

## THE “MICROBITECH” DIGITAL PLATFORM: ARCHITECTURE AND PURPOSE

P. S. Demenkov, A. M. Mukhin, V. A. Ivanisenko, S. A. Lashin, N. A. Kolchanov

Kurchatov Genome Center of the Institute of Cytology and Genetics of the Siberian Branch  
of the Russian Academy of Sciences (KGC ICG SB RAS)  
630090, Novosibirsk, Russia

---

---

DOI: 10.24412/2073-0667-2024-4-27-36

EDN: HXJHCY

The article examines the architecture of the developed digital platform “Microbitech” for solving a wide range of problems in systems and structural biology, discusses the use of software integrated into the platform for processing and analyzing large volumes of genetic information, as well as for predicting the structure and function of proteins. The use of the “Microbitech” digital platform allows increasing the productivity of research, improving the accuracy of data analysis, and contributing to the development of new research methods.

**Key words:** information platform, systems biology, bioinformatics.

### References

1. Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N. Digital business strategy: toward a next generation of insights // *MIS quarterly*, 2013. P. 471–482.
2. Yoo, Y., Henfridsson, O., & Lyytinen, K. Research commentary — the new organizing logic of digital innovation: an agenda for information systems research // *Information systems research*, 2010. N 21(4). P. 724–735.
3. Goecks, J., Nekrutenko, A., Taylor, J., & Galaxy Team. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences // *Genome biology*, 2010. N 11(8), R86.
4. Gruning, B. A., Rasche, E., Rebollo-Jaramillo, B., Eberhard, C., Houwaart, T., Chilton, J., ... & Backofen, R. Jupyter and Galaxy: Easing entry barriers into complex data analyses for biomedical researchers // *PLoS computational biology*, 2017. N 13(5), e1005425.
5. Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Cech, M., ... & Goecks, J. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update // *Nucleic acids research*, 2016. N 44(W1), W3–W10.
6. Lowe, R., Shirley, N., Bleackley, M., Dolan, S., & Shafee, T. Transcriptomics technologies // *PLoS computational biology*, 2017.
7. Gruning, B., Chilton, J., Koster, J., Dale, R., Soranzo, N., van den Beek, M., ... & Backofen, R. Practical computational reproducibility in the life sciences // *Cell systems*, 2019. N 8(3). P. 183–188.

---

This work was funded by the Ministry of Science and Higher Education of the Russian Federation within the framework of the project “Kurchatov Center for World-Class Genomic Research” No 075-15-2019-1662 dated 10/31/2019.

8. Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Cech, M., ... & Blankenberg, D. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update // *Nucleic acids research*, 2018. N 46(W1), W537–W544.
9. Kluyver, T., Ragan-Kelley, B., Perez, F., Granger, B.E., Bussonnier, M., Frederic, J., ... & Ivanov, P. Jupyter Notebooks — a publishing format for reproducible computational workflows // In *ELPUB*, 2016. P. 87–90.
10. Zaharia, M., Xin, R. S., Wendell, P., Das, T., Armbrust, M., Dave, A., ... & Ghodsi, A. Apache Spark: a unified engine for big data processing // *Communications of the ACM*, 2016. N 59(11). P. 56–65.
11. Varia, J., & Mathew, S. Overview of Amazon Web Services. Amazon Web Services, 2014. N 16.
12. Chee, B. J., Franklin, J. C., & Chee, B. J. Cloud computing: Technologies and strategies of the ubiquitous data center // CRC Press, 2009.
13. Pronozin A. Y., Salina E. A., Afonnikov D. A. GBS-DP: a bioinformatics pipeline for processing data coming from genotyping by sequencing // *Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding*. 2023; N 27(7). P. 737–745. DOI 10.18699/VJGB-23-86
14. Ivanisenko V. A., Saik O. V., Ivanisenko N. V., Tiys E. S., Ivanisenko T. V., Demenkov P. S., Kolchanov N. A. ANDSsystem: an associative network discovery system for automated literature mining in the field of biology // *BMC Syst Biol*. 2015;9 Suppl 2(Suppl 2):S2. DOI: 10.1186/1752-0509-9-S2-S2.
15. Ivanisenko V. A., Demenkov P. S., Ivanisenko T. V., Mishchenko E. L., Saik O. V. A new version of the ANDSsystem tool for automatic extraction of knowledge from scientific publications with expanded functionality for reconstruction of associative gene networks by considering tissue-specific gene expression // *BMC Bioinformatics*. 2019. N 20(Suppl 1). P. 34. DOI: 10.1186/s12859-018-2567-6.

