

РЕЦЕНЗИИ

DOI: 10.31116/tsitol.2018.06.10

В. Е. Стефанов, А. А. Тулуб, Г. Р. Мавропуло-Столяренко. Биоинформатика. М.: «Издательство Юрайт», 2016. 252 с.

В самом начале будет уместным заметить, что биологические науки — науки преимущественно экспериментальные, их достижения, выводы и методология основываются на эксперименте. Теоретическим сторонам и направлениям традиционно отводилось не слишком большое место.

Развитие информационных и компьютерных технологий за последние 40 лет оказало на эти науки существенное влияние, сблизив их с физикой, химией и математикой. Из них наибольшее пересечение интересов произошло в области молекулярной биологии, для которой передача, хранение и обработка информации играют ключевую роль. В результате произошедшего заметно возросла значимость теоретической биологии. Сегодня она охватывает различные разделы, включая нейронные сети, теорию групп, нечеткие множества, квантовую биохимию, кибернетику, теорию информации и т. д. Наиболее востребованным разделом теоретической биологии является биоинформатика.

Поначалу термин «биоинформатика», предложенный Б. Хеспером и П. Хогевегом в 1970 г., подразумевал «изучение информационных процессов в биотических системах» (Hesper B., Hogeweg P., 1970. *Kameleon*. 1 (6) : 28—29). Позже, в начале «геномной революции», термин «биоинформатика» был переоткрыт и стал обозначать «создание и техническое обслуживание базы данных для хранения биологической информации» (Hogeweg, 2011. *PLoS Comput. Biol.* 7 (3) : e1002021).

Сегодня биоинформатика — это не только самостоятельная специальность (у нас в России она имеет уже даже отдельный шифр ВАК), но и важнейший компонент любых современных исследований в области молекулярной генетики, молекулярной биологии, нейрофизиологии, биофизики, экологии и антропологии. Область применения данной специальности чрезвычайно широка. Она интегрирована в решение вопросов протеомики и геномики, сравнительной палеонтологии и биологической эволюции, биотехнологии, компьютерного моделирования структур и процессов на молекулярном, органном и организменном уровнях. Это исключительно важный момент, так как моделирование белковых и других молекулярных структур, скажем, методами молекулярной динамики позволяет выявить и (или) предсказать в том числе и такие особенности, которые недоступны прямому экспериментальному наблюдению. Поэтому компьютерный эксперимент уже сейчас может быть равноправен (конеч-

но, при правильно заданных исходных параметрах) эксперименту реальному.

Замечу, что на Западе издано уже довольно большое число книг по биоинформатике. Правда, они преимущественно академической и прикладной направленности для специалистов (см. книги: Gu J., Bourne P. E., 2011. *Structural Bioinformatics*. Wiley-Blackwell Publ. USA; Pevsner J., 2013. *Bioinformatics and Functional Genomics*. New York: Springer; Bergeron B., 2015. *Bioinformatics Computing*. MIT Publ. USA), хотя имеются издания студенческо-аспирантской адресованности. Например, не так давно вышел в свет учебник Дж. Кейта (Keith J. M., 2017. *Bioinformatics*. London: Humana Press). Отечественный же книжный рынок грешит отсутствием подобных изданий по биоинформатике. Поэтому публикацию учебника сотрудников С.-Петербургского государственного университета В. Е. Стефанова, А. А. Тулуба, Г. М. Мавропуло-Столяренко в сентябре 2016 г. следует всячески приветствовать. Он написан в соответствии с новыми образовательными стандартами, адресован в первую очередь студентам бакалавриата, однако ряд обсуждаемых в нем вопросов может быть интересен и даже необходим студентам магистратуры, а также и начинающим исследователям.

С момента выхода учебника и написания настоящей рецензии минуло больше года, что в определенной мере послужило качественной проверкой реальной эффективности рассматриваемых в настоящем издании понятий, положений, закономерностей, методических приемов, рекомендаций и т. д.

Учебник невелик и состоит из шести разделов, снабженных большим числом иллюстраций, введения и заключения. Каждый из разделов сопровождается списком цитированной литературы с перечнем вопросов и небольших заданий, призванных закрепить у читателя полученные знания.

Все начинается с общих свойств биологических систем (структура азотистых оснований, правила спаривания, организация ДНК и РНК молекул, структура белков), а также вопросов переноса информации на клеточном уровне. Четко и понятно очерчены подходы и задачи, стоящие перед биоинформатикой.

