Development of computer programs for drug discovery needs qualified scientific staff, not only software development and technical support. Growth of biomedical data volume and availability of new experimental technologies give solid background for complex computer modeling limiting rather by human resources than hardware. Bioinformatics education faces new challenges related to changing educational standards, distant education on online meeting formats.
COVID-19 changed research and development directions in 2020-2021. We still have consequences (distal conferences and disrupted contacts)

Professor of Sechenov University
Minister of Health of the Russian Federation Mikhail Murashko

Importance of the digital solutions for medicine was strengthen by COVID-19 and related lockdowns

At the forefront of health care
Monument to the feat of medical workers in the fight against COVID-19 (монумент «Подвигу медицинских работников в борьбе с COVID-19») (open 17.09.2021 in campus of Sechenov University in Moscow)

Rector Prof., Acad. P.V.Glybochko

New challenges in digital medicine and education
Scientific challenges in connection with the coronavirus pandemic have raised research and educational problems, changes in the methodology for mastering scientific disciplines.

Medical universities use new E-health technologies, one of the global areas of which is telemedicine. Medical teleconsultations make it possible to increase the availability of medical care for the population of remote areas, elderly and inactive patients, which is especially relevant for monitoring the spread of coronavirus infection.

The First Moscow State Medical University of the Ministry of Health of Russia (Sechenov University) and the Institute of Digital Medicine deal with the problems of digitalization of medicine, provide a platform for discussing existing issues in the development of medical technologies, online conferences, and new educational programs.

Publication activity - The Russian Journal of Telemedicine and E-Health continues series of publication on this topic (https://jtelemed.ru/). We have arranged series of international journal issues on gene expression regulation as well.
Training a neural network to recognize bladder malady

3D positioning

Fundus pictures

Peripheral and bone marrow smear images

Dr. I.A. Shaderkin
Laboratory of electronic healthcare, Sechenov University

Approaches at Sechenov University

- Teaching on informatics, IoT, Machine Learning courses
- Master-classes for students
- Specialized Russian journal on telemedicine (RSCI - РИНЦ)
Network medicine conception – extension of gene network, drug-disease network, genotype-phenotype network terms

The idea is to compare diseases by comparing the functions of genes, symptoms, drug compounds - the network approach.

Analysis of genes associated with a disease, assessment of their place in the gene network (connectivity) allows us to evaluate them as target genes for drug effects.

**Standard for students’ diploma** and course works at Sechenov University on gene networks for complex human diseases – cancers, metabolic syndrome, Parkinson’s disease.
Work on gene network of metabolic syndrome with students


DOI: 10.18097/PBMC20216703213

Presented at previous “Way2drug” conference in 2021
Publication on bioinformatics in co-authorship with the students (*Journal of Integrative Bioinformatics*, Q1)

Due to distant education format in Moscow in 2020-2021 we set priority for online bioinformatics tools

Natayla V. Gubanova, Nina G. Orlova, Arthur I. Dergilev, Nina Y. Oparina and Yurii L. Orlov*

**Glioblastoma gene network reconstruction and ontology analysis by online bioinformatics tools**

https://doi.org/10.1515/jib-2021-0031

Received August 31, 2021; accepted October 18, 2021; published online November 16, 2021

<table>
<thead>
<tr>
<th>No.</th>
<th>Description</th>
<th>Category</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Glioma</td>
<td></td>
<td>1.882 x 10^{-25}</td>
</tr>
<tr>
<td>2</td>
<td>Pilocytic astrocytoma</td>
<td></td>
<td>3.147 x 10^{-25}</td>
</tr>
<tr>
<td>3-4</td>
<td>Adult pilocytic astrocytoma/childhood pilocytic astrocytoma</td>
<td></td>
<td>1.829 x 10^{-22}</td>
</tr>
<tr>
<td>5</td>
<td>Malignant glioma</td>
<td></td>
<td>3.741 x 10^{-21}</td>
</tr>
<tr>
<td>6</td>
<td>Mixed gliomas</td>
<td></td>
<td>3.741 x 10^{-21}</td>
</tr>
<tr>
<td>7</td>
<td>Neurofibromatosis 1</td>
<td></td>
<td>6.750 x 10^{-21}</td>
</tr>
<tr>
<td>8</td>
<td>Malignant neoplasm of soft tissue</td>
<td></td>
<td>2.560 x 10^{-20}</td>
</tr>
<tr>
<td>9</td>
<td>Ganglioglioma</td>
<td></td>
<td>4.231 x 10^{-20}</td>
</tr>
<tr>
<td>10-11</td>
<td>Childhood oligodendroglioma/adult oligodendroglioma</td>
<td></td>
<td>4.616 x 10^{-20}</td>
</tr>
<tr>
<td>12</td>
<td>Sarcoma</td>
<td></td>
<td>2.272 x 10^{-19}</td>
</tr>
</tbody>
</table>

