moscow, Russia
e-mail: i.erokhin@inbox.ru

Key words: flowering plants, cereals, ontogenesis, junk DNA, Genome Tree Theory

The meristem of plants contains cells with different genomes. The genome of somatic cells differs from that of generative cell after crossing-over. The genome of zygote and embryo cells also differs from that of somatic cells. Some alleles of the polyploid endosperm cells genome, in its turn, differs from the haploid component of genomes of somatic cells, zygote or embryo cells. Despite the fact that the steps of Genome Tree [1], which determine the ontogenesis of apical generative cells and apical somatic cells are located on different branches of the Genome Tree (and may belong to somewhat different variants of the Genome Tree), these cells are located close to each other in apical meristem in terms of space. Moreover, in the process of branching, they are divided simultaneously and continue to stay close to each other in each new offshoot.

In the process of branching, the Genome Tree step of the mother somatic cell in apical meristem transfers the management in each daughter cell to the root (in terms of graphs theory) step of the Genome Tree branch responsible for offshoot ontogenesis. Such complete looping of the Genome Tree fragment may ensure "endless" plant growth. Similarly, the Genome Tree fragment of the mother generative cell in the apical meristem is also fully looped in the process of branching. Grafting a veneer that contains such an offshoot forms a whole new plant. Spontaneous mutations of regulatory non-coding genes that change the looped somatic branch of the Genome Tree change the phenotype of the plant. This property is the basis of fruit trees domestication. At the same time, generative branch remains unchanged and the plant that grows from the planted seed preserves the old unchanged phenotype. The mutations that change the generative branch of the Genome Tree do not change the current phenotype of the plant, but influence the phenotype of net generation plants that grow from the seeds.

If looping of the Genome Tree fragment is limited by the external loop that contains several steps, the plant growth will not be "endless". For example, a wheat stem consists of several metamers, and a head also contains several spikelets. Introduction of mutations that extend the external loop of the Genome Tree fragment, responsible for head formation, should lead to increase in the number of grains in the head and thus better wheat yield. Spontaneous mutations of non-coding regulatory genes may lead to appearance of plants of larger size. If such a phenotype change is associated with mutations in the somatic branch of the Genome Tree, the next generation of plants loses these beneficial properties. New properties can be saved in several ways. Modeling the structure of looped fragments and other branches of Genome Tree that are responsible for ontogenesis of separate organs and groups of plant cells is of great interest, but this is the subject of other studies.

References

1. Erokhin I. (2016) Genome tree theory. International Conference on the Mathematical Modeling and High-Performance Computing in Biomedicine and Biotechnology, August 29–September 2, Novosibirsk, Russia.

DOI 10.18699/BGRSSB-2018-133

Digital collection of morphological variability of wheat spike

M. Genaev[1]*, E. Komyshev[1], F. Hao[2, 3], V. Koval[1], N. Goncharov[1, 4], D. Afonnikov[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Heilongjiang University, Harbin, China*
[4] *Novosibirsk State Agrarian University, Novosibirsk, Russia*
** e-mail:mag@bionet.nsc.ru*

Key words: wheat, spike characteristics, phenotyping, annotation, data integration

Motivation and Aim: The structure of the ear is one of the most important features of cereals associated with such agronomically important traits as productivity, resistance to environmental factors and pests, threshebility. Ears differ in shape, size, density, ostectomy, color, etc. Analysis of the ear traits requires visual inspection, manual measurements and is very time-consuming.
Methods and Algorithms: The effectiveness of ears' phenotyping can be improved by the introduction of an automated image processing technology, storage of information in databases, use of machine learning algorithms to analyze this information. In this work, we present a new approach for collecting, storing and analyzing of information about morphometric characteristics of ears of wheat. Two protocols for obtaining digital images of the ear have been developed.
Results: The computer-aided information system SpikeDroidDB has been developed, which allows you to store digital images of the ear, annotate their phenotypic features (14 features, including plant variety description, links to parent genotypes, generation, plant sowing number, ear morphology description). The interface provides a flexible query system to access the data. SpikeDroidDB represents an interconnected representation between genotype, phenotype, location, and growing conditions. The web interface of SpikeDroidDB is available at http://spikedroid.biores.cytogen.ru and allows you to work with the system as with desktop computers or mobile devices. We used SpikeDroidDB for the digitization and annotation of a collection of ears of F2 hybrids from crosses between the Australian cultivar of common wheat Triple Dirk and accession KU506 of Chinese wheat *Triticum yunnanense*. This experiment includes analysis of 103 plants, 230 spike images.
Conclusion: The analysis of the variability of ears in form, length, and other traits allowed determination of the type of their genetic control: compactness is controlled by two recessive genes, awn type and hairiness at the site of attachment of the spikelet to the axis is controlled by single dominant gene type, hairiness on the axis of the spike is controlled by two dominant genes.
Acknowledgements: The work supported by the RSF grant No. 17-74-10148.

DOI 10.18699/BGRSSB-2018-134

Altering barley agronomic traits via targeted mutagenesis

S.V. Gerasimova[1,*], A.M. Korotkova[1], C. Hertig[2], I. Otto[2], S. Hiekel[2], J. Kumlehn[2], A.V. Kochetov[1, 3], E.K. Khlestkina[1, 3, 4]

[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

[2] *Leibniz Institute of Plant Genetics and Crop Plant Research* (*IPK*)*, Gatersleben, Germany*

[3] *Novosibirsk State University, Novosibirsk, Russia*

[4] *N.I. Vavilov All-Russian Research Institute of Plant Genetic Resources* (*VIR*)*, Saint-Petersburg, Russia*

** e-mail: gerson@bionet.nsc.ru*

Key words: barley, *Hordeum vulgare*, genome editing, crop improvement, Cas9, gRNA, spike architecture

Motivation and aim: Target-specifically customizable endonucleases are a new molecular tool, which opens a huge variety of novel opportunities for life sciences, biotechnology and breeding. Crop improvement is one of the most promising applications of this technology. The aim of study is to modify agronomical traits of barley (*Hordeum vulgare* L.) exploiting RNA-guided endonuclease. Targeted knockout of *Nud* and *Vrs1* genes should lead to phenotype switches from hulled to naked barley and two-rowed to six-rowed spike, respectively. To extend the utilization of the technology from the model to local elite cultivars, we decided to evaluate the regeneration ability of immature embryos from the ten most prospective spring barley cultivars of the Siberian collection to include the best performing ones in our approach.

Materials and Methods: *In vitro* multiple shoot formation was tested across several barley cultivars and lines (“Biom”, “Talan”, “Vorsinskiy 2”, “Aley”, “Acha”, “Signal”, “L-421”, “Kolchan”, “V-1”, “Krasnoyarskiy 91”, “Golden Promise”). The structure of *Nud* and *Vrs1* genes was confirmed by sequencing. Adequate guide RNAs for targeting *Nud* and *Vrs1* barley genes were selected with available online-tools. Transformation vectors with different gRNAs and Cas9 genes were assembled for transient and stable expression in barley. The activity of gRNAs was tested *in vivo* using the methods of mutational restoration of reporter gene functionality [1] and protoplast transfection with subsequent T7E1 assay [2]. Selected genetic constructs were introduced into the barley genome by *Agrobacterium*-mediated gene transfer to immature embryos.

Results: Regeneration analysis of Siberian barley cultivars revealed few cultivars with high *in vitro* regeneration capacity. Cultivar “Aley” was selected for further experiments aiming at site-directed mutagenesis. A set of different gRNAs for knockout of *Nud* and *Vrs1* genes was established. *In silico* and *in vivo* evaluation of gRNA activity revealed highly active gRNAs for both genes. Genetic constructs containing gRNA and Cas9 expression cassettes were introduced into the genome of barley and primary transgenic plants were obtained.

Conclusion: RNA-guided endonuclease targeting *Nud* and *Vrs1* genes was shown to be highly active in barley cells of cv. “Aley” and “Golden Promise”, genetic transformation of the Siberian cv. ”Aley” is being established.

Acknowledgements: The study is supported by the RSF (No. 16-14-00086).

References

1. Budhagatapalli N. et al. (2016) A simple test for the cleavage activity of customized endonucleases in plants. Plant Methods. BioMed Central. 12(1):1-10.
2. Shan Q. et al. (2014) Genome editing in rice and wheat using the CRISPR/Cas system. Nat. Protoc. 9:2395-2410.

DOI 10.18699/BGRSSB-2018-135

Organization and evolution chalcone synthase gene family in bread wheat and related species

A.Yu. Glagoleva[1, 2]*, E.K. Khlestkina[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: glagoleva@bionet.nsc.ru*

Key words: *Chs gene, chalcone synthase, gene duplication, flavonoid biosynthesis, Triticum, Aegilops*

Motivation and Aim: Flavonoids are secondary plants metabolites, the most of these compounds are related with coloration traits in plants, which makes them a convenient model for genetics studies. Chalcone synthase (CHS) is a key enzyme of the flavonoid biosynthesis pathway and is involved in the biosynthesis of all classes of flavonoid compounds. The genes encoding this enzyme are usually represented in the plants genomes in many copies and in most species are described in detail. Nevertheless, in bread wheat (*Triticum aestivum* L.) these genes have not been explored yet. The purpose of this study was to investigate the structural and functional organization of the chalcone synthase gene family, and its evolution in bread wheat and relative species.
Methods and Algorithms: Homologous sequences search was performed using BLAST algorithm in two databases (https://urgi.versailles.inra.fr/blast/blast.php, www.ncbi.nlm.nih.gov/Database/) within genomic sequences of *T. aestivum* and its tetraploid (*T. durum*) and diploid (*T. monococcum*, *T. urartu, Aegilops speltoides*, *Ae. sharonensis*, and *Ae. tauschii)* relatives. Multiple sequence alignment was done with MULTALIN 5.4.1. Cluster analysis was performed with MEGA v6.06 software using Neighbor-Joining algorithm. To design a set of copy-specific primer pairs PrimerQuest Tool (https://eu.idtdna.com/Primerquest/Home/Index) was used. These primers were exploited for PCR from DNA of nulli-tetrasomic and deletion lines and RT-PCR from cDNA for pericarp, coleoptile and root.
Results: The nucleotide sequences of the five *Chs* copies in *T. aestivum* were identified. Among them three homeologous gene copies in A- and D-genomes (*Chs-A1*, *Chs-B1* and *Chs-D1*) and two paralogous gene copies in B-genome (*Chs-B2*, *-B3*). It was shown that all *Chs* gene copies are located on the distal regions of 2AS, 2BS and 2DS chromosomes. All copies with the exception of *Chs-B2* transcribed in colored pericarp and coleoptile, but they were not transcriptionally active in colorless pericarp and root. *Chs-B2* was transcribed in colored coleoptile only. Analysis of transcriptional activity and the structure of the promoters of copies of the *Chs* copies showed that they participate in the synthesis of different classes of flavonoid compounds in different (both optimal and stressful) conditions. To clarify the origin of paralogous *Chs* duplications in the B genome, we compared sequences of *Chs* genes in different *Triticum* and *Aegilops* species and calculated the time of divergence of the paralogues.
Conclusion: First *Chs* duplication event took place in common ancestor of *Triticum* and *Aegilops* about 10 MYA, then one of the copies was again duplicated 6-7 MYA in the ancestor of the B-genome, while other copy likely pseudogenized in *T. aestivum* 2A and 2D chromosomes. The specialization of individual copies seems to be the reason for maintaining five *Chs* gene copies in the genome of *T. aestivum*.
Acknowledgements: The study was partially supported by RFBR (16-04-00912).

DOI 10.18699/BGRSSB-2018-136

Identification of genes, associated with black pigmentation of seeds in cereals, based on transcriptomic analysis

A.Yu. Glagoleva[1, 2]*, N.A. Shmakov[2], O.Yu. Shoeva[2], G.V. Vasiliev[2], N.V. Shatskaya[2], A. Börner[3], D.A. Afonnikov[1, 2], E.K. Khlestkina[1, 2, 4]

[1] *Novosibirsk State University, Novosibirsk, Russia*

[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

[3] *Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany*

[4] *N.I. Vavilov All-Russian Research Institute of Plant Genetic Resources (VIR), Saint-Petersburg, Russia*

** e-mail: glagoleva@bionet.nsc.ru*

Key words: barley, near-isogenic lines, phytomelanin, Blp, RNA-seq, differetial expression

Motivation and Aim: Some plant species have 'melanin-like' black seed pigmentation. However, the chemical and genetic nature of this 'melanin-like' black pigment has not yet been fully explored due to its complex structure and ability to withstand almost all solvents. Nevertheless, identification of genetic networks participating in trait formation is key to understanding metabolic processes involved in the expression of 'melanin-like' black seed pigmentation. The aim of the current study was to identify differentially expressed genes (DEGs) in barley near-isogenic lines (NILs) differing by allelic state of the *Blp* (black lemma and pericarp) locus.

Methods and Algorithms: We used FastQC to estimate sequencing quality, cutadapt to remove adapter sequences and prinseq to filter sequences by quality and length. Filtered libraries were mapped to H. vulgare genome assembly version 32,608 v 1.33 from the Ensembl Plants database (http://plants.ensembl.org). Mapping was performed with TopHat2 tool. The resulting alignments were processed with the Cufflinks v 2.2.1 pipeline. Pathway analysis was performed using PlantCyc database (http://www.plantcyc.org/). After that selected DEGs were verified using qRT-PCR.

Results: Firstly, we excluded involvement of metabolic pathways, known for other types of pigments (i. e. widespread flavonoid pigments). Our data demonstrated that none of the key flavonoid biosynthesis genes is expressed significantly in black-grained barley during the pigment formation. Then, we used RNAseq approach to reveal DEGs in black-colored and uncolored-grained barley. A total of 957 genome fragments had statistically significant changes in expression levels between lines BLP and BW, with 632 fragments having increased expression levels in line BLP and 325 genome fragments having decreased expression. Genes with high level of differential expression in BLP line were identified in following pathways: suberin monomer biosythesis, diterpene phytoalexins precursors biosynthesis, cutin biosynthesis, cutular wax biosynthesis, and phenylpropanoid biosynthesis, initial reactions.

Conclusion: The number and diversity of metabolic pathways confirms our earlier assumption about the pleiotropic nature of the *Blp* locus. Besides forming a black color of the seeds, *Blp* locus is associated with the resistance to oxidative stress (possibly due to the increased content of antioxidants, including ferulic acid), as well as increased resistance to pathogens (due to the synthesis of phytoalexins and cuticular waxes). The black color of the seeds is presumably associated with the action of the enzyme polyphenol oxidase (PPO).

Acknowledgements: The study was supported by the Russian Science Foundation (No. 16-14-00086).

DOI 10.18699/BGRSSB-2018-137

Genetic variability of tea plant (*Camellia sinensis* (L.) Kuntze) on the Black Sea coast of Caucasus

M.V. Gvasaliya[1], L.S. Samarina[1]*, Yu.L. Orlov[2], M.-V. Hanke[3], H. Flachowsky[3]
[1] *Russian Research Institute of Floriculture and Subtropical Crops, Sochi, Russia*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[3] *Julius Kühn-Institut (JKI), Federal Research Institute for Cultivated Plants Institute for Breeding Research on Fruit Crops, Dresden, Germany*
** e-mail: samarinalidia@gmail.com*

Key words: Camellia sinensis, inheritance, somatic mutagenesis, breeding, chromosome aberrations

Motivation and Aim: It is known that the frequency of spontaneous hereditary variability depends on plants species. There is a high genetic variability and plasticity observed in tea plant (*Camellia sinensis* (L.) Kuntze) grown in natural conditions [1]. Moreover, the increase of genetic variability of tea is influenced by various environmental factors, such as temperature shocks, high radiation, high doses of fertilizers etc. The spectrum of somatic mutations of tea is represented by morphological, physiological or plastid modifications. Deviations in length, width, shape, surface (smooth, bubbly), and degree of leaf serration, leap positions, and the length of the internodes are occurred most often among the morphological traits. Physiological changes are manifested as the leaf color change from yellow-green to dark-green, as well as various degree of anthocyanin color, lengthening the vegetative period, reducing generative activity, increasing yield. In turn, the plastid mutations are represented by sectorial and periclinal chimeras [2]. The aim of current work is to reveal genetic variability in two tea genotypes on the Black Sea coast of Caucasus.

Methods: Studies were carried out on tea cultivar Kolkhida, and population Kimyn (large-Chinese varieties) grown in the region of the Black Sea coast seeds were germinated in a thermostat at a temperature of 27–30 °C. Cytological and karyological analysis were performed on root meristems using the Carnoy fixation, and staining with acetocarmine.

Results and conclusion: The highest frequency of chromosomal rearrangements was noted in the Kymyn population of 2.2–5.9 %. Positive correlation between the frequency of somatic mutations and the frequency of cells with aberrations was noted. The Kolkhida variety was characterized by a lower frequency of altered anaphases. Moreover, a higher frequency of chromosome rearrangements was noted in regions with increased use of chemical fertilizers: the differences were statistically significant in the varieties of the Kymyn population ($P < 0.001$) and in the Kolkhida variety ($P < 0.05$). Thus, the frequency and spectrum of chromosomal aberrations of tea depend on the genotype and area of growth. Among the studied forms, the variety Kimyń is more variable. High genetic variability of can serve as a material for obtaining new tea cultivars.

References

1. Gvasaliya M.V. (2016) Breeding of tea (Camellia sinensis (L.) O. Kuntze), achievements and perspectives Horticulture and Viticulture of Russia. XXXXVII:94-97.
2. Gvasaliya M.V. (2014) Frequency and spectrum of chromosomal aberrations in varieties and mutant forms of Camellia sinensis (l.) Kuntze under spontaneous mutagenesis. Agricultural Biol. 3:65-69.

DOI 10.18699/BGRSSB-2018-138

Resistivity estimation for a row of potato starches of domestic collection

L. Gvozdeva*, V. Khlestkin
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: ungersy@mail.ru*

Key words: diabetes, genome-wide association studies, potato starch, resistant starch

Motivation and Aim: Production of starch with desired characteristics is an important and urgent task related to dietetic food (resistant starch) and (bio)chemical processing on an industrial scale. Resistant to digestion starch passes through small intestine, remaining intact, and then enters large intestine and serves as an excellent nutrient medium for beneficial local microflora. In contrast to conventional starch, the resistive one increases sensitivity of body cells to insulin and lowers the level of glucose in a blood [1, 2]. The results of the work will make possible to choose more accurately a healthy diet for prevention and/or control of a number of diseases (e. g. diabetes) as well as more efficient use of starch in industrial processing.
Methods and Algorithms: We isolated starch from different varieties of potato mainly of Russian origin. Starch sample was subjected to digestion by α-amylase for 16 h at 37 °C. Precipitate obtained was split by glucosidase for 1 h at 60 °C. We measured the amount of glucose formed with a test-kit «Glucose Olveks»at a wavelength of 500 nm and calculated the amount of resistive starch as a percentage of the total amount of starch taken for analysis.
Results: 90 varieties were investigated for the content of resistive raw starch. The average content of resistive starch was 70 % of the total amount of starch. The contrast forms have been identified and will be used at further stages for the genome-wide association studies. That will allow identification of loci and associated DNA-markers associated with starch resistance to digestion in the upper sections of the gastrointestinal tract.
Conclusion: The first comprehensive study of the starch resistance of domestic potato varieties is performs. The data obtained will be used for associations search after genotyping. Thus, it is possible to quickly select breeding varieties rich in resistive or digestible starch that can be used in the treatment / prevention of a number of diseases (related to the digestibility of starch) or to increase the efficiency of industrial processing of starch.
Acknowledgements: The work supported by a grant from the RFBR (No. 17-29-08006). Samples of potatoes were obtained from the collection of the Genetic Control Center «GenAgro» (IC&G SB RAS).

References

1. Robertson D. et al. (2005) Insulin-sensitizing effects of dietary resistant starch and effects on skeletal muscle and adipose tissue metabolism. American Society for Clinical Nutrition.
2. Shen R.L., Zhang W.J., Dong L.J. (2016) Preparation, structural characteristics and digestibility of resistant starches from highland barley, oats and buckwheat starches. Journal Food Nutrition Research. 55(4):303-312.

DOI 10.18699/BGRSSB-2018-139

Genome-wide prediction of transcription factor binding sites in cassava via phylogenetic footprinting between plants in Euphorbiaceae family

S. Kalapanulak[1, 2]*, C. Rangsiwutisak[3], T. Saithong[1, 2]

[1] *Bioinformatics and Systems Biology Program, School of Bioresources and Technology, King Mongkut's University of Technology Thonburi, Bang Khun Thian, Bangkok, Thailand*

[2] *Systems Biology and Bioinformatics Research Laboratory, Pilot Plant Development and Training Institute, King Mongkut's University of Technology Thonburi, Bang Khun Thian, Bangkok, Thailand*

[3] *Bioinformatics and Systems Biology Program, School of Bioresources and Technology, and School of Information Technology, King Mongkut's University of Technology Thonburi, Bang Khun Thian, Bangkok, Thailand*

** e-mail: saowalak.kal@kmutt.ac.th*

Key words: cassava, regulatory elements, transcription factor binding site, phylogenetic footprinting

Motivation and Aim: Discovering the list of all transcription factor biding sites (TFBSs) on promoter regions in the genome is vital for a complete understanding of transcriptional regulation inside the cell. In order to unravel the systems regulation, several computation approaches have been applied to predict TFBSs and their transcription factors (TFs), for example, TFBS scan based on sequence similarity between known TFBSs of model organisms and promoter sequences of interested organisms. To overcome the limitation of TFBS information, phylogenetic footprinting approach was proposed under the hypothesis that the regulatory elements will be conserved across the related species via evolutionary conservation [1, 2].

Methods and Algorithms: Therefore, in this work, the phylogenetic footprinting approach was applied to identify all putative transcription factor binding sites (TFBSs) in cassava, the emphasized plant for food and energy security in the 21st century. Firstly, the related plants in the same Euphorbiaceae family as cassava were selected, i. e. physic nut and castor bean. The 10,890 orthologous groups between cassava and the other two related plants were identified via bi-directional BLASTp. The upstream sequences of each orthologous group were retrieved from the three plant genomes in the range up to 2,000 bps from translation start site without the overlapping coding sequences with previous gene for identifying TFBSs via Multiple Em for Motif Elicitation (MEME) and Analysis of Motif Enrichment (AME) tool.

Results: Finally, 12,925 candidate TFBSs were discovered from 7,769 cassava genes functioning as several biological components such as enzymes and regulatory proteins.

Conclusion: These results will be useful for proposing regulatory elements found in non-coding DNA, and further hypothesizing a transcriptional regulation in cassava plant.

Acknowledgements: Supported by Thailand research cluster (NSTDA, TRF, ARDA, STI, HSRI and NRCT) with grant ID: P-17-51609 and King Mongkut's University of Technology Thonburi, Thailand for the travelling grant in BGRS\SB-2018 conference.

References

1. Katara P., Grover A., Sharma V. (2012) Phylogenetic footprinting: a boost for microbial regulatory genomics. Protoplasma. 249(4):901-907.
2. Hu J. et al. (2014) Genome-wide identification of transcription factors and transcription-factor binding sites in oleaginous microalgae Nannochloropsis. Sci Rep. 26(4):5454.

DOI 10.18699/BGRSSB-2018-140

Database on molecular identification of genes for resistance in wheat (MIGREW)

F.V. Kazantsev[1, 2]*, E.S. Skolotneva[1], E.A. Salina[1], S.A. Lashin[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: kazfdr@bionet.nsc.ru*

Key words: wheat, stem rust, resistant genes, avirulent genes, database

Accessibility: migrew.sysbio.cytogen.ru
Motivation and Aim: The breeding for stem rust immunity has been recently relaunched in Western Siberia since the infection was occurred in wheat fields. Taking into account the evolution of host-pathogen interactions, the genetic diversity of both, wheat and fungus, must be a permanent object to monitor. The information on avirulent genes (Avr) in migrating fungal populations together with the resistant genes and translocations (R) in wheat lines is important to be involved into breeding programs. However, this pool of data is hard to use without any system where Avr of pathogens and R genes wheat would be agreed upon. There is a lack of available database for wheat cultivars and breeding lines genotyped for resistant genes. Resistance is provided by combinations of specific and non-specific resistant genes with different effectiveness against stem rust between regions. In order to catalogue this information, we have developed the MIGREW database.
Methods and Algorithms: The MIGREW database is designed with PostgreSQL. User interface as a WEB application is designed with DRUPAL libraries. For the purpose of direct data accessing the REST API has been developed via spring.io libraries. The MIGREW data were manually extracted from the publications.
Results: The MIGREW database has been developed for keeping information on fungi-wheat objects. It contains data on (1) pathogen resistance genes, their localization and molecular markers; (2) molecular marker protocols; (3) rust disease resistance genes effectiveness. Public access to the MIGREW is possible in two ways: via WEB application as the main user interface; via direct data access through the REST API. The REST API makes real integration into bioinformatics pipelines.
The data on the polymorphism of Avr genes from the West Siberian population of stem rust, as well as the data on the expression of resistance genes for wheat when infected with monopustule isolates, have been already deposited into the Migrew.
Acknowledgements: The study has been supported by the RFBR grant No. 17-29-08018.

DOI 10.18699/BGRSSB-2018-141

The first study of genetically regulated potato starch biochemical characteristics for *S. tuberosum* cultivars of russian selection

V. Khlestkin[1, 2]*, T. Erst[1], L. Gvozdeva[1], Yu. Khoroshavin[1, 2], I. Totsky[1], E. Khlestkina[1, 2]

[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

[2] *Novosibirsk State University, Novosibirsk, Russia*

* *e-mail: khlestkin@bionet.nsc.ru*

Key words: potato starch, morphology, amylose, amylopectin, resistant starch, phosphorylation

Motivation and Aim: Starch is a well-known sustainable plant resource, which is important for food, textile, paper and other industries. Starch consists of two polysaccharides – linear amylose and branched amylopectin, packed in semicrystalline granules in plant plastids. Ease of isolation makes it an economical and readily available source of pure carbohydrates. Exact ratio and structure of the polysaccharides significantly influence the macroscopic practically useful physical and chemical properties of starch gels. Potato is the most reasonable source of starch in Siberia. Thus, knowledge about gene networks regulating potato starch biosynthesis is a key to growth and manufacturing starch with target properties.

Methods and Algorithms: 90 Cultivars of potato (*Solanum tuberosum* L.) mostly of Russian selection from ecological-geographical testing and "GenAgro" collection grown in 2017 in Novosibirsk region were treated to isolate DNA and starch. Wide SNP-genotyping of the DNA isolated has been outsourced for further association studies. Some important biochemical parameters of starch were evaluated. Thus, we developed a procedure for microscopy analysis of starch granules morphology. We also compiled and tuned a procedure for evaluation of amylose and amylopectin content in potato starch by spectrophotometry. "Resistant starch" value for a raw potato starch for various cultivars and phosphorylation of starch polysaccharides were also estimated.

Results: For the first time an advanced biochemical analysis of a wide range of potato cultivars of Russian selection has been performed. Morphological traits, amylose\amylopectin ratio, resistant starch, glucose units phosphorylation were evaluated. We were able to identify contrasting forms on every biochemical parameter studied for the given set of *S. tuberosum* cultivars. These data will be used for both DNA markers elaboration for marker – assisted selection and future genome-wide association studies. Some assumptions on DNA loci, responsible for starch biochemical traits in the studied cultivars, may be done based on the literature data.

Conclusion: We shown that amylose content of the studied set of *S. tuberosum* cultivars varies within 13–30 %, phosphorus content – within 0.05–0.1 %, resistant starch – within 40–99 %. Average starch granule value is in significant correlation with starch preparative yield. Moderate correlations of resistant starch content with Feret's diameter and phosphorylation with preparative starch yield are observed.

Acknowledgements: Supported by RFBR (17-29-08006). Potato cultivars obtained from plant collection "GenAgro" (ICG SB RAS).

DOI 10.18699/BGRSSB-2018-142

The synthesis of antioxidant enzymes in potato plants under biotic and abiotic stress conditions

I. Kirgizova[1]*, A. Gajimuradova[2], R. Omarov[2]
[1] *Omsk State Technical University, Omsk, Russia*
[2] *L.N. Gumilyov Eurasian National University, Astana, Kazakhstan*
* *e-mail:irina.kz-89@mail.ru*

Key words: potato, superoxidedismutase, catalase, peroxidase, salt stress, viruses

Motivation and Aim: Potato is one of the important crops in the world. Many different viruses that can reduce yield and tuber quality can infect it. Infection of potato with virus and salinity stress leads to the generation of reactive forms of oxygen (ROS) and the development of an "oxidative burst". The high amount of ROS production is correlated by an equally high activation of antioxidant enzymes [1]. In some potato varieties, the levels of catalase, peroxidase and superoxide dismutase are enhanced with salinity, and it might alleviate the adverse effect of stress in potato [2].
Methods and Algorithms: The activity of catalase, superoxide dismutase and peroxidase were determined by spectrophotometric methods. Determination of the activity of antioxidant enzymes was carried out by native gel electrophoresis.
Results: The analysis of recent research investigation on the functional interaction of superoxide dismutase, catalase, peroxidase under the influence of viral infection and salinity stress significantly expanded the understanding of the mechanisms of early potato plant responses to stressful environmental conditions. Inoculation of potato with viruses Y and S combination with salinity (10, 20, 30, 40, 50 mM NaCl) stress lead to the high superoxide dismutase, catalase, peroxidase activities in potato.
Conclusion: Under salinity stress and Y and S virus infection the level of superoxide dismutase in Y-virus infected potato increased up to 31 % , in S-virus infected plants it enhanced up to 27 %, catalase level in Y-virus plants increased up to 43 %, in S-virus infected plants – 67 %, peroxidase activity in Y-virus infected plants was 33 %, in S-virus – 29 %. The results suggest that the Y-virus is more aggressive in potato plants and has a huge influence to immune response in potato under salinity stress, especially in early varieties of potato.

References

1. Yergaliyev T.M. et al. (2016) The involvement of ROS producing aldehyde oxidase in plant response to Tombusvirus infection. Plant Physiology Biochemistry. 109:36-44.
2. Aghaei K. et al. (2009) Potato responds to salt stress by increased activity of antioxidant enzymes. Journal Integrative Plant Biology. 51(12):1095-1103.

DOI 10.18699/BGRSSB-2018-143

Impact of DNA reorganization on recombination suppression of 5b chromosome during wheat evolution

A.A. Kiseleva[1]*, A.F. Muterko[1], Z. Frenkel[2], A. Korol[2], E.A. Salina[1]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Institute of Evolution, University of Haifa, Haifa, Israel*
** e-mail: antkiseleva@bionet.nsc.ru*

Key words: wheat, genetic maps, evolution

Motivation and Aim: A comprehensive study of the evolution of wheat chromosome 5B based on the genetic and physical mapping data and analysis of the primary structure of the chromosome regions that influence the ability of this chromosome to moderate recombination in interspecific *T. aestivum* × *T. dicoccoides* crosses is described.
Methods and Algorithms: Genetic maps were constructed using the MultiPoint UltraDense software. Consensus maps were developed using the "MapFuser" R package. Comparative analysis of DNA was performed using UGENE, "Tandem Repeats Finder" and MAFFT 7.311 software. In order to identify genes and mobile genetic elements, the ENA database of plant protein coding sequences and MIPS Repeat Element Database were used respectively. Similarity searching was performed using BLASTn.
Results: The performed comparative genetic analysis comprises construction of the 5B map using a cross of CS and CS-5Bdic and comparison of the resulting map to the 5B chromosome maps of *T. durum* × *T. dicoccoides*, *T. durum* × *T. durum* and the maps built based on several crosses of common wheat cultivars using the same approaches. The regions of considerable recombination suppression on the 5B chromosome maps of *durum*/*dicoccoides* and *aestivum*/*aestivum* have been identified by comparing to CS/CS-5Bdic and can be explained by the divergence between the analyzed tetraploid and hexaploid wheat accessions. The same suppression regions as in CS and CS-5Bdic have been observed in all three studied populations. These regions of the 5B chromosome are actually the sites of rearrangements involved in the tetraploid–hexaploid evolutionary transition. No suppression events have been undetectable in the region of the Ph locus, the emergence of which was associated with the evolutionary stage prior to or during formation of the wild emmer. To unravel the causes of recombination suppression, the primary structure of the detected recombination suppression regions of 5B *T. aestivum* and *T. dicoccoides* chromosomes has been analyzed. In this analysis the inversion, large insertions/deletions, and extended clusters of 119.2 tandem repeats were found, which differ studied homoeologous chromosomal regions of both species.
Conclusion: To study the causes of recombination suppression, the analysis of the primary structure of the discussed regions was performed and the mechanisms underlying their contribution of the detected inversions, large insertions/deletions, and extended clusters of 119.2 tandem repeats to recombination suppression in 5BS are discussed. The results suggest that the detected rearrangements happened during the divergence of tetraploid emmer wheat.
Acknowledgements: This study was supported by the RSF (No. 14-14-00161).

DOI 10.18699/BGRSSB-2018-144

Detection of genes involved in regulation of wheat flowering time

A.A. Kiseleva*, A.B. Shcherban, I.N. Leonova, E.A. Salina
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: antkiseleva@bionet.nsc.ru*

Key words: wheat, flowering time, SNP, putative transcription factors

Motivation and Aim: Variability of flowering time may assist in wheat adaptation to local environments. Thereafter, discovery of new flowering time determinants is important for cereal improvement. In this study, we used common wheat cultivar Chinese Spring (CS) and the substitution line of CS with 5B chromosome from *T. dicoccoides* (CS-5Bdic), different in their flowering time by two weeks, to detect determinants of flowering time on 5B chromosome.
Methods and Algorithms: To ascertain the loci determining flowering time difference, a set of 116 recombinant inbred 5B chromosomal lines as a result of hybridization of CS with CS-5Bdic were developed and their flowering time was estimated. Genotyping was performed using the Illumina Infinium 15k Wheat platform and a set of SSR markers. Genetic linkage map was developed using the MultiPoint UltraDense software.
Using the QTL-analysis locus on 5B chromosome was revealed. To detect candidate genes functional annotation of genes, associated with SNP markers, based on synteny with other plant species was performed.
Results: Phenotype (flowering time) – genotype association analysis revealed 79 markers in pericentromeric region of 5B chromosome significantly associated with flowering time variation. Based on SNP sequences and synteny with crop genomes we identified the four best candidate genes: *WRKY*, *ERF/AP2*, *FHY3/FAR1* and *ELF4*, known to be involved in flowering time modulation in model species.
Conclusion: We identified the four best candidate genes: *WRKY*, *ERF/AP2*, *FHY3/FAR1* and *ELF4*. These genes were previously shown to be involved in flowering time modulation. We propose that the probable cause of flowering time differences may be due to differences in the origin of interacting flowering time pathways and putative transcription factors located on 5B might modulate. Contribution of *FHY3/FAR1* in flowering pathways was shown in further experiments.
Acknowledgements: This study was supported by the RSF (Project No. 14-14-00161).

DOI 10.18699/BGRSSB-2018-145

Global gene expression in organ-specific cold stress response in *Arabidopsis thaliana*

A. Klepikova[1, 2]*, I. Kulakovsky[3, 4], A. Kasianov[2, 3], M. Logacheva[2, 5, 6], A. Penin[1, 2, 5]
[1] *Institute for Information Transmission Problems RAS, Moscow, Russia*
[2] *Lomonosov Moscow State University, Moscow, Russia*
[3] *N.I. Vavilov Institute of General Genetics RAS, Moscow, Russia*
[4] *Engelhardt Institute of Molecular Biology RAS, Moscow, Russia*
[5] *Kazan Federal University, Kazan, Russia*
[6] *Skolkovo Institute of Science and Technology, Skolkovo, Russia*
* *e-mail: annklepikova@gmail.com*

Key words: calcium overload, rhythm disturbances, cardiac mechanics

Motivation and Aim: Plants face unavoidable environmental perturbations because of their sessile nature and have evolved various tolerance mechanisms to stress conditions. The process of increasing tolerance to freezing – cold acclimation – is well-studied in *Arabidopsis thaliana*. The main regulatory genes controlling cold response, such as *CBF* genes, were identified by classic genetics methods [1] and lots of transcriptomic studies were conducted on cold treated plants [2]. However, most of these papers were focused on whole aerial parts of plants or leaves and there have been no attempts to compare the expression profiles in different parts of plants during cold exposure

Methods and Algorithms: For this study we have chosen six most different biological samples of *A. thaliana*. The samples were collected from plants under normal condition as control and after 3 and 27 hours of cold (+4 °C) treatment. Expression features of cold response in various organs were analyzed.

