

БИОИНФОРМАТИКА РЕГУЛЯЦИИ И СТРУКТУРЫ ГЕНОМА

16-22 июля Институт цитологии и генетики СО РАН провел в Доме ученых Сибирского отделения 5-ю Международную конференцию «Биоинформатика регуляции и структуры генома — 2006» (Bioinformatics of genome regulation and structure — BGRS-2006).

Людмила Юдина, «НВС»

Минуло два года. Снова лето. Снова июль. И опять в новосибирский Академгородок устремились специалисты из научных центров, университетов, исследовательских институтов многих городов России и разных стран мира, чтобы познакомиться с наиболее интересными достижениями в области компьютерной биологии, обсудить перспективы развития новых методов компьютерного и теоретического анализа молекулярно-генетических систем и процессов, обменяться опытом их приложения к решению фундаментальных и прикладных задач геномики, эволюционной и системной биологии, биомедицины и биотехнологии, а также других областей науки, связанных с исследованиями закономерностей структурно-функциональной организации и эволюции молекулярно-генетических систем человека, животных, растений и микроорганизмов.



Директор Института цитологии и генетики академик Владимир Шумный, участник всех состоявшихся ранее конференций по биоинформатике, в своем вступительном слове еще раз подчеркнул огромное значение подобных форумов, выводящих в конечном счете на решение многих сложных и важных для человечества проблем. Тепло поприветствовав гостей, он не преминул отметить, что некоторые из «иностранцев», прибывших на конференцию, начинали научную карьеру в Институте цитологии и генетики СО РАН. В. Шумный пожелал участникам конференции успехов в предстоящей сложной и напряженной работе.

Работа, действительно, была напряженной — все семь дней. Пленарные доклады замечательных специалистов неизменно поддерживали высокий интерес слушателей. Не снижался накал дискуссий на секциях — компьютерная структурная и функциональная геномика; компьютерная структурная и функциональная протеомика; компьютерная эволюционная биология; компьютерная системная биология; новые подходы к анализу биомолекулярных данных и моделированию процессов.

Комплекс проблем, рассмотренных в ходе международной конференции, включал самые острые современные темы — структурные и функциональные характеристики ДНК, РНК и белков; регуляцию процессов считывания информации с ДНК; моделирование и предсказание структуры и функции белков; молекулярная динамика биологических макромолекул; эволюция геномов и белков; моделирование и анализ генных сетей и метаболических путей; моделирование электронной (виртуальной) живой клетки, разработка информационно-

компьютерных технологий для хранения, аннотирования систематизации и анализа колоссальных объемов экспериментальных данных, получаемых в результате крупномасштабных молекулярно-биологических и молекулярно-генетических экспериментов и многие другие вопросы, подпадающие под определение «современное состояние биоинформатики и перспективы дальнейших исследований». Устные сообщения сменялись стендовыми сессиями, компьютерными демонстрациями и демонстрациями программного обеспечения.

Как подтверждение тезиса о планомерной смене поколений математических биологов — прошедшая перед 5-й конференцией по биоинформатике, под ее эгидой, международная школа молодых ученых «Эволюционная биология и высокопроизводительные вычисления в биоинформатике».



Председатель Международного программного комитета конференции член-корреспондент РАН, зам. директора ИЦиГ СО РАН **Николай КОЛЧАНОВ** рассказал об истории научного форума, который в течение последних лет стал заметным событием в международной и российской биоинформатике, о его тематике, традициях и планах на будущее:

— Первую конференцию по компьютерной биологии мы провели в Новосибирске в 1984 году совместно с выдающимся молекулярным биологом-кибернетиком, создателем школы математической биологии и биоинформатики ИЦиГ СО РАН, профессором Вадимом Александровичем Ратнером. Одним из своих учителей В. Ратнер считал член-корреспондента Академии наук СССР Алексея Андреевича Ляпунова, которого называют отцом отечественной кибернетики и одним из основоположников кибернетики вообще

