

УДК 612.17.001

О.Е. Баксанский,
д.ф.н., ведущий научный сотрудник Института
философии РАН, профессор кафедры теории и
технологии обучения в высшей школе Первого МГМУ
им. И.М. Сеченова

O.E. Baksansky,
Doctor of philosophy, leading researcher of the Institute
of philosophy (RAS), prof. of the chair of theory and
technology of learning in higher school of the
I.M. Sechenov First MSU

БИОИНЖЕНЕРИЯ И БИОИНФОРМАТИКА: КОНВЕРГЕНТНЫЕ ТЕХНОЛОГИИ

BIOENGINEERING AND BIOINFORMATICS: CONVERGENT TECHNOLOGIES

КОНТАКТНАЯ ИНФОРМАЦИЯ:

Олег Евгеньевич Баксанский, профессор кафедры
теории и технологии обучения в высшей школе
Адрес: 119991, г. Москва, ул. Б. Пироговская, д. 2,
стр. 10
Телефон: 8 (499) 248–06–29
E-mail: obucks@mail.ru
Статья поступила в редакцию: 30.11.2014
Статья принята к печати: 03.02.2015

CONTACT INFORMATION:

Oleg Evgenievich Baksansky, prof. of the chair of theory
and technology of learning in higher school
Address: 2/10 B. Pirogovskaya str., Moscow, 119991
Tel.: 8 (499) 248–06–29
E-mail: obucks@mail.ru
The article received: 30.11.2014
The article approved for publication: 03.02.2015

Аннотация. Цель настоящего исследования — аналитический обзор современного понимания биоинженерии и биоинформатики и перспективы дальнейшего развития. Быстрое развитие этих областей за последние годы позволило ученым перейти от простого исследования природных биообъектов к их изменению и усовершенствованию, улучшению их полезных свойств, к созданию совершенно новых биологических объектов, не существующих в природе. Нано-, био-, инфо-, когнитивные технологии — основа развития науки и технологии будущего, перехода на принципиально новые, возобновляемые ресурсы и технологии, созданные по образцу живой природы, с использованием современных технологических достижений.

Annotation. The purpose of this study is the analytical overview of the current understanding of bioengineering and bioinformatics and prospects for further development. The rapid development of these areas in recent years has allowed scientists to go from the simple study of natural biological objects to their change and improvement, improve their beneficial properties to the creation of entirely new biological objects that do not exist in nature. Nano-, bio-, info-, cognitive technologies are the basis of the development of science and technology of the future, the transition to a fundamentally new line — renewable resources and technology, modeled on nature, using the latest technological advances.

Ключевые слова. Биоинженерия, биоинформатика, вычислительная биология, конвергентные технологии.
Keywords. Bioengineering, bioinformatics, computational biology, convergent technologies.

В современных перспективных исследованиях эксперты выделили следующие основные тенденции научно-технологического развития:

- усиление конвергенции технологий;
- усиление проникновения современных высоких технологий в различные сектора производственной сферы;
- растущее значение мультидисциплинарности научных исследований;
- усиление воздействия новых технологий на управление, стимулирующее развитие гибких сетевых структур.

В рамках каждой из этих тенденций формируются многообещающие новые технологии и области науки с точки зрения их потенциального применения в различных сферах человеческой деятельности.

Эти технологии потенциально являются ответами на глобальные вызовы и формируют новый технологический образ мира, что является предметом рассмотрения в настоящей работе.

С 2010 г. на фармацевтическом факультете Первого МГМУ им. И.М. Сеченова, помимо традиционной специальности «Фармация», подготовка ди-

пломированных специалистов осуществляется и по специальности «Биоинженерия и биоинформатика», срок обучения — 5 лет.

Развитие биологии привело к широкому использованию совершенно новых практических подходов для решения проблем здравоохранения, для разработки принципиально новых технологий и материалов в различных отраслях.