Results: We have analyzed differential expression between cold-treated and control samples for each organ and different number of DE genes for each sample and 15,459 DE genes at all. The majority of the genes were not identified as stress-response in Gene Ontology (GO). Upregulated in all samples genes were enriched in stress-related GO terms and were "core" genes for cold resistance. Unique for samples genes were enriched in organ-specific manner. We found specific features of expression patterns in whole plant for organ-specific DE genes, such as activation of pollen-specific genes in leaves under cold treatment. We found new TFs to be involved in cold response.

Conclusion: Our results showed that the gene expression changes were highly organ-specific. When only a small number of cold-response genes were common in all samples acting as "core", there were a great variety of organ-specific genes with distinct expression specificities. Thus, while the mechanisms of cold stress response are common in all plants, in every organ they are modified in a unique fashion, including the recruitment of genes that are expressed in other organs in non-stress conditions. The results are summarized in the TraVa public database.

Acknowledgements: Supported by the RSF (project No. 14-50-00150).

References

1. Gilmour S.J. et al. (1998) Low temperature regulation of the Arabidopsis CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression. Plant J. 16:433-442.
2. Kilian J. et al. (2007) The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. Plant J. 50:347-363.

DOI 10.18699/BGRSSB-2018-146

Wheat ear recognizing algorithm for high throughput wheat phenotyping

E. Komyshev[1]*, M. Genaev[1], S. Tumanyan[1], N. Goncharov[1, 3], D. Afonnikov[1, 2], V. Koval[1]

[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Novosibirsk State Agrarian University, Novosibirsk, Russia*
** e-mail:komyshev@bionet.nsc.ru*

Key words: spike, wheat ear, image processing, plant phenotyping

Motivation and Aim: Qualitative and quantitative analysis of morphological characteristics of wheat yield is an important step in breeding studies and developing new high-yielding varieties. A complex analysis of morphological features allows a more detailed understanding plant biology. The structure of the spike is one of the most important trait in cereals. Traditionally, spike analysis performed manually by experts. It includes the definition of the spike type, awnedness, dimensions, color, number of grains, etc. Defining, documenting, storing, and organizing a lot of different parameters manually is a time-consuming task. Therefore, it is important to develop computational tools for fast and efficient morphometry of wheat traits.
Methods and Algorithms: In this work, we present methods for spike phenotyping using image processing. We developed an algorithm for recognizing spike on the image, estimation of its characteristics and evaluate models based on it. The algorithm recognizes a color checker with a scale and wheat spike on a blue background, defines spike contour and performs its morphometry. Input images were obtained using two different protocols. The first protocol involved the horizontal placement of the ear on the surface without the use of any additional fixture. This approach is simpler, but it does not allow to obtain images of the ear in all projections. The second approach involved the vertical position of the ear with a special tripod. This makes it possible to obtain images of the ear in 4 projections.
Results: The extracted morphometric parameters are used to construct several shape recognition algorithms. The involved parameters include an ear profile, its length, width, area, color, awnedness etc. We compared the obtained algorithms to assess their applicability in problems of automated classification and clustering. The constructed models include parameters such as the profile of spike contour, the Fourier transform of the contour profile parameters, the trapezoidal model, the model based on the distances from the point of the center of mass to the points of the contour.
Conclusion: Wheat spike recognition algorithm implemented on the Java programming language. The OpenCV library for image processing was used [1] for algorithms implementation. Input program data is image of wheat spike located on a blue background. Developed wheat spike recognition algorithm can be used for high-throughput phenotyping of wheat spikes.
Acknowledgements: Supported by the RSF (No. 17-74-10148).

References

1. Bradski G., Kaehler A. (2008) Learning OpenCV: Computer vision with the OpenCV library. O'Reilly Media, Inc.

DOI 10.18699/BGRSSB-2018-147

Sequencing and assembly of transcriptome in non-photosynthetic plant Lathraea squamaria and its dynamics during the light and dark period of the day

M.D. Logacheva[1, 2]*, A.V. Klepikova[3], M.I. Schelkunov[3], A.S. Kasianov[4], E.S. Gerasimov[1, 3], A.A. Penin[1, 3]

[1] *Lomonosov Moscow State University, Moscow, Russia*
[2] *Skolkovo Institute of Science and Technology, Moscow, Russia*
[3] *Institute for Information Transmission Problems, Moscow, Russia*
[4] *Vavilov Institute of General Genetics, Moscow, Russia*
* *maria.log@gmail.com*

Key words: non-photosynthetic plants, RNA-seq, circadian clock

Motivation and Aim: Photosynthesis is one of the most prominent features of plants. However, several species have adapted to parasitic lifestyle and lost ability to photosynthesis. Currently little is known about the genomic changes that mediate this adaptation. We studied this using as a model Lathraea squamaria, parasitic plant from the family Orobanchaceae, with focus on genetic networks that regulate plant interaction with light. As we have previously shown, L. squamaria is characterized by rather recent switch to heterotrophy and thus allows inferring first stages of heterotrophy-related genome modification.

Methods and Algorithm: Plant material was collected in two replicates at 15 and 18 hours and then to the midnight with a 1-hour intervals. In order to represent non-polyadenylated transcripts total RNA after depletion of rRNA was used for library construction. Libraries were sequenced with 150 + 150 bp length on Nextseq (Illumina) platform. Assembly and gene expression analysis was performed using CLC Genomics Workbench; transcripts were annotated using blast2go.

Results: As expected, the transcripts for most proteins involved in photosynthesis (components of photosystems I and II, cytochrome b6/f complex, light-harvesting complex, rubisco) are either absent or represented by expressed pseudogenes (in particular, plastid-encoded components of electron transfer chain). In contrast, the transcripts encoding proteins that are involved in non-photosynthetic interaction with light (e. g. phytochromes, cryptochromes, circadian clock associated proteins) are present. We found that several groups of genes significantly increase their expression after the sunset. Unexpectedly among them, there are the genes involved in cell division and translation.

Conclusion: The results obtained suggest that non-photosynthetic plants, even those that recently transitioned to heterotrophy, the genome undergoes profound changes that result in the loss (physical loss of gene or the downregulation of expression) of photosynthesis-related genes. Other aspects of interaction with light, in particular those that are involved in photomorphogenesis, are likely to remain unchanged. Also, novel mechanisms, not typical for photosynthetic plants, exist in heterotrophic plants; they presumably reflect the interaction of parasite with its host plant.

DOI 10.18699/BGRSSB-2018-148

De nova genome sequence of Karnal bunt pathogen (*Telitia indica*) of wheat

S. Marla[3]*, A. Kumar[1], V. Pandey[1], M. Singh[1], M.S. Saharan[2], P. Mishra, R. Maurya
[1] *G.B. Pant University of Agriculture & Technology, Pantnagar, India*
[2] *IARI, New Delhi, India*
[3] *ICAR.NBPGR, New Delhi, India*
** soma.marla@icar.gov.in*

Key words: wheat, genome, sequencing, disease, Karnal bunt

Karnal bunt of wheat is a serious quarantine disease caused by hemibiotrophic fungus, *Tilletia indica*. Despite its economic importance, little knowledge is known about molecular pathogenesis including various pathogenic determinants and avirulnce factors. Besides studies are complicated as reported genes of *T. indica* do not share high degree of homology with other closest basidiomycete fungi. *T. indica* genome was sequenced employing hybrid approach of PacBio Single Molecule Real Time (SMRT) and Illumina HiSEQ 2000 sequencing platforms (1). The genome was assembled into 10,957 contigs, 7,87 scafolds with total size of 31.83 Mbp. We predicted 11,535 putative genes, which were annotated employing Gene Ontology databases. The improved draft version of the assembly was achieved using Metaassembler, (v.1.5), by merging draft monotelesporic sequence-based assemblies from DAOM 236416 and RAKB_UP_1 isolates with the improved and reassembled hybrid assembly. Repeated genome assembly, gap filling and polishing on merged assembly employing Gapfiller tool (V.1.10) resulted in an improved coverage of 107 x. Secretome analysis of pathogen and the predicted putative Functional annotation of Karnal bunt pathogen genome and classification of identified effectors into protein families revealed interesting functions related to pathogenesis. Several biological, cellular and molecular functions were detected that include some related to pathogenesis, which include mating, zoospore development, host surface attachment, cell wall degrading enzyme, translocation and several others. Work is in progress to oomprove genome coverage and identification of potential effectors that could serve as molecular targets for development of diagnostic markers and new fungicide markers.

Acknowledgements: We gratefully acknowledge the support of Department of Biotechnology, Government of India Grant. BT/PR14987/AGR/21/914/2015), administration and Division of Genomic resources, ICAR, NBPGR, New Delhi, G.B. Pant University of Agriculture & Technology, Pantnagar, India for providing necessary laboratory facilities.

Reference

1. Kumar A., Mishra P., Maurya R., Mishra A.K., Vijai K., Pramod W., Marla S. (2018) Improved draft genome sequence of a monoteliosporic culture of the karnal bunt (*Tilletia Indica*) pathogen of wheat. Genome Announc. 6:e00015-18. https://doi.org/10.1128/genomeA.00015-18.

DOI 10.18699/BGRSSB-2018-149

Genome wide analysis of quantitative disease resistance against Verticillium wilt in the model legume *Medicago truncatula*

M. Mazurier[1], M. Toueni[1], A. Sbeiti[1], M.C. Tardin[2], M.C. Gras[2], M. Rickauer[1], L. Gentzbitte[l], C. Ben[1]*

[1] *EcoLab, Université de Toulouse, CNRS, Toulouse INP, UPS, Toulouse, France*

[2] *R2n – Centre de Recherche Le Bourg, Druelle, France*

** e-mail: ben@ensat.fr*

Key words: GWAS, quantitative trait loci (QTL), candidate gene, root disease, biodiversity, functional validation, biotechnology

Motivation and Aim: Legumes are one of the most important crop families with high nutritional value thanks to their nitrogen-fixing symbiosis with rhizobia. However, they are prone to many diseases which reduce yields. The interaction between the model legume *Medicago truncatula* and the root pathogen *Verticillium alfalfae* is studied to investigate the genetic mechanisms involved in quantitative disease resistance against Verticillium wilt in Legumes.

Methods and Algorithms: Genome-wide association studies (GWAS) were conducted on disease parameters assessing disease development and plant colonization in a collection of 242 *M. truncatula* accessions. The disease parameters were obtained after fitting non-linear mixed models to the time-course disease symptoms. The analyses on adjusted means from an augmented-block experimental design were performed with the TASSEL v5 software by implementing a linear mixed model including kinship as the variance-covariance matrix of the random genetic effects, and population structure as fixed effects (EMMAX algorithm). A set of 5M SNPs was used.

Results: Symptom scoring and fungus reisolation in 242 *M. truncatula* ecotypes highlighted a large biodiversity of the response to *V. alfalfae* (*Va* V31-2 strain) with a continuous range from fully resistant to susceptible lines. The genome wide association study (GWAS) on various modeled disease parameters pinpointed quantitative trait loci (QTL) on chromosome 1, 7 and 8. Both phenotypic and genetic analyses thus suggest the occurrence of different resistance mechanisms in *M. truncatula* populations towards *V. alfalfae.* Among five candidate genes localized under a previously described major QTL for *V. alfalfae* resistance on chromosome 7 [1], only two genes, encoding for proteins involved in ubiquitination and lipid metabolism, were found to be expressed in response to *V. alfalfae.* The latter one was validated as a disease susceptibility gene towards *Va* V31.2 in *M. truncatula*. Gene silencing using artificial microRNA in A17 (resistant) and F83005.5 (susceptible) lines decreases *Va* V31-2 colonization rate on transgenic roots whereas overexpression of the gene increases the colonization rate in A17.

Conclusion: Understanding Verticillium resistance and its genetic control in the model legume *M. truncatula* will help to develop breeding strategies for legume plants which are important components of sustainable agriculture and ecosystems.

Acknowledgements: M. Mazurier was supported by a PhD scholarship from the French "Ministère de la Recherche et de l'Enseignement Supérieur», M. Toueni by a doctoral grant from Mayotte department, and A. Sbeiti by a PhD grant from "Association des Jeunes de KFARSIR", managed by Toufik Sbeiti et Mariam Malek, Lebanon. We thank J.M. Prosperi for providing seeds and maintaining a large part of the *Medicago truncatula* collections. We thank the Medicago Hapmap international consortium, mainly funded by the Plant Genome Program of the National Science Foundation (USA), for providing *M. truncatula* genome resequencing data.

References

1. Ben C. et al. (2013) Natural diversity in the model legume Medicago truncatula allows identifying distinct genetic mechanisms conferring partial resistance to Verticillium wilt. J Exp Bot. 64(1):317-332.

DOI 10.18699/BGRSSB-2018-150

MetaRE: search for cis-regulatory elements via meta-analysis of transcriptomic data

D.D. Novikova[1, 2, 3]*, P.A. Cherenkov[1], K.U. Tkachev[2], V.G. Levitsky[1, 2], V.V. Mironova[1, 2]

[1] *Novosibirsk State University, Novosibirsk, Russia*

[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

[3] *Department of Agrotechnology and Food Sciences, Subdivision Biochemistry, Wageningen University and Research Center, Wageningen, Netherlands*

** e-mail: novikovadd@bionet.nsc.ru*

Key words: transcriptomes, gene expression, transcription factors, regulatory elements

Meta-analysis of biological data is becoming more and more popular. This tendency is due to accumulation of the big data in omics, like transcriptome, metabolome and proteome profiles, etc. More than two decades passed from uprise of whole-genome expression profiling methods. Nowadays thousands of transcriptomes are publicly available. Moreover, in many cases several experiments on the same phenomena could be found, providing plenty of material for the meta-analysis. One of the tasks is identification of cis-regulatory elements associated with expression changes in response to a stimulus, which allows to study complicated processes controlled by a set of transcription factors (TFs). We developed an R package MetaRE which provides search for the cis-regulatory elements enriched in the promoters of differentially expressed genes (DEGs) in response to a stimulus [1]. Via meta-analysis of multiple expression profiles, cis-regulatory elements are identified as associated with gene up- or down-regulation in response to a stimulus. Limma, edgeR and GEOquery are built-in in the package to search for DEGs in the transcriptome experiments from GEO database. C++ used to fasten slow components with the built-in Rcpp package to integrate the C++ code into R.

MetaRE package could be applied to any organism with sequenced genome, for which a number transcriptome experiments on a particular topic are available. The package was tested on Arabidopsis and Zebrafish to identify the cis-elements associated with response to auxin [1] and cold stress, correspondingly.

Acknowledgements: The research was supported by President Foundation (MK-1297.2017.4) and Russian Foundation of Basic Research (18-04-01130).

References

1. Cherenkov P. et al. (2017) Diversity of cis-regulatory elements associated with auxin response in Arabidopsis thaliana. Journal Experimental Botany. 69(2):329-339.

DOI 10.18699/BGRSSB-2018-151

Study of the effect of nanocomposites based on humic substances of different nature on the causative agent of ring rot of *Clavibacter michiganensis* ssp. *sepedonicus* potato plants

O.A. Nozhkina[1]*, A.I. Perfilyeva[1], I.A. Graskova[1], B.G. Sukhov[2]
[1] *Siberian Institute of Physiology and Biochemistry of Plants SB RAS, Irkutsk, Russia*
[2] *Irkutsk Institute of Chemistry of A.E. Favorskii SB RAS, Irkutsk, Russia*
** e-mail: alla.light@mail.ru, smallolga@mail.ru*

Key words: silver nanocomposites, ring decay, humic substances, activity level peroxydase

Motivation and Task: Plants of potatoes transfer the dangerous bakterilny disease caused *Clavibacter michiganensis* ssp. *sepedonicus*. Bacteria complicate food of plants, causing withering and death, thereby bring to a yield loss. There are no ecologically safe measures for fight against this disease, all measures come down only to disinfecting. Therefore, there was a need for search of safe and effective remedies with a disease. Argentiferous nanocomposites (NC) based on natural components of various humic substances (HS) can be such substances. They not toxicity, are steady for a long time, are factors of growth of plants and a soil-forming biota, etc.
Methods and Algorithms: Were investigated AC 1405 strain bacteria, three types of the NC silver packed into HS and their predecessors by HS (NCHS – dirt/$AgNO_3$, NCNS – coals/$AgNO_3$, NCNS – slates/$AgNO_3$, HS – dirt, HS – coals, HS – slates and nitrate of silver) and plants of potatoes of a grade of Lukyanovsky. All substances were synthesized at the Irkutsk institute of chemistry by it A.E. Favorskii, soluble in water is good and the water decisions are convenient in use. Bacteria grew up 2–3 days, after incubated the NC and its predecessors. Studying of bactericidal effect of the NC and formation of the biomovie is executed by measurement of optical density of bacterial suspension. plants grew up on nutrient medium of Murasige-Skuga in faktorostatnykh conditions within 14 days. Further NC and HS incubated. Each 2 days took measurements of biometric parameters of plants and activity of peroxidase.
Results and Discussion: The bakteriostatic effect was shown by the predecessor of NC - $AgNO_3$ considerably suppressed growth of bacteria from the beginning of the experiment. So, HS-cl and HS-sl and their NC inhibited growth of bacteria and ability of a biofilm formation. HS-dt and its NC stimulated reproduction of bacteria, and reduced formation of biofilms. HS-cl stimulated reproduction of a bacterium. After processing of plants of the NC and their predecessors, was analyzed peroxidase in potatoes plants. that was revealed $AgNO_3$, HS-dt, HS-sl, NCHS-dt/$AgNO_3$, NCHS-sl/$AgNO_3$ more than twice reduced activity of peroxidase. HS-cl and NCHS-cl/$AgNO_3$ stimulated activity of peroxidase of potatoes and on biofilm formation of bacteria.
Conclusion: Studying influence of antibacterial activity of NC on Cms, showed us that use of NC based on HS as safe means of fight against a bacterial disease of plants of potatoes is possible.
Acknowledgements: Work is performed with financial support of the Russian Federal Property Fund and Government of the Irkutsk region (project No. 17-416-380001 of a r_).

DOI 10.18699/BGRSSB-2018-152

Mathematical modeling of chilling stress induced changes in *Arabidopsis thaliana* root meristem

M.S. Savina[1, 2]*, J.H. Hong[3], Jian Xu[3], V.V. Mironova[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Department of Biological Sciences and Centre for BioImaging Sciences, National University of Singapore, Singapore, Singapore*
** e-mail: savina@bionet.nsc.ru*

Key words: mathematical modeling, developmental biology, chilling stress, hormone regulation

Motivation and Aim: Plant hormone auxin guides many physiological processes, including of stem cell niche maintenance in a changing environments. In Arabidopsis, chilling stress (24 h under 4 °C) leads to DNA damage predominantly in root stem cells and their early descendants. However, only newly generated/differentiating columella stem cell daughters (CSCDs) preferentially die in a programmed manner. Inhibition of the DNA damage response in these CSCDs prevents their death but makes the stem cell niche more vulnerable to chilling stress. We studied the protective effect of CSCD death on the stem cell niche under chilling stress in a mathematical model.
Methods and Algorithms: As a readout, we analyzed the expression levels of DR5::GFP, PIN::PIN-GFP, and PIN::GUS at 4 and 22 °C. Chilling stress significantly alters expression of auxin transporters and lead to decrease in auxin response in the QC. For this study, we carry out computer modeling of the auxin distribution taking into account the positive and negative feedbacks between the auxin and its PIN transporter in the Arabidopsis thaliana root tip [1].
Results: Consistent with the experimental findings, in silico analysis showed that differential changes in the expression of PINs led to a new steady-state equilibrium of auxin distribution in chilling-stressed root, despite an overall decline in auxin levels. In this new steady state, however, the division of CSCs caused a loss of auxin maximum in the QC, which could be restored only if the death of newly generated CSCDs occurred. By contrast, CSC division at normal temperature had no effect on the maintenance of auxin maximum in the QC. We thus concluded that CSCD death was the strategy used by the root to sustain the auxin maximum.
Conclusion: In agreement with our model prediction, that CSCD death increases the auxin concentration in the QC, roots with chilling stress-induced CSCD death displayed a higher DR5::GFP and WOX5::GFP expression in the QC than those without. Together, our findings demonstrate the importance of auxin level in the protection of root stem cell integrity and indicate that chilling stress-induced death of CSCDs results in an increase of auxin levels in the root stem cell niche, which helps prepare the root to withstand the accompanying environmental stresses and to recover faster when returned to optimal temperatures [2].
Acknowledgements: The work was supported by MK-1297.2017.4.

References

1. Mironova V.V. et al. (2012) Combined in silico/in vivo analysis of mechanisms providing for root apical meristem self-organization and maintenance. Annals of Botany. 110(2):349-360.
2. Hong J.H., Savina M. et al. (2017) A sacrifice-for-survival mechanism protects root stem cell niche from chilling stress. Cell. 170(1):102-113.

DOI 10.18699/BGRSSB-2018-153

Establishment of a regional centre for DNA-barcoding of rare and endangered plant species based on the DNA bank of the Republic of Belarus

N. Savina[1]*, S. Kubrak[1], E. Kuzminova[1], E. Mikhalenko[1], L. Milko[1], A. Kolbas[2], N. Matusevich[2], E. Makeyeva[1] , A. Kilchevskiy[1]
[1] *Institute of Genetics and Cytology NASB, Minsk, Belarus*
[2] *Brest State University named after A.S. Pushkin, Brest, Belarus*
** e-mail: N.Savina@igc.by*

Key words: DNA-barcoding, ITS2, rbcL, psbA-trnH

Motivation and Aim: Rare and endangered plant species are characterized by a lower ability to survive in the face of climate change and the pressure of anthropogenic factors, leading to loss of valuable genotypes and a decrease in biodiversity in general. In view of this, a unit "DNA Bank of Rare and Endangered Species of Wild Flora and Fauna" was established at the Republican DNA Bank of a human, animals, plants and microorganisms for comprehensive study and conservation of biological diversity of rare and endangered species listed in the Red Book of the Republic of Belarus.
Methods and Algorithms: The employees of Brest State University collect biomaterial of rare plants and compile a Database "Cartographic Web-application of rare plants' populations in Brest Region" that contains a morphological, ecological and geographic description of every plant [1]. The staff of the Institute of Genetics and Cytology, NAS of Belarus, perform identification of the collected material using a DNA-barcoding technique. The information obtained is entered in the Database of the Republican DNA Bank of a human, animals, plants and microorganisms.
Results: To date, 35 rare and endangered plant species (I-IV protection categories) collected in the National Parks "Narochansky" and "Belovezhskaya Pushcha" have been examined. DNA-barcoding was carried out using ITS2, rbcL, psbA-trnH marker sequences. The resolving power of rbcL and psbA-trnH chloroplast markers proved insufficient for independent use as barcodes. At the same time, the use of ITS2 nuclear region significantly increased the efficacy of species identification of rare plants both for independent use and along with rbcL and psbA-trnH markers.
Conclusion: The obtained results allow to use DNA-barcoding as an ecologic monitoring tool for wildlife and in nature conservation activity. It is planned to hold a training course for Central and Eastern Europe and Central Asia's specialists on DNA-barcoding use and to establish a Regional Centre for DNA-barcoding of rare and endangered plant species.
Acknowledgements: Supported by the Presidium of NASB (No. 629 of August 13, 2016).

References

1. http://arcg.is/2hssjpw .

DOI 10.18699/BGRSSB-2018-154

The novel tandem repeat of 646 BP identifies the subtelomeric region of wheat 5BS chromosome

E.M. Sergeeva*, I.G. Adonina, M.A. Nesterov, E.A. Salina
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: sergeeva@bionet.nsc.ru*

Key words: BAC FISH, *Triticum aestivum*, *Triticum dicoccoides*, tandem repeat

Motivation and Aim: For study of plant chromosomes the fluorescence *in situ* hybridization with large inserts of genomic DNA cloned in Bacterial Artificial Chromosome (BAC-FISH) is broadly used. Most often the BAC-FISH on wheat chromosomes demonstrates the dispersed pattern because of high mobile elements content [1, 2]. The most specific observed BAC-FISH pattern on wheat chromosomes was the distinct signals on several chromosomal pairs [1]. Our aim was to identify 5B chromosome-specific DNA sequence for study of reorganization of this chromosome during evolution of hexaploid *T. aestivum* and tetraploid *T. dicoccoides* wheats.
Methods and Algorithms: The 20 BAC-clones were randomly selected from 5B chromosome specific BAC-library of *Triticum aestivum* cv. Chinese spring. BAC-FISH on *T. aestivum* (AABBDD, $2n = 6x = 42$) and *T. dicoccoides* (AABB, $2n = 4x = 28$) chromosomes was carried out according to [3]. BAC-end sequences were obtained with Sanger method and used for BLASTn search over the reference genomic sequences: *T. aestivum* cv. CS RefSeq v1.0 (International Wheat Genome Sequencing Consortium repository) and publicly available sequence of *T. dicoccoides* cv. Zavitan. The isolated regions were annotated using BLASTn with TREP and NCBI databases.
Results: The clone 030N24 demonstrated the spot FISH signal in the subtelomeric part of *T. aestivum* and *T. dicoccoides* chromosomes. The BLASTn search with the BAC-end sequences of 443 and 604 bp defined the region of 93,576 bp on the distal part of *T. aestivum* 5BS chromosome. The region contains the cluster of 10,060-bp length formed by 16 units of tandem repeat (unit length is 646 bp). In other regions of the *T. aestivum* genome these repeats were absent. The corresponding region on *T. dicoccoides* chromosome also localized at the distal part and contains the 31 units of 646-bp tandem repeat.
Conclusion: A 5B-chromosome specific region which represents a 10,060 bp cluster formed by 646-bp tandem repeats was identified. The probability of its inheritance from a diploid progenitor or its amplification during the formation of the first allotetraploid of the Emmer group *T. dicoccoides* is discussed.
Acknowledgements: Supported by RSF No. 14-14-00161.

References

1. Zhang P. et al. (2004) BAC-FISH in wheat identifies chromosome landmarks consisting of different types of transposable elements. Chromosoma. 112(6):288-99.
2. Salina E.A. et al. (2011) The impact of Ty3-gypsy group LTR retrotransposons Fatima on B-genome specificity of polyploid wheats. BMC Plant Biology. 11:99.
3. Salina E.A. et al. (2006) Phylogenetic reconstruction of Aegilops section Sitopsis and the evolution of tandem repeats in the diploids and derived wheat polyploids. Genome. 49(8):1023-1035.

DOI 10.18699/BGRSSB-2018-155

The evaluation of reproduction type of *Puccinia graminis* f. sp. *tritici* population prevailing in West Siberia

E.M. Sergeeva[1]*, E.S. Skolotneva[1], V.N. Kelbin[1], S.I. Konkova[2], M.A. Nesterov[1], E.A. Salina[1]

[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

[2] *Novosibirsk State University, Novosibirsk, Russia*

** e-mail: sergeeva@bionet.nsc.ru*

Key words: stem rust, bread wheat

Motivation and Aim: Stem rust is a fungal disease of grasses (including agronomically important bread wheat *Triticum aestivum* L.) caused by *Puccinia graminis*, resulting in a significant decrease in grain production. The species *P. graminis* subdivides into specialized forms (f. sp.) passed through the adaptation to the specific range of host plants. The sexual process is optional for the fungal life cycle but leads to novel combinations of virulent genes as well as determines high genetic variability. In the case of sexual reproduction, an alternative host plant barberry *Berberis vulgaris* is served. The aim of the study is to establish the reproduction type (sexual/asexual) of the West Siberian population of *P. graminis* f sp. *tritici* occured on wheat.

Methods and Algorithms: The samples of infected leaves of barberry and stems of grasses (*Triticum aestivum*, *Elytrigia repens* and *Dactylis glomerata*) were collected in Novosibirsk area during June-August in 2016 and 2017. DNA was extracted according to modified CTAB method [1]. First to confirm the identity to *Puccinia* species the 45S rDNA internal transcribed spacer (ITS) regions were amplified and sequenced from the samples with the rust-specific primers pair ITS1RustF10d/ITS1rustR3c [2]. Then the samples were analyzed by PCR using the set of 16 publicly available SSR markers developed for *P. graminis*.

Results: The ITS analysis confirmed the stem rust pathogen in all 15 samples. The SSRs analysis showed the differences between the amplification spectra of *P. graminis* from barberry and wheat. The SSRs patterns from barberry and other grasses were identical. This means the local barberry is alternative host not for f. sp. *tritici* but for other special form of the pathogen. Thus, we suggest an asexual type of reproduction for wheat stem rust population prevailing in West Siberia.

Conclusion: Since asexual reproduction type was established for *P. graminis* f. sp. *tritici* population prevailing in West Siberia, the source of inoculum in the region was considered as exogenous. The virulence structure of asexual wheat stem rust population migrating over West Siberia as well as determining the source of inoculum require further study.

Acknowledgements: Supported by RFBR (No. 17-29-08018).

References

1. Justesen F. et al. (2002) The recent history of *Puccinia striiformis* f.sp. *tritici* in Denmark as revealed by disease incidence and AFLP markers. Plant Pathology. 51:13-23.
2. Barnes C.W., Szabo L.J. (2007) Detection and identification of four common rust pathogens of cereals and grasses using real-time polymerase chain reaction. Phytopathology. 97:717-727.

DOI 10.18699/BGRSSB-2018-156

Investigating genetic control of pigmentation in mutant barley lines with RNA-seq

N. Shmakov[1]*, A. Glagoleva[1, 2], A. Doroshkov[1], G. Vasiliev[1], N. Shatskaya[1], O. Shoeva[1], D. Afonnikov[1, 2], E. Khlestkina[1]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: shmakov@bionet.nsc.ru*

Key words: RNA-seq, transcriptomics, barley, de novo assembly, differential expression

Motivation and Aim: Plant pigments are compounds that colorize plant organs and tissues. The most important plant pigment is chlorophyll that takes part in photosynthesis – the process that impacts all life on earth. Phytomelanin is ather plant pigments that gives black coloration to certain plant organs. Phytomelanin containing barley lines are believed to be more resistant to fungal pathogens [1].
Plant nearly-isogenic lines (NILs) with abnormal pigmentation are an appropriate object to investigate genetic machinery behind plant pigments synthesis and distribution. Considering genetic similarity of NILs, RNA-seq method could be implemented to specifically discover genes responsible for plant pigmentation control.
Methods and Algorithms: Total RNA was extracted from plants of *Hordeum vulgare* (barley) NILs Bowman, i:Bw*Alm* with partial chlorophyll deficiency and BLP with partial melanism. Sequencing was carried out with IonTorrent platform. Libraries were filtered and quality-controlled with Prinseq tool. *De novo* assembly was performed using Trinity software. Assembled contigs were then aligned to *H. vulgare* reference genome. Fragments that showed no significant homology to the genome were functionally annotated using TransDecoder and Trinnotate tools.
Libraries were mapped to the reference genome using TopHat2 and STAR software. Differential expression was assessed with Cufflinks pipeline and EdgeR package. DEGs were functionally analyzed with AgriGO and PlantCyc databases.
Differential expression of a list of genes was verified with RT-qPCR. Expression of novel assembled genes was verified with PCR.
Results: Novel transcripts and lists of DEGs were discovered for each of the studied NILs. Genes up-regulated in i:Bw*Alm* line are involved in vesicular transport, while genes up-regulated in BLP line participate in several metabolic pathways, including cutin and suberin monomers biosynthesis. Novel genes predicted with *de novo* assembly include several disease resistance genes for line i:Bw*Alm* and autophagy-related gene for line BLP.
Conclusions: This work provides new insight on genetic mechanisms underlying partial albinism and melanism phenotypes formation in barley NILs. Our results demonstrate interaction between chloroplast and nuclear genomes in i:Bw*Alm* line and activation of several metabolic pathways in BLP line. Finally, we predict several genes not encountered in *H. vulgare* before.
Acknowledgements: The work is supported by the RSF (16-16-04073).

References

1. Zhou X., Chao M., Liang X. (1991) Screening and testing of barley varieties for scab resistance (in Chinese with English abstract). Acta Phytophylacica Sinica. 18:261-265.

DOI 10.18699/BGRSSB-2018-157

Regulation of metabolic pathway underlying anthocyanin pigmentation of barley pericarp

O. Shoeva[1]*, T. Kukoeva[1], E. Gordeeva[1], A. Glagoleva[1], E. Khlestkina[1, 2, 3]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *N.I. Vavilov All-Russian Research Institute of Plant Genetic Resources* (*VIR*)*, Saint-Petersburg, Russia*
* *e-mail: olesya_ter@bionet.nsc.ru*

Key words: *Hordeum vulgare*, pericarp, structural genes, regulatory genes, transcription

Motivation and Aim: Anthocyanins are secondary metabolites determining pigmentation in plants. These molecules and their uncolored precursors have important physiological roles including regulation of plant growth and development and adaptation to stress [1]. Two groups of genes underlay anthocyanin biosynthesis (AB): structural genes, encoding enzymes of the pathway, and regulatory genes, encoding MYB, bHLH or WD40 transcriptional factors (TFs). The TFs form complex that activates transcription of the structural genes. The interaction between regulatory genes at transcriptional level has been recently observed in wheat [2]. The aim of the current study was to reveal possible interactions between the genes underlying the pigmentation in barley.
Methods and Algorithms: The *Ant1* (encodes the MYB TF, chromosome 7H) and *Ant2* (bHLH, 2H) genes, predetermining purple pigmentation of barley pericarp, were sequenced in Bowman near-isogenic lines (NILs) differing by grain color. Primers distinguishing the dominant and recessive alleles of the genes were designed by OLIGO [3] and used for marker-assisted selection of lines with different combinations of the *Ant* genes (*Ant1Ant1ant2ant2* and *ant1ant1Ant2Ant2*). Expression of the structural (*Chs*, *Chi*, *F3h*, *F3'h*, *Dfr*, *Ans*) and regulatory (*Ant1*, *Ant2*) genes in pericarp of the NILs was determined by qPCR or qRT-PCR.
Results: We showed that the dominant and recessive alleles of the *Ant* genes have differences in regulatory regions that prevent expression of the recessive alleles in non-colored grain pericarp. The NILs with different combination of the dominant and recessive *Ant1* and *Ant2* genes were developed based on the Bowman NILs. Transcription activity assay of the structural and regulatory genes in the NILs revealed that (1) the dominant alleles of *Ant1* and *Ant2* genes are required for activation of the structural genes expression in purple pericarp and (2) *Ant1* affects expression of *Ant2*: transcription of *Ant2* was up-regulated in the presence of dominant *Ant1* only, whereas *Ant2* was not expressed when *ant2* was recessive.
Conclusion: The data on revealed interaction between AB regulatory genes is base for further genetic regulatory network reconstruction.
Acknowledgements: Supported by the RFBR (No. 16-34-60052).

References

1. Lev-Yadun S., Gould K.S. (2008) Role of anthocyanins in plant defence. In: Winefield C., Davies K., Gould K. (eds), Anthocyanins. Springer New York. pp. 22-28.
2. Gordeeva E.I. et al. (2015) Marker-assisted development of bread wheat near-isogenic lines carrying various combinations of *Pp* (*purple pericarp*) alleles. Euphitica. 203:469-476.
3. Offerman J.D., Rychlik W. (2003) Oligo primer analysis software. In: Krawetz S.A., Womble D.D. (eds), Introduction to Bioinformatics: a Theoretical and Practical Approach. Humana Press; New Jersey, USA. pp. 345-361.

