



# Editorial: Bioinformatics of Genome Regulation, Volume II

Yuriy L. Orlov<sup>1,2\*</sup>, Anastasia A. Anashkina<sup>1,3</sup>, Tatiana V. Tatarinova<sup>4</sup> and Ancha V. Baranova<sup>5,6</sup>

<sup>1</sup>The Digital Health Institute, I.M.Sechenov First Moscow State Medical University (Sechenov University), Moscow, Russia,

<sup>2</sup>Agrobiotechnology Department, Agrarian and Technological Institute, Peoples' Friendship University of Russia (RUDN University), Moscow, Russia, <sup>3</sup>Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russia,

<sup>4</sup>Natural Science Division, La Verne University, La Verne, CA, United States, <sup>5</sup>School of Systems Biology, George Mason University, Fairfax, VA, United States, <sup>6</sup>Research Centre for Medical Genetics, Moscow, Russia

**Keywords:** bioinformatics, transcriptomics, plant science, gene networks, gene expression regulation, genetics, computational genomics

## Editorial on the Research Topic

### Bioinformatics of Genome Regulation, Volume II

This Research Topic Volume II, “Bioinformatics of Genome Regulation,” continues the studies in the field of bioinformatics of gene expression presented initially at *Frontiers in Genetics* journal (<https://www.frontiersin.org/research-topics/8383/bioinformatics-of-genome-regulation-and-systems-biology>) and then in Volume I (<https://www.frontiersin.org/research-topics/14266/bioinformatics-of-genome-regulation-volume-i>). The materials presented here were discussed in a BGRS\SB (Bioinformatics of Genome Regulation and Structure Systems Biology) conference series in Novosibirsk, Russia (<https://bgrssb.icgbio.ru/2020/>). The BGRS is the central event in the computational genetics field, held in Novosibirsk, Russia, every other year since 1998 (Orlov et al., 2015). The publications were later completed by new studies on computational methods of gene expression analysis regulation. Starting in 2018, materials of the conference materials in genetics and genomics were presented in *Frontiers in Genetics*, and due to popular demand in 2021 it was extended as the Volume II. The BGRS conference series have been presented at special journal issues earlier (Orlov et al., 2016; Tatarinova et al., 2019; Orlov et al., 2015; Orlov et al., 2019a; Orlov et al., 2019b; Baranova et al., 2019) and recently (Tatarinova et al., 2020; Orlov and Baranova, 2020; Orlov et al., 2020; Orlov et al., 2021a; Orlov et al., 2021b). We have to acknowledge “Bioinformatics of Gene Regulations and Structure” special issue at *MDPI IJMS* ([https://www.mdpi.com/journal/ijms/special\\_issues/Bioinformatics\\_Genomics](https://www.mdpi.com/journal/ijms/special_issues/Bioinformatics_Genomics)), as well as *PeerJ* journal BGRS-2020 collection (<https://peerj.com/collections/72-bgrs-sb-2020>).

This research topic presents seventeen papers on medical genomics applications, new bioinformatics tools and applications to laboratory animal models, and plant sciences.

Biomedical papers start from applications to cancer studies. Xiaorong Yang et al. discussed the interplay between human diseases at multiple biological levels. The authors show the role of schizophrenia in the pathological development of myocardial infarction, suggesting its role in promoting the development and progression of myocardial infarction at different levels, including genes, small molecules, and complex molecules. Pathway analysis revealed nine genes connecting these diseases.

Maxim Sorokin et al. proposed an algorithm that identifies functional roles of the pathway components and applied it to annotate 3,044 human molecular pathways extracted from the Biocarta, Reactome, KEGG, Qiagen Pathway Central, NCI, and HumanCYC databases. The resulting knowledgebase may be used to calculate the levels of activation for individual pathways and to establish large-scale profiles of the signaling, metabolic, and DNA repair regulation using high throughput gene expression data, as was presented recently (Wang et al., 2021).