Биоинженерия (англ. bioengineering) — это одно из современных направлений науки, возникшее на стыке физико-химической биологии, биофизики, генной инженерии и компьютерных технологий. Быстрое развитие этих областей за последние годы позволило ученым перейти от простого исследования природных биообъектов к их изменению и усовершенствованию, улучшению их полезных свойств, созданию совершенно новых биологических объектов, не существующих в природе. Сфера деятельности биоинженерии простирается от создания искусственных органов для компенсации сниженных или утраченных физиологических функций (биомедицинская инженерия) до разработки генетически модифицированных организмов, например, сельскохозяйственных растений и животных (генетическая инженерия). Среди задач биоинженерии — искусственные белки, выполняющие заданные функции, новые клеточные структуры, обладающие полезными свойствами, и даже целые живые организмы, сконструированные для нужд человека.

Спектр научных направлений биоинженерии довольно широк:

- молекулярное моделирование различных биологических объектов и изучение динамики макромолекул;
- структурные особенности и взаимодействие макромолекул;
- дизайн искусственных белков с заданными свойствами, а также синтез и изучение свойств таких белков;
- генетические маркеры выносливости и работоспособности человека;
- исследование внутриклеточного транспорта токсичных молекул и т. д.

Биоинформатика (или *вычислительная биология*) — дисциплина, исследующая использование компьютеров для решения биологических задач. Под биоинформатикой понимают любое использование компьютеров для обработки биологической информации. На практике, иногда это определение более узкое, под ним понимают использование компьютеров для характеристики молекулярных компонентов.

Термины *биоинформатика* и *вычислительная биология* часто взаимозаменяются, хотя последний чаще указывает на разработку алгоритмов и конкретные вычислительные методы. Считается, что не

все использование вычислительной биологии является биоинформатикой, например, математическое моделирование — это не биоинформатика, хотя и связана с биологическими задачами.

Биоинформатика использует методы прикладной математики, статистики и информатики. Исследования в вычислительной биологии нередко пересекаются с системной биологией. Основные усилия исследователей в этой области направлены на изучение геномов, анализ и предсказание структуры белков, анализ и предсказание взаимодействий молекул белка друг с другом и другими молекулами, а также моделирование эволюции. Биоинформатика и ее методы используются также в биохимии и биофизике. Основная линия в проектах биоинформатики — это использование математических средств для извлечения полезной информации из «шумных» или слишком объемных данных о структуре ДНК и белков, полученных экспериментально.

Современная биоинформатика — весьма молодая, бурно развивающаяся и довольно непривычная для математиков и информатиков наука. Фактически это собрание различных математических моделей и методов в помощь биологам (в основном молекулярным биологам) для решения чисто биологических задач (предсказание пространственной структуры белков, расшифровка структуры ДНК, хранение, поиск и аннотирование биологической информации и т. п.). Несмотря на наличие корня «информатика» в названии этой науки, она настолько пропитана биологией, что без довольно глубоких биологических знаний делать математику в ней нечего. В то же время область приложения математических знаний, в т. ч. и самых современных математических теорий, здесь просто необъятна.

Биоинформатика ставит своей целью анализ и систематизацию огромного объема информации о живой природе. Эта задача требует широкого применения математических методов. Анализ огромного объема информации о структуре геномов «простейших» микроорганизмов, а тем более генома человека, поиск взаимосвязей в живых организмах, изучение эволюции, поиск гомологий в ключевых признаках — решение всех этих вопросов требует создания новых математических подходов.

Достижения в этой области позволяют расшифровывать структуру и функции генетического аппарата клетки и разрабатывать планы создания биологических объектов, просто сидя за компьютером. Развитие биоинформатики позволит установить взаимосвязь между структурой и функцией в генетическом аппарате клетки и разработать новые пути управления и реконструкции биологических объектов, что должно предоставить невиданные до сих пор возможности для развития медицины и новых технологий для различных отраслей народного хозяйства. Как отмечает М.С. Гельфанд, по сути, био-

информатика представляет собой несколько удобных технологий и набор конкретных задач, которые решают с их помощью [1].

