BGRS\SB-2018

The 11th International Conference

Integrative Bioinformatics and Systems Biology, WIBSB-2018

First Sino-Russian Workshop

WIBSB-2018
NOVOSIBIRSK, RUSSIA
22–23 AUGUST, 2018

CONF.BIONET.NSC.RU/SRW2018

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

INTEGRATIVE BIOINFORMATICS AND SYSTEMS BIOLOGY (WIBSB-2018)

First Sino-Russian Workshop

Abstracts

22–23 August, 2018
Novosibirsk, Russia

Novosibirsk
ICG SB RAS
2018

УДК 575
I73

Integrative Bioinformatics and Systems Biology (WIBSB-2018) : First Sino-Russian Workshop (22–23 Aug. 2018, Novosibirsk, Russia); Abstracts / Novosibirsk State University; Institute of Cytology and Genetics, Siberian Branch of Russian Academy of Sciences. – Novosibirsk: ICG SB RAS, 2018. – 78 pp. – ISBN 978-5-91291-042-5.

Contacts

The Federal Research Center
Institute of Cytology and Genetics
The Siberian Branch of the Russian Academy of Sciences
Russian Federation
630090 Novosibirsk, Lavrentyeva, 10
Tel: +7(383) 363-49-80
Fax: +7(383) 333-12-78

ISBN 978-5-91291-042-5

Organizers

- Novosibirsk State University (NSU)
- Institute of Cytology and Genetics SB RAS (ICG SB RAS)
- The Federal Agency for Scientific Organizations (FASO Russia)

Sponsors

Publication of the Proceedings is supported by the Russian Foundation of Basic Research, Grant No. 18-04-20044

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

Publisher

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

DOI 10.18699/WIBSB-2018-01

Brief introduction of bioinformatics education in China

Q. Dai[1, 4], W. Liu[2,4], M. Chen[3, 4]*
[1] *School of Life Science, Zhejiang Sci-Tech University, P.R.China*
[2] *College of Mathematics, Physics and Electronic Information Engineering, Wenzhou University, P.R.China*
[3] *College of Life Sciences, Zhejiang University, P.R.China*
[4] *Bioinformatics Society of Zhejiang Province, P.R.China*
** e-mail: mchen@zju.edu.cn*

Key words: education, bioinformatics, international cooperation

As an interdisciplinary field of science and technology, bioinformatics has become an important part of many areas of biology. Bioinformatics in China has grown significantly in the past decade despite a delayed and patchy start at the end of the 1980s by a few pioneer scientists from other disciplines, most noticeably physics and mathematics, where China's traditional strength has been. In the late 1990s and early 2000s, rapid expansion of this field was fueled by the Internet boom and genomics boom worldwide.

In China, more and more researchers are converted into bioinformatics from other disciplines. Meanwhile, with the rapid development of economy and attraction of excellency policy, more and more scientists return China after their formal training overseas and join bioinformatics research. There are more than 30 bioinformatics related societies/organizations are set up to promote the bioinformatics development. More and more scientists from other countries, especially European countries, are drawn to work in China by the improved research and funding environment. International cooperation research institutes/centers are established, e. g. the Max Planck–Chinese Academy of Science Partners Institute in Computational Biology in Shanghai, which employs a number of European scientists and plays key roles in facilitating international collaborations.

Today, more than 30 universities offer undergraduate majors in bioinformatics, and nearly 20 universities offer bioinformatics graduate programs at the PhD and Master's levels. The number of bioinformatics students is increasing every year. According to a market estimation, average monthly salary for bioinformatics graduates with bachelor degree is about 8,000 RMB Yuan. The country has educated more than 4,000 students, but this is far from meeting the needs of more than 20 thousand of bioinformatics and related companies in China.

In Zhejiang province, Zhejiang University is offering bioinformatics degree programs at the PhD, Master's, and Bachelor's levels. Many other universities have carried out the bioinformatics course, such as Zhejiang Sci-Tech University, Zhejiang University of Technology, Hangzhou Normal University, Zhejiang Chinese Medical University, Wenzhou University, etc. In addition to formal courses, most of the universities offer bioinformatics training program to bench biologists.

The Bioinformatics Society of Zhejiang Province was established in 2013. It is a registered NGO that meets all the legal and financial requirements given by the government. The goal of the society is to provide a platform for bioinformatics scientists, researchers, young students and publics, to create a network between institutions and companies, and to encourage interdisciplinary communication and in-depth discussion. We welcome members from board fields and support members and other persons interested in obtaining advanced knowledge in the field, present current progress and

expand the general knowledge and applications of bioinformatics. The society will work with foreign societies, contribute to the further education of its members and organize seminars, conferences and workshops. Our next annual meeting will be held this October in Lishui city, Zhejiang.

We work on cooperation in bioinformatics education with Russian universities. We had joint Sino-Russian grant project RFBR-NSCFC (finished in 2016), applied for new research grants in the frames of BRICS initiative in 2018. Recent review papers and special journal issue publications were organized by the international conferences in Russia [1, 2].

Acknowledgements: The participation of MC at BGRS-2018 multiconference has been supported by Russian Ministry of Science project 28.12487.2018/12.1.

References

1. Orlov Y.L., Fernandez-Masso J.R., Chen M., Baranova A.V. (2018) Medical genomics at Belyaev Conference – 2017. BMC Medical Genomics.11(Suppl 1):11. DOI 10.1186/s12920-018-0324-3.
2. Chen M., Harrison A., Shanahan H., Orlov Y. (2017) Biological Big Bytes: Integrative Analysis of Large Biological Datasets. J Integr Bioinform. 14(3). DOI 10.1515/jib-2017-0052.

DOI 10.18699/WIBSB-2018-02

Statistical approaches for analysis of mapping quality for single-cell sequencing data

I.I. Abnizova[1, 2]*, R. te Boekhorst[3], N. Beka[3], F.M. Naumenko[4], A.V. Tsukanov[4, 5], Y.L. Orlov[4, 5, 6]*

[1] *Wellcome Trust Sanger Institute, Cambridge, UK*
[2] *Babraham Institute, Cambridge, UK*
[3] *University of Hertfordshire, Hertfordshire, UK*
[4] *Novosibirsk State University, Novosibirsk, Russia*
[5] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[6] *Institute of Marine Biology Researches of the RAS, Sevastopol, Russia*
* *e-mail: irina_abnizova@yahoo.co.uk;orlov@bionet.nsc.ru*

Key words: Next-generation sequencing, DNA alignment, read density distribution.

Motivation and Aim: Bioinformatics analysis is essential in providing biological insights for single-cell experiments, such as detecting variants, quantifying gene expression, and subpopulation detection. However, conventional tools developed for bulk-cell genomics cannot be directly applied to single-cell sequencing data [1].

Methods and Algorithms: This low coverage characteristic of single-cell sequencing data has posed difficulties in the variant calling procedure. Most bioinformatics tools employ sequence read density to call variants [2]. Single nucleotide polymorphisms and small insertions/deletions with low read support are excluded in conventional bioinformatics tools. In genome assemblies, the low coverage and heterogeneity of single-cell sequencing data also bring substantial disadvantages, leading to truncated sequences with high numbers of sequencing artefacts. Recently, single-cell assemblers such as SPAdes and IDBA-UD have been specifically developed to overcome the challenge of amplification artefacts in single-cell sequencing and generate more precise single-cell genomic assemblies. Common gene expression metrics such as Fragments Per Kilobase Million/Reads Per Kilobase Million (FPKM/RPKM) do not address these 3′-end biases and thus have a limited application for scRNA sequencing. Using own scripts we investigated chromosome mapping quality and possible artefacts [3].

Results and conclusion: We applied or approaches to study Differentially Chromatin accessed regions (DARs) and Diff Methylated Regions (DMR). The generation of artificial data by mapping of generated reads to a reference genome is justified from the point of view of reducing the benchmarking time. We will review current state of art of mapping programs in this research area.

Acknowledgements: The research has been supported by RFBR. Computing done at Siberian Supercomputer center SB RAS was supported by budget project No. 0324-2018-0017.

References

1. Yuan Y. et al. (2018) Single-Cell Genomic Analysis in Plants. Genes (Basel). 9(1):E50.
2. Valihrach L. et al. (2018) Platforms for Single-Cell Collection and Analysis. Int J Mol Sci. 19(3). pii: E807.
3. Naumenko F.M. et al. (2018) Novel read density distribution score shows possible aligner artefacts, when mapping a single chromosome. BMC Genomics. 19(Suppl 3):92.

DOI 10.18699/WIBSB-2018-03

miRNA interaction with 5'UTR, CDS, 3'UTR mRNA candidate genes of breast cancer subtypes

D. Aisina*, R. Niyazova, Sh. Atambayeva, A. Ivashchenko
al-Farabi Kazakh National University, Almaty, Kazakhstan
** e-mail: dana.aisina03@gmail.com*

Key words: breast cancer, subtypes, genes, miRNAs

Breast cancer subtypes are distinguished by a set of candidate genes involved in the development of this disease. The expression of many genes is regulated by binding of their mRNAs with miRNAs. It is required to identify which candidate genes can interact with miRNAs. The MirTarget program defines the following features of binding: start of the initiation of miRNA binding to mRNAs; localization of miRNA binding sites in 5′UTRs, CDSs and 3′UTRs; free energy of binding; schemes of nucleotide interactions between miRNAs and mRNAs. mRNAs of many genes have miRNA binding sites with overlapping nucleotide sequences (clusters) located in 5′UTR, CDS, 3′UTR. There are cluster of three sites of different miRNAs in the 5′UTR mRNA *EPOR*, *MAZ* and *NISCH* candidate genes (her2 subtype), cluster of 11 sites in the CDS mRNA *MAZ* gene, clusters of three sites and 17 sites in the 3′UTR mRNA *BRCA2* gene and *CDK6* genes, respectively. Candidate genes of the triple-negative subtype are targets: in the 5′UTR mRNA CBL gene are 11 sites, mRNA *MMP2* gene – five sites, mRNA *RAB5A* gene are two cluster each of three sites, in the 3′UTR mRNA *SFN* gene – 18 sites. Candidate genes of luminal A and B subtypes are targets: in the 5′UTR mRNA *FOXA1* gene are 19 sites, mRNA *HMGA2* gene – 12 sites, mRNA *TGFB1* gene – two clusters of three and four sites. There are clusters of four sites and three sites in the CDS mRNA *ITGB1* and *SOX4* genes, respectively; clusters of three sites, four sites and five sites in the 3′UTR mRNA *SMAD3*, *SOX4* and *GFB1* genes, respectively. The organization of binding sites into clusters several times reduces the proportion of binding sites in nucleotides in 5′UTR, CDS and 3′UTR. Based on the results, associations of miRNAs and mRNAs candidate genes are recommended for developing methods of breast cancer subtypes diagnostics.

DOI 10.18699/WIBSB-2018-04

De novo sequencing, assembly and annotation of *Armillaria borealis* genome

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, TX 77843-2138, USA*
** e-mail: vfedotova@sfu-kras.ru*

Key words: genome assembly, functional annotation, fungal pathogenicity

Motivation and Aim: Massive forest decline as a result of negative anthropogenic and climatic effects, often aggravated by pests, fungi and other phytopathogens, has been observed almost everywhere. Environmental changes can weaken trees and make fungi more destructive. Forest conservation has become a serious issue, since the scale of tree death caused by phytopathogenic fungus is enormous. *Armillaria borealis* (Marxm. & Korhonen) is a fungi from the *Physalacriaceae* family 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 sequence *de novo*, assemble and characterize the genome of *Armillaria borealis* and to obtain data that can be used to identify the fungi virulence factors, such as target genes. We also intend to provide population genetics with reference material to study forest populations of *Armillaria spp*.
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 SibFU. To evaluate the completeness of the gene set and assembly, BUSCO was performed using *Basidiomycota* odb9 base. Coding regions were identified in the genome using Exonerate; the EVidenceModeler and Augustus software were used to predict genes. Finally, a functional annotation was done using predictions, protein and transcript alignments, and assignments based on PFAM, InterPro and GO ontology.
Results: The *A. borealis* genome assembly contained ~69 Mbp and was comparable with 60 and 84 Mbp for the *A. ostoyae* and *A. gallica* genomes, respectively. The N50 for contigs equaled 15,659 bp. BUSCO results showed that 94.8 % of reference genes were captured as complete single-copy BUSCOs. Functional annotation revealed 6,703 protein coding genes, which was comparable with 7,797 and 8,261 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 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 work was supported by research grant No. 14.Y26.31.0004 from the Government of the Russian Federation.

DOI 10.18699/WIBSB-2018-05

Developing the protein-concentrating nanofluidic chips for early diagnostics of neurodegenerative disorders

T.G. Amstislavskaya[1, 2, 3]*, Y.-L. Yang[5], C.-P. Jen[4]
[1] *Scientific Research Institute of Physiology and Basic Medicine, Novosibirsk, Russia*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[3] *Novosibirsk State University, Novosibirsk, Russia*
[4] *Department of Mechanical Engineering, National Chung Cheng University, Chia-Yi, Taiwan ROC*
[5] *Department of Biochemical Science and Technology, National Chia-Yi University, Taiwan ROC*
** e-mail: amstislavskayatg@physiol.ru*

Key words: neurophysiology, cognitive disorders, Alzheimer's disease, nanofluidic chips, animal model, international collaboration

Motivation and Aim: Mainly due to the worldwide aging problem, neurodegenerative disorders have accounted for more than 70 % of all dementia and has obviously become a serious health problem to be coped with. Here we present a bilateral collaborative project in neurophysiology supported by SB RAS (Russia) and MOST (Taiwan) and entitled "Developing the protein-concentrating nanofluidic chips for early diagnostics of neurodegenerative disorder". The goal of the project is to develop electrokinetic protein preconcentration in nanofluidic channels which could be used to detect Aβ, Aβ-like proteins and Lcn2 at low concentrations in plasma for early diagnosis of disease.
Methods and Algorithms: Registration of certain proteins at very low concentration in peripheral plasma is expected to be a promising diagnostic approach at early stages of neurodegenerative disorders (such as Alzheimer's disease or Parkinson's disease).
Results: We developed electrokinetic concentration of proteins in nanofluidic channels which could be used to detect low concentrations of molecular biomarkers (Aβ, Aβ-like proteins and Lcn2) in plasma. The chips achieve the enrichment of proteins basing on the exclusion-enrichment effect in a nanofluidic channel and are expected to preconcentrate a sample up to 10^3–10^6-fold that would allow using routine immunodetection methods for determining biomarker levels.
Conclusion: The proposed device is easy to operate and compact. We plan to use highly sensitive immunoassay technology to detect different forms of Aβ, Aβ-like proteins and their ratio and Lcn2 levels in different tissues in several models of neurodegenerative damage. Research groups of highly qualified biologists and physicists from SB RAS and Taiwan were involved into the project and collaborated effectively [1, 2].
Acknowledgements: The research has been supported in part by MOST-SB RAS Joint Research Projects in 2018-2020, RFBR and Computing was done at Siberian Supercomputer center SB RAS (SSCC) with support by ICG SB RAS budget project (No. 0324-2018-0017).

References

1. Chen Y.H., Wu H.F., Amstislavskaya T.G., Li C.Y., Jen C.P. (2016) A simple electrokinetic protein preconcentrator utilizing nano-interstices. Biomicrofluidics. 10(2):024121.
2. Wu Y.J., Chen Y.L., Tang T.H., Ng M.C., Amstislavskaya T.G., Tikhonova M.A., Yang Y.L., Lu K.T. (2017) Unilateral stimulation of the lateral division of the dorsal telencephalon induces synaptic plasticity in the bilateral medial division of zebrafish. Sci Rep. 7(1):9096.

DOI 10.18699/WIBSB-2018-06

Systems biology approaches for analysis of dementia with Lewy bodies in mouse models

T.G. Amstislavskaya[1, 4]*, M.A. Tikhonova[1, 4], H. Bai[3], Y.L. Orlov[2, 4], M. Chen[3, 5]
[1] *Scientific Research Institute of Physiology and Basic Medicine, Novosibirsk, Russia*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[3] *Inner Mongolia University for the Nationalities, China*
[4] *Novosibirsk State University, Novosibirsk, Russia*
[5] *Zhejiang University, Hangzhou, China*
* *e-mail: amstislavskayatg@physiol.ru*

Key words: neurodegenerative diseases, dementia with Lewy bodies, mouse models, systems biology

Motivation and Aim: We present joint research project dedicated to the fundamental problem of neurophysiology and metabolism related to the elucidation of the physiological and molecular mechanisms of the disorders caused by aging and neurodegenerative processes. The estimated economic burden caused by dementia related problems was US$ 18 billion in 2015 and would rise above $1 trillion by 2018. The common causes of dementia in elderly are the Alzheimer's disease, vascular dementia and dementia with Lewy bodies (DLB). Our aim was to identify the opportunities to correct these neurological pathologies and their related dysfunctions using experimental molecular-biological and bioinformatics approaches.
Methods and Algorithms: The dementia has complex genetic background and shares high comorbidity with metabolic diseases, diabetes, bipolar disorder, and alcohol addiction, thus presenting socially important problem in Asian populations. There is evidence of multi-factorial nature of these diseases. Based on population genetics analysis and network models we select target genes regulating the diseases networks. We validate gene targets experimentally as specific candidates for combined treatment. To understand the genetic etiology of major neurodegenerative processes we model dementia with Lewy bodies in mice to test drug effects on gene targets.
Results: We created original experimental model of DLB based on the combined effects of α-syn and Aβ. Mice of transgenic B6.Cg-Tg (Prnp-SNCA*A53T)23Mkle/J strain overexpressing a mutant form of the human α-syn were administered with Aβ into the lateral ventricles of the brain. We used several steps of systems biology analysis concluded by construction of experimental model in mouse. The original experimental models of various forms of proteinopathy associated with the pathogenesis of DLB will be used to study the immune characteristics and the severity of neuro-inflammation in different brain areas that accompany neurodegeneration.
Conclusion: The results broaden the modern notion about the role of neuro-inflammation in the pathogenesis of DLB as well as to evaluate the prognostic significance of the concentrations of pro-inflammatory cytokines (TNF-alpha, IL-1, IL-6) in the blood as biomarkers for the differential diagnosis of dementia with proteinopathy.
Acknowledgements: The research has been supported by Russian RFBR, Indian DST and Chinese NSCF. Computing was supported by ICG SB RAS budget project 0324-2018-0017.

DOI 10.18699/WIBSB-2018-07

DNA methylation studies in plants based on sequencing technologies

P. Arrigo[1]*, A.A. Anashkina[3], N.G. Esipova[3], O.B. Dobrovolskaya[4, 5], Y.L. Orlov[4, 5]*
[1] *ISMAC Institute, Genova, Italy*
[3] *Engelhardt Instittue of Molecular Cell Biology RAS, Moscow, Russia*
[4]*I nstitute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[5] *Novosibirsk State University, Novosibirsk, Russia*
* *e-mail: orlov@bionet.nsc.ru*

Key words: epigenetics, methylation, plants, next-generation sequencing, nucleosome, ncRNA

Motivation and Aim: The epigenetic mechanisms regulating phenotype expression could be studied now using novel sequencing technologies. Based on the commonly accepted definition, this term refers to all phenotypic variation that does not require a nucleotide change in the DNA. The epigenetic regulation varies from molecular mechanism such as DNA methylation to high order modification such as nucleosome positioning o chromosome unfolding. In the last decade it is emerged the critical role of epigenetic modification of long term exposure to environmental stress factors and stress response in plants. In an environmental sustainability perspective, the alteration of epigenetic regulation can also impact biodiversity.
Methods and Algorithms: The harmful effect of pollutants on human epigenetic control is largely demonstrated, however the effects on plants and, more generally on a specific ecological community, requires more studies. Several plants such as Arabidopsis, are became a system model to study different molecular mechanisms. The high complexity of plants genome organization requires to dedicate more studies in order to tackle the different epigenetic regulation in these biological systems. This kind of research has great impact in order to optimize the crop production under the pressure of climate change and pollution. The new high-throughput molecular techniques have allowed to enhance the capability to investigate the events that affect the epigenetic regulation. We guess that a particular relevance can be assigned to the chemical modifications of informational macromolecules (DNA and RNA) that could be considered as starting point of epigenetic regulation.
Results and conclusion: The molecular investigation of epigenetic process is greatly supported extensively by immunoprecipitation sequencing (ChIP-seq) and methylome microarray. These high-throughput techniques permit to maps a large number of methylated sites on DNA and to evaluate their role in transcriptional processes. The long non coding RNAs seem to be involved in the control of trasncriptional process [1]. It is also important to underline that ncRNA are also capable to contribute to nucleosome positioning. We will review current state of art in this area in the frames of new joint collaborative project.
Acknowledgements: The research has been supported by RFBR No. 18-54-00037. Computing done at Siberian Supercomputer center SB RAS was supported by budget project No. 0324-2018-0017.

References

1. Wang J., Meng X., Dobrovolskaya O.B., Orlov Y.L., Chen M. (2017) Non-coding RNAs and their roles in stress response in plants. Genomics Proteomics Bioinformatics. 15(5):301-312.