## OPEN ACCESS

### Edited and reviewed by:

Richard D. Emes,  
University of Nottingham,  
United Kingdom

### \*Correspondence:

Yuriy L. Orlov  
orlov@d-health.institute

### Specialty section:

This article was submitted to  
Computational Genomics,  
a section of the journal  
*Frontiers in Genetics*

Received: 14 October 2021

Accepted: 25 October 2021

Published: 08 November 2021

### Citation:

Orlov YL, Anashkina AA, Tatarinova TV  
and Baranova AV (2021) Editorial:  
Bioinformatics of Genome Regulation,  
Volume II.  
*Front. Genet.* 12:795257.  
doi: 10.3389/fgene.2021.795257

Anna E. Letiagina et al. presented a new bioinformatics application on massively parallel reporter assays (MPRAs). The assays are based on the construction of reporter plasmid libraries. The authors present a pipeline for processing raw MPRA data obtained by NGS for reporter construct libraries with a priori unknown sequences of a region of interest and a barcode in the plasmid, located outside and within the transcription unit. The pipeline robustly identifies unambiguous barcodes, calculates the normalized expression level, and provides a graphical visualization of the processed data.

Maria M. Litvinova et al. have studied genes polymorphisms associations with calcium urolithiasis development. Authors found statistically significant associations between calcium urolithiasis and the polymorphisms in *CASR* rs1042636, *CALCR* rs1801197, and *ORAI1* rs6486795.

Yi He et al. explored the regulatory network among circRNA-miRNA-mRNA in osteosarcoma. The authors interrogated NCBI GEO osteosarcoma datasets and extracted circRNA, microRNA, and mRNA expression profiles. Downstream network analysis was performed with CircInteractome (Circular RNA Interactome) and mirDIP (microRNA Data Integration Portal) databases. Authors show that quinacridine, thalidomide, and zonisamide may be potential drugs for the treatment of osteosarcoma.

Stepan Nersisyan et al. investigated the effects of hypoxia on pathophysiological processes, including cancer progression and metastasis formation. This work continued the earlier studies on miRNA regulatory interactions in cancer (Shkurnikov et al., 2019). The landscape of hypoxia-induced miRNA and mRNA expression alterations was studied in human colorectal cancer cell lines (HT-29 and Caco-2) to show that miR-148a downregulation contributes to poor survival due to overexpression of the *ITGA5* and *PRNP* genes.

Sergey Nikulin et al. have studied gene expression in breast cancer cells. The authors showed that malignant breast tumors with reduced expression of the *ELOVL5* and *IGFBP6* genes could metastasize at a higher probability due to a more efficient invasion of tumor cells. In addition, a set of novel computational techniques was developed for deciphering gene expression regulation.

Igor B. Rogozin et al. discussed fundamental biochemical mechanisms mutations in cancer driver genes. A large fraction of these mutations arose due to the off-target activity of DNA/RNA editing cytosine deaminases, followed by the replication/repair of edited sites by DNA polymerases. Using methylation data from malignant lymphomas, the authors showed that driver genes are subject to different (de)methylation processes than non-driver genes.

Natalya V. Klimova et al., using previously published web-tool SNP\_TATA\_Comparator (Ponomarenko et al., 2017), conducted a genome-wide study of single-nucleotide polymorphisms (SNPs) within core promoters of 68 human rheumatoid arthritis-related genes. They show that the disruptive natural selection of human immunostimulatory and immunosuppressive genes concurrently elevates and reduces the risk of rheumatoid arthritis. The authors hypothesize that rheumatoid arthritis in humans could be a self-domestication syndrome referring to evolution patterns in domestic animals (Chadaeva et al., 2021).

James Sweet-Jones et al. applied genetic tools to livestock breeding on the example of the Welsh mountain sheep breeds. Genotyping data from 317 individuals representing 15 Welsh sheep breeds were used alongside the whole-genome resequencing data of 14 species from the same set to scan for the signatures of selection and candidate genetic variants using haplotype SNP-based approaches. The authors found new variants in genes with potential functional consequences to the adaptation of local sheep to their environments in Wales. The study continues in new markers search in animal adaptation to cold climate (Igoshin et al., 2021).

