

ПЕРСПЕКТИВЫ БИОИНФОРМАТИКИ

С.А. Грязнов, канд. пед. наук, доцент
Самарский юридический институт ФСИН России
(Россия, г. Самара)

DOI:10.24412/2500-1000-2021-6-2-100-102

Аннотация. Биоинформатика – это одна из самых динамических наук в наше время. Стимулирующие факторы роста мирового рынка биоинформатики включают растущий спрос на нуклеиновые кислоты и секвенирование белков, рост инициатив со стороны государственных и частных организаций, ускорение роста протеомики и геномики, а также увеличение исследований в области молекулярной биологии и открытия лекарств. В данной статье рассмотрены перспективы лечения болезни Альцгеймера в связи с ростом развития биоинформатики, а также навыки, которыми должен обладать современный специалист в данной области.

Ключевые слова: биоинформатика, большие данные, геномные данные, секвенирование генома, анализ ДНК, омики.

Многие века единственным способом исследования живых существ для биологов было наблюдение, где основным инструментом являлся скальпель. Пока не был открыт геном, считалось, что организмы состоят только из органов. Позже ученые продвинулись в исследованиях до уровня ДНК, а главным их инструментом стал компьютер. Таким образом, на стыке биологии и математического моделирования возникла наука – биоинформатика.

Историю биоинформатики следует вести от Фредерика Сэнгера, английского ученого, получившего в 1980 году Нобелевскую премию по химии за открытие способа прочтения последовательности ДНК. С тех пор, с каждым годом методы чтения последовательностей совершенствуются, однако метод «секвенирования по Сэнгеру» послужил основой для всех дальнейших исследований в этой области.

Сегодня геном можно прочитать разными методами, одним из которых является так называемое «секвенирование нового поколения». Для проведения этой процедуры нужно сначала получить биологические образцы. В каждой клетке организма геном одинаковый, поэтому чаще всего для чтения генома берут кровь (это проще всего). После этого клетки разрушают и отделяют ДНК от всего остального. Затем, полученную ДНК дробят на множество маленьких кусочков и «пришивают» к

каждому из них специальные адаптеры – искусственно синтезированные известные последовательности нуклеотидов. Потом цепочки ДНК разделяют, и однонитевые цепочки с помощью адаптеров присоединяют к специальной плашке, на которой проводится секвенирование. В ходе секвенирования к последовательности ДНК присоединяются комплементарные флуоресцентно меченые нуклеотиды. Каждый меченый нуклеотид при присоединении испускает пучок света определенной длины волны, что фиксируется на компьютере. Так компьютер прочитывает короткие последовательности исходной ДНК, которые потом с помощью специальных алгоритмов собираются в исходный геном [1].

Геном – это находящийся в каждой клетке организма набор числа генов. Чтобы понять, как именно функционируют живые организмы на уровне клеток, необходимо проанализировать тысячи собираемых по всему миру геномов, в том числе древних людей, животных или растений. Биологи-экспериментаторы получают из ДНК геном, а биоинформатики его расшифровывают и обрабатывают полученные данные. Благодаря вторым первые могут сравнивать древние и современные геномы. Это позволяет изучать эволюцию видов и пытаться предсказывать дальнейшее развитие живых существ. В перспективе анализ генома позволит научиться

лечить сложные болезни и значительно увеличить продолжительность жизни.

Биоинформатика, как все науки, делится на прикладную и фундаментальную. Фундаментальная биоинформатика – это эволюционная молекулярная биология. Прикладная биоинформатика – это расшифровка данных последовательностей геномов и структуры белков. С ее помощью можно ответить на вопрос, как работает клетка в целом, что делает определенный белок, как регулируется нужный ген и прочее.

В свое время популяризатор эволюционной доктрины Эрнст Геккель подделал рисунки эмбрионов животных, чтобы увеличить их сходство с человеком и подчеркнуть тем самым родственные связи организмов [2]. Сегодня можно просто ввести информацию в компьютер и сравнить (увидеть), как молекулярные изменения в ранних зародышах транслируются потом в морфологические различия. Побочным продуктом будет понимание развития уродств, генная терапия и тому подобное – все это прикладная биоинформатика.

Для того чтобы изобрести лекарство от болезни, необходимо точно выявить причину этой болезни. Так, ученые всего мира активно занимаются исследованием болезни Альцгеймера, от которой в настоящее время нет лекарства. Именно развитие биоинформатики позволяет предположить, что в скором времени лекарство будет найдено. Болезнь Альцгеймера – сложное нейродегенеративное заболевание, вызываемое несколькими генетическими и эпигенетическими факторами. Они обусловлены генетической изменчивостью, изменениями в экспрессии генов, а также взаимодействиями генов и окружающей среды, которые могут быть вовлечены в начало болезни.