DOI 10.18699/BGRSSB-2018-158

Genome-wide association study between chickpea accessions from VIR collection and phenotypic data

A.B. Sokolkova[1]*, M.A. Vishnyakova[2], S. Bulyntsev[2], P.L. Chang[3, 4], N. Carrasquilla-Garcia[4], M.G. Samsonova[1], S.V. Nuzhdin[1, 3]

[1] *Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia*

[2] *Federal Research Centre All-Russian N.I. Vavilov Institute of Plant Genetic Resources* (*VIR*), *St. Petersburg, Russia*

[3] *University of Southern California, Los Angeles, CA, USA*

[4] *Department of Plant Pathology, University of California, Davis, CA, USA*

** e-mail: alyonasok@yandex.ru*

Key words: chickpea (*Cicer arietinum* L.), GWAS analysis, candidate genes

The Vavilov Institute of Plant Genetic Resources (VIR), in St. Petersburg, Russia, houses a unique genebank of chickpea (*Cicer arietinum* L.). Genotyping by sequencing of VIR's 428 chickpea accessions from different countries identified 56,855 segregating single nucleotide polymorphisms (SNP). For these SNP calls we implemented inclusion criteria: minor allele frequency (MAF) more than 3 %, genotype call-rate more than 90 %. In addition, we performed imputation of missing values. We performed Genome-wide association study (GWAS) to find associations (for data with and without imputation) between VIR's chickpea accessions and phenotypic data obtained in Kuban experimental station of VIR in 2016 and 2017. GWAS analysis identified a large number of genome intervals and potential gene candidates that may affect important agronomic traits. Besides that, we made a link between traits measured in GWAS and specific climatic variables at collection sites.

DOI 10.18699/BGRSSB-2018-159

Regulatory genes of anthocyanins biosynthesis in the barley grain

K.V. Strygina[1]*, E.K. Khlestkina[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: pushpandzhali@bionet.nsc.ru*

Key words: gene divergence, gene duplication, Hordeum, flavonoids biosynthesis, MBW, transcription factor

Motivation and Aim: Many higher plants including an important agricultural crop barley (*Hordeum vulgare* L.) synthesize secondary metabolites flavonoids. For activation of the structural flavonoid biosynthesis genes the expression of genes coding transcription transcription factors MYC, MYB and WD40 (forming the MBW regulatory complex) is necessary. The aim of this research was the identification, comparison and analysis of full-length sequences of duplicated MBW gene copies.
Methods and Algorithms: The search of homologous sequences was made in databases for not annotated barley sequences using BLAST. Cluster analysis using MEGA software was based on the UPGMA algorithm. Promoters of the genes were analyzed with PLACE. Primers design for PCR, RT-PCR, qRT-PCR and sequencing was performed using OLIGO. Genetic mapping was performed in DOMxREC mapping population using MAPMAKER program.
Results: Two *MYC* genes, three *MYB* genes and two *WD40* genes were found in barley genome. The MYC-coding gene *Myc2* was mapped precisely (tightly linked to marker XBmac186-4H). Exon-intron organization of all identified genes is similar to the structure of another MBW-coding genes in genomes of dicot and monocot plant species. The transcription activity of the detected genes in various parts of plant varied among the copies.
Conclusion: MYB, bHLH and WD40 genes, controlling the accumulation of blue anthocyanins in the aleurone layer, were first discovered. We showed that the main regulator of the blue pigmentation of the aleurone layer is the bHLH-encoding *Myc2* gene, localized on the chromosome 4HL. Identified major allele of the gene is common for the German and Siberian barley populations.
Acknowledgements: This study was supported by the Russian Science Foundation (No. 16-14-00086).

DOI 10.18699/BGRSSB-2018-160

Genes determining anthocyanin pigmentation in *Solanum tuberosum* L.

K.V. Strygina*, E.K. Khlestkina, A.V. Kochetov
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: pushpandzhali@bionet.nsc.ru*

Key words: gene duplication, flavonoids biosynthesis, MBW, Solanaceae, transcription factor

Motivation and Aim: Anthocyanin pigments are plant secondary metabolites, having multiple biological functions in plants including potato *Solanum tuberosum* L. Activation of anthocyanin synthesis occurs with MBW complex, which is formed by transcription factors MYB, bHLH and WD40. Certain genes and their allelic variants affecting phenotype (purple or red tuber skin, purple or red tuber flash, purple leaves, stems) have been not yet identified. The aim of our research is to find allelic differences underlying phenotypic variation by coloration traits in potato and to develop allele-specific DNA-markers for future markers-assisted potato breeding programs.
Methods and Algorithms: The search for sequences and genes structure identification was performed using the PGSC database. The search for functionally significant gene domains was carried out using InterPro. Development of diagnostic primers for the alleles determination as well as primers for the amplification of full-length transcripts of genes was performed using the OLIGO program. A total of 36 *S. tuberosum* L. varieties and hybrids were studied. Eight of them, differring by anthocyanin pigmentation, were selected for further analysis of the expression of anthocyanin biosynthesis genes. Molecular cloning was performed using a pDrive cloning vector.
Results: The information about known MYB-, bHLH- and WD40-encoding genes was exploited to develop allele-specific markers for analysis of the regulatory genes. Diagnostic DNA markers were developed for each of the identified genes and their allelic variants. Using RT-PCR on eight genotypes differing by anthocyanin coloration we compared activity of certain *MYB*, *bHLH* and *WD40* gene copies. The regulatory genes of *MYB* (*StAN1*, *StMYBA1*, *StMYB113*), *bHLH* (*StJAF13* and *StbHLH1*), *WD40* (StWD40) families were isolated. We showed that the main regulator of anthocyanin coloration in leaves and stems among the MYB-encoding genes is *StAN1*. The expression of other MYB-encoding genes as well as that of the bHLH- and WD40- encoding genes did not correlate with the potato coloration.
Conclusion: We have developed intragenic markers, by which differences in *StAN1* alleles can be detected. However, due to the high allelic diversity in this locus, the markers developed could not be used alone for prediction of potato stem and tuber anthocyanin coloration. To develop effective markers further analysis of nucleotide sequences in the promoter region and the identification of genetic polymorphisms association with the transcription level is required. Overall, the results of the study are important for understanding the mechanisms underlying tissue specific regulation of anthocyanin synthesis.
Acknowledgements: This study was supported by the Russian Science Foundation (No. 16-16-04073).

DOI 10.18699/BGRSSB-2018-161

Application of SRAP fingerprinting for analysis of *Lupinus angustifolius* cDNA

E. Sysoliatin*, N. Anisimova, A. Kilchevsky
Institute of Genetics and Cytology, NAS of Belarus, Minsk, Belarus
** e-mail: E.Sysoliatin@igc.by*

Key words: cDNA fingerprinting, differential expression, disease response

Motivation and Aim: Narrow-leaved lupine (*Lupinus angustifolius* L.) is a valuable legume crop capable of nitrogen fixation. There is a need to expand a list of molecular markers of the genes of economically valuable traits and to study the molecular mechanisms that determine the traits. SRAP (sequence-related amplified polymorphism) is based on the amplification of open reading frames. This method proved to be suitable in differential expression studies of various plant cultures [1]. The aim of this work was to assess the feasibility of using SRAP for cDNA analysis of narrow-leaved lupine samples.
Methods and Algorithms: RNA was isolated from the roots of ten-day seedlings of narrow-leaved lupine varieties Aschadny, Yorrel, Frost, Nemchinovsky 846. There were two groups of seedlings: intact and exposed to *Fusarium* culture. cDNA was synthesized on the RNA matrix and was used in a series of PCR reactions with 12 combinations of 3 forward (Me8, f12, f16) and 4 reverse (Em5, Em12, r14, r9) SRAP primers [2, 3]. SRAP fragments associated with resistant seedlings were isolated and sequenced. Their putative function was assigned based on alignment to known sequences and conserved domain search.
Results: The results showed that 12 primer combinations produced a total of 234 clear bands. The proportion of polymorphic bands varied in the range from 66,7 to 100 %. The SRAP fragment (f12-r14-290) associated with seedlings resistant to *Fusarium* isolates (Frost, Aschadny) were found. Fragment sequencing and subsequent analysis showed the presence of SNARE domain, which is characteristic of syntaxin-binding protein (STBP).
Conclusion: The results obtained indicate that the SRAP method is suitable for the analysis of narrow-leaved lupine cDNA and can be used to study the differential gene expression of this culture. The SRAP fragment associated with resistant seedlings showed the presence of SNARE domain, characteristic of STBP involved in exocytosis and mechanisms of plant nonspecific resistance to pathogens.
Acknowledgements: Supported by the SPSR “Biotechnology”, subprogram “Structural and functional genomics”, task: “Biotechnology 2.02. Evaluation of narrow-leaved and yellow lupine gene pools by resistance to fungal diseases based on DNA-marking”.

References

1. Que Y. et al. (2012) cDNA-SRAP and Its Application in Differential Gene Expression Analysis: A Case Study in *Erianthus arundinaceum*. Journal Biomedicine Biotechnology. Article ID 390107.
2. Mutlu N. et al. (2008) Development of SRAP, SRAP-RGA, RAPD and SCAR markers linked with a *Fusarium* wilt resistance gene in eggplant. Theor Appl Genet. 117(8):1303-1312.
3. Ma J.-X. et al. (2013) Genetic diversity of wild *Medicago sativa* by sequence-related amplified polymorphism markers in Xingjiang region, China. Pak. J. Bot. 45(6):2043-2050.

DOI 10.18699/BGRSSB-2018-162

Algorithms for prediction and analysis of regulatory regions

T. Tatarinova[1]*, V. Solovyev[2], N. Alexandrov[3], A. Baranova[4], A. Kel[5, 6]
[1] *University of La Verne, La Verne, CA, USA*
[2] *Softberry, Inc. Mount Kisco, NY, USA*
[3] *Inari, Inc, Boston, MA, USA*
[4] *George Mason University, VA, USA*
[5] *geneXplain GmbH, Wolfenbuettel, Germany*
[6] *Institute of Chemical Biology and Fundamental Medicine, Novosibirsk, Russia*
** e-mail: ttatarinova@laverne.edu*

Key words: promoter, deep learning

Motivation and Aim: They say that the chain is only as good as its weakest link. Current methods of genome annotation are capable of accurate prediction of coding regions, but are failing in promoter prediction. Accurate identification of transcription start sites and core promoter regions remains an unsolved problem.
Methods and Algorithms: We will present a comprehensive analysis of genomic features associated with promoters in several plant genomes, and demonstrate how probabilistic integrative algorithms succeed in accurate prediction of transcription start sites. We developed models using distributions of sequence polymorphisms, RNA sequencing reads on genomic DNA, methylated nucleotides, transcription factor binding sites, as well as relative frequencies of nucleotides and their combinations.
Results: Accuracies of promoter-prediction methods differ between species and functional classes of genes, and we will present an approach to select the optimal method for promoter prediction for a studied genome. We have identified three distinct classes of TFBS that show different positional preference with respect to TFBS. We have also demonstrated evolutionary conservation of distribution of TFBS between plant species.

DOI 10.18699/BGRSSB-2018-163

Discordant evolution of YUCCA family proteins demonstrated along sequence

I. Turnaev[1]*, K. Gunbin[1], V. Suslov[1], D. Afonnikov[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: turn@bionet.nsc.ru*

Key words: auxin biosynthesis pathway, horizontal gene transfer, discordant evolution

Motivation and Aim: The origin of the tryptophan dependent pathway of auxin biosynthesis in terrestrial plants is a subject for discussion. The answer to this question related to the evolution of two pathway's enzymes, TAA and YUCCA. Here we perform phylogenetic analysis of the evolution of the YUCCA protein family in plants.
Methods and Algorithms: Proteins of YUCCA family and their closest homologues were extracted from NCBI database using protein BLAST with a threshold e-value < 1e–5. Homologs of protein YUC2 (AT4G13260) of *Arabidopsis thaliana* were the BLAST queries; only full-size proteins were taking into account. Promals used for alignment reconstruction. The phylogenetic tree was built using different methods: Maximum Likelihood, Bayesian, Approximated Maximum Likelihood (AML). To assess the degree of anomaly of evolution (the degree of qualitative differences between the topologies of phylogenetic trees of various alignment fragments and known species tree) along the alignment we reconstructed phylogenetic trees for the alignment fragments within 60 aa sliding windows. We group obtained phylogenetic trees by their sliding window position from 1 to 50, 25 to 75, 50–100 etc. For each of these groups we build a phylogenetic network with the SplitsTree program. We compared network for these sequence segments with the species tree topology and identified the deviation from the species tree.
Results: As result, we identified two regions with discordant evolution incompatible with the species tree topology: aa 95–185 and 210–280. Phylogeny for those regions of the YUCCA protein is significantly different from the known species phylogeny. We propose that the cause of this discordancy may be the horizontal transfer events of fragments of the coding part of genes from non-plant taxa (perhaps symbionts of ferns or gymnosperms plants) to the ancestral proteins of the subfamily YUC10-11 terrestrial plants in the above-named regions of protein.
Acknowledgements: The work was supported by ICG SB RAS budget project (No. 0324-2018-0017).

DOI 10.18699/BGRSSB-2018-164

Dissecting the mechanisms of EIN3-dependent regulation of ethylene response in *Arabidopsis thaliana*

E.V. Zemlyanskaya[1, 2]*, D.Y. Oshchepkov[1], V.V. Mironova[1, 2], V.G. Levitsky[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: ezemlyanskaya@bionet.nsc.ru*

Key words: transcription factor, epigenetic regulation, ETHYLENE-INSENSITIVE3, EIN3 binding site (EBS), ChIP-seq, RNA-seq

Motivation and Aim: The plant hormone ethylene regulates numerous developmental processes and stress responses [1]. Ethylene signaling proceeds via a linear pathway, which activates EIN3 transcription factor [2]. EIN3 influences gene expression upon binding to a specific sequence in gene promoters. However, in many cases EIN3 binding to gene promoter is not enough to trigger transcriptional response [3]. Here we perform whole genome bioinformatics study to dissect the factors essential for EIN3 functioning.
Methods and Algorithms: We extract EIN3 binding regions (EBRs) from publicly available ChIP-Seq data on EIN3 binding in *Arabidopsis thaliana* [3] and use RNA-seq data on ethylene-induced transcriptomes [3] to determine ethylene responsiveness of corresponding genes. We further investigate the impact of DNA-binding context, its position relative TSS and epigenetic status on the ethylene sensitivity of genes bound by EIN3. We use previously published genome-wide map of nine chromatin states in *A. thaliana* [4] to characterize EIN3 binding regions with respect to the epigenetic status.
Results: The analysis of ChIP-seq data showed bimodality of distribution of EIN3 binding regions in gene promoters. We found that the implicit distal peak was associated with a specific chromatin state (referred to as chromatin state 4 in the primary source), which was just poorly represented in the pronounced proximal peak. EIN3 binding regions corresponding to the chromatin state 4 were significantly associated with ethylene response, unlike the others representing the overwhelming majority of EBRs related to the explicit proximal peak. Moreover, we found that specific EIN3 binding sequences predicted with previously described model were enriched specifically in the EBRs mapped to the chromatin state 4, but not to the rest ones [5].
Conclusion: These results allow us to conclude that the interplay of genetic and epigenetic factors might cause the distinct modes of EIN3 regulation.
Acknowledgements: The work was supported by Russian Foundation for Basic Research and the government of Novosibirsk region, grant No. 18-44-540039.

References

1. Abeles F.B. et al. (1992) Ethylene in plant biology. San Diego: Acad. Press.
2. Merchante C. et al. (2013) Ethylene signaling: simple ligand, complex regulation. Curr. Opin. Plant Biol. 16:554-560.
3. Chang K.N. et al. (2013) Temporal transcriptional response to ethylene gas drives growth hormone cross-regulation in Arabidopsis. Elife. 2:e00675.
4. Sequeira-Mendes J. et al. (2014) The functional topography of the Arabidopsis genome is organized in a reduced number of linear motifs of chromatin states. Plant Cell. 26:2351-2366.
5. Zemlyanskaya E.V. et al. (2017) The interplay of chromatin landscape and DNA-binding context suggests distinct modes of EIN3 regulation in Arabidopsis thaliana. Front. Plant. Sci. 7:2044.

DOI 10.18699/BGRSSB-2018-165

Computational modelling of wheat leaf growth and morphogenesis based on data from 3D LSM-images

U. Zubairova[1]*, A. Doroshkov[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
* *e-mail: ulyanochka@bionet.nsc.ru*

Key words: wheat leaf, epidermal patterning, computational modelling, confocal laser scanning microscopy, morphogenesis

Motivation and Aim: Leaf epidermis of a monocotyledonous plant is a widely used model system for studying morphogenesis. The epidermis of cereals (wheat, barley, rice) leaf is a complex tissue consisting of different cell types organizing in parallel cell rows. For such leaves, a unidirectional growth occurring for a long time enables to observe a series of successive morphogenetic stages at one snapshot. In this work, we propose the concept for using a growing wheat leaf to study dynamical changes in morphogenesis, including stress-induced changes. Linear leaf of a wheat, during its formation for a long time, maintains a phase of steady growth. Therefore, it is possible to observe a series of successive events of morphogenesis fixed in the cellular architecture of a mature leaf.
Methods and Algorithms: High-resolution 3D LSM-images allow extracting quantitative characteristics describing the cellular structure of leaf epidermis. However, to obtain a large amount of statistical data methods of high throughput computer based image segmentation should be used. We developed a workflow for detection of structural properties of leaf epidermis from 3D images obtained from confocal laser scanning microscopy. The workflow includes the protocol of sample preparation, image processing ImageJ-plugin and data extraction algorithms. The data on the cellular architecture further will act as a basis for the elaboration and verification of spatial models accounting for structural features of leaves. For the leaf epidermis of cereals, a brickwork-like pattern combined with unidirectional growth allows to reduce the dimension and use a quasi-one-dimensional representation of the cellular ensemble in the model. This idea was realized in the model [1] growth of a linear leaf blade. The model allows for fitting of the visible cell length using the experimental cell length distribution along the longitudinal axis of a leaf epidermis.
Results: In this work, we assume a unidirectional growing cell ensemble starting from a meristem-like layer of generative cells and then generating parallel cell rows from every cell of the initial layer. We considered the growth zone of the leaf includes division and elongation zones; in addition, the division zone includes a zone of asymmetric divisions forming specialized cells (trichomes and stomata). The model was verified on qualitative and quantitative data on cold stress induced disturbances of morphogenesis in the epidermis of wheat leaf.
Acknowledgements: The work was supported by the project 17-44-543384 from Russian Foundation for Basic Research.

References

1. Zubairova U., Nikolaev S., Penenko A., Podkolodnyy N., Golushko S., Afonnikov D., Kolchanov N. (2016). Mechanical behavior of cells within a cell-based model of wheat leaf growth. Frontiers in plant science. 7:1878.

ANIMAL GENETICS

DOI 10.18699/BGRSSB-2018-166

FTO haplotyping underlines high obesity risk for European populations

R.O. Babenko, V.N. Babenko*, L.N. Trut, A.L. Markel
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: bob@bionet.nsc.ru*

Key words: population specific obesity risk, metabolic syndrome, 1000G populations, FTO gene, haplotype profiling, ancestor allele

We analyzed the population specific haplotype profiles of the FTO gene genomic locus hotspot identified by Genome Wide Association Studies (GWAS) for the high obesity risk by scrutinizing eighteen 1000G populations from 4 continental groups. The hotspot is located in FTO gene intron 1 spanning around 40kb. We reconstructed the ancestral state of the locus, which comprised 'healthy' major allele found in all populations, and two minor 'risky' alleles, each one specific for African and European populations, correspondingly. The allele locus structure and frequency distribution underscores the high risk allele specifically for European population. South Asian populations take the second place on the 'risky' allele frequency, while East Asian populations have the minimal ratio of risky allele. African populations specific allele was only 'partially' risky, while the majority of GWAS SNPs were manifested by healthy alleles'. These observations corroborate the previous reports on the FTO locus implication in population specific manner as well as WHO BMI index population distribution. Thus, the conclusions presented imply FTO locus analyzed is rather a major genetic determinant of the genetic obesity risk from the GWAS SNPs set.

DOI 10.18699/BGRSSB-2018-167

Studying the impact of sex on the molecular mechanisms of liver adaptation to fasting in mice

N. Bazhan[1, 2]*, N. Sitnikova[2], T. Iakovleva[1], A. Dubinina[1], E. Makarova[1]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: bazhan-nm@yandex.ru*

Key words: carbohydrate-lipid metabolism, fasting-refeeding, liver, gene expressions

Motivation and Aim: Fundamental aspects of metabolic homeostasis are known to be regulated differently in males and females. A key adaptive response to fasting is the shift from glucose to lipid oxidation. The liver plays a crucial role in mobilizing energy during nutritional deprivation [1]. It remains unclear if there are sex-dependent peculiarities in molecular mechanisms of liver adjustment to fasting. The aim of this study was to evaluate the expression patterns of key genes involved in lipid and glucose metabolism in liver and their relationship with blood parameters in response to food restriction.
Methods and Algorithms: Male and female C57Bl mice were studied under different feeding conditions: feeding state and 24 h fasting. Blood parameters and the expression of genes involved in glucose metabolism (gluconeogenesis, glycolysis) and lipid metabolism (lipogenesis, fatty acid oxidation) in liver were analyzed. mRNA levels were measured by RT-PCR.
Results: As an effect of fasting, there was decrease in circulating glucose and leptin levels, increase in FGF21 and adiponectin levels both in male and female mice. There were no differences between male and female mice in fasting induced changes in expression of genes involved in glucose metabolism: mRNA levels of genes controlling gluconeogenesis (*Pck1*, *G6P*, *Pgc1*) were increased, and glycolysis (*Gk*) was decreased compared to those in fed animals. Transcriptional regulation of hepatic lipid metabolism was sex-dependent. Expression of *Ppar*-α gene, which is involved in fatty acid oxidation and concomitant production of ketone bodies, was higher in females compare to males both in fed and fasting states. Fasting decreased the expression of lipogenic gene (Fas), and increased the expression of genes related with fatty acid oxidation (*Ppar*-α, *Fgf21 Cpt1*) in mice of both sexes, but fasting induced increase in the *Fgf21*and *Cpt1* gene expressions were more pronounced in female than in male mice.
Conclusion: Thus, in mouse liver, the effect of fasting on the expression of genes involved in glucose metabolism was independent on sex, and on genes involved in fat metabolism was dependent on sex. Increased fatty acid oxidation favors the production of ketone bodies – the main energy source under fasting conditions. Increased expression of hepatic *Fgf21*and *Cpt1* genes in fasted females possibly reflects a higher ability for the female mice to respond to increased energy demands.
Acknowledgements: Supported by the Russian Science Foundation, grant No. 17-15-01036.

References

1. Potthoff M.J. et al. (2009) FGF21 induces PGC-1alpha and regulates carbohydrate and fatty acid metabolism during the adaptive starvation response. Proc Natl Acad Sci USA. 106(26):10853-10858. DOI 10.1073/pnas.0904187106.

DOI 10.18699/BGRSSB-2018-168

Sequencing of reindeer (*Rangifer tarandus*) genomes: insights into evolution, domestication and adaptation

M.W. Betsha[1, 2], K. Pokharel[1], M. Yao[3], M. Honkatukia[1], N. Mazzullo[4], J. Peippo[1], T. Reilas[1], P. Soppela[4], A. Stammler-Grossman[4], F. Stammler[4], K.H. Røed[5], J. Kantanen[1]*

[1] *Natural Resources Institute Finland, Jokioinen, Finland*
[2] *Department of Environmental and Biological Sciences, University of Eastern Finland, Kuopio, Finland*
[3] *BGI-Tech, Guangdong, China*
[4] *Arctic Center, University of Lapland, Rovaniemi, Finland*
[5] *Norwegian University of Life Sciences, Oslo, Norway*
** e-mail: juha.kantanen@luke.fi*

Key words: *de novo* sequencing, mitochondrial DNA, resequencing, subspecies, taxonomy

Motivation and Aim: Semi-domesticated reindeer (*Rangifer tarandus*) have pivotal economic, societal and cultural value for indigenous people and pastoralists in northern and subarctic regions in Eurasian. Currently, there exist several semi-domesticated and wild *Rangifer*-populations, the taxonomic status of which has been actively debated. To examine genetic diversity, domestication history, taxonomy and adaptation, we deep-sequenced and *de novo* assembled one reindeer genome, resequenced 23 other *Rangifer* sp. samples and analysed whole mitochondrial DNA of these individuals.
Materials and Methods: Genomic DNA of a 1-year-old Finnish male reindeer was sequenced at high coverage (100X) on the Illumina HiSeq2500 and 4000 platforms. Seven paired-end DNA libraries with insert sizes ranging from 170bp to 20kb were constructed. The genome assembly, annotation, and orthology analysis were conducted using a robust bioinformatics pipeline. In the resequencing approach, 23 semi-domesticated and wild reindeer and caribou were sequenced (10X) using Illumina HiSeq2500 platform. The data were subjected e. g. for the principal component analysis. The complete mitochondrial genome for reindeer was assembled using the MITObim v.1.9 software.
Results: A total of 300.5 Gb of clean data was assembled using SOAP denovo resulting into 256,454 scaffolds (N50 = 502 Kb) with cumulative scaffold length of 2.66 Gb and spanning 90 % of the estimated (2.9 Gb) genome size of reindeer. Using a homology based approach, the reindeer genome was predicted to harbour 27,332 protein coding genes, 98 % of which were functionally annotated. Reindeer displayed a heterozygosity level, which is 2.3 and 1.5 times higher than that found in taurine cattle or yak. The resequenced animals grouped into two main clusters: northern European and northern Russian/northern American. Our study provides new information on the architecture of mitochondrial genome of reindeer (16,451 bp).
Conclusion: The draft quality of the reference genome along with the annotations will provide important insights into the evolution and demographic history of the reindeer and taxonomy of *Rangifer* sp. Our findings suggest that there have been at least two domestication events in the history of reindeer.
Acknowledgements: The present study is supported by the Academy of Finland (Grant No. 286040).

DOI 10.18699/BGRSSB-2018-169

Synapsis and recombination in intra- and interspecies hybrids between two voles species *Microtus* (*Alexandromys*) *evoronensis* and *M. maximowizcii*

T. Bikchurina[1, 2]*, T. Vasil'eva[3], M. Pavlenko[3], I. Sheremet'eva[3], I. Kartavtseva[3]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Federal Scientific Center of the East Asia Terrestrial Biodiversity FEB RAS, Vladivostok, Russia*
** e-mail: bikchurina@bionet.nsc.ru*

Key words: chromosomal rearrangements, pachytene, recombination, voles

Motivation and Aim: In the early stages of speciation, reproductive isolation can occur due to genetic incompatibility and/or gradual accumulation of different chromosomal rearrangements. Heterozygotes for the chromosomal rearrangements form complex multivalents in meiotic prophase I, which result in high occurrence of non-homologous synapsis and asynapsis of homologous chromosomes. These aberrations result in apoptosis of the hybrid gametocytes. However, cytological basis of these processes remains poorly understood. Intra- and interspecies hybrids between parental forms differed by chromosomal rearrangements provide a good model to analyze different degree of meiotic abnormalities at the early stages of speciation.
Methods and Algorithms: Using immunolocalization of SYCP3 (the main protein of lateral element of synaptonemal complex), MLH1 (miss-match protein marking late recombination nodules) and centromere proteins at pachytene we analyzed synapsis and recombination in female and male hybrids between different chromosomal races of *M. evoronensis* and between *M. evoronensis* and closely related *M. maximowizcii*.
Results: We found a significant difference in the degree of meiotic abnormalities between female and male interspecies hybrids. In females, we observed simple multivalents with recombination at the homologously paired arms. In males, most cells contained complex multivalents with completely suppressed recombination. Intraspecies hybrids showed different level of multivalent complexity. The number of recombination event per cell was slightly reduced compared to the parental species.
Conclusion: The severity of meiotic abnormalities in analyzed hybrid voles increased with increasing degree of chromosomal divergency. In the interspecies hybrids, the complexity of multivalents was higher and the recombination was affected stronger compared to intraspecies hybrids. In accordance to the Haldane rule, the meiotic disturbances were more pronounced in the heterogametic sex.
Acknowledgements: Supported by the by Russian Foundation for Basic Research (Grants No. 17-29-08019 and No. 15-04-03871) and the Federal Agency for Scientific Organizations via the Institute of Cytology and Genetics (Grant No. 0324-2018-0019).

DOI 10.18699/BGRSSB-2018-170

Do rodent species adopt to underground lifestyle by different ways?

O. Bondareva[1]*, A. Kasianov[2], N. Abramson[1]
[1] *Zoological Institute RAS, Saint Petersburg, Russia*
[2] *Vavilov Institute of General Genetics, Moscow, Russia*
** e-mail: olga.v.bondareva@gmail.com*

Key words: underground rodents, adaptations, gene ontology, underground lifestyle

Motivation and Aim: Up to now the majority of studies on the analysis of genetic diversity within species and populations was carried out using a small number of molecular markers. However, this approach does not account for the molecular basis of adaptive variation, it remains unclear how many genes involved in particular adaptation, what is the origin of genetic diversity responsible for formation of adaptation. Comparison of genomes of phylogenetically close taxa, but the contrast in adaptations and phylogenetically distant but with similar adaptive traits will recover convergence and parallelisms at the molecular level. This allows testing the hypothesis on the origin of mutations that lead to similar phenotype effects; reveal the velocity at which mutation in the DNA may cause a phenotypic effect.
Methods and Algorithms: We used annotated genomes of terrestical and underground rodents from the Ensembl genome browser (ensembl.org). Seven species with well assembled genomes were used: *Cavia porcellus*, *Chinchilla lanigera*, *Rattus norvegicus*, *Mus musculus*, *Fukomys damarensis*, *Nannospalax galii*, *Heterocephalus glaber*. Among listed organisms three are underground rodents. Ortholog genes were identified with *protheinortho* program and dN/dS values were determined in PAML codeml program for each orthogroup. GO enrichment analysis was performed with Webgestalt software (http://www.webgestalt.org/option.php).
Results: First of all we tried to find out ortholog genes, for which directions of selection within analyzed groups is the same but differs between underground and terrestical species. Unfortunately, we could not identify such examples. Next we excluded from the analisis genomes of terrestical species and identified ortholog genes that have the same direction of selection in all underground species. These genes are enriched with several GO terms, connected with many essential processes: “RNA binding”, “metabolic processes”, “regulation” and “transcription activity” etc. The special interest represents groups of terms linked to mitochondria or oxidation processes, because they may be involved in the adaptation to hypoxia. This fact is good agreement that mitochondrial genes are under strong positive selection.
Conclusion: Obtained results demonstrate that adaptations to the underground lifestyle may be linked to changes in different essential processes.
Acknowledgements: Supported by the RFBR No. 18-04-00730.

DOI 10.18699/BGRSSB-2018-171

High *Wolbachia* infection rate in four-eyed fir bark beetle (*Polygraphus proximus*) populations of Tomsk province

R. Bykov[1]*, I. Kerchev[2], M. Yudina[1, 3], G. Yurlova[1], A. Tikhomirova[3], Yu. Ilinsky[1, 3]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Institute of Systematics and Ecology of Animals SB RAS, Novosibirsk, Russia*
[3] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: bykovra@bionet.nsc.ru*

Key words: *Polygraphus proximus*, *Wolbachia*, pest, Siberia

Motivation and Aim: The four-eyed fir bark beetle (*Polygraphus proximus*) is invasive pest of the Siberian fir [1]. Studying of its biology, in particular associations with any bacterial symbionts, may be important for biological control of this pest. Bacteria of the genus *Wolbachia* are common symbionts of many insects [2]. In some cases, these bacteria play significant role in host biology. This is the first study of *Wolbachia* infection in *P. proximus* populations.
Methods and Algorithms: The collection includes 152 samples of *P. proximus* from five regions of Tomsk province. Total DNA was individually extracted from whole beetles. *Wolbachia* detection was performed by PCR with primers specific to two housekeeping loci of *Wolbachia*. Infection rates were estimated for each studied locality. The multilocus sequence typing approach was performed to characterize *Wolbachia* isolates.
Results: *Wolbachia* symbiont was found in all studied localities of Tomsk province, in particular in Chainsky (44 %), Molchanovsky (42 %), Bakcharsky (23 %), Krivosheinsky (77 %) and Tomsky (51 %) regions. The average infection rate was 47 % (95 % confidence interval, 39–56 %).
Conclusion: Here we firstly report on *Wolbachia* infection in *P. proximus*. Our data suggests of high *Wolbachia* infection rate in the populations of four-eyed fir bark beetle of Tomsk province. Further analysis of *Wolbachia – P. proximus* interactions may be used to develop new approach to control four-eyed fir bark beetle populations.
Acknowledgements: Supported by the RFBR (18-316-00099) and by the RSF (15-14-10014).

References

1. Kerchev I.A. (2014) Ecology of four-eyed fir bark beetle *Polygraphus proximus* Blandford (Coleoptera; Curculionidae, Scolytinae) in the West Siberian region of invasion. Russian Journal of Biological Invasions. 5(3):176-185.
2. Zug R., Hammerstein P. (2012) Still a host of hosts for Wolbachia: analysis of recent data suggests that 40% of terrestrial arthropod species are infected. PloS One. 7(6):e38544.

DOI 10.18699/BGRSSB-2018-172

Wolbachia and mtDNA diversity and distribution in palearctic *Drosophila melanogaster* populations

R. Bykov[1]*, I. Mazunin[2], M. Yudina[1, 3], Yu. Ilinsky[1, 2, 3]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Immanuel Kant Baltic Federal University, Kaliningrad, Russia*
[3] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: bykovra@bonet.nsc.ru*

Key words: *Wolbachia*, *Drosophila melanogaster*, mtDNA

Motivation and Aim: *Wolbachia* symbionts are found in *Drosophila melanogaster* populations all over the world [1, 2]. Genetic diversity of the symbiont in *D. melanogaster* is subdivided into several clades that have strong association with certain host mtDNA clades [3, 4]. Here we represent results of comprehensive surveys on cytoplasmic inheritance (*Wolbachia* and mtDNA) of fruit fly populations in a vast Palearctic territory. We address to the symbiont prevalence and symbiont genetic pattern as well as mtDNA pattern.
Methods and Algorithms: In total 1550 *D. melanogaster* samples were collected from different Palearctic localities. Samples were screened by PCR for *Wolbachia* infection and mitochondrial haplotypes.
Results: *Wolbachia* infection was found in every studied *D. melanogaster* populations from Western Europe to Far East. The average rate of infection was 0.56 (95 % confidence interval, 0.54–0.59). Infection rates are not dependent on longitude or latitude. Five mtDNA clades were found, where two of them (III and V clades) were predominant, that confirmed with previous data: the III clade is widely distributed in the world while the V clade is found only in Palearctic region.
Conclusion: We demonstrate wide *Wolbachia* distribution in Palearctic *D. melanogaster* populations. According to our data on symbiont diversity and distribution the fly populations of many regions in temperate zone renew after cold season. High frequency of the V clade of mtDNA in Palearctic populations may be explained by it having originated in this region.
Acknowledgements: Supported by the RFBR (16-04-00980).

References

1. Solignac M. et al. (1994) Widespread occurence of the proteobacteria *Wolbachia* and partial cytoplasmic incompatibility in *Drosophila melanogaster*. Comptes rendus de l'Académie des sciences. Série 3. Sciences Vie. 317(5):461-470.
2. Riegler M. et al. (2005) Evidence for a global *Wolbachia* replacement in *Drosophila melanogaster*. Current Biology. 15(15):1428-1433.
3. Richardson M.F. et al. (2012) Population genomics of the *Wolbachia* endosymbiont in *Drosophila melanogaster*. PLoS Genetics. 8(12):e1003129.
4. Ilinsky Yu. (2013) Coevolution of *Drosophila melanogaster* mtDNA and *Wolbachia* genotypes. PLoS One. 8(1):e54373.

DOI 10.18699/BGRSSB-2018-173

Gene expression related to aggressive behavior on rat model

I.V. Chadaeva[1, 2]*, A.O. Bragin[1], R.V. Kozhemyakina[1], N.V. Klimova[1], Yu.L. Orlov[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: ichadaeva@bionet.nsc.ru*

Key words: laboratory rats, aggressive behaviour, dopaminergic transmission, gene expression, RNA-seq

Motivation and Aim: Aim of this study was to study the mechanisms of hereditary-mediated aggressive behavior on laboratory animal models based on transcriptome profiling. To establish mechanisms of aggression at the molecular level, we used unique experimental model of grey rats (Rattus norvegicus) developed at the ICG SB RAS for more than 80 generations.
Methods and Algorithms: Rats have been subjected to selection in two directions – tolerant behavior towards human and aggressive behavior. We estimated the gene expression in rat brain areas based on RNA-seq data [1] and verification it by RT-PCR.
Results: We focused on genes presumably associated with the manifestation of aggressive behavior: Gad2, Drd2, Cacna1b, Egr1, Gbrd, Pomc, Gria2, Mapk1, Syn1, Cacna2d3, Nos1, Oxt. RNA-profiling experiments revealed the lists of differentially expressed genes in the brain samples.
Conclusion: A set of synapse associated genes have statistically significant deviation in splicing depending on brain regions and behavioral models (tolerant/aggressive) of rats. The genetic factors exert a strong influence to the phenotypic variation of aggressive behavior in populations.
Acknowledgements: The research has been supported by RFBR (grant 18-34-00496).