Lilit Nersisyan et al. investigated telomere maintenance mechanisms for studying cancers and designing therapies.

In Brief Research Report, Elena V. Ignatieva et al. presented a new database - a catalog of human genes associated with pathologies of the sperm. It contains data genes related to male fertility and their functional annotation based on the literature data and clinical trials (Kolmykov et al., 2021). Functional annotation of genes from the catalog showed that spermatogenic failure could be associated with mutations in genes that control biological processes essential for spermiogenesis (such as DNA metabolism, cell division). Azoospermia can be caused by mutations in genes that control cellular responses to unfavorable conditions (stress factors, including oxidative stress and exposure to toxins).

Ekaterina Ilgisonis et al. considered a problem of transcriptome annotation by sequencing. They compared the results obtained from different transcriptome analysis platforms (quantitative polymerase chain reaction, Illumina RNASeq, and Oxford Nanopore Technologies MinION) for the transcriptome encoded by human chromosome 18 using the same sample types. The combination of Illumina RNASeq and MinION nanopore technologies reduced the probability of false-positive detection of low-copy transcripts due to the simultaneous confirmation of the presence of a transcript by the two fundamentally different technologies: short reads essential for reliable detection and long-read sequencing data.

The next group of articles in this Research Topic performed gene expression analysis in plants. This science field was presented at the bioinformatics conference series in Novosibirsk (Computer plant biology Session of BGRS conference, and the Plantgen conference series, <https://conf.icgbio.ru/plantgen2021/>) (Orlov, 2019; Orlov et al., 2019b).

Nann Miky Moh Moh et al. studied miRNAs and lncRNAs from the mango (*Mangifera indica* L.). Although mango is a popular food having pharmacological potential, its non-coding RNA data were limited. For the first time, a large-scale study identified nearly a hundred miRNAs and over 7,000 temperature-responsive lncRNAs. Characterization of target genes for these ncRNAs was performed.

Ekaterina M. Dvorianinova et al. presented the application of sequencing technologies to study plant pathogens of flax (*Linum usitatissimum* L.). Genome Assembly of the pathogen *Fusarium oxysporum* f. sp. *lini* was presented (Krasnov et al., 2020). Due to *F. oxysporum* f. sp. *lini* includes many genotypes, it is of high significance to study the origins of pathogenicity at the molecular level. This work mainly focused on genome sequencing of strains

of the flax pathogen *F. oxysporum* f. sp. *lini*, possessing diverse pathogenicity degrees, on two platforms—Oxford Nanopore Technologies and Illumina. Sequencing using these two platforms proved to effectively achieve high-quality genome assemblies for complex plant genomes (Melnikova et al., 2021).

Victoria A. Scobeyeva et al. studied patterns of evolution in plant genome on the example of *Allium* species. *Alliums* are widespread and diversified; they are adapted to various habitats, from shady forests to open steppes. The genes present in chloroplast genomes (plastomes) play fundamental roles for photosynthetic plants. Plastome traits could thus be associated with geophysical abiotic characteristics of habitats. The authors compared their data with previously published plastomes and provided our interpretation of *Allium* plastome genes' annotations. They can hypothesize that adaptive evolution in genes, coding subunits of NADH-plastoquinone oxidoreductase could be driven by abiotic factors of alpine habitats, especially by intensive light and UV radiation.

Elena N. Pushkova et al. studied Allele-Specific Expression in plants. Transcriptome sequencing of plant tissues from the male and female trees of *Populus × sibirica* and genome sequencing of the same plants were performed first. Targeted sequencing of sex-determining region (SDR) genes such as *CLC* (Chloride channel protein *CLC-c* on a representative set of trees confirmed the sex-associated allele-specific expression in generative and vegetative tissues of *P. × sibirica*.

