Минобрнауки России

Федеральное государственное бюджетное образовательное учреждение

высшего образования

«САРАТОВСКИЙ национальный исследовательский ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра информатики и программирования

Реферат

БИОИНФОРМАТИКА

студентки 1 курса 151 группы  
направления 09.03.04 Программная инженерия

факультета компьютерных наук и информационных технологий

Анофриковой Анны Сергеевны

Саратов 2024

**СОДЕРЖАНИЕ**

ВВЕДЕНИЕ………………………………………………………………………….. 3

1 История возникновения биоинформатики………………………………………. 4

2 Цели и методы биоинформатики...………………………………..……………... 4

3 Основные направления биоинформатики………………….……………………. 5

4 Используемые в биоинформатике информационные технологии……...……… 6

5 Применение биоинформатики…………………………………..………………... 7

6 Последние открытия в области биоинформатики………………………….…… 8

7 Профессия биоинформатик…………………...………………………………….. 8

ЗАКЛЮЧЕНИЕ…………………………………………………………………….. 10

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ……………………………… 11

**ВВЕДЕНИЕ**

Биоинформатика – одна из перспективно развивающихся в настоящее время наук, объединяющая в себе несколько дисциплин (биологию, информатику, математику, статистику) и изучающая применение информационных технологий для управления биологическими данными и процессами.

Возникновение биоинформатики связано с началом изучения людьми генома – совокупности наследственного материала, заключенного в клетке любого живого организма. Геном человека состоит из более чем трех миллиардов знаков и представляет собой огромный массив информации. Кроме того, объем исследуемых биологический данных ежегодно значительно увеличивается, следовательно, без использования современных информационных технологий проанализировать их невозможно.

Биоинформатика позволяет обрабатывать большие объемы генетических данных и получать на их основе конкретные научные результаты. Возможность владеть наследственной информацией открывает для человечества новые возможности, поэтому эта наука играет важную роль в развитии генетики, медицины, сельского хозяйства и других областей науки.

**История возникновения биоинформатики**

Термин «биоинформатика» был введен как изучение информационных процессов в биотических системах в 1970 году Полиной Хогевег.

Современная биоинформатика возникла во второй половине 1970-х годов в связи с появлением эффективных методов секвенирования генома (расшифровки нуклеотидных последовательностей ДНК, в которых хранится информация о живых существах). Важным этапом в развитии биоинформатики стал проект по секвенированию генома человека. Первые несколько сотен расшифрованных последовательностей белков были опубликованы в виде книги-атласа [1].

Однако вскоре их число возросло настолько, что стала очевидна необходимость специальных методов информационного анализа для обработки и сравнения последовательностей. В связи с этим интенсивное развитие данной науки происходило одновременно с распространением компьютерных технологий. Другим значимым шагом в становлении биоинформатики стало возникновение и развитие Интернета, так как в связи этим событием исчезла необходимость независимой разработки программных продуктов в каждой исследовательской лаборатории [2].

Временем выделения биоинформатики в отдельную научную область считается 1980 год, когда началось издание журнала «Nucleic Acids Research», посвященного компьютерным методам анализа биологических данных. Также в 1980-х годах начали разрабатывать первые базы данных для хранения генетической информации. В 1990-х годах начались масштабные проекты по секвенированию геномом различных организмов. Подобные проекты требовали использования информационных технологий для обработки и анализа огромных объемов генетической информации [3].

Таким образом, информационные технологии стали неотъемлемой частью биоинформатики и сыграли решающую роль в ее развитии.

**Цели и методы биоинформатики**

Биоинформатика – наука о хранении, извлечении, организации, анализе, интерпретации и использовании биологической информации.

Целями биоинформатики являются:

* организация и сохранение биологических данных для обеспечения доступа к ним исследователей и возможности внесения новых сведений;
* разработка программных средств и создание специализированных информационных ресурсов для управления данными и их анализа;
* автоматизация анализа биологических данных, их интерпретация и использование для получения биологического смысла [4].

Методы и инструменты биоинформатики включают в себя большое количество математических и компьютерных способов, алгоритмов и информационных ресурсов. Математические и компьютерные методы, применяемые в данной сфере, включают в себя линейную алгебру, теорию графов, статистику и машинное обучение и используются для обработки данных, полученных в процессе экспериментов и исследований. С помощью алгоритмов, используемых в рассматриваемой науке, можно изучать значительные объемы информации с высокой скоростью и точностью [5].