## ЦИФРОВАЯ ПЛАТФОРМА «МИКРОБИОТЕХ»: АРХИТЕКТУРА И НАЗНАЧЕНИЕ

П. С. Деменков, А. М. Мухин, В. А. Иванисенко, С. А. Лашин, Н. А. Колчанов

Курчатовский Геномный Центр Института цитологии и генетики Сибирского отделения  
Российской академии наук (КГЦ ИЦиГ СО РАН)  
630090, Новосибирск, Россия

УДК 004.4

DOI: 10.24412/2073-0667-2024-4-27-36

EDN: НХЈНСУ

В статье рассматривается архитектура разработанной цифровой платформы «Микробиотех» для решения широкого круга задач системной и структурной биологии, обсуждается использование программного обеспечения, интегрированного в платформу для обработки и анализа больших объемов генетической информации, а также для предсказания структуры и функции белков. Использование цифровой платформы «Микробиотех» позволяет увеличить производительность исследования, улучшить точность анализа данных и способствовать развитию новых методов исследования.

**Ключевые слова:** информационная платформа, системная биология, биоинформатика.

**Введение.** Современные исследования в области управления информационными технологиями все чаще фокусируются на концепции интегрированных цифровых платформ (ИЦП). Данные платформы представляют собой комплексные программно-аппаратные решения, обеспечивающие интеграцию и координацию различных цифровых технологий и сервисов в рамках единой системы [1, 2].

Основными характеристиками таких платформ являются:

- 1) Интеграция: платформа интегрирует различные приложения, данные и процессы в единую систему, обеспечивая их взаимодействие и обмен информацией.
- 2) Централизованное управление: платформа предоставляет единую точку управления для всех подключенных компонентов, упрощая администрирование и контроль.
- 3) Масштабируемость: платформы обладают высокой масштабируемостью, позволяя добавлять или удалять функциональные модули по мере необходимости.
- 4) Гибкость: интегрированные платформы отличаются гибкостью, позволяя адаптировать их под конкретные потребности пользователя.
- 5) Эффективность: объединение различных систем в единую платформу повышает общую эффективность и производительность.
- 6) Аналитика и отчетность: платформы предоставляют инструменты для сбора, анализа и визуализации данных, что помогает принимать обоснованные решения.

Данная работа финансировалась Министерством науки и высшего образования Российской Федерации в рамках проекта «Курчатовский центр геномных исследований мирового уровня» № 075-15-2019-1662 от 31.10.2019.

Интегрированные цифровые платформы обладают рядом преимуществ:

- Упрощение доступа к вычислительным ресурсам и аналитическим инструментам для ученых без специальной подготовки [3];
- Возможность воспроизводимости научных исследований за счет документирования рабочих процессов [4];
- Интеграция различных программных компонентов в единую среду, что повышает эффективность работы [5].

К недостаткам ИЦП стоит отнести:

- Зависимость от стабильности и доступности платформы, что может создавать проблемы при сбоях или отключениях [6];
- Необходимость в обучении пользователей работе с платформой, что может быть барьером для некоторых ученых [7];
- Ограниченность функциональности платформы по сравнению с использованием отдельных программных инструментов [8].

Одним из ярких примеров таких платформ является Galaxy project [8]. Galaxy project представляет собой веб-ориентированную платформу с открытым исходным кодом, которая позволяет ученым выполнять широкий спектр биоинформатических анализов без необходимости глубоких познаний в программировании.

Помимо Galaxy project, существует ряд других примеров интегрированных цифровых платформ, используемых в научных исследованиях:

1. Jupyter Notebook [9]. Это веб-приложение, позволяющее создавать и обмениваться документами, содержащими программный код, визуализации и повествовательный текст. Jupyter Notebook широко используется в различных областях науки, включая машинное обучение, анализ данных и визуализацию.