Overall, we are proud of the continuing Research Topic at Frontiers in Genetics we collated. Biomedical applications for gene

expression studies in chronic diseases are presented in ongoing Research Topic “High-throughput sequencing-based investigation of chronic disease markers and mechanisms” (<https://www.frontiersin.org/research-topics/21036/high-throughput-sequencing-based-investigation-of-chronic-disease-markers-and-mechanisms>). We hope you will find this paper collection a stimulating reading and consider coming to the next BGRS\SB conferences in Novosibirsk, Russia (<https://bgrssb.icgbio.ru/2022/>).

## AUTHOR CONTRIBUTIONS

YO, AA, TT, and AB organized the Research Topic as guest editors, supervised the reviewing of the manuscripts. All the authors wrote this Editorial paper. All authors contributed to the article and approved the submitted version.

## ACKNOWLEDGMENTS

The guest editors are grateful to the authors contributing to this special issue papers collection and thank all the reviewers who helped improve the manuscripts. The BGRS-2020 conference organization was supported by Novosibirsk State University and the Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia. The publication has been prepared with the support of the RUDN University Strategic Academic Leadership Program (recipient YO).

## REFERENCES

- Baranova, A. V., Klimontov, V. V., Letyagin, A. Y., and Orlov, Y. L. (2019). Medical Genomics Research at BGRS-2018. *BMC Med. Genomics*. 12 (Suppl. 2), 36. doi:10.1186/s12920-019-0480-0
- Chadaeva, I., Ponomarenko, P., Kozhemyakina, R., Suslov, V., Bogomolov, A., Klimova, N., et al. (2021). Domestication Explains Two-Thirds of Differential-Gene-Expression Variance between Domestic and Wild Animals; the Remaining One-Third Reflects Intraspecific and Interspecific Variation. *Animals*. 11 (9), 2667. doi:10.3390/ani11092667
- Igoshin, A., Yudin, N., Aitnazarov, R., Yurchenko, A. A., and Larkin, D. M. (2021). Whole-Genome Resequencing Points to Candidate DNA Loci Affecting Body Temperature Under Cold Stress in Siberian Cattle Populations. *Life*. 11 (9), 959. doi:10.3390/life11090959
- Kolmykov, S., Vasiliev, G., Osadchuk, L., Kleschev, M., and Osadchuk, A. (2021). Whole-Exome Sequencing Analysis of Human Semen Quality in Russian Multiethnic Population. *Front. Genet.* 12, 662846. doi:10.3389/fgene.2021.662846
- Krasnov, G. S., Pushkova, E. N., Novakovskiy, R. O., Kudryavtseva, L. P., Rozhmina, T. A., Dvorianinova, E. M., et al. (2020). High-Quality Genome Assembly of *Fusarium Oxysporum* F. Sp. *Lini*. *Front. Genet.* 11, 959. doi:10.3389/fgene.2020.00959
- Melnikova, N. V., Pushkova, E. N., Dvorianinova, E. M., Beniaminov, A. D., Novakovskiy, R. O., Povkhova, L. V., et al. (2021). Genome Assembly and Sex-Determining Region of Male and Female *Populus × Sibirica*. *Front. Plant Sci.* 12, 625416. doi:10.3389/fpls.2021.625416
- Orlov, Y. L. (2019). 5th International Scientific Conference of "Plant Genetics, Genomics, Bioinformatics, and Biotechnology" (24-29 June 2019, Novosibirsk, Russia). *J. Food Qual. Hazards Control*. 6, 41. doi:10.18502/jfqhc.6.1.458