(А.А.Ляпунов — один из трех советских граждан, награжденных Computer Society Золотой медалью «Computer Pioneer»). Интерес Алексея Андреевича к биологии зародился с 30 гг. XX в, когда он по инициативе А. Н. Колмогорова проводил статистические исследования расщепления признаков у гибридов. Позднее, в пятидесятых годах, уже будучи признанным математиком с мировым именем, А. Ляпунов вернулся в биологию. Сложность биологических систем и процессов их эволюции привлекала его как объект приложения методов исследования, характерных для дескриптивной теории множеств. Его целостный взгляд на организм хорошо совпадал с воззрениями великих русских эволюционистов-морфологов И. И. Шмальгаузена и А.Н. Северцова, но слабо сочетался с только что законченной тогда на Западе синтетической теорией эволюции (СТЭ), построенной, прежде всего, на достижениях популяционной генетики. Нельзя сказать, что кибернетика и СТЭ противоречили друг другу. Они просто говорили на разных языках. В экспериментальных исследованиях по доместикация животных академик Д. К. Беляев показал важнейшую роль регуляторных систем в эволюции. Поэтому ему импонировали кибернетические представления В. А. Ратнера, сформулировавшего представление о МГСУ — молекулярно-генетических системах управления. Таким образом, в 60-80 годы прошлого века в ИЦиГ СО РАН был проведен изящный синтез кибернетики, математики, популяционной и экспериментальной генетики.

Я был в числе выпускников факультета естественных наук НГУ 1971 года по специальности «Математическая биология», выполняя дипломную работу под руководством чл.-корр. РАН А. А. Ляпунова и проф. М. Г. Колпакова, а после аспирантуры — работал в ИЦиГ СО РАН под руководством проф. В. А. Ратнера, где на моих глазах созданное им направление превращалось в мощную научную школу, характерной чертой которой был широкий диалог между специалистами — математиками с одной стороны и биологами-экспериментаторами с другой.

При организации конференций мы продолжаем эти традиции с учетом прогресса как экспериментальных, так и информационных технологий, методов моделирования и анализа данных. Задачу нынешней конференции, как и предыдущих, Оргкомитет видит, прежде всего, в создании пространства для широкого обмена мнениями между теоретиками, а также поиска областей сотрудничества с биологами-экспериментаторами, как применяющими в своей работе компьютерные методы, так и интересующимися прикладными или теоретическими аспектами биологии.

Но вернемся к 1984 году. Именно тогда стало ясно, что в нашей стране накопилась достаточная «критическая масса» специалистов в области математической биологии, которым не только полезно, но и необходимо регулярно общаться на конференциях для того, чтобы более плодотворно решать возникающие проблемы. Именно в это время, собственно говоря, и началась биоинформатика как наука, никогда не оставлявшая, и это я хотел бы особенно подчеркнуть, союза с математикой — источником многих замечательных идей, эффективных алгоритмов и методических подходов.

В 1986-м и 1988-м годах было организовано еще две конференции по математической биологии — в статусе общесоюзных. Ещё через два года в 1990 году наш форум стал международным. Потом — перерыв по причинам, всем известным, связанным с распадом великой державы. И затем — с 1998 г., когда мы вновь стали накапливать силы, снова начались международные конференции (с интервалом в два года). Так что, по сути, это уже девятая конференция по биоинформатике. И если подсчитать общее количество участников этих конференций, оно заведомо переваливает за две с половиной тысячи.

— Какие обстоятельства стимулируют развитие биоинформатики?

— Прежде всего появление качественно новых экспериментальных технологий: геномики, транскриптомики, протеомики, метаболомики, клеточной биологии, которые позволяют с высокой эффективностью исследовать молекулярно-генетические системы и процессы. Рутинными становятся такие задачи, как полная расшифровка геномов, пространственных структур белков, изучение метаболических профилей организмов, исследование и картирование полиморфизмов (генетических вариаций геномной ДНК в популяциях). В результате в последние десять лет в молекулярной биологии и генетике произошел информационный взрыв, сопровождаемый экспоненциальным ростом объемов экспериментальных данных. Даже просто прочесть все статьи, содержащие эту информацию, не говоря уже об ее осмыслении, физически невозможно. Поэтому создание высокопроизводительных информационно-компьютерных технологий, предназначенных для анализа и интерпретации экспериментальных данных, стало в настоящее время необходимым для решения фундаментальных задач биологии, а тем более — для практического использования в биомедицине и биотехнологии.