Методы биоинформатики дают возможность не только обрабатывать большие массивы данных, но и выявлять в них закономерности, которые нельзя заметить при проведении обычного эксперимента. Это позволяет предсказывать функции генов и зашифрованных в них белков, строить модели взаимодействия генов в клетке.

**Основные направления биоинформатики**

В настоящее время биоинформатику разделяют на две части: структурную и последовательную.

Первая из них занимается анализом пространственной структуры белковых молекул. В данном случае с помощью компьютерных технологий проводятся сложные расчеты по вычислению пространственного расположения аминокислот, из которых состоят белковые молекулы. Компьютерные модели молекул белков позволяют предсказывать результат их взаимодействия, например, с молекулой лекарства. Знания, полученные в результате таких исследований, дают возможность создания и контроля качества создаваемых искусственных органический веществ (лекарств, новых материалов) [6].

Биоинформатика последовательностей работает с более высоким уровнем организации живой материи: с нуклеотидными и белковыми последовательностями. С помощью различных технологий специалисты собирают в банки данных экспериментальную информацию, осуществляют ее компьютерный анализ, ищут закономерности и предсказывают новые свойства живых систем. Например, последовательная биоинформатика может объяснить, с какими мутациями (изменениями генома) связаны различные заболевания. Кроме медицинских данных эта часть рассматриваемой науки изучает роли и функции конкретных генов, закономерности распространения организмов по земле, популяционные различия между группами животных [7].

**Используемые в биоинформатике информационные технологии**

Данные, полученные в ходе экспериментов, для биоинформатиков представляются в виде последовательностей символов, массивов вещественных чисел, матриц из нулей и единиц. Статистический анализ информации является неотъемлемой частью научного исследования.

Наиболее часто используемыми инструментами и технологиями в этой области являются:

* языки программирования (Java, R, C, C#, C++, Python);
* статистические пакеты – программные продукты, предназначенные для статистической обработки данных (SPSS, SAS, MatLab) [1].

Базы данных, используемые для хранения большего числа биоинформационных данных, делятся на:

* таксономические, которые хранят информацию о биологических таксонах – группах объектов, объединяемых на основании общих свойств и признаков;
* нуклеотидные, которые хранят данные о расшифрованных нуклеотидных последовательностях ДНК (например, GenBank);
* белковые, которые хранят информацию о аминокислотных последовательностях (например, PIR);
* пространственных структур макромолекул, которые хранят данные о пространственной организации белков и ДНК (например, PBD) [8].

Одной из наиболее распространенных баз биологических данных является GenBank, содержащая нуклеотидные последовательности ДНК и РНК. Обмен данными там осуществляется ежедневно. GenBank поддерживается NCBI (National Center for Biotechnological Information), который занимается разработкой новых информационных технологий для исследования молекулярно-генетических процессов. Эта организация предоставляет и другие ресурсы, например, поисковую систему по биомедицинским исследованиям PubMed и программу для сравнительного анализа последовательностей BLAST [9].

**Применение биоинформатики**

Методы биоинформатики широко применяются в различных областях биологии и медицины для анализа и использования биологических данных. Например, в геномике информационные технологии используют для секвенирования генома и анализа полученных результатов, а фармакологии – для исследования взаимодействия лекарственных препаратов с мишенями в организмах живых существ.

Одной из наиболее перспективных и быстро развивающихся областей биоинформатики является конструирование лекарств направленного действия. Действие таких препаратов нацелено на конкретный белок в организме возбудителя болезни, при этом белки пациента не подвергаются изменениям, что означает отсутствие побочных эффектов [10].

Благодаря информационным технологиям можно получить ответ на ряд вопросов, касающихся эволюционных, биохимических и биофизических характеристик исследуемых молекул.

В медицине методы биоинформатики позволяют выяснять причины болезней и создавать новые подходы к лечению заболеваний, обрабатывая большие объемы медицинской информации и обнаруживая различные закономерности [2].

В сельском хозяйстве биоинформатика используется, например, для создания пестицидов, исследования свойств растений, выбора лучших организмов для скрещивания, что способствует повышению эффективности производства сельхозпродукции.

**Последние открытия в области биоинформатики**

За последние годы в области биоинформатики произошло немало важных открытий, которые могут в будущем значительно повлиять на развитие как самой науки, так и других сфер человеческой деятельности.

Например, в 2022 году российские, американские и датские биоинформатики разработали новую методику анализа геномных данных, которая точно определяет время и количество волн миграций людей и раскрывает историю перемешивания популяций.