2. Apache Spark [10]. Это распределенная вычислительная платформа, предназначенная для обработки больших объемов данных. Spark интегрирует различные компоненты, такие как Spark SQL для работы с структурированными данными, Spark Streaming для обработки потоковых данных и MLlib для машинного обучения.

3. Amazon Web Services (AWS) [11]. Это облачная платформа, предоставляющая широкий спектр услуг, включая вычислительные ресурсы, хранилища данных, аналитические инструменты и средства разработки. AWS позволяет ученым масштабировать вычислительные мощности и хранилище по мере необходимости.

4. Microsoft Azure [12]. Это еще одна облачная платформа, предлагающая аналогичные услуги, что и AWS. Azure предоставляет интегрированные инструменты для анализа данных, машинного обучения и визуализации. Каждая из этих платформ имеет свои преимущества и недостатки, и выбор подходящей платформы зависит от конкретных потребностей исследователей и особенностей их проектов.

В целом, интегрированные цифровые платформы предоставляют ученым мощные инструменты для проведения исследований, но также требуют определенной адаптации и преодоления технических барьеров.

### **1. Архитектура интегрированной цифровой платформы «Микробиотех».**

Огромные объемы и сложность накопленных в настоящее время больших генетических данных требуют создания информационно-программных комплексов на основе нового поколения методов биоинформатики и системной компьютерной биологии, использующих вычислительные конвейеры, реализующих сложные сценарии анализа и интегрирующих

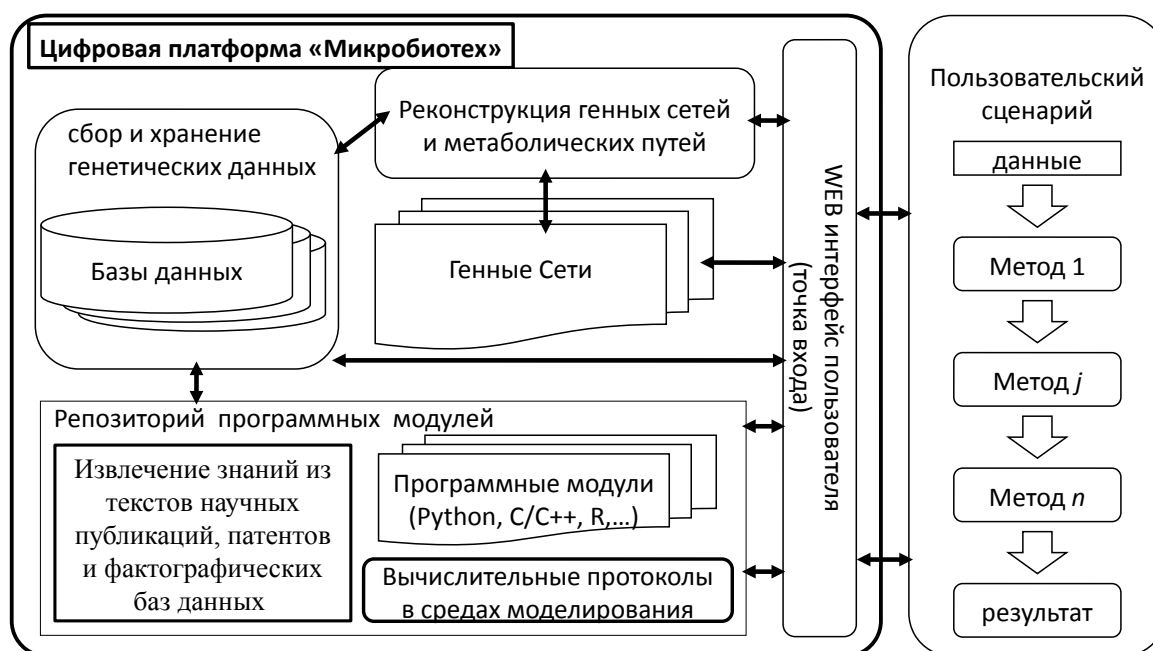


Рис. 1. Принципиальная схема организации ИЦП «Микробиотех»

большое количество различных программных продуктов и баз данных, в том числе основанных на использовании методов искусственного интеллекта.