Однако, это только первый срез проблемы. Молекулярная биология и генетика, как экспериментальные науки, до последнего времени носили ярко выраженный аналитический характер, концентрируя свое внимание на изучении отдельных аспектов организации сложнейших по своей сущности биологических систем. Развитие в последнее десятилетие высокоэффективных экспериментальных технологий молекулярной биологии и генетики привело к накоплению огромных объемов первичных данных, касающихся различных уровней организации жизни (от геномов и кодируемых ими биологически макромолекул через молекулярно-генетические системы клеток к тканям, органам, организмам, вплоть до популяций, биологических видов и экосистем). Эти данные слабо связаны, плохо структурированы, имеют разную степень полноты и сами по себе не позволяют реконструировать полноценный портрет любой изучаемой биологической системы или процесса. Именно поэтому в настоящее время в биоинформатике на первое место выходят

задачи синтеза — интеграции и систематизации первичных экспериментальных данных и задачи продукции знаний на основе современных информационных и математических технологий.

— Конференции можно назвать многокомпонентными, или на каждой из них доминирует определенная тема?

— Хотя биоинформатике как науке нет и 20 лет, в ней, как и в любой сложившейся науке, существуют свои традиционные направления, такие как компьютерный анализ ДНК, РНК и белковых последовательностей, распознавание функциональных сайтов, реконструкция пространственных структур биополимеров, теоретический и компьютерный анализ структурно-функциональной организации геномов и белков. К традиционным направлениям следует отнести и такую важную область, как развитие баз данных по молекулярно-генетической тематике. Эффективное структурирование экспериментальных данных уже позволяет извлечь полезную информацию, обнаружить значимые корреляции, применить классические статистики. Такие традиционные направления всегда были представлены на наших конференциях.

Однако, традиционное — не значит рутинное, застывшее. Адекватный анализ разнородных экспериментальных биологических данных представляет собой вызов теории анализа данных и машинного обучения (Data Mining and Machine Learning). Эффективное структурирование данных уже позволяет извлечь полезную информацию, обнаружить значимые корреляции, применить классические статистики и даже, как в таблице Менделеева, заполнить пробелы, например, предсказав неизученные характеристики функциональных районов по их окружению.

В Институте математики им. С. Л. Соболева под руководством Е. Витяева разрабатываются методы автоматической продукции знаний из данных. Подход, лежащий в основе названных методов, успешно применялся для решения многих задач в медицине, финансах (см. [website http://www.math.nsc.ru/AP/ScientificDiscovery](http://www.math.nsc.ru/AP/ScientificDiscovery)). Сейчас в сотрудничестве с ИЦиГ СО РАН он используется для решения актуальных задач биоинформатики: распознавание и анализ регуляторных районов транскрипции генов эукариот. В рамках этого направления разработаны две программные системы: Discovery (обнаружения логических закономерностей) и NatClass (построения естественных классификаций). Полученные результаты были доложены на BGRS'2006 и вызвали большой интерес у зарубежных коллег. Достигнуты предварительные договоренности о сотрудничестве по применению программ.

Направление работ в современной биологии меняется очень быстро. Появляются новые методы, задачи, проблемы, что находит естественное отражение в тематике. Поэтому на каждой конференции помимо традиционных направлений рассматриваются наиболее «горячие» в настоящий момент области.

Сейчас такой областью становится системная биология, которая и предназначена для осуществления нового синтеза в биологии. Применительно к биоинформатике ее задача — построение математических моделей биологических систем, функционирующих на основе информации, закодированной в геномах и взаимодействующих со средой (клеток, тканей, организмов и их объединений). Системная биология — самая горячая точка текущей конференции. Как уже говорилось выше, до последнего времени биологические науки имели преимущественно аналитический характер — они рассекали живую систему и анализировали ее отдельные уровни (ген, белок, геном, клетка). Сейчас самое важное — объединять все эти уровни и строить реальные модели сложных систем.