References

1. Babenko V.N. et al. (2016) Analysis of differential gene expression by RNA-seq data in brain areas of laboratory animals. Journal of Integrative bioinformatics. 13(4):292.

DOI 10.18699/BGRSSB-2018-174

Sex-specific effect of leptin on gene expression in placentas and fetal tissues in mice

E. Denisova*, E. Makarova
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: elena_nsib@list.ru*

Key words: leptin, placenta, fetal brain, pregnancy, gene expression, mice

Aim: According to the DOHAD hypothesis, the health of individuals depends on the conditions of prenatal development [1]. The hormone leptin is elevated in obese animals. Hyperleptinemia during pregnancy was shown to affect the metabolic phenotype of the offspring, and this effect may depend on the offspring sex [2]. Leptin receptors were found in the placenta and fetal brain and liver. Programming leptin effects may be mediated via leptin influence on the placental functions and development of fetuses. The goal of the work was to examine how leptin administration to the mice at the end of pregnancy affects expression of genes encoding signal and transport proteins in placentas and genes involved in energy homeostasis regulation in the brain of male and female fetuses.
Methods: Leptin or saline (control) were administered to C57Bl mice at the day 17 of pregnancy. Weight of fetuses and placentas, relative expression of genes encoding IGF1, IGF2 and IGF2R, glucose transporters GLUT1 and GLUT3, amino acid transporters SNAT1, SNAT2 and SNAT4 in placentas, IGF1, IGF2 and IGF2R in liver, and MCR4, AgRP, NPY and POMC in brain were measured in male and female fetuses within 3 and 7 hours after injection. Leptin plasma concentrations were measured in fetuses and pregnant females within 1 h after injection.
Results: Through 1 h after leptin administration, leptin concentrations in fetal and mother blood were significantly higher than in saline treated mice. Through 3 h after leptin administration, gene expression was changed only in female foetuses: the expression of IGF2 was decreased in female placentas and expression of MCR4 was increased in fetal female brain. Through 7 h after leptin administration, IGF2R gene expression was increased in liver of female fetuses, AgRP gene expression was decreased in brain of male fetuses, MCR4 gene expression in brain was increased and fetal weight was decreased in fetuses of both sexes.
Conclusion: Leptin administration retards fetal growth rate via the sex-specific molecular mechanisms. In female fetuses, these mechanisms include inhibition of IGF2 expression in placentas and activation of IGF2R expression in liver. In male and female fetuses, leptin administration differently affects the expression of genes that regulate energy homeostasis at maturity. Gender-specific programming effect of maternal leptin may be associated with different gene expression response to leptin in placentas, brain and liver of male and female fetuses.
Acknowledgements: Supported by the RFBR (17-04-01357A and 14-04-00694A).

References

1. Wadhwa P.D. et al. (2009) Developmental origins of health and disease: brief history of the approach and current focus on epigenetic mechanisms. Semin Reprod Med. 27(5):358-368.
2. Makarova E.N. et al. (2013) Influence of abnormally high leptin levels during pregnancy on metabolic phenotypes in progeny mice. Am J Physiol Regul Integr Comp Physiol. 305:R1268-R1280.

DOI 10.18699/BGRSSB-2018-175

Selected rat strains HT, LT as a model for the study of dysadaptation states dependent on the level of excitability of the nervous system

N.A. Dyuzhikova*, A.I. Vaido
Pavlov Institute of Physiology of the RAS, St. Petersburg, Russia
** e-mail: dyuzhikova@mail.ru*

Key words: selection, excitability, rat strains, stress, anxiety disorders

To elucidate the interconnections between the functional state of the nervous system (excitability), the functioning of the brain and a wide range of behavioral characteristics, a selection program with a primary goal of generating rat strains differing in thresholds of excitability of the nervous system has been launched in the 70s of the last century [1]. Wistar rats (Stolbovaya) were used as a starting material. The selection has been carried on the magnitude of the threshold of electric excitability (rectangular electric impulses, 2 ms) of the tibial nerve (n. tibialis). Four rat strains with different gradation of excitability thresholds have been breeded: HT1, LT1, HT2, LT2 (high and low thresholds, 1, 2 – numbers of breeding programs) [1]. Currently, two strains that have passed more than 70 generations of breeding are maintained – HT1 (HT) and LT2 (LT) having mostly contrasting thresholds of excitability (3.5 ± 0.3 and 0.70 ± 0.04 V, respectively). Over this long research period, the rats of these strains have been manifesting differences not only in tibial nerve excitability, but also in the thresholds of excitability in other parts of the nervous system, both the peripheral and CNS (mainly subcortical structures). The impact of the nervous system excitability on a wide spectrum of conditioned and unconditioned behavioral characteristics has been revealed [2]. In these rat strains, alterations have been detected in various parts of hormonal regulation systems, neurotransmission, the ion channels functioning, structural and functional properties of the nerve cell membranes. Strains have manifested different stress-reactivity in tests of sleep deprivation, immobilization, response to short and long-term emotional-painful stress. Prolonged post-stress behavioral manifestations persist for 6 months in the both HT and LT rats. Disorders of higher nervous activity have strain-specific manifestations: formation of a depressive-like behavior in the low-excitable HT strain, an increase in excitability, aggression, the disturbance of plastic processes, and an increase in stereotypic behavior (*jactatio capitis*) in the highly excitable LT strain. These characteristics allow the usage of these strains as model objects for studies on the post-stress anxiety disorders, in particular, post-traumatic stress disorder (PTSD) and compulsive disorder (CD). The long-term effects of stress are based on the morphological changes of neurons in various brain structures, differential chromatin modifications in neurons and other somatic cells associated with epigenetic DNA and histone modifications [3, 4]. The horizons of using the HT and LT strains for the elucidating genetic and epigenetic mechanisms of dysadaptation in response to environmental factors in terms of a personified medicine with a glance on the nervous system excitability are discussed.
Acknowledgements: Supported by Presidium of UB RAS (0134-2018-0003).

References

1. Vaido A.I., Sitdikov M.Kh. (1979) Selection of rat strains by long-term threshold of excitability of neuromuscular apparatus. Genetica. XV(1):144-148.
2. Levina A.S. et al. (2018) Comparative behavioral characteristics of two rat strains differing in the nervous system excitability threshold in a spatial learning task in the Morris water maze. Zh. Vyssh. Nerv. Deiat im I.P. Pavlova. 68(2):1-12.
3. Vaido A.I. et al. (2009) Systemic control of the molecular, cell and epigenetic mechanisms of long-lasting consequences of stress. Genetica. 45(3):342-348.
4. Dyuzhikova N.A. et al. (2015) Epigenetic mechanisms in post-stress states. Usp Fiziol Nauk. 46(1):47-75.

DOI 10.18699/BGRSSB-2018-176

The hormonal mechanism of heat stress effect on the carbohydrate metabolism in *Drosophila melanogaster* females

M.A. Eremina*, E.K. Karpova, N.E. Gruntenko
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
* *e-mail: eremina@bionet.nsc.ru*

Key words: heat stress, dopamine, octopamine, juvenile hormone, 20-hydroxyecdysone, *Drosophila*

The *Drosophila* studies have revealed a strong evolutionary conservatism of the insulin/insulin-like growth factor-like signaling (IIS) and its involvement in the regulation of metabolic homeostasis and resistance to various types of stressors. The two main forms of circulating carbohydrates in *Drosophila* are glucose and trehalose (glucose disaccharide). The use of the evolutionary conservatism of I/IGF makes it possible to analyze the mechanisms underlying the development of diabetes mellitus and allows the experimental study of the influence of factors that can have a provoking effect on it, in cases where human research is impossible. One of these factors is stress – the universal response of living organisms to any adverse influences. Various hormones have been implicated in the stress response of the *Drosophila* adult, in particular biogenic amines – dopamine (DA) and octopamine (OA), which perform neurotransmitter, neuromodulatory and neurohormonal functions, juvenile hormone (JH) and 20-hydroxyecdysone (20E), playing a fundamental role in the control of reproductive function in adult insects. The purpose of this work is to study the effect of heat stress on the carbohydrate metabolism of *D. melanogaster* females of the wild type Canton S (CS) strain combinated with changes in the level of stress-related hormones (DA, OA, JH, 20E). We show that:
1) DA and JH have an inhibitory effect on carbohydrate metabolism under normal conditions;
2) OA and 20E exert a stimulating effect on the level of the main carbohydrate of insects, trehalose, but an inhibitory effect on the level of glucose under normal conditions;
3) DA, JH, 20E have an inhibitory effect on the level of both carbohydrates, OA has an inhibitory effect on the level of glucose and a stimulating effect on the level of trehalose in case of heat stress.
Our data suggest that the hormones involved in the neuroendocrine stress response of *D. melanogaster* are involved in the regulation of carbohydrate metabolism. Further work will focus on correcting carbohydrate metabolism in *D. melanogaster* with strains disrupted IIS disorders.

DOI 10.18699/BGRSSB-2018-177

IGNG1-IGNG3 locus and its possible role in the multiple sclerosis

M. Fridman[1]*, P. Bycadorov[1], N. Oparina[2], V. Makeev[1]
[1] *Vavilov Institute of General Genetics, RAS, Moscow, Russia*
[2] *Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden*
** e-mail: marina-free@mail.ru*

Key words: multiple sclerosis, SNP, population rates

Motivation and Aim: A large amount of GWAS data on multiple sclerosis (MS) has been obtained recently almost exclusively for populations dominated by Caucasians. Unfortunately, SNPs that display significant association with MS may only be linked to those that are mechanistically related to the disease development. In addition, GWAS does not provide any idea on the mechanisms of the SNP influence. Such information can obtained from eQTLs, but an eQTL also usually marks only a fairly large locus. Therefore, to study genome segments functionally associated with disease developement it is relevant to identify a rather long genome segments containing several SNPs and perform enrichment analysis for the features associated with the segment.
Methods and Algorithms: The idea of our approach is to take into account SNPs displaying low association with the target feature, and discard the corresponding genome segments, thus reducing the target regions. We have generated a set of loci enriched with SNPs associated with the MS development. Then, we created two SNP lists, statistically associated and non-associated with the development of MS according to all GWAS data. SNPs strongly linked with SNPs associated with MS were added to the target list, whereas SNPs linked with SNPs displaying low MS association were discarded. The second list contained SNPs simultaneously linked with two SNPs, statistically associated with MS according to all GWAS data set, the third list contained SNPs linked with three MS associated SNPs. Surprisingly, all SNPs in the final list were found in one locuse in chromosome 14, containing IGHG1 and IGHG3 genes and in several loci in chromosome 6 (containing HLA genes).
Results: Out of 18302 SNPs not associated with MS or linked with non-associated SNPs only 13 are very frequent in Europeans (> 0.3) and very rare in Africans and Asians (< 0.03) in population frequencies. Conversly, out of 7524 SNPs associated with MS or linked with associated SNPs 14 displayed such population frequency bias, and of these all but one were found in the locus we identified in the chr 14 or in its immediate vicinity. More to the point, the given locus is associated with IgG index (the ratio of concentrations of IgG in the cerebrospinal fluid and serum as compared with the same ratio for albumin). According to the data of (1) 5 out of 6 SNPs associated with the IgG index, are on this site
Conclusion: As for the mechanism, we propose that IgG increases antigen presentation by interacting with FcRγ-receptors (2), and stimulates B-cell secondary immune responce to IgG synthesis by activating T-helper cells of antigenic presenting complex. As it is known, the MS frequency is about ten-fold higher in countries with a predominance of the Caucasian population. In addition, the disease development sometimes is different for non-Europeans. For instance, the abnormal intrathecal synthesis of IgG, reflected by cerebrospinal fluid oligoclonal IgG bands and increased IgG index, is much less frequent in Japanese (3). We assume that the given locus is responsible for the corresponding differences. We suppose that at this locus there was a positive selection during the resettlement to the high latitudes of Europe.

References

1. Buck D., Albrecht E., Aslam M. et al. (2013) Genetic variants in the immunoglobulin heavy chain locus are associated with the IgG index in multiple sclerosis. Ann Neurol. 73(1):86-94.
2. Getahun A., Heyman B. (2004) IgG- and IgE-mediated antigen presentation on MHC class II. Immunol. Lett. 92:33-38.
3. Yoshimura S., Isobe N., Matsushita T. et al. (2014) Genetic and infectious profiles influence cerebrospinal fluid IgG abnormality in Japanese multiple sclerosis patients. PLoS One. 9:e95367.

DOI 10.18699/BGRSSB-2018-178

Genotype of recipient mother modulate body composition and immunocompetence of transferred progenies

L. Gerlinskaya[1]*, E. Litvinova[1], G. Kontsevaya[1], N. Feofanova[1, 2], K. Achasova[1], M. Anisimova[1], S. Maslennikova[1], M. Zolotykh[1], M. Moshkin[1]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Institute of Immunology*
** e-mail: lgerlinskaya@gmail.com*

Key words: embryos, mothers, immunogenetic dialogue, modulations, offspring properties

Motivation and Aim: Among the causes of persistent phenotypic modifications, the genotype of the nurturing mother, which as a rule differs from the genotype of the transplanted embryo. During fetal development between the mother and the embryo there is a bilateral exchange of immunologically significant information the nature of which is determined by the degree of antigenic differences between mother and embryo, in particularly between genes of the main histocompatibility complex (MHC) [1]. In experiments with C57BL mice ($H2^b$ haplotype) and BALB / c ($H2^d$ haplotype) were previously showed, that offspring obtained by interlinear embryo transfers are different from those obtained by intrastrain embryo transplantation by a less adrenocortical response to various variants of social stress [2]. However, this study does not cover the whole variety of immunogenic interactions between the fetus and the mother and the whole variety of potential modulations of the offspring properties caused by the conditions of intrauterine development that can substantially vary depending on the genotype of the nurturing female.
Methods and Algorithms: In the presented study, transplantation of two-cell inbred strain embryos coinciding or different in the MHC (H2) haplotype from the recipient female was used to model three variants of the immunogenic dialogue in the mother-fetus system. 1. A bidirectional immunogenic dialogue – embryos C57BL/6J ($H2^b$ haplotype) and nurturing BALB/c mothers ($H2^d$ haplotype). 2. One-way dialogue – embryos C57BL/6J and not capable to effectively respond to foreign antigens and immunodeficient NOD.SCID mothers ($H2^{g7}$ haplotype). 3. Absence of immunogenetic dialogue – embryos and mothers of the C57BL/6J mice of the same $H2^b$ haplotype. In the offspring, at the time of weaning from mothers and at reaching the sexually mature age, the body weight, its composition, as well as the parameters of nonspecific and specific immune responses to the action of *Anthrax* and *Helicobacter hepaticus* were measured
Results: Adult descendants born to mothers of the same genetic strain were characterized by a high fat content compared to that of offspring that were born to females of a different genotype. C57BL/6J progenies nurtured by immunodeficient NOD SCID mothers had a more pronounced humoral immune response to oral infection with *H. hepaticus* and less *in vitro* macrophage cytokine reaction on the anthrax exotoxin in comparison with the offspring nurtured by the C57BL and BALB/c mothers.
Conclusion: These results let us to conclude that the influence of the genotype dependent mother-fetus dialogue on the fat metabolism and the immune system in adult offspring.
Acknowledgements: Supported by the FPI (6/099/2017ав). The studies are implemented using the equipment of the Center for Genetic Resources of Laboratory Animals at ICG SB RAS, supported by the Ministry of Education and Science of Russia (Unique identifier of the project RFMEFI62117X0015).

References

1. Gerlinskaya L.A., Evsikov V.I. (2001) Influence of genetic dissimilarity of mother and fetus on progesterone concentrations in pregnant mice and adaptive features of offspring. Reproduction. 121(3): 409-17.
2. Bracamonte-Baran W., Burlingham W. (2015) Non-inherited maternal antigens, pregnancy, and allotolerance. Biomed J. 38(1): 39-51.

DOI 10.18699/BGRSSB-2018-179

Selection for behavior, intermale confrontations, and corticosterone and testosterone levels in the blood of norway rats

R. Gulevich[1]*, R. Kozhemyakina[1], M. Konoshenko[2], S. Shikhevich[1]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Institute of Chemical Biology and Fundamental Medicine, Novosibirsk, Russia*
** e-mail: gulevich@bionet.nsc.ru*

Key words: aggressiveness, tameness, selection, rats, confrontations, neutral area

Motivation and Aim: Aggression is believed to be promoted by various factors: hereditary, environmental, developmental, hormonal, and neurotransmitter-related. In some scientists' view, testosterone contributes to aggression manifestation, and this contribution increases with an imbalance of other factors, such as low cortisol levels. This work concerns the effect of selection for attitude towards humans on the dynamics of blood testosterone, corticosterone after the test for intermale aggression in an unfamiliar cage and on the parameters of agonistic contacts under the same conditions in tame, aggressive, and unselected rats.
Methods and Algorithms: Experiments were conducted with two-month-old rat males (*Rattus norvegicus*) of the 79th generation of selection for absence and enhancement of the aggressive–fearful response to humans. Intermale aggression was tested in an unfamiliar transparent plastic cage, The cage was partitioned into two compartments. Male rats were placed by ones into the compartments, and the partition was removed. Eighteen tests were done: six between unselected animals, six between aggressive, and six between tame. Behavior parameters were recorded for 5 min: attack latency; number and durations of attacks, chasings, kickings, upright postures, pinning, wrestling, aggressive grooming, lateral threats. Blood was sampled from the tail tip immediately and within 30, 60, and 120 min after the intermale aggression test.
Results: The attack latency is longer in tame males than in aggressive or unselected ones, whereas the number and duration of aggressive behaviors are less. Confrontations between unselected males in a neutral area lasted for longer than between aggressive ones. After confrontations, the level of corticosterone increases, being higher in unselected animals than in aggressive or tame within two hours after the test, as well as in the intact state. Although blood testosterone levels show no significant differences among the three groups prior to the test, tame rats are inferior in this index to unselected and aggressive animals immediately after the test, and only to aggressive animals within 30 min after a confrontation.
Conclusion: Our data on the parameters of behavior in a neutral area indicate that aggressiveness in tame males is reduced in comparison to not only aggressive animals, but also to unselected ones. The fact that confrontations between unselected males are longer than between aggressive may be determined by the selection criterion itself, because the original response to a glove was related to the latency time rather than aggression duration. Post-confrontation corticosterone in unselected rats was higher than in others and post-confrontation testosterone in tame rats was lower than in other ones.
Acknowledgements: Supported by the State Budgeted Project 0324-2018-0016.

DOI 10.18699/BGRSSB-2018-180

RNA-Seq for *Danio rerio*, exposed to pulp and paper wastewaters

E. Gusakova
Northern Arctic Federal University, Arkhangelsk, Russia
e-mail: helengusakova@gmail.com

Key words: *Danio rerio*, wastewaters, effluents, transcriptome, RNA

Motivation and Aim: The wood processing industry (mainly pulp and paper complexes) occupy third place in the amount of waste water produced [1]. The waste water of the wood processing industry is contaminated with both organic and non-organic compounds (lignins, tannins, chlorine derivatives etc.) [2] typically passed mechanical, physical, chemical as well as microbiological treatment to remove or decrease the abundance of these compounds. However, imperfect treatment regularly leads to the leakage of contaminants, which may have a significant impact on the environment, especially on neighboring water systems. The associated risks particularly for fish include the growth of harmful bacteria decreasing dissolved oxygen levels and leading to fish anoxia, low visibility because of low water transparency, a reduction in the abundance of feedstock (invertebrates) and hence fish population densities, the accumulation of trace and toxic elements in fish, which could be transferred to human, physiological changes [3, 4], as well as evolutionary responses particularly in immune and metabolic pathways. The goal of the project is to characterize the transcriptome response of zebrafish to the contamination from the wood processing wastewaters.
Methods and Algorithms: Fishes are exposed to wastewaters with variable concentrations. For acute experiment fishes are exposed for 96 hours to detect acute toxicological effect on fishes mortality, and limiting concentration of the wastewaters. For chronical toxicological effect, fishes are exposed to wastewaters with concentrations less then acute concentration for 30 days. By the end of exposure, survived fishes are prepared for RNA-Sequencing.
Results and Conclusion: in process.

References

1. Savant D.V., Abdul-Rahman R., Ranade D.R. (2006) Anaerobic degradation of adsorbable organic halides (AOX) from pulp and paper industry wastewater. Bioresour Technol. 97(9):1092-1104.
2. Buzzini A.P., Miwa D.W., Motheo A.J., Pires E.C. (2006) Use of electrochemical oxidation process as post-treatment for the effluents of a UASB reactor treating cellulose pulp mill wastewater. Water Sci Technol. 54(2):207-213.
3. ESAMP (1991) Reducing environmental impacts of coastal aquaculture. IMO/FAO/Unesco/WMO/WHO/IAEA/UN/UNEP Joint Group of Experts on the Scientific Aspects of Marine Pollution. 47:35.
4. Richardson K. (2003) Anthropogenically induced changes in the environment: Effect on fisheries. Responsible Fisheries in the Marine Ecosystem. 275-289.

DOI 10.18699/BGRSSB-2018-181

Hippocampal differentially expressed genes between tame and aggressive foxes are included in pathways associated with stress, behavior and adult neurogenesis

Yu. Herbeck*, M. Genaev, G. Vasiliev, R. Gulevich, S. Shikhevich, D. Shepeleva, D. Afonnikov, L. Trut
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: herbek@bionet.nsc.ru*

Key words: domestication, stress, NGS

Motivation and Aim: The experimental domestication of foxes has demonstrated that essential mechanism of domestication is the decrease of stress response in particular toward human. Previously it was shown reduced plasma basal and stress-induced cortisol levels in tame foxes [1]. However, the molecular genetic mechanisms of these changes remain unclear. The aim of this work is to identify differentially expressed genes (DEG) in the hippocampus of domesticated ('tame') and aggressive foxes. The hippocampus is the key site of glucocorticoid negative feedback [2].
Methods and Algorithms: Single-read 75 bp sequencing of hippocampal cDNA libraries was performed using Illumina HiSeq ($n = 3$ foxes per behavior group, on average 40 mln reads per sample). The DEG analysis was performed by the Cufflink (cole-trapnell-lab.github.io/cufflinks). The dog genome was used as a reference, and Ensemble rev. 86 was used as an annotation. Over-representation analysis was performed using web-based gene set analysis toolkit (WebGestalt) (www.webgestalt.org).
Results: Analysis of 496 DEGs detected in the hippocampus of tame and aggressive foxes revealed only 10 pathways significantly associated with DEG. Most of these pathways are associated with the nervous system, behavior, aggression, stress and neurogenesis.
Conclusion: Apparently, the differences in gene expression in the hippocampus between tame and aggressive foxes are primarily associated with signal transmission (cAMP and calcium signaling pathways), neurogenesis and axon guidance, adrenergic signaling, opioid system (amphetamine addiction). These pathways and individual DEGs play an important role in stress reactivity, behavior and adult neurogenesis.
Acknowledgements: Supported by the RSF (No. 16-14-10216).

References

1. Trut L.N. et al. (2009) Animal evolution during domestication: the domesticated fox as a model. BioEssays. 31(3):349-360.
2. Jacobson L., Sapolsky R.M. (1991) The role of the hippocampus in feedback regulation of the hypothalamo-pituitary-adrenocortical axis. Endocr. Rev. 12:118-134.

DOI 10.18699/BGRSSB-2018-182

FGF21 signaling and brown adipose activity gene expressions in male and female mice under fasting and refeeding states

T. Iakovleva[1]*, N. Balybina[2], N. Makarova[1], N. Bazhan[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: jakov@bionet.nsc.ru*

Key words: brown adipose tissue, FGF21 signaling, starvation, refeeding, *C57BL mice*

Motivation and Aim: Brown adipose tissue (BAT) is involved in the body adaptation to starvation and increased food consumption. The main function of BAT is to produce heat. Thermogenesis in BAT is mediated through the BAT-specific uncoupling protein 1 (UCP1). Fasting suppresses UCP1 expression in BAT [1]. Several studies showed that females have more active BAT because sex hormones affect UCP1 expression as well as the adrenergic receptor levels, lipolytic activity and lipid accumulation in brown adipocytes [2]. Fibroblast growth factor 21 (FGF21) recently discovered regulator of lipid and glucose metabolism activates brown adipocytes inducing thermogenic gene expression and caloric expenditure during drug treatment [3]. It is unknown how sex hormones affect on FGF21 signaling in various metabolic states. We investigated the effects of fasting and refeeding on expression of genes of FGF21 signaling and BAT activity in males and females mice.

Methods and Algorithms: Males and females of C57BL mice were decapitated after 24 hours of fasting or after 24 hours of fasting plus 6 hours of refeeding or in the fed state (controle). Plasma hormone concentrations were measured using commercial kits, glucose blood level was determined using a glucometer OneTouch Select, gene expression was evaluated using the qPCR method.

Results: There were no sex differences in alterations of blood hormone levels and BAT mRNA levels in response to fasting. Fasting increased *Ppargc1a* mRNA levels and did not affect the expression of other genes equally in males and females. There were sex differences in hormonal reaction to refeeding because insulin levels were greater and leptin levels were lower in males compared with females although there were no sex differences of the hormone levels in control. Refeeding decreased *Ppargc1a* and *Slc2a1* mRNA levels both in males and females. Furthermore refeeding increased FGF21 mRNA levels and it was significantly lower in males compared to females.

Conclusion: There were sex differences only in FGF21 mRNA levels in response to refeeding. Differential expression of FGF21 might contribute to sex differences in BAT activity, but further research is needed.

Acknowledgements: Supported by the Russian Science Foundation, grant No. 17-15-01036.

References

1. Nedergaard J. et al. (2001) UCP1: the only protein able to mediate adaptive non-shivering thermogenesis and metabolic ineffiency. Biochimica Biophysica Acta. 1504:82-106.
2. Monjo M. et al. (2003) Direct effects of testosterone, 17 beta-estradiol, and progesterone on adrenergic regulation in cultured brown adipocytes: potential mechanism for gender-dependent thermogenesis. Endocrinology. 144(11):4923-30.
3. Fisher M. et al. (2012) FGF21 regulates PGC-1a and browning of white adipose tissues in adaptive thermogenesis. Genes Development. 26:271-281.

DOI 10.18699/BGRSSB-2018-183

Genome-wide association study for body temperature maintenance under the cold stress in Siberian cattle

A.V. Igoshin[1]*, N.M. Belonogova[1], A.A. Yurchenko[1], N.S. Yudin[1, 2], D.V. Petrovsky[1], D.M. Larkin[1, 3]

[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Royal Veterinary College, University of London, London, UK*
** e-mail: igoshin@bionet.nsc.ru*

Key words: body temperature, GWAS, cattle, adaptation, cold climate

Motivation and Aim: There are over 1000 cattle breeds existing worldwide, some of which dwell in local environments characterised by very cold winter temperatures suggesting selection pressure for cold climate tolerance. While being of economic significance, the genetic basis of cold tolerance in cattle is not well understood. Therefore, we performed a genome-wide association study (GWAS) for body temperature maintenance under cold stress in Hereford and Kazakh Whitehead cattle breeds bred in Siberia.
Methods and Algorithms: A total of 183 dams and bulls were used for hourly ear canal temperature measurements for two weeks in February 2017. The area under the curve of body temperature over the period of five coldest days (from –20 to –30 °C) within the two-week interval was considered as the phenotype for each animal. DNA was extracted from blood and genotyped on the GeneSeek Bovine GGP HD150K commercial SNP array. After data filtering in plink: --maf 0.05, --chr1-29, 108298 SNPs were used in association studies. We performed a single SNP GWAS using EMMAX software with sex and breed being taken as covariates. Then the data were phased with fastPhase software, haploblocks were defined in Haploview (D' > 0.8) and the haplotype trend regression (HTR) analysis (R package "gap", function "htr") was performed to test for association between haploblocks and the range of phenotype measurements in the dataset applying the same covariates. In addition, signatures of selections were identified in the whole dataset of 183 individuals using the de-correlated composite of multiple signals (DCMS) framework combining the H1, H12, Tajima's D and nucleotide diversity statistics.
Results: Out of two SNPs found above the suggestive significance threshold (q-value < 0.1) from the EMMAX results, one (BovineHD1500000472) was found within one of two haploblocks reported significant by HTR (q-value < 0.05). The same interval on BTA15: 1.58–2.03 Mbp was reported as a putative signature of selection by DCMS. This interval contains four genes of which two: *MSANTD4* and *GRIA4* are functional candidates for body temperature-related traits.
Conclusion: A genomic interval of 450 Kbp on BTA15 is associated with temperature maintenance in Hereford and Kazakh Whitehead cattle in Siberia under the cold temperature stress. This region contains two relevant functional candidate genes and requires further studies (e. g., resequencing and RNASeq analysis of relevant tissues) to confirm its role and to identify causative variants.
Acknowledgements: This study was supported by the grant from the Russian Science Foundation (project No. 16-14-00090).

DOI 10.18699/BGRSSB-2018-184

Exploring neuroprotective potential of astrocytes

S. Kasparov
University of Bristol, United Kingdom
** e-mail: bristolbabajaga@gmail.com*

Key words: oxidative stress, astrocytes, neuroprotective pathways, ligand-receptor interaction

Our laboratories are interested in signaling systems and novel receptors expressed by astrocytes which could be used as potential novel drug targets. Discovery of neuroprotective pathways is one of the major priorities for neuroscience. Astrocytes are the natural neuroprotectors and it is likely that brain resilience can be enhanced by mobilizing their protective potential. Among G-protein coupled receptors expressed by astrocytes, two highly related receptors, GPR37L1 and GPR37, are of particular interest. Previous studies suggested that these receptors are activated by a peptide Saposin C and its neuroactive fragments (such as prosaptide TX14), which were demonstrated to be neuroprotective in various animal models by several groups. However, pairing of Saposin C or prosaptides with GPR37L1/GPR37 has been challenged and presently GPR37L1/GPR37 have regained their orphan status. Here we demonstrate that in their natural habitat, astrocytes, these receptors mediate a range of effects of TX14, including protection from oxidative stress. The Saposin C/GPR37L1/GPR37 pathway is also involved in the neuroprotective effect of astrocytes on neurons subjected to oxidative stress. The action of TX14 is at least partially mediated by Gi-proteins and the cAMP-PKA axis. On the other hand, when recombinant GPR37L1 or GPR37 are expressed in HEK293 cells, they are not functional and do not respond to TX14, which explains unsuccessful attempts to confirm the ligand-receptor pairing. Therefore this study identifies GPR37L1/GPR37 as the receptors for TX14, and, by extension of Saposin C, and paves the way for the development of neuroprotective therapeutics acting via these receptors.

DOI 10.18699/BGRSSB-2018-185

Germ cell migration under GAGA-factor control

A. Khruscheva[1, 2], K. Ahmetova[1], S. Fedorova[1, 2], E. Baricheva[1], N. Dorogova[1]*
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk National Research State University, Novosibirsk, Russia*
** e-mail: dorogova@bionet.nsc.ru*

Key words: germ cells, cell migration, GAGA-factor, Trl gene, Drosophila

Motivation and Aim: Germ cells migration is an important stage of gonadogenesis and normal development of precursors of germ cells. This process occurs in early embryogenesis and it is carried out by the ability of the germ cells (GC) to move individually and actively, using specialized structures and penetrating the surrounding cellular barrier (epithelium). Molecular-genetic nature of GCs migration is evolutionarily conservative and has high homology in different cell types, organs and tissues, and it is similar in many aspects with the movement of the immune system cells and metastatic tumors. This circumstance explains the interest in this research topic and confirms its relevance.
Methods and Algorithms: We used the genetic model of Drosophila melanogaster, cytological methods and bioinformatic analysis.
Results: In this study, we performed a detailed cytological analysis of the early stages of germ cell development in the mutant background of the Trl gene that encodes GAGA-factor. We showed that the mutation causes premature morphological transformation of the early germ cells into actively migrating cells and, consequently, promote their ectopic migration inside the embryo. The first signs of active migration in mutants observed already at the stage of the cellular blastoderm. Part of the germ cells becomed amoeboid and migrated through a single-layered epithelium. These cells migrated chaotically, disoriented and did not reach the gonadal region. Premature migration leads to a reduction in the germ cells number in the adult flies. The character of Trl gene expression indicates its activity in epithelial blastoderm cells, some of which contact with primordial germ cells. It is the most likely that anomalies in the migration are associated with a mutant effect in surrounding somatic epithelial cells. Since the GAGA-factor regulated the expression of diversity genes, its effect mediated through the activation of target genes. We analyzed both expression patterns and regulatory regions of genes that expressed during embryonic GC migration and revealed 25 potential GAGA-target genes.
Conclusion: Thus, the GAGA-factor influences the migration of the embryonic GCs through the regulation of target gene expression in their somatic environment. The mutant effect in somatic cells led to an early activation of the migration program and the premature trans-epithelial GCs migration. The phenomenon of premature migration demonstrates that the somatic environment, in particular the epithelial cells, do not simply form a substrate for GCs migration, but can also influence this process in a regulatory manner.
Acknowledgements: Supported by the RFBR 18-34-00321mol_a (A. Khruscheva, K. Ahmetova) and the State Project No. 0324-2018-0019 (S. Fedorova, E. Baricheva, N. Dorogova).

DOI 10.18699/BGRSSB-2018-186

Oxidative stress and nitric oxide synthesis in ISIAH rats with inherited stress-induced arterial hypertension

L. Klimov[1, 2]*, M. Zhuravina[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: maple1708@mail.ru*

Key words: ISIAH rats, hypertension, NO generation

Motivation and Aim: Hypertensive disease is widespread in modern society. It can be the cause of myocardial infarction, stroke and heart failure. One of the factors in the development of the disease is chronic emotional stress. The ISIAH rat strain (Inherited Stress-Induced Arterial Hypertension) is a rat model with the genetically determined enhanced responsiveness to stressful stimulation. The current population of the ISIAH rats is characterized by elevatedof both the basal arterial BP which reachesup to 175.0 ± 3.5 mmHg in males.

In addition to the four major systems responsible for BP control, there are number of local systems, in particular, the paracrine vasodilator – nitric oxide. Previously negative correlation of the bioavailability of nitric oxide in plasma and the level of BP in ISIAH rats was shown. The bioavailability of nitric oxide and its content in the kidneys and vessels is affected not only by the expression of the eNOS gene, but also by the concentration of reactive oxygen species (ROS) within the cell. The constitutive formation of ROS is necessary for the normal functioning of cells. Oxidative stress arises from an imbalance between antioxidant enzymes and enzymes that metabolize ROS, which leads to an increase in the formation of ROS. As a result, endothelium-dependent vasodilation and water-salt homeostasis can change, as ROS decreases the bioavailability of NO. Also previously it was shown a decrease in the activity of the antioxidant enzyme superoxide dismutase in blood plasma and a reduced concentration of the reduced form of glutathione in the blood plasma of ISIAH rats.

Methods and Algorithms: The relative amount of target mRNA was measured by qPCR. Lymphocytes were analyzed in 3-month old ISIAH and WAG rats (6 animals in each group). Lymphocytes were isolated from the blood plasma by centrifugation with LSM. Total RNA was extracted from lymphocytes and caudal arteries using the TRI reagent (Molecular research center, USA). Remaining traces of genomic DNA were removed from the RNA samples using DNase I (Promega, USA) treatment, according to the manufacturer's instructions. The value for the target gene was further normalized against the qPCR level of the reference gene.

Results: We found changes in the level of mRNA genes responsible for the synthesis of ROS in the lymphocytes and vessels of the ISIAH rats. Therefore we can assume a possible imbalance in the synthesis and metabolism of ROS, and as a result, the participation of oxidative stress mechanisms in the development of a stress-induced form of GB.

