Департамент образования города Москвы

Государственное бюджетное общеобразовательное учреждение города Москвы «Школа №1505

«Преображенская»»

**РЕФЕРАТ**

на тему

**Расшифровка ДНК**

Выполнила :
Пашинцева Анастасия Валерьевна

Руководитель:
Шалимова Елена Георгиевна

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (подпись руководителя)

Рецензент:
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (подпись рецензента)

Москва

2017/2018 уч.г.

|  |  |  |
| --- | --- | --- |
|  | ОГЛАВЛЕНИЕ |  |
|  | Введение………………………………………………. | 3 |
| 1 | 1 Методы секвенирования старого поколения…………………………………………….. | 5 |
|  | 1.1 ДНК-проба, или ядерная проба…………………………………………………... | 5 |
|  | 1.2 Метод Сэнджера………………………………….. | 8 |
|  | 1.3 МетодМаксама**-**Гилберта………………………... | 9 |
|  | 1.4 Словарь терминов первой главы………………… | 10 |
| 2 | 2 Проект «Геном человека»………………………… | 12 |
|  | 2.1 Типы карт…………………………………………. | 13 |
|  | 2.2 Два подхода к картированию геномов………….. | 15 |
|  | 2.3Банкиданных……………………………………... | 16 |
|  | 2.4 Краткая характеристика результатов……………. | 16 |
|  | 2.5 Будущее проекта………………………………….. | 18 |
|  | 2.6 Словарь терминов второй главы………………… | 18 |
| 3 | 3 Новые методы секвенирования и применение знаний расшифрованного генома человека на практике………………………………………… | 19 |
|  | 3.1 Новые методы секвенирования. Секвенирование «нового поколения»…………………………… | 19 |
|  | 3.1.1 Пиросеквенирование | 20 |
|  | 3.2 Практическое применение знаний расшифрованного генома человека………… | 21 |
|  | 3.3 Словарь терминов третьей главы………………... | 26 |
| 4 | Заключение…………………………………….…. | 28 |
| 5 | Список литературных источников……………………………………….... | 29 |

Введение.

Все организмы состоят из белков, липидов, углеводов и нуклеиновых кислот. Липиды выполняют энергетическую функцию, углеводы – запасную, белки – строительную, а нуклеиновые кислоты отвечают за хранение и передачу наследственной информации, ее реализацию. Иными словами, строение организма определяется белками. Информация о структуре белков зашифрована в молекулах ДНК. ДНК представляет собой две спирально закрученные цепей. Остов цепи образован остатками фосфорной кислоты и дезоксирибозы. Внутрь спирали направлены азотистые основания, они соединяются между собой с помощью водородных связей по принципу комплементарности. Участок молекулы ДНК, содержащий информацию о структуре одной молекулы белка-фермента, называется геном. Совокупность наследственного материала организма, называется геномом. В белках встречаются 20 аминокислот, их последовательность определяет структуру и свойства белков. Правила перевода нуклеотидной последовательности в нуклеиновой кислоте в аминокислотную последовательность белка называют генетическим кодом. Он был расшифрован в 60-х годах XX века в результате ряда экспериментов и математических расчетов. Если прочитать этот код, можно понять индивидуальные особенности организма. Если прочитать его у разных организмов, можно понять насколько близки разные этнические группы, как люди расселялись по планете, причины наследственных заболеваний. Первая расшифровка человеческого генома продолжалась в течении 10 лет, в 2003 году эта задача была решена. С совершенствованием технологии определения порядка нуклеотидов, то есть секвенирования, этот процесс пошел быстрее.

**Актуальность моего исследования:** С генетической точки зрения все люди одинаковы более чем на 99%, разница в 1% делает людей уникальными. Этим числом определяются наши физические возможности, предрасположенность к болезням и реакции на лекарства и т.д. Из-за этой разницы одно лекарство подходит одним пациентам, а другим