Например, огромный интерес вызывает построение математических моделей бактериальных клеток, для которых в настоящее время накоплено огромное количество экспериментальных данных. На таких моделях можно изучать функционирование генов, геномов, регуляцию метаболизма, что представляет самостоятельный интерес, а также имеет большое значение для прикладной биотехнологии. Использование математического моделирования позволяет, в частности, проектировать создание бактериальных систем — суперпродуцентов биологически активных веществ.

Практически возможным стало использование методов биоинформатики для проектирования бактериальных генно-инженерных конструкций (мини-генных сетей), способных усиливать на много порядков рецептируемые бактериями из окружающей среды сигналы о наличии неблагоприятных (опасных) веществ. На этой основе возможно создание биосенсоров нового поколения, предназначенных для выявления в окружающей среде и продуктах питания вредных и опасных для человека веществ (например, мутагенов и канцерогенов).

Еще одна важная область системной компьютерной биологии микроорганизмов — построение детальных портретных моделей, описывающих взаимодействие инфекционных агентов с организмом человека и его защитными системами, в частности, иммунной системой.

Например, в лаб. теоретической генетики ИЦиГ СО РАН реконструирована сеть молекулярно-генетических взаимодействий, происходящих при размножении вируса гепатита С в клетке. На основе этой сети группы В. Иванисенко и В. Лихошвая из ИЦиГ совместно с профессором В. Антюфеевым (ИВМиМГ СО РАН) построили модели, отражающие взаимодействие основных процессов, происходящих при размножении вируса в клетке. Их назначение — изучить возможные сценарии развития инфекции в клетке. В связи с тем, что число вирусных частиц и их компонентов в клетке ограничено, наиболее адекватно описывать процесс репродукции вируса не только химико-кинетическими моделями, а также использовать стохастическое моделирование. Эти работы были доложены на конференции и вызвали большой интерес у участников.

Для ряда патогенов (например, туберкулезной палочки *Mycobacterium tuberculosis*) характерно очень быстрое возникновение устойчивости к антибиотикам. Практически возможным становится создание информационно-компьютерных подходов для решения задачи оптимального управления инфекционным процессом на уровне отдельного организма. При этом существенно, что мишенями оптимального управления могут быть не только гены (белки) бактерии, но и компоненты молекулярно-генетической машины человека. В ближайшем будущем это позволит создавать генотип-специфические стратегии лечения инфекционных заболеваний. Не менее актуальным является создание компьютерных методов прогнозирования развития эпидемий.

Одна из «горячих» проблем биоинформатики на сегодняшний день — компьютерная поддержка экспериментов по генотипированию. Эта задача перешла в разряд первоочередных в мировой науке и практической медицине в связи с тем, что в настоящее время стало возможным осуществление крупномасштабных исследований по массовому (до десятков тысяч человек) генотипированию населения с целью выявления генетических компонент (полиморфизмов, мутаций), обуславливающих предрасположенность к мультифакториальным заболеваниям.

Перед участниками таких проектов стоит несколько задач. Во-первых, необходимо оценить степень связи известных полиморфизмов в ряде функционально значимых генов с риском заболевания. Во-вторых — провести поиск новых генов, полиморфизмы в которых вносят свой вклад в предрасположенность к заболеванию.

В России также планируется осуществление крупных проектов по генотипированию. Как все другие проекты, связанные с исследованием больших массивов геномных данных, эти работы нуждаются в компьютерной поддержке и, прежде всего, со стороны специалистов-биоинформатиков. Две задачи, решения которых ждут от биоинформатики — планирование крупномасштабных экспериментов по генотипированию и интерпретация результатов экспериментов.

В настоящее время специалисты ряда лабораторий в ИЦиГ СО РАН (Т. Аксенович, А. Ромащенко, Т. Меркулова, Е. Игнатьева) разрабатывают ряд биоинформационных и экспериментальных подходов к решению этих проблем.