DOI 10.18699/WIBSB-2018-08

Computer analysis of alternative splicing events by RNA-seq data in brain cells

V.N. Babenko[1], N.V. Gubanova[1], A.V. Tsukanov[1], S.S. Kovalev[2], A.O. Bragin[1], G.V. Vasiliev[1], Yu.L. Orlov[1, 2]*

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

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

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

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

Key words: transcriptome, glioblastoma, alternative splicing, differential splicing, cancer stem cells

Motivation and Aim: Alternative splicing is a critical mechanism for expanding regulatory and functional diversity from a limited number of genes. It is particularly complex in the mammalian brain [1]. Glioblastoma is the most common and aggressive type of primary brain tumor, accounting for 80 % of malignant astrocytomas. Therefore, it is critical that the genetic pathways underlying the development of this type of cancer are defined.
Methods and Algorithms: The gene expression profiles of glioblastoma were obtained on cell culture samples of primary glioblastoma isolated and processed for RNA extraction. Transciptome profiling of normal brain samples and glioblastoma were done by Illumina sequencing. We used set of computer tools applied recently to analysis of gene expression in laboratory animals to study differential splicing events. To analyze alternative splicing events in the transcriptomics data MATS (multivariate analysis of transcript splicing) and rMATS (replicate MATS) were used as tools.
Results: We identified gene loci with highly significant differential isoforms expression. The major GO entries for alternatively spliced genes were cytoskeleton and intracellular (cytoplasmic) related genes. We found also genes of nuclear pore complex as differentially expressed in NGB cell culture sample. During the analysis of differential splicing events we found significant differences in splicing of three cancer associated genes, in particular: APP (amyloid beta precursor protein), CASC4 (cancer susceptibility candidate 4) and TP53.
Conclusion: Multiple alternative splicing transcripts have been identified as progression markers, including generalized splicing abnormalities and tumor- and stage-specific events.
Acknowledgements: The work was supported by RFBR, ICG SB RAS budget project 0324-2018-0017.

References

1. Babenko V.N. et al. (2017) Computer analysis of glioma transcriptome profiling: alternative splicing events. J. Integrative Bioinformatics. 14(3). pii: /j/jib.2017.14.issue-3/jib-2017-0022/jib-2017-0022.xml.

DOI 10.18699/WIBSB-2018-09

Graduate certificate programs provide both motivation and flexibility for careers in bioinformatics and biomedicine: experience of George Mason University, Virginia, USA

A. Baranova, V. Chandhoke
School of Systems Biology, George Mason University, Fairfax, VA, United States
** e-mail: aancha@gmail.com*

Key words: graduate certificates, non-degree programs, bioinformatics, laboratory and data management, career development

Motivation and Aim: The mission of the School of Systems Biology (SSB) at George Mason University (GMU) is to provide diverse population of post-baccalaureate and graduate students with a research and educational environment that allows integration of the various areas of molecular and cellular biosciences from a systems perspective. To this end, the SSB offers research-based degree programs whose graduates go on to pursue careers in academy, government agencies, private industry, higher education, and health care.

Since mid-2000, SSB GMU offers MS and PhD programs in Biology/Biosciences as well as in Bioinformatics and Computational Biology. In all four of these programs together, current headcounts reached more than 150 students, with 40 students graduating each academic year. Fifteen years long experience with managing these programs led us to understanding that "traditional" MS (30-32 credits) and PhD (72 credits) programs do not always fit the goals of post-baccalaureate students.

Results: To this end, two specialized non-degree Certificate programs were developed: 1) Graduate Certificate in Personalized Medicine (15 credits); 2) Graduate Certificate in Bioinformatics and Computational Biology (15 credits). In each program, twelve out of 15 credits may be transferred to any "traditional" MS or PhD program offered by SSB.

As an example, the following groups are identified as target groups for enrolling in Certificate Program in Personalized Medicine:

1) Non-Physician Healthcare professionals currently employed as healthcare administrators, physician assistants, nurses and healthcare technicians, nutritionists, dieticians, etc;
2) Recent BS or BA graduates seeking admission to medical graduate programs (pre-med and pre-dental students);
3) Healthcare IT professionals, computational biologists and bioinformaticians with attained degrees;
4) R&D professional in the Drugs and Diagnostics fields;
5) Sales and other professional in biomedical and insurance industry;
6) Bench scientists in different branches of biology (predominantly at BS and MS levels);
7) Health industry economists, Government policy makers and regulators, Patient advocates and media experts;
8) Employees and associates of biomedicine-related investment funds.

Both programs were successful in collecting substantial amount of applications: Certificate in Personalized Medicine (N = 14) was more attractive than Certificate in Bioinformatics and Computational Biology (N=2). In 2017, Certificate was then re-evaluated and expanded into Professional (non-Thesis) Masters in Bioinformatics

Management, also made available as a hybrid online degree. Superior attractiveness of Certificate in Personalized Medicine for potential applicants is explained by initial lack of advertising and the fact that an inception of the programs, only a few educational institutions offered training in Personalized Medicine. Given that this program responds to a need in educated workforce that is able to understand, interpret and implement recent discoveries in genomics, proteomics, and metabolomics in both basic and clinical research settings, it provides its graduates a competitive edge over BS in Biology alumni when they apply for professional jobs, even in absence of graduate degree.

By interviewing the students accepted to the Certificate in Personalized Medicine, we found that an overwhelming majority of these students see the Certificate program as a stepping stone toward the Master's or PhD degree, which they typically intend to complete within the same University.

Conclusion: Graduate certificate programs provide both motivation and flexibility for careers in bioinformatics and biomedicine, and facilitate educational access for professional populations of students by providing step-wise re-immersion into academic environment.

Acknowledgements: Supported by the College of Science, GMU, Virginia, USA.

DOI 10.18699/WIBSB-2018-10

Computer analysis of genes expression, involved in the serotonergic and dopaminergic systems work, in the ventral tegmental brain area of aggressive and non-agressive rats

A.O. Bragin*, A.L. Markel, R.V. Kozhemyakina, Y.L. Orlov
Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
** e-mail: ibragim@bionet.nsc.ru*

Key words: serotonergic system, dopaminergic systems, aggressive rats, computer analysis, gene expression

Motivation and Aim: Aggression remains one of the important problems even in developed countries. It is known, that aggressive behavior can be caused not only by social factors, but also by genetic factors. For this reason, the analysis of gene expression patterns – associated with aggressive behavior – is an urgent task. For the analysis of the genetic basis of aggressive behavior in the Institute of Cytology and Genetics SB RAS two rat lines were developed: aggressive and tolerant (non-aggressive).
Methods and Algorithms: In the work, the ventral tegmental area of the rat brain was analyzed, containing dopaminergic neurons and responsible for motivation. RNA-sequencing of the samples of the ventral tegmental region of the brain of aggressive and non-aggressive rats was performed [1].
Results: As a result of the computer analysis of RNA-seq data, the expression of protein genes – associated with dopamine and serotonin in the brain tissue of aggressive and non-aggressive rats was assessed. Significant differences in the expression of the serotonin and dopamine receptors, the dopamine transporter and the enzymes – responsible for the synthesis and catabolism of serotonin and dopamine – were found.
Conclusion: The obtained results may be of interest both for the problem of aggressiveness in fundamental science and for industry, for example, for breeders of animals.
Acknowledgements: The research has been supported by RFBR (grant No. 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 Integrative Bioinformatics. 13(4):292.

DOI 10.18699/WIBSB-2018-11

Epigenetic correlation of interleukin expression between cigarette smoking and the therapeutic efficiency of periodontitis

H.-M. Chang[1]*, T.-Y. Renn[2], N.-C. Teng[3], Y.-K. Huang[4]
[1] *Department of Anatomy and Cell Biology, College of Medicine, Taipei Medical University, Taipei, Taiwan*
[2] *Graduate Institute of Medical Sciences, College of Medicine, Taipei Medical University, Taipei, Taiwan*
[3] *School of Dentistry, College of Oral Medicine, Taipei Medical University, Taipei, Taiwan*
[4] *School of Oral Hygiene, College of Oral Medicine, Taipei Medical University, Taipei, Taiwan*
** e-mail: taiwanzoo@gmail.com*

Key words: biomedicine, epigenetics, methylation, periodontitis, interleukin

Cigarette smoking is widely considered as the most important environmental risk factor for the development and progression of periodontitis (PD). Epigenetic modulation caused by cigarette smoking may serve as the underlying mechanism for the increased susceptibility of PD. However, although the impact of cigarette smoking on exaggerating the severity of PD has been well-documented, the condition that why patients with the same clinical phenotype responded differently to therapeutic efficiency has still largely unknown. Concerning interleukins (ILs) is the major cytokine participated in the modulation and progression of numerous inflammatory diseases; the present study is aimed to determine whether cigarette smoking would alter the methylation status of ILs, and therefore, contribute to the therapeutic discrepancy following PD. A total of 167 patients consisting of 79 males and 88 females were recruited in this study. Well-trained interviewers carried out the standardized personal interviews based on a structured questionnaire. Both the saliva samples and the clinical measurements [including the plaque index (Pi), bleeding on probing (Bop), and pocket depth (Pd)] were taken at the baseline and after the treatment of PD. The genomic DNA isolated from the gingival tissues was modified firstly by sodium bisulfite and then analyzed for DNA methylation levels of IL-1β, IL-6, and IL-8 genes with direct sequencing. The levels of the cytokines were further determined by the enzyme-linked immunosorbent assay. The results indicated that patients suffered from PD with smoking behavior exhibited a higher percentage of hypomethylation of IL-1b gene and expressed high levels of IL-1β, IL-6, and IL-8 than that of non-smokers. Moreover, the degree of Pi, Bop, and Pd after treatment was more evident in smoking patients, suggesting that the therapeutic efficiency of PD was negatively correlated with the extent of epigenetic change. As changes in methylation profile and subsequent increase in the expression of IL may be causally related to the poor response of treatment following PD, our findings thus not only provide a useful strategy for easy identification of patients at risk for poor therapeutic efficiency, but also helps to reduce the treatment costs that form the health economic perspective for the society.

DOI 10.18699/WIBSB-2018-12

Roles of non-coding RNAs in stress response in plants

M. Chen[1]*, J. Wang[1], O.B. Dobrovolskaya[2, 3], V.N. Babenko[2, 3], Y.L. Orlov[2, 3], Y. Liu[1], L.S. Samarina[4]*

[1] *Zhejiang University, Hangzhou, China*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[4] *FSBSI Russian Research Institute of Floriculture and Subtropical Crops, Sochi, Russia*
** e-mail: mchen@zju.edu.cn, q11111w2006@ya.ru*

Key words: non-coding RNA, transcriptomics, stress response, crop plants

Motivation and Aim: Eukaryotic genomes encode thousands of non-coding RNAs (ncRNAs), which play crucial roles in transcriptional and post-transcriptional regulation of gene expression. The computer analysis of transcription regulation in stress response in crop plants is a challenging problem. Accumulating evidence indicates that ncRNAs, especially microRNAs (miRNAs) and long ncRNAs (lncRNAs), have emerged as key regulatory molecules in plant stress responses. We have summarized the current progress on the understanding of plant miRNA and lncRNA identification, characteristics, bioinformatics tools, and resources, and provided examples of mechanisms of miRNA- and lncRNA-mediated plant stress tolerance [1]. Although remarkable progress has been made in explaining the role of plant miRNAs and lncRNAs in plant adaption to stress, mechanistic details are still limited.

Methods and Algorithms: With the advantage of the next-generation sequencing technologies and bioinformatics approaches, a great number of ncRNAs have been identified and characterized in plants, especially miRNAs and lncRNAs. miRNAs and lncRNAs are two important types of ncRNAs in plants, which play important roles in various biological processes. Rapid progress in high-throughput sequencing and advancement of bioinformatics tools provide revolutionary ways for identification and prediction of novel ncRNAs.

Results: In this work, we summarized the common bioinformatics tools and resource of miRNAs and lncRNAs. In addition, recently-developed single-cell sequencing and single-molecule sequencing will offer more opportunities to increase the number of ncRNAs. Therefore, it is necessary to develop new bioinformatics methods for the identification and functional analysis of ncRNAs.

Conclusion: Recent works show role of alternative splicing events in stress (draught) response in bread wheat [2]. We continue work on integration of stress-response data and gene sets using available data sources.

Acknowledgements: The research has been supported by RFBR 18-04-00483. Computing done at Siberian Supercomputer center SB RAS was supported by budget project 0324-2018-0017.

References

1. Wang J., Meng X., Dobrovolskaya O.B., Orlov Y.L., Chen M. (2017) Non-coding RNAs and their roles in stress response in plants. Genomics Proteomics Bioinformatics. 15(5):301-312.
2. Liu Z. et al. (2018) Global profiling of alternative splicing landscape responsive to drought, heat and their combination in wheat (Triticum aestivum L.). Plant Biotechnol J. 16(3):714-726.

DOI 10.18699/WIBSB-2018-13

Versatile interactions and bioinformatics analysis of noncoding RNAs

Q. Chen, X. Meng, Q. Liao, M. Chen*
Department of Bioinformatics, College of Life Sciences, Zhejiang University, Hangzhou, P. R. China
** e-mail: mchen@zju.edu.cn*

Key words: noncoding RNAs, ncRNA transcription, ncRNA–RNA interaction, ncRNA–DNA interaction, ncRNA, protein interaction, bioinformatics resources

Advances in RNA sequencing technologies and computational methodologies have provided a huge impetus to noncoding RNA (ncRNA) study. Once regarded as inconsequential results of transcriptional promiscuity, ncRNAs were later found to exert great roles in various aspects of biological functions. They are emerging as key players in gene regulatory networks by interacting with other biomolecules (DNA, RNA or protein). Here, we provide an overview of ncRNA repertoire and highlight recent discoveries of their versatile interactions. To better investigate the ncRNA-mediated regulation, it is necessary to make full use of innovative sequencing techniques and computational tools. We further describe a comprehensive work- flow for in silico ncRNA analysis, providing up-to-date platforms, databases and tools dedicated to ncRNA identification and functional annotation.

DOI 10.18699/WIBSB-2018-14

Cell-free RNA studies in cancer based on T oligo-primed polymerase chain reaction (TOP-PCR) technology

K.-P. Chiu[1], N.V. Gubanova[2], A.V. Tsukanov[2], S.S. Kovalev[2, 3], Y.L. Orlov[2, 4]*
[1] *Genomics Research Center, Academia Sinica, Taipei, Taiwan*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[3] *Novosibirsk State Medical University, Novosibirsk, Russia*
[4] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: orlov@bionet.nsc.ru*

Key words: sequencing technology, transcriptome, glioblastoma, splicing, cancer

Motivation and Aim: We suggest joint research project devoted to the development of cell-free RNA (cfRNA) analysis technology using TOP-PCR. The project relies on the expertise of Taiwanese group in TOP-PCR technology development and the experience of Russian science team in gene expression and transcriptomics data analysis and network reconstruction. The TOP-PCR technology [1] will be applied for analysis of biomarkers in cancer and other diseases. Body fluid nucleic acid sequencing is a powerful noninvasive approach for the diagnosis of genetic defects, infectious agents and diseases. The success relies on the quantity and quality of the DNA samples.
Methods and Algorithms: Numerous clinical samples are either at low quantity or of poor quality due to various reasons. To overcome these problems, Taiwanese group have developed T oligo-primed polymerase chain reaction (TOP-PCR) for full-length nonselective amplification of minute quantity of DNA fragments. TOP-PCR adopts homogeneous "half adaptor" (HA), generated by annealing P oligo (carrying a phosphate group at the 5′ end) and T oligo (carrying a T-tail at the 3′ end), for efficient ligation to target DNA and subsequent PCR amplification primed by the T oligo alone.
Results: The preliminary results also showed that TOP-PCR is a superior method for detecting apoptosis and outperforms the method adopted by Illumina for DNA amplification.
Conclusion: We will analyze cancer mutations in cell-free DNA and develop methods of gene network reconstruction and computer models of non-coding RNA interactions with focus on cell-free RNA. Gene expression analysis will allow find biomarkers associated with cancer, especially for glioblastoma [2]. The group will use text mining of science literature and associate gene network reconstruction methods to find related genes and RNA targets. Associative gene network reconstruction will be based on ANDVisio and ANDSystem tools; gene network modeling will be based on TRRD and GeneNet databases developed at ICG SB RAS. The project has fundamental role for biomarkers studies and could be applied for wide range of biotechnology problems.
Acknowledgements: The work was supported by RFBR No. 18-29-09105 and ICG SB RAS budget project No. 0324-2018-0019.

References

1. Nai Y.-S. et al. (2017) T Oligo-Primed Polymerase Chain Reaction (TOP-PCR): A Robust Method for the Amplification of Minute DNA Fragments in Body Fluids. Scientific Reports. 7:40767.
2. Babenko V.N. et al. (2017) Computer Analysis of Glioma Transcriptome Profiling: Alternative Splicing Events. Journal of Integrative Bioinformatics. 14(3). pii: /j/jib.2017.14.issue-3/jib-2017-0022/jib-2017-0022.xml.

DOI 10.18699/WIBSB-2018-15

Quantifying genome sequence repeatability by repeater

M. Dai[1, 2], C. Feng[1], M. Chen[1]*
[1] *Department of Bioinformatics, the State Key Laboratory of Plant Physiology and Biochemistry, College of Life Sciences, Zhejiang University, Hangzhou, China*
[2] *Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China*
** e-mail: mchen@zju.edu.cn*

Key words: DNA repeats, repeatability map

Motivation and Aim: DNA repeats are abundant in eukaryotic genomes and illustrated to play a crucial part in genome evolution and regulation. To identify various repeats and classify them into different families are the key aspects of previously related researches. And a large number of approaches are proposed to detect and annotate repeat elements in the genome. While a few de novo repeats identification methods are able to generate quantitative repetitiveness map of genome sequences based on word counting algorithms. However, characteristics of such repetitiveness pattern are not well recognized and the applications are still limited. Therefore we developed a software named Repeater that quantifies the sequence repeatability, which is defined as a genome property to reflect the sequence repetitive level.
Methods and Algorithms: Repeater mainly consists of three parts. In the first part, we presented min tree, a modified suffix tree data structure, to compute all the exact repeats in the genome. The min tree is realized in Java and contains less redundant information than the suffix tree and able to run in linear time and space. Next, we transformed the fragmented sequence data originated from repeater-core to repeatability map. Both Single Nucleotide Repeatability (SNRP) and Sequence Average Repeatability (SARP) can be generated to represent the repeatability at single base and sequence region level. After we got the genome sequence repeatability map, MACS2 was utilized to capture the peaks which represent the highly repetitive regions.
Results: The analysis shows that Repeater performed well at identifying the highly repetitive sequences in the genome by determining the peak regions of repeatability map. And we found that repeatability map is complementary to the mappability track, so it may also useful in reads mapping check. Combined with ChIP-seq and DNase-seq data, highly repetitive regions identified by our tool are related to epigenetic modifications and chromatin accessibility. That indicates the potential capability of repeatability to serve as a genome feature for further structure and function predictions. The software and an online testing program are freely available at http://bis.zju.edu.cn/repeater. And we also provide repeatability track of common model species implemented with JBrowse on the website.
Conclusion: We developed a software Repeater to quantify the genome sequence repeatability. Highly repetitive regions in the genome can be efficiently detected on the basis of repeatability map. And we demonstrated that repeatability may be useful for reads mapping check and chromatin accessibility prediction.
Acknowledgements: We are grateful to the members of Ming Chen's laboratory for helpful discussions and comments.

DOI 10.18699/WIBSB-2018-16

Computer analysis of the influence of the presence of transcription factors on plant genome evolution

A.I. Dergilev[1, 2, 3], A.V. Tsukanov[2, 4], Y.L. Orlov[1, 2]*
[1] *Novosibirsk State University, Novosibirsk, Russia*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[3] *Université Grenoble Alpes, Grenoble, France*
[4] *Perm National Research Polytechnic University, Perm, Russia*
** e-mail: orlov@bionet.nsc.ru*

Key words: bioinformatics, full-genome analysis, binding sites, PWM, big data

Motivation and Aim: The great evolution of plants led to increasing during plant complexification: 5,700 species of green algae combined in 360 genera exist today. With changing of complexification of plants the specific transcription factors (TF) in plant genomes also have been changing with time as well as TF binding sites [1]. The question to figure out is: "How the presence of a TF changes the genome sequence?" What did evolved first? Does the transcription factor change first and then this created binding sites or are there already binding sites in the genome of the ancient one and the transcription factor evolved the binds of these one?

The idea is to look for TFBS in genomes with or without the TF and test whether enrichment is detectable overall [2]. The work requires full-genome analysis of the families of certain plants as well as highlighting the best results (best TFBS scores among all).

Methods and Algorithms: Although many algorithms for recognizing TFBS exist, tools for using the DNA binding models they generate are relatively scarce and their use is limited among the biologist community by the lack of flexible and user-friendly tools. We use a suite of tools "Morpheus" to analyse transcription factor binding sites (TFBS) on DNA sequences [3]. As input, the program uses a set of sequences in FASTA format and a PWM (Position Weight Matrix) and Python language to write scripts. We use library Matplotlib to get a graphic interpretation of our results (histograms for best TFBS representing at density and arrangement in their vicinity (probability of detecting others, relative position [like ER-IR-DR or 10N distance for ARFBS, high density for LFY or MADS also 10N]).

Availability: "Morpheus" is available at http://biodev.cea.fr/morpheus/

References

1. Rensing S.A. et al. (2008) "The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. Science. 319:64.
2. Wasserman W.W., Sandelin A. (2004) Applied bioinformatics for the identification of regulatory elements. Nat. Rev. Genet. 5:276-287. PMID 15131651. DOI 10.1038/nrg13152.
3. Minguet E.G., Segard S., Charavay C., Parcy F. (2015) MORPHEUS, a Web tool for transcription factor binding analysis using position weight matrices with dependency. PLoS One. 10(8):e0135586. DOI 10.1371/journal.pone.0135586.