А в 2023 году российский ученый сделал открытие, которое изменило современные представления о ДНК. Более 70 лет считалось, что ДНК хранит и обрабатывает информацию за счёт структуры двойной спирали. Максим Никитин же доказал, что для эффективной обработки генетической информации ДНК совершенно не обязательно образовывать двойную спираль.

В этом же году ученые из Швеции разработали набор инструментов, которые смогут улучшить изучение групп крови человека и способствовать дальнейшим изучениям роли групп крови в развитии различных заболеваний.

**Профессия биоинформатик**

Существуют различные специалисты в области биоинформатики: биоинформатик-биохимик, биоинформатик-биофизик, биоинформатик-медик, биоинформатик-эколог, биоинформатик-фармаколог, биоинформатик-генетик, биоинформатик в области сельского хозяйства. Названия специализаций отображают сферу деятельности человека, занятого в той или иной из них. Например, биоинформатики-медики занимаются вопросами наследственных заболеваний, а биоинформатики-биохимики изучают химические процессы в клетках [5].

От специалистов – программистов в данной сфере часто требуется:

* знание языков программирования (Python, R);
* умение работы с базами данных;
* знание основ биоинформатики, биологии и генетики;
* знания в области машинного обучения.

При этом существуют вакансии, требующие глубоких знаний в сфере информационных технологий и владения лишь основами биологии, и наоборот. Следовательно, войти в область биоинформатики можно как с биологической базой, так и со знанием программирования и математики. Для математиков и информатиков это возможность использовать знания информационных технологий в биологии и медицине.

**ЗАКЛЮЧЕНИЕ**

Таким образом, биоинформатика является междисциплинарной научной областью, которая сочетает в себе знания сразу нескольких наук.

Сегодня человечество находится на начальном этапе использования генетической информации о живой материи, но развитие все более эффективных методов биоинформатики способствует лучшему пониманию строения и механизмов функционирования живых систем. Открытия, сделанные в этой довольно молодой науке, уже привели ко многим прорывам в изучении генома, создании новых методов лечения, предотвращении наследственных заболеваний и разработке новейших методов диагностики и лечения.

В настоящее время биоинформатика продолжает активно развиваться, а информационные технологии играют ключевую роль в улучшении человеческих знаний о биологических процессах.

**СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ**

1. Романова В.В., Ивина О.А. Биоинформатика и средства компьютерного анализа // Инновации в пищевой биотехнологии: сборник трудов Международного симпозиума / под. общ. ред. А. Ю. Просекова; ФГБОУ ВО «Кемеровский государственный университет». – Кемерово, 2018. – с. 354 – 357.
2. Основы биоинформатики: учеб. пособие / А.Н. Огурцов. – Харьков: НТУ «ХПИ», 2013. – 400 с.
3. Бурнашев Р.Ф. Роль информационных технологий в развитии современной биоинформатики [Электронный ресурс] // Universum: химия и биология: электрон. научн. журн. 2023. – 8(110). – с. 31 – 36. – URL: <https://7universum.com/ru/nature/archive/item/15785> (дата обращения: 07.05.2024).
4. Биоинформатика: курс лекций / В.В. Янковская. – Витебск: ВГУ им. П.М. Машерова, 2022. – 83 с.
5. Назаренко М.Д., Субботина В.В., Воробьев И.Ю., Сафоновa Т.В., Мокряк А.В. Биоинформатика: исследование жизни с помощью алгоритмов и данных // Международный журнал информационных технологий и энергоэффективности – Смоленск: Изд. центр «Открытая наука», 2024. – Т. 9. – № 2. – с. 79 – 84
6. Еремина Д.В., Мартемьянова Д.Д. Биоинформатика: современное состояние и перспективы // Научный журнал «Эпоха науки» – Ачинск: Издательство Красноярского государственного аграрного университета, 2024. – № 37. – с. 12 – 16.
7. Колесников А.А., Селюкова Г.П. Биоинформатика // LV Студенческая научно-практическая конференция «Актуальные вопросы науки и хозяйства: новые вызовы и решения» – Тюмень: Издательство Государственного аграрного университета Северного Зауралья, 2021. – с. 471 – 474
8. Биоинформатика: учебно-методическое пособие. / Н.Ю. Часовских – Томск: Издательство СибГМУ, 2015. – 109 с.
9. Биоинформатика: учебник / Н.Ю. Часовских – Москва: Изд. Группа «ГЭОТАР-Медия», 2020. – 344 с.
10. Межгорин Д.С., Петкилев А.А., Митрохин М.А., Гурин Е.И. Информационные технологии в биоинформатике // Международный симпозиум «Надежность и качество» – Пенза: Издательство ПГУ, 2020. – с. 287– 288.