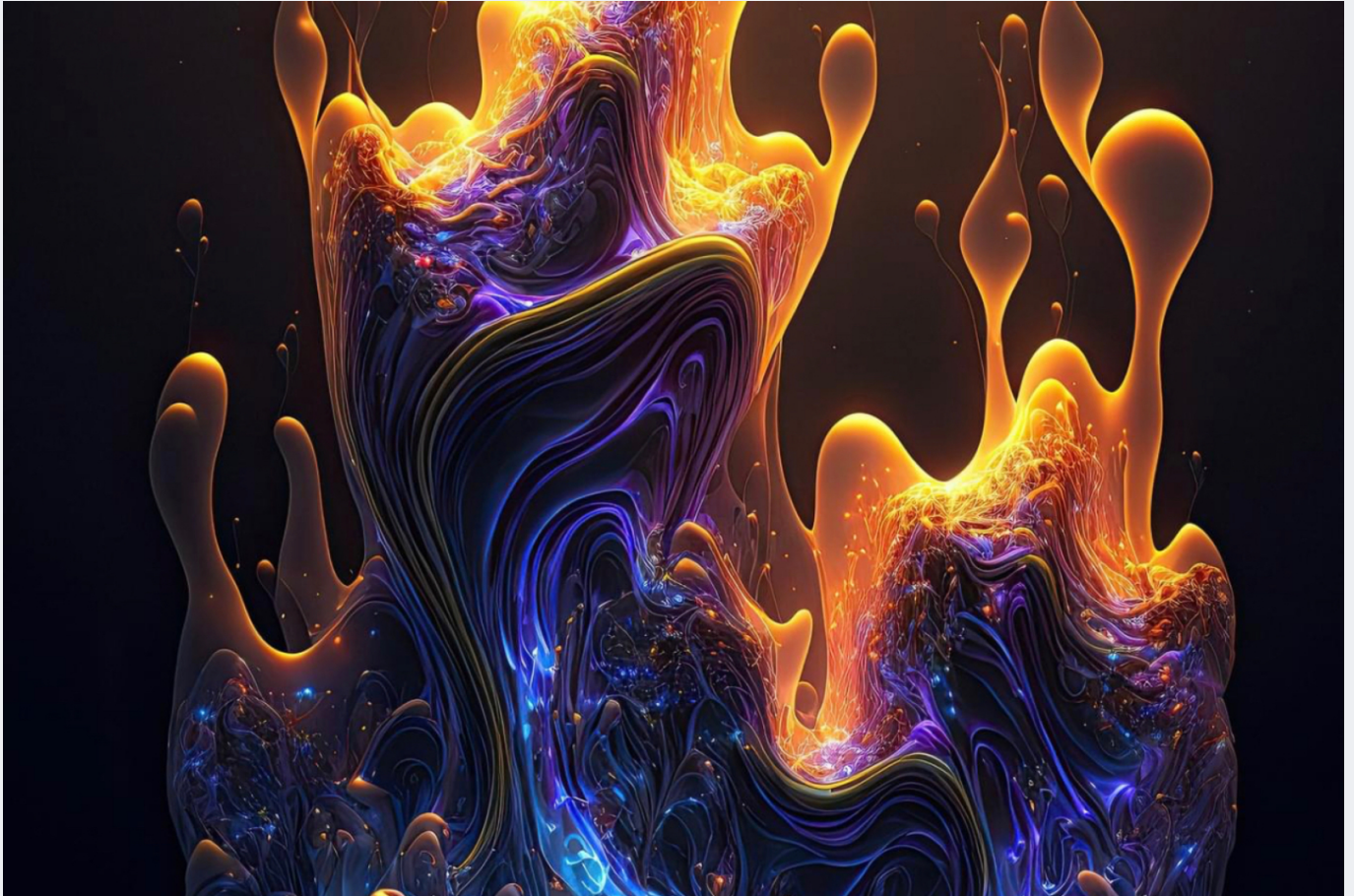


«Центр вычислительной биологии»: информатика живого



Информационные технологии перестали быть уделом программистов и математиков, что привело к расширению круга биологических задач, решаемых с помощью компьютеров. Масштабные проекты в области фундаментальных исследований и для создания инфраструктурных решений для обработки, анализа и хранения омиксных данных в биомедицине и других областях реализуются в научно-исследовательской лаборатории «Центр вычислительной биологии».