Acknowledgements: This work was supported by the Ministry of Education and Science of Russia (project No. 0324-2018-0016).

DOI 10.18699/BGRSSB-2018-187

Comparative experimental analysis of the reproductive potential and sexual behavior of house mice from the transcaucasian hybrid zone and *Mus musculus*

E. Kotenkova*, A. Maltsev, A. Ambaryan
Severtsov Institute of Ecology and Evolution RAS, Moscow, Russia
** e-mail: evkotenkova@yandex.ru*

Key words: house mice, reproduction, closely related species, hybridization, hybrid zone

Motivation and Aim: Traditionally hybridization is considered as negative process because hybrids have genomic disruption and as result lower fitness than either parental genotypes. We proceeded from the hypothesis of Milishnikov et al. (2004) that there is large and genetically complicate hybrid zone in Transcaucasia. In this region, in different historical periods, three forms of house mice were hybridized: the first one is genetically close to *Mus musculus* s. str. preserved a relict gene pool; differentiated *M. musculus* and *M. domesticus*. In the laboratory, we investigated the fecundity, fertility and sexual behavior of house mice from Transcaucasian populations and *M. musculus* to evaluate the fitness of natural hybrids and development of reproductive isolating mechanisms.
Methods and Algorithms: The breeding parameters were investigated by the method of experimental hybridization. Additionally we investigated weight of testicles, sperm quality and concentration. 90 min dyadic encounters of con- and heterospecific males and estrous females were conducted. Behavior of sexual partners was recorded by a video camera and analyzed by means of Observer Video-Pro, Version 4.1.
Results: Reproductive intensity and indicators of fertility of house mice from Moscow, the Moscow region and from the zone of hybridization in Transcaucasia was similar. *M. musculus* and Transcaucasian mice were crossed easy in the laboratory. Viability and fertility of F_1 were ordinary, mortality was higher in one variant of crosses. The sperm quality was reduced lightly only in F_2 hybrids. In conspecific encounters males, as well as females of *M. musculus* exhibited affiliative behavior frequently and longer than females of Transcaucasian mice. Differences in duration and frequency of copulatory behavior in total and elements of this behavior in two forms of house mice were not statistically significant. In heterospecific encounters, females as well as males of *M. musculus*, exhibited affiliative behavior less than in encounters with conspecific partners. This also concerns of grooming, naso-nasal contacts and overall duration and frequency of affiliative behavior in female *M. musculus*.
Conclusion: The data did not support the reduced fitness of Transcaucasian hybrids and the effect of isolating mechanisms between mice from Transcaucasia and *M. musculus*. The results confirm genetic relationship of these forms.
Acknowledgements: Supported by the Russian Foundation for Basic Research project 16-04-00149.

DOI 10.18699/BGRSSB-2018-188

Effect of early experience on neuronal and behavioral responses to con- and heterospecific odors in three closely related Mus taxa: epigenetic contribution in formation of precopulatory isolation

E. Kotenkova[1]*, A. Maltsev[1], A. Ambaryan[1], A. Romachenko[2]
[1] *Severtsov Institute of Ecology and Evolution RAS, Moscow, Russia*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
** e-mail: evkotenkova@yandex.ru*

Key words: *Mus musculus*, *M. spicilegus*, cross-fostering, early olfactory experience, main olfactory bulb, accessory olfactory bulb

Motivation and Aim: The individual learned phenotypic traits of parents and/or siblings, such as olfactory ones, result in the learner being able to discriminate its own species and sex of conspecifics. Olfactory plasticity to main social olfactory cues is limited to a critical period when exposure to odor might change responses to con- and heterospecific odors. The olfactory system thus provides an attractive model to investigate processes involving interplay between genetic and epigenetic influences and their role in evolutionary process, especially development of precopulatory reproductive isolation. We alter maternal environment by cross-fostering. The objective was to evaluate the influence of early olfactory experience on the neuronal and behavioral response of males to con- and heterospecific odors of receptive females in two species *M. musculus* (subspecies musculus, wagneri) and *M. spicilegus* and thus, to determine the potential role of epigenetic contribution in formation of precopulatory isolation.
Methods and Algorithms: Males were reciprocally cross-fostered shortly after the birth and were tested for response to con- and heterospecific urine odors of estrous females using two-choice tests at 70–85 days of age. Neuronal activity of non- and cross-fostered males were evaluated at 90–110 days of age in main (MOB) and accessory olfactory bulbs (AOB) to con- and heterospecific female odor using fMRI (MEMRI).
Results: Non-fostered males of three taxa demonstrated a strong preference for odor of conspecific females. Male spicilegus raised by female musculus spent significantly more time investigating odor sources of heterospecific females. Wagneri-nursed spicilegus did not demonstrate significant choice of con – or heterospecific female odor. Non-fostered male spicilegus investigated longer odor of conspecific females in comparison with wagneri-nursed males. Wagneri-nursed spicilegus investigated more time urine odor of female wagneri in comparison with non-fostered male spicilegus. The level of MRI signal obtained from the evaluation of manganese accumulation in AOB neurons was significantly higher when the odor of urine of conspecific estrus females was exposed compared to urine exposure of heterospecific females. The response pattern changed to the opposite in males raised by heterospecific females. The maternal environment, including odor, had a greater effect on the level of MRI signal in the AOB than the genetic relationships of the recipient and the donor of the odor stimulus.
Conclusion: Behavioral and neuronal responses to con- and heterospecific odors changed in three closely related taxa of Mus as result of early experience. We demonstrated importance of early leaning in mate choice in adulthood in mice and possibility of epigenetic contribution in formation of precopulatory isolation.
Acknowledgements: Supported by the Russian Science Foundation grant No. 16-14-10269.

DOI 10.18699/BGRSSB-2018-189

Features of the autophagy process during its induction by 48-h fasting and inhibition by chloroquine in the rat liver

O.S. Kozhevnikova*, D.V. Telegina, E.A. Zolotenkova, M.A. Tyumentsev, N.G. Kolosova
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: oidopova@bionet.nsc.ru*

Key words: liver, autophagy, chloroquine, senescent-accelerated OXYS rats

Motivation and Aim: Autophagy is a multi-stage process of delivery of cytoplasmic material to lysosomes for the subsequent degradation. It is the main mechanism of degradation of long-lived proteins and the only one – for organelle degradation. This process is involved in the maintenance of cellular homeostasis, the removal of damaged proteins and organelles, and is necessary to maintain cell metabolism in conditions of energy and nutrient deficiency. With age, the intensity of autophagy decreases. It is known that the pathogenesis of many age-related diseases, including neurodegenerative diseases, is associated with a disruption of autophagy, but the underlying mechanisms of these disorders have not been adequately studied. The liver plays a central role in controlling glucose and lipid homeostasis in response to fasting and feeding. The present work is aimed at studying the contribution of autophagy changes in the early development of age-dependent diseases in OXYS rats, a unique model of premature aging. Its purpose is to study the features of the process of autophagy when it is induced and inhibited in the liver of OXYS rats.
Methods: The work was performed on male OXYS and Wistar (control) rats at the age of 4 months (n = 60). 48-h fasting was used for the induction of autophagy, and the chloroquine injections (50 mg/kg body mass) were used for its inhibition. The content of protein markers of autophagy (ATG7, p62 and LC3) was assessed by Western blot analysis and immunohistochemistry, autophagosome formation by means of electron microscopy (Microscopy Centre of ICG SB RAS).
Results: It was shown that the administration of chloroquine reduced the liver mass of rats of both strains, the body weight - only in OXYS rats. The chloroquine injections on a fasting background slowed down body weight reduction in Wistar rats, but not in OXYS rats. By the Sanger sequencing we confirmed the presence of non-synonymous single nucleotide substitution in the Pik3c2b gene in the genome of OXYS rats. The product of this gene is involved in the regulation of autophagy. According to our preliminary data, the activity of autophagy in the liver of OXYS rats has been reduced already at the young age.
Acknowledgements: Supported by the RSF 18-75-00031.

DOI 10.18699/BGRSSB-2018-190

Aldosterone mineralocorticoid receptors expression in male Ay mice

N.S. Logvinenko*, L.E. Katkova, G.S. Baturina
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: ninlo@bionet.nsc.ru*

Key words: aldosterone, kidney, obesity, mineralocorticoid receptor

Motivation and Aim: Mice with melanocortin type of obesity, heterozygous for the dominant lethal mutation Agouti-yellow, (Ay-mice) are a convenient model for studying molecular mechanisms of obesity development. Obesity and complications associated with it often lead to the chronic renal failure, which is one of the main causes of death in the late stages of the syndrome. However, the basic molecular characteristics of the mineralocorticoid system at this type of genetically determined obesity have not been adequately studied. Therefore, the aim of this work was to study the expression of mineralocorticoid receptor and aldosterone level in Ay mice.
Methods and Algorithms: Using the real-time PCR method the mRNA level of the mineralocorticoid receptors (MR) in the hypothalamus, kidneys, heart and adipocytes in adult male mice of the standard C57Bl/6J and the Ay mice at the age of 29–30 weeks was investigated. At this age, Ay mice show a non-dietary type of obesity. The level of aldosterone in the blood of both lines was studied using the enzyme immunoassay method (Mouse Aldosterone (ALD) ELISA kit).
Results: We have detected a higher level of MR mRNA in the kidney cortex and in the heart left ventricle in C57Bl/6J mice, compared to the line Ay ($p \leq 0.05$). Perhaps it is due to a more active aldosterone-dependent protein regulation in these target tissues in control animals. No significant differences in MR mRNA level in the pituitary glands and adipocytes in mice of both lines have been identified. We have shown that there was no difference in the plasma level of aldosterone in the blood of both lines (124.1 ± 25.1 and 102.8 ± 16.5 pg/ml in the male Ay and C57Bl/6J, respectively, $p \geq 0.05$).
Conclusion: We suggest that the absence of significant differences in the expression of mineralocorticoid receptors in adipocytes and the pituitary gland and in the level of aldosterone in blood of male Ay and C57Bl/6J is due to the gender peculiarities of the mineralocorticoid system participation in the development of melanocortin type of obesity.
Acknowledgements: Supported by the RFBR (17-04-00912) and Budget Project No. 0324-2018-0016.

DOI 10.18699/BGRSSB-2018-191

Methylation of DNA in Colorado potato beetle *Leptinotarsa decemlineata* (Say)

Yu.M. Nikonorov, G.V. Benkovskaya*
Institute of Biochemistry and Genetics, Ufa Federal Research Centre of RAS, Ufa, Russia
** e-mail: bengal2@yandex.ru*

Key words: DNA methylation, Colorado potato beetle, ecdysone receptor gene

Motivation and Aim: Enzymatic methylation-demethylation of DNA is of great importance for gene activity regulation in plants and animals. For a number of investigated insect species it is typical the presence in genome of genes with both low and high levels of methylation in coding regions, that is connected to a certain extent of reducing the elements of methylation DNA systems. Our aim is to investigate the methylation status of some genes of Colorado potato beetle *Leptinotarsa decemlineata* and to evaluate their transcriptional activity at the separate phases of winter diapause.
Methods and Algorithms: S-adenosyl methionine (SAM) and 20-hydroxyecdysone (20E) have been applied in concentration of 1×10^{-7} M for treatment of fresh potato leaves used to feed Colorado potato beetle adults from native excerpts during 7 days. DNA and RNA extracted from tissues of thorax and gonads of adults at the stages of winter diapause initiation and diapause termination by phenol-detergent method. Relative abundance of mRNAs of genes ecdysone receptor *ecr*, enzyme acetylcholinesterases *ache* and protein transporter ferritin *fer* has been evaluated by qRT-PCR with reference gene *rp18*. Level of DNA methylation in coding regions of mentioned genes determined by MSRE-PCR (methylation-sensitive restriction enzyme PCR).
Results: SAM and 20E addition to the food at the stage of diapause initiation had an adaptogenic effect under partial starvation (due to substitution of potato leaves to honey solution) and intoxication. The reducing of gene *ecr* mRNA ratio registered between tissues of gonads and thorax of active males and females fed by food with SAM and 20E addition.
Stimulatory action of SAM revealed toward the gene *fer* mRNA content in gonads of adults treated by insecticides. Content of 5-methylcytosine (5 mC) in DNA sites of restrictase HpaII localized in investigated genes *ache* and *fer* coding regions amounted no more than 5 %. Proximal site against 5′-end of *ecr* gene has been methylated to a greater extent than the distal site and in males it has been methylated to a greater extent than in females. SAM became apparent as an inductor of DNA demethylation process in proximal site of *ecr* gene in males. High correlation level ($r = 0.86$) determined between methylation of DNA level in proximal site of *ecr* gene and abundance of its mRNA.
Conclusion: We revealed the coordinated change of methylation status of ecdysone receptor gene *ecr* in Colorado potato beetle adults and its transcriptional activity during the winter diapause and the sensitivity of insect to presence in the food of methyl groups.
Acknowledgements: Investigations funded by RFBR No. 15-04-04801-a and 17-44-020347-r_a.

DOI 10.18699/BGRSSB-2018-192

Lethal yellow (A^Y) mutation in the *agouti* gene causes the depressive-like alterations in the mouse brain and behavior

A.V. Plyusnina[1]*, N.V. Khotskin[1], E.A. Kulikova[1], E.Y. Bazhenova[1], D.V. Fursenko[1], I.E. Sorokin[1], I. Kolotygin[2], O.B. Shevelev[1], A.V. Kulikov[1]
[1] *Institute of Cytology and Genetic SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
* *e-mail: plyusninaav@bionet.nsc.ru*

Key words: lethal yellow, brain, behavior, MRI, mice, depressive-like alterations

Motivation and Aim: The *agouti* gene is expressed only in cells of the hair follicle and it induces yellow pigment (phaeomelanine) synthesis by melanocytes [1]. The lethal yellow (A^Y) mutation results from the large deletion in the promotor of the mouse agouti gene that put the agouti gene under control of the promotor of a ubiquitously expressed Raly gene. In addition, A^Y mutation causes ectopic expression of the agouti protein in many tissues including the brain, adipose and other tissues. The agouti protein is an inhibitor of the melanocortin-4 receptors involving in the regulation of total metabolism and feeding behavior. So, A^Y mutation causes obesity and diabetes II alteration in mice [2]. The aim of this study is the effect of the A^Y mutation on the brain and behavior.
Methods and Algorithms: The experiments were carried out on adult (11–12 weeks old) males of A^Y/a mice and their wild-type counterparts (*a/a*).
Results: Mice of A^Y/a and *a/a* genotypes did not differ in their home cage activity, sleep, food and water consumption, learning ability in the Morris water maze, anxiety in the open field and elevated plus-maze, as well as in the level and metabolism of monoamines and expression of some proinflammatory genes in the brain. At the same time, A^Y/a showed elevated fat mass ($F_{1.14} = 46.3$, $p < 0.0001$) and depressive-like behavior in the forced swim test ($F_{1.14} = 11.85$, $p < 0.01$) compared with *a/a* mice. MRI revealed a reduction of cortex volume ($F_{1.14} = 13.65$, $p < 0.01$), while MR spectroscopy showed a shift the balance between excitatory and inhibitory substances to excitatory substances in the hippocampus in A^Y/a mice. The level of mRNA of *Ptpn5* gene encoding striatal enriched protein tyrosine phosphatase in the frontal cortex of A^Y/a mice was elevated compared with their wild-type counterparts ($F_{1.13} = 10.71$, $p < 0.01$).
Conclusion: In the present study we first investigate the effects of A^Y mutation on the mouse brain and behavior. So, the A^Y mutation precipitated depressive-like alterations in the behavior and brain functions and A^Y/a mice are a promising model of depressive disorders induce by metabolic dysfunction.
Acknowledgements: Supported by the RSF (grant No. 17-15-01032) with using of equipment of the Research Center supported by the Russian Ministery of Education and Science (project No. RFMEFI62117X0015). The maintenance of mice was supported by the basic research project No. 0324-2018-0016.

References:

1. Perry W.L., Copeland N.G., Jenkins N.A. (1994) The molecular basis for dominant yellow agouti coat color mutations. BioEssays. 16(10):705-707.
2. Boston B.A., Blaydon K.M., Varnerin J., Cone R.D. (1997) Independent and additive effects of central POMC and leptin pathways on murine obesity. Science. 278(5343):1641-1644.

DOI 10.18699/BGRSSB-2018-193

The moonlighting functions of the NON3 protein in *Drosophila melanogaster*

J. Popova[1, 2]*, G. Pavlova[1], E. Andreyeva[1], A. Ogienko[1], A. Yushkova[1, 3], A. Ivankin[1], E. Kozhevnikova[1, 2], A. Pindyurin[1, 2, 3]

[1] *Institute of Molecular and Cellular Biology SB RAS, Novosibirsk, Russia*

[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

[3] *Novosibirsk State University, Novosibirsk, Russia*

** e-mail: popova@mcb.nsc.ru*

Key words: mitosis, chromatin, *Drosophila melanogaster*, microtubules, mitotic spindle, kinetochore-driven microtubule growth, NON3, CID, nucleolus

Motivation and Aim: It was previously shown that a number of proteins, which are components of the nucleolus and necessary for the biogenesis of ribosomes, are also involved in the mitotic spindle assembly in Drosophila S2 cells [1]. The molecular mechanism of the latter process currently is not fully understood. Our aim was to clarify the role of *D. melanogaster* NON3 (Novel nucleolar protein 3) protein on kinetochore-driven microtubule growth.

Methods and Algorithms: We generated polyclonal antibodies specific to NON3. Using *P*-element imprecise excision, we obtained a set of new *Non3* mutations. They were sequenced and analyzed by complementation analysis. To study microtubule regrowth from kinetochores in cultured S2 cells, we performed RNA-interference (RNAi) to deplete NON3 and used colcemid treatment.

Results: We describe the viability and fertility of generated *Non3* alleles: null-mutations are early embryonic recessive lethal, hypomorphic mutations survive to adult stage and are semi-sterile, precise excisions (control) are fully viable and fertile. We found that NON3 protein is a component of nucleolus and pericentric regions of chromosomes. NON3 depletion after RNAi in S2 cells results in (1) a short mitotic spindle, (2) formation of metaphases with disorganized spindle and (3) appearance of pseudo-anatelophases (anatelophases, in which the spindle looks like at the last stages of division, but sister chromatids didn't separate to the poles of the spindle). The colcemid treatment of NON3-depleted S2 cells affects microtubule regrowth from kinetochores. A lack of NON3 protein (due to mutations or RNAi) affects the loading of drosophila CENP-A protein (CID) at centromeric regions of chromosomes, which may be a reason of the short spindle in NON3-depleted S2 cells.

Conclusion: Our results suggest that the NON3 protein affects kinetochore-driven microtubule growth due to the disruption of CID loading to centromeric chromatin.

Acknowledgements: Supported by the RFBR (18-34-00699).

References

1. Moutinho-Pereira S. et al. (2013) Genes involved in centrosome-independent mitotic spindle assembly in Drosophila S2 cells. Proc. Natl. Acad. Sci. USA. 110:19808-19813.

DOI 10.18699/BGRSSB-2018-194

Taurine affects the predisposition to audiogenic epilepsy in PM rats

O.I. Prokudina[1]*, T.A. Alekhina[1], A.V. Solovyeva[2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Altai State University, Barnaul, Russia*
** e-mail: petrenko@bionet.nsc.ru*

Key words: audiogenic seizure, epilepsy, taurine, animal model, PM rats

Motivation and Aim: Epilepsy is one of the most common serious neurological disorders, affecting about 1 % of the world population. The searching for anticonvulsants and drugs for adjunctive therapy is required. PM rat strain had a high frequency of audiogenic seizure and can used as an animal model of some features of the human epilepsy. It was shown that PM rats had a lower taurine level in hippocampus [1]. Taurine (2-aminoethanesulfonic acid) is a natural amino acid with wide occurrence which play inhibitory role in the neural system. There are some works which show the decrease in predisposition to epilepsy after the taurine administration. The changed brain taurine level was demonstrated at the animal model of epilepsy [2]. The aim of this work was to investigate the potential effect of the taurine to predisposition to epilepsy in PM rats.
Methods and Algorithms: The GC rat strain was used in the experiments. PM rat strain was breeding at the Institute of Cytology and Genetics (Novosibirsk, Russia) from a Wistar outbred population. Breeding rats for increased stereotypic hyperkinesis in the form of lateral rocking of the head and body (pendulum movements, PM) resulted in a heightened frequency of audiogenic epilepsy (about 80 %), whereas in the Wistar rats this frequency was about 25 %. Experiment was performed on adult male rats. Taurine (i.p. 800 mg/kg in saline) were given three times at 24 h intervals. After 1h after last injections rats were tested for audiogenic seizures by exposing them to a sound of 109 dB intensity (90 sec duration) at special Plexiglas cube box. All procedures were carried out in accordance with the international guidelines for animal care and use (Ethical principles and guidelines for experiments on animals, Experientia 1995 Jan 15; 51(1):1-3). The results were statistically analyzed using Statistica 6.0.
Results: The acute taurine administration induced significant changes in seizure score severity in PM rats (1.4 ± 0.4 vs. 2.58 ± 0.15 in control group, $p < 0.001$). No significant effect of the taurine administration on the post-ictal catalepsy, on the wild running time and on the latency time from audiogenic stimulus, onset to the initiation of wild running or clonus was observed.
Conclusion: The light anticonvulsive actions of taurine preliminary was shown in PM rats. The further research is required for determine the suitability of the taurine as a drug for adjunctive therapy.
Acknowledgements: Supported by the RFBR No. 17-04-01631-a.

References

1. Akulov A.E. et al. (2014). Selection for catatonic reaction in rats: a study of interstrain differences by magnetic resonance imaging. Zh Vyssh Nerv Deiat im I.P. Pavlova. 64(4):439-47.
2. Oja S.S., Saransaari P. (2013). Epilepsy Res. 104(3):187-94.

DOI 10.18699/BGRSSB-2018-195

The roles of Asp and Patronin in mitotic spindle formation in *Drosophila*

A. Razuvaeva[1, 2]*, G. Pavlova[1], J. Popova[1, 3], L. Yarinich[1, 2], M. Lebedev[1, 2], E. Andreyeva[1], A. Anders[1,2], M. Gatti[4], A. Pindyurin[1, 2, 3]
[1] *Institute of Molecular and Cellular Biology SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[4] *Sapienza University of Rome, Rome, Italy*
** e-mail: alena.razuvaeva@mcb.nsc.ru*

Key words: microtubule, mitotic spindle, kinetochore, colcemid, *Drosophila*

Motivation and Aim: Proper formation of a functional mitotic spindle in centrosome-containing cells requires the concerted action of several microtubule- (MT-) biding proteins. They include proteins that bind the MTs lateral walls (e. g., Dgt6, a subunit of the augmin complex), MT plus end-associated proteins and MT minus end-binding proteins (e. g., Asp and Patronin). While the roles of the plus end-binding proteins in spindle assembly are rather well known, the precise functions of the minus end-binding factors are still poorly defined. Here, we analyzed the relationships between Asp, Dgt6 and Patronin in *Drosophila* spindle assembly and our results suggest a model on how these proteins cooperate to ensure proper spindle formation and functioning.
Methods and Algorithms: To understand the roles of Asp, Patronin and Dgt6, we performed RNAi in *Drosophila* S2 cells against the genes that encode these proteins and examined the ensuing mitotic phenotypes. To define the functional relationships among these genes, we also performed double RNAi against gene pairs. In addition, we generated S2 cell lines that stably express Asp-eGFP or Patronin-eGFP fusion proteins. For Dgt6 localization we used a previously generated antibody that specifically reacts with Dgt6.
Results: Immunostaining experiments revealed that Patronin-eGFP preferentially associates with the kinetochore fibers during mitosis showing a localization pattern rather similar to that of Dgt6. In contrast, Asp-eGFP accumulates at the minus ends of the spindle pole MTs, consistent with previous results. Moreover, results of RNAi experiments together with localization of eGFP fusion proteins suggest that Asp and Patronin functions during mitosis are at least in part independent. Double RNAi-treatment for *asp* and *Patronin* genes resulted in a much stronger phenotype than that observed after RNAi against each of the genes. In double *asp* and *Patronin* RNAi cells we observed failure of sister chromatid separation and metaphase arrest.
Conclusion: Our findings suggest that Asp and Patronin are involved in different pathways required for spindle formation. They also raise the possibility that Patronin specifically binds and stabilize the minus ends of MTs that grow from the augmin-coated lateral walls of preexisting MTs. However these conclusions are the moment only working hypotheses, as they require additional experimental support.
Acknowledgements: Supported by the RFBR (18-34-00688).

DOI 10.18699/BGRSSB-2018-196

Genetic control of the stress-sensitivity in hypertensive ISIAH rats

O. Redina[1]*, S. Smolenskaya[1], L. Fedoseeva[1], A. Markel[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: oredina@ngs.ru*

Key words: hypertension, stress, adrenal gland, plasma corticosterone concentration, differentially expressed genes, QTL analysis, ISIAH rat strain

Motivation and Aim: The study of the genetic control of blood pressure and body composition in ISIAH rats being a model of the stress-sensitive hypertension revealed the quantitative trait locus (QTL) on chromosome X associated with the adrenal gland weight. This locus was also associated with the elevation of blood pressure level and the elevation of the plasma corticosterone concentration under the influence of mild emotional stress [1]. These data suggested that adrenal gland weight may serve as an intermediate phenotype for two other traits related to the enhanced stress-sensitivity of hypertensive ISIAH rats. The current study was directed to the identification of the genes differentially expressed in adrenal glands from hypertensive ISIAH and normotensive WAG rats and localized in the described QTL on Chr. X. The goal of the study was to identify the candidate genes for the elevation of blood pressure level and the elevation of the plasma corticosterone concentration under the influence of mild emotional stress.
Methods: The transcriptional profiling (RNA-Seq) of the adrenal glands from ISIAH and WAG male rats was performed to detect the differentially expressed genes (DEGs).
Results: The results of RNA-Seq analysis revealed 9 DEGs localized in QTL on Chr. X, which was earlier associated with both the adrenal glands weight and the increase in blood pressure level and plasma corticosterone concentration under the emotional stress. The knowledge-driven filtering of the list of these genes suggested *Sms* gene encoding spermine synthase as a positional candidate gene, which may be related to the enhanced stress-sensitivity of hypertensive ISIAH rats.
Conclusion: *Sms* gene encoding spermine synthase may be considered as a positional candidate gene associated with the enhanced stress-sensitivity of hypertensive ISIAH rats. However, the functions of several other discussed DEGs are poorly studied and their further investigation may probably reveal additional candidate genes in the locus.
Acknowledgements: Supported by the Budget project No. 0324-2018-0016.

References

1. Redina O.E. et al. (2013) The genetic control of blood pressure and body composition in rats with stress-sensitive hypertension. Clin Exp Hypertens. 35(7):484-95.

DOI 10.18699/BGRSSB-2018-197

Impact of early life stress on cognition, behavior and hippocampal neuronal plasticity in female mice

V. Reshetnikov[1]*, A. Kovner[1], J. Ryabushkina[1, 2], A. Lepeshko[1, 2], N. Bondar[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: vasiliyreshetnikov@bionet.nsc.ru*

Key words: early life stress, hippocampus, cognitions, neurogenesis, neuronal plasticity, maternal care

Motivation and Aim: Hippocampus is a crucial part of limbic system that involved both in the cognitive processing such as memory and in the regulation of responses to stress. In the rodent, the first postnatal weeks are crucial time for the hippocampal development. Adverse experiences in early life can disrupt neural and behavioral development and impairment of the HPA-axis responsive to subsequent stressors. However, how early-life stress lead to delayed behavioral impairments in adult remains relatively uncertain.
Methods and Algorithms: In our study, two types of early life stress were used: prolonged separation of pups from their mothers (for 3h/day, maternal separation-MS) and brief separation (for 15min/day, handling-HD). In first part of our study, we analyzed the effects of early-life stress on cognition (by using Morris water maze and Novel object recognition tests). As markers of neuronal and synaptic activities number of mature neurons (NeuN+ cells) and level of expression of immediate early genes (qPCR) in the hippocampus as well as number of maturing neurons (DCX+ cells) and number of proliferating neurons (Ki676+ cells) in the dentate gyrus were measured. We examined only female mice, since they are much less investigated, but often more sensitive to stress than males. In second part of our study, we investigated the level of maternal care of females with history of early life stress as a key female behavior, which is dependent on both hippocampus and HPA-axis.
Results: We found that adult female mice in the MS group demonstrated reduced locomotor activity, spatial long-term and recognition memory impairments and reduced level of maternal care, while the HD group showed mild changes. Additionally, MS in early-life resulted in reduced number of mature neurons in the CA3 area of the hippocampus that is crucial for learning and memory.
Conclusion: Thus, prolonged maternal separation but not brief leads to memory and behavioral impairments and reduced number of neurons in the CA3.
Acknowledgements: Supported by the Russian Science Foundation, grant (16-15-10131).

DOI 10.18699/BGRSSB-2018-198

Inheritance of the acoustic signal characters in interspecific hybrids of bank (*Clethrionomys glareolus*) and tianshan (*C. centralis*) voles

M. Rutovskaya
A.N. Severtsov Institute of Ecology and Evolution RAS, Moscow, Russia
e-mail: desmana@yandex.ru

Key words: sound communication, rodents, inheritance

Motivation and Aim: The continuity of behavioral responses in the traits of generation can be interpreted ambiguously, because animal behaviour can be transmitted from generation to generation genetically or can be trained. Genetic inheritance of sound characteristics in mammals has been shown by a number of authors, for example, in common [1] or bank voles [2], and it can also correlate with the inheritance of morphological features of the species [3]. The study of inheritance of the sound signals characteristics in the absence of directional selection is of particular interest, and that was the purpose of the present work.
Methods and Algorithms: Comparisons of distress signals of hybrids, born of TienShan voles females (*C. centralis*) from Kyrgyzstan and bank vole males from the Tver region (*C. glareolus suecicus*) and the sounds of the parent species were made. Distress signals of voles were recorded in laboratory, using a Tascam NoDA-P1 professional digital tape recorder (United States) and SENNHEISER K6 microphone (Germany). Sounds were analyzed using the AviSoft SASLab pro professional program (version 4.2). To perform the spectral analysis of signals in building spectrograms, a fast Fourier transform length of 512 points and an overlap of 100 % for the frequency axis and 88 % for the temporal axis were used. Acoustic signals of the hybrids and parent forms were compared using the variance and discriminant analysis.
Results: The distress signals of the TienShan voles were shorter and the dominant frequency had lower values than in the signals of the bank voles. The dominant frequency of hybrids signals was closer to that of the bank voles. However, the duration of the hybrids sounds, on the contrary, is closer to that of the TienShan voles. The expression of the noise component in the sounds of the hybrids occupies an intermediate position. The discriminant analysis of distress signals in the the TienShan and bank voles showed 93 % of correct attributions.
Conclusion: The hybrids signals were different from the sounds of parental species by its characters. The father's genome (bank voles) has had great influence on characters of the sound signals in hybrids.
Acknowledgements: Supported by the RFBR (16-04-00149)

References

1. Frank F. (1967) Twittering – a voice mutation in *Microtus*. Heredity. 22(4):607-610.
2. Rutovskaya M.V., Osipova O.V., Soktin A.A. (2017) Acoustic signal characters in interspecific hybrids of bank and red-backed voles. Biology Bulletin. 44(3):296-306.
3. Nikol'skii A.A. (1984) Acoustic signals of mammals in the evolutionary process. Moscow: Nauka.

DOI 10.18699/BGRSSB-2018-199

Expression of *Th, Comt* and *Maoa* genes and brain catecholamine levels in rats with genetic catatonia

M. Ryazanova[1, 3]*, T. Latysheva[2], V. Plekanchuk[1, 3], O. Prokudina [1], M. Gilinsky[2], T. Alekhina[1]

[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*

[2] *Research Institute of Physiology and Fundamental Medicine, Novosibirsk, Russia*

[3] *Novosibirsk State University, Novosibirsk, Russia*

** e-mail:ocean-2006@yandex.ru*

Key words: genetic catatonia, genes expression, catecholamines

Motivation and Aim: The work was performed on the GC rats strain with genetic catatonia, which can be considered as a model of the catatonic syndrome observed in schizophrenia, affective disorders and other diseases [1]. It is known that brain dopaminergic system is involved in development of catatonia and schizophrenia and its involvement is complex. Many drugs for treatment of psychosis have an effect on dopamine system function. The noradrenergic system can modulate dopamine system, as well as directly affect on motor activity and emotional sphere. The aim of this work was to study mRNA expression of genes catecholamines synthesis – tyrosine hydroxylase (Th) and degradation – catechol-O-methyltransferase (Comt) and monoamine oxidase A (Maoa), and also to estimate the catecholamines and 3,4- dihydroxyphenylacetic acid (DOPAC) content in brain structures in inbred GC rats.

Methods and Algorithms: Using RT-qPCR Th, Comt, Maoa mRNA level in brain structures (hypothalamus, striatum, amigdala, midbrain, medulla oblongata) of GC an WAG rats was estimated. High performance liquid chromatography with electrochemical detection was used for investigation the level of catecholamines (dopamine and noradrenaline) and DOPAC in same brain structures.

Results: The study revealed a low level of *Maoa* mRNA in striatum of rats with genetic catatonia. The mRNA level of *Com*t, *Th* in brain structures GC rats did not differ from the control. The high level of norepinephrine in hypothalamus, midbrain and striatum of GC rats was shown.

Conclusion: The high level of norepinephrine in striatum of GC rats may be due to the low mRNA expression of *Maoa* in this structure, however, the increase norepinephrine level in the hypothalamus and in midbrain are not associated with the altered expression of the mRNA gene. The revealed changes in the level of norepinephrine in the brain can contribute to the increased excitability and nervosity of GC rats, which they shown before [2] along with the akinetic reactions of freezing.

Acknowledgements: Supported by the RFBR (grant No. 16-04-01325).

References

1. Kolpakov V.G. et al. (2004) Catatonia or depression? The GC rat strain is a genetic animal model of psychopathology. Genetics. 40(6):827-834.
2. Ryazanova M.A. (2012) Increase in the share of “nervous” animals during the selection for catatonia: participation in catatonic reactions of central adrenoreceptors. Genetics. 48(11):1141-1147.

DOI 10.18699/BGRSSB-2018-200

Drosophila mutational models for Huntington's disease, Parkinson's disease with dementia and levy bodies and genomic disorder Williams-Beuren syndrome

E.V. Savvateeva-Popova
Pavlov Institute of Physiology of the RAS, St. Petersburg, Russia
e-mail: esavvateeva@mail.ru

Key words: kynurenine pathway, cascade of actin remodeling, LIM-kinase 1-dependent cognitive phenotypes, hypersociability, domestication

Pavlov Institute Multi-access Center "Biocollections" exploits the concepts of N. Vavilov's (1920) Law of Homologous Rows of Hereditary Variability and T. Dobzhansky assuming that the evolutionary conservation both of gene function and of elemental behavioral mechanisms guarantees that much of what we learn in one organism will inform our understanding of behavior in all animals, including humans. This insight permits behavior-geneticists to choose organisms based on experimental tractability for a given scientific question and to develop experimental model systems to probe the causes, consequences and mechanisms of pathology leading to human disease. Traditionally, we are doing this together with Novosibirsk Institute of Cytology and Genetics, the first example is neurochemical study of the kynurenine pathway in Drosophila and the honey bee [1]. This has resulted in developing models both for Huntigton's Disease [2], aging [3] and in silico bioinformatics analysis of antioxidant properties of kynurenines, as a cause, having neurodegeneration and cataract as consequence [4]. The second example is molecular biologic study which has allowed to develop Drosophila model for genesis of LIM-kinase 1-dependent cognitive phenotypes both in Parkinson's Dementia with Levy Bodies (DLB) and genomic disorder Williams-Beuren Syndrome (WBS) [5, 6]. At the same time, structural variants in genes associated with human WBS appear to underlie hypersociability in domestic dogs [7] explaining both the Pavlovian selection on different types of higher nervous activity and D. Belyaev's domestication [8].

Acknowledgements: Supported by Presidium of UB RAS (0134-2018-0003).