DOI 10.18699/WIBSB-2018-17

SNPs associated with accelerated senescence in OXYS rats

V. Devyatkin[1, 2]*, O. Redina[1], N. Muraleva[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 diseases. However, little is known about a genetic overlap between complex age-related diseases. The senescence-accelerated OXYS rats selected in the ICG SB RAS (Novosibirsk) are a good model to identify the pathways that modulate the onset and progression of multiple age-related diseases as these rats develop 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 aim of our work is to investigate the transcriptome of OXYS rats and to identify the mutations (SNPs) in genes associated with cataract, which can potentially contribute to the development of accelerated aging.
Methods and Algorithms: We used the RNA-Seq data obtained from sequencing of 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. The mutation was considered as reliable SNP if it was detected 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, and KEGG Disease databases.
Results: In the genome of OXYS rats 52539 SNPs overlapped with 11684 transcripts representing 8012 genes. 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 543 described and 614 novel SNPs related to 255 genes that can be associated with cataract development in OXYS rats. 4 of this genes, *Pex2*, *Nbn*, *Rab18* and *Prss56* have SNPs (rs198310567, rs105362013, rs106234270 and rs106604882, respectively), which are expected to exert a deleterious effect on the structure or function of the encoded proteins. These polymorphisms are also described for SHR/OLAIPCV and SD rat strains, which were not earlier tested for signs of cataract. It is known that mutations in these genes are associated with mitochondrial diseases, nervous and cardiovascular disorders, consistent with the complex manifestation of the senile phenotype in OXYS rats.
Conclusion: The results of the study may serve as a background for further verification of SNPs contribution 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/WIBSB-2018-18

A Kolmogorov – Smirnov based approach for predicting targets of transcription

M.J. Djordjevic[1]*, M.R. Djordjevic[2], E.Zdobnov[3]
[1] *Faculty of Biology, Institute of Physiology and Biochemistry, University of Belgrade, Belgrade, Serbia*
[2] *Institute of Physics Belgrade, University of Belgrade, Belgrade, Serbia*
[3] *Swiss Institute of Bioinformatics and Department of Genetic Medicine and Development, University of Geneva, Geneva, Switzerland*
** e-mail: dmarko@bio.bg.ac.rs*

Key words: regulatory element detection, Kolmogorov–Smirnov, transcription regulation, ChIP-seq

Motivation and Aim: Accurately predicting direct targets of transcriptional regulators is necessary to understand gene expression regulation. Predicting these targets typically leads to a large number of false positives, and usually corresponds to searching for high-scoring binding sites in the upstream genomic regions. In contrast to the common approach, we here propose a novel concept, where overrepresentation of the scoring distribution that corresponds to the entire searched region is assessed, as opposed to predicting individual binding sites.
Methods: As opposed to predicting individual binding sites, we here propose a novel concept, where the entire searched region is scored, and overrepresentation of this scoring distribution is assessed [1]. We implement this concept through both Kolmogorov–Smirnov (KS) and Anderson–Darling (AD) tests, where both allow straightforwardly predicting *P*-values for each target.
Results: We first apply this approach to pleiotropic bacterial regulators, including σ^{70} (housekeeping bacterial σ factor) whose target prediction is a classical bioinformatics problem characterized by high number of false positives. We show that KS based approach is both more accurate and faster compared to AD, departing from the current paradigm of AD being more accurate (though slower). Moreover, KS has a significantly higher accuracy compared to the standard approach, while straightforwardly assigning well established *P*-values to potential targets. Secondly, we apply the method to ChIP-seq data analysis, to test how it can predict bacterial transcription targets *in vivo*. While we find a good correspondence between computational predictions and *in vitro* binding data, both of them correlate significantly worse with *in vivo* data [2].
Conclusion: New KS based method proposed here, which assigns P-values to fixed length upstream regions, provides a fast and accurate approach for predicting bacterial transcription targets. Binding of transcription factors in vivo may be significantly influenced by factors other than binding energy, even in bacteria where this binding does not happen in the chromatin context.
Acknowledgement: This work is supported by the Swiss National Science Foundation under SCOPES project number IZ73Z0_152297 and by the Ministry of Education and Science of the Republic of Serbia under Project number ON173052.

References

1. Djordjevic M.J., Djordjevic M.R., Zdobnov E. (2017) Scoring targets of transcription in bacteria rather than focusing on individual binding sites. Front. Microbiol. 8:2314.
2. Djordjevic M.J., Djordjevic M.R. (2018) Regulation by bacterial transcription factors: relationship between binding energy and binding site functionality. (in preparation).

DOI 10.18699/WIBSB-2018-19

Computer studies of miRNA in abiotic stress response in plants

O.B. Dobrovolskaya[1, 2], A.M. Spitsina[2], A.G. Galieva[2], A.V. Tsukanov[1, 2], A.G. Bogomolov[1, 2], Y.L. Orlov[1, 2]*, M. Chen[3]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Zhejiang University, Hangzhou, China*
** e-mail: orlov@bionet.nsc.ru*

Key words: non-coding RNA, transcriptomics, stress response, crop plants

Motivation and Aim: Computer studies of miRNA in stress response in plants present a challenging problem. Eukaryotic genomes encode thousands of non-coding RNAs (ncRNAs), which play crucial roles in transcriptional and post-transcriptional regulation of gene expression. The computer analysis of transcription regulation in stress response in crop plants is important. ncRNAs, especially microRNAs (miRNAs) and long ncRNAs (lncRNAs), have emerged as key regulatory molecules in plant stress responses. We summarized the current progress on the understanding of plant miRNA and lncRNA identification, analysis, usage of bioinformatics tools and resources [1]. Although remarkable progress has been made in explaining the role of plant miRNAs and lncRNAs in plant adaption to stress, mechanistic details are still limited.
Methods and Algorithms: We used available databases for this review. With the advantage of the next-generation sequencing technologies and bioinformatics approaches, a great number of ncRNAs have been identified and characterized in plants, especially miRNAs and lncRNAs. miRNAs and lncRNAs are two important types of ncRNAs in plants, which play important roles in various biological processes. Rapid progress in high-throughput sequencing and advancement of bioinformatics tools provide revolutionary ways for identification and prediction of novel ncRNAs.
Results: In this work, we summarized the common bioinformatics tools and resource of miRNAs and lncRNAs. In addition, recently-developed single-cell sequencing and single-molecule sequencing will offer more opportunities to increase the number of ncRNAs. Therefore, it is necessary to develop new bioinformatics methods for the identification and functional analysis of ncRNAs.
Conclusion: Recent works show role of alternative splicing events in stress (draught) response in bread wheat [2]. We continue work on integration of stress-response data and compiled first database on stress response genes in different crop plants.
Acknowledgements: The research has been supported in part by RFBR. Computing done at Siberian Supercomputer center SB RAS was supported by budget project No. 0324-2018-0017.

References

1. Wang J., Meng X., Dobrovolskaya O.B., Orlov Y.L., Chen M. (2017) Non-coding RNAs and their roles in stress response in plants. Genomics Proteomics Bioinformatics. 15(5):301-312.
2. Liu Z. et al. (2018) Global profiling of alternative splicing landscape responsive to drought, heat and their combination in wheat (Triticum aestivum L.). Plant Biotechnol J. 16(3):714-726.

DOI 10.18699/WIBSB-2018-20

Selective killing of cancer cells by ultrasound assisted by biopolymer-coated metal nanoparticles

Y.A. Dyakov[1, 2]*, O.K. Kosheleva[1, 2], T.C. Lai, M. Hsiao[1], N.G. Chen[3], C.H. Chen[1]*
[1] *Genomics Research Center, Academia Sinica, Taipei, Taiwan*
[2] *Karpov Institute of Physical Chemistry, Moscow, Russia*
[3] *Institute of Biomedical Engineering, Department of Electrical and Computer Engineering, College of Electrical and Computer Engineering, National Chiao-Tung University, Hsinchu, 300, Taiwan*
* *e-mail: yuri_dyakov@mail.ru;winschen@gate.sinica.edu.tw*

Key words: Metal nanoparticles, biopolymer coating of nanoparticles, ultrasound treatment of cancer.

Motivation and Aim: Cancer is a leading cause of death worldwide. Traditional physical methods of cancer treatment, such as radiation therapy, including electromagnetic ionizing radiation, elementary particle ionizing beam radiation (electrons, protons, and neutrons), and non-ionizing radiation (photons, microwaves, and radio waves) have numerous undesirable side effects and cannot provide fully efficient treatment of the decease. The ionizing radiation is intended to be directed only at the tumor. Moreover, ionizing radiation itself may cause DNA mutation in normal cells, causing these cells to become cancerous. Non-ionizing radiation therapies are mainly based on hyperthermia in tumors as a result of the higher sensitivity of tumor cells to heat than their normal counterparts. But these methods are also does not have sufficient selectivity.
Methods and Algorithms: In our work we use high intensity ultrasound treatment enforced by addition of biopolymer-coated various metal nanoparticles in tissue or cell culture. In our study we used gold, silver, palladium, mercury, platinum, bismuth, wolfram, and some other metal nanoparticles covered by biologically inertial specific polymer shells.
Results: Combination of ultrasound and metal nanoparticles leads simultaneously to the strengthening of the destructive effect of ultrasound irradiation, and the selective toxicity of nanoparticles for cancer cells. This allows decreasing the power of the ultrasound radiation and expands the irradiated area without damaging the surrounded normal tissues. Proper selection of biopolymer shells, as far as the size and composition of metal cores provide the selective killing of cancer cells in the main tumor and in the metastases without negative impact to the surrounded normal tissues and perceptible side effects. This work is the continuation of previous studies [1,2].
Conclusion: We found that nanoparticle-assisted therapy of cancer by the high intensity ultrasound irradiation significantly increase the efficiency of the treatment, allows reducing the intensity of irradiation, and decreasing the side effects.
Acknowledgements: This work was supported by Ministry of Science and Technology of Taiwan (MOST 102-2113-M-001-002-MY5 and NHRI-Ex107-10603EI).

References

1. Kosheleva O.K. et al. (2016) Selective killing of cancer cells by nanoparticle-assisted ultrasound. Journal of Nanobiotechnology 14: Art.No. 46.
2. Kosheleva O.K. et al. (2017) Nanoparticle-Assisted Ultrasound for Cancer Therapy. Journal of Biomedical Science and Engineering 10(05):54-58.

DOI 10.18699/WIBSB-2018-21

The comparison of brain stem transcriptomes in rats from hypertensive ISIAH and normotensive WAG strains

L.A. Fedoseeva[1], L.O. Klimov[1, 2], N.I. Ershov[1], V.M. Efimov[1, 2], A.L. Markel[1, 2], Y.L. Orlov[1, 2], O.E. Redina[1]*

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

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

** e-mail:oredina@ngs.ru*

Key words: stress-sensitive hypertension, brain stem, transcriptional profiling, RNA-Seq, ISIAH rats

Motivation and Aim: Neurons of the brain stem control the activity of the sympathetic nervous system and play an important role in the blood pressure (BP) regulation. The purpose of this work was to elucidate the central mechanisms that contribute to the genetic predisposition for the development of a stress-sensitive form of hypertension in ISIAH rats.

Methods: Using the RNA-Seq method, a comparative analysis of transcriptomes of the brain stems (medulla oblongata and pons) of hypertensive ISIAH rats and normotensive WAG rats at the age of 3 months was carried out. To identify the differentially expressed genes (DEGs), the Cufflinks / Cuffdiff software package was used, the PLS-DA (partial-least squares discriminant analysis) was employed to detect the DEGs, which may contribute the most to the inter-strain differences.

Results: A total of 224 DEGs were found in the brain stems of ISIAH and WAG rats. Their annotation in the databases made it possible to identify genes associated with hypertension, as well as with diseases of the central nervous system and with a number of diseases that often accompany the development of hypertension (immune system diseases, insulin resistance, diabetes). Functional annotation of DEGs has revealed a number of biological processes that can influence the development of arterial hypertension in ISIAH rats. The most statistically significant was the group of DEGs described by the term 'hormone metabolic process', which includes several genes that control the biosynthesis of steroid hormones. The most essential for the present work were the groups of genes described by terms 'blood circulation', 'regulation of blood pressure', 'regulation of blood vessel size', 'muscle contraction' and 'tissue remodeling'. The detection of groups of genes associated with a response to various stimuli, such as 'response to endogenous (hormone) stimulus', 'response to external stimulus', and 'response to stress' (oxidative stress) underlines the existence of the homeostatic problems in the brain stem of ISIAH rats. Using the PLS-DA method, the genes possibly making the maximum contribution to inter-strain differences have been identified.

Conclusion: The results obtained contribute to an understanding of the central mechanisms of genetic predisposition to the development of a stress-sensitive form of hypertension.

Acknowledgements: The study has been supported by Russian Science Foundation and budget project No. 0324-2018-0016.

DOI 10.18699/WIBSB-2018-22

Promoter-pathway analysis approach to interpretation of microarray data of the antitumor peptide CIGB-552

J.R. Fernández-Massó[1]*, D. Palenzuela-Gardon[1], D. Vázquez-Blomquist[2], L. Novoa[2], M. Vallespí[1]

[1] *Pharmaceutical Department, Center for Genetic Engineering and Biotechnology, Cuba*

[2] *Department of System Biology, Center for Genetic Engineering and Biotechnology, Cuba*

** e-mail: julio.fernandez@cigb.edu.cu*

Key words: CIGB-552, cytotoxic peptide, promoter analysis, pathway analysis

Motivation and Aim: CIGB-552 is a novel synthetic peptide derived from the antimicrobial peptide LALF32-51 (Limulus sp) which has been shown to be a potential candidate for the anticancer therapy and one of its useful property is the cell-penetrating capacity. COMMD1 is a newly recognized pleotropic protein that plays an important role in inflammation, hypoxic response and cell survival. Others have also demonstrated that COMMD1 levels are reduced in some cancers, which is associated with reduced survival. A strategy is presented that allows a causal analysis of co-expressed genes, which may be subject to common regulatory influences. Promoter analysis for potential transcription factor (TF) binding sites in combination with a network-based analysis of the upstream pathways that control the activity of these TFs is shown to lead to hypothetical master regulators.

Methods and Algorithms: Transcriptomics studies were conducted for the identification of molecular mechanism and cellular targets of CIGB-552 peptide. TF analysis was conducted using iRegulon and geneXplain. Networks were visualized in Cytoscape.

Results: The result of the promoter analysis comprises enriched TF-binding motifs for each cluster of up- and down-regulated genes. When we followed the upstream activation pathways of the TFs potentially involved in the (co-)regulation of the differentially expressed genes, we found the potential master regulator of the up-regulated genes. Altogether, we noticed that the suggested master regulators are involved in promoting tumor progression and/or apoptosis.

Conclusion: In this study, we show that CIGB-552 regulates pathways that are known to play essential roles in apoptosis or cancer development.

References

1. Rekin's J. et al. (2014) iRegulon: from a gene list to a gene regulatory network using large motif and track collections. PLoS Comput Biol. 10(7):e1003731.
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 10.18699/WIBSB-2018-23

The cellular role of *Drosophila hyperplastic discs* gene

Y.A. Galimova[1], E.U. Bolobolova[2], A.S. Khruscheva[2], N.V. Dorogova[2]*
[1] *Institute of Molecular and Cellular Biology SB RAS, Novosibirsk, Russia, Novosibirsk, Russia*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
** e-mail: dorogova@bionet.nsc.ru*

Key words: tumor suppressor Hyd, ubiquitin ligase, *Drosophila*

Motivation and Aim: The tumor suppressor Hyd (hyperplastic discs), also known as EDD (E33 isolated by differential display) or Rat100, is a member of the ubiquitin ligase (E3) family containing the HECT domain (homologous to E6-associated protein carboxyl terminus) that provides specific ubiquitin-mediated proteolysis. A highly conservative ubiquitin-dependent proteolysis system controls the degradation of many key regulatory proteins. The tumor suppressor nature of Hyd protein was supported by the studies in mammalians [1]. Hyd is required for the regulation of cell proliferation during development [2], and mutations in the *hyd* gene resulting in developmental abnormalities that include adult sterility caused by germ cell defects [3].
Methods and Algorithms: We used Bloomington stocks: $kni^{ri\text{-}1}$ hyd^{15} $e^{1}/TM3$, Sb^{1} (3718) and y^{1} w^{1118}; *PBac{3HPy+}*$hyd^{C017}/TM3$, Sb^{1} Ser^{1} (16256) as the source of *hyd* alleles. Isolated testes and ovaries were examined by the methods of light, fluorescence and electron microscopy.
Results: The role *hyd* in *Drosophila* germinal line has been studied. It is shown that mutations of this gene cause multiple cell division abnormalities in spermatogenesis, which included: abnormal chromosome segregation and absence of meiotic cytokinesis resulting in formation of polyploid spermatids, and also defects in spermatid elongation. However, these abnormalities do not lead to hyperplasia of generative tissue. In contrast, in oogenesis, this gene manifests the function of a tumor growth suppressor, the disturbances of which cause an excessive number of cell divisions and overproliferation.
Conclusion: The question arises, why the same protein has different effects on the cell division in the generative tissue of males and females. Even earlier studies have shown that tumor suppressor genes can function in many tissues, but only in some of them they cause hyperplasia [4]. Probably, in oogenesis, the *hyd* gene manifests as a tumor suppressor, and in spermatogenesis as a ubiquitin ligase.

References

1. Cho E. et al. (2006). Delineation of a Fat tumor suppressor pathway. Nat. Genet. 38:1142-1150.
2. Mansfield, E.A. et al. (1994). Genetic and molecular analysis of hyperplastic discs, a gene whose product is required for regulation of cell proliferation in *Drosophila melanogaster* imaginal discs and germ cells. Dev. Biol. 165(2):507-526.
3. Callaghan M.J. et al. (1998). Identification of a human HECT family protein with homology to the *Drosophila* tumor suppressor gene hyperplastic discs. Oncogene. 17:3479-3491.
4. Gateff E. (1994). Tumor suppressor and overgrowth suppressor genes of *Drosophila melanogaster*: Developmental aspects. Int. J. Dev. Biol. 38:565-590.

DOI 10.18699/WIBSB-2018-24

Comparative analysis of transcriptional regulation of several cell types' migration in *Drosophila* embryo

Y.A. Galimova[1], D.E. Maslov[2], E.D. Shvedkina[2], Y.A. Fedorova[2], A.S. Khruscheva[2, 3], S.A. Fedorova[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*
* *e-mail: fsveta@bionet.nsc.ru*

Key words: transcription regulation, cell migration, embryo, *Drosophila*, GAF

Motivation and Aim: Molecular-genetic nature of cell migration is evolutionarily conservative and has high homology in different cell types, organs and tissues, and is similar in many aspects with the movement of the immune system cells and metastatic tumors. The *Drosophila melanogaster* embryo is a good model for studying the mechanisms of cell migration. First, primordial germline cells (PGC) start to migrate individually through midgut towards somatic gonadal precursor cells at stage 9. Beginning from stage 10, caudal visceral mesoderm (CVM) cells move as a streaming collective. Hemocytes (HCs) initiate a posterior migration along the ventral midline at stage 11. In contrast to CVM, HCs appear to move more independently, as individuals. All migration types involve loose streams of cells that appear to be controlled in their movement.

Methods and Algorithms: Using the Flybase data, gene sets with tissue-specific expression at the 1-11 stages of embryogenesis were formed. Among these, we select genes exhibit specific expression, confirmed by in situ hybridization experiments from Y-K Bae and coauthors study [1], through the BDGP (Berkeley Drosophila Genome Project) database, or by our studies. Transcriptional regulation data were obtained from DroId (A comprehensive database of gene and protein interactions) database.

Results: In this work, we compared the transcriptional regulation of 3 migrating cell types from the Drosophila embryo: PGC, CVM (precursors of longitudinal muscles of the gut), and HCs (the Drosophila equivalent of blood cells). We found that among 24 genes, supported the PGC migration, expression of 23 genes is regulated by transcription factor GAF. Also, among 73 genes shared gene expression profiles in both migrating CVM cells and HCs [1], expression of 64 genes are regulated by transcription factor GAF. Moreover, we revealed, that PGC migration is affected in the mutants of Trl gene that encodes the Drosophila GAF, as well as in mutants of GAF targets *shg*, *upd2* and *tll*.

Conclusion: Comparison of transcriptional regulation of genes shared in common across different migrating cell types suggested that transcription factor GAF is the potential key regulator of cell migration in Drosophila embryo.

Acknowledgements: The research has been supported by RFBR No. 18-34-00321mol-a.

References

1. Bae Y.K., Macabenta F., Curtis H.L., Stathopoulos A. (2017) Comparative analysis of gene expression profiles for several migrating cell types identifies cell migration regulators. Mech Dev. 148:40-55.

DOI 10.18699/WIBSB-2018-25

Influence of homo-repeats on the aggregation properties of proteins from eukaryotic and bacterial proteomes and codon usage for homo-repeats in human proteome

O.V. Galzitskaya[1]*, G.S. Novikov[2], N.V. Dovidchenko[1], M.Yu. Obanov[1]
[1] *Institute of Protein Research, Institutskaya Street 4, Pushchino, Russia*
[2] *St. Petersburg Academic University, St. Petersburg, Russia*
** e-mail: ogalzit@vega.protres.ru*

Key words: homo-repeat, codon usage, proteome, splicing site; disease

The influence of homo-repeats with lengths larger than four on the aggregation properties of proteins has been studied across 122 eukaryotic and bacterial proteomes. It has been found that proteins with homo-repeats are on average longer than in the whole database. The ability of proteins with homo-repeats to aggregate cannot be explained only by the presence of long homo-repeats in the proteins. There should be other characteristics of proteins increasing the aggregation property including such as appearance of homo-repeats in pairs in the same protein.

We have found the biases for codon usages for some amino acids in homo-repeats for human proteome and for all amino acids when the same codon is used for homo-repeats. Similar results are obtained for human proteins with homo-repeats associated with diseases. Moreover, for proteins associated with diseases (from the HraDis database), the fraction of proteins for which the same codon is used for homo-repeats is larger than for proteins from the human proteome. We are the first who demonstrated for human proteome that in some cases the splicing sites correspond to the homo-repeats and these sites more often appear at the *C*-terminal part of the homo-repeats.

Acknowledgements: This work was supported by the programs “Molecular and Cellular Biology” (01201353567) and by the Russian Science Foundation (No. 18-14-00321).