- Orlov, Y. L., Anashkina, A. A., Klimontov, V. V., and Baranova, A. V. (2021a). Medical Genetics, Genomics and Bioinformatics Aid in Understanding expression studies in chronic diseases are presented in ongoing Research Topic “High-throughput sequencing-based investigation of chronic disease markers and mechanisms” (<https://www.frontiersin.org/research-topics/21036/high-throughput-sequencing-based-investigation-of-chronic-disease-markers-and-mechanisms>). We hope you will find this paper collection a stimulating reading and consider coming to the next BGRS\SB conferences in Novosibirsk, Russia (<https://bgrssb.icgbio.ru/2022/>).
- Molecular Mechanisms of Human Diseases. *Int. J. Mol. Sci.* 22, 9962. doi:10.3390/ijms22189962
- Orlov, Y. L., Galieva, A. G., Orlova, N. G., Ivanova, E. N., Mozyleva, Y. A., and Anashkina, A. A. (2021b). Reconstruction of Gene Network Associated With Parkinson Disease for Gene Targets Search. *Biomed. Khim.* 67 (3), 222–230. doi:10.18097/PBMC20216703222
- Orlov, Y. L., and Baranova, A. V. (2020). Editorial: Bioinformatics of Genome Regulation and Systems Biology. *Front. Genet.* 11, 625. doi:10.3389/fgene.2020.00625
- Orlov, Y. L., Baranova, A. V., and Markel, A. L. (2016). Computational Models in Genetics at BGRS\SB-2016: Introductory Note. *BMC Genet.* 17 (Suppl. 3), 155. doi:10.1186/s12863-016-0465-3
- Orlov, Y. L., Baranova, A. V., and Tatarinova, T. V. (2020). Bioinformatics Methods in Medical Genetics and Genomics. *Int. J. Mol. Sci.* 21 (17), 6224. doi:10.3390/ijms21176224
- Orlov, Y. L., Hofestädt, R. M., and Kolchanov, N. A. (2015). Introductory Note for BGRS\SB-2014 Special Issue. *J. Bioinform. Comput. Biol.* 13, 1502001. doi:10.1142/S0219720015020011
- Orlov, Y. L., Hofestädt, R., and Tatarinova, T. V. (2019a). Bioinformatics Research at BGRS\SB-2018. *J. Bioinform. Comput. Biol.* 17, 1902001. doi:10.1142/S0219720019020013
- Orlov, Y. L., Salina, E. A., Eslami, G., and Kochetov, A. V. (2019b). Plant Biology Research at BGRS-2018. *BMC Plant Biol.* 19 (Suppl. 1), 56. doi:10.1186/s12870-019-1634-0
- Ponomarenko, P., Chadaeva, I., Rasskazov, D. A., Sharypova, E., Kashina, E. V., Drachkova, I., et al. (2017). Candidate SNP Markers of Familial and Sporadic Alzheimer's Diseases Are Predicted by a Significant Change in the Affinity of TATA-Binding Protein for Human Gene Promoters. *Front. Aging Neurosci.* 9, 231. doi:10.3389/fnagi.2017.00231
- Shkurnikov, M., Nikulin, S., Nersisyan, S., Poloznikov, A., Zaidi, S., Baranova, A., et al. (2019). LAMA4-Regulating miR-4274 and its Host Gene SORCS2 Play a Role in IGF1BP6-Dependent Effects on Phenotype of Basal-Like Breast Cancer. *Front. Mol. Biosci.* 6, 122. doi:10.3389/fmolb.2019.00122

- Tatarinova, T. V., Baranova, A. V., Anashkina, A. A., and Orlov, Y. L. (2020). Genomics and Systems Biology at the "Century of Human Population Genetics" Conference. *BMC Genomics*. 21 (Suppl. 7), 592. doi:10.1186/s12864-020-06993-1
- Tatarinova, T. V., Chen, M., and Orlov, Y. L. (2019). Bioinformatics Research at BGRS-2018. *BMC Bioinformatics*. 20 (Suppl. 1), 33. doi:10.1186/s12859-018-2566-7
- Wang, Y., Tong, Z., Zhang, W., Zhang, W., Buzdin, A., Mu, X., et al. (2021). FDA-Approved and Emerging Next Generation Predictive Biomarkers for Immune Checkpoint Inhibitors in Cancer Patients. *Front. Oncol.* 11, 683419. doi:10.3389/fonc.2021.683419

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2021 Orlov, Anashkina, Tatarinova and Baranova. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*