В современной биологической науке накопление новой информации происходит настолько стремительно, а объемы данных таковы, что без применения информационных технологий интерпретировать их практически невозможно. Более того, компьютерные мощности уже отстают от темпов, с которыми новые технологии производят биологические данные (например, прогресс в области методик секвенирования геномов). Изучением и моделированием процессов в больших биологических системах занимается вычислительная биология. О том, зачем в Политехе создана научно-исследовательская лаборатория «Центр вычислительной биологии», рассказывает ее руководитель, ведущий научный сотрудник, кандидат биологических наук Александр Канапин.

- Когда и по чьей инициативе в Политехе создан Центр вычислительной биологии? Кто работает в этом научном подразделении?

- Лаборатория была создана в 2020 году по инициативе директора Института биомедицинских систем и биотехнологий (ИБСиБ) Андрея Владимировича Васина и проректора Виталия Владимировича Сергеева для развития новых направлений в области вычислительной биологии в Политехе. Сейчас у нас семь сотрудников, примерно треть из них – молодые ученые до 39 лет. Также есть иностранный ученый из Японии Хирата Дайскэ, с которым мы давно и плодотворно работаем. Наш японский коллега, к слову, недавно выиграл грант РФ по геномике белых медведей. Геномика и популяционная генетика входят в число направлений, которыми занимается наша лаборатория. Этот проект позволит изучить, как менялась популяция животных в ходе освоения русской Арктики, и его результаты могут быть использованы для разработки природоохранных стратегий.

- В чем уникальность лаборатории, есть ли аналоги в нашей стране?

- Биоинформатика развивается довольно быстро, сейчас и в Петербурге, и в целом по стране довольно много лабораторий и центров этого направления. Наша уникальность в том, что мы интегрируем самые современные методики и решаем мультидисциплинарные задачи в самом широком спектре – от медицины до сельского хозяйства. Мы принесли с собой опыт работы в ведущих мировых центрах, таких как Оксфорд, Гарвард, Европейский институт биоинформатики, а также компетенции, необходимые для реализации широкомасштабных геномных проектов. Ну и наконец, уникальность в том, что у нашего Центра два руководителя – помимо меня, еще один ведущий научный сотрудник нашей лаборатории Анастасия Самсонова. Мы дополняем

опыт и умения друг друга.

- С какими российскими и зарубежными научными центрами вы сотрудничаете?

- Прежде всего, это исследовательские проекты в области онкоиммунологии совместно с Оксфордским университетом и проекты по разработке вакцин совместно с НИИ Гриппа Минздрава. Мы взаимодействуем с Томским национальным исследовательским медицинским центром РАН и другими центрами медицинской генетики, разрабатываем концепцию экосистемы больших медико-генетических данных.

- Расскажите, пожалуйста, чем еще занимаются сотрудники вашей лаборатории?

- Центр вычислительной биологии – это междисциплинарная лаборатория. Мы занимаемся исследованиями и разработками в сфере аналитических, алгоритмических и инфраструктурных решений для вычислительной биологии, трансляционной биоинформатики и агрогеномики. Работаем в таких областях, как иммуноонкология, механизмы поддержания стабильности и целостности генома, биология мобильных элементов, популяционная и медицинская генетика. Участвуем в нескольких проектах по генетике сельскохозяйственных культур, направленных на достижение высокоточной селекции с помощью генетических маркеров и определение механизмов взаимодействия хозяина и патогена. Поскольку работа с большими данными имеет важное значение для экспериментальной науки, мы разрабатываем экосистемы больших данных и аналитики для стимулирования и поддержки биомедицинских, а также агрогеномных исследований на новом уровне.

- Да, сейчас развитие сельского хозяйства, особенно создание новых сортов, переживает качественный скачок. И связано это с возможностью анализировать индивидуальные геномы растений. Уже сегодня получают гибриды с теми или иными хозяйственными признаками, скрещивая известные линии растений, которые этими признаками не обладают. Это может быть пшеница, устойчивая к засухе или вредителям, или свекла, устойчивая к вирусу, вызывающему болезни этого растения, из-за чего можно потерять большую часть урожая.