Reference

1. Kamyshev N.G., Savvateeva E.V., Kudriavtseva N.N., Lobacheva I.I. (1981) Serotonin content in mutants of the metabolic pathway tryptophan – xanthommatin in *Drosophila melanogaster* r. Dokl. Akad. Nauk USSR. 256(5):1237-1238.
2. Savvateeva E., Popov A., Kamyshev N., Bragina J., Heisenberg M., Senitz D., Kornhuber J., Riederer P. (2000) Age-dependent memory loss, synaptic pathology and altered brain plasticity in the Drosophila mutant cardinal accumulating 3-hydroxykynurenine. J Neural Transm. 107:581-601.
3. Savvateeva-Popova E.V., Popov A.V., Heinemannt, Riederer P. (2003) Drosophila Mutants of the kynurenine pathway as a model for ageing studies. Advances Experimental Medicine Biology. 527:713-22.
4. Zhuravlev A.V., Zakharov G.A., Shchegolev B.F., Savvateeva-Popova E.V. (2016) Antioxidant properties of kynurenines: density functional theory calculations. PLOS Computational Biology. DOI 10.1371/journal.pcbi.1005213.
5. Savvateeva-Popova E.V., Peresleni A.I., Scharagina L.M., Medvedeva A.V., Korochkina S.E., Grigorieva I.V., Dyuzhikova N.A., Popov A.V., Baricheva E.M., Karagodin D., Heisenberg M. (2004) Architecture of the X chromosome, expression of LIM Kinase 1, and recombination in the agnostic mutants of Drosophila: a model for human Williams syndrome. Russian Journal Genetics. 40(6):605-624.
6. Savvateeva-Popova E.V., Zhuravlev A.V., Brázda V., Zakharov G.A., Kaminskaya A.N., Medvedeva A.V., Nikitina E.A., Tokmatcheva E.V., Dolgaya J.F., Kulikova D.A., Zatsepina O.G., Funikov S.Y., Ryazansky S.S., Evgen'ev M.B. (2017) Drosophila model for the analysis of genesis of LIM-kinase 1-dependent Williams-Beuren syndrome cognitive phenotypes: INDELs, transposable elements of the Tc1/mariner superfamily and microRNAs. Frontiers in Genetics. Section Behavioral Psychiatric Genetics. 8:123. DOI 10.3389/fgene.2017.00123.
7. Von Holdt B.M., Shuldiner E., Koch I.J., Kartzinel R.Y., Hogan A., Brubaker L., Wanser S., Stahler D., Wynne C.D., Ostrander E.A., Sinsheimer J.S., Udell M.A. (2017) Structural variants in genes associated with human Williams-Beuren syndrome underlie stereotypical hypersociability in domestic dogs. Sci Adv. 19;3(7):e1700398. DOI 10.1126/sciadv.1700398.eCollection 2017.
8. Trut L., Oskina I., Kharlamova A. (2009) Animal evolution during domestication: The domesticated fox as a model. Bioessays. 31:349-360.

DOI 10.18699/BGRSSB-2018-201

Astrocyte-to-neurone lactate communication in the brain

A.G. Teschemacher*, B.V. Cardoso, V. Mosienko, B.H. Liu, S. Kasparov
University of Bristol, United Kingdom
** e-mail: anja.teschemacher@bristol.ac.uk*

Key words: brain metabolism, central metabolic signalling, astrocytes, extracellular L-lactate, release of noradrenaline

Astrocytes are thought to be the main source of extracellular L-lactate (LL) in the brain under physiological conditions. LL levels respond dynamically to neuronal network activity and metabolism. LL may serve as an additional energy substrate during periods of high activity. There is evidence that resting concentration of LL in astrocytes is higher than in neurones, allowing gradient-driven transfer of LL from astrocytes to neurones. This so called "lactate shuttle hypothesis" postulates that, under conditions of high energy demand, LL can be used by neurones as a preferred energy substrate. However, this idea is not universally accepted. Therefore, why astrocytes produce more LL than they can consume and whether the LL which astrocytes release is indeed used by neurones as fuel is not clear at present. In addition, increasing evidence indicates that LL has a signalling role in the brain. We demonstrated that LL released from astrocytes may release noradrenaline from noradrenergic neurones located in the brainstem and have been searching for the potential new receptor for LL in that area (Tang et al., Nature Communications, 2014). LL-mediated release of noradrenaline could be important for regulation of sleep/wake and attention states, learning and memory, and cardiorespiratory control.

However, almost all information we have about the physiological significance of astrocytic LL comes from acute experiments, and we currently lack the means for selective manipulation of LL release from astrocytes over longer time periods for further investigation of its actions in the brain of living animals. We therefore set out to develop an array of astrocyte-selective viral vectors that will allow us to specifically limit LL release by expressing enzymes which break down or decrease LL synthesis. Such enzymes can be 'borrowed' from bacteria. We show, using imaging, fluorimetry, and amperometry, that these constructs significantly decrease intracellular LL pool sizes. We anticipate that astrocyte-selective expression of these novel constructs will help to clarify the role of astrocytic LL in brain metabolism and central metabolic signalling.

DOI 10.18699/BGRSSB-2018-202

High-density genotyping of 15 native Russian sheep breeds reveals genomic regions under selection related to domestication, acclimation and economically important traits

A.A. Yurchenko[1], T.E. Deniskova[2], A.V. Dotsev[2], N.S. Yudin[1, 3], H. Reyer[4], K. Wimmers[4], G. Brem[2, 5], N.A. Zinovieva[2], D.M. Larkin[1, 6]*

[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *L.K. Ernst Federal Science Center for Animal Husbandry, Podolsk, Moscow region, Russia*
[3] *Novosibirsk State University, Novosibirsk, Russia*
[4] *Institute of Genome Biology, Leibniz Institute for Farm Animal Biology, Dummerstorf, Germany*
[5] *Institute of Animal Breeding and Genetics, University of Veterinary Medicine, Vienna, Austria*
[6] *Royal Veterinary College, University of London, London, UK*
** e-mail: dmlarkin@gmail.com*

Key words: sheep native breeds, selection, adaptation, genotyping, Russia

Motivation and Aim: There are over 1000 sheep breeds existing worldwide, many of which are adapted to local environments characterised by a variety of conditions: e. g. climates, parasite profiles, etc. Breeds were selected by humans for various traits, including wool quality/quantity, milk and meat production. Therefore, 9,000–11,000 years of domestication and breed design led to formation of local breeds that contain signatures of natural adaptation and artificial selection in their genomes. To identify genomic intervals that respond to selection and relate to adaptions to a variety of conditions found in Russian Federation we analysed ~300 genomes of 15 Russian native sheep breeds.

Methods and Algorithms: A total of 312 representatives from 15 Russian native sheep breeds (ranging 20–24 samples per breed) were genotyped on the Illumina HD Ovine SNP Array, containing ~600 M SNPs. The data were filtered in plink: --maf 0.05, --chr 1–26 and analysed with Admixture and PCA algorithms to identify groups of related breeds. We then performed detection of signatures of putative selection in the combined set of breeds and in two sets of related breeds separately using the hapFLK software.

Results: A total of 79 regions under putative selection ranging from 39 Kbp to 5 Mbp in size were identified in the combined set of 15 Russian breeds. The number of regions in two sets of related breeds was 104 and 64. A total of 556 genes were found within these intervals. The strongest signals of selection (p-value < e^{-23}) were detected near the genes: *KIT, MC1R, MITF* (coat color), *PXFP2* (polled/horns), LCORL/ NCAPG (growth). We further detected strong signatures of selection in the regions containing genes known to be related to wool quality (*PRLR*, *IRF2BP2*), virial resistance (*TMEM154*), milk production (*ABCG2*, *SOCS2*), carcass (*BMP2*), and reproduction (*VEGFA, NR3C1*). In addition, candidate genes related to adaptation to cold stress/climate were identified in the selected regions, including: *MMRN1*, *ADD1*, *ATF4*, *MCM8*, and *TSHR*.

Conclusion: Multiple regions under putative selection were identified in the set of 15 Russian native sheep breeds forming a basis for future genomics-based selection and targeted breeding of Russian sheep.

Acknowledgements: This study was supported by the grants from the Russian Science Foundation (projects No. 16–14-00090 and 14-36-00039).

DOI 10.18699/BGRSSB-2018-203

Genotyping of nine native Russian cattle breeds combined with the 1000 bull genome project data reveals signatures of selection and adaptation in Russian cattle genomes

A.A. Yurchenko[1], N.S. Yudin[1, 2], D.M. Larkin[1, 3]*
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Royal Veterinary College, University of London, London, UK*
* *e-mail: dmlarkin@gmail.com*

Key words: cattle native breeds, selection, adaptation, genotyping, sequencing, Russia

Motivation and Aim: In Russian Federation, there are ~20 native cattle breeds currently being registered. Our previous data suggest that these breeds form four distinct phylogenetic clusters when clustered together with cattle breeds from around the world. In the present study we focused on detecting signatures of adaptation/selection that shaped the Russian cattle genomes during domestication, breed formation, selective breeding, and adaptation to local environments. We combined our genotyping data from nine genetically distinct Russian native cattle breeds with the sequence-based genotypes of four European and two Asian breeds and analysed the combined datasets to spot adaptive genomic changes common among the breeds and those that are characteristics of the Russian breeds only.
Methods and Algorithms: A total of 172 representatives from nine Russian native cattle breeds (ranging 18–30 samples per breed) previously genotyped on the GGP HD150K Bovine SNP array containing ~139K SNPs were combined with sequence/HD array genotyping data from six foreign breeds (130 animals total). The data were filtered in Plink and signatures of putative selection in three sets of related breeds were detected separately using the hapFLK software. In parallel, signatures of genomic selection were identified for each breed using the de-correlated composite of multiple signals framework combining the F_{ST}, H1, H12, Tajima's D and nucleotide diversity statistics.
Results: A total of 999 regions under putative selection ranging from 1 bp to 15 Mbp in size were identified in the combined set of all breeds. A total of 1,578 genes were found within these intervals. The strongest putative signals of selection were detected near the genes: *LCORL/NCAPG*, *HMGA2*, *IMPAD1* (growth), *KIT* (coat colour), *PLAG1* (reproduction). We further detected signatures of selection in Russian breeds related to domestication (*KITLG*, *EDN3*, *COPA*), feed intake (*XKR4*, *TMEM68*), milk production (*DGAT1*, *GHR*, *ABCG2*, *GLI2*, *LAP3*, *TRPV5*, *FKBP2*), and reproduction (*CSF2*, *BCL2*, *ANXA10*, *NPBWR1*). In addition, strong candidate genes for adaptation to cold climate and local environment were found under selection in the Yakut cattle: *RETREG1*, *RPL7*, *TNKS*, Kholmogory: *AQP5*, and multiple Russian breeds: *ARRDC3*, *RAD50*, *SYK*.
Conclusion: Multiple regions under putative selection were identified in the set of nine Russian native cattle breeds forming a basis for future genomics-based selection and targeted breeding of Russian cattle.
Acknowledgements: This study was supported by the grant from the Russian Science Foundation (project No. 16–14-00090).

DOI 10.18699/BGRSSB-2018-204

The impact of rapid decrease of *Aporia crataegi* (Lepidoptera: Pieridae) population size on *Wolbachia* infection rate

G. Yurlova[1]*, R. Bykov[1], M. Yudina[1, 2], A. Tikhomirova[2], O. Kosterin[1, 2], Yu. Ilinsky[1, 2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: yurlova@bionet.nsc.ru*

Key words: *Wolbachia*, *Aporia crataegi*, population, mtDNA

Motivation and Aim: *Wolbachia* endosymbionts are widely distributed among insects. The high level of infection was observed in populations of black-veined white (*Aporia crataegi*) of Novosibirsk province in 2006. Rapid population decrease had occurred in 2007 due to late mid-May frosts. Here we examined *Wolbachia* infection rate in Novosibirsk *A. crataegi* population that restored its abundance in 2015–2016 seasons. Also we try to compare data of infection rates, *Wolbachia* and mtDNA diversity derived from different populations of black-veined white.
Methods and Algorithms: The collection includes 246 *A. crataegi* specimens from Novosibirsk and Kemerovo provinces, Altai Republic, and Yakutia. *Wolbachia* infection status was determined by PCR with primers to *coxA* and *16SrRNA Wolbachia* loci. The barcoding region of *COI* gene of both infected and uninfected specimens was sequenced. The phylogenetic tree *COI* gene was reconstructed in Mega7 using all available sequences from BOLD database and our data.
Results: *Wolbachia* symbionts were not found in *A. crataegi* from Altai Republic, Novosibirsk and Kemerovo provinces, while several infected specimens were found in Yakutia. Their *Wolbachia* symbionts had *coxA-6* allele. New alleles of *A. crataegi* mitochondrial gene ... were found.
Conclusion: *Wolbachia* infection rate in black-veined white of Novosibirsk province dramatically decreased after a severe drop of host population in 2007. No *Wolbachia* infection was detected at neighboring provinces as well. Also we found association of *Wolbachia* infection with a certain mtDNA allele of *A. crataegi.*
Acknowledgements: Supported by the RFBR (16-04-00980, 18-316-00099).

DOI 10.18699/BGRSSB-2018-205

Larks, owls, swifts and woodcocks among the fruit flies: heritable individual differences in the sleep-wake pattern induced by long, and sometimes, hot Siberian summer days

L.P. Zakharenko[1]*, D.V. Petrovsky[1], A.A. Putilov[2]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Institute of Molecular Biology and Biophysics, Novosibirsk, Russia*
** e-mail: zakharlp@bionet.nsc.ru*

Key words: morning-evening preference; chronotype; circadian rhythm; photoperiod; temperature; locomotor activity; morning and evening oscillators; migration out of Africa

Motivation and Aim: *Drosophila melanogaster* and our own species have many things in common including history of rapid out-of-Africa dispersal. In Eurasia, we both faced the problem of adjustment of our circadian rhythms to seasonal variation in day length, and each of us usually responds to night sleep disturbances by prolongation of siesta. To further explore similarity between two species, a division of flies into chronotypes was examined.
Methods: By testing the circadian rhythms of locomotor activity in constant darkness, 4 strains originating from three wild populations of Africa, Europe and the USA were selected as the representatives of 4 distinct chronotypes [1]: "larks" (early morning and evening peaks), "owls" (late morning and evening peaks), "swifts" (early morning and late evening peaks) and "woodcocks" (late morning and early evening peaks). The locomotor rhythms and sleep-wake pattern of selected chronotypes were tested under either long day condition (L:D 20:4) at 20 °C or combination of high temperature (29 °C) with L:D 20:4.
Results: Despite identity of such an experimental condition for 4 chronotypes, their circadian rhythms showed 4 distinct patterns of maladaptive response to it. For instance, as we expected for L:D 20:4, the flies of any chronotype failed to increase the phase angle between morning and evening peaks of activity beyond 16 hours (i. e., to remain in synch with transitions from darkness to light and back from light to darkness that were divided by 20-h interval). Instead, "woodcocks" became fully arrhythmic. "Owls" exhibited as many as three prominent peaks of activity. In "larks" the reaction to the light off was blunted but their response to the light switch was very strong, while the opposite was true for "swifts".
Conclusion: An experimental study of heritable chronotypes in the fruit fly can deepen our understanding the genetic underpinnings of individual variation in vulnerability to maladaptive sleep-wake behavior, circadian misalignment and sleep disorders.
Acknowledgements: LPZ and DVP were partially funded by the Federal Research Program (the project number 0324-2018-0016).

Reference

1. Putilov A.A. (2017) Owls, larks, swifts, woodcocks and they are not alone: A historical review of methodology for multidimensional self-assessment of individual differences in sleep-wake pattern. Chronobiol Int. 34:426-437.

SYSTEMS BIOLOGY OF AGING

DOI 10.18699/BGRSSB-2018-206

Functional characterization of the conservative protein CG17337 in *Drosophila melanogaster*

E. Andreyeva[1]*, A. Ogienko[1], E. Kozhevnikova[1, 2], A. Ivankin[1], L. Yarinich[1], G. Pavlova[1], J. Popova[1, 2], A. Alekseeva[1], T. Dubatolova[1], S. Kopyl[1], A. Pindyurin[1]
[1] *Institute of Molecular and Cellular Biology SB RAS, Novosibirsk, Russia*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
** e-mail: andreeva@mcb.nsc.ru*

Key words: metallopeptidase M20, CG17337, CNDP2, cell cycle, cell death, cancer, lifespan, senescence

Motivation and Aim: Eukaryotic peptidases play an important role in protein and peptide metabolism. Currently, close attention is being paid to the human *CNDP2* gene encoding a protein of metallopeptidase M20 family. The CNDP2 is a tumor suppressor protein; changes of its expression level stimulate cell proliferation and are used as a biomarker of cancer. Our aim was to clarify the function of drosophila *CG17337* gene, which is the ortholog of *CNDP2,* at the level of cells and whole organism.
Methods and Algorithms: First, we have raised polyclonal antibodies specific to CG17337, which allowed us to visualize this protein in drosophila tissues and cultured cells by confocal microscopy. Second, using CRISPR/Cas9-mediated homologous recombination, we generated null mutant of the *CG17337* gene. The *CG17337* null-mutation was checked by genotyping PCR, DNA sequencing and Western Blot analysis. Also, we generated transgenic flies carrying genomic copy of the *CG17337* gene for rescue experiments. We estimated influence of the *CG17337* null-mutation on cell cycle in different tissues as well as on the lifespan of drosophila.
Results: We provide first insights on the role of the CG17337 protein in drosophila. First, we found that CG17337 is not only extracellular protein as was reported earlier [1], but also plays a role both in the cytoplasm and nucleus (chromatin). Particularly, we found CG17337 in a co-immunoprecipitation assay as a putative interactor partner of the SUUR (Suppressor of Underreplication) protein in embryonic nuclear extracts. The nuclear localization and chromatin association of CG17337 was observed in cell cycle-dependent manner. We determined by Fly-FUCCI approach that the CG17337 protein is enriched in G2 cells. Second, we have shown that depletion/lack of CG17337 in cultured S2 cells or in drosophila tissues leads to changes in mitotic cell cycle, whereas endocycle is not affected. Finally, the *CG17337* is a dominant and haploinsufficient gene. The obtained null-mutation appeared to be homozygous viable and is characterized by an elongated female lifespan. Unexpectedly we found that females null mutant for the *CG17337* gene display a prolonged period of fertility probably due to suppression of cell death in middle oogenesis during senescence.
Conclusion: Altogether, our results indicate that the drosophila CG17337 protein is necessary for proper mitotic cell cycle progression and cell death control.
Acknowledgements: The work was supported by RFBR grant No. 16-04-01598.

References

1. Karlsson et al. (2004). Proteomic analysis of the *Drosophila* larval hemolymph clot. J. Biol. Chem. 279(50):52033-52041.

DOI 10.18699/BGRSSB-2018-207

The role of compensatory reactions in life-span determining program realization: investigations in model strains of house fly

G.V. Benkovskaya*, T.T. Akhmetkireeva, Yu.M. Nikonorov
Institute of Biochemistry and Genetics, Ufa Federal research center of RAS, Ufa, Russia
** e-mail: bengal2@ru*

Key words: house fly, life span, stress, compensatory reaction

Motivation and Aim: The aim of investigation is evaluate the developmental regularities of compensatory reactions to stressors and revealing the role of compensatory reactions in life span determination.
Methods and Algorithms: Objects of investigations: house fly strains selected by life span from strain Cooper (heterogeneous strains *Sh gen* – short-living and *L gen* – long-living flies; homogeneous strains *Sh 28* short-living and *L 2* long-living flies. Experimental stresses – stress of selection and short-term heat stress single or repeated at the each developmental stage (larva, pupa, adult). Registered indices were life span of adults, duration of larval and pupal development, weight of larvae and pupa, frequencies of morphological deviations. Dynamics evaluated of tyrosinase, DOPA-oxidase, acetylcholinesterase activity under the stress conditions and compensatory process. Geometric morphometrics applied for shape and size of adult's wings to detect the delayed consequences of stress. Change of transcriptional activity of genes *hsp70* and transposone hermes transposase *hem* estimated by real time qPCR.
Results: The indirect feedback mechanism revealed limiting life span changes in homogenous housefly strains, clearly displaying under the selection for life span diminishing and early reproduction. Effects of compensatory reactions appeared as the spike of fitness indices variability evidenced the genomic stress, confirmed by transposone *hermes* DNA copy number enhancement in genome of selected strains vs. origin strain. The monitoring of development in the homogenous and heterogeneous housefly strains and analysis of fitness indices changes followed the heat stress at the larval stage showed formation of adaptive ontogenic reactions of two types, distinguished for short-living and long-living strains. Common for all strains inadaptive reactions disclosed followed by compensative changes of development rate and resulted in reproductive disturbances and adult's life span decrease. The longitudinal observations of strains with altered life span undergoing to the repetitive heat stress at each developmental stage allowed revealing compensatory processes at the population level. There is trend of return to the reaction norm of all fitness traits in the origin strain. The biochemical parameters of stress-reaction and their variability dynamics allowed concluding that the repetitive events of stress during ontogenesis caused the adaptive reaction development directed to maintain the stable level of phenoloxidases system activity and catecholamines concentration in haemolymph as well as to suppress the transposone reproduction. Stress-reaction intensity decreased as the variability of its parameters evidenced the metabolic processes synchronization.
Conclusion: The necessity of immediate reaction rearrangement led to the significant developmental changes. The destabilization forced by environmental stress factors influence as showed via the analysis of wing geometric morphometry in undergoing to heat stress laboratory strains of house fly. All results obtained suggested the significance of compensatory adaptive reactions developed after the eustress as the inadaptive reactions for definition of life strategies and homeostasis maintenance at the levels of genome, organism and population or intrapopulation groups.
Acknowledgements: Investigations completed in the framework AAAA-A16-116020350032-1 and partially funded by RFBR No. 12-04-01450-a and 15-04-04801-a.

DOI 10.18699/BGRSSB-2018-208

Alterations of neurogenesis during development of Alzheimer's disease-like pathology in OXYS rats

A.O. Burnyasheva*, N.G. Kolosova, N.A. Stefanova, E.A. Rudnitskaya
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: burn.alena2505@mail.ru*

Key words: neurogenesis, Alzheimer's disease, OXYS rats

Motivation and Aim: Neurogenesis is one of the major mechanisms of neuronal plasticity and it is crucial for successful learning. Neurogenesis is altered in a couple of pathological conditions including Alzheimer's disease (AD). Alteration of neurogenesis in turn results in further loss of neuronal plasticity triggering vicious circle of neurodegeneration. To investigate the link between changes of neurogenesis and development of AD we used OXYS rats which are considered as a suitable model of the most common sporadic form of AD.
Methods and Algorithms: 3- and 18-months-old male OXYS and Wistar (control) rats were used. Immunohistochemistry was used to identify the number of neuronal cells at different stages of maturation and amyloid-β (Aβ) deposition in the dentate gyrus (DG) of the hippocampus. Animal's learning and memory were evaluated in the Morris water Maze (MWM).
Results: We have shown that 3-months-old Wistar rats were able to find hidden platform in the MWM already at the 2nd day of training while OXYS rats failed to remember the location of the platform thus demonstrating unsuccessful learning. After relocation of the platform to the opposite quadrant on the 6th trial day Wistar rats successfully found it already at the 7th training day while OXYS rats failed thus demonstrating alterations of reversal learning. Moreover, OXYS rats demonstrated alterations of reference memory: they spent much less time in the target quadrant at 11th trial day, when the platform was removed from the pool. Altered learning and spatial memory may reflect changes of neuronal plasticity in the hippocampus. Indeed, we shown decreased density of neuroblasts in the DG of 3-months-old OXYS rats compared to Wistar rats. The density of neuroblasts and immature neurons dramatically decreased with age in the DG of both strains. However, the density of immature neurons was higher in the DG of 18-months-old OXYS rats compared to Wistar rats against background of pronounced Aβ deposition. Increased density of immature neurons may reflect alterations of neuronal maturation in the DG of OXYS rats.
Conclusion: Manifestation of AD-like pathology in OXYS rats occur against background of significant learning and spatial memory deteriorations as well as alterations of hippocampal neurogenesis. Progression of AD-like pathology in OXYS rats occur against background of pronounced Aβ deposition and altered neuronal maturation in the DG of hippocampus.
Acknowledgements: This work was supported by grant from the Russian Foundation for Basic Research (project No. 18-015-00320).

DOI 10.18699/BGRSSB-2018-209

Search of single-nucleotide polymorphisms associated with accelerated senescence in OXYS rats

V. Devyatkin[1, 2]*, N. Muraleva[1], O. Redina[1]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: devyatkin@bionet.nsc.ru*

Key words: aging, age-related diseases, OXYS rats, cataract, SNP

Motivation and Aim: Aging is the single largest risk factor for chronic disease, still little is known about a genetic overlap between complex age-related diseases. For search pathways that modulate the onset and progression of multiple age-related diseases here we used senescence-accelerated OXYS rats derived from Wistar rats in the ICG SB RAS (Novosibirsk) and developing a phenotype similar to human geriatric disorders including cataract, age-related macular degeneration-like retinopathy and neurodegenerative pathology of the brain with features of Alzheimer's disease. We hypothesize that the senile cataract development can serve as a biomarker of systemic changes associated with aging. The purpose of our work is to identify in the genome of OXYS rats mutations in genes associated with cataract, which can potentially contribute to the development of signs of accelerated aging.
Methods and Algorithms: We used the data of RNA-Seq from prefrontal cortex, retina and hippocampus of senescence-accelerated OXYS and Wag (control) rats. Positions of SNPs within the aligned reads relative to the reference genome (Rnor 6.0) were identified using SAMtools (v. 0.1.17) utilities. Each mutation was present in at least 3 OXYS rats in homozygous state and was not present in any of the Wag rats. The effect of an amino acid substitution on protein function was predicted by the Variant Effect Predictor Web service; the consequence type, SIFT score, and prediction were obtained for each variant. The list of genes associated with cataracts was compiled according to NCBI, Cat-Map, KEGG Disease databases.
Results: In the genome of OXYS rats 52539 SNPs in the homozygous state, not presented in the genome of Wag rats, overlapped with 8012 genes and 11684 transcripts were revealed. In 328 cases the substitutions can result in significant structural rearrangements (high impact effect) of the transcripts. Among the non-synonymous substitutions 254 have a deleterious effect on the structure or function of the protein product according to the SIFT algorithm. We revealed SNPs in 255 genes that can be associated with cataract development in OXYS rats and contained 543 described and 614 novel SNPs. 4 of this genes, *Pex2*, *Nbn*, *Rab18* and *Prss56* have SNPs (rs198310567, rs105362013, rs106234270 and rs106604882, respectively) with a deleterious result according to SIFT, although these polymorphisms are described for SHR/OLAIPCV and SD rat strains without signs of cataracts. However, it is known that mutations in these genes are associated with a number of mitochondrial diseases, nervous and cardiovascular disorders, consistent with the complex manifestation of the senile phenotype in OXYS rats against the background of cataract development.
Conclusion: Genes with mutations revealed in OXYS rats are promising for further verification of the contribution of found polymorphisms to the development of complex age-related diseases.
Acknowledgements: Supported by the RFBR (project No. 18-015-00336, No. 18-315-00216).

DOI 10.18699/BGRSSB-2018-210

Influence of insertion/deletion and transcriptional activity of Alu-elements on human longevity

V. Erdman*, D. Karimov, T. Nasibullin, I. Tuktarova, O. Mustafina
Institute of Biochemistry and Genetics, Ufa Federal Innovation Center, RAS, Ufa, Russia
** e-mail: danivera@mail.ru*

Key words: human longevity, Alu-element, insertion/deletion polymorphism, transcriptional activity

Motivation and Aim: It is hypothesized that genome instability can affect the lifespan [1]. Alu-insertions are one of the causes of such instability. This element can contain regulatory sites, such as silencers, and therefore influence transcriptional activity (TA) of genes [2]. Our aim was to search association of longevity with insertion/deletion and TA of Alu-elements in some genes, whose protein products involved potentially in age-associated processes – apoptosis, immune reaction, intra- and extracellular signaling and others.
Methods and Algorithms: Total group was formed of 2000 unrelated individuals aged from 21 to 113 years, ethnic Tatars (residents of Republic of Bashkortostan, Russia). *PLAT*, *COL13A1*, *ACE*, *LAMA2*, *EVI5*, *STK38L*, *PKHD1L1*, *HECW1*, *SOX5*, *CDK11A*, *NOTCH2*, *PTPRO*, *SEMA6A* and *TEAD1* genes were selected for the analysis of Alu-insertion/deletion polymorphism and TA. DNA was isolated from lymphocytes of peripheral venous blood by phenol-chloroform method. Alu-polymorphism was detected with PCR. Total RNA was extracted from peripheral blood leukocytes using TRIzol reagent following manufacturer's instructions. First strand cDNA was synthesized from RNA template using reverse transcription PCR. TA of Alu-elements in selected genes was detected by RT-PCR. Logistic regression analysis was used to assess age dynamics of genotype and allele frequency. $\Delta\Delta C_t$-method was applied to analyze gene TA.
Results: It was found that the chances to achieve longevity were increased in females with *COL13A1**D/*D, *LAMA2**I/*D, *TEAD1**I/*I, *PKHD1L1**I/*I genotypes, in males with *PKHD1L1**I/*D genotype, and were decreased in females with *ACE**D/*D and *LAMA2**D/*D genotypes. Carriers of *CDH4**D/*D genotype had higher chances to become centenarians. TA of *COL13A1* gene progressively decreased during aging, but peaked up again at very old ages. Expression level of *LAMA2* gene was higher among carriers of *LAMA2**D allele compared to individuals with *LAMA2**I allele.
Conclusion: Alu-polymorphism of *COL13A1* (Ya5ac1986), *ACE* (Ya5ACE), *LAMA2* (Ya5-MLS19), *PKHD1L1* (Yb8AC702), *CDH4* (Yb8NBC516) and *TEAD1* (Ya5ac2013) genes is likely associated with chances of longevity. Alu-polymorphism of *LAMA2* gene influences its TA among long-livers. It can be suggested that Alu(I/D) polymorphism of selected genes can be influence on the development of a number of age-dependent pathologies through a change in the gene expression level.
Acknowledgements: Supported by RFBR (No. 17-44-020735).

References

1. Mustafina O.E. (2012) The possible roles of human Alu elements in aging. Papers of the conference on genetics of aging and longevity. Frontiers E-books: 129.
2. Cardelli M., Marchegiani F. (2013) Good, bad, mobile elements: genome's most successful "parasites" as emerging players in cell and organismal aging. Current Pharmaceutical Design. 19(9):1739-1752.

DOI 10.18699/BGRSSB-2018-211

Systems biology approach reveals the mystery of aging origin

A. Khalyavkin[1, 2]*, V. Krut'ko[2, 3]
[1] *Emanuel Institute of Biochemical Physics RAS, Moscow, Russia*
[2] *Institute for Systems Analysis FRC CSC RAS, Moscow, Russia*
[3] *Sechenov First Moscow State Medical University, Moscow, Russia*
** e-mail: antisenesc@mail.ru*

Key words: origin of aging, systems approach, self-maintenance, environmental influence

Motivation and Aim: An overview of our current knowledge of mortality statistics and biology of aging can provide better understanding of some general peculiarities of macro-systems in different environments and might lead us to the development of useful approaches to the origin of aging and its control and especially explain why we age despite having potentially ageless somatic stem cells that can even reverse aging.
Methods and Algorithms: We used the critical analysis of the array of published experimental findings and our particular interpretation of these findings in order to unite at the one conception many separate and various data obtained from the molecular level up to the level of population.
Results: We suggested that the control systems of a potentially ageless organism are able to sustain a physiological regimen of complete self-maintenance strictly within a certain range of values and changes in external conditions known as 'environmental pressure'. The bell-shaped and the U-shaped curves characterizing 'dose – effect' relationship are well-known features, which describe the regularities of interaction between biological systems and their environments. For this reason even a potentially non-senescent body must start to age under inadequate condition (like a non-melting piece of ice taken out from the deepfreeze inevitably starts to melt at the temperatures above zero Celsius). This conception is totally consistent with existing patterns of mortality and with agelessness potential of somatic stem cells [1–6].
Conclusion: There is no need to build up and explore too complicated "systems models of intrinsic aging" to understand the origin of this mainly extrinsic root cause of natural aging. For this reason a simple phenomenological black-box approach with *Input-Output analysis* is ample. Here *Input* refers to the environmentally dependent initial force of mortality, whereas *Output* is a rate of age-related increase of mortality force.

References

1. Khaliavkin A.V. (2001) Influence of environment on the mortality pattern of potentially non-senescent organisms. General approach and comparison with real populations. Advances in Gerontology. 2:46-49.
2. Khalyavkin A.V., Yashin A.I. (2007) Nonpathological senescence arises from unsuitable external influences. Annals of the New York Academy of Sciences. 1119:306-309.
3. Khalyavkin A.V. (2012) From macro- to nano-systems and back in search of the primary cause and control of aging. Proceedings of International Conference "Instabilities and control of Excitable Networks: From Macro- to Nano-Systems", 69-76 (Dolgoprudny). DOI 10.13140/2.1.4408.0326
4. Khalyavkin A.V. (2013) Phenoptosis as genetically determined aging influenced by signals from the environment. Biochemistry (Moscow). 78(9):1001-1005. DOI 10.1134/S0006297913090058
5. Khalyavkin A.V., Krut'ko V.N. (2014) Aging is a simple deprivation syndrome driven by a quasi-programmed preventable and reversible drift of control system set points due to inappropriate organism-environment interaction. Biochemistry (Moscow). 79(10):1133-1135.
6. Khalyavkin A.V., Krut'ko V.N. (2018) How regularities of mortality statistics explain why we age despite having potentially ageless somatic stem cells. Biogerontology. 19(1):101-108.

DOI 10.18699/BGRSSB-2018-212

On the choice of control objects in experimental gerontological research

A.N. Khokhlov*, A.A. Klebanov, G.V. Morgunova
School of Biology, Lomonosov Moscow State University, Moscow, Russia
** e-mail: khokhlov@mail.bio.msu.ru*

Key words: aging, senescence, geroprotectors, control groups, cell cultures, experimental animals

Recently, a large number of papers have appeared that describe the successful use of various biologically active compounds (mitochondrial antioxidants, antidiabetic biguanides, mimetics of dietary restriction, etc.) as geroprotectors. However, in our opinion, in most cases, the positive results of such studies are determined by a "successful" selection of control objects. As such, animals with some abnormalities are often used [1, 2], so that any favorable effect on the corresponding pathological processes leads to an increase in life span. Besides, control animals can be normal, i. e. wild type, but placed in some extreme conditions, which can be overcome precisely by certain biologically active compounds. Thus, treatment of pathologies is present, and not an effect on the fundamental processes of aging [3]. There is a point of view, according to which the results of Clive McKay's experiments, which have significantly prolonged the life of rats by limiting caloric intake, are determined, firstly, by the fact that the control animals fed *ad libitum* (which is not at all characteristic of animals in the wild), and secondly, because the Fisher-344 rats used in experiments are short-lived. It should be said that the above considerations seem to concern also gerontological experiments on cultured cells. In particular, we sometimes hear from our colleagues remarks about the model of stationary phase aging of cell cultures used in our laboratory due to the fact that most of the experiments are carried out on transformed rather than normal cells. However, this approach seems to us quite justified, because the phenomenon of stationary phase/chronological aging is common to a wide variety of cells, including bacteria, yeasts, cyanobacteria, mycoplasmas, animal and plant cells [4]. Cells with unlimited mitotic potential do not change either from experiment to experiment or during long-term cultivation without subcultivation [3], which cannot be said of normal diploid fibroblasts, whose telomeres are shortened with each division (and from the moment of seeding of the cells at low density to their entering the stationary phase of growth they can divide up to 10 times!).

Conclusion: To search for effective geroprotectors, which provide an impact on the fundamental mechanisms of aging, it is necessary to conduct studies on "maximally healthy" animals or on "maximally stable" model systems.