Институт цитологии и генетики работает над задачами компьютерной поддержки экспериментов по массовому генотипированию и интерпретации их результатов в тесном сотрудничестве как с российскими коллегами (Институт терапии СО РАМН, Новосибирск, чл.-корр. РАМН М. Воевода; ФГУ НИИ Физико-химической медицины Росздрава, Москва, проф. В. Говорун; Центр «Биоинженерия» РАН, Москва, проф. К. Скрябин, а также с зарубежными коллегами (Медицинский центр «Эразмус», Роттердам, Нидерланды, профессор Корнелия ван Дуин).

От прикладных задач перейдем к фундаментальным. Эволюционная проблематика в молекулярно-биологических исследованиях является традиционной. Вспышка интереса к ней в последнее время связана с огромным массивом накопленной информации по первичным последовательностям ДНК и белков, с одной стороны, и реконструкциями генных сетей, контролирующими процессы развития организмов, с другой. Еще 10 — 15 лет назад молекулярно-филогенетические исследования преимущественно охватывали сравнительно небольшую группу генов, прежде всего гены рибосомальной ДНК. Выполняя фундаментальные, но идентичные функции у всех организмов, эти гены очень консервативны. Таким образом, их эволюция идет в основном в нейтральном режиме, хорошо маркируя расхождение таксонов на эволюционном древе, но ничего не говоря об эволюции морфологии, физиологии и др. систем этих таксонов. Между тем, эта эволюция должна быть связана с эволюцией генов-регуляторов генных сетей развития. В результате программа развития организма меняется. Перед биоинформатикой стоит захватывающая перспектива разыскать следы мутаций в последовательностях генов регуляторов, затем, используя математические модели, реконструировать процессы изменения генных сетей в филогенезе и, наконец, сопоставить реконструированную картину эволюции с данными палеонтологии, эмбриологии, сравнительной анатомии и др. дисциплин.

В нашей лаборатории проведен анализ генов компонентов путей передачи сигналов, детерминирующих процессы морфогенеза (Hh, BMP, WNT-каскадов сигналов), с целью поиска участков филогенетических деревьев, характеризующихся движущим (адаптивным) режимом эволюции.

Формообразование в ходе роста живых организмов — морфогенез — процесс, в котором генная регуляция проявляется в морфотипе организма. В этом процессе структура и функция определяют друг друга так явно, как ни в каком другом. Крупные ароморфозы как правило связаны с изменениями программ морфогенеза. В процессе роста и морфогенеза, как и в процессе поддержания «стационарного» состояния формы организма, большую роль играют стволовые клетки, популяции которых поддерживаются в самых разных органах и тканях организма. Стволовые клетки плюрипотентны, то есть сохраняют способность дифференциации во все или почти все клетки данного органа. Это общий принцип всех сложных многоклеточных — животных и растений, но изучать стволовые клетки растений значительно проще. Изучение и моделирование механизмов поддержания устойчивой структуры апикальной

меристемы растений — предмет международного сотрудничества групп специалистов ИЦиГ СО РАН (Н. А. Омелянчук, С. В. Николаев, В. А. Лихошвай), ИМ СО РАН (С.И. Фадеев), ИСИ СО РАН (Д.К. Пономарев), Института геномики и биоинформатики (проф. Эрик Мьелнесс, Калифорнийский университет) и Калифорнийского технологического института (проф. Эллиотт Мейровиц).

Как уже сказано выше, стволовые клетки «поддерживаются» другими клетками, которые своими сигнальными молекулами «формируют нишу» для последних. Взаимное расположение стволовых клеток и клеток ниши по-видимому важно для поддержания стабильного пула стволовых клеток. Именно в силу важности указанной проблемы на конференции BGRS'2006 было представлено несколько наших совместных с зарубежными коллегами работ по моделированию регуляции структуры меристемы.

Кстати, раз уж мы заговорили о стволовых клетках, должен согласиться с крылатым высказыванием Больцмана, что «нет ничего практичнее хорошей теории». Исследование механизмов морфогенеза, а также их нарушений и эволюции имеет практическое значение для широкого круга дисциплин — от медицины до биотехнологии.