В цифровой платформе «Микробиотех» в качестве базового средства интеграции ресурсов взята платформа Galaxy (<https://galaxyproject.org/>), предоставляющая всю необходимую программную инфраструктуру для создания биоинформатических конвейеров и обмена данными между модулями. Galaxy представляет собой проект с открытым исходным кодом, при этом лицензия допускает свободный академический доступ — Academic Free License version 3.0. (<https://galaxyproject.org/admin/license/>). Серверная часть программного комплекса Galaxy реализована на языке программирования Python3. Клиентская часть реализована с использованием платформы VueJS. Также для системы были развернуты базы данных в СУБД PostgreSQL для сохранения состояния системы и логирования задач. Система Conda использована для задач создания узлов вычислительных конвейеров и запуска биоинформатических программ. Сервер Galaxy был развернут в двух вариантах соответственно на двух аппаратных системах:

1. «Рабочая среда», то есть непосредственно на головном узле вычислительного кластера ИЦиГ СО РАН. Развернут с использованием готового образца Docker от разработчиков.
2. Тестовый физический сервер для отладки новых модулей и программ. У него есть доступ к вычислительному кластеру по сети ИЦиГ СО РАН. Система развернута без использования систем виртуализации, таких как KVM, Xen, LXC или Docker. Доступ к сети кластера ИЦиГ СО РАН позволяет использовать сетевое хранилище и вычислительные узлы для задач отладки новых биоинформатических конвейеров.

На текущий момент на обеих системах установлен базовый набор биоинформатических программ, а также некоторые дополнительные инструменты, которые используются на основном сервере Galaxy. Принципиальная схема организации ИЦП «Микробиотех» представлена на рис. 1.

Одним из крупных разделов ИЦП «Микробиотех» является репозиторий тематически ориентированного программного обеспечения, алгоритмов, программ и баз данных для решения задач биоинформатики и системной компьютерной биологии. Программные продукты ИЦП используются как модули в пользовательских сценариях или конвейерах обработки больших генетических данных (геномных, транскриптомных, протеомных, метаболомных). В ИЦП интегрированы базы больших генетических данных, как внешние, так и собственные, разработанные в ИЦиГ СО РАН. Блок «Генные сети» содержит генные сети молекулярно-генетических систем и метаболических путей микроорганизмов, соответствующих различным иерархическим уровням организации живых систем (геномам, генам, РНК, белкам, организмам, популяциям, экосистемам).

Разрабатываемая ИЦП «Микробиотех» обладает следующим функционалом:

- сбор и хранение больших генетических данных (геномных, транскриптомных, протеомных, метаболомных);
- реализация сложных сценариев конвейерной (и сетевой) обработки и анализа больших генетических данных, в том числе, с использованием технологий искусственного интеллекта (ИИ);
- математическое и компьютерное моделирование молекулярно-генетических систем и процессов;
- компьютерная реконструкция генных сетей, молекулярно-генетических систем и метаболических путей человека, животных, растений и микроорганизмов;
- автоматическое извлечение геномных и генетических знаний из текстов научных публикаций и фактографических баз данных на основе методов искусственного интеллекта (ИИ) и глубокого машинного обучения.

**2. Архитектура модуля исполнения пользовательских сценариев обработки данных.** ИЦП «Микробиотех» — это веб-ориентированная платформа для биоинформатических исследований, которая позволяет ученым выполнять сложные вычислительные задачи без необходимости глубокого понимания интерфейса командной строки или написания кода. Когда пользователь запускает анализ в ИЦП, система должна обеспечить эффективное распределение вычислительных ресурсов для выполнения этой задачи. В ЦКП «Биоинформатика» используется кластер с системой управления заданиями Slurm. Схема работы ИЦП «Микробиотех» выглядит следующим образом:

1. Пользователь запускает анализ в веб-интерфейсе ИЦП.
2. ИЦП отправляет запрос на выполнение задачи в систему Slurm.
3. Slurm получает запрос и анализирует доступные вычислительные ресурсы (CPU, память, время и т. д.).
4. Slurm выделяет необходимые ресурсы для выполнения задачи и запускает ее на соответствующем узле кластера.
5. Задача выполняется на выделенных ресурсах, используя программное обеспечение и инструменты, установленные в ИЦП.
6. По завершении задачи, Slurm возвращает результаты обратно в ИЦП.
7. ИЦП отображает результаты анализа в веб-интерфейсе для пользователя.