После этого следует рассмотрение математического аппарата биоинформатики. Тут уже появляются статистические модели типа оценок параметров случайных величин, метода скользящего окна и др. Подробно рассмотрены понятия «информационной энтропии» и «совместной информации». Не забыты вопросы «выравнивания» аминокислотных последовательностей, подробно обсуждаются алгоритмы «выравнивания» (глобальный,

локальные, множественные). Достаточно подробно рассмотрены «цепи Маркова» и «скрытые модели Маркова», «случайные блуждания», приведены алгоритмы, применяемые при работе со скрытыми моделями Маркова и т. д. В виде отдельного раздела выделены «искусственные нейронные сети», представляющие собой математические модели «с памятью». Удачным, как мне показалось, и уместным оказываются здесь и введение понятия «обучение нейрона», а также алгоритм работы с этим объектом. Рассмотрены «сети Хопфилда» как разновидность «архитектуры» обучения нейрона.

Более того, достаточно много внимания в книге уделено понятию «байесовской» вероятности и статистики с алгоритмами их реализации. Нельзя пройти мимо подробно рассмотренных современных методов Монте-Карло в их применении к биологическим структурам. Детально также представлены «эволюционные алгоритмы» и «кластерный» анализ иерархических структур. Мне представляется, что это одна из наиболее трудных в изложении и восприятии часть учебника. Она требует немалых усилий по преодолению математического барьера со стороны студентов-биологов, математическая подготовка которых значительно уступает таковой физиков и математиков. Считаю, тем не менее, что авторы учебника блестяще справились со стоявшей перед ними непростой задачей, изложив материал в легкой и доступной для биолога форме.

Один из разделов книги посвящен моделям в биологии. Он целиком опирается на изложенные на предыдущих страницах сведения и раскрывает сущность процесса моделирования при поиске оптимальных структур. Часть материала учебника отведена современным «базам данных» и поиску по ним информации. Для будущих биологов это краеугольный камень знания. А если говорить строго, то умелое общение с базами данных позволяет начинающему исследователю выборочно пользоваться необходимыми сведениями, накопленными наукой к настоящему времени, т. е. является превосходным инструментом учебного процесса в целом.

Еще один из разделов книги посвящен биоинформационным программам и сервисам. По сути он служит

продолжением многого того, что было уже известно ранее. Особенно следует остановиться на части, которая во введении к учебнику почему-то названа авторами факультативной. Тем не менее ее значение крайне важно. Здесь, как я думаю, сделана попытка придать транслируемой в клетке информации глубокий физический смысл. Этот смысл основывается на квантово-механических представлениях о молекулярных структурах, в которых имеет место соприкосновение поверхностей потенциальной энергии. В областях соприкосновения происходит (или не происходит) изменение спиновых состояний (Тулуб, Стефанов, 2010. Спинтроника нуклеотидов. СПб.: Наука), которые определяют течение последующих биохимических процессов. Параллельно направленные спины (триплетные состояния) запрещают молекулярным доменам взаимодействовать, т. е. «узнавать» друг друга. Напротив, противоположно направленные спины (синглетные состояния) способствуют распознаванию доменов. Эта уникальная способность спинов позволяет биологическим объектам «видеть» друг друга на клеточном и даже межклеточном уровне, порождая информационные ансамбли с архитектурой, включающей логические вентили «да» и «нет». Такой подход предполагает наличие у доменов «спиновой фазы», являющейся прямым следствием законов квантового мира. Спиновая фаза управляет физиологическими процессами и направляет их в нужное русло. Помимо этого, спиновая фаза порождает квантовый компьютеринг, который, по мнению ряда специалистов, базируется на нецелочисленном кодировании информации.

В целом следует признать учебник состоявшимся и крайне полезным для студентов-биологов, специализирующихся в области молекулярной биологии, нейробиологии и физиологии организма в целом. По своему подходу и уровню излагаемого материала учебник существенно обогащает довольно ограниченный пока набор учебной литературы в бурно развивающейся и широко востребованной области математической биологии, имеющей фундаментальный и прикладной потенциал.

© А. Д. Ноздрачев

V. E. Stefanov, A. A. Tulub, G. R. Mavropulo-Stolyarenko. Bioinformatics. Moscow: Yurait Publ., 2016. 252 p.

A. D. Nozdrachev