

## **О РАБОТЕ НАУЧНОГО СОВЕТА СО РАН ПО БИОИНФОРМАТИКЕ**

*Научно-образовательная деятельность в области биоинформатики и системной компьютерной биологии.* В рамках подготовки профильных специалистов по этим наукам на ММФ НГУ разработана новая магистерская программа «Алгоритмы анализа больших биологических данных» ММФ НГУ (<https://mca.nsu.ru/aabbd/>). Программа включает обязательные дисциплины из учебного плана направления «Прикладная математика и информатика» механико-математического направления НГУ, а также специальные дисциплины, посвященные изучению задач системной биологии и биоинформатики. Специальные курсы разработаны при тесном сотрудничестве с кафедрой информационной биологии ФЕН НГУ, а некоторые из них преподаются для совместных групп математиков и биологов. Это позволяет извлечь практический опыт взаимодействия математиков и биологов при решении междисциплинарных задач биоинформатики.

*Организация конференций и школ молодых ученых.* С 4 по 9 октября 2021 года ФИЦ ИЦиГ СО РАН провел Мульти-школу молодых ученых, объединившую 13-ую Школу «Системная биология и Биоинформатика» SBB-2021 («Systems biology and Bioinformatics») и Школу молодых ученых «Генетика, геномика, биоинформатика и биотехнология растений» PlantGen School 2021 («Plant genetics, genomics, bioinformatics and biotechnology»). На Школе были рассмотрены различные информационные технологии в генетике: компьютерная статистическая геномика, 3D структура геномов, математические модели геномной селекции, компьютерная транскриптомика, компьютерный анализ и моделирование пространственных структур белков, анализ генетических последовательностей, компьютерное моделирование метаболических путей бактерий, интеграция транскриптомных и

метаболомных данных как метод селекции растений, математика и системная биология COVID-19.

***Вычислительная инфраструктура на базе ЦКП»Биоинформатика» ФИЦ ИЦиГ СО РАН.*** За 2021 год выполнено 216278 заданий по обработке генетических данных и моделированию молекулярно-генетических систем и процессов. Количество активных пользователей кластера превысило 100 человек. Прирост данных на системах хранения составил порядка 150 Тб. На базе инфраструктуры ЦКП проведено несколько конференций и научных школ. На базе научно-образовательного комплекса ФИЦ ИЦиГ СО РАН «Биодема» создан современный компьютерный класс и развернута беспроводная сеть WiFi-6. Распределенная система виртуализации на базе Xen-ng обеспечивает работу более 100 виртуальных серверов, обеспечивающих выполнение задач по предоставлению различного рода услуг для внешних пользователей и для служб ФИЦ ИЦиГ. В рамках пилотного проекта запущено корпоративное облако NextCloud с интегрированной системой совместной работы над документами. В сотрудничестве с компанией Huawei, ведется работа по адаптации биоинформатического программного обеспечения к архитектуре ARM. В дополнение к полному стеку биоинформатического ПО, предоставляемому репозиториями bioconda и conda-forge, были закуплены лицензии на ПО для визуализации филогенетических деревьев iTOL и система для функционального анализа омиксных данных OmixVox от компании BioVam.

***Экспертно-аналитическое заключение.*** Биоинформатика и системная компьютерная биология относятся к числу важнейших сквозных генетических технологий генерации новых знаний в области наук о жизни, т.к. они обеспечивают: (1) хранение, обработку и анализ различных больших данных, получаемых с помощью методов геномики, транскриптомики, протеомики, метаболомики и других высокопроизводительных технологий; (2) планирование экспериментов на основе интеграции и интерпретацию больших биологических данных и

методов математического и компьютерного моделирования. Биоинформатика и системная компьютерная биология имеют важнейшее значение для решения задач сельского хозяйства, биомедицины, фармакологии, биотехнологий, экологии, а также для решения междисциплинарных задач, которые всё более интенсивно выполняются при взаимодействии наук о жизни с математикой, информационными технологиями, физикой, химией и гуманитарными науками. Для эффективного решения задач биоинформатики и системной компьютерной биологии лимитирующим фактором является недостаточная мощность вычислительных систем Новосибирского научного центра. В связи с этим исключительно важное значение имеет создание СКЦ «Лаврентьев», который планируется построить в рамках программы «Академгородок 2.0».