Таким образом, обеспечивается эффективное распределение вычислительных ресурсов кластера для выполнения задач, запускаемых через ИЦП. Это позволяет пользователям использовать мощные вычислительные ресурсы без необходимости управления ими напрямую.

## Схема запуска конвейеров на кластере

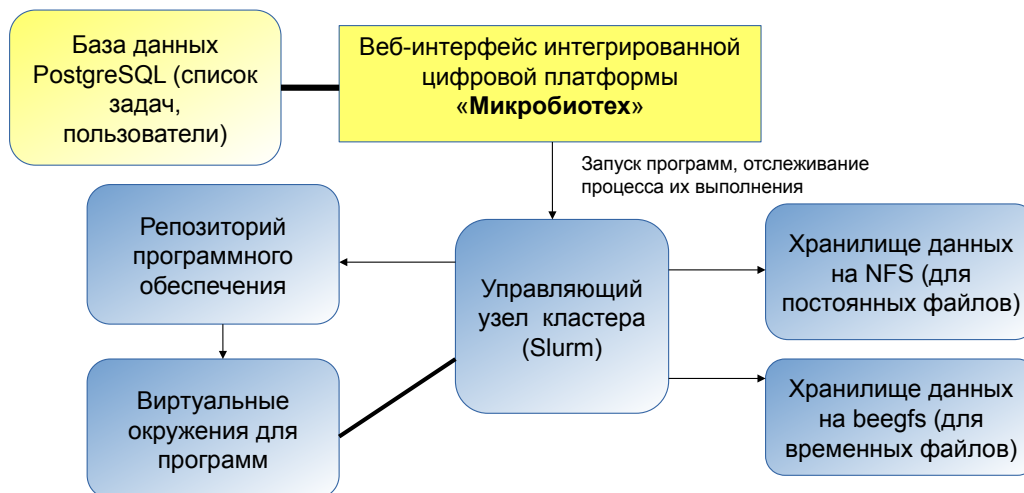


Рис. 2. Схема запуска конвейеров на кластере

**3. Основные программные модули.** Всем включаемым в ИЦП «Микробиотех» программным продуктам приписываются тематические разделы системной биологии, в которых данный программный продукт используется.

Тематический раздел **«Структурная и регуляторная компьютерная геномика»** включает в себя:

- вычислительный конвейер GBS-DP для обработки данных, полученных методом GBS (генотипирование путем секвенирования). Конвейер применяется для быстрого генотипирования образцов и идентификации генетической изменчивости (<https://ipcgncytogen.ru/modules/genomics/gbsdpr>) [13].

- вычислительный конвейер AntiNoise для решения задач поиска мотивов de novo в данных массового картирования ChIP-seq сайтов связывания транскрипционных факторов (<https://ipcgncytogen.ru/modules/genomics/antinoise>).

- вычислительный конвейер MultiDeNA для задач совместного применения различных моделей de novo поиска мотивов в данных ChIP-seq (<https://ipcgncytogen.ru/modules/genomics/multidena>).

В тематическом разделе **«Компьютерная протеомика и компьютерная фармакология»** представлены:

- программа FastProtMutStab для оценки влияния аминокислотных замен, ассоциированных с патологиями человека, на структуру и функцию белка (<https://ipcgncytogen.ru/modules/peptomics/fastprotmutstab>).

- программный конвейер AIPeptideDesigner для дизайна пептидов, имитирующих сайты связывания в белках, с применением методов искусственного интеллекта (<https://ipcgncytogen.ru/modules/peptomics/aipeptidedesigner>).

- программный конвейер предсказания функциональных сайтов в пространственной структуре белков. Раздел «Системная компьютерная биология: реконструкция, компьютерный анализ и моделирование генных сетей микроорганизмов» включает в себя:

- базу знаний ANDSystemB, построенную на технологии ANDSystem [14, 15], включающую информацию о молекулярно-генетических объектах (белки, гены, метаболиты и т. д.) микроорганизмов и различных типах взаимосвязях между ними;
- программный модуль для свернутого представления и визуализации результатов массовых вычислительных экспериментов, проводимых при моделировании биологических систем (<https://ipcgn.cytogen.ru/modules/sysbio/tsoced>).