Биоинформатика, по своей сути — интегративная наука. Это находит отражение и в широком круге сотрудничества ИЦиГ СО РАН. Наш институт является одним из узлов биоинформационной сети Академгородка и Сибирского отделения РАН. В пределах этой сети ИЦиГ сейчас взаимодействует с ИХБиФМ СО РАН (структурная компьютерная биология и компьютерная физика биополимеров), ИМ СО РАН, ИВТ СО РАН, ИВМиМГ СО РАН, ИСИ СО РАН (интеграция баз данных по молекулярной биологии и генетике, методы моделирования молекулярно-генетических систем и процессов, суперкомпьютерные вычисления), ИК СО РАН (автоматический анализ текстов научных публикаций — text mining), ИТПМ СО РАН (теория управления), ИТФ СО РАН (физика биополимеров), ИЯФ СО РАН (геносенсоры и микрофлюидные системы), НГУ (образовательный процесс и решение широкого круга содержательных научных задач). Появляются новые партнеры. Особенно перспективным представляется расширение сотрудничества с ИСиЭЖ СО РАН, ЦСБС СО РАН, ИПА СО РАН для решения задач на стыке классической биологии, молекулярной биологии и генетики и биоинформатики.

— Как в настоящее время решается проблема подготовки специалистов в области биоинформатики.

— Институт имеет более чем 45-летнюю историю развития математической биологии и информатики. Сейчас в нем трудится уже третье поколение биоинформатиков, формируется четвертое поколение. Ситуация уникальная. Пожалуй, в России нет ни одного другого НИИ, где бы в течение четырех с половиной десятилетий было накоплено столько знаний, столько подходов и умения эти знания передавать — прежде всего студентам. Математическая биология стартовала у нас в 1968 году, затем превратилась в информационную биологию.

Следуя велению времени — необходимости подготовки высоко квалифицированных специалистов в области биоинформатики — в 2003 году в НГУ была создана Кафедра информационной биологии (КИБ), которую я возглавляю. Профессорско-преподавательский коллектив кафедры представлен сотрудниками ИЦиГ СО РАН и других институтов ННЦ, взаимодействующих с нами. К чтению лекций привлекаются также приглашенные профессора из научных и образовательных учреждений Москвы, Санкт-Петербурга, других научных центров. КИБ — выпускающая кафедра, на которой проходят специализацию студенты ФЕНа, решившие связать свою научную судьбу с биоинформатикой. На кафедре информационной биологии студенты получают углубленные знания как по специальным биологическим

дисциплинам, так и по специальным разделам математики, в том числе теории графов, методам решения обратных задач, теории классификации и распознавания образов, дискретного анализа и комбинаторики. Студенты кафедры изучают структурную и функциональную организацию ДНК, РНК и белков, принципы организации и механизмы функционирования геномов, генных сетей и метаболических путей, учатся формализованному описанию и моделированию молекулярно-генетических процессов и сложных биологических систем, интеграции и анализу огромных массивов данных, получаемых в результате использования мощных современных методов исследования живых систем. Первые выпускники кафедры получили дипломы в 2005 г. Большая часть продолжает обучение выбранной специальности в аспирантуре ИЦиГ и НГУ.

— Известно, что ИЦиГ СО РАН имеет обширную сеть сотрудничества с зарубежными коллегами в области биоинформатики. Получило ли это сотрудничество дополнительный импульс на конференции?