DOI 10.18699/WIBSB-2018-26

Genome admixture components accurately predict quantitative functional traits in plants

L. Gentzbitte[1]*, C. Ben[1], M. Mazurier[1], M.G. Shin[2], T. Lorenz[3], M. Rickauer[1], P. Marjoram[2], S.V. Nuzhdin[2], T.V. Tatarinova[3]

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

[2] *Department of Biological Sciences, University of Southern California, Los Angeles, USA*

[3] *Department of Biology, Division of Natural Sciences, University of La Verne, La Verne, CA, USA*

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

Key words: admixture, genomic prediction, prediction accuracy, WhoGEM, genomic selection, molecular ecology, quantitative disease resistance, adaptation, *Medicago truncatula*

Motivation and Aim: Linking plant genomic and phenotypic variation for quantitative characteristics associated with fitness, functional traits or tolerance to biotic and abiotic stress provides tools to understand natural variation and optimize breeding schemes. Being indifferent to the source of genetic similarity between populations of a species (demography and local adaptation or selective breeding), we hypothesize that a large proportion of phenotypic variation between individuals may be best explained by population admixture across several ancestral populations originally separated.

Methods and Algorithms: Here we propose and test a method named "WhoGEM" to explain variation in genetically complex traits using population admixture proportions of plant individuals. The overall goal is to predict quantitative phenotypes rather than identify causative variations or candidate genes, or inferring the relative parts of demography and selection in the evolution of quantitative phenotypes. The WhoGEM analysis program is developed around three key data inputs, data preprocessing (genotypes), ProvenancePredictor (GPS coordinates), and phenotypic characterization (quantitative functional traits). The multicriteria approach is used to determine the optimal number of population subdivisions for the plant species under study, to assign an admixture proportion vector to each accession and to characterize/predict quantitative phenotypes. Geographical coordinates are used as covariates to optimize the definition of number of ancestral populations using the ProvenancePredictor algorithm. The relationships between genome components and phenotypes were estimated using linear models with systematic search for the best minimum model using the leaps R package to cope with dependencies among the predictors (the proportions of genome components must sum to one).

Results: The population structure of the circum-Mediterranean model legume *Medicago truncatula* was assessed by ADMIXTURE, DAPC and ProvenancePredictor. We summarized the genome variations by the eight-dimensional vector of admixture components. The results show that admixture components are significantly correlated with climate and geography. Then, the WhoGEM method was used to assess variation in genome admixture proportions, and revealed that it explains most of the phenotypic variation for quantitative disease resistances and quantitative functional traits. The method's prediction accuracy outperforms or equates current algorithms for Genomics Selection (GS). WhoGEM analysis of *M. truncatula* produced the first evidence that quantitative phenotypes can be well-predicted using genome-wide patterns of admixture.

Conclusion: The development of WhoGEM demonstrates that population admixture can integrate the effects of demography (i.e., gene flow and genetic drift) and of natural

selection towards adaptation, and thereby explains more phenotypic variation than GS- or Quantitative Trait Loci (QTL)-based approaches. This insight contributes to understanding of phenotypic variation in evolutionary biology, and can accelerate plant and animal breeding and biomedical research programs by linking phenotypic traits.

Acknowledgements: M. Mazurier was supported by a PhD scholarship from the French "Ministère de la Recherche et de l'Enseignement Supérieur" and a "Visiting Student" fellowship from Toulouse INP for a stay at USC. L. Gentzbittel and C. Ben were supported by a fellowship from Toulouse INP for a stay at USC and a "Visiting Scholar" fellowship from the US Feed the Future Innovation Lab "Climate Resilient Chickpea". T.V. Tatarinova was supported by a "Visiting Scholar" fellowship from Toulouse INP for a stay at Toulouse INP and by the NSF Division of Environmental Biology award No. 145663. 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.

DOI 10.18699/WIBSB-2018-27

Alternative splicing in transcriptomes of glioma cell cultures

N.V. Gubanova[1, 2], A.O. Bragin[2], A.G. Bogomolov[1, 2], S.S. Kovalev[2], Y.L. Orlov[1, 2]*
[1] *Novosibirsk State University, Novosibirsk, Russia*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
** e-mail: orlov@bionet.nsc.ru*

Key words: genomics, transcriptomics, cancer, glioma, RNA-seq, alternative splicing, database

Motivation and Aim: Detection of genes responsible for glioma progression is important biomedical problem. Glioma is the most common malignant tumor in the central nervous system. The identification of sensitive and specific biological markers that would help identify patients at a higher or lower risk of death from glioma is of crucial importance [1]. We plan ollection of such candidate in a specialized database.
Methods and Algorithms: We performed RNA-seq experiments on primary glioma cell cultures. The primary cell culture samples from normal brain and secondary glioblastoma were processed for RNA extraction. This was followed by RNA-sequencing, data processing and filtration of reads. The algorithm of the pipeline consists of several stages: 1) pre-processing of the reads; 2) mapping reads to the reference genome; 3) identification of differential expression and cases of alternative splicing; 4) annotation of obtained results. For assessment of gene expression level and finding differently expressed genes we used Cufflinks on own scripts for data processing. Set of computer tools were used for alternative splicing analysis [2].
Results: Set of differently expressed gene in normal brain and glioma cell cultures and analyzed alternative splicing variants was contracted by experimental data. We found set of hormone transporter genes overexpressed in the glioblastoma cell culture. SLCO1C1 mediates the Na+-independent high affinity transport of organic anions such as the thyroid hormones thyroxine. Multiple alternative splicing transcripts have been identified as progression markers. We found set of differentially alternatively spliced transcripts. 73 of the differentially expressed genes were found in OMIM as related to glioma (of 193 loci) including GLI1, GLI3 (GLI-Kruppel family member 3); GLM4 (GLIOMA SUSCEPTIBILITY 4), GLTSCR1 (Glioma tumor suppressor candidate region gene 1) and others. We have developed database of genes with alternative splicing in glioma including APP, CASC4 and TP53.
Conclusion: The RNA-seq analysis of the cells cultures of normal brain and glioma confirmed association of these genes with tumor progression. We continue work on integration of data on glioma genes using available data sources developing own database.
Acknowledgements: The research has been supported in part by RFBR and by the budget project No. 0324-2018-0019.

References

1. Liu M. et al. (2017) The identification of key genes and pathways in glioma by bioinformatics analysis. Journal of Immunology Research. 2017:Article ID 1278081.
2. Babenko V.N. et al. (2017) Computer analysis of glioma transcriptome profiling: alternative splicing events. J Integr Bioinform. 14(3). pii: /j/jib.2017.14.issue-3/jib-2017-0022/jib-2017-0022.xml

DOI 10.18699/WIBSB-2018-28

Analyzing the genes related to Alzheimer's disease via a nework and pathway-based approach

Y. Hu[1, 2]*, J. Wang[2]**, M. Chen[1]
[1] *Department of Bioinformatics, College of Life Sciences, Zhejiang University, Hangzhou, China*
[2] *School of Biomedical Engineering, Tianjin Medical University, Tianjin, China*
** e-mail: yanshihu@zju.edu.cn; ** wangju@tmu.edu.cn*

Key words: Alzheimer's disease, functional enrichment analysis, network analysis, pathway crosstalk

Motivation and Aim: Our understanding of the molecular mechanisms underlying Alzheimer's disease (AD) remains incomplete. Previous studies have revealed that genetic factors provide a significant contribution to the pathogenesis and development of AD. In the past years, numerous genes implicated in this disease have been identified via genetic association studies on candidate genes or at the genome-wide level. However, in many cases, the roles of these genes and their interactions in AD are still unclear. A comprehensive and systematic analysis focusing on the biological function and interactions of these genes in the context of AD will therefore provide valuable insights to understand the molecular features of the disease.
Methods and Algorithms: In this study, we collected genes potentially associated with AD by screening publications on genetic association studies deposited in PubMed. The major biological themes linked with these genes were then revealed by function and biochemical pathway enrichment analysis, and the relation between the pathways was explored by pathway crosstalk analysis. Furthermore, the network features of these AD-related genes were analyzed in the context of human interactome and an AD-specific network was inferred using the Steiner minimal tree algorithm.
Results: We compiled 430 human genes reported to be associated with AD from 823 publications. Biological theme analysis indicated that the biological processes and biochemical pathways related to neurodevelopment, metabolism, cell growth and/or survival, and immunology were enriched in these genes. Pathway crosstalk analysis then revealed that the significantly enriched pathways could be grouped into three interlinked modules–neuronal and metabolic module, cell growth/survival and neuroendocrine pathway module, and immune response-related module–indicating an AD-specific immune-endocrine-neuronal regulatory network. Furthermore, an AD-specific protein network was inferred and novel genes potentially associated with AD were identified.
Conclusion: By means of network and pathway-based methodology, we explored the pathogenetic mechanism underlying AD at a systems biology level. Results from our work could provide valuable clues for understanding the molecular mechanism underlying AD. In addition, the framework proposed in this study could be used to investigate the pathological molecular network and genes relevant to other complex diseases or phenotypes.
Acknowledgements: Supported in part by grants from the National Key Research and Development Program of China (No. 2016YFC0906300), the National Natural Science Foundation of China (No. 31271411), and the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry of China.

DOI 10.18699/WIBSB-2018-29

Alternative cleavage and polyadenylation sites in colorectal cancer

Y. Kang[1]*, X. Yang[1], W. Xu[1], Z. Mao[1], J. Sun[2]
[1] *School of Biomedical Engineering, Bio-ID Center, Shanghai Jiao Tong University, Shanghai, China*
[2] *Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai, China*
** e-mail: kangyani@sjtu.edu.cn*

Key words: alternative polyadenylation, colorectal cancer, 3T-seq, 3'UTR

Motivation and Aim: Alternative polyadenylation (APA) is an important post-transcriptional regulation in eukaryotic cells. It plays considerable roles in many biological processes and diseases, such as cell differentiation, proliferation and cancer. Colorectal cancer is one of the most common malignancies worldwide, which is among the top five in incidence and mortality of all cancers in China, even though China is not a high prevalence area. Studies have shown that ~70 % of protein-coding genes have conserved microRNA (miRNA) target sites within their 3'UTRs. Moreover, during transformation, the cell starts using the poly (A) sites (PAS) most proximal to the open reading frame (ORF) to generate a short 3'UTR, which makes the mRNA resistant to miRNA by eliminating miRNA-binding sites. Recent studies reported a widespread preferential usage of proximal PAS in cancers, such as breast, lung, liver, and colorectal cancers. Even if, the role of APA in transformation and cancer is still not very clear. Although there have been some studies on the APA of colorectal cancer, the normal and carcinoma samples used for genome-wide profiling were not matched. The purpose of this study was to obtain genes with switched 3' untranslated region (UTR) that may be associated with intracellular regulation of colorectal cancer by analyzing APA patterns of strict control groups from clinical patients.
Results: We used a robust approach, 3T-seq to profile global APA sites in three patients and observed hundreds genes exhibit shortened 3'UTR, and some of them have been reported play a key role in cancer. Overall, we identified 35,076 poly (A) sites in total. Compared to the matched normal tissues, we detected 350, 405 and 375 genes with significantly APA-mediated 3'UTR alteration in cancer tissues of 3 patients, respectively. Forty-seven genes with switched 3'UTR were shared in all three patients. In addition, most of these genes have shortened 3'UTRs, some of which were associated with cancers, such as *GPI*.
Conclusion: Our studies found several genes with switched 3'UTR in colorectal cancer patients, which may provide some important clues for more in-depth study of the cellular regulation in colorectal cancer from the perspective of post-transcriptional regulation. It may also help in the search for new biomarkers of colorectal cancers.
Acknowledgements: This work was supported by National Natural Science Foundation of China (31671299), Meng Minwei foundation, Medical engineering cross fund (YG-2017ZD15, YG2015QN35) and Laboratory Innovative Research Program of Shanghai Jiao Tong University (17SJ-18).

DOI 10.18699/WIBSB-2018-30

Application of bioinformatics for identification of candidate genes conferring tolerance to drought and salinity in chickpea (*Cicer arietinum* L.) in the environment of Kazakhstan

G. Khassanova[1]*, A. Zhubatkanov [1], S. Jatayev[1], A. Turbekova[1], Y. Shavrukov[2]
[1] *S. Seifullin Kazakh AgroTechnical University, Astana, Kazakhstan*
[2] *Flinders University, Adelaide, Australia*
** e-mail: khasanova-gulmira@mail.ru*

Key words: bioinformatics, plant biology, gene expression, chickpea, drought resistance

Introduction and Aim: Chickpea is gaining popularity as a crop in Kazakhstan, with its arid climate and soils with localised, patchy salinity. The two chickpea ecotypes, Desi and Kabuli, differ in various growth and seed production traits and supply diverse markets reflecting differing consumer preferences. Based on the results of field trials, genotypes for both ecotypes were selected for two groups with either high or low seed yield, 1000 seed weight, and lowest pod height. The selected genotypes were used for further bioinformatic and molecular genetic analyses. The aim of this study was to identify and analyse possible candidate genes controlling better tolerance to drought and salinity and test for their association with higher seed yield in selected chickpea genotypes via bioinformatics.
Methods: International and local germplasm collections comprising 250 chickpea samples were tested over two years in field trials in Northern and Central Kazakhstan. Additional experiments with drought and salinity treatments were carried out in greenhouses. Bioinformatics and systems biology methods were applied in this study to identify important candidate genes and their possible involvement in tolerance to abiotic stresses and seed yield in the selected chickpea genotypes.
Results: Eight possible candidate genes encoding various structural and regulatory polypeptides were identified in chickpea using bioinformatic analyses with comparisons to other plant species. Expression analyses revealed the maximum level of expression in *CaRABC1*, which controls cell membrane trafficking and is associated with better adaptation of plants to abiotic stress conditions. Different levels of expression were found between the two groups of chickpea genotypes with high and low seed yield, the two ecotypes and following exposure to drought or salinity stress. *CaRABC1* expression seems to be regulated by Transcription factor *CaDREB2*, but this hypothesis will need to be verified by further molecular experiments.
Conclusion: The candidate gene *CaRABC1*, possibly regulated by *CaDREB2*, was identified with the use of bioinformatic and molecular genetic analyses. The genetic polymorphism of these candidate genes can be used for molecular breeding of chickpea genotypes with high seed yield and better adaptation to the environment of Kazakhstan.
Acknowledgements: This study was supported by the Ministry of Education and Science. Kazakhstan, Research program BR05236500.

DOI 10.18699/WIBSB-2018-31

Widely-expressed and conserved long noncoding RNAs LINC00493 and LINC01420 influence on cell physiology

D.O. Konina[1], P.A. Sparber[2], A.Yu. Filatova[2], I.A. Krivosheeva[2], M.Yu. Skoblov[1, 2, 3]
[1] *Department of Biological and Medical Physics, Moscow Institute of Physics and Technology (State University), Dolgoprudny, Moscow region, Russia*
[2] *Research Centre of Medical Genetics, Moscow, Russia*
[3] *Far Eastern Federal University, Vladivostok, Russia*
* *e-mail: darya.konina@phystech.edu*

Key words: Long noncoding RNA, RACE, MTT assay

Motivation and Aim: According to the date of GENCODE project (v 24), 15,767 long non-coding RNA (lncRNA) genes were implicated in humans. But functions have been revealed only for a small part of them. Long noncoding RNAs can act as regulators at many different levels of gene expression, including chromatin organization, transcriptional regulation and posttranscriptional control. Moreover, lncRNAs play important roles in disease. To date, it has been confirmed that there are about 3000 lncRNA-disease entries in accordance with lncRNADisease database. Improving our knowledge about lncRNAs will lead to more complete understanding of the mechanisms of biological processes at the molecular level in the body and help to optimize and systematize the diagnosis of hereditary human diseases, as well as develop approaches for their therapy. This work is devoted to the determination of the structures and functions of the new long non-coding RNAs LINC01420 and LINC00493.
Methods: Nucleotide sequences of the studied genes were found in following databases: RefSeq, Ensemble, GENCODE. Conservation and expression level in various human cell lines and tissues were analyzed using data from the UCSC genomic browser. The reverse transcription reaction, PCR, RACE-PCR was used to determine the structures of the RNAs. To study the expression profile, RT-q PCR was performed on RNA samples from various human cell lines. To define functions of genes, we investigated its subcellular localization by using soft lysis and analyzing cytoplasmic compartment, nucleoplasm and chromatin fraction of RNA. In addition, we performed effective knockdown of lncRNAs using lipofection with 2 siRNAs for each gene in A375 cell line. Then, we observed the variation of cell proliferation by MTT assay. To study cell migration *in vitro* wound-healing assay was performed.
Results: Bioinformatic analysis showed that these LINCs are highly expressed in many human cell lines and tissues. The profile of the distribution of LINCs expression in human tissues allows to attribute these genes to the group of "housekeeping genes". RT-PCR analysis revealed that LINC01420 has 3 exons and 2 isoforms: short-major, long-minor. LINC00493 has 3 exons and 1 main isoform. We determined the exact 5′ and 3′-ends of the studied RNAs by RACE. We established that the transcripts of LINCs have cytoplasmic localization. Besides, we performed LINCs knockdown experiments followed by proliferation and migration assays. Results of MTT-assay indicated that the cell proliferation was attenuated after *LINC00493* knockdown, while downregulation of *LINC01420* didn't affect cell viability. On the other hand, wound healing assay showed that the migration ability was decreased with knockdown of studied genes.
Conclusion: In our study, the exact structure of two lncRNAs was obtained. We revealed main LINC's isoforms and number of exons. We also received data on their subcellular localization, the effect on viability and migration of cells. Thus, we have come closer to understanding function and mechanism of action of the studied genes.

DOI 10.18699/WIBSB-2018-32

Modeling the glycolysis dynamics

A. Kotipalli[1], C.K. Mitra[2]*
[1] *Centre for Development of Advanced Computing, Pune, India*
[2] *University of Hyderabad, Hyderabad, India*
** e-mail: c_mitra@yahoo.com*

Key words: metabolic dynamics, kinetics model, reaction pathway, systems biology, glycolysis

Motivation and Aim: Glucose concentration must be regulated relatively tightly within the cell. In the glycolytic pathway, the sequence of biochemical reactions is regulated at several places: early at hexokinase, somewhere in the middle during the conversion of fructose-6-phosphate to fructose bis-phosphate by phosphofructose kinase and in the late stages the concentration of pyruvate is regulated by pyruvate kinase. Although studied as a stand-alone pathway, glycolysis is tightly coupled with the TCA cycle, the pentose phosphate pathway and the glycogen metabolism pathway. Different cells have their own metabolic profiles and several cell types are also under external hormonal controls.
Methods and Algorithms: In this study we have used a simulation (based on Octave) and used an extension of Michealis-Menten equation for the reaction kinetics. We also used an (empirical) extension of the Quasi-Steady-State (QSS) approximation to reduce the effective number of parameters. In the current simulation, we have not considered the coupling with other networks nor the effect of the gene regulation on glycolysis. We have kept the enzyme concentrations (molar numbers) equal for all the enzymes in the pathway. The substrate concentrations were often expressed in terms of the relevant K_M (the dissociation constant of the enzyme substrate complex) to reduce the apparent complexity. The final set of differential equations were solved by Octave [1, 2].
Results: The results of the present simulations are generally in line with expected results and we do not expect an accurate match with the cellular level concentrations of the different metabolites.
Conclusion: Although the methodology uses approximations and the quantitative effects of the approximations have not been studied here, we believe the tools presented in this study are effective in general understanding of the roles played by different regulators. Additional cellular level (genetic level) or external induced (hormonal effects) have not been considered.
Acknowledgements: Financial Support from the University Grants Commission (https://www.ugc.ac.in) to one of the authors (CKM) is gratefully acknowledged (2015–2017).

References

1. Korla K., Mitra C.K. (2014) Modeling the Krebs cycle and oxidative phosphorylation. J. Biomol. Struct. Dyn. 32(2):242-256.
2. Korla K., Mitra C.K. (2014) Kinetic modeling of coupled transport across biological membranes. Ind. J. Biochem. Biophys. 51(2):93-99.

DOI 10.18699/WIBSB-2018-33

Aging of rat retina: transcriptome study

O.S. Kozhevnikova*, D.V. Telegina, N.G. Kolosova
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
** e-mail: oidopova@bionet.nsc.ru*

Key words: aging, retinopathy, senescent-accelerated OXYS rats

Motivation and Aim: The aging process is the major risk factor for age-related diseases, including age-related macular degeneration (AMD). Age-related macular degeneration (AMD) is a complex disease leading to visual dysfunction through a variety of mechanisms and complications. The senescence-accelerated OXYS rats develop retinopathy with clinical and morphological manifestations similar to AMD. Clinical, histological and molecular manifestations of retinopathy in OXYS rats, the results of identification quantitative trait loci (QTLs) and transcriptome studies will be presented. OXYS rats first demonstrated fundoscopic changes at 1.5 months of age with the appearance of atrophic areas in the RPE and choriocapillaris. Older rats developed thickening of Bruch's membrane, drusen and RPE detachments. By 12 months, some animals demonstrated photoreceptor atrophy, decreased ERGs, destruction of the choriocapillaris with fibrosis and in some cases hemorrhagic detachment of the retina due to neovascularization.
Methods: In order to identify the impact of aging process on in response to normal aging and progression of AMD-like retinopathy, we compared gene expression profiles of retina from young and old OXYS and control Wistar rats by RNA sequencing (RNA-Seq).
Results: The majority of DE genes are related to the immune system and extracellular matrix turnover. Little age-regulated genes were common for the two strains, suggestive of different rates and mechanisms of aging. Hundreds genes showed significant differences in expression between the two strains. These genes are involved in disease-associated pathways such as immune response, inflammation, apoptosis, Ca^{2+} homeostasis, and oxidative stress. We can conclude that the development of retinopathy in OXYS rats is associated with an imbalance in immune and inflammatory responses. Aging had significant effects on the expression of inflammatory genes but their composition was different in the retina of OXYS and Wistar rats. The retinopathy development in OXYS rats accompanied by downregulation of immune response genes in the retina. This indicates that any disturbances in immune defenses can accompany retinal disease, not only upregulation, but also downregulation, which can be explained within the framework of immunosenescence theory. Our data support the view that the genetic background has a profound impact on AMD development.
Acknowledgements: Supported by the RSF No. 18-75-00031.