Предмет биоинформатики — анализ геномов и создание модели обмена веществ. Геном — это «текст», состоящий из нескольких миллиардов «букв», их вводят в память компьютера, и можно, получая картинки, сравнивать геномы разных живых существ: человека и шимпанзе, мышки и собаки, одной и другой бактерии.

И биоинженер, и биоинформатик должны владеть математическими методами, иметь химические и биологические знания, компьютерные навыки.

Что требуется для проведения научных исследований в этих областях?

Тому, кто занимается биоинформатикой, прежде всего, нужен мощный суперкомпьютер. А биоинженерия — очень дорогая наука, и требуется дорогостоящее приборное оснащение.

Наиболее известной и наиболее эффективной областью применения биоинформатики в настоящее время является анализ геномов, тесно связанный с анализом последовательностей.

Этот раздел биоинформатики изучает нуклеотидные и белковые последовательности. В настоящее время разработаны эффективные экспериментальные методы определения нуклеотидных последовательностей. Определение нуклеотидных последовательностей стало рутинной хорошо автоматизированной процедурой.

Первый тип биоинформатической задачи — это задачи поиска в нуклеотидных последовательностях особых участков, кодирующих первичную аминокислотную последовательность белков, РНК, участков связывания с регуляторными белками и др. И это не всегда простые задачи; например, гены эукариотических организмов состоят из чередующихся «осмысленных» и «бессмысленных» фрагментов (экзонов и интронов), и расстояние между «осмысленными» фрагментами может достигать тысяч нуклеотидов.

Пусть ген найден. Что он кодирует? Зачем он нужен?

Если речь идет об участке ДНК, кодирующем белок, то с помощью весьма простой операции — трансляции с использованием известного генетического кода — можно получить аминокислотные (белковые) последовательности. Из известных на сегодня 4 273 512 белков около 94% последовательностей — это именно такие гипотетические трансляты, и больше о них ничего не известно. Скорость поступления информации с автоматических секвенаторов превышает скорость нашего понимания ее смысла!

Генетические тексты — тексты с большой долей шума; сравнивая родственные последовательности, в ряде случаев удается отфильтровать шум и

выявить сигнал, например, короткую последовательность нуклеотидов, способную связываться с белком-регулятором, или аминокислотные остатки в ферменте, отвечающие за связывание субстрата. Чтобы быть уверенными в результате, специалисты в области биоинформатики используют теорию вероятности и математическую статистику.

Создание новых экспериментальных технологий ставит перед биоинформатикой целый ряд новых задач. Например, развитие масс-спектрометрии позволяет (пока в принципе) в одном эксперименте проанализировать весь набор белков, присутствующий в клетке. Для решения этой задачи необходим совместный анализ спектров масс и геномов. Открытие новых биологических явлений и механизмов также приводит к появлению новых задач. Хорошим примером служит открытие РНК-интерференции, за которую в 2006 г. присудили Нобелевскую премию по физиологии. Это открытие способствовало появлению целой серии биоинформатических работ, посвященных поиску участков связывания микро-РНК. Многие находки были затем подтверждены экспериментально.

Каждый белок, помимо своей уникальной последовательности аминокислот, из цепочки которых состоит его молекула, обладает еще и уникальным способом укладки этой цепочки в пространстве. Задачу предсказания укладки по последовательности можно, в принципе, тоже считать задачей биоинформатики, но это задача в общем виде еще слишком далека от своего решения. Поэтому структурная биоинформатика занимается анализом пространственных структур, уже определенных экспериментально.

Структур белков известно намного меньше, чем последовательностей белков. Это связано с тем, что экспериментальные процедуры для определения структуры намного сложнее, дороже, и к тому же (в отличие от секвенирования) не являются «рутинными», т. е. их результат вовсе не гарантирован. Например, большинство лекарств представляют собой химические вещества, молекулы которых способны связываться — образовывать комплексы — с молекулами тех или иных белков (как правило, в результате такого связывания белок оказывается неспособен выполнять свою природную функцию, что и обеспечивает эффект лекарства). Исследование механизма действия лекарств имеет большое практическое значение, поэтому определением структуры комплексов молекул белков с молекулами лекарств занимаются многие экспериментальные группы.