Недавний прогресс в высокопроизводительных технологиях привел к новому сценарию исследования патологий, названному «эрой Омики», который объединяет возможность сбора больших объемов данных и информации на молекулярном и белковом уровнях вместе с разработкой новых вычислительных и компью-

терных технологий, статистических инструментов, которые могут анализировать и фильтровать такие данные. Если геном изучает наука геномика, то по аналогии стали возникать другие «омики» – науки о совокупности биологических данных (протеомика, эпигеномика) [3].

Достижения в области массивов гено-типирования, секвенирования следующего поколения, технологии масс-спектрометрии и биоинформатики позволили одновременно провести крупномасштабное исследование тысяч генов (геномика), факторов эпигенетики (эпигеномика), РНК (транскриптомика), метаболитов (метаболомика) и белков (протеомика) с возможностью интеграции нескольких типов омических данных («мультикомика»). Все эти технологические инновации изменили подход к изучению сложных заболеваний, таких как болезнь Альцгеймера, тем самым представляя собой многообещающий инструмент для исследования взаимосвязи между несколькими молекулярными путями болезни Альцгеймера, а также с другими патологиями.

Было обнаружено три мутации одного гена, связанные с ранним началом болезни Альцгеймера: белок-предшественник амилоида (APP), пресенилин 1 (PSEN1), пресенилин 2 (PSEN2). Мутации в этих генах приводят к выработке аномальных белков, связанных с заболеванием. Каждая из этих мутаций играет роль в распаде APP, белка, точная функция которого еще полностью не изучена. Этот распад является частью процесса, в результате которого образуются вредные формы амилоидных бляшек, что является признаком болезни Альцгеймера.

Немаловажным аспектом является то, что разработки в области биоинформатики и, в частности, массового параллельного секвенирования для анализа ДНК позволили извлекать бесценную информацию без высоких затрат. Некоторые области ДНК, анализ которых раньше считался трудоемким, теперь можно легко изучать, что дает дополнительную информацию, которая в противном случае была бы потеряна [4].

Профессия биоинформатик сложна, но и необычайно увлекательна. Особенностью профессии является то, что успешный специалист в этой области должен знать много вещей, которые традиционно не сгруппированы вместе. Эти темы варьируются от биологии и статистического анализа до программирования и машинного обучения, а также включают следующее:

- хранение и управление большими наборами данных;
- кластерный анализ;
- анализ секвенирования РНК, ДНК и белков;
- разработка алгоритмов;
- знакомство с Linux, особенно для управления кластерами серверов;
- различные языки программирования, включая Python, R и Perl;

- надзор за системами управления лабораторной информацией;
- создание визуализаций данных для использования в отчетах;
- анализ изображений;
- разработка процессов мониторинга.

В дополнение к этим техническим навыкам, критическое мышление и адаптируемость, также являются ведущими – необходимо знать, как подойти к проблеме и разложить ее на подкомпоненты, а затем соединить их вместе, чтобы получить ответ.

Сегодня в специальных виртуальных банках в открытом доступе находится достаточное количество «сырых» биологических данных, поэтому молодым ученым есть на чем ставить эксперименты и делать свои научные открытия.

Библиографический список

1. Бородин А.Г., Манойлов В.В. Поколения методов секвенирования ДНК. [Электронный ресурс] // Режим доступа: <https://elibrary.ru/item.asp?id=44238826> (дата обращения: 24.06.2021)
2. E. van Niekerk Countering revisionism—part 1: Ernst Haeckel, fraud is proven. [Электронный ресурс] // Режим доступа: <https://creation.com/haeckel-fraud-proven> (дата обращения: 24.06.2021)
3. Свирдлов Е.Д. Взгляд на жизнь через окно генома. Т. 2: Очерки современной молекулярной генетики. М.: «Наука», 2019. – 495 с.
4. Mohit Mazumder Finding Omics Data to Study Alzheimer's Disease [Электронный ресурс] // Режим доступа: <https://edu.t-bio.info/blog/2021/01/15/finding-omics-data-to-study-alzheimers-disease/> (дата обращения: 24.06.2021)

PROSPECTS OF BIOINFORMATICS

S.A. Gryaznov, *Candidate of Pedagogical Sciences, Associate Professor*
Samara Law Institute of the Federal Penitentiary Service of Russia
 (Russia, Samara)

Abstract. *Bioinformatics is one of the most dynamic sciences in our time. Stimulating factors for the growth of the global bioinformatics market include the growing demand for nucleic acids and protein sequencing, the growth of initiatives from public and private organizations, the acceleration of the growth of proteomics and genomics, as well as increased research in the field of molecular biology and drug discovery. This article discusses the prospects for the treatment of Alzheimer's disease in connection with the growing development of bioinformatics, as well as the skills that a modern specialist in this field should have.*

Keywords: *bioinformatics, big data, genomic data, genome sequencing, DNA analysis, omics.*