— Хорошим примером международного сотрудничества является прошедшее 21 июля в рамках конференции расширенное рабочее совещание участников российско-германской сети по биоинформатике. На этом совещании присутствовали проф. Р. Хафестафт (Билфилдский университет), проф. К. Лонгман (Кельнский университет) и проф. Й. Гроссе (Университет Галле). На совещании обсуждали вопросы, связанные с организацией совместных проектов в области образования (участвовали зав. кафедрой биоинформатики ФЕН НГУ проф. Н. Колчанов, декан ФИТ НГУ проф. М. Лаврентьев, декан факультета повышения квалификации НГУ проф. С. Загребельный); сообщение о недавно состоявшемся первом виртуальном образовательном симпозиуме по биоинформатике, организованном Азиатским обществом исследований и образования в области биоинформатики (ABREN) и объединившем около 600 студентов из стран Юго-Восточной Азии (в котором активное участие принял ИЦиГ СО РАН) сделано к.б.н. доц. А. Кочетовым. Центральной темой дискуссии были совместные проекты в области биоинформатики растений. Недавно директор центра «Биоинженерия» РАН проф. К. Скрыбин предложил исследователям, объединенным российско-германской сетью взаимодействия в области биоинформатики, принять участие в 7-й рамочной программе ЕС по биотехнологии растений «Plants for future». Применение методов и подходов биоинформатики и системной биологии может в значительной степени увеличить эффективность экспериментальных исследований. Эта инициатива нашла поддержку крупных ученых, отвечающих за организацию программы, и в настоящее время проф. Р. Хафестафт готовит доклад для программной комиссии. Поэтому совместные проекты в области биоинформатики растений были одной из основных тем, обсуждаемых на рабочем совещании.

Следует особо подчеркнуть, что в биоинформационном совещании приняли участие биологические экспериментаторы, специалисты в области генетики растений (проф. Л. Першина, д.б.н. Н. Гончаров, д.б.н. Е. Салина и др.) и эволюции геномов (к.б.н. А. Блинов). В частности, дискутировалась возможность участия экспериментальных подразделений ИЦиГ СО РАН в проведении совместных исследований со специалистами в области биоинформатики в рамках в ходе выполнения 7-й рамочной программы, а также организация совместных работ с Институтом генетики растений в г. Гатерслебен (проф. Й. Гроссе), центральным исследовательским институтом в этой области в Германии. Совместные дискуссии биологов-экспериментаторов и биоинформатиков, проходившие на рабочем совещании, наглядно иллюстрируют стратегическую тенденцию развития современной биологии — необходимость объединения экспериментаторов и теоретиков для решения крупных задач современной науки.

В результате работы расширенного рабочего совещания принято несколько конкретных решений, направленных на организацию процесса дистанционного образования в области

биоинформатики между НГУ и несколькими университетами в Германии, а также посвященных разработке проектов в области биоинформатики растений.

— Напрашивается вывод, что возможности биоинформатики просто неограничены?

— У каждой науки — свои пределы. В том числе — у информатики и информационной биологии. Сейчас становится ясным, что имеется ряд аспектов организации живого, которые мы раньше не то чтобы не понимали, а серьезно недооценивали, не могли найти адекватного математического аппарата. Я имею в виду, прежде всего, проблему сложности биологической организации. Представьте только: в одном организме в течение жизни координированно функционируют триллионы клеток, каждая из них имеет десятки тысяч генов и содержит огромное количество РНК, белков, метаболитов. при этом многоклеточный организм характеризуется удивительной устойчивостью по отношению к внешним факторам среды и внутренним флуктуациям. За этим стоят какие-то новые принципы биологической организации, которых мы до сих пор не знаем. Каким образом из простых элементов, причем, не однородных, но иерархически организованных, каждый из которых имеет свою собственную программу действий, свои собственные риски (слабые звенья), собственные сбои, приводящие порой к патологиям, как из такой гетерогенности возникает консолидированная система с достаточно предсказуемым, по крайней мере, на уровне организма, фенотипом? Если представить, что каждый ген, в каждой клетке, в каждом геноме может быть мишенью неблагоприятных воздействий, и при этом все риски минимизируются, флуктуации и внешние воздействия каким-то образом нейтрализуются, то очевидно, что здесь присутствуют какие-то неизвестные принципы, неизученные механизмы, которые, отработываясь в течение миллиардов лет эволюции, создали стройную, гармоничную систему, обеспечивающую целостность структурной организации и поведения при огромном разнообразии элементов.

Следует заметить, что мы не имеем примеров подобной гармоничной организации для социальных систем. Изучая биологические системы, можно, по-видимому, понять, чего же нам не хватает в социальной организации и что можно использовать из опыта природы. Конечно, нельзя сводить социальную организацию к биологической, ибо у каждого человека есть свобода выбора, свобода воли. Тем не менее, опыт трех с половиной миллиардов лет эволюции, накопленный в геномах, чего-то стоит. Его надо изучать. И биоинформатика, изучая живые системы, как мне представляется, сможет внести свой достойный вклад в изучение принципов организации сложных систем, возникающих в результате эволюции.