Примеры задач структурной биоинформатики:

- определение участков белковой молекулы, важных для той или иной функции данного белка;
- сравнительный анализ структур родственных белков, классификация белков на основе их пространственной структуры;

— анализ структур комплексов двух или нескольких молекул белка, комплексов молекул белка с другими молекулами; предсказание воздействия молекул химических веществ на молекулы белков;

— предсказание структуры белка по структуре белка с похожей последовательностью.

В настоящее время определены полные или почти полные последовательности геномов многих организмов. Прочтение полной нуклеотидной последовательности какого-либо генома не является самоцелью. На самом деле это является первым шагом для исследования функционирования той или иной клетки. Анализ геномов бактерий проводится для того, чтобы исследовать метаболизм бактерий и, в случае патогенных организмов, найти потенциальные мишени для лекарств. Экспериментальное определение функции только одного гена требует интенсивной работы одной лаборатории как минимум в течение нескольких месяцев. Компьютерный же анализ позволяет с известной степенью точности охарактеризовать несколько тысяч генов силами небольшой группы примерно за неделю. Разумеется, компьютерный анализ не исключает экспериментальную проверку, однако в этом случае экспериментальная работа существенно упрощается.

Компьютерный анализ геномов состоит из следующих основных элементов:

— Предсказание генов в последовательностях. При этом в некоторых случаях удается даже найти ошибки в последовательности.

— Предварительная аннотация по сходству и другим особенностям белковых последовательностей.

— Сравнительный анализ геномов.

— Исследование регуляции работы генов.

Другой класс исследований, проводимых компьютерной геномикой — полногеномный анализ и исследование эволюции. В частности, с помощью массового анализа было обнаружено, что альтернативный сплайсинг в генах человека является скорее правилом, чем исключением. Эволюционный взгляд на проблему позволяет выдвинуть гипотезу о том, что сплайсинг, в частности альтернативный сплайсинг, является эффективным механизмом для эволюции, позволяющим без значительного риска для генома перебирать варианты последовательностей.

Существует широкий спектр методов и инструментов для компьютерного анализа биологических данных. Однако важно понимать границы применимости тех или иных методов — важна четкость постановки и необходимы соответствующие контроли.

Иногда существующие программы недостаточны для решения поставленных задач: или существующие программы имеют не достаточную точность, или для интересующей исследователя биологической задачи нет подходящих средств, или появился

новый тип данных. В этом случае приходится разрабатывать новые алгоритмы и программы.

Особенностями специальностей являются значительное увеличение цикла математических и естественно-научных дисциплин (математика, информатика, физика, общая и неорганическая химия, аналитическая химия, органическая химия, физическая химия) и общепрофессиональных дисциплин (ботаника, зоология, молекулярная биология, биохимия, биоинформатика, микробиология, генетика, молекулярная генетика, вирусология, клеточная биология, иммунология, цитология, эмбриология, нейробиология и др.).

На старших курсах, согласно учебным планам, будут изучаться дисциплины биотехнологического и биоинженерного профиля (генная инженерия, биоинженерия микроорганизмов, растений и животных, клонирование и трансплантация клеток, белковая инженерия, инженерная энзимология и др.), а также дисциплины, связанные с проектированием и оборудованием предприятий биотехнологической промышленности.

Кроме теоретической подготовки, большое внимание уделено практической составляющей подготовки специалистов. В учебном процессе будут активно использоваться деловые игры.