## Список литературы

1. Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N. Digital business strategy: toward a next generation of insights // *MIS quarterly*, 2013. P. 471–482.
2. Yoo, Y., Henfridsson, O., & Lyytinen, K. Research commentary — the new organizing logic of digital innovation: an agenda for information systems research // *Information systems research*, 2010. N 21(4). P. 724–735.
3. Goecks, J., Nekrutenko, A., Taylor, J., & Galaxy Team. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences // *Genome biology*, 2010. N 11(8), R86.
4. Gruning, B. A., Rasche, E., Rebolledo-Jaramillo, B., Eberhard, C., Houwaart, T., Chilton, J., ... & Backofen, R. Jupyter and Galaxy: Easing entry barriers into complex data analyses for biomedical researchers // *PLoS computational biology*, 2017. N 13(5), e1005425.
5. Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Cech, M., ... & Goecks, J. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update // *Nucleic acids research*, 2016. N 44(W1), W3–W10.
6. Lowe, R., Shirley, N., Bleackley, M., Dolan, S., & Shafee, T. Transcriptomics technologies // *PLoS computational biology*, 2017.
7. Gruning, B., Chilton, J., Koster, J., Dale, R., Soranzo, N., van den Beek, M., ... & Backofen, R. Practical computational reproducibility in the life sciences // *Cell systems*, 2019. N 8(3). P. 183–188.
8. Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Cech, M., ... & Blankenberg, D. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update // *Nucleic acids research*, 2018. N 46(W1), W537–W544.
9. Kluyver, T., Ragan-Kelley, B., Perez, F., Granger, B. E., Bussonnier, M., Frederic, J., ... & Ivanov, P. Jupyter Notebooks — a publishing format for reproducible computational workflows // *In ELPUB*, 2016. P. 87–90.
10. Zaharia, M., Xin, R. S., Wendell, P., Das, T., Armbrust, M., Dave, A., ... & Ghodsi, A. Apache Spark: a unified engine for big data processing // *Communications of the ACM*, 2016. N 59(11). P. 56–65.
11. Varia, J., & Mathew, S. Overview of Amazon Web Services. Amazon Web Services, 2014. N 16.
12. Chee, B. J., Franklin, J. C., & Chee, B. J. Cloud computing: Technologies and strategies of the ubiquitous data center // CRC Press, 2009.
13. Pronozin A. Y., Salina E. A., Afonnikov D. A. GBS-DP: a bioinformatics pipeline for processing data coming from genotyping by sequencing // *Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding*. 2023; N 27(7). P. 737–745. DOI 10.18699/VJGB-23-86
14. Ivanisenko V. A., Saik O. V., Ivanisenko N. V., Tiys E. S., Ivanisenko T. V., Demenkov P. S., Kolchanov N. A. ANDSystem: an associative network discovery system for automated literature mining in the field of biology // *BMC Syst Biol*. 2015;9 Suppl 2(Suppl 2):S2. DOI: 10.1186/1752-0509-9-S2-S2.
15. Ivanisenko V. A., Demenkov P. S., Ivanisenko T. V., Mishchenko E. L., Saik O. V. A new version of the ANDSystem tool for automatic extraction of knowledge from scientific publications with expanded functionality for reconstruction of associative gene networks by considering tissue-specific gene expression // *BMC Bioinformatics*. 2019. N 20(Suppl 1). P. 34. DOI: 10.1186/s12859-018-2567-6.



**Деменков Павел Сергеевич** — канд. техн. наук, научный сотрудник Института цитологии и генетики СО РАН. Окончил НГУ в 2005 году по специальности «Прикладная математика и информатика». Защитил кандидатскую диссертацию в 2008 году. Области научных интересов: биоинформатика, генные сети, системная биология, большие геномные данные, ИИ в биологии, text-mining. E-mail: demps@bionet.nsc.ru.

**Pavel Sergeevich Demenkov** — PhD in Computer Science, researcher at the Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences. Areas of scientific interest: bioinformatics, gene networks, systems biology, big genomic data, AI in biology text mining. E-mail: demps@bionet.nsc.ru.