DOI 10.18699/WIBSB-2018-33

Study of essential hypomutated genes in skin melanoma

I. Krivosheeva[1], A. Filatova[1], M. Pyatnitskiy[2, 3], S. Moshkovskii[2, 4], M. Skoblov[1, 5, 6]*
[1] *Research Centre of Medical Genetics, Moscow, Russia*
[2] *nstitute of Biomedical Chemistry, Moscow, Russia*
[3] *Higher School of Economics, Moscow, Russia*
[4] *Pirogov Russian National Research Medical University, Moscow, Russia*
[5] *Department of Biological and Medical Physics, Moscow Institute of Physics and Technology (State University), Dolgoprudny, Russia*
[6] *Far Eastern Federal University, Vladivostok, Russia*
** e-mail: mskoblov@gmail.com*

Key words: medical genomics, cancer, melanoma, knockdown

Motivation and Aim: Current methods of studying tumors are aimed at finding new proto-oncogenes, tumor suppressors, markers of cancer, etc. Information about the origination of the tumor and its genetic causes may be useful in the diagnostics, but its use for cancer therapy is limited. The searching for the genes essential for cancer cell surviving could help to develop effective cancer therapy.
Methods: For *in vitro* analysis we used melanoma cell line A375. We performed knockdown (KD) by lipofection with siRNAs, assessed the efficiency of KD by RT-qPCR. Cell proliferation status was analyzed by MTT-test and flow cytometry. Induction of apoptosis was measured by Caspase-Glo® 3/7 Assay System. Cell migration was assessed by wound-healing assay.
Results: Using the principle named after Abraham Wald our colleagues conducted a bioinformatic analysis of mutations in the genome of the melanoma tumors and identified 91 genes not affected by considerable mutations (1). 12 out of 91 genes were selected for further analysis. We analyzed the expression level of these genes in melanoma cell line A375 by RT-qPCR. For three genes with highest expression level (*UNC45A*, *RHPN2* and *ZNFx1*) we performed knockdown (KD) experiments on the A375 cell line followed by proliferation, apoptosis and migration assays. We achieved KD efficiency for the *UNC45A* gene about 70 %, for *RHPN2* – 55 %, and for *ZNFx1* – 60 %. Caspase activity raised up to 150 %, 125 % and 200 % comparing to control, respectively. Whereas proliferation assay did not reveal significant changes. Data of the migration of the cells is in progress.
Conclusion: Our results demonstrate that knockdown of chosen genes do affect caspase activity of the melanoma cells, but do not cause changes in viability of the cells. Further investigations are needed to fully confirm these results.

References

1. Pyatnitskiy M., Karpov D., Poverennaya E., Lisitsa A., Moshkovskii S. (2015) Bringing down cancer aircraft: searching for essential hypomutated proteins in skin melanoma. Chammas R., ed. PLoS ONE. 10(11):e0142819. DOI 10.1371/journal.pone.0142819.

DOI 10.18699/WIBSB-2018-34

In silico approach of finding inhibitor of Nedd8-activating enzymes for cancers

P. Kumar*, S. Piramanayagam
Department of Bioinformatics, Bharathiar University, Coimbatore, India
** e-mail:sirajane@gmail.com*

Key words: NEDD8, pharmacophore, docking, Schrodinger, Symyx draw

Motivation and Aim: Cancer normally refers to diseases in which abnormal cells will divide without control and it can able to invade the other tissues. NEDD8 (Neural Precursor Cell Expressed Developmentally Downregulated Protein 8)-Activating Enzyme is an essential component of the NEDD8 conjugation pathway that controls the activity of the cullin RING subtype of ubiquitin ligases, which regulates the turnover of a subset of proteins upstream of the proteosome. This work is focused on identifying the inhibitors for NEDD8 with already identified first-class inhibitor MLN4924 [1].
Methods and Algorithms: The pharmacophore studies and ADME-Tox analysis was used to screen the compounds to find out the optimal lead molecules. The molecular docking studies was carried out to find out the molecular binding affinity of protein and ligand molecules.
Results: In the first phase 200 ligands were generated and evaluated by using Ramachandran plot. ADME-Tox was carried out and found 100 molecules had toxic substances. Finally 18 lead molecules was generated as a potent ligands with druggable properties.
Conclusion: This work was carried out based on Fragment Based Drug Design. Based on the Pharmacophore analysis and ADMET studies, the optimal lead molecules were identified. These ligands were compared with already available anti cancer drugs. The NEDD8-Activating enzyme lead molecules are generated by using Fragment Based Drug Design and it's active against angiogenesis
Acknowledgements: I acknowledge UGC-RGNF and DBT-BIF, New Delhi for providing computational facilities to carried out this research work.

References

1. Qiang Xu et al. (2017) MLN4924 neddylation inhibitor promotes cell death in paclitaxel-resistant human lung adenocarcinoma cells. Oncol. Lett. 15(1):515-521.

DOI 10.18699/WIBSB-2018-35

Knowledge mining from large scale protein-protein interaction datasets at the era of big data

D. Li
State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences (*Beijing*)*, Beijing Institute of Life Omics, Beijing, China*
e-mail: lidong.bprc@foxmail.com

Key words: proteomics, protein-protein interaction, big data, data mining, knowledge discovery

Motivation and Aim: Throughout the history of natural science, it is definite that our knowledge and disciplines come from accumulated discoveries, which are triggered by the unprecedented scale and speed of big data and achieved by efficient mathematical strategies, taking the discovery of Kepler's laws as an example. In our laboratories, simple machine learning strategies, such as Naïve Bayesian network, have been used to find the instinct features of proteome organization, especially the protein interactions [1–4].
Methods and Algorithms: By using naïve Bayesian network, reliability was assigned to the human protein-protein interactions identified by high throughput experiments by combining multiple heterogeneous biological evidences. Then domain enrichment ratio was introduced to measure the direction between interacting proteins, resulting in an integrated human directional protein interaction network. Next, logistic regression was taken to integrate six representative features, to develop a proteome-wide prediction model of self-interacting proteins.
Results and Conclusion: Recently, we developed a naïve Bayesian classifier to combine multiple heterogeneous biological evidences to investigate the human E3-substrate interactions which determine the high specificity of ubiquitination. UbiBrowser is now provided as an integrated bioinformatics platform to predict and present the proteome-wide human E3-substrate interaction network.
It is believed that the era of big data will bring in new insights in life sciences and present new opportunities in research. Artificial intelligence strategy will play dominant roles in the coming knowledge discovery. Our colleagues are now engaged to develop an automatic knowledge discovery highway (ProDiGy), integrating the OMICS datasets and literature information into a biomedical knowledge graph (a heterogamous information network) followed by the feature extraction and deep learning for grand knowledge.

References

1. Li Y., Xie P., Lu L., Wang J., Diao L., Liu Z., Guo F., He Y., Liu Y., Huang Q., Liang H., Li D., He F. (2017) An integrated bioinformatics platform for investigating the human E3 ubiquitin ligase-substrate interaction network. Nat. Commun. 8:347.
2. Li D., Liu W., Liu Z., Wang J., Liu Q., Zhu Y., He F. (2008) PRINCESS: a PRotein INteraction Confidence Evaluation System with multiple data Sources. Mol. Cell. Proteomics. 7:1043-1052.
3. Liu Z., Guo F., Zhang J., Wang J., Lu L., Li D., He F. (2013) Proteome-wide prediction of self-interacting proteins based on multiple properties. Mol. Cell. Proteomics. 12(6):1689-1700.
4. Liu W., Li D., Wang J., Xie H., Zhu Y., He F. (2009) Proteome-wide prediction of signal flow direction in protein interaction network based on interacting domains. Mol. Cell. Proteomics. 8(9):2063-2070.

DOI 10.18699/WIBSB-2018-59

Digestion-Ligation-Only Hi-C, a simple, cost-effective, and highly efficient method for chromosome conformation capture

D. Lin[1#], P. Hong[2#], G. Li[2]*, G. Cao[1]*

[1] *State Key Laboratory of Agricultural Microbiology, Huazhong Agricultural University, Wuhan, China*

[2] *National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, China*

** e-mail: guoliang.li@mail.hzau.edu.cn* (*G.L.*), *gcao@mail.hzau.edu.cn* (*G.C.*)

Key words: genomics, sequencing technology, chromosome conformation, Hi-C

Chromosome conformation capture technologies open an avenue to investigate the three-dimensional (3D) structures of genomes. However, high noise, high cost, and lack of straightforward noise evaluation in current methods impede the advancement of 3D genomic research. Here, we developed a simple Digestion-Ligation-Only Hi-C (DLO Hi-C) technology to explore the 3D landscape of the genome. This method requires only two rounds of digestion and ligation without biotin-labeling and pull-down for reducing the cost. The noise DNA was efficiently removed in a cost-effective step by purifying specific linker-ligated DNA fragments. Notably, random ligation could be quickly evaluated in an early quality-control step before sequencing. Moreover, we performed an *in situ* version of DLO Hi-C method based on 4-cutter restriction enzyme. We applied DLO Hi-C to delineate the genomic architecture of THP-1 and K562 cells and uncovered chromosome translocations. This technology may facilitate the investigation of genomic organization, gene regulation, and (meta-) genome assembly.

# These authors contributed equally to this work.

DOI 10.18699/WIBSB-2018-36

Analysis of possible sequence aligner artefacts using novel read density distribution

F.M. Naumenko[1]*, I.I. Abnizova[2, 3], N. Beka[4], Y.L. Orlov[1, 5, 6]**
[1] *Novosibirsk State University, Novosibirsk, Russia*
[2] *Wellcome Trust Sanger Institute, Cambridge, UK*
[3] *Babraham Institute, Cambridge, UK*
[4] *University of Hertfordshire, Hertfordshire, UK*
[5] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[6] *Institute of Marine Biology Researches of the RAS, Sevastopol, Russia*
** e-mail: fedor.naumenko@gmail.com; ** orlov@bionet.nsc.ru*

Key words: next-generation sequencing, DNA alignment, read density distribution

Motivation and Aim: The use of artificial data to evaluate the performance of aligners and peak callers not only improves its accuracy and reliability, but also makes it possible to reduce the computational time. One of the natural ways to achieve such time reduction is by mapping a single chromosome.
Methods and Algorithms: Using own program scripts we investigated whether a single chromosome mapping causes any artefacts in the alignments' performances [1]. We applied our benchmarking tests on 7 open source DNA sequencing mapping tools, namely Bowtie (1.1.1), Bowtie2 (2.2.4), BWA (0.7.5 and 0.7.12 applying two algorithms), MAQ (0.7.1), MOSAIK (2.2.3), SMALT (0.7.6).
Results: In this paper, we compared the accuracy of the performance of seven aligners on well-controlled simulated benchmark data which was sampled from a single chromosome and also from a whole genome. The generation of artificial data by mapping of reads generated from a single chromosome to a reference chromosome is justified from the point of view of reducing the benchmarking time. The proposed quality assessment method allows to identify the inherent shortcoming of aligners that are not detected by conventional statistical methods, and can affect the quality of alignment of real data.
Conclusion: We found that commonly used statistical methods are insufficient to evaluate an aligner performance, and applied a novel measure of a read density distribution similarity, which allowed to reveal artefacts in aligners' performances. We also calculated some interesting mismatch statistics, and constructed mismatch frequency distributions along the read. The approach could be used for analysis of transcriptomics data in plants [2].
Acknowledgements: The research has been supported by RFBR 18-04-00483. Computing done at Siberian Supercomputer center SB RAS was supported by budget project 0324-2018-0017.

References

1. Naumenko F.M. et al. (2018) Novel read density distribution score shows possible aligner artefacts, when mapping a single chromosome. BMC Genomics. 19(Suppl 3):92.
2. Wang J. et al. (2017) Non-coding RNAs and their roles in stress response in plants. Genomics Proteomics Bioinformatics. 15(5):301-312.

DOI 10.18699/WIBSB-2018-37

The optimal feedbacks in the mathematical model of chemotherapy for a nonmonotonic therapy function

N. Novoselova
Krasovskii Institute of Mathematics and Mechanics, Ural Branch of the RAS, Yekaterinburg, Russia
Ural Federal University, Yekaterinburg, Russia
e-mail: n.g.novoselova@gmail.com

Key words: therapy function, optimal control, optimal feedbacks, Hamilton-Jacobi-Bellman equation, Cauchy method of characteristics, Rankin-Hugoniot line

Motivation and Aim: We investigate a pharmacokinetic problem for a deterministic nonlinear system with piecewise monotonic dynamics describing the process of chemotherapy of a malignant tumor. We consider the case when the therapy function, which describes the effect of the drug on the cell growth rate, has two maxima.
Methods and Algorithms: The work presents results of numerical calculation of the optimal result (the value function) and optimal positional strategy of therapy (optimal feedbacks) in a corresponding optimal control problem. The construction use the fact that the value function is the unique minimax (viscosity) solution [1, 2] of the Cauchy problem for the basic Hamilton–Jacobi–Bellman (HJB) equation. By means of the continuous gluing of a finite number of smooth functions obtained by the Cauchy method of characteristics for auxiliary linear HJB equations, the continuous function ϕ is constructed. The paper [3] proves that the constructed function ϕ coincides with the value function.
Results: A new element of the construction is the construction of a line of nonsmooth gluing using the Rankin-Hugoniot conditions [4, 5]. This line plays a key role for the optimal feedback strategy, because it determines its discontinuity line. The results of numerical calculations of the Rankin-Hugoniot line are exposed. Comparison with the results for the case of a single maximum in the therapy function in this model [6] is given.
Acknowledgements: Supported by the RFBR (No. 17-01-00074).

References

1. Crandall M.G., Evans L.C., Lions P.-L. (1984) Some properties of viscosity solutions of Hamilton-Jacobi equations, Trans. Amer. Math. Soc. 282(2):487-502.
2. Subbotin A.I. (1995) Generalized Solutions of First-Order PDEs: The Dynamical Optimization Perspective. Birkhauser: Boston.
3. Subbotina N.N., Novoselova N.G. (2017) Optimal result in a control problem with piecewise monotone dynamics. Journal "Proceedings of the Institute of Mathematics and Mechanics". 23(4):265-280.
4. Goritskiy A.Yu., Kruzhkov S.N., Chechkin G.A. (1999) First-Order Partial Differential Equations (Study Guide). Moscow: Lomonosov Moscow State University, Faculty of Mechanics and Mathematics.
5. Subbotina N.N., Kolpakova E.A., Tokmantsev T.B., Shagalova L.G. (2013) The Method of Characteristics for Hamilton Jacobi Bellman equations. RIO UrO RAN, Yekaterinburg.
6. Chumerina E.S. (2009) Choice of optimal strategy of tumor chemotherapy in Gompertz model. Journal Computer Systems Sciences International. 48(2):325-331.

DOI 10.18699/WIBSB-2018-38

T7-like bacteriophage promoters: stress-induced duplex destabilization suggests a role in replication

M. Orlov*, A.Ryasik, E. Zykova, A. Sorokin
Institute of Cell Biophysics SB RAS, Pushchino, Russia
** e-mail: orlovmikhailanat@gmail.com*

Key words: bacteriophage T7, SIDD, DNA physics, genomics, replication

Motivation and Aim: The fact that replication and transcription are coupled in prokaryotes has been known for many years, but a solid mechanistic explanation for it is lacking. T7-like bacteriophages are well-suited model systems to study the phenomenon since their small genomes contain both multiple promoters (being recognized by two different RNA-polymerases) and replication origins. Moreover, appreciable fraction of the promoter regions is known to be involved in transcription initiation [1].
Methods and Algorithms: DNA-protein interactions that govern various regulatory events are due to physicochemical properties of the molecules. Among them, one of the most important is stress-induced DNA duplex destabilization (SIDD). Sites predicted to be susceptible to duplex destabilization are found to be significantly associated with regulatory regions. SIDD profiles are calculated by means of statistical mechanics method evaluating the equilibrium distribution among states of denaturation of a DNA molecule under a defined level of superhelicity imposed [2]. Here we report analysis of genome-wide SIDD profiles for T7-like bacteriophages with the respect to regulatory site locations.
Results: Among T7-DNA native promoters ones functioning as secondary replication origins were found to be highly destabilized. The latter includes *phiOL* and *phiOR* located at both termini of several T7-like bacteriophage DNA (including T7). In all cases the two promoters were shown to coincide with highly destabilized duplex regions. For comparison purposes, selected T7-like genomes that lack *phiOL* and *phiOR* were additionally considered. In some cases sharp SIDD profile maxima were found to be located near one or both DNA termini, which implies presence of promoters analogous to *phiOL* and *phiOR*. Similar promoters might also be present in other bacteriophages.
Conclusion: The connection between promoter SIDD as a DNA duplex feature and specific involvement in replication has been evidenced for T7-like bacteriophage promoters involved in initiation of replication. This provides additional insight into mechanisms underlying link between transcription and replication and highlights importance of DNA duplex stability.

References

1. Wang H.Q., Benham C.J. (2006) Promoter prediction and annotation of microbial genomes based on DNA sequence and structural responses to superhelical stress. BMC Bioinformatics. 7:248-262.
2. Dunn J.J., Studier F.W., Gottesman M. (1983) Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. J. Mol. Biol. 166:477-535.

DOI 10.18699/WIBSB-2018-39

Systems biology research at the SBB young scientists school series

Y.L. Orlov[1, 2]*, L. Wong[3], T.V. Tatarinova[4], R. Hofestädt[5], N.A. Kolchanov[1]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *National University of Singapore* (*NUS*)*, Singapore*
[4] *University of La Verne, USA*
[5] *Bielefeld University, Bielefeld, Germany*
** e-mail: orlov@bionet.nsc.ru*

Key words: education, bioinformatics, international conferences, BGRS conference series

The Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences (ICG SB RAS) hosts the International Multi-conference on Bioinformatics of Genome Regulation and Structure\Systems Biology (BGRS\SB) every two years, beginning from 1998. From BGRS\SB'2008 onwards, the Young Scientists School on Systems Biology and Bioinformatics (SBB) runs as a satellite event following the BGRS\SB conference or as a standalone annual event. Since the first meeting, the SBB has grown into a large international event. Gradually, the initial focus has been extended from systems biology and classical bioinformatics topics to gene network analysis and reconstruction, and omics technologies [2].

The Journal of Bioinformatics and Computational Biology (JBCB) publishes special issues on bioinformatics, algorithms, network analysis dedicated to BGRS\SB. The first JBCB special issue in 2006 highlighted BGRS\SB-2006 [3]. Then JBCB published special issues on the 2012, 2014, and 2016 conferences [4–6]. Additionally, the journal publishes reports from the SBB schools. For instance, JBCB has published proceedings of SBB-2015 on modeling of gene network based on material presented at earlier BGRS\SB meetings [3, 5, 6].

The Institute of Cytology and Genetics SB RAS organized in 2017 the Belyaev Conference-2017 on genetics and evolution, dedicated to the 100th anniversary of Academician, Professor Dmitry K. Belyaev (1917–1985), an outstanding scientist, evolutionist and geneticist [7]. The works reported at Belyaev Conference-2017 and SBB'2017 School on Computational Biology were recently covered in the special issues of several international journals: BMC Evolutionary Biology [7], BMC Genetics [8], BMC Plant Biology, BMC Genomics, BMC Neuroscience, Vavilov Journal of Selection and Breeding (http://vavilov.elpub.ru/jour/issue/view/32/showToc). Meanwhile, a recent JBCB special issue presents materials from SBB'2017 School on Systems Biology and Bioinformatics (SBB'2017), organized in June 2017 in Yalta, Russia [1].

Integration of bioinformatics data is central point of the meetings and Schools organized in Novosibirsk. Journal of Integrative bioinformatics (https://www.degruyter.com/view/j/jib) collects papers on Integration of Data, Methods and Tools, Big Data [9].

Overall, we continue the organization of special journal issues collected works on bioinformatics and biomedical applications presented by young scientists, thus extending traditions of post-conference special journal issues of the BGRS\SB conference series. This year's series will highlight materials presented at BGRS\SB-2018 multiconference; special issues are planned at several BMC journals, and Frontiers in Genetics, and JBCB. The JBCB special issue on BGRS-2018 is currently open for submission at the journal

web-site; it is our pleasure to invite authors to submit materials to the JBCB BGRS-2018 special issue.

Acknowledgements: We shall acknowledge all the international reviewers who helped make JBCB special issues better, improve the science and presentation of the materials of the authors. The organization of education round table at BGRS-2018 multiconference has been supported by Russian Ministry of Science project 28.12487.2018/12.1.