Computer pipeline – using only online tools – appropriate for students and distal education format
The growth in the volume of genetic data provides a basis for searching for associations with diseases, which is reflected in the replenishment of such databases as OMIM (https://omim.org/), GeneCards (https://www.genecards.org/). The development of experimental sequencing technologies leads to an increase in transcriptomic data, which allows the reconstruction of gene networks / signal transduction pathways based on co-expression. Existing online bioinformatics tools allow solving many practical tasks for the reconstruction of gene networks without using additional software (used in the training course for students of the Sechenov University - First Moscow State Medical University of the Ministry of Health of the Russian Federation named after I.M. Sechenov).

Figures are from http://griza.nevrologica.ru/parkinson

The research interest in the study of Parkinson's disease is due to the fact that this disease is a medical and economic problem for society and at the moment there are no treatments that can stop or reverse the neurodegenerative process accompanying this disease.
Online tools for gene ontology analysis based on gene list

http://www.pantherdb.org/

The PANTHER (Protein ANalysis THrough Evolutionary Relationships) Classification System is a unique resource that classifies genes by their functions, using published scientific experimental evidence and evolutionary relationships to predict function even in the absence of direct experimental evidence. Proteins are classified by expert biologists according to:

- **Gene families and subfamilies**, including annotated phylogenetic trees
- **Gene Ontology classes**: molecular function, biological process, cellular component
- **PANTHER Protein Classes**
- **Pathways**, including diagrams

PANTHER is part of the **Gene Ontology Reference Genome Project**.

PANTHER is supported by a research grant from the National Institute of General Medical Sciences [grant GM081084] and maintained by the Thomas lab at the University of Southern California.
The tool for gene list manipulations and gene ontology analysis

Welcome to DAVID 6.8

2003 - 2019

The Database forAnnotation, Visualization and Integrated Discovery (DAVID) v6.8 comprises a full Knowledgebase update to the sixth version of our original web-accessible programs. DAVID now provides a comprehensive set of functional annotation tools for investigators to understand biological meaning behind large list of genes. For any given gene list, DAVID tools are able to:

- Identify enriched biological themes, particularly GO terms
- Discover enriched functional-related gene groups
- Cluster redundant annotation terms
- Visualize genes on BioCarta & KEGG pathway maps
- Display related many-genes-to-many-terms on 2-D view.
- Search for other functionally related genes not in the list
- List interacting proteins
- Explore gene names in batch
- Link gene-disease associations
- Highlight protein functional domains and motifs
- Redirect to related literatures
- Convert gene identifiers from one type to another

What's Important in DAVID?

- Cite DAVID
- IDs of Affy Exon and Gene arrays supported
- Novel Classification Algorithms
- Pre-built Affymetrix and illumina backgrounds
- User's customized gene background
- Enhanced calculating speed

Statistics of DAVID

DAVID Citations (2003-2018)

> 38,000 Citations
The Internet resource OMIM (Online Mendelian Inheritance in Man) (https://omim.org/) was used to search for genes of susceptibility to the disease. A list of 229 genes was found, and the categories and analysis of gene ontologies were calculated using the PANTHER resource (Protein ANalysis THrough Evolutionary Relationships) (http://pantherdb.org/) (Mi et al., 2013).

Of the 229 original genes, 170 identifiers were recognized, 59 identifiers were not recognized or could not be unambiguously mapped. In total, 20851 genes were used in the PANTHER reference genome. With the help of PANTHER, we built a table of ontologies for categories of biological processes, in order to obtain the most informative results, the p-values were limited to E-20.