— Помнится, на прошлой конференции один из выступающих заметил, что биоинформатика не заменяет классических подходов биологии и способов, с помощью которых биологи осмысливают действительность. Вы согласны с этим утверждением?

— Конечно же, не заменяет! Но обогащает и предоставляет ученым качественно новые возможности! Биоинформатика предоставляет исследователям свои средства, углубляющие процесс познания — информационные технологии описания, хранения, анализа, визуализации, классификации первичных экспериментальных данных по структурно-функциональной организации геномов, генов, РНК, белков, генных сетей, метаболических путей, путей передачи сигналов, процессам развития, молекулярной эволюции и т.п., извлечения из этих данных закономерностей и знаний; принципиально важно, что в настоящее время она предоставляет также высокоэффективные методы синтеза целостной картины изучаемых биологических систем, объектов и процессов на основе интеграции данных и знаний и методов математического моделирования.

Например, выше уже говорилось о проблеме обвального роста публикаций. В нашем институте разработан информационно-программный комплекс, позволяющий осуществлять:

- 1) построение ассоциативных сетей на основе машинного анализа текстов и БД;
- 2) реконструкцию генных сетей физиологических процессов в норме и патологии;
- 3) анализ возможных эффектов полиморфизмов на транскрипционном уровне регуляции экспрессии путем выявления потенциальных сайтов связывания транскрипционных факторов в районе локализации полиморфизма;
- 4) сравнение трансляционной активности полиморфных вариантов 5'-нетранслируемых последовательностей путем оценки стабильности вторичной структуры 5'-НТП мРНК;
- 5) анализ возможных эффектов полиморфизмов на функции белка, который включает оценку термодинамической стабильности белков, функциональной активности мутантных белков, структуры функциональных сайтов в белках;
- 6) накопление данных о мутациях и полиморфизмах, значимых для исследования предрасположенности к мультифакториальным, наследственным и инфекционным заболеваниям.

Таким образом, биоинформационные технологии позволяют исследователю быстро сориентироваться в море публикаций, выявить наиболее перспективные для него темы, а также обратить внимание на те или иные смежные области исследований. Однако здесь возможности биоинформатики кончаются — связи, выявленные на основе анализа журнальных публикаций, исследователь должен проинтерпретировать сам, а также проверить экспериментально.

Фактически, получается так, что в настоящее время биоинформатика становится неотъемлемым атрибутом исследований в различных науках о жизни, позволяя ученым-биологам идти все дальше и дальше, в неведомые миры, поднимая на поверхность пласты скрытой информации. Биоинформатика приобретает сейчас особую ценность именно потому, что становится эффективным инструментом анализа огромных потоков экспериментальных данных, получаемых исследователями, и синтеза получаемых в результате этого знаний.

Важнейшей миссией биоинформатики становится создание экспериментально-компьютерных технологий анализа биологических систем и процессов, обеспечивающих планирование эксперимента и строгую количественную интерпретацию его результатов. В этом смысле биология повторяет тот путь, по которому более ста лет назад пошла физика, в недрах которой сформировались теоретическая, математическая и вычислительная физика, без которых в настоящее время невозможно представить созвездие физических наук, получивших выдающиеся результаты, изменившие во многом окружающий нас мир и представление о нем.

Негласно уровень любой из конференций определяется участием в ней именитых отечественных и зарубежных участников. Продолжая мысль, можно заключить, что международная конференция по биоинформатике регуляции и структуры генома прошла на высоком уровне — знаменитых гостей было достаточно. Сопредседателем программного комитета и на этот раз выступал профессор Ральф Хофстадт из Швейцарии.

Участники любой из конференций, покидая Сибирь, всегда подчеркивают, что любят приезжать сюда на научные мероприятия. Были они единомышленны и на этот раз.

Фото Владимира Новикова и Игоря Глотова