Демонстрационный материал будет представлен с помощью мультимедийной техники и видеофильмов. В помощь студентам на кафедрах созданы электронные учебники и электронные библиотеки. С помощью компьютеров будет проходить обучение, контроль и самоконтроль усвоения материала. В процессе обучения студенты будут выполнять и защищать курсовые работы, дипломные проекты и дипломные (квалификационные работы). Базой выполнения дипломных работ являются кафедры факультета и другие профильные институты.

Исследования в области биоинформатики и вычислительной молекулярной биологии проводятся в следующих областях:

— Разработка алгоритмов (распознавание генов и регуляторных сайтов, предсказание структуры РНК);

— Факультет биоинженерии и биоинформатики МГУ, Москва;

— Институт физико-химической биологии им. Белозерского МГУ, Москва;

— Институт проблем передачи информации им. Харкевича РАН, Москва;

— Институт математических проблем биологии РАН, Пушкино;

— Институт цитологии и генетики СО РАН, Новосибирск;

— ГНЦ Институт генетики и селекции промышленных микроорганизмов ГосНИИГенетика, Москва;

— Аннотация генов и геномов:

- Институт проблем передачи информации им. Харкевича РАН, Москва;
- Анализ регуляторных сигналов:
- Факультет биоинженерии и биоинформатики МГУ, Москва;
- Институт проблем передачи информации им. Харкевича РАН, Москва;
- Институт цитологии и генетики СО РАН, Новосибирск;
- ГНЦ Институт генетики и селекции промышленных микроорганизмов ГосНИИГенетика, Москва;
- Транскриптомика и анализ экспрессии генов:
- Физический факультет МГУ, Москва;
- Институт общей генетики им. Вавилова РАН, Москва;
- Институт биомедицинской химии им. Ореховича РАН, Москва;
- Структура белка:
- Институт белка РАН, Пушкино;
- Институт молекулярной биологии им. Энгельгардта РАН, Москва;
- Институт цитологии и генетики СО РАН, Новосибирск;
- Молекулярная динамика пептидов, белков и РНК:
- Институт физико-химической биологии им. Белозерского МГУ, Москва;
- Биологический факультет МГУ, Москва;
- Институт биоорганической химии им. Шемякина и Овчинникова РАН, Москва;
- Белок-ДНКовые взаимодействия:
- Факультет биоинженерии и биоинформатики МГУ, Москва;
- Институт физико-химической биологии им. Белозерского МГУ, Москва;
- Институт проблем передачи информации им. Харкевича РАН, Москва;
- Моделирование метаболических и регуляторных сетей:
- Институт физико-химической биологии им. Белозерского МГУ, Москва;
- Институт цитологии и генетики СО РАН, Новосибирск;
- Санкт-Петербургский государственный политехнический университет;
- Рациональное проектирование лекарств:
- Институт биомедицинской химии им. Ореховича РАН, Москва;
- Молекулярная эволюция и таксономия:
- Институт физико-химической биологии им. Белозерского МГУ, Москва;
- Институт проблем передачи информации им. Харкевича РАН, Москва;
- Институт цитологии и генетики СО РАН, Новосибирск;
- Биологический факультет МГУ, Москва;
- Институт общей генетики им. Вавилова РАН, Москва;

- Институт цитологии РАН, Санкт-Петербург;
- ГНЦ Институт генетики и селекции промышленных микроорганизмов ГосНИИГенетика, Москва;
- Институт физико-химической медицины, Москва.

Совместные исследования проводятся с группами из Германии, Франции, Великобритании, США, Испании, Италии, Швеции, Индии, Китая. Эти исследования поддерживаются двусторонними программами, координируемыми РФФИ. В прошлом ряд совместных проектов поддерживался INTAS. Попытки участия в Шестой и Седьмой рамочных программах пока были безуспешными, однако они продолжаются.

В России действуют филиалы нескольких компаний, работающих в области автоматического анализа литературных данных по белок-белковым взаимодействиям и сигнальным путям и приложения этого анализа к задаче поиска мишеней для лекарств (GeneGo, США; Ariadne Genomics, США; Biobase, Германия; Biomax, Германия), а также поиска потенциальных лекарств путем анализа взаимодействия белков и малых молекул (Algodign, США).