References

1. Orlov Y.L., Tatarinova T.V., Zakhartsev M.V., Kolchanov N.A. (2018) Introduction to the 9th Young Scientists School on Systems Biology and Bioinformatics (SBB'2017). Journal Bioinformatics Computational Biology. 16(1):1802001 (5 pages). DOI 10.1142/S0219720018020018.
2. Baranova A.V., Orlov Y.L. (2016) The papers presented at 7th Young Scientists School "Systems Biology and Bioinformatics" (SBB'15): Introductory Note. BMC Genetics. 17(Suppl 1):20.
3. Kolchanov N.A., Hofestädt R., Orlov Y.L. (2007) Introduction to the selected papers of BGRS'06. Journal Bioinformatics Computational Biology. 5(2b):vii-viii.
4. Kolchanov N.A., Orlov Y.L., (2013) Introductory note for BGRS-2012 special issue. Journal Bioinformatics Computational Biology. 11(1):1302001.
5. Orlov Y.L., Hofestädt R., Kolchanov N.A. (2015) Introductory note for BGRS\SB-2014 special issue. Journal of Bioinformatics and Computational Biology. 13(1):1502001.
6. Orlov Y.L., Kolchanov N.A., Hofestädt R., Wong L. (2017) Editorial – Bioinformatics development at the BGRS\SB conference series: 10th anniversary. Journal Bioinformatics Computational Biology. 15(2):1702001.
7. Orlov Y.L., Baranova A.V., Herbeck Y.E. (2017) Evolutionary Biology at Belyaev Conference-2017. BMC Evol Biol. 17(Suppl 2):260.
8. Orlov Y.L., Baranova A.V., Tatarinova T.V., Kolchanov N.A. (2017) Genetics at Belyaev Conference-2017: introductory note. BMC Genetics. 18(Suppl 1):116.
9. Chen M., Harrison A., Shanahan H., Orlov Y. (2017) Biological Big Bytes: Integrative Analysis of Large Biological Datasets. J Integr Bioinform. 14(3). pii: /j/jib.2017.14.issue-3/jib-2017-0052/jib-2017-0052.xml.

DOI 10.18699/WIBSB-2018-40

Evolutionary and genetic features of drug targets

Y. Quan[1], Z.Y. Wang[1, 2], X.Y. Chu[1], H.Y. Zhang[1]*
[1] *Hubei Key Laboratory of Agricultural Bioinformatics, College of Informatics, Huazhong Agricultural University, Wuhan, P. R. China*
[2] *Center for Molecular Biology of the University of Heidelberg* (*ZMBH*)*, Heidelberg, Germany*
** e-mail: zhy630@mail.hzau.edu.cn*

Key words: drug discovery, drug target, genetics, evolutionary biology

Motivation and Aim: In the modern drug discovery pipeline, identification of novel drug targets is a critical step. Despite rapid progress in developing biomedical techniques, it is still a great challenge to find promising new targets from the ample space of human genes. This fact is partially responsible for the situation of "more investments, fewer drugs" in the pharmaceutical industry. To escape this predicament, novel concepts should be introduced in this field.

A series of recent researches revealed that successfully targeted genes share some common evolutionary and genetic features [1–3], which means that the knowledge accumulated in modern evolutionary biology and genetics is very helpful to identify potential drug targets and to find new drugs as well. In this article, we comprehensively summarize the links between human drug targets and genetic diseases and their evolutionary origins, with an attempt to introduce these novel concepts and their medical implications to the biomedical community.

Results: Genetics-derived disease genes, especially those strongly linked with disease phenotypes, are a rich source of drug targets. Besides, the genetic knowledge about the pathogenesis, that is, whether the diseases arise from loss-of-function (LOF) or gain-of-function (GOF) mutations of the responsible genes, provides critical information about drug mode of action. More interestingly, the evolutionary background of human genes is also valuable in drug target identification. The genes originated in certain evolutionary stages are more druggable and ohnolog genes have higher probability to serve as drug targets, because of their sensitivity to dosage variance.

Conclusion: It seems that during the paradigm shift of pharmaceutical industry in this omics era, modern genetics and evolutionary biology will play a more and more important role.

Acknowledgements: Supported by the Fundamental Research Funds for the Central Universities (Grants No. 2662015PY004 and No. 2662017PY115).

References

1. Wang Z.Y., Fu L.Y., Zhang H.Y. (2012) Can medical genetics and evolutionary biology inspire drug target identification? Trends Mol. Med. 18(2):69-71.
2. Sun H.Y., Hou T.J., Zhang H.Y. (2014) Finding chemical drugs for genetic diseases. Drug Discov. Today. 19(12):1836-1840.
3. Kamb A., Harper S., Stefansson K. (2013) Human genetics as a foundation for innovative drug development. Nat. Biotechnol. 31(11):975-978.

DOI 10.18699/WIBSB-2018-41

Molecular investigation of mutation in PncA gene for pyrazinamide resistance in *Mycobacterium tuberculosis*: an *in silico* approach

D. Ruban Durairaj*, P. Shanmughavel
Department of Bioinformatics, Bharathiar University, Coimbatore, India
** e-mail: ruban.bioinfo@gmail.com*

Key words: pyrazinamide, PncA gene, mutations, *Mycobacterium tuberculosis*, *in silico*

Motivation and Aim: PncA is a gene coding for the enzyme pyrazinamidase in Mycobacterium [1].The pyrazinamidase enzyme converts the pyrazinamide into its active form pyrazinoic acid [2]. There is significant relation between the mutations in PncA gene and the pyrazinamide resistance of *M. tuberculosis*. There are several mutations in PncA gene identified and reported all over the world from the year 1996 to till date [2]. The main aim of this study is to investigate those mutations at molecular level by applying *in silico* approaches.
Methods and Algorithms: We modelled the structure of the enzyme pyrazinamidase encoded by PncA gene, with I-Tasser tool. The modelled structure is validated by using Ramachandran plot. Then the modelled structure is subjected to molecular docking with the pyrazinamide drug by utilizing Schrodinger Maestro software. Then the apo structure and docked complex were subjected to molecular dynamics simulation by using GROMACS software package.
Results: The mutation at the nucleotide position 202, mutates the residue W in wild protein to R in mutant which is disfavoured substitution. This may leads to the pyrazinamide resistance. In molecular docking study we identified that the pyrazinamide drug is binding well with the wild protein whereas it lacks its binding activity in the mutant protein. The binding scores are –4.34 in wild protein and –3.21 in mutant protein. The stability of the complexes were further checked with molecular dynamics simulation studies.
Conclusion: The results obtained in this study suggesting that the mutation at the nucleotide position 202 leads to the pyrazinamide resistance in Mycobacterium tuberculosis. Further experimental investigation at that mutation will give us fruitful results in future. Hence we are suggesting that the PncA mutant protein is resistance to pyrazinamide drug, we can consider it as active drug target for finding the drug alternate of pyrazinamide to treat tuberculosis.
Acknowledgements: We acknowledge DBT-Centre for Bioinformatics, Bharathiar University, Coimbatore, Tamilnadu, India for providing all the computational facilities to carry out this work.

References

1. Whitfield M.G., Soeters H.M., Warren R.M., York T., Sampson S.L., Streicher E.M., van Helden P.D., van Rie A. (2015) A global perspective on pyrazinamide resistance: systematic review and meta-analysis. PLoS One. 10(7):e0133869. DOI 10.1371/journal.pone.0133869.
2. Juréen P., Werngren J., Toro J.-C., Hoffner S. (2016) Pyrazinamide resistance and pncA gene mutations in Mycobacterium tuberculosis. Antimicrob. Agents Chemother. 52(5):1852-1854. DOI 10.1128/AAC.00110-08.

DOI 10.18699/WIBSB-2018-42

Development of National Information System (NIS) for sustainable management of biodiversity

P. Shanmughavel*, P. Kumar
Department of Bioinformatics, Bharathiar University, Coimbatore, India
** e-mail: shanmughavel@buc.edu.in*

Key words: National Information System (NIS), document and conserve biodiversity, sustainable management of biological resources

Motivation and Aim: The need to document and conserve biodiversity has become a necessity in the wake of increased threats from deforestation, alteration in land use, soil degradation, pollution and climatic change. Current electronic inventories provide limited information on biodiversity, its environmental status, bioclimatic, evolutionary history and management units of conservation [1, 2]. This paper explains a system that provides species-specific and site-specific information with improved genetic data semantics to help sustainable management of biological resources. The functions and features of NBGIS are also given.
Methods: NIS can be developed by the authors provides detailed information on bio resources and their environment. The approach employed is to map independently the environmental attributes of a biological species and its spatial distribution over space and time and then integrate them with community data [1, 2]. The major tasks proposed in the global initiative for biodiversity conservation is to document floral, faunal and microbial diversity with detailed accounts of species of all regions and developing an integrated interactive information system.
Results: The developed NIS found to be useful for biodiversity documentation and analysis. It is a powerful tool for the preparation of species database, atlases, derivative maps for identification of biological hotspots and preparation of habitat-wise conservation plans and for implementing national biodiversity strategies and action plans [1, 2].
Conclusion: The information generated on Biodiversity sector could be used for a variety of purposes including to obtain an overview of forestry situation, to formulate policy decisions, to take strategic decisions in forestry and ecosystem management, to plan development programmes in forestry and ecosystem management.
Acknowledgements: The research has been supported Department of Biotechnology, New Delhi, India.

References

1. Jamison D.C. (2003) Bioinformatics. 12:679.
2. Soberón J., Peterson A.T. (2004) Bioinformatics. 29:1444.

DOI 10.18699/WIBSB-2018-43

Computer genomics of regulatory single nucleotide polymorphisms in neurodegenerative diseases based on metabolic pathways models

P. Shanmughavel[1]*, R. Sathishkumar[1], M.P. Ponomarenko[2, 3], M.R. Khantemirova[3], L.E. Tabikhanova[3], Y.L. Orlov[2, 3]

[1] *Department of Bioinformatics, Bharathiar University, Coimbatore, Tamilnadu, India*

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

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

** e-mail: shanmughavel@buc.edu.in*

Key words: non-coding RNA, transcriptomics, stress response, crop plants

Motivation and Aim: Neurodegenerative diseases present challenging problem for growing population in Asian countries demanding joint research efforts based on genetics data and bioinformatics tools. Our work will focus on integration of national sources and computer tools developed for regulatory single nucleotide polymorphisms (SNP) analysis [1]. Object of the study is neurodegenerative diseases in context of regulatory effects of SNPs. SNPs play important role in neurodegenerative diseases such as Parkinson's disease, Alzheimer's disease, Huntington's disease, Amyotrophic lateral sclerosis (ALS), Motor neuron disease, Schizophrenia.

Methods and Algorithms: The modeling approaches is based on integration of data, analysis of sequence contexts, estimation of regulatory effects of single nucleotide polymorphisms in metabolic pathways and gene networks. We develop novel software for supercomputer analysis of omics data, regulatory SNP prediction, integrated databases on nucleotide polymorphisms in neurodegenerative diseases. Aim is computational prognostic effects of polymorphysms and treatment of neurodegenerative diseases based on personalized genome data. We will use several tools for SNP prediction. The tolerated and deleterious SNPs will be predicted using SIFT (Sorting Intolerant from Tolerant (SIFT, version 2). For regulatory SNP (rSNP) effects we use tool SNP_TATA_comparator developed at the Institute of Cytology and Genetics SB RAS, Novosibirsk [2].

Results: We analyze the SNP responsible for these neurodegenerative diseases and compiled available data and computer tools to a web-resource.

Conclusion: The modeling approaches will be based on integration of genomics data, analysis of sequence contexts, estimation of regulatory effects of single nucleotide polymorphisms in metabolic pathways and gene networks. We continue work on integration of stress-response data and gene sets using available data sources.

Acknowledgements: The research has been supported by RFBR and Indian DST. Computing done at Siberian Supercomputer center SB RAS was supported by budget project 0324-2018-0017.

References

1. Sachidanandam R. et al. (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature. 409:928-933.
2. 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 10.18699/WIBSB-2018-44

Long non-coding RNAs in human hereditary diseases

P. Sparber[1]*, A. Filatova[1], M. Skoblov[1, 2, 3]
[1] *Research Centre for Medical Genetics, Moscow, Russia*
[2] *Moscow Institute of Physics and Technology, Dolgoprudny, Russia*
[3] *Far Eastern Federal University, Vladivostok, Russia*
** e-mail: psparber93@gmail.com*

Key words: long non-coding RNA, lncRNA, inherited disease, gene regulation

Motivation and Aim: About 15'000 of long non-coding RNA (lncRNA) genes are annotated in the human genome. Recent studies showed the key role of lncRNAs in a variety of fundamental cellular processes. Dysregulation of lncRNAs can drive tumorigenesis and they are now consider to be a promising therapeutic target in cancer. However, how lncRNA contribute to the development of hereditary diseases in human is still mostly unknown. This review is focused on hereditary diseases in the pathogenesis of which long non-coding RNAs play an important role.
Methods: Analysis of recent studies in the field of lncRNAs function, both in human or model organism systems can shed light upon the role of lncRNAs in the pathogenesis of hereditary diseases.
Results: Many different diseases, from imprinting to blood disorders are described, in which lncRNAs dysregulation is a crucial event in disease progression. LncRNAs can participate in the development of hereditary diseases by different mechanism. The most common is recruitment of choromatin-modifying complexes, such as polycomb repressive complex 2 (PRC2) to different target loci across the genome. Other mechanisms include antisense transcription, splicing regulation, miRNA-dependent mechanism and others.
Conclusion: LncRNAs can play an important role in the development of hereditary diseases in human, regulating the expression of their targets genes by numerous transcriptional and post-transcriptional mechanisms. We assume that further studies in this field will expand the spectrum of hereditary diseases that involve lncRNA dependent mechanism in their pathogenesis. A deeper understanding of the pathogenesis of hereditary diseases is crucial for the development of novel therapeutic strategies and lncRNA looks very promising in this area.

DOI 10.18699/WIBSB-2018-45

Walking pathways in cancer

Ph. Stegmaier[1], O.V. Kel-Margoulis[1], E. Wingender[1, 2], F.A. Kolpakov[3], A.E. Kel[1, 4]*
[1] *geneXplain GmbH, Wolfenbuettel, Germany*
[2] *Institute of Bioinformatics, Goettingen University, Goettingen, Germany*
[3] *Biosoft.ru, Novosibirsk, Moscow, Russia*
[4] *Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia*
** e-mail: alexander.kel@genexplain.com*

Key words: promoters, signal transduction, genomics, transcriptomics, cancer, TRANSFAC

Motivation and Aim: Regions of non-coding DNA in genomes are the source of high adaptability of molecular genomic systems of multicellular eukaryotic organisms (such as human) to varying external conditions. Such high adaptability is provided and maintained, first of all, through structural plasticity of gene regulatory networks. Binding of highly variable complexes of transcription factors to the fluctuating opened chromatin regions in genome (due to non-coding genomic and epigenomic variations) underlies the fundamental basis for such structural plasticity of gene regulatory networks. One of the most important results of the plasticity of the cancer pathways is the acquired resistance of tumors to chemotherapy. In this talk we will describe our results of analysis of non-coding regions of differentially expressed genes using TRANSFAC database can help to elucidate molecular mechanisms of cancer drug resistance and identify new potential therapeutic targets.
Methods and Algorithms: Empirical information about the interaction of transcription factors and the regulated target genes, obtained by either conventional or high-throughput methods, has been collected in the TRANSFAC database since 30 years. We used TRANSFAC to predict transcription factors (TFs) that regulate drug resistance genes in cancer. Analysis of signaling pathways that control the activity of TFs signal transduction network was done using most comprehensive signal transduction database TRANSPATH. With the help of geneXplain platform (http://www.genexplain.com) we identified structural changes in regulatory pathways with positive feedback loops, which helps to decipher the molecular mechanisms of the emergence of drug resistance and reveal potential drug targets .
Results: We analysed multiple "-omics" data in the cancer cells resistant to chemotherapy by methotrexate (MTX), including RNA-seq, ChiP-seq and phosphoproteomics. We identified the following potential drug targets against induced resistance of cancer cells: TGFalpha, IGFBP7, alpha9-integrin. Application of chemoinformatics tool PASS to revealed targets helped to identify the following chemical compounds: zardaverine, divalproex and human metabolite nicotinamide N-oxide that potentially may be used to sensitization of cancer cells against MTX.
Conclusion: We think, that often non-reversible structural changes of the regulatory networks due to an epigenomic "evolution" of gene regulatory regions cause transformations in the system switching to a disease state. We call such structural network changes as "walking pathways". The analysis of this phenomenon helps us to to identify prospective drug targets to treat cancer.
Acknowledgements: The research has been supported in part by FASIE grant MirCol No. 340ГР/24467.

DOI 10.18699/WIBSB-2018-46

Ontology of homologous series

V. Suslov
Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
e-mail: valya@bionet.nsc.ru

Key words: parallel evolution, Vavilov's law of homologous series

Motivation and Aim: It is assumed that the Vavilov's law of HS must unite all events of parallelisms, but the description of the intrinsic trends of autoadaptation and stabilization of homologous series (HS) raised the question of the types of HS.
Results: We can talk about the bouquet of HS laws. In the literature described HS for inherited traits of adult forms living syn- (Vavilov) and diachronously (Sobolev), for successions of transformational trends of these traits (Cope's homo- and heterologous series) and structural completeness for these trends (Kammerer's Gesetz der Serie and Meyen's refrains); HS for norm and pathology (Krenke, Rostand), between pathologies (Rostand), norm and stress (Suslov) and norm of reaction in norm and under the stress (Gorban) and between norms (Walsh); HS for sexual dimorphism (Efremov and Geodakjan's rules); HS for inherited expression/location of the traits in body space (Ugolev's universal blocks) and the same for inherited and non-inhereted expression/location (Iordansky) and finally HS for onto- and phylogenesis (Mueller-Haeckel and von Baer laws), for ontogenesies (von Baer laws) for inherited traits and non-inherited modification (Lamarck) and HS between living and nonliving nature (Meyen's refrains). Four concepts of these HS explicated. 1) Vavilov's idea. HS evolution reduced to the evolution of traits in a limited space of possibilities[1]. 2) Kammerer's idea. HS is a consequence of the regulation of several subsystems by a supersystem: a priori[2], permanently[3] or not – HS can exist by inertia[4]. Meyen's refrain can be found in all three subtypes. 3) The HS as program (transposition etc.). 4) The HS as result of autoadaptation. This HS-4 is logically incompatible with the HS-2, but not HS-1 and 3 and can explain the incorrectness in them.
Acknowledgements: Supported by the grant 0324-2018-0017 and IP 0324-2018-0021.

[1] For example, classic Vavilov's HS based on the geographical space of centers of diversity and Zavarzin HS based on the Zavarzin-Sax's space of logical possibilities.
[2] Periodic Mendeleev system and other correlative nonliving, suborganismal and superorganismal systems functioning on the Maupertuis principle, but not linearly additive as HS-4.
[3] For example, Belyaev's HS of domesticated forms usually accepted as Vavilov's HS. In sensu stricto this is not so. Vavilov's HS includes both domesticated and wild forms and originates within centers of diversity that characterized by an abundance of rare dominant alleles. Belyaev's HS revealed only between domesticated forms: traits revealed during the domestication of foxes are parallel to those of domesticated mink, horse, cattle and, finally, deer but not found in any wild species of Carnivora and ungulates. De novo traits in Belyaev's HS revealed as semi-dominant but not dominant. Belyaev's HS is more logical to compare with the Rostand and Krenke HS, if we consider domestication as a pathology.
[4] Most purely represented in mechanics. Unsuccessfully searched by Kammerer and Lyubishchev (as style phenomenon) in biology among the inherited traits of the archetype. However, were successfully found in biology by Gorban (the non-inherited effect of group stress) and Efremov and Geodakjan (inherited traits of sexual dimorphism without archetype).

DOI 10.18699/WIBSB-2018-47

Polymorphism *CYP2D6* gene for xenobiotic biotransformation in populations of Buryats and Russians of Eastern Siberia

L.E. Tabikhanova[1]*, L.P. Osipova[1, 3], E.N. Voronina[2, 3], M.L. Filipenko[2, 3]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia*
[3] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: tabikhan@bionet.nsc.ru*

Key words: population genomics, Asian populations, Buryats, polymorphism, genes for xenobiotic biotransformation, databases, CYP2D6

Motivation and Aim: The study of the gene polymorphism of the system of biotransformation of xenobiotics is an important area of modern medical and genetic research. The *CYP2D6*3 (2549del)* and *CYP2D6*4 (1846A)* variants are associated with risks of drugs side effects and cancer. The variants are annotated as regulatory polymorphisms *rs35742686* and *rs3892097*, correspondingly. The aim of this work is to study *CYP2D6* gene polymorphism in different ethnic groups.
Methods and Algorithms: This study was performed on Eastern (N = 132) and Western (N = 278) Buryats, Russians of East Siberia (N = 122) and Metis, the progeny of mixed marriages of Buryats with Russians (N = 56). Genotyping was performed using real-time PCR with competitive TaqMan allele-specific probes.
Results: The *CYP2D6*3 (2549del)* allele was not detected in Buryat cohorts, among Russians it was 0.4 %, and it was 2.7 % among Métis. The frequency of the *CYP2D6*4 (1846A)* in Eastern and Western Buryats was 5.3 % and 4.3 %, respectively. These data correspond to the frequency range found in Eastern Asian populations [1]. It was significantly higher in the Russian population (12 %), and among Métis (9.8 %).
Conclusion: The obtained data makes it possible to predict a reduced risk of side effects of drugs and cancer associated with *CYP2D6*3 (2549del)* and *CYP2D6*4 (1846A)* in the Buryat population. However, metisation introduces new polymorphic variants into indigenous populations, shifts gene frequencies and changes the degree of risks.
Acknowledgements: Authors are grateful to Dr Ming Chen and Haihua Bai for science discussion. Previous researches were supported by RFBR-NSFC grant project. The work is supported by budget project No. 0324-2018-0016.

References

1. The 1000 Genomes Project Consortium (2012) An integrated map of genetic variation from 1,092 human genomes. Nature. 7422:56.