- Мы принимаем участие в проекте Лаборатории математической биологии и биоинформатики по изучению механизмов устойчивости льна к фузариозному увяданию. Это заболевание растений, вызываемое различными видами грибов *Fusarium*, которое наносит серьезный экономический ущерб хозяйствам и агрохолдингам, занятым выращиванием льна. Система хозяин-паразит дает возможность детального изучения молекулярных механизмов их взаимодействия, включая инфекционный процесс, козволюцию льна и патогена как фактора изменения генетического разнообразия. Наши исследования вносят вклад в глобальные усилия, направленные на ликвидацию всплеск заболеваний растений, помогая в выведении новых устойчивых сортов сельскохозяйственных культур.

- Какой проект реализует ваша научная группа в рамках программы «Приоритет-2030»? В чем его суть и какой ожидается результат?

- В рамках программы «Приоритет-2030» мы развиваем как фундаментальные, так и прикладные исследования. Наш проект называется «Аналитика больших биомедицинских данных». Как я уже говорил, вместе с коллегами из Оксфордского университета мы работаем над проектами в новой перспективной области – онкоиммунологии. Целью исследования является разработка противораковых вакцин. В этом году мы подали в печать статью, в которой описан новый класс опухолевых антигенов – будущих кандидатов для вакцин. Кроме фундаментальных есть и прикладная задача, мы создаем программный комплекс для поиска таких антигенов применительно к любым заболеваниям, включая инфекционные. Также мы разрабатываем экосистемы больших биомедицинских данных, чтобы интегрировать все этапы их анализа – от получения «сырых» экспериментальных данных до автоматически сгенерированных отчетов для ученых или клиницистов.

- Так называемые омиксные технологии изучают, как устроен геном и как реализуется закодированная в нем информация. Они являются одним из главных инструментов геномной и постгеномной медицины, которая, как считают ученые, позволит улучшить качество диагностики и прогнозирования протекания болезни, а врач будет формировать индивидуальный набор лекарств, опираясь на «генетический паспорт» пациента. Пока это совсем новое направление в медицине, ему всего несколько лет. Ваша научная группа занимается омиксными технологиями? Что, на ваш взгляд, требуется для введения постгеномной медицины в практику наших больниц и как скоро это произойдет?

- Омиксные технологии – понятие более широкое, оно относится к новым методам изучения живых объектов и получения детальной информации об их молекулярной структуре, и геномная медицина – лишь одна из областей, где могут применяться эти методологии. Пока основные направления, в которых применяются методы геномики в медицине, – это диагностика наследственных заболеваний и стратификация пациентов в онкологии. До генетического паспорта пациента пока достаточно далеко. Наш Центр занимается исследованиями омиксных данных, то есть аналитикой тех результатов экспериментов, которые ставят биологи и клиницисты. Если говорить о практической стороне внедрения геномики в рутинную медицинскую практику, на наш взгляд, первостепенной задачей является разработка стандартов методов анализа данных и законодательных нормативов, регулирующих эту область деятельности.

- На форуме «Передовые цифровые и производственные технологии» в конце прошлого года вы участвовали в круглом столе, посвященном применению цифровых технологий, в частности цифровых двойников, для создания новых сортов культурных растений. Несмотря на то, что это очень важный фактор, для агробиологии, по вашему мнению, он довольно сильно отстает.

- Да, это верно. Математика в этой области пока что обгоняет биологию. Существующие двойники – это некий внешний цифровой скелет, за которым пока не стоит каких-то биологических механизмов. Необходимо найти за этими базами данных и статистическими таблицами реальные биологические процессы, и уже начинать создавать их модели – вот это будет настоящий цифровой двойник растения. А пока у нас есть только внешняя, достаточно далекая от биологической реальности, его оболочка. Которая все же имеет свою область практического применения и используется на практике, например в поиске селекционных

маркеров.

- В одной из своих статей вы пишете, что очередных прорывов в биологии ждать осталось недолго. Однако без биоинформатиков - этих «бухгалтеров от биологии», как их назвал нобелевский лауреат Сидни Бреннер, такие прорывы точно невозможны.

- Наверно, я был излишне оптимистичен восемь лет назад, когда писалась эта статья. Накопление данных идет огромными темпами, но пока мы все еще «ищем под фонарем», то есть преимущественно используем геномные данные, которые легче и дешевле получать. Они, безусловно, важны, но дают статичную картину, описание генотипа организма. Данные, которые позволяют посмотреть на динамику биологических процессов - транскриптомика, протеомика, метаболомика, эпигеномика и другие, - их пока мало, по сравнению с геномами и этих объемов недостаточно для того, чтобы полноценно охарактеризовать живой организм. Однако их накопление тоже идет, и прорыв - просто вопрос времени.