



Н.А. Колчанов



Т.А. Бухарина

Уважаемые коллеги!
Предлагаем вашему вниманию очередной выпуск «Вавиловского журнала генетики и селекции», посвященный вопросам компьютерной биологии.

В последние 20 лет происходило стремительное развитие методов секвенирования геномов. Стоимость их расшифровки упала почти на пять порядков (для генома человека, например, со 100 000 до 500 долларов). Стремительно развивались транскриптомика, протеомика, метаболомика и другие омиксные технологии, возникло новое поколение методов визуализации биологических объектов на геномном, клеточном, тканевом, организменном уровнях организации живых систем. В результате в генетике произошел информационный взрыв: она стала главным источником больших данных, обогнав по темпам накопления информации не только другие науки, но и мировые социальные сети. Действительно, сейчас за год в науках о жизни продуцируется до 40 экабайт геномных данных, тогда как крупнейшая социальная платформа YouTube дает только 2 экабайта данных, т. е. в 20 раз меньше.

Анализ больших генетических данных привел к формированию новой парадигмы современной генетики. Ее

центральным объектом становятся генные сети – группы координированно функционирующих генов, которые взаимодействуют друг с другом через продукты – РНК, белки, метаболиты, обеспечивая на основе информации, закодированной в геномах, формирование фенотипических характеристик организмов (молекулярных, биохимических, клеточных, физиологических, морфологических, поведенческих и т. д.). Была выявлена огромная сложность регуляции генных сетей. Проявляется она в том, что функционирование любого отдельного элемента генной сети может контролироваться десятками и сотнями элементарных регуляторных процессов. Это относится как к регуляции транскрипции генов, которую контролируют десятки транскрипционных факторов, взаимодействующих с сайтами связывания в промоторах генов, функционированию белков, активность которых модулируется взаимодействием с большим количеством лигандов – аллостерических регуляторов, так и к метаболическим путям, в которых количество элементарных управляющих процессов может быть на порядок больше, чем число биохимических реакций. Еще одна фундаментальная особенность живых систем, выявленная при анализе больших генетических данных, – исключительно высокий уровень геномной изменчивости в популяциях человека, животных, растений и микроорганизмов.

Для анализа больших генетических данных необходима разработка нового поколения методов, способных работать со сверхбольшими объемами информации. Это новые методы биоинформатики для реконструкции, анализа и моделирования структурной организации и молекулярно-генетических механизмов функционирования геномов, генов и кодируемых ими генетических макромолекул – РНК и белков. Это также новые методы системной компьютерной биологии для реконструкции, анализа и

моделирования генных сетей и молекулярно-генетических систем, функционирующих на уровне клеток, тканей, органов и целостных организмов.

Новая эра больших генетических данных, в которую входят науки о жизни, требует трансформации базовых подходов биоинформатики и системной компьютерной биологии. Каковы же здесь базовые тренды? Прежде всего, это интеграция классических методов биоинформатики и системной компьютерной биологии с методами искусственного интеллекта и глубокого машинного обучения. Разработка на этой основе нового поколения гибридных информационно-программных систем для ин-

терпретации больших генетических данных, а также, что особенно важно, для планирования экспериментов по проверке результатов компьютерных предсказаний, полученных из анализа больших генетических данных. Движение в этом направлении будет означать принципиальную трансформацию базовой парадигмы современных исследований: дополнение науки, направляемой гипотезами, новой наукой, направляемой результатами анализа больших данных.

И такое движение идет во всех науках, но именно биоинформатика и системная компьютерная биология находятся на его острие.

Научные редакторы выпуска:
академик Н.А. Колчанов,
научный руководитель ФИЦ ИЦиГ СО РАН
Т.А. Бухарина, ученый секретарь
отдела системной биологии ФИЦ ИЦиГ СО РАН