DOI 10.18699/WIBSB-2018-48

Bioinformatics tools for 3D chromosome contacts analysis

O. Thierry[1, 2], A.G. Galieva[2], A.V. Tsukanov[2, 3], Y.L. Orlov[2, 3]*
[1] *University of Bordeaux, France*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
** e-mail: orlov@bionet.nsc.ru*

Key words: genomics, sequencing technologies, gene expression regulation, 3D genome structure, Hi-C, ChIA-PET, database

Motivation and Aim: Chromatin interactions play a critical role for gene expression regulation. Series of post-genome technologies have been developed to study the binding of transcription factors for transcription regulation, such as chromatin immunoprecipitation arrays (ChIP-Seq) [1]. Correspondingly, set of software tool for processing of such data has been developed. Another challenge is to detect functional contacts of target gene promoters via chromosome loops or attracting RNA polymerase II complex for gene transcription. Identification of genome-wide distal chromatin interactions that lead the regulatory elements to their target genes may provide novel insights into the study of transcription regulation. Chromatin Interaction Analysis with Paired-End-Tag sequencing (ChIA-PET) method for such analysis requires development of specialized software [2]. The aim of the work was to review existing tools for 3D genome structure develop a computer program for statistical data analysis and test it on CTCF binding sites, genes and spatial topological domains [3].
Methods and Algorithms: The data have been obtained via available data sources containing experimental information from ChIP-seq, Hi-C, ChIA-PET tests using different sequencing platforms. Gene annotation was obtained from UCSC Genome Browser (http://genome.ucsc.edu). We reviewed existing software and created a database prototype of bioinformatics tools for 3D genome structure analysis.
Results and Conclusion: We tested program for analysis of ChIA-PET experimental data. The result of the program is a distribution of CTCF transcription factor binding sites on domains on the human chromosomes. The distributions of human genes relative CTCF binding sites and a randomly generated list of such sites as the program output were used to estimate statistical significance of the associations found. With the rapidly increasing resolution of Hi-C datasets, the size of the chromatin contact map will soon exceed the memory capacity of general computers. The same problem related to ChIA-PET and subsequent data integration has to be solved by our software development.
Acknowledgements: OT internship at Novosibirsk University was supported by French educational exchange program. The research has been supported in part by RFBR and by the budget project No. 0324-2018-0017.

References

1. Li G. et al. (2014) Chromatin Interaction Analysis with Paired-End Tag (ChIA-PET) sequencing technology and application. BMC Genomics. 15(Suppl 12):S11.
2. Fullwood M.J. et al. (2009) An oestrogen-receptor-alpha-bound human chromatin interactome. Nature. 462(7269):58-64.
3. Li G. et al. (2012) Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell. 148(1-2):84-98.

DOI 10.18699/WIBSB-2018-49

Fast search of long approximate repeats in DNA sequences with bounded indel density

S. Tsarev[1]*, M. Senashova[2], M. Sadovsky[1, 2]
[1] *Siberian Federal University, Krasnoyarsk, Russia*
[2] *Institute of computational modelling SB RAS, Krasnoyarsk, Russia*
** e-mail: sptsarev@mail.ru*

Key words: algorithms, Vernier pattern, edit distance, gauge repeat, mutation

Motivation and Aim: The search of common strings in two or several symbol sequences makes a core in bioinformatics and up-to-date molecular biology. The problem is far from a completion, in spite of a long story. In general, the problem is the following: given sequences T_1, T_2, …, T_k of symbols from some finite alphabet, find all possible common substrings (i. e. coherent subsequences) occurred in the sequences, maybe, with some mismatches. Previously, a new algorithm for the fast search of common substrings in two or several symbol sequences had been reported [1]. The algorithm was originally implemented for the exact matching strings search, while it allows some extensions for error tolerant search of substrings. The algorithm [1] for search of exactly matching substrings is much faster compared to the brute force search methods; it is based on a simple idea of rarefied dictionaries and uses the classical Vernier scale, cf. for example [2].
Results: A novel algorithm to find all sufficiently long repeating nucleotide substrings in one or several DNA sequences is proposed. The algorithm searches approximately matching strings very fast with given level of local mutation density. Also, the extended version of the method to identify all sufficiently long repeating nucleotide substrings in one or several DNA sequences with indel mismatches is proposed. The method based on a specific gauge applied to DNA sequences that guarantees the identification of all repeating substrings. The method allows the matching substrings to contain a given level of errors of all types. The gauge is based on the development of a heavily sparse dictionary of repeats, thus drastically accelerating the search procedure [1–3]. Some biological applications illustrate the method.
Acknowledgements: This study was supported by a research grant No. 14.Y26.31.0004 from the Government of the Russian Federation (M.G. Sadovsky) and the grant from Russian Ministry of Education and Science to Siberian Federal University, contract No. 1.1462.2014/K (S.P. Tsarev).

References

1. Tsarev S., Sadovsky M. (2016) New error tolerant method for search of long repeats in DNA sequences. LNBI. 9702:171-182.
2. https://en.wikipedia.org/wiki/Vernier_scale
3. Tsarev S., Senashova M., Sadovsky M. (2018) Fast Algorithm for Vernier Search of Long Repeats in DNA Sequences with Bounded Error Density. LNBI. 10849:88-99.

DOI 10.18699/WIBSB-2018-50

Extended clusters of transcription factor binding sites in embryonic stem cells

A.V. Tsukanov[1]*, A.I. Dergilev[2, 3], Y.L. Orlov[1, 3]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *University Grenoble-Alps, Grenoble, France*
[3] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: ya.cukanton@yandex.ru*

Key words: genomics, embryonic stem cells, transcription factors, binding sites, ChIP-seq, gene expression regulation

Motivation and Aim: Computer study of transcription regulation presents important problem in molecular biology challenging growing volume of sequencing data. ChIP-Seq allows to detect interactions between DNA and protein factors regulation in genome-wide scale [1]. Additionally, ChIP-seq data with RNA-seq data allows construct transcriptional regulatory network. It can help us better understand the molecular mechanisms of organism development and cellular reprogramming, which is very important for generating various cell types for regenerative therapies [2]. Thus, it is critical to annotate clusters of binding sites of different transcription factors in distal region of genome that may function as enhancers.
Methods and Algorithms: We used own R scripts and some R packages from Bioconductor (http://www.bioconductor.org/) for TFBS analysis. The ChIP-seq data of H1 stem cell was downloaded from Cistrome database (http://cistrome.org/db/#/).
Results: We have developed set of scripts with using packages from Bioconductor and CRAN (GenomicRanges, AnnotationHub, fastcluster, ggplot2) for transcription factor binding sites analysis. It implements methods for processing, analyzing, and visualizing ChIP-seq data. In particular, next functions are: 1) calculating enrichment of peaks across the genome; 2) calculating distribution density of binding peaks near transcription start site (TSS); 3) generating random peaks; 4) hierarchical clustering of binding peaks. The clusters of transcription factors were obtained. It was shown that clusters contained four transcription factors and more can't be obtained by chance. We have generated the heatmap of co-association between transcription factors, and distribution of binding peas near gene TSS. We have shown that transcription factors mostly enrich the regions near TSS in embryonic stem cells.
Conclusion: Construction of clusters of transcription factor binding sites with RNA-seq data allowed reconstruct gene regulatory networks. We continue work on integration of ChIP-seq transcription factor binding sites data using available data sources.
Acknowledgements: The research has been supported by RFBR. Computing at Siberian Supercomputer center SB RAS was supported by budget project 0324-2018-0017.

References

1. Chen X., Bhadauria V., Ma B. (2018) ChIP-Seq: A powerful tool for studying protein-DNA interactions in plants. Curr. Iss. Mol. Biol. 27:171-180.
2. Bing H., Kai T. (2016) Understanding transcriptional regulatory networks using computational models. Curr. Opin. Genet. Dev. 37:101-108.

DOI 10.18699/WIBSB-2018-51

Genome-wide analysis of long non-coding RNAs responsive to multiple nutrient stresses in *Arabidopsis thaliana*

J. Wang[1, 2, 3], Q. Chen[1], W. Wu[1], Y. Chen[1, 4], Y. Zhou[1], Y. Liu[1], M. Chen[1, 2]*
[1] *Department of Bioinformatics; the State Key Laboratory of Plant Physiology and Biochemistry, College of Life Sciences, Zhejiang University, Hangzhou, P. R. China*
[2] *James D. Watson Institute of Genome Sciences, Zhejiang University, Hangzhou, P. R. China*
[3] *Center for Stem Cell and Regenerative Medicine, Zhejiang University School of Medicine, Hangzhou, P. R. China*
[4] *College of Life Science, Inner Mongolia University for the Nationalities, Tongliao, P. R. China*
** e-mail: mchen@zju.edu.cn*

Key words: long non-coding RNAs, nutrient stresses, *Arabidopsis thaliana*

Motivation and Aim: Nutrient stress is a most important environmental stress that limits plant growth and development. Although recent evidence highlights the vital functions of long non-coding RNAs (lncRNA) in response to single nutrient stress in some model plants, a comprehensive investigation of the effect of lncRNAs in response to nutrient stress has not been performed in *A. thaliana*. So we presented the identification and characterization of lncRNAs under seven nutrient stress conditions.
Methods and Algorithms: Raw data from previously RNA-seq datasets were downloaded from NCBI SRA, including 14 nutrient stress samples and 12 untreated control ("normal") samples. After processing these raw data, we identify lncRNAs, construct ceRNA network based on the crosstalk of miRNAs and their target, built a stress-related co-expression network based on 14 stress samples.
Results: The expression pattern analysis revealed that aberrant expression of lncRNAs is a stress-specific manner under nutrient stress conditions, and that lncRNAs are more sensitive to nutrient stress than protein coding genes (PCGs). Moreover, competing endogenous RNA (ceRNA) network and lncRNA-mRNA co-expression network (CEN) were constructed to explore the potential function of these lncRNAs under nutrient stress conditions. We further combined different expressed lncRNAs with ceRNA network and CEN to selected key lncRNAs in response to nutrient stress.
Conclusion: Our results suggest that lncRNAs play a significant role in the response of nutrient stress in *A. thaliana*. The integrative analysis of ceRNA network and CEN may provide important information for further insights into the role of lncRNAs in response to stress in plants.
Acknowledgements: We are grateful to the other members of Ming Chen's laboratory for helpful discussions and helpful comments.

DOI 10.18699/WIBSB-2018-52

Genome-wide cell-free DNA methylation profiling in lung cancer patients

W. Xu[1], Q. Zhao[1], C. Xia[1], J. Lu[2], Y. Kang[1], X. Zhao[3]*
[1] *Bio-ID Center, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China.*
[2] *Department of Pulmonary Medicine, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai, China*
[3] *Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai, China.*
* *E-mail: xiaodongzhao@sjtu.edu.cn*

Key words: cfDNA, DNA methylation, MeDIP, next generation sequencing, non-small cell lung cancer

Motivation and Aim: As a non-invasive blood testing, the detection of cell-free DNA (cfDNA) methylation in plasma is raising increasing interest due to its diagnostic and biology applications [1]. Although extensively used in cfDNA methylation analysis, bisulfite sequencing is less cost-effective. Through methylated DNA immunoprecipitation coupled with deep sequencing (MeDIP-seq), we aimed to characterize cfDNA methylome in cancer patients.
Methods and Algorithms: In this study, we investigated the cfDNA methylation patterns in lung cancer patients by MeDIP-seq. MEDIPS package was used for the identification of differentially methylated regions (DMRs) between patients and normal ones.
Results: Overall, we identified 128 differentially methylated regions (DMRs) in gene promoter regions, 21 hypermethylation and 107 hypomethylation respectively, by comparing lung cancer patiens and healthy individuals as controls. 21 hypermethylation regions represent 20 genes. Some of the genes had been previously reported to be associated with lung cancers, such as CPXM1 and C1orf210.
Conclusion: Taken together, our study provided an alternative method of cfDNA methylation analysis in lung cancer patients with potential clinical applications.
Acknowledgements: This work was supported by Development Program for Basic Research of China (2014YQ09070904), National Natural Science Foundation of China (31671299), Shanghai Science and Technology Committee Program (17JC1400804), Medical engineering cross fund (YG2017ZD15, YG2015QN35) and Laboratory Innovative Research Programme of Shanghai Jiao Tong University (17SJ-18).

References

1. Feng H., Jin P., Wu H. (2018). Disease prediction by cell-free DNA methylation. Briefings in bioinformatics.

DOI 10.18699/WIBSB-2018-53

Circular RNAs in plants

C. Ye, Q. Chu, L. Chen, X. Zhang, L. Fan*
Institute of Crop Science and Institute of Bioinformatics, Zhejiang University, Hangzhou, China
** e-mail: fanlj@zju.edu.cn*

Key words: plant genomics, circular RNA, Oryza sativa, *Arabidopsis thaliana*, splicing, database

Circular RNAs (circRNAs) are covalently closed loops derived from back-splicing of precursor mRNAs. Recently, a large number of circRNAs have been detected among diverse organisms. For the first time in plants, we have identified 12,037 and 6,012 circRNAs in *Oryza sativa* and *Arabidopsis thaliana*, respectively. We demonstrated that circRNAs are widespread in plants and revealed the common and distinct features of circRNAs between plants and animals. We found that there are a certain number of circRNAs with sequence and splicing position conservation in plants, demonstrating the important role of circRNAs in plant growth, development and evolution. Regarding to the novelty of plant genomes/genes, a software for prediction of plant circRNAs, termed PcircRNA_finder, which is more sensitive and precise in detecting plant circRNAs than other frequently used programs. Another software, termed circseq-cup, was also developed, which can assemble the full-length sequences of circRNAs using the back-splicing RNA-Seq reads and their corresponding paired-end reads. We for the first time identified full-length sequences of nearly 3,000 circRNAs from *O. sativa* genome. Based on the full-length sequences, we showed that alternative circularization of circRNA is a common feature in *O. sativa* and, surprisingly, found that the junction sites of a large number of circRNAs are flanked by diverse non-GT/AG splicing signals in *O. sativa*. Moreover, we have created a database of plant circRNAs, PlantcircBase (http://ibi.zju.edu.cn/plantcircbase/). We have collected publicly available circRNAs identified in recent years by bioinformatics prediction and/or experimental validation, as well as circRNAs newly identified by our own lab in rice and Arabidopsis. Apart from the detailed and comprehensive information for each circRNA entry, PlantcircBase also provides basic tools including browsing, searching and downloading as well as advanced tools, including visualizing structures of circRNAs and predicting circRNAs. Part of our results have been published in Ye et al., 2015, New Phytologist; Chen et al., Bioinformatics; Ye et al., 2017, RNA Biology; Chu et al., 2017, Molecular Plant.

DOI 10.18699/WIBSB-2018-54

Computer program for visualization of gene ontology results based on transcriptomics data

A.A. Zaporozhchenko[1], A.G. Galieva[1], Y.L. Orlov[1, 2]*
[1] *Novosibirsk State University, Novosibirsk, Russia*
[2] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
* *e-mail: orlov@bionet.nsc.ru*

Key words: genomics, transcriptomics, gene expression, annotation, gene ontology, bioinformatics tool

Motivation and Aim: Analysis of gene ontology results based on selected gene lists presents technical problem for processing of transcriptomics data. New annotation categories, such as presence of miRNA, non-coding RNA, transcripts demand development of new mathematical functional for data processing [1]. Previously, program complex ICGenomics has been designed for storage, mining, and analysis of high-throughput sequencing experiments [2]. We sought extend potential of the tools by new functional properties for results presentation and visualization.
Methods and Algorithms: The developed program complex able fulfill the following tasks: ChIP-seq analysis; functional annotation of gene regulatory regions in nucleotide sequences; prediction of nucleosome positioning; and structural and functional annotation of proteins, including prediction of allergenicity parameters, as well as estimates of evolution changes in protein families. We developed set of new computer scripts for integration into the program complex. It implements both standard and modern methods for processing, analyzing, and visualizing sequencing data and functional annotation of genome regions.
Results and Conclusion: The program complex allows process nucleotide sequences from next generation sequencing data, contract gene expression tables, calculate percentage of genes related to defined categories (ontology terms, parameters of expression in different tissues – tissue-specificity). Correlation of any parameters of gene groups could be visualized as heat map. The work has to be extended to the analysis of transcription factor binding sites in eukaryotic genomes.
Acknowledgements: The research has been supported by the budget project No. 0324-2018-0017.

References

1. Wang J., Meng X., Dobrovolskaya O.B., Orlov Y.L., Chen M. (2017) Non-coding RNAs and their roles in stress response in plants. Genomics Proteomics Bioinformatics. 15(5):301-312.
2. Orlov Y.L. et al. (2012) ICGenomics: program complex for symbol sequence analysis in genomics, Vavilov journal of genetics and breeding. 16(4/1):732-741. (In Russian).

DOI 10.18699/WIBSB-2018-55

An overview of existing clinical severity scales and genotype-phenotype studies in FSHD

N.V. Zernov[1]*, M.Yu. Skoblov[1, 2, 3]
[1] *Research Centre for Medical Genetics, Moscow, Russia*
[2] *Moscow Institute of Physics and Technology, Dolgoprudny, Russia*
[3] *Far Eastern Federal University, Vladivostok, Russia*
** e-mail: nzernov01@gnmail.com*

Key words: FSHD, genotype-phenotype correlation, clinical severity score, D4Z4 array size, D4Z4 array methylation, disease prognosis, family counselling

Motivation and Aim: Facial-scapular-humeral myodystrophy (FSHD) is an autosomal dominant myodystrophy, the basis of its pathogenesis is ectopic expression of the transcription factor DUX4 in skeletal muscle. *DUX4* is encoded by each D4Z4 repeat unit on 4th chromosome [1]. The prevalence of the disease in different populations varies from 1: 20000 to 1: 14,000 [2]. In this paper, we reviewed existing studies about impact of genetic, epigenetic and gender differences on clinical severity and the possibility of their use for disease prognosis and family counselling.
Methods: To date, there are no specific, fully validated scores for assessing the clinical severity of the FSHD. The most frequently used clinical severity scale was published by Ricci et al. [3] and corrected to age of onset by van Overveld et al. [4].
Results: In most studies of the influence of genetic factors on phenotype, the inverse correlation between the number of the D4Z4 repeats on the 4th chromosome and the severity of clinical manifestations was established [3]. For both types of FSHD, the inverse dependence of the severity of clinical manifestations on the level of methylation of the D4Z4 repeats on 4qA chromosome was shown [5].
Conclusion: Reviewed works allow concluding that the penetrance and the severity of FSHD are in inverse correlation with the number of D4Z4 repeats on the 4qA haplotype and the level of methylation. Moreover, the smaller the residual number of the D4Z4 repeats, the smaller the variability of the clinical picture. In groups of asymptomatic carriers and patients with minimal clinical manifestations are dominated by females. Most researchers conclude that the most accurate prediction could be given with the following data: age of onset, gender, family examination data; clinical severity scoring; the length of D4Z4 repeats on the 4qA chromosome in FSHD1, additionally for FSHD2 – presence and type of mutations in the gene SMCHD1; haplotype and methylation status of the array of D4Z4 repeats of chromosome 4.

References

1. Gabriels J. et al. (1999) Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. Gene. 236(1):25-32.
2. Lemmers R.J. et al. (2010) Worldwide population analysis of the 4q and 10q subtelomeres identifies only four discrete interchromosomal sequence transfers in human evolution. Am J Hum Genet. 86(3):364-77.
3. Ricci E. et al. (1999) Progress in the molecular diagnosis of facioscapulohumeral muscular dystrophy and correlation between the number of KpnI repeats at the 4q35 locus and clinical phenotype. Ann Neurol. 45(6):751-7.
4. van Overveld P.G. et al. (2005) Variable hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy. Ann Neurol. 58(4):569-76.
5. Lemmers R.J. et al. (2015) Inter-individual differences in CpG methylation at D4Z4 correlate with clinical variability in FSHD1 and FSHD2. Hum Mol Genet. 24(3):659-69.

DOI 10.18699/WIBSB-2018-56

An overview of FSHD diagnostic methods

N.V. Zernov[1]*, M.Yu. Skoblov[1, 2, 3]
[1] *Research Centre for Medical Genetics, Moscow, Russia*
[2] *Moscow Institute of Physics and Technology, Dolgoprudny, Russia*
[3] *Far Eastern Federal University, Vladivostok, Russia*
** e-mail: nzernov01@gnmail.com*

Key words: FSHD diagnostic, D4Z4 array, Southern blotting, molecular combing; nanopore sequencing, genotyping, methylation profiling

Motivation and Aim: Facial-scapular-humeral myodystrophy Landouzy-Dejerine (FSHD) is an autosomal dominant myodystrophy. The prevalence of the disease in different populations varies from 1: 20000 to 1: 14,000 [1–4], conceding only to Duchenne muscular dystrophy and myotonic dystrophy. The basis of FSHD pathogenesis is ectopic expression of the transcription factor DUX4 in skeletal muscle encoded by D4Z4 repetitive units of subtelomeric region of chromosome 4q35 [5]. The D4Z4 repeats are widely presented in human genome and have high similarity between different chromosomes. This fact makes diagnostic of FSHD highly complicated for basic molecular diagnostic laboratories. In this paper, we review existing methods for FSHD molecular diagnostic thus conduce to choose the most suitable diagnostic technique in a basic molecular genetic laboratory.
Methods: Here we present a comprehensive analysis of current and possible methods for FSHD diagnostic: Southern blotting [6]; molecular combing [7]; nanopore sequencing [8]; STR haplotyping [9]; 4qA haplotyping [10]; methylation profiling [11].
Results: The most informative of currently existing methods is demonstrated by molecular combing; however, a price-informative ratio is the best for nonradioactive Southern blotting.
Conclusion: Based on the criteria of specificity, sensitivity, informative and price for today, the most suitable method for FSHD diagnostic in basic molecular genetic laboratory is the nonradioactive Southern blotting. This method is allow to diagnose about 90 % of all FSHD cases, however it couldn't diagnose cases with low mosaicism, complex rearrangements between 4qA/10qA allels. These issues resolve by the molecular combing method, but because of price of reagents and special equipment this method not suitable for basic molecular genetic diagnostic laboratory. The nanopore sequencing now have high price, not enough reads length and quality for FSHD diagnostic. However, the nanopore sequencing technology may be method of choice for the diagnostic in future because it has potential to resolve all limitations of currently existing methods.