**Мухин Алексей Максимович** — младш. науч. сотрудник Института цитологии и генетики СО РАН. Окончил ФИТ НГУ в 2019 году, в 2023 — аспирантуру ИЦиГ СО РАН. С 2017 г. сотрудник ИЦиГ СО РАН. В сфере его научных интересов — программное обеспечение в области биологии. E-mail: mukhin@bionet.nsc.ru.

**Mukhin Aleksey Maksimovich** — graduated from Faculty of Information Technology of the Novosibirsk State University in 2019, in 2023 — postgraduate program of ICG SB RAS. Since 2017 he has been an employee of ICG SB RAS. His research interests include software in the field of biology.



**Иванисенко Владимир Александрович** — кандидат биологических наук, доцент, заведующий лабораторией компьютерной протеомики и лабораторией искусственного интеллекта и больших геномных данных ИЦиГ СО РАН. Области научных интересов: биоинформатика, генные сети, структурная биология, системная биология, большие геномные данные, ИИ в биологии, text-mining. E-mail: salix@bionet.nsc.ru.

**Vladimir Alexandrovich Ivanisenko** — PhD in Biology, associate professor, the head of the Laboratory of Computational Proteomics and the Laboratory of Artificial Intelligence and Big Genomic Data at the Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences. Areas of scientific interest: bioinformatics, gene networks, structural biology, systems biology, big genomic data, AI in biology text mining. E-mail: salix@bionet.nsc.ru.



**Лашин Сергей Александрович** — канд. биол. наук, ведущий научный сотрудник Института цитологии и генетики СО РАН. Области научных интересов: биоинформатика, математическое и компьютерное моделирование биологических систем, разработка программ и баз данных. E-mail: lashin@bionet.nsc.ru.

**Lashin Sergey Aleksandrovich** graduated in 2003 from the Faculty of Math of the Novosibirsk State University, PhD in Biology, specialist in mathematical and computer modeling of biological systems — molecular-genetic, population-genetic, ecological, development of bioinformatic methods, software and databases.



**Колчанов Николай Александрович** — доктор биологических наук, профессор, академик РАН, научный руководитель Федерального исследовательского центра «Институт цитологии и генетики СО РАН», заведующий отделом системной биологии Института, заведующий кафедрой информационной биологии ФЕН НГУ. Специалист в области биоинформатики и системной компьютерной биологии. В 1971 году окончил НГУ по специальности «Биология». Поступил на работу в ИЦиГ СО АН СССР (ИЦиГ СО РАН), где прошел путь от аспиранта до директора ФИЦ ИЦиГ СО РАН. В н/вр. — научный руководитель Института и заведующий отделом системной биологии. В 1975 году защитил кандидатскую, в 1989 — докторскую диссертации. Научные интересы: анализ больших генетических данных, реконструкция, компьютерный анализ

и моделирование структурно-функциональной организации и эволюции генных сетей, геномов, генов, РНК, белков человека, животных, растений и микроорганизмов на основе методов биоинформатики и системной компьютерной биологии. E-mail: kol@bionet.nsc.ru.

**Kolchanov Nikolay Aleksandrovich** — Doctor of Biological Sciences, Professor, Academician of the Russian Academy of Sciences, Scientific Director of the Federal Research Center Institute of Cytology and Genetics SB RAS, Head of the Department of Systems Biology of the Institute, Head of the Department of Information Biology of the NSU Natural Sciences Department. Specialist in bioinformatics and systems computer biology. In 1971 he graduated from NSU with

a degree in biology. He joined the Institute of Cytology and Genetics SB RAS, where he rose from a graduate student to the director of the FRC ICG SB RAS. At present — Scientific Director of the Institute and Head of the Department of Systems Biology. In 1975 he defended his candidate's dissertation, in 1989 — his doctoral dissertation. Research interests: analysis of large genetic data, reconstruction, computer analysis and modeling of the structural and functional organization and evolution of gene networks, genomes, genes, RNA, proteins of humans, animals, plants and microorganisms based on the methods of bioinformatics and systems computer biology. E-mail: kol@bionet.nsc.ru.

*Дата поступления* — 01.11.2024