При этом развитие биоинженерии и биоинформатики требует применения новых исследовательских подходов, в качестве которых могут выступать конвергентные технологии.

Конвергентные технологии сегодня — это не просто новое научное направление, они представляют собой переход к новой парадигме развития науки и общества — новому качеству науки и техники, а также новому состоянию экономики и технологии. Сегодня мировая наука стоит перед необходимостью отказываться от узкой специализации и переходить к междисциплинарным исследованиям. Нано-, био-, инфо-, когнитивные технологии — основа развития науки и технологии будущего, перехода на принципиально новые, возобновляемые ресурсы и технологии, созданные по образцу живой природы, с использованием современных технологических достижений. Причем это не просто аддитивное соединение одной технологии с другой, а их конвергенция, взаимопроникновение, ориентация на единую цель. Все это требует тщательного философско-методологического осмысления.

Исследования и разработки в сфере когнитивных технологий могут быть осуществлены учеными и специалистами, владеющими, прежде всего, универсальным знанием, математической культурой и системным мышлением, а это отличительные черты и конкурентные преимущества ученых, входящих в группу российской науки. Первоначальными шагами в данном направлении могли бы стать:

- анализ методологии конвергентных технологий, включая реферирование имеющихся подходов;

- создание центра компетенции по конвергентным технологиям и определение технологического облика России в данной сфере;

- организация разработки нового класса продуктов и механизмов коммерциализации конвергентных технологий;

- разработка новых учебных программ и внедрение образовательных практик в сфере конвергентных технологий;

- разработка методологии управления инновациями и координация деятельности в сфере конвергентных технологий.

Конвергентные технологии задают новую стратегию развития цивилизации и в этом качестве нуждаются во всестороннем социально-гуманитарном осмыслении, для чего необходимо эффективно использовать имеющийся интеллектуально-кадровый потенциал, традиции фундаментальной науки и способности к производству универсального знания для создания когнитивного стандарта как одного из ключевых направлений глобального развития конвергентных технологий.

ЗАКЛЮЧЕНИЕ

В настоящее время в западной научной литературе закрепился термин «конвергенция технологий» или «конвергентные технологии» под которым понимается широкий круг процессов — как конвергенция отдельных областей наук, так и непосредственно технологий. При этом высказываются две диаметрально противоположные точки зрения на существо самого процесса конвергенции:

- простая междисциплинарная конвергенция на основе горизонтального влияния нанотехнологии на другие технологии;

- появление полностью новых направлений науки и технологии, которые в будущем будут развиваться по своим собственным траекториям.

Подтверждением идущих процессов конвергенции могут служить государственная стратегия финансирования новых направлений. Библиометрические исследования свидетельствуют о том, что за последние 10 лет резко возросло число публикаций в сферах «пересечения» областей конвергентных технологий.

Наибольшие ожидания эксперты связывают с развитием нанотехнологии, которая становится стержнем формирования новых отраслевых комплексов. В связи этим выделяют несколько видов кластеров:

- нанотехнология + ИКТ;
- нанобиотехнология + ИКТ;
- когнитивные науки + ИКТ;
- нанотехнология + материаловедение + ИКТ.

Развитие нанотехнологии основано на интеграции целого ряда дисциплин: химии, физики, механики, материаловедения, электроники и т.д. В краткосрочной перспективе применение нанотехнологий скажется, прежде всего, на традиционных отраслях, в долгосрочной перспективе принципиальные достижения дадут толчок к появлению новых секторов и рынков.

При этом следует обратить внимание на то, что конвергенция технологий носит взаимонаправленный характер.

Список литературы

1. Гельфанд М.С. Что может биоинформатика? // *Химия и жизнь*. 2009; 9. <http://elementy.ru/lib/430895>
[Gelfand M.S. What can bioinformatics? // *Khimiya i zhizn'*. 2009; 9. <http://elementy.ru/lib/430895>]