References:

1. Padberg G.W. et al. (1995) Muscle Nerve Suppl. (2):S81-4.
2. Flanigan K.M. et al. (2001) Neuromuscul Disord. 11(6-7):525-9.
3. Mostacciuolo M.L. et al. (2009) Clin Genet. 75(6):550-5.
4. Norwood F.L. et al. (2009) Brain. 132(11):3175-86.
5. Snider L. et al. (2010) PLoS Genet. 6(10):e1001181.
6. Stefan J. (2007) White and Stuart Cantsilieris (eds.). Genotyping: Methods and Protocols. Methods in Mol Biol. 1492:107-125. DOI 10.1007/978-1-4939-6442-0_7.
7. Nguyen K., Walrafen P. et al. (2011) Ann Neurol. 70:627-633.
8. Mitsuhashi S., Nakagawa S. (2017) Sci Rep. 7:14789.
9. Barat-Houari M. et al. (2010) Eur J Hum Genet. 18(5):533-538.
10. Papanikos F., Skoulatou C. et al. (2014) Clin Chim Acta. 429:96-103. DOI 10.1016/j.cca.2013.11.032.
11. Jones T., Yan C. et al. (2014) Clin Epigenetics. 6(1):23. DOI 10.1186/1868-7083-6-23.eCollection 2014.

DOI 10.18699/WIBSB-2018-57

Systems genetics-based drug discovery

H.-Y. Zhang
College of Informatics, Huazhong Agricultural University, Wuhan, P.R. China
e-mail: zhy630@mail.hzau.edu.cn

Key words: systems biology, medical genetics, drug discovery, gene networks, disease genes, gene predication

The past decade has witnessed unprecedented progress in biotechnology, particularly in omics technologies. However, these great technical advances have made limited contributions to the advancement of the pharmaceutical industry. It still requires $2.6 billion and approximately 13.5 years on average for a new molecular entity (NME) drug to enter the market. Therefore, we urgently need methods that can efficiently synthesize and utilize biological big data to streamline the drug discovery pipeline.

In recent years, medical genetics has found important applications in drug discovery. However, most complex diseases are caused by multiple interacted causative factors; hence, agents that only hit single pathogenic factors are not very efficacious. Thus, to utilize medical genetic information more efficiently in drug development, we should aim at multiple disease genes, especially interacted pathogenic factors, to find potential drugs.

Systems genetics is a thriving area of study that aims to understand genetic interactions on a genome-wide scale. One of the representative algorithms in this field is the HotNet diffusion-oriented subnetworks (HotNet2) algorithm. HotNet2 is based on an insulated heat diffusion kernel algorithm that considers the heats (reflecting genetic importance) of individual genes as well as the topology of gene-gene interactions. This method can reveal functionally interacted genes within subnetworks, thus overcoming the limitations of traditional GWAS/PheWAS methods for elucidating disease pathogenesis.

In this study, we demonstrate that integrating HotNet2 calculation with GWAS/PheWAS can substantially promote the drug discovery efficiency. Besides, this method is also efficient to find drug combinations.

DOI 10.18699/WIBSB-2018-58

The interplay of drought and dehydration with the duration of plant growth: Application of bioinformatics for candidate gene identification in bread wheat

L. Zotova[1]*, S. Jatayev[1], V. Shvidchenko[1], Y. Shavrukov[2]
[1] *S. Seifullin Kazakh AgroTechnical University, Astana, Kazakhstan*
[2] *Flinders University, Adelaide, Australia*
** e-mail: lupezo_83@mail.ru*

Key words: plant genetics, wheat, drought, abiotic stress, transcription factors

Introduction and Aim: Drought can affect plants at any stage of ontogenesis. Late- and early-maturing cultivars of wheat differ in their development in ways that can advantage grain production in the unpredictable occurrence of rainfall or drought. In the environment of Northern Kazakhstan, late-maturing genotypes of wheat are preferable in the absence of rainfall during the spring/early-summer season. In contrast, higher grain yield is typically shown by wheat genotypes with an early-maturing habit in conditions of terminal drought. Therefore, high grain yield represents an integrative trait, determined by not only yield potential but also the length of plant growth and reactions to drought. The aims of this study were to identify and analyse the most suitable gene candidates among Transcription factors regulating plant reactions to drought during the development of wheat plants and to carry out comparative bioinformatic analysis of similar genes in other plant species.
Methods: In this study, grain yield among an International germplasm collection of 200 bread wheats was studied in field trails in the Akmola region of Northern Kazakhstan as well as in laboratory treatments for drought and dehydration. Bioinformatics and systems biology approaches were used to identify and analyse the most important candidate genes during plant growth for higher grain yield.
Results: Two genes encoding different Transcription factors, *TaDREB5* (Drought Responsive Element Binding) and *TaNFYC-A7* (Nuclear Factor Y), showed the highest levels of gene expression in plants exposed to drought and dehydration, and were strongly associated with higher grain yield compared to controls. Similar results in other plant species in response to abiotic stresses were found and confirmed through bioinformatics analysis.
Conclusion: The results support the use of the above candidate genes for advanced selection of bread wheat genotypes for high grain yield with improved tolerance to drought and dehydration through adaptation to the required length of plant growth.
Acknowledgements: This study was supported by the Ministry of Education and Science. Kazakhstan, Research program BR05236500.

DOI 10.18699/WIBSB-2018-59

Promoter sequences driving inducible gene expression in human T- and NK-cells

S. Kulemzin[1] *, D. Matvienko[1, 2], A. Sabirov[1, 2], A. Sokratyan[1, 2], D. Chernikova[1, 2], T. Belovezhets[1, 2], A. Chikaev[1], A. Taranin[1, 2], A. Gorchakov[1, 2]
[1] *Institute of Molecular and Cellular Biology SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: skulemzin@mcb.nsc.ru*

Key words: chimeric antigen receptor, cancer, reporter, inducible promoter, NK cell lines

Motivation and Aim: Cytotoxic activity of T- and NK-cells can be efficiently retargeted against cancer cells using chimeric antigen receptors (CARs) and rTCRs. In the context of solid cancers, use of armored CAR T- and NK cells secreting additional anti-cancer molecules such as cytokines, chemokines, antibodies, BiTEs, inverted cytokine receptors, and checkpoint inhibitors, appears particularly promising, as this may help overcome immunosuppressive tumor microenvironment, attract bystander immune cells and boost CAR T/NK-cell persistence. Placing the expression of such molecules under the transcriptional control downstream of CAR-mediated T/NK-cell activation offers the advantage of targeted delivery, high local concentration, and reduced toxicity. Several canonic DNA sequences that are known to function as activation-inducible promoters in human T and B cells have been described to date and typically encompass the multimers of NFkB and NFAT binding sites. However, relatively little is known about the DNA sequences that may function as activation-driven switches in the context of NK cells. We set out to compare the functionality of several activation-inducible promoters in primary human T cells, as well as in NK cell lines NK-92 and YT.
Methods and Algorithms: To this end, lentiviral constructs were engineered to express two fluorescent reporters: mCherry under NFATx4, NFkBx2, NFkBx5, NFkBx10, NFkBx20, as well as two variants of CD69 promoter and copGFP under the strong constitutive promoter of the human EF1a gene. Reporters lacking the inducible promoter or driven by the weak constitutive PGK promoter served as the controls. Pseudotyped lentiviral particles obtained using these constructs were transduced into primary human T cells and NK-92 and YT cell lines expressing a CAR specific for PSMA.
Results: The transgenic cells obtained were activated by conventional TCR cross-linking (T cells) or via a CAR (CAR-NK cell lines). Promoter activity before and after activation was assayed using FACS analysis. In T cells, the CD69 promoter encompassing CNS1 and CNS2 regions displayed the highest signal/noise ratio. Intriguingly, in the context of CAR-YT cell line neither of the seven promoters tested displayed acceptable activation profile. In CAR-NK-92 cells, the largest fold activation (which was modest) was achieved with the NFkBx20 promoter, however its expression was clearly leaky in “resting” non-activated cells.
Conclusion: Unlike in T cells, the robust activation-driven inducible expression of genetic cassettes in NK cells requires unbiased identification of promoter sequences.
Acknowledgements: Supported by the RSF (16-14-10237) and the Basic scientific research program project (0310-2018-0012).

DOI 10.18699/WIBSB-2018-60

Drug reposition as a promising strategy in therapy of cognitive deficits at neurodegenerative disorders

M.A. Tikhonova[1, 2]*, T.G. Amstislavskaya[1, 2, 3], Y.J. Ho[4]
[1] *Scientific Research Institute of Physiology and Basic Medicine, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[4] *Chung Shan Medical University, Taichung City, Taiwan, R.O.C.*
** e-mail: tikhonovama@physiol.ru*

Key words: cognitive deficits, Parkinson's disease, Alzheimer's disease, neuroinflammation, rats

Motivation and Aim: Drug repositioning (or drug repurposing / drug reprofiling) is a process of redeveloping a compound for application in a different pathology and finding new therapeutic indications for the existing drugs. The premise of repositioning is that the drugs that have previously passed clinical trials will minimize the risk of failure in future late-stage clinical trials due to toxicity and thus lead to faster drug approvals. It has been growing in importance in the last few years and becoming mainstream area in drug research and industry. This strategy appeared to be quite effective approach in psychopharmacology as well. Our project was focused on antimicrobial drug with neuroprotective properties, ceftriaxone (CEF). We evaluated its potential to restore cognitive deficits in animal models of neurodegenerative disorders (Alzheimer's disease (AD), Parkinson's disease (PD)).
Methods and Algorithms: Wistar rats were injected with a selective neurotoxin of dopaminergic neurons MPTP into SNc to establish a PD model. Mice of C57Bl/6J strain injected bilaterally i.c.v. with amyloid-beta fragment 25-35 were used as pharmacological AD model while rats of OXYS strain were used as a genetic model of sporadic AD. To evaluate the effects of CEF, the animals were treated with the drug (100 mg/kg/day, i.p., 14-36 days) and then underwent behavioral testing for cognitive function. Neuronal activity *in vivo* was measured with ME-MRI. Brain neuromorphology was evaluated with Nissl staining and immunohistochemical analysis.
Results: The treatment with CEF improved cognitive deficits in MPTP-induced rat PD model, model of sporadic AD in OXYS rats with accelerated senescence, and murine AD model induced by i.c.v. Aβ injections. We showed that MPTP lesioning resulted in decreased neuronal activity and density in the nigrostriatal dopaminergic (DAergic) system and the hippocampal CA1, CA3, and dentate gyrus, while CEF treatment prevented those disturbances. Neuromorphologically, control OXYS rats exhibited a lowered neuronal density in the hippocampal CA1 area whereas CEF restored this parameter. CEF also de-creased amyloid accumulation and neuroinflammation in Aβ-treated mice.
Conclusion: The results suggest CEF as a promising pharmacological tool for the prevention of cognitive decline at neurodegenerative disorders and gave new insights into mechanisms of its neuroprotective effects.
Acknowledgements: Supported partially by grants No. 15-04-05593-a, 15-54-52029_ HHC-a from the Russian Foundation for Basic Research (Russia) and No. MOST 104-2923-H-040-001-MY3 from the Ministry of Science and Technology (Taiwan, R.O.C.); by NSU: Academic Strategic Unit "Neuroscience in Translational Medicine".

Conclusion: international education programs and round table

Y.L. Orlov[1, 2]*, A.N. Savostyanov[1, 2, 3], T.G. Amstislavskaya[2, 3], L.I. Aftanas[2, 3]
[1] *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*
[3] *Scientific Research Institute of Physiology and Basic Medicine, Novosibirsk, Russia*
** e-mail: orlov@bionet.nsc.ru*

Key words: integrative bioinformatics, education, databases, international science exchanges

The round table on education and integration of bioinformatics data together with science cooperation will conclude the First Sino-Russian Workshop on Integrative Bioinformatics and Systems Biology. Russian and international experts in the field of bioinformatics, education programs and neurobiology from top world universities will discuss education programs. Science groups from China by Prof Xiaodong Zhao (Shanghai Jiao Tong University), Prof Ming Chen (Zhejiang University), and Prof Hong-Yu Zhang (Huazhong Agricultural University) will discuss integration of bioinformatics data and education programs. Prof M.Chen will present professional initiatives in bioinformatics in China, work of Zhejiang Bioinformatics Society.

Dr. Fyodor Kondrashov, Institute of Science and Technology, Klosterneuburg, Austria will tell about experience in organization of Young Scientists Schools on bioinformatics in Russia. The invited speakers from Russia will continue the discussion about national programs in fundamental medicine – presented by Irina M. Larina, Institute of Mathematical Problems of Biology RAS, Alexey A. Lagunin, Russian National Research Medical University named after N.I. Pirogov, Maria G. Samsonova, Saint-Petersburg Polytechnic University.

Dr Olga F. Krebs (Heidelberg Institute for Theoretical Research, Heidelberg, Germany) and Dr Maxim V. Zakharsev (Norwegian University of Natural Sciences, Oslo) will give a talk about integration of data in bioinformatics and open source platforms on computer biology in Europe. Prof Ralf Hofestädt (Bielefeld University, Germany) will discuss German-Chinese network on bioinformatics and education. Dr Marko Djordjevic (Belgrade University, Serbia) will tell about the integrating computational systems biology in research and education. Prof Tatiana Tatarinova (University of La Verne, USA) will give the educational program review "Bioinformatics: science of a toolbox?" The neurobiology experts from Academia Sinica, Taiwan ROC – Arthur C. Tsai, Hung-Ming Chang, Ya-Ling Yang – will share their expertise in research.

Bioinformatics, as an interdisciplinary field, is a case example for analysis of science history perspectives. During the seminar the analytical materials on increasing the effectiveness of the implementation of the activities of the roadmap projects of the National Technological Initiative in Russia will be discussed. We aim to develop methodological recommendations on educational projects including the usage in educational process based on the example of the Department of Medicine and Psychology and the Humanitarian Institute of Novosibirsk State University.

Acknowledgements: The participation of the incited speakers at the BGRS conference and the workshop has been supported by Russian Ministry of Science project 28.12487.2018/12.1.

Author index

浙江省生物信息学学会
Bioinformatics Society of Zhejiang Province

The Bioinformatics Society of Zhejiang Province was established in 2013. It is a registered NGO that meets all the legal and financial requirements given by the government. The goal of the society is to provide a platform for bioinformatics scientists, researchers, young students and publics, to create a network between institutions and companies, and to encourage interdisciplinary communication and in-depth discussion. We welcome members from board fields and support members and other persons interested in obtaining advanced knowledge in the field, present current progress and expand the general knowledge and applications of bioinformatics. The society will work with foreign societies, contribute to the further education of its members and organize seminars, conferences and workshops. Website: http://www.zjbioinformatics.org/

ZJBS

遂真 Life Real

Hangzhou Lifereal Biotechnology

Founded in September 2011, Hangzhou Lifereal Biotechnology Co., Ltd. focuses on the R&D and production for the life science and clinical diagnostic products. LifeReal has settled its production line with the occupation of 2000 sqM and strictly followed up the ISO 9001 quality management system. The company has got ISO13485 certificate in 2017. The company has a full team of engineers to provide the original designing service, including electronic design, structural design, software design, etc. Viewing to the worldwide cooperation, LifeReal has built up the business channels into overseas countries, such as US, Japan, Germany, UK, etc. Mostly of our products are designed with RoHS complied and CE marked. The company welcomed OEM/ODM cooperation based on the customer requirements.

Other enterprise members:

浙江天科 ZHEJIANG TIANKE

Gene 基因

厚泽生物 HOUZE BIO-TECH

联川生物 掌握基因科技核心技术

可帮基因 Canhelp Genomics

维讯生物

基因帮 .com genebang

SeeGene 圣庭基因

G-BIO 晶佰

PUTIAN

iGeneTech 艾吉泰康

中翰金诺医学检验所 JOINGENOME DIAGNOTICS

杭州高六博生物科技有限公司 GLOBAL BIOTECHNOLOGY & BIOINFORMATICS CO. LTD.

华东医药（杭州）基因科技有限公司 EAST CHINA PHARMACEUTICAL (HANGZHOU) GENE TECHNOLOGIES CO.,LTD.

APPLIED PROTEIN TECHNOLOGY 中科新生命

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.

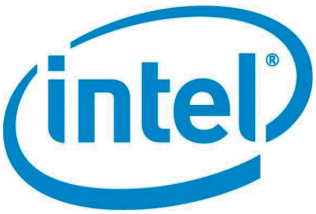
Имея богатый опыт и технические знания в области НИОКР, Huawei придерживается стратегии тесного сотрудничества и интеграции с корпоративными заказчиками и предоставляет им широкий спектр высокоэффективных клиентоориентированных ИКТ-решений и услуг, на базе глубокого понимания их потребностей. Согласно 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).

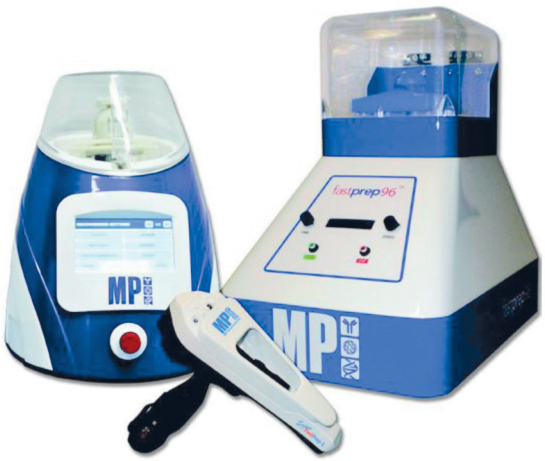
Центр исследований и разработок Intel в Нижнем Новгороде

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

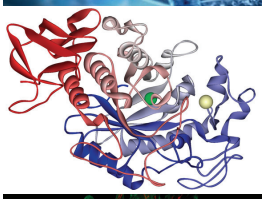
PerkinElmer For the Better
Leica MICROSYSTEMS
BioTek
BD

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.

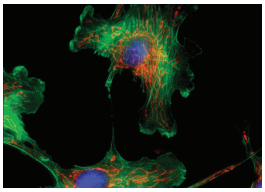
Молекулярно-генетические исследования

- 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



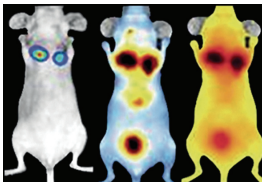
Протеомные исследования

- 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

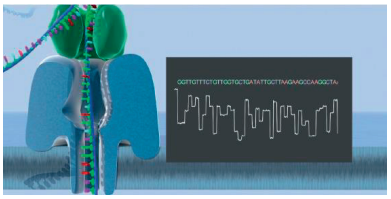
Москва, тел.: +7 (800) 555 49 40
Новосибирск, тел.: +7 (383) 227 09 63
Екатеринбург, тел.: +7 (343) 287 32 49
Владивосток, тел.: +7 (423) 201 18 08
Нижний Новгород, тел.: +7 (831) 278 61 47
Ростов-на-Дону, тел.: +7 (863) 268 99 32
Казань, тел.: +7 (843) 570 66 88
Самара: +7 (846) 246-06-54

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.	ул. Парижской	ул. Профессора	пер. Семашко, 114	Представитель	тел./факс:
тел./факс:	Лаврентьева, 6/1	Коммуны, д. 6	Попова, 23	тел/факс:	в УФО	(473) 232-4412
(495) 747-0508	тел./факс:	тел/факс:	тел./факс:	(863) 250-0006	тел./факс:	voronezh@dia-m.ru
sales@dia-m.ru	(383) 328-0048	(843) 210-2080	(812) 372-6040	rnd@dia-m.ru	(342) 202-2239	
	nsk@dia-m.ru	kazan@dia-m.ru	spb@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.

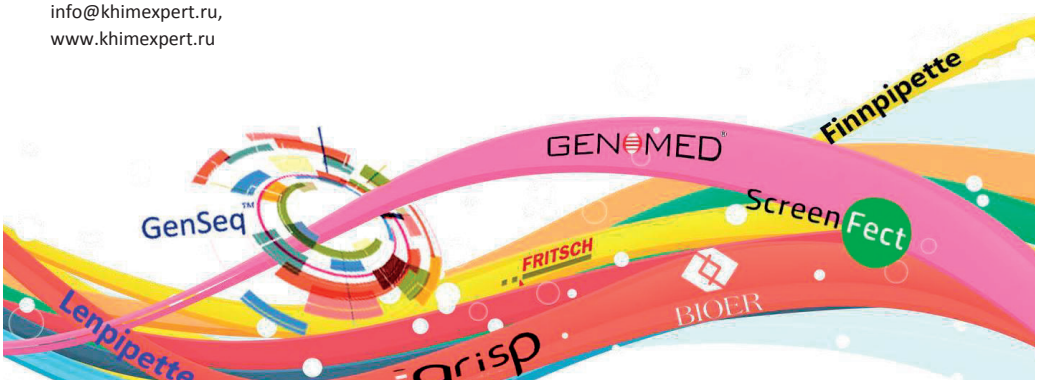
Химэксперт предлагает оборудование для анализа ДНК и РНК, в том числе и методами NGS, фундаментальных протеомных и цитологических исследований, фармацевтики и биотехнологий, прикладного тестирования, включая идентификацию личности и установление родства в криминалистике и судебно-медицинской экспертизе.

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!

Научное издание

INTEGRATIVE BIOINFORMATICS AND SYSTEMS BIOLOGY (WIBSB-2018)

First Sino-Russian Workshop

Abstracts

Printed without editing

ИНТЕГРАТИВНАЯ БИОИНФОРМАТИКА И СИСТЕМНАЯ БИОЛОГИЯ (WIBSB-2018)

Первое российско-китайское совещание

Тезисы докладов

Публикуется в авторской редакции

Публикация сборника поддержана РФФИ (грант 18-04-20044).

Выпуск подготовлен информационно-издательским отделом ИЦиГ СО РАН

Подписано к печати 02.08.2018. Формат 70 × 108 $^{1}/_{16}$. Усл. печ. л. 7,7.
Тираж 100 экз. Заказ № 190

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