The most significant categories for the genes of Parkinson's disease are general regulation of cell death, regulation of cell death of neurons, regulation of apoptosis and programmed cell death, negative regulation of cell death, which confirms the etiology of the disease - death of substantia nigra neurons.

These data confirm the key etiological features of the disease, among which the central aspect of the pathophysiology of Parkinson's disease is the progressive death of dopamine neurons in the midbrain and their axonal projections.

(Published recently at “Biomedical chemistry” (2021) )
Some findings on gene ontologies for Parkinson’s disease using these tools

<table>
<thead>
<tr>
<th>GO molecular functions</th>
<th>Number of genes</th>
<th>P-value (corrected)</th>
</tr>
</thead>
<tbody>
<tr>
<td>enzyme binding</td>
<td>59</td>
<td>2,48E-12</td>
</tr>
<tr>
<td>protein binding</td>
<td>150</td>
<td>2,22E-11</td>
</tr>
<tr>
<td><strong>ubiquitin protein ligase binding</strong></td>
<td>22</td>
<td>7,08E-11</td>
</tr>
<tr>
<td><strong>ubiquitin-like protein ligase binding</strong></td>
<td>22</td>
<td>1,94E-10</td>
</tr>
<tr>
<td>signaling receptor binding</td>
<td>46</td>
<td>2,15E-09</td>
</tr>
<tr>
<td>heatshock protein binding</td>
<td>15</td>
<td>2,41E-09</td>
</tr>
<tr>
<td>binding</td>
<td>164</td>
<td>2,80E-08</td>
</tr>
<tr>
<td>identical protein binding</td>
<td>46</td>
<td>9,35E-08</td>
</tr>
<tr>
<td><strong>tau protein binding</strong></td>
<td>9</td>
<td>1,69E-06</td>
</tr>
<tr>
<td>kinase binding</td>
<td>26</td>
<td>4,88E-06</td>
</tr>
<tr>
<td>protein domain specific binding</td>
<td>25</td>
<td>6,55E-06</td>
</tr>
<tr>
<td>catalytic activity</td>
<td>86</td>
<td>1,38E-05</td>
</tr>
<tr>
<td>protein kinase binding</td>
<td>22</td>
<td>1,99E-04</td>
</tr>
</tbody>
</table>

The most significant are the categories of enzyme binding, protein binding, binding of ubiquitin ligase and ubiquitin-like proteins, and binding of signaling proteins and heat shock proteins.

Autophagy is one of the main pathways for intracellular degradation of α-synuclein, and current research shows that dysfunctional autophagy in Parkinson’s disease is one of the main risk factors for the development of the disease (Hale et al. 2016).
Plots of pointwise values of the categories of gene ontologies of Parkinson's disease genes, calculated using the GOST program for a list of 293 genes.

The most significant categories of this list of genes include the death of neurons, Lewy bodies, regulation of cell death, and the somatodendritic compartment.

Findings on gene ontologies for Parkinson's disease using these tools

http://biit.cs.ut.ee/gprofiler/gost
GeneMANIA and STRING-DB resources were used to reconstruct the gene network of interactions between Parkinson's disease genes. The figure shows a gene network of 187 genes for Parkinson's disease, reconstructed using GeneMANIA.

In the center of the constructed network are genes (proteins) that have a large number of connections with other elements - **SNCA, CASP3, GFRA1, HTT, PARK7**. This trend is supported by current studies of candidate gene associations (Billingsley et al. 2018), in which the most statistically significant signals associated with Parkinson's disease are common variants located close to SNCA, LRRK2 and MAPT, as well as low-frequency coding variants in GBA.
Gene network for the same disease reconstructed using STRING-DB

- Number of nodes: 268
- Number of edges: 2079
- Average node degree: 15.5
- Avg. local clustering coefficient: 0.428
- Expected number of edges: 942
- PPI enrichment p-value: < 1.0e-16

[Link to STRING-DB: https://string-db.org/]
According to the GeneCards resource, the following 10 genes are the most significant for Parkinson’s disease:

<table>
<thead>
<tr>
<th>Gene</th>
<th>Full name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SNCA</td>
</tr>
<tr>
<td>2</td>
<td>LRRK2</td>
</tr>
<tr>
<td>3</td>
<td>PRKN</td>
</tr>
<tr>
<td>4</td>
<td>PARK7</td>
</tr>
<tr>
<td>5</td>
<td>PINK1</td>
</tr>
<tr>
<td>6</td>
<td>MAPT</td>
</tr>
<tr>
<td>7</td>
<td>ATP13A2</td>
</tr>
<tr>
<td>8</td>
<td>GBA</td>
</tr>
<tr>
<td>9</td>
<td>APOE</td>
</tr>
<tr>
<td>10</td>
<td>APP</td>
</tr>
</tbody>
</table>

- **Gene**: SNCA
- **Full name**: Synuclein Alpha
- **Score**: 157.73

The first place in this hierarchical list is occupied by the SNCA gene encoding the alpha-synuclein protein, mutations in this gene lead to the development of autosomal dominant forms of the disease, the severity of the disease correlates with the number of copies of the SNCA gene. Mutations in the LRRK2 gene have been identified as the causes of the autosomal dominant nature of Parkinson's disease as the most common monogenic form of the disease identified to date (Paisán-Ruiz et al., 2004; Zimprich et al., 2004). Genetic variants in LRRK2 are associated with most of all known inherited manifestations of Parkinson's disease.
Approach for Diseases network reconstruction by genes and back - diseasome bipartite network

The human disease network
PNAS, 2007 104 (21) 8685-8690;
https://doi.org/10.1073/pnas.0701361104

Generating human disease–drug networks

Online tools (some not free)
GeneGO (www.genego.com), Ingenuity (www.ingenuity.com) and Biocarta (www.biocarta.com).

The human disease network
PNAS, 2007 104 (21) 8685-8690;
https://doi.org/10.1073/pnas.0701361104
Reconstruction of gene networks associated with Parkinson's disease leads to the identification of network structures. Such a discovery of functional connections opens the way to the creation of new drugs. The same network modeling approach is applied to series on complex diseases.

Acknowledgements

Sechenov University, Novosibirsk State University

Students:

A.G.Galieva  
E.N.Ivanova  
N.A.Makarov  
A.I.Dergilev

The development of education programs using online bioinformatics tools was supported in 2022 by Potanin Foundation (grant ГК22-000797):

The development of education programs using online bioinformatics tools was supported in 2022 by Potanin Foundation (grant ГК22-000797):
Special journal issue – **MPDI International Journal of Molecular Sciences**

For the conference participants!

[https://www.mdpi.com/journal/ijms/special_issues/Medical_Genetics_2022](https://www.mdpi.com/journal/ijms/special_issues/Medical_Genetics_2022)
Post-conference publication is a way to connect fix the conference results, present the work for wider audience.

Special journal issues help in faster publications and free publications as well.

**Vavilov Journal of Genetics and Breeding** is a Platinum Open Access peer-reviewed scholarly journal, which does not charge author fees. The journal has been published since 1997 (until 2011, as The Herald of Vavilov Society for Geneticists and Breeding Scientists) eight issues per year. The journal publishes works on all fields of genetics, breeding, and related sciences. The scope of the journal includes: plant genetics, animal genetics, plant breeding, animal breeding, human genetics, medical genetics, neurogenetics, paleogenetics, microbial genetics and breeding, symbiogenetics.

**Journal of Integrative Bioinformatics**

ISSN: 1613-4516
First published: January 1, 2017
Editors-in-chief: Falk Schreiber, Ralf Hofestäd't
Edited by: Jan Baumbach, Ming Chen, Yuriy Orlov, Jens Allmer
Managing Editor: Björn Sommer

Publications in Russian journals (free)

https://vavilov.elpub.ru/jour

http://pismavavilov.ru/

https://vavilov.elpub.ru/jour

https://www.sechenovmedj.com/jour/index

https://vavilov.elpub.ru/jour
Welcome to the conference in live format – BGRS-2022 (Bioinformatics of Genome Regulation and Structure) Novosibirsk, Russia 4-8 July 2022

XXVIII Symposium on Bioinformatics and Computer-Aided Drug Discovery

https://bgrssb.icgbio.ru/2022/

Longest conference series on bioinformatics in Russia, in Novosibirsk (biannually since 1998)

26/05/2022