References

1. Shabalina I.G. et al. (2017) Improved health-span and lifespan in mtDNA mutator mice treated with the mitochondrially targeted antioxidant SkQ1. Aging (Albany NY). 9(2):315-336.
2. Takeda T., Hosokawa M., Higuchi K. (1997) Senescence-accelerated mouse (SAM): a novel murine model of senescence. Exp. Gerontol. 32(1-2):105-109.
3. Khokhlov A.N., Klebanov A.A., Morgunova G.V. (2018) Anti-aging drug discovery in experimental gerontological studies: from organism to cell and back. In: Aging: exploring a complex phenomenon, Ed. by Sh.I. Ahmad, Boca Raton: Taylor & Francis. 577-595.
4. Morgunova G.V., Klebanov A.A., Khokhlov A.N. (2016) Some remarks on the relationship between autophagy, cell aging, and cell proliferation restriction. Moscow Univ. Biol. Sci. Bull. 71(4):207-211.

DOI 10.18699/BGRSSB-2018-213

Approximation of the growth and survival curve of a non-subcultured cell culture within the framework of the stationary phase aging model

A.A. Klebanov, G.V. Morgunova, A.N. Khokhlov*
School of Biology, Lomonosov Moscow State University, Moscow, Russia
** e-mail: khokhlov@mail.bio.msu.ru*

Key words: stationary phase aging, cell cultures, growth curve, survival curve, approximation

Motivation and Aim: The most reliable way to test geroprotectors is to construct the survival curves of organisms. For a long time the phenomenon of our stationary phase aging model was considered as accumulation in cultured cells, the proliferation of which is restricted by some way (for example, contact inhibition), of "age-related" lesions similar to age-related changes *in vivo*. At the same time, in our cell-kinetic model, only the growth rate of the culture and the saturating density reached, which characterizes the "biological age" of the cells, were estimated, and the kinetics of their death was not considered. However, we subsequently showed that during "stationary phase aging" the cells, accumulating "age-related" damage, die out in accordance with the Gompertz equation, i. e. really grow old, and the probability of their death increases with time exponentially, as in aging animals/humans. Thus, the growth and death curves of a non-subcultured cell culture could be approximated by two different equations. At the same time, a number of data showed that both parts of the curve can be interrelated. We hypothesized that the creation of a single equation describing both the initial growth of the culture and the death of cells in the stationary phase could significantly improve the efficiency of testing various geroprotectors and geropromoters on this model system.
Methods and Algorithms: The corresponding growth and death curves of the non-subcultured cell culture were obtained using transformed Chinese hamster cells of the B11-dii FAF28 line which were incubated at 37 °C in Dulbecco's modified Eagle's medium supplemented with 10 % bovine serum in hermetically sealed vials for 40–55 days. For subsequent mathematical processing, new data on the influence of the buffer capacity of the growth medium and exogenous 8-oxo-2′-deoxoguanosine on the growth and stationary phase aging of cells were used as well as earlier results concerning the effects of C_{60}-fullerene and Quinton Marine Plasma in this model system. Approximation was carried out using either SigmaPlot 12.0 (Systat Software Inc.) built-in equations or our own ones.
Results: It turned out that, from the equations embedded in SigmaPlot, our experimental data can best be approximated using the equation of the 4-parameter log-normal distribution. However, herewith the "plateau" practically disappears on the growth curve. This problem can be avoided by using the equation developed by us, which is based on the much modified Verhulst-Pearl equation. As a result, the reliability of comparing the control and experimental curves is significantly increased. In particular, we managed to demonstrate a significant effect on the kinetics of cell growth and survival of the culture medium buffer capacity and the absence of such an effect in the case of exogenous 8-oxo-2′-deoxoguanosine, which was not obvious without the use of approximation.

DOI 10.18699/BGRSSB-2018-214

Contribution of neuroinflammation to the development of Alzheimer's disease-like pathology in OXYS rats

O.Yu. Koneva[1, 2]*, E.A. Rudnitskaya[1], Yu.V. Rumyantseva[1], N.A. Stefanova[1]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: okoneva94@gmail.com*

Key words: Alzheimer's disease, neuroinflammation, RNA-Seq, OXYS rats

Motivation and Aim: Brain aging is central to Alzheimer's disease (AD), although the mechanisms by which it occurs are not fully understood. In the last decade, accumulating evidence in support of a role of age-related dysregulation of neuroinflammation and immune system in AD. It remains unclear whether astrocytes and immune cells – primarily microglia – influence disease onset, progression or both. Using senescence-accelerated OXYS rats, which simulate key characteristics of sporadic AD, we evaluated contributions of neuroinflammation to the disease development.
Methods and Algorithms: At preclinical (age 20 days), early (5 months), and advanced (18 months) stages of AD-like pathology, in the prefrontal cortex of OXYS and Wistar (control) rats, we evaluated (i) the state of microglial and neuronal cells by histological and immunohistochemical analysis; (ii) differences in the cell-specific (astrocytes, microglia and neurons) gene expression according to RNA sequencing (RNA-Seq) data; (iii) the expression of the main genes of cytokines, chemokines and components of the complement system by RT-PCR.
Results: The number of microglial cells and the expression of the genes involved in immune processes in the cortex of OXYS rats at preclinical stage of the disease indicate the absence of inflammation and increased phagocytic activity, possibly associated with the elimination of "transitional" cells by apoptosis during the formation of interneuronal contacts. By the age of 5 mo the level of mRNA of most of the genes in OXYS rats was increased on the background of neurodegenerative changes. In addition, the changes in expression of microglia-specific genes in OXYS rats were functionally associated with innate immune response and NF-κB pathway. By the age of 18 mo the number of microglial cells was increased and the changes in expression of microglia-specific genes in OXYS rats were associated with innate immune response, response to cytokine and phagocytosis. Notably, 18-mo animals did not differ in the density of microglia and the number of activated microglial cells, which indicates absence of hyperactivation of microglia in OXYS rat cortex at advanced stage of AD-like pathology. However, the number of macrophage cells in OXYS rats was higher than in Wistar rats which may reflect increased need for microglia in conditions of progressive neurodegeneration. This is indicated by a decrease in the expression of the immune system genes and by changes in expression of cell-specific genes associated with regulation of angiogenesis and signal transduction in astrocytes and synaptic transmission in neurons.
Conclusion: These results suggest that dysregulation of the immune system function and the loss of neuroprotective properties of microglia and astrocytes may contribute to the development of AD-like pathology in OXYS rats.
Acknowledgements: This work was supported by grant from the Russian Foundation for Basic Research (project No. 18-015-00320).

DOI 10.18699/BGRSSB-2018-232

Neurogenesis in the hippocampus of senescence-accelerated OXYS rats in the juvenile period

T.A. Kozlova*, N.A. Stefanova, N.G. Kolosova, E.A. Rudnitskaya
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: kozlovatanya21@gmail.com*

Key words: neurogenesis, Alzheimer's disease, OXYS rats

Motivation and Aim: The increasing quality of life and advances of medicine in developed countries lead to increased lifespan and as a consequence growth of percentage of elderly in population. Aging is accompanied by a couple of diseases including Alzheimer's disease (AD). In recent years it was shown, that AD-related neurodegenerative changes are associated with alterations of the neuronal plasticity. Neurogenesis is one of the major mechanisms of neuronal plasticity, and its alterations already at the young age may contribute to AD manifestation. However, the precise mechanisms of the process are not fully understood. To investigate the contribution of variations of juvenile neurogenesis to pathogenesis of AD we used OXYS rats.

Methods and Algorithms: 3-10-, 20- and 45-days-old male OXYS (model of sporadic form of AD) and Wistar (control) rats were used. Immunohistochemistry was used to identify the number of neuronal cells at different stages of maturation. Development of the forelimb grasp, righting, negative geotaxis and cliff avoidance reflexes was investigated. Animal behavior and learning were evaluated in the Open field, Dark cylinder, Elevated plus maze and Morris water maze tests.

Results: OXYS pups demonstrated delayed development of negative geotaxis and righting reflexes as compared to Wistar pups. Locomotor activity of rats of both strains increased from 10 to 45 days of age; however it was lower in OXYS rats compared to Wistar rats at all analyzed ages. Moreover, 45-days-old OXYS rats demonstrated increased anxiety as well as learning and memory deficits. Behavioral and cognitive changes may reflect alterations of neuronal plasticity. Indeed, in dentate gyrus (DG) of the hippocampus of 10-day-old OXYS rats the neuronal cell density was almost twice as large as in Wistar rats due to the increased content of neuroblasts and immature neurons. From 10 to 20 days of age the neuronal cell density in DG of Wistar rats did not change while in OXYS rats it decreased. From 20 to 45 days of age the density of neuronal cells increased only in the DG of Wistar rats due to the increased density of immature neurons reflecting activation of neurogenesis in Wistar rats.

Conclusion: Delay of reflexes' development and alterations of behavior and cognitive function at young age may reflect retardation of brain development in OXYS rats. Indeed, we showed increased density of immature neurons in the DG of 10-day-old OXYS rats which may be the result of delayed neuronal maturation and/or delay of neonatal apoptosis of the neurons. Moreover, activation of neurogenesis in the DG of 45-day-old OXYS rats did not occur. These alterations of neuronal plasticity in the DG of OXYS rats occur before detection of any signs of neurodegeneration. Therefore retardation of brain development early in life may be considered as one of the risk factors associated with manifestation of AD-like pathology in adult OXYS rats.

Acknowledgements: supported by grant from the Russian Foundation for Basic Research (project no. 18-015-00320).

DOI 10.18699/BGRSSB-2018-215

The influence of environmental and social factors on the parameters dynamics of the Gompertz function

A.V. Krementsova
Emanuel Institute of Biochemical Physics, RAS, Moscow, Russia
e-mail: akrementsova@mail.ras.ru

Key words: life-span, the Gompertz force of mortality, Strehler-Mildwan correlation

It is known that the life span distribution of humans and laboratory animals is well fitted by the Gompertz function. The analysis of demographic data makes it possible to evaluate both the contribution of natural environmental factors that have a time localized, for example, a crop failure or epidemic, and socially significant factors, for example wars or the achievement of medicine in the dynamics of the life span distribution. In the work, demographic data of the survival during the last 200 years of a several European countries were analyzed. Based on these data, the parameters of the Gompertz function were identified. The Strehler-Mildvan correlation dependences between these parameters were obtained. It is shown that, despite the fact that environmental factors can strongly influence to the dynamics of the Gompertz function parameters, they do not violate the Strehler-Mildvan correlation, while the influence of social factors leads to its violation. Currently there is an active discussion in the literature about the origin of the Strehler-Mildvan correlation. Data are given that the correlation is a mathematical artifact, and it is not a reflection of real demographic patterns. To verify this statement, the parameters of the Gompertz function in the variation proposed by Gumbel were identified. It is shown that in this case also the correlation dependence between the parameters of the Gompertz function were observed. The results obtained make it possible to conclude that the dynamics of the life span distribution is not accidental, but it is a reflection of the internal patterns of aging.

DOI 10.18699/BGRSSB-2018-216

Molecular signatures of Alzheimer's disease and aging in the *TOMM40-APOE-APOC1* locus

A.M. Kulminski*, L. Shu, Y. Loika, I. Culminskaya
Biodemography of Aging Research Unit, Social Science Research Institute, Duke University, Durham, USA
** e-mail: Alexander.Kulminski@duke.edu*

Key words: Alzheimer's disease, aging, linkage disequilibrium, molecular signatures

Motivation and Aim: Enduring interest to the apolipoprotein E (*APOE*) region is driven by remarkably strong associations of variants from this region with Alzheimer's diseases (AD) and pleiotropic associations with multiple aging-related traits. The role of this region in pathogenesis of AD and aging remains, however, poorly understood. Elucidating genetic predisposition to aging-related traits characteristic for post-reproductive period is hampered by the uncertain role of evolution in establishing their molecular mechanisms. This uncertainty is inevitable source of natural-selection-free genetic heterogeneity in predisposition to AD and aging-related traits.
Methods and Algorithms: We examined linkage disequilibrium (LD) structures characterized by nine single nucleotide polymorphisms (SNPs) from *TOMM40-APOE-APOC1* locus, including rs429358 and rs7412 SNPs coding the *APOE* ε4 and ε2 alleles, in 2,661 AD-affected and 16,079 AD-unaffected subjects and in 570 short-lived (<75 years, SL) and 1,999 long-living (85+ years, LL) subjects from four independent studies.
Results: The LD structures, being heterogeneous, are significantly different in subjects with and without AD, $p < 2 \times 10^{-4}$. The pattern of the significant difference represents molecular signature of AD comprised of SNPs from these genes. We identified 31 of 36 SNP pairs with pair-wise estimates of the LD difference between subjects with and without AD significant after Bonferroni correction, $p < 1.3 \times 10^{-3}$. In contrast, differences in LD between SL and LL subjects attained only marginal Bonferroni-adjusted significance ($p = 3.5 \times 10^{-3}$) for one SNP pair (rs405509 [*APOE*] and rs439401 [intergenic]), nominal significance ($p < 5 \times 10^{-2}$) for three pairs and suggestive significance ($p < 10^{-1}$) for three more pairs. For all these seven SNP pairs, LD changed in the same directions in AD/no-AD and LL/SL groups.
Conclusion: Significant and highly heterogeneous molecular signature of AD provides evidence on complex polygenetic predisposition to AD in the *TOMM40-APOE-APOC1* locus. Significant differences in pair-wise LD in subjects with and without AD indicate SNPs, or their proxies, likely involved in AD pathogenesis. The same directions of the LD differences in AD/no-AD and LL/SL groups suggests heterogeneous, partly overlapping molecular mechanisms for AD and aging, defined as survival to old ages, in the *APOE* region.
Acknowledgements: Supported by Grants P01 AG043352 and R01AG047310 from the National Institute on Aging.

DOI 10.18699/BGRSSB-2018-217

Neurotrophin signaling pathway in development of AMD-like retinopathy

M.D. Matytsin*, O.S. Kozhevnikova, E.A. Rudnitskaya, D.V. Telegina
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: telegina@bionet.nsc.ru*

Key words: aging, neurotrophic factors, retinopathy, AMD, NGF, BDNF, OXYS rats

Motivation and Aim: Age-related macular degeneration (AMD) is complex, multifactorial disease that is the main cause of vision loss in people over than 60 years old. As neurotrophins manage neuronal survival, death and synaptic plasticity, we hypothesized that the neurodegenerative changes in the retina during aging and AMD development may be accompanied by alterations in neurotrophin signaling pathway/supplementation, which may be caused by alterations of their level or their bioavailability. To investigate a link between age-related alterations of neurotrophin signaling pathway and progression of AMD we used senescence-accelerated OXYS rats as a suitable model of AMD.
Methods: The RNA-seq data obtained from the retinas of 20 day-, 3-, and 18-month-old OXYS and Wistar (control) rats were used to analyze changes in neurotrophin signaling pathway. Immunohistochemistry (IHC) was used to localize neutrophic factors – proBDNF and NGF, and their corresponding receptors $p75^{NTR}$ and TrkA.
Results: Based on comparative analysis of the transcriptome of retina of OXYS and Wistar rats, differently expressed genes from neurotrophin signaling pathway were identified at the preclinical (20 d), the early (3 mo) and the advanced (18 mo) stages of retinopathy development in OXYS rats. These groups were functionally annotated using bioinformatics databases DAVID and KEGG. At the age of 20 d the expression of nine genes were changed: 6 genes were downregulated (*Calm2*, *Rap1a*, *Rap1b*, *Rps6ka3*, *Rps6ka6*, *Sos2)* and 3 genes were upregulated (*Atf4*, *Mapk13*, *Mapk14)*. At the age of 3 mo the expression of 7 genes were changed: 6 genes downregulated (*Foxo3*, *Nfkb1*, *Nfkbia*, *Pdk1*, *Rps6ka2*, *Shc3*) and *Pik3r2* were upregulated. During the advanced stage of retinopathy (18 mo), the expression of 19 genes was altered: 12 genes were downregulated (*Arhgdib*, *Camk2g*, *Irak2*, *Jun*, *Mapk13*, *Nfkb1*, *Nfkbia*, *Ntf3, Rap1a*, *Rap1b*, *Rps6ka1* and *Ywhah*) and 7 genes were upregulated (*Mapk12*, *Nfkbib*, *Ngfr*, *Ntrk2*, *Rapgef1*, *Rps6ka2* and *Sos1*). The products of these genes were associated with the development of the nervous system and signal transduction. According to the IHC, we revealed the increased colocalization of proBDNF with its receptor $p75^{NTR}$, in the retina of OXYS rats at 20 d and 18 mo, indicating shift in neurotrophic balance toward an increase of cell death. Importantly, we found the increased colocalization of NGF and TrkA in 18 mo associated with gliosis of Muller cells.
Conclusion: Considerable alterations of neurotrophin signaling pathway were found at the advanced stage of AMD-like retinopathy development in OXYS rats.
Acknowledgements: The experiments in OXYS and Wistar rats at the age of 3 and 18 mo were supported by the budget project No. 0324-2018-0019, whereas the analysis of 20-day-old rats was supported by the RFBR according to the research project No. 18-315-00216.

DOI 10.18699/BGRSSB-2018-218

Maintaining pH of the culture medium in cytogerontological experiments: effect on the cell viability and the shape of the cells' survival curve

G.V. Morgunova*, A.A. Klebanov, A.N. Khokhlov
School of Biology, Lomonosov Moscow State University, Moscow, Russia
** e-mail: morgunova@mail.bio.msu.ru*

Key words: growth medium pH, buffer capacity, cell cultures, stationary phase aging, chronological aging

Motivation and Aim: One of the problems arising when modeling stationary phase/chronological aging of mammalian/yeast cells in a non-subcultured culture is the acidification of the growth medium at later stages. Moreover, some researchers believe that the chronological aging of yeast is associated with the accumulation over time of acetic acid in the medium and, as a consequence, acidifying it to pH 4 and lower [1]. The aim of this work was to assess the effect of different methods of maintaining pH on the growth and survival of a non-subcultured culture of mammalian cells.
Methods and Algorithms: Transformed Chinese hamster cells were incubated at 37 °C in Dulbecco's modified Eagle's medium supplemented with 10 % bovine serum. The contribution of various ways of maintaining pH of the growth medium to viability of the culture as well as its growth and survival was analyzed. For this purpose, HEPES buffer solution was added to the medium in hermetically sealed vials to a final concentration of 20 mM, or the cells in non-hermetical vials were cultured at 5 % CO_2.
Results: It was found that the optimal way to maintain pH was to cultivate the cells not at 5 % CO_2, but in hermetically sealed vials with the addition of HEPES buffer to the growth medium. If the buffer was present in the medium from the moment of seeding, the culture quickly reached the "plateau" and its saturating density was greater than in the control group, with the shape of the curve in both groups being the same and all cells' dying out by 52nd day. The groups differed in pH only at the initial stages, after reaching the "plateau" this parameter remained the same in both control and experimental group until the end of the experiment.
Conclusion: Growth medium pH affects the growth rate and saturation density of the culture, but one can not name the acidification of the medium as the main cause of cell death in the stationary phase. Adding a buffer only provides optimal conditions at the initial stages of culture growth. As for works in which the acidification of the medium is called the main cause of the chronological aging of yeast, from such papers' data it is obvious that in the absence of pH-supporting factors, yeasts simply die out "exponentially", i. e. there is no aging [2]. It should be noted that in our experiments on the study of stationary phase aging of mammalian cells in hermetically sealed culture vials, they "grow old according to Gompertz".
Acknowledgements: G.V. Morgunova's studies have been supported by the RF President's Scholarship (SP-4224.2018.4).

References

1. Fabrizio P., Wei M. (2011) Conserved role of medium acidification in chronological senescence of yeast and mammalian cells. Aging (Albany NY). 3(12):1127-1129.
2. Morgunova G.V., Klebanov A.A., Marotta F., Khokhlov A.N. (2017) Culture medium pH and stationary phase/chronological aging of different cells. Moscow Univ. Biol. Sci. Bull. 72(2):47-51.

DOI 10.18699/BGRSSB-2018-219

MAPK signaling pathway in development of Alzheimer's disease

N. Muraleva
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
e-mail: Myraleva@bionet.nsc.ru

Key words: αB-crystallin, phosphorylation, Alzheimer's disease, OXYS rats

Motivation and Aim: Alzheimer's disease (AD) is an aging-related neurodegenerative disease and accounts for the majority of human dementia. One of the major neuropathological hallmarks of AD is the impairment of proteostasis in the form of deposition of amyloid β-protein (Aβ) and hyperphosphorylated tau in the brain. The molecular chaperones include αB-crystallin (CryAB) are able to prevent aggregation or misfolding of proteins and enable their correct refolding. Mitogen-activated protein kinase signaling pathway (MAPKsp) is regulate of CryAB activity through its phosphorylation. To investigate a link between alteration of MAPKs pathway and development of sings AD we compared the gene expression profiles of frontal cortex of OXYS rats and control Wistar rats. Next we compared this result with the CryAB and pS59-CryAB proteins expression and its localization in the brain.
Methods and Algorithms: We used 20-day-old, 5 and 18 mo-old OXYS rats at the pre-symptomatic, symptomatic, and progressive stage of AD-like pathology, respectively, and Wistar rats as the control (n = 6). Analyze of changes in MAPKsp was obtained, using the date of the frontal cortex RNA-seq. CryAB and CryABS59 expressions in brain were indicated by western blot and immunostaining analyzis.
Results: According to KEGG pathway, 27 genes from the MAPKs pathway were changed of which 17 genes were upregulated in cortex of 20 days-old OXYS rats compared to Wistar rats. OXYS rats showed changes in expression of 22 genes, including 15 genes with upregulation, related to MAPKsp at the age of 5 months. The number of genes with altered expression increased to 65 (44 up- and 21 downregulation) by the age of 18 months. Importantly, the mRNA of the MAPKsp activator genes increased with age in OXYS rats. Thus, these results are exactly the same as data of the cortex RNA-seq of the cerebral in humans with AD from open databases. Analysis of protein content in the cortex showed that CryAB level did not differ between OXYS and Wistar rats and did not change with age in the both protein fractions. Phosphorylation of CryaB was absent in detergent – soluble protein fraction of OXYS and Wistar rats. PS59-CryAB was detected only in detergent-insoluble fraction of rats and its level increased with age in the both strain. Importantly, pS59-CryAB level was higher in OXYS rats compared with the Wistar at the age of 5 and 18 mo. This result was confirmed by immunostaining analysis. The pS59-CryAB was localized with Aβ and its co-localization was increased in OXYS rats at 18 mo of age.
Conclusion: We showed that the development of the signs of AD in OXYS rat takes place during to alteration of MAPKsp activity in brain cortex. Increase in the level of phosphorylation with age and its co-localization with Aβ may be considered as a compensatory response to the accumulation of aggregates of proteins in OXYS rats. Importantly, mRNA expressions of MAPKsp genes in OXYS rats at the age of 18 mo were same as for humans with AD. These results demonstrate the role of the MAPKsp in the AD suggest that MAPKs could be targeted for novel treatments in the future.
Acknowledgements: This study was supported by the RFBR (project No. 18-015-00336).

DOI 10.18699/BGRSSB-2018-220

Expression of neurogenesis-associated genes during development of Alzheimer's disease-like pathology in OXYS rats

E.A. Rudnitskaya*, N.G. Kolosova, N.A. Stefanova
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: ekaterina.rudnitskaia@gmail.com*

Key words: neurogenesis, neurotrophin, Alzheimer's disease, OXYS rats

Motivation and Aim: Alzheimer's disease (AD) is the most common type of age-related dementia worldwide. However, the precise mechanisms of its pathogenesis are not fully understood. One of the processes that may contribute to neurodegeneration is alteration of neuronal plasticity. Neurogenesis is one of the major mechanisms of neuronal plasticity, and neurotrophic supply is crucial for it. Using OXYS rats as a suitable model of AD previously we have shown that development of AD signs is accompanied by changes in expression of genes involved in neurotrophic signaling pathway. Thus in this work we investigated a link between changes in expression of neurogenesis-associated genes and development of AD-like pathology in OXYS rats.
Methods and Algorithms: 20-days-, 3-5- and 18-months-old male OXYS and Wistar (control) rats were used. The RNA-seq data obtained for hippocampus were used to analyze differentially expressed genes involved in neurogenesis according to MANGO (Mammalian Adult Neurogenesis Gene Ontology) database. These genes were functionally annotated using DAVID (Database for Annotation, Visualization and Integrated Discovery). ELISA was used to quantify the levels of nerve growth factor (NGF) and brain-derived neurotrophic factor (BDNF), TrkA and $p75^{NTR}$ receptors, western-blot analysis was used to quantify levels of TrkB and phosphorylated TrkB (phTrkB) receptors in the hippocampus.
Results: We found that to the age of manifestation of AD signs in OXYS rats (from 20 days to 5 months) changes in expression of 22 genes were involved in the hippocampal neurogenesis. Functionally, these genes were associated with neuronal precursor cell proliferation, cellular response to glucose stimulus and positive regulation of angiogenesis. Changes in gene's expression might be due to activation of neurotrophic signaling observed in OXYS rats at the age of 3 months: levels of BDNF, pro-survival TrkA receptor of NGF and pro-apoptotic $p75^{NTR}$ receptor were higher compared to Wistar rats. Progression of AD-like pathology in OXYS rats (from 5 to 18 months of age) is accompanied by changes in expression of 25 genes involved in neurogenesis. These genes were associated with angiogenesis, response to oxidative stress and negative regulation of cell death. The changes occurred against background of depletion of neurotrophic supply: the levels of BDNF and activation of its receptor (phTrkB/TrkB ratio) decreased in OXYS rats. However, the levels of NGF increased with age in both rat strains and may be considered as a compensatory response to slow down neuronal loss of function.
Conclusion: Alterations of neurotrophic supply in the hippocampus occur during development of AD-like pathology in OXYS rats and may result in disturbances of hippocampal neurogenesis.
Acknowledgements: This work was supported by grant from the Russian Foundation for Basic Research (project No. 18-015-00320).

DOI 10.18699/BGRSSB-2018-221

Longevity in mammals: lost genes as a determinant

A.V. Seliverstov[1]*, G.A. Shilovsky[2], L.I. Rubanov[1], O.A. Zverkov[1], V.A. Lyubetsky[1, 2]
[1] *Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute), Moscow, Russia*
[2] *Lomonosov Moscow State University, Moscow, Russia*
** e-mail: slvstv@iitp.ru*

Key words: lifespan, longevity, neoteny, naked mole-rat, synteny, orthology, gene loss

Motivation and Aim: The genetic propensity of certain species for longevity and anti-ageing is a challenging problem of vertebrate biology etc. Following our general hypothesis on the evolutionary significance of gene loss, we tried to identify lost genes in mammals with a greater lifespan than could be expected from their body size: mole rats, and primates (including the naked mole rat, little brown bat, capuchin monkeys, gibbons, western gorilla, bonobo, chimpanzee, and human). Species living under different conditions were analyzed to reduce the habitat impact on the phenomenon studied.
Methods and Algorithms: This study relies on our previous bioinformatics method experimentally approved while testing the hypothesis for regeneration potential and endbrain development. The method takes into account the similarity and local synteny of genes and generates amazingly short lists of lost genes.
Results: The list of lost mouse genes includes *Abca14*, *Ace3*, *Csn1s2a*, *Gm17416*, *Hist1h2af*, *Prss43*, *Slc6a21*, *Smpd5*, *Spint5* (a Kunitz type 5 serine protease inhibitor), *Ttc41*, *Wap*, *Wfdc16* (a peptidase inhibitor affecting cell proliferation), *2310003L06Rik* expressed in the tongue, and certain vomeronasal and olfactory receptor genes.
Discussion: The mouse *Smpd5* gene encodes the neutral sphingomyelinase responsible for breaking down sphingomyelin, a critical factor of the integrated mitochondrial reticulum formation in the skeletal muscles; and its development is notably decelerated in the naked mole rat relative to mouse. A homologous sphingomyelinase is involved in apoptosis initiation through the formation of pores in the mitochondrial membrane. The mouse *Tmbim7* encodes a transmembrane protein, a putative apoptosis inhibitor; it is actively expressed in the mouse testis and corresponds to the human TMBIM7P pseudogene. The gene corresponds to the *ENSGL00000028944* pseudogene in the female naked mole rat; without regard to synteny, the gene aligns with the *TMBIM6* gene, which is overexpressed in certain cancer types and suppresses *Bax*-induced apoptosis.
Conclusion: The revealed mouse genes demonstrate specific expression in reproduction-associated tissues, which agrees with the Williams' hypothesis concerning the reallocation of physiological recourses of an organism between the self-maintenance and reproduction. The loss of some revealed vomeronasal and olfactory receptor genes in human and naked mole rat conforms to special anatomical features. We suggest that the loss of certain genes in evolution is a determinant of lifespan elongation and ageing deceleration including neoteny.
Availability: http://lab6.iitp.ru/en/lossgainrsl/
Acknowledgements: The research was supported by the Russian Science Foundation, project No. 14-50-00150. The research was carried out using supercomputers at Joint Supercomputer Center of the Russian Academy of Sciences (JSCC RAS).

DOI 10.18699/BGRSSB-2018-222

Is retardation of development of retina and brain a predictor of age-related macular degeneration and Alzheimer's disease?

N.A. Stefanova*, O.S. Kozhevnikova, D.V. Telegina, N.G. Kolosova
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: stefanovan@bionet.nsc.ru*

Key words: age-related macular degeneration, Alzheimer's disease, RNA sequencing, OXYS rats

Motivation and Aim: Alzheimer's disease (AD) and age-related macular degeneration (AMD) are two complex incurable neurodegenerative disorders the common pathogenesis of which is actively discussed. They manifest in the accumulation of amyloid beta (Aβ) in abnormal extracellular deposits: senile plaques in the brain of AD patients and drusen in the eyes of AMD patients. AD and AMD are complex diseases, formation of which is controlled by a variety of interacting genetic and environmental factors. Their manifestation occurs after the unfolding of the underlying events at the molecular level that are difficult to study in humans. Using senescence-accelerated OXYS rats, which simulate key characteristics of AMD and sporadic AD, we evaluated molecular and genetic background of the development of these diseases.
Methods and Algorithms: At preclinical (age 20 days), early (3–5 mo), and advanced (18 mo) stages of AMD- and AD-like pathologies in the retina and prefrontal cortex of OXYS and Wistar (control) rats, we evaluated differences in gene expression according to RNA sequencing (RNA-Seq) data to identify the metabolic processes and pathways, the changes in the activity of which lie at the basis of the transition from healthy aging to the development of AMD and AD.
Results: The manifestation of signs of AMD- and AD-like pathologies in OXYS rats (3–5 mo) is associated with change in the expression of genes associated with the immune system, inflammation, oxidative stress, calcium homeostasis, and apoptosis. Progression of the pathologies (3–18 mo) occurs against the background of alteration in expression of genes associated with the metabolic pathway of AD, including those related to Aβ precursor protein processing, aggregation and degradation of Aβ, synaptic processes, and mitochondrial dysfunction – in retina and cortex. Our results are consistent with RNA-Seq data on changes in the transcriptome of the prefrontal cortex of AD patients. Importantly, at the preclinical period in the retina and the cortex of OXYS rats, the expression of genes whose products are involved in the development of the central nervous system, synaptic transmission and neuronal plasticity is changed.
Conclusion: Our results suggest that retardation of the development of retina and brain may predict the emergence of the signs of AMD and AD in OXYS rats and, possibly, of these diseases in humans. This is supported by the data on the delayed hippocampal neurogenesis and reflex formation in early postnatal period as well as reduced duration of pregnancy (5 % shorter than in Wistar, $p < 0.0003$) in OXYS rats (Kozlova et al., 2018). Our results raise the question that preterm neonates may have a higher chance of AMD and AD development in the future even in the absence of cognitive impairments in early ontogenesis and retinopathy in prematurity.
Acknowledgements: This work was supported by a Russian Scientific Foundation grant (16-15-10005).

DOI 10.18699/BGRSSB-2018-223

Apoptosis and autophagy alterations are involved in the development of Alzheimer's disease-like pathology in OXYS rats

G.K. Suvorov*, D.V. Telegina, N.A. Stefanova
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: telegina@bionet.nsc.ru*

Key words: apoptosis, autophagy, Alzheimer's disease, OXYS rats, RNA-seq

Motivation and Aim: Alzheimer's disease (AD) is the most common neurodegenerative disorder and is a cause of dementia that is characterized by accumulation of neurotoxic forms of the peptide amyloid-β (Aβ), a hyperphosphorylation of the tau protein, a deficit of synapses, inflammation, mitochondrial dysfunction, oxidative stress, and neuronal death. Autophagy and apoptosis are basic physiologic processes contributing to the maintenance of cellular homeostasis. The proper function and balance in the action of these two type of cell death are especially important in neurons and other long-lived cells. Hence, their dysfunction contributes to pathogeneses of neurodegenerative diseases such as AD but the information on their alterations during development of AD is very limited. Our aim is to researcher the dynamics of apoptosis and autophagy in prefrontal cortex of senescence-accelerated OXYS rats that simulate key aspects of sporadic AD. We investigated three stages of the disease (pre-symptomatic, 20 days; symptomatic, 5 month; and progressive stage, 18 month) in OXYS rats, using RNA-Seq technique and immunohistochemical analysis.
Results: Our results show that the development of the signs of AD (between ages 20 days and 5 months) in OXYS rats takes place during changes in mRNA expression of the 7 genes that are mostly related to processes of autophagy, such as Atg12, Atg7, and Atg8 (regulators of elongation). In addition, changes in mRNA expression of the 21 genes were related to apoptosis (proapoptotic genes and inhibitors) in the prefrontal cortex of OXYS rats between ages 20 days and 5 months. In OXYS rats, with progression of disease, 24 genes related to apoptosis and 7 genes related to autophagy change their expression. Importantly, Wistar rats show changes in expression of 21 genes related to apoptosis only between ages 20 days and 5 months. We also indicated the upregulation of 5 proapoptotic genes in 20-day-old OXYS rats compared Wistar rats. At the age of 5 and 18 months in OXYS rats, the balance between pro- and anti-apoptotic genes were changed (compared to Wistar rats). Using TUNEL kit, we demonstrated that level of apoptosis was increasing in OXYS rats compare age-matched Wistar rats (control) at the age of 20 days, 5 and 18 mo. The analysis of the results of immunohistochemical staining of the brain sections showed that the LC3 A/B (marker of autophagy) was predominant in the glial cells in prefrontal cortex of 20 day-old and 5 mo rats of both lines. The content of LC3 protein in the prefrontal cortex of OXYS and Wistar rats increased by the age of 18 mo and LC3 localized predominantly in neuron. In addition, the extensive colocalization was observed between Mfn1-labeled mitochondria and LC3 A/B at the age of 18 mo in both rats lines that indicated the increasing mitophagy with age.
Conclusion: Accordingly, we demonstrated that the development of AD-like pathology in OXYS rats is related to the alterations in processes of autophagy and apoptosis.
Acknowledgements: Supported by the Russian Science Foundation (project No. 161510005).

DOI 10.18699/BGRSSB-2018-224

Neuronal transcriptional networks in life span control

A.V. Symonenko[1], N.V. Roshina[1, 2], A.V. Krementsova[1, 3], E.R. Veselkina[1], O.Y. Rybina[1, 4], E.G. Pasyukova[1]*

[1] *Institute of Molecular Genetics of RAS, Moscow, Russia*
[2] *Vavilov Institute of General Genetics of RAS, Moscow, Russia*
[3] *Emmanuel Institute of Biochemical Physics of RAS, Moscow, Russia*
[4] *Moscow State Pedagogical University, Institute of Biology and Chemistry, Moscow, Russia*
** e-mail: egpas@rambler.ru*

Key words: Life span, transcription factors, the nervous system, *Drosophila*

Motivation and Aim: The nervous system has long been suggested as a key tissue which defines life span. Numerous and diverse interactions between the nervous system and life span are reciprocal and intimately linked. Several pathways play key roles both in longevity control and in shaping the nervous system during development. The question remains whether their impact on lifespan is related to their developmental functions, or whether it is explained by other pleiotropic influences later in life.

To get closer to solving this issue, we assess the effects of the genes encoding neuronal transcription factors on life span using genetic approaches combined with RT-qPCR, RNA-seq and immunohistochemical analyses.

Results: We previously demonstrated that mutations/reversions and tissue-specific RNAi knockdown and overexpression of several genes encoding transcription factors involved in the development of the nervous system (*stc*, *Lim3*, and others) affect *Drosophila melanogaster* life span. Interestingly, embryonic knockdown of both *stc* and *Lim3* increased life spans in adult flies, indicating that it is a decrease in transcription at the developmental stages that matters for longevity. These results made us assume that the long-term effects of alterations in embryonic gene expression may be epigenetically inherited or may affect transcriptional cascades that predetermine properties of the adult. Unexpectedly, alterations in *stc* and *Lim3* expression which either decreased or increased life span were associated with reduced synaptic function. These preliminary results indicated that the effects of *stc* and *Lim3* on life span and the development of the nervous system might be uncoupled. To better understand mechanisms providing *stc* and *Lim3* impact on development and aging, we assessed effects of these transcription factors on mRNA levels of possible primary and secondary target genes in embryos. Our results show that *stc* and *Lim3* affect neuronal development, protein metabolism, mitochondrial function and energy homeostasis, which indicates probable molecular mechanisms providing their impact on longevity.

Naturally occurring polymorphisms located in target sites of regulatory proteins in the regulatory regions of *stc* and *Lim3* were associated with variation in life span and gene transcription. Effects of naturally occurring variation on gene transcription were directly confirmed in experiments with cell culture and flies. X-ChIP technique was used to show that Polycomb/Trithorax response element was located in the *Lim3* distal promoter and ensured physical and functional interactions with the Polycomb group (PcG) proteins. Based on these findings, we suggested that PcG proteins are directly involved in the regulation of *Lim3* transcription and might be important for lifespan control.

Conclusion: Systemic regulation of neuronal transcriptional networks is proposed as one of the mechanisms regulating life span.

Acknowledgements: Supported by the RFBR grant No. 18-04-01127-a and Russian State Project (01201355486).

DOI 10.18699/BGRSSB-2018-225

Molecular and cellular mechanisms of age-related macular degeneration: evidences from OXYS rats

D.V. Telegina, A.Z. Fursova, O.S. Kozhevnikova*
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: oidopova@bionet.nsc.ru*

Key words: AMD, retina, transcriptome, autophagy, glial cells, RPE, OXYS rats

Motivation and Aim: Age-related macular degeneration (AMD) is a complex neurodegenerative progressive eye disease, resulting in severe loss of central vision in the aging population. A number of genetic and non-genetic factors influences on the pathogenesis of AMD, which remains poorly understood. A major obstacle for understanding of the pathophysiology of AMD is its complexity and a lack of an animal model that can adequately replicate key features of the human disease. Here we present data of the analysis of clinical, histological and molecular manifestations of AMD-like retinopathy in OXYS rats, which simulate key features of AMD.
Methods: Ophthalmoscopy, analysis of RNA-Seq data, gene ontology and pathway annotation, western-blot analysis, immunohistochemistry, confocal microscopy.
Results: Using retinal mRNA profiles generated by RNA-seq we found hundreds differentially expressed (DE) genes at the preclinical (20 d), the early (3 mo) and the advanced (18 mo) stages of retinopathy in OXYS rats. Functional analysis was suggestive of a developmental process, signal transduction, and cell differentiation as the most enriched biological processes among DE genes at 20 d. Functional groups that were significantly enriched for DE genes at 3 and 18 mo included immune response, inflammation, apoptosis, Ca2+ homeostasis and oxidative stress. We showed that pathological processes in the retina of OXYS rats develop against the background of atrophic changes in RPE cells: the hypertrophy, the increase in variability in cell size, the increase in proportion of multinucleated cells, disruptions of the cell form, and irregular immunolabeling. The cell death in the retina of OXYS rats is realized by apoptosis, necrosis and autophagy against the background of phagocytic dysfunction and a decrease in the elimination of dead cells, as indicated by the absence of a migration of macrophages and microglia into the photoreceptor layer. The estimation of age-related alterations of autophagy process in the retina has shown the increased levels of LC3A/B, Atg7, and Atg12 proteins in the OXYS retina during manifestation stage at the age of 3 months. By contrast, in the retina of rats with a progressive stage of retinopathy, we revealed significantly decreased protein levels of autophagy markers. Simultaneously with perturbation of the autophagic response, the necrosome subunits Ripk1 and Ripk3 were detected in the OXYS retina. The downregulation of autophagy markers coincided with amyloid accumulation (Moab-2) in the retinal pigment epithelium and choroid.
Conclusion: Our study emphasizes the importance of autophagic pathway, imbalance in immune and inflammatory responses, aberrant migration of macrophages and microglia in the pathogenesis of AMD and supports the view that the genetic background has a profound impact on AMD development.
Acknowledgements: Supported by the RFBR No. 18-315-00216.

DOI 10.18699/BGRSSB-2018-226

Analysis of metabolic pathways associated with neurogenesis in the retina of OXYS rat – model of age-related macular degeneration

D.V. Telegina*, O.S. Kozhevnikova, N.G. Kolosova
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: telegina@bionet.nsc.ru*

Key words: retina, AMD, neurogenesis, OXYS rats, RNA-seq

Motivation and Aim: Age-related macular degeneration (AMD) is a complex, multifactorial neurodegenerative disease that represents the most common cause of irreversible blindness. In recent years, AMD has grown younger: late stages of the disease are detected in relatively young people. It is known that pathogenesis of AMD based on the age-related structural and functional changes in the retina but the mechanisms that trigger the conversion of the normal age-related changes to the pathological process as well as the background of early development of AMD remain unclear. *De novo* neurogenesis in the adult mammalian retina very limited. Thereby the structural and functional features formed during the period of its maturation and formation can have long-term effects on the further ontogenesis of the tissue. The aim of this project is to study the possible contribution of the changes/disturbances of the postnatal retinal neurogenesis in the early development of AMD on the model of this disease – senescence-accelerated OXYS rats.
Methods and results: We used RNA-Seq to compare gene expression profiles of retinas of senescence-accelerated OXYS rats and age-matched control Wistar rats at the age of 20 days, which is the period of completion of postnatal development of the organ of vision data of the retina. We determined the metabolic pathways controlling processes of neurogenesis based on the analysis of the transcriptome between OXYS and Wistar rats. We obtained a list of 349 DE genes (at pval < 0.01) associated with neurogenesis. Of them, 213 were downregulated, and 136 were upregulated. These genes involved in the 22 functional groups such as nervous system development, neurogenesis, neuron projection development, neuron migration, axon development, dendrite development and eye development (according DAVID databased, FDR < 0.05, EASE > 1.3). As for the KEGG pathway analysis, we found that in OXYS rats compared Wistar rats, the DE genes were associated with a Axon guidance (*Efnb2*, *Rasa1*, *Nrp1* et al.), Hedgehog signaling pathway (*Ptch1*, *Csnk1g3*, *Hhip* et al.), Wnt signaling pathway (*Fzd4*, *Wnt5a*, *Tcf7* et al.) and Notch signaling pathway (*Dtx3*, *Notch2*, *Jag2* et al.) (FDR < 0.05). Using GeneMANIA, we identified 8 hub DE genes with the node degree > 55. Of them, 6 were downregulated (*Col5a2*, *Itga1*, *Dab2*, *Prrx1*, *Fzd2*, *Calu*) and 2 were upregulated (*Ptch1* and *Dlg4*). These genes probably play essential roles in the relevant biological systems. Thereby, the OXYS rats is characterized by disturbance of neurogenesis in retina.
Conclusion: Our data proposed that changes in retina's development precede the appearance of the AMD signs in OXYS rats and probably that the disruption of neurogenesis of retina significantly affects AMD pathogenesis.
Acknowledgements: This work was supported by the RFBR according to the research project No. 18-315-00216.

DOI 10.18699/BGRSSB-2018-227

Reduced expression of *shaggy*, the gene encoding protein kinase GSK3, in dopaminergic neurons increases *Drosophila melanogaster* lifespan

E.R. Veselkina[1]*, M.V. Trostnikov[1, 2], A.V. Krementsova[1, 3], N.V. Roshina[1, 2], E.G. Pasyukova[1]

[1] *Institute of Molecular Genetics of RAS, Moscow, Russia*
[2] *Vavilov Institute of General Genetics of RAS, Moscow, Russia*
[3] *Emmanuel Institute of Biochemical Physics of RAS, Moscow, Russia*
** e-mail: veselkinaer@gmail.com*

Key words: life span, protein kinase, GSK3, the nervous system, dopaminergic neurons, *Drosophila*

Motivation and Aim: Phosphorylation cascades play crucial roles in systemic regulation of numerous metabolic processes. Various protein kinases are well recognized as basic participants of these cascades. A serine-threonine protein kinase GSK3 (glycogen synthase kinase-3) is a well-known protein involved in many functionally different signaling pathways and metabolic processes. In particular, it is important for neurogenesis and neuronal function. We have previously shown that several genes involved in asymmetric neuroblast division, including *shaggy* (*sgg*) encoding *Drosophila melanogaster* GSK3, affect *Drosophila* lifespan. In order to better understand the role of GSK3 in the control of aging and neuronal function, we assessed the effects of tissue-specific increase and decrease in GSK3 expression on lifespan. Our results demonstrated that differential expression of GSK3 is one of the mechanisms involved in a complex regulation of lifespan and aging.

To get closer to understanding molecular mechanisms underlying tissue-specific effects of GSK3 on lifespan, we studied effects of *shaggy* misexpression in dopaminergic neurons in more detail, using genetic approaches combined with RT-qPCR and immunohistochemical analyses.

Results: Dopaminergic neurons were the most sensitive to *shaggy* function: strong *shaggy* overexpression was lethal, moderate *shaggy* overexpression decreased male and female lifespan, while moderate decrease in GSK3 amounts increased mean female lifespan by 20 % and reduced mean male lifespan by 20 %. The results were highly reproducible over several years. Different techniques (different *Drosophila* driver lines) were also used to double check the impact of reduces GSK3 amounts on female lifespan; positive effects were confirmed. Analysis of effects of *sgg* misexpression in dopaminergic neurons on lifespan of mated flies is currently under way. We expect that these data will shed light on the role of fly physiology, sex and mating status in shaping GSK3 effects on lifespan. Moderate *shaggy* overexpression caused a severe reduction of the number of dopaminergic neurons in the brains of adult flies. Of note, decrease in the amount of dopaminergic neurons in the brain is an indication of a number of neurodegenerative diseases in humans, and analysis of molecular mechanisms involved in maintenance of the proper neuron amount and their integrity is of significant interest.

Conclusion: Alterations in the level of expression of GSK3 in dopaminergic neurons underlie systemic sex-specific changes in molecular mechanisms involved in a complex regulation of lifespan and aging.

Acknowledgements: Supported by the RFBR grant No. 18-34-00934 мол_а.

Author index

HUAWEI

Компания Huawei является ведущим мировым поставщиком ИКТ-решений. Благодаря установлению взаимовыгодных отношений с нашими партнерами и заказчиками компании Huawei удалось добиться существенных преимуществ в сфере операторских сетей, корпоративного и потребительского бизнеса, а также в сфере облачных технологий. Мы стремимся создавать максимальные преимущества для операторов связи, предприятий и потребителей путем разработки конкурентных ИКТ-решений и услуг. Оборудование и решения Huawei используются в более чем 170 странах мира. Компания обслуживает более трети населения земного шара.

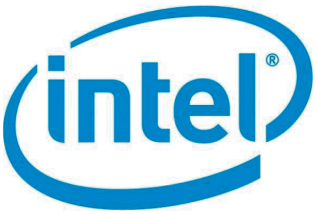
Имея богатый опыт и технические знания в области НИОКР, Huawei придерживается стратегии тесного сотрудничества и интеграции с корпоративными заказчиками и предоставляет им широкий спектр высокоэффективных клиентоориентированных ИКТ-решений и услуг, на базе глубокого понимания их потребностей. Согласно этой стратегии Huawei предлагает широкий выбор передовых ИКТ-решений в сфере государственного управления, общественного сектора, финансов, транспорта, электроэнергетики, крупных предприятий, а также малых и средних предприятий (SME). Эти решения охватывают корпоративные сети, универсальные системы связи и взаимодействия (UC&C), системы облачных вычислений и центры данных, системы корпоративной беспроводной связи, сетевого электропитания, а также инфраструктурные услуги.

ООО «Техкомпания Хуавэй»
Филиал в СФО:
630112, Новосибирск, ул. Фрунзе, 242,
11-й этаж
Тел.: +7(383) 328 00 70
Факс: +7(383) 328 00 71
E-mail: Kroshin.Fyodor@huawei.com
URL: e.huawei.com/ru

Huawei is a leading global ICT solutions provider. Through our dedication to customer-centric innovation and strong partnerships, we have established end-to-end capabilities and strengths across the carrier networks, enterprise, consumer, and cloud computing fields. We are committed to creating maximum value for telecom carriers, enterprises and consumers by providing competitive ICT solutions and services. Our products and solutions have been deployed in over 170 countries, serving more than one third of the world's population.

By leveraging our strong R&D capabilities and comprehensive technical expertise, Huawei's strategy in the enterprise domain focuses on close cooperation and integration with partners to deliver a wide range of highly efficient customer-centric ICT solutions and services that are based on a deep understanding of customer needs. In line with our strategy, we offer a broad portfolio of innovative ICT solutions that cater to global vertical industry and enterprise customers across government and public sector, finance, transportation, energy, large enterprises and small and midsize enterprises (SMEs). Our portfolio covers enterprise networking, unified communications & collaboration (UC&C), cloud computing & data center, enterprise wireless, network energy and infrastructure services.

HUAWEI Technologies Co Ltd., Russia
Siberia office:
630112, Russia, Novosibirsk, Frunze Str., 242
Business Center "New Height"
Tel.: +7(383) 328 00 70,
Fax: +7(383) 328 00 71
Email: Kroshin.fyodor@huawei.com
URL: e.huawei.com/ru



Корпорация Intel

Корпорация Intel была основана в 1968 году Робертом Нойсом и Гордоном Муром. На протяжении 50 лет Intel создает инновационные технологии, открывающие новые возможности для людей.

Корпорация Intel является мировым лидером в области микроэлектроники и информационных технологий. Intel создает технологии для умного мира эпохи больших данных. Основное внимание корпорация уделяет созданию интеллектуальных решений для умного мира, от устройств Интернета вещей и пользовательских ПК до коммуникационной инфраструктуры, технологий для центров обработки данных и суперкомпьютеров.

Штаб-квартира корпорации расположена в г. Санта-Клара, шт. Калифорния. Общий штат Intel насчитывает более 100 тыс. сотрудников в более, чем 60 странах по всему миру. Главным исполнительным директором корпорации является Роберт Свон (Robert Swan).

Intel в России

Первое представительство Intel в России было открыто в 1991 году в Москве. Сегодня в российских офисах Intel в Москве и Нижнем Новгороде работают более 800 человек.

В московском офисе компании представлены отделы маркетинга и развития бизнеса, группы по разработке программного обеспечения, юридический отдел.

В НИОКР центре Intel в Нижнем Новгороде создаются новые и инновационные продукты для разработки ПО. Сегодня он является одним из крупнейших центров исследований и разработок Intel в Европе. Более **700 специалистов и инженеров** разрабатывают программные инструменты и приложения для архитектур Intel. В Нижнем Новгороде также размещаются различные группы поддержки бизнеса (например, административно-хозяйственная часть, финансовый отдел, отдел ИТ, отдел кадров).

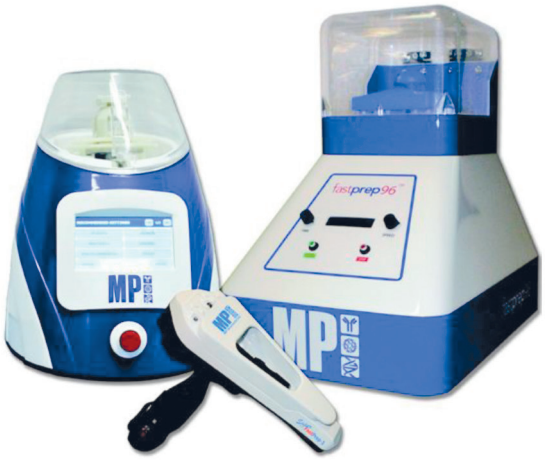
Центр исследований и разработок Intel в Нижнем Новгороде

Нижегородский офис Intel был является центром экспертизы корпорации в области высокопроизводительных вычислений, разработки программного обеспечения в области численных методов и беспроводной связи.

✉ MP Biomedicals
(ООО «МПБА диагностика»)
Адрес: 109147, г. Москва, ул. Марксистская, д. 3,
стр. 2, оф. 2.1.20/2
☎ Тел./факс:
+7(495)604-13-44
E-mail rus@mpbio.com

WEB mpbio.com; mpbio.ru

Компания ООО «МПБА диагностика» является дочерней компанией MP Biomedicals, ранее известной как ICN Biomedicals, основанной в 1959 году, признанного лидера в области производства широкого спектра химических реактивов, оборудования для пробоподготовки (система для гомогенизации FastPrep) и наборов реагентов. Каталог продукции компании MP Biomedicals включает более 55000 наименований высококачественных продуктов для проведения биохимических исследований, фармацевтического и биотехнологического производства, для различных отраслей иммунологии и генетики.



ООО «Рош Диагностика Рус» – официальный импортер продукции Roche в России и лицензиат компании F.Hoffmann–La Roche Ltd.

Roche Sequencing Solutions, подразделение Roche, ориентированное на решения для NGS, а в частности на пробоподготовку к NGS, предлагает:

-**Наборы KAPA Biosystems для приготовления библиотек ДНК** (включают баркодированные адаптеры, частицы для очистки, наборы для оценки концентраций ДНК и библиотек методом ПЦР в реальном времени).

-**Наборы для направленного отбора генов перед NGS:**

NimbleGen SeqCap EZ –гибридизационное обогащение панелей генов, экзомов, транскриптомов и метиломов;

HEAT-seq - амплификационное обогащение панелей генов, в том числе и панелей онкогенов;

AVENIO –гибридизационное обогащение панелей онкогенов из внеклеточной опухолевой ДНК и анализ данных.

ООО «Рош Диагностика Рус» предлагает комплексные решения, включающие в себя не только оборудование и реагенты, но и технический сервис, обучение персонала и постоянную методическую поддержку.

Unlock the Potential of Every Sample

ООО «Рош Диагностика Рус
115114, Россия, Москва
Ул. Летниковская, д. 2, стр. 2
Тел:+7 (495) 229-69-99
Бизнес-центр «Вивальди Плаза»
ras.russia@roche.com
sequencing.roche.com

BioLine
группа компаний

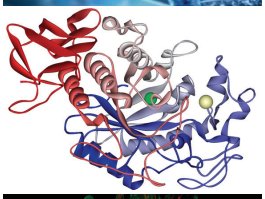
PerkinElmer For the Better
Leica MICROSYSTEMS
BioTek
BD

ООО «БиоЛайн» является официальным дистрибьютором компаний **BD Biosciences, Leica Microsystems, PerkinElmer, BioTek,** ведущих мировых производителей приборов и реагентов для биомедицинских исследований.

Благодаря уникальному портфолио продукции и опыту наших специалистов мы выполняем поставки и внедрение комплексных решений для разнообразных задач в области молекулярной и клеточной биологии.

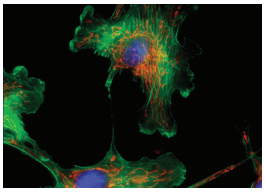
Молекулярно-генетические исследования

- Системы для выделения и молекулярного анализа одиночных клеток Becton Dickinson
- Станции для выделения ДНК, оборудование PerkinElmer для подготовки и контроля библиотек для NGS
- Наборы Nextflex для подготовки библиотек NGS PerkinElmer: полногеномное и таргетное секвенирование, транскриптомика, эпигенетика, метагеномика



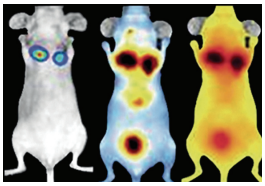
Протеомные исследования

- Передовые оптические технологии компании BioTek Instruments для биохимических исследований, идентификации и количественной оценки аналитов, исследования взаимодействия биомолекул
- Реагенты и расходные материалы PerkinElmer для протеомных исследований



Клеточные исследования

- Системы для проточной цитометрии и сортировки клеток компании BD Biosciences
- Оптическая визуализация клеток для моделирования процессов в клеточных культурах и на 3D сфероидах: решения PerkinElmer и BioTek Instruments
- Системы для конфокальной микроскопии Leica Microsystems



Исследования на животных

- Приборы для оптической визуализации ***in vivo*** Spectrum и Lumina, системы для КТ и ПЭТ компании PerkinElmer
- Оборудование для исследований на животных Leica Biosystems

Официальные дистрибьюторы BD Biosciences, Leica Microsystems, PerkinElmer, BioTek в России — компания «БиоЛайн»

ООО «БиоЛайн»
Россия, 197101, Санкт-Петербург
Пинский пер., д. 3, Лит. А
тел.: +7 (812) 320 49 49
факс: +7 (812) 320 49 40
e-mail: main@bioline.ru
www.bioline.ru

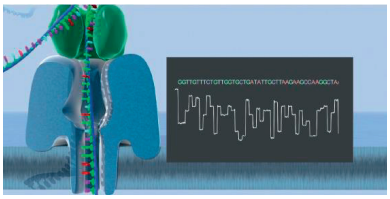
Москва, тел.: +7 (800) 555 49 40
Новосибирск, тел.: +7 (383) 227 09 63
Екатеринбург, тел.: +7 (343) 287 32 49
Владивосток, тел.: +7 (423) 201 18 08
Нижний Новгород, тел.: +7 (831) 278 61 47
Ростов-на-Дону, тел.: +7 (863) 268 99 32
Казань, тел.: +7 (843) 570 66 88
Самара: +7 (846) 246-06-54

Единый бесплатный номер сервисной службы для всех регионов России:

8 800 333 00 49

Компания Диаэм – крупнейший поставщик современного лабораторного оборудования на Российском рынке. Каталог компании насчитывает более 500 000 наименований приборов, реагентов и расходных материалов для медицинских и научно-исследовательских лабораторий.
В каталоге компании представлена продукция ведущих мировых производителей, как: Abcam, Applied Biosystems, Binder, Bio-Rad, Corning, Eppendorf, Illumina, Ion Torrent, Lexogen, Oxford Nanopore Technologies, Panasonic (Sanyo), Sage Sciences, Sigma-Aldrich, Thermo Fisher Scientific, Qiagen:

- Наборы для подготовки библиотек, для высокопроизводительного секвенирования NGS, для исследовательских работ и, в онкологии, репродуктивной медицине, в изучении наследственных заболеваний, реагенты и наборы для капиллярного секвенирования.
- Секвенаторы капиллярные и высокопроизводительные NGS, оборудование для анализа качества НК для NGS, роботизированные станции для подготовки библиотек и секвенирования.
- Все для ПЦР, реагенты,наборы, пластик, амплификаторы.
- Нанопоровые секвенаторы Oxford Nanopore Technologies, наборы для секвенирования ДНК и РНК.


Секвенирование теперь доступно каждому!

Диаэм сегодня представляет продукцию Oxford Nanopore Technologies – это секвенаторы третьего поколения – MinION, GridION, PromethION.
Технология секвенирования Oxford Nanopore Technologies позволяет делать прямое прочтение цепей ДНК или РНК в режиме онлайн, длина рида ограничена только длиной фрагмента, а портативность оборудования и быстрая подготовка библиотек дает возможность секвенировать даже в полевых условиях с минимальными требованиями к генетической лаборатории. С Oxford Nanopore Technologies секвенировать теперь может каждый, даже тот, кто ранее и не задумывался о секвенировании - это просто и доступно.
Секвенирование третьего поколения не заменяет и не отменяет применение капиллярных секвенаторов по Сэнгеру или платформ NGS второго поколения, наоборот, сочетание трех поколений генетического анализа открывает новые возможности получения ранее неизвестных данных.
Специалисты Диаэм прошли обучение в Oxford Nanopore Technologies, осуществляют профессиональное консультирование и техническую поддержку, помогут спланировать эксперимент и подобрать необходимые наборы реагентов для решения конкретной задачи независимо от бюджета лаборатории.
Обращайтесь в любой из наших офисов или пишите на sales@dia-m.ru

ООО «Диаэм» www.dia-m.ru

Москва	Новосибирск	Казань	Санкт-Петербург	Ростов-на-Дону	Пермь	Воронеж
ул. Магаданская, 7/3	пр. Акад. Лаврентьева, 6/1	ул. Парижской Коммуны, д. 6	ул. Профессора Попова, 23	пер. Семашко, 114	Представитель в УФО	тел./факс: (473) 232-4412
тел./факс: (495) 747-0508	тел./факс: (383) 328-0048	тел/факс: (843) 210-2080	тел./факс: (812) 372-6040	тел/факс: (863) 250-0006	тел./факс: (342) 202-2239	voronezh@dia-m.ru
sales@dia-m.ru	nsk@dia-m.ru	kazan@dia-m.ru	spb@dia-m.ru	rnd@dia-m.ru	perm@dia-m.ru	

Компания АЛЬБИОГЕН — официальный дистрибьютор illumina и Lucigen

Компания ООО «АЛЬБИОГЕН» с 2015 года является эксклюзивным (единственным) официальным торговым представителем и дистрибьютором компании illumina на территории Российской Федерации, Республики Беларусь, Республики Казахстан и Республики Узбекистан.

Нашей задачей является обеспечение полного доступа клиентов к передовым технологиям и сервисам illumina, включая современные системы NGS и анализа ДНК-биочипов, программное обеспечение для биоинформатики и весь спектр реактивов.

ООО «АЛЬБИОГЕН» предоставляет полный комплекс услуг, связанных с продажей, технической поддержкой и сервисным (гарантийным и постгарантийным) обслуживанием продукции компании Illumina, а также обучением пользователей работе на данном оборудовании.

Инновационная и стремительно развивающаяся компания illumina Inc., являющаяся мировым лидером в области геномных технологий, заключила соглашение с компанией АЛЬБИОГЕН, специализирующейся на поставках оборудования и расходных материалов для секвенирования нового поколения (NGS) и анализа на ДНК-биочипах.

Новейшие продукты компании illumina, создаваемые совместно с ведущими мировыми учеными, позволяют изучать геном на очень глубоком уровне и дают возможность для новаторских достижений в науке, медицине, сельском хозяйстве и потребительской геномике. Более 90% научных статей, связанных с технологиями секвенирования нового поколения, сделаны при помощи оборудования Illumina.

Сотрудничество с компанией АЛЬБИОГЕН направлено на то, чтобы сделать технологии NGS и анализа ДНК-биочипов более доступными на территории Российской Федерации и в странах СНГ.

Компания АЛЬБИОГЕН использует свой обширный опыт в области продаж и продвижения продукции, знания передовых технологий и сеть региональных представителей для обеспечения быстрой, эффективной и бесперебойной работы лабораторий клиентов illumina.

Компания АЛЬБИОГЕН также является официальным дистрибьютором компании Lucigen, основными продуктами которой являются ферменты и реагенты для секвенирования нового пколения и молекулярной диагностики.

Компания СкайДжин предлагает к поставке со склада в Москве и под заказ наборы реагентов, оборудование, расходные материалы, реактивы, а также специализируется на сервисном обслуживании и поверке дозаторов, лабораторных весов различных производителей. Мы предлагаем гибкие условия работы и очень большой ассортимент продукции.

Поставляемая нашей компанией продукция широко используется в научно-исследовательских лабораториях и R&D центрах, лабораториях секвенирования, при решении практически любых молекулярно-биологических задач.

Большая часть производителей в нашем портфолио - это прямые, эксклюзивные поставки. Мы являемся первым звеном в поставках для таких компаний как New England Biolabs, Agilent Technologies, Oxford Nanopore Technologies, QIAGEN, 10x Genomics, NIMAGEN, Integrated DNA Technologies, Thermo Fisher Scientific, SIGMA-ALDRICH, BioSan, Gilson.

К флагманским продуктам наших линеек относятся:

- Набор для пробоподготовки образцов от New England Biolabs ULTRA II FS с интегрированной системой фрагментации и другие наборы серии ULTRA для образцов ДНК, РНК и микроРНК;
- Digital NGS: готовые панели и наборы для обогащения на основе ПЦР от QIAGEN с мономолекулярным баркодированием;
- Специализированные наборы для работы с микроРНК и анализа экспрессии от QIAGEN-Exiqon;
- Нанопоровые секвенаторы третьего поколения: портативный секвенатор MinION, высокопроизводительный секвенатор GridION;
- Уникальная система Chromium производства 10x Genomics для автоматической пробоподготовки геномов и транскриптомов единичных клеток.

За дополнительной информацией о производителях, товарах, ценах и условиях поставки обращайтесь к нашим квалифицированным специалистам.

Будем рады ответить на Ваши вопросы и помочь выбрать качественное и недорогое решение для Ваших задач!

ООО «СкайДжин»
Адрес: 115093, Москва,
ул. Люсиновская, д. 36, стр. 1
Тел: 8 (495) 215 02 22
info@skygen.com
www.skygen.com

ХИМЭКСПЕРТ

Агентство Химэксперт

Информация о компании:

Компания Химэксперт существует 16 лет и давно зарекомендовала себя, как надежный поставщик приборов, реактивов и расходных материалов для молекулярной биологии. Мы собрали для своих клиентов самые интересные и перспективные бренды, большинство из которых в России можно приобрести только у нас.

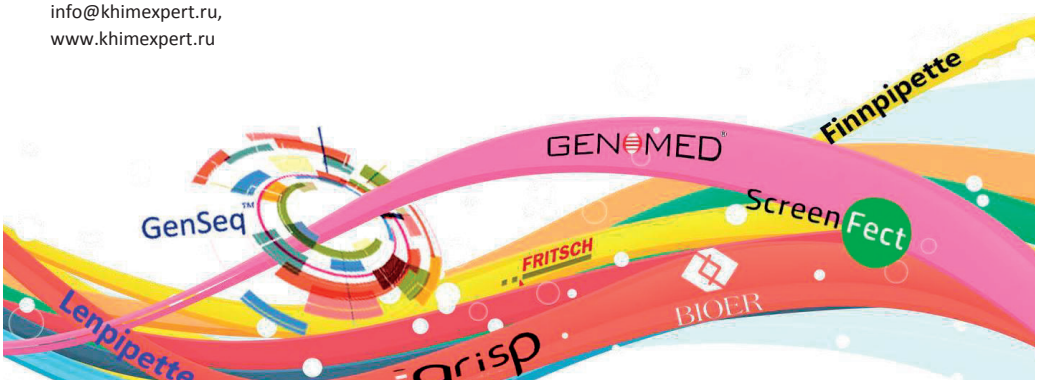
Химэксперт предлагает оборудование для анализа ДНК и РНК, в том числе и методами NGS, фундаментальных протеомных и цитологических исследований, фармацевтики и биотехнологий, прикладного тестирования, включая идентификацию личности и установление родства в криминалистике и судебно-медицинской экспертизе.

Наши клиенты выбирают Химэксперт потому что:

- Химэксперт всегда находит самые прогрессивные решения в области Life Sciences. Наша компания постоянно расширяет свое портфолио и в курсе последних веяний в области молекулярной биологии
- Химэксперт осуществляет полную техническую и методическую поддержку наших клиентов: обратившись к нам, вы получаете помощь квалифицированных сотрудников в подборе оборудования и реагентов под поставленные задачи и их последующем использовании
- Химэксперт стремится идти навстречу заказчикам и осуществлять быстрые поставки, так как скорость и четкость исполнения заказов очень важна.

Обратившись к нам, вы можете быть уверены в будущем своего эксперимента.
Начните сотрудничество с компанией Химэксперт и убедитесь в этом на своем опыте!

ООО «Агентство Химэксперт» 125009, г. Москва, Страстной б-р, д. 4, оф. 101
Тел: +7 (495) 629 28 69, 650 36 66
info@khimexpert.ru,
www.khimexpert.ru



www.genexplain.com

The geneXplain GmbH is glad to welcome you at the BGRS/SB'2018 conference and is proud to introduce you the following software and database solutions for the needs of bioinformatics, systems biology and systems medicine:

geneXplain platform – is a high-performance tool for multi-omics data analysis, which allows identification of new therapeutic targets and biomarkers. A unique feature of the geneXplain platform is its Upstream Analysis. You can register and immediately receive access to a free account.

TRANSFAC database – is a unique collection of transcription factors, their experimentally validated binding sites (TFBS) and a widely known library of positional weight matrices (PWMs). The database has its own integrated methods for TFBS search. It can also be used as an integral part of the geneXplain platform. TRANSFAC is available online or can be downloaded as a set of flat files.

TRANSPATH database – is one of the biggest and most famous collections of signaling and metabolic pathways, which counts over 489000 reactions. The database can be applied for master-regulators search within the geneXplain platform. TRANSPATH is also available online in one package with HumanPSD database or can be downloaded as a set of flat files.

HumanPSD database – is a collection of genes, proteins and micro-RNAs, which includes information about disease biomarkers and clinical trials for various diseases. Besides the detailed biomarkers data, the database contains information about drugs.

BRENDA database – is a comprehensive enzyme and enzyme-ligand information system. Its manually derived core contains over 3 million data points about 77,000 enzymes annotated from 135,000 literature references.

PASS – is a software tool for evaluating the general biological potential of organic compounds based on their structural formula. This program predicts main and side pharmacological effects, molecular mechanisms of action, specific toxicities, and antitargets, actions associated with the metabolism and transport of pharmaceutical substances and their influence on gene expression.

PharmaExpert – is a software tool for analysis of the biological activity spectra of substances predicted by PASS and selecting compounds with the desirable set of biological activity, for analyzing the relationships between biological activities, drug-drug interactions and for multiple targeting of chemical compounds.

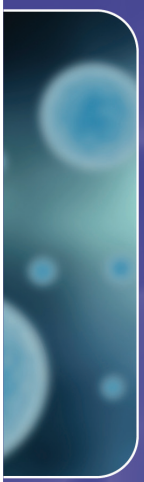
GUSAR – is a software tool for analysis of quantitative structure-activity/structure-property relationships (QSAR/QSPR) based on the structural formulas of the compounds and data on their activity/property, and for prediction of activity/property for new compounds. GUSAR can be easily applied to different routine QSAR/QSPR tasks, for building multiple models, and for prediction of the different quantitative values simultaneously.

If you got interested in any of the products, provided by GeneXplain, or you have any questions, please contact us by e-mail info@genexplain.com. We will be glad to help you!

ISSN 1386-6338

In Silico Biology

Journal of Biological Systems Modeling and Multi-Scale Simulation



```
function [ t, x ] = firstReactionMethod( ...
    stoich_matrix, propensity_fcn, tspan, x0,...
    rate_params, output_fcn, max_out)

if ~exist('rate_params', 'var')
    rate_params = [];
end

num_rxns = size(stoich_matrix, 1);
num_species = size(stoich_matrix, 2);

%Simulation loop
while t(rxnCount) <= max(span)
    % Step 1: calculate propensities
    a  = propensity_fcn(X(rxnCount,:), rate_params);
    % Step 2: identify the reaction that will occur
    r = rand(1,num_rxns);
    taus = -log(r)./a;
    [tau, mu] = min(taus);
    % Update time and execute reaction mu
    rxnCount = rxnCount + 1;
    T(rxnCount)    = T(rxnCount-1)    + tau;
    X(rxnCount,:) = X(rxnCount-1,:) + stoich_matrix(mu,:);

    if rxnCount > max_out
        warning('SSA:ExceededCapacity','');
        return;
    end
end

% Simulation completed
t = T(1:rxnCount-1);
x = X(1:rxnCount-1,:);
end
```

IOS Press

Научное издание

BIOINFORMATICS OF GENOME REGULATION AND STRUCTURE\SYSTEMS BIOLOGY

The Eleventh International Conference

Abstracts

Printed without editing

БИОИНФОРМАТИКА РЕГУЛЯЦИИ И СТРУКТУРЫ ГЕНОМА\СИСТЕМНАЯ БИОЛОГИЯ

Одиннадцатая международная конференция

Тезисы докладов

Публикуется в авторской редакции

Выпуск подготовлен информационно-издательским отделом ИЦиГ СО РАН

Подписано к печати 01.08.2018. Формат 70 × 108 $^{1}/_{16}$. Усл. печ. л. 24.0.
Тираж 100 экз. Заказ № 182

Федеральный исследовательский центр
«Институт цитологии и генетики Сибирского отделения Российской академии наук»
630090, Новосибирск, проспект Академика Лаврентьева, 10

Отпечатано в типографии ФГУП «Издательство СО РАН»
630090, Новосибирск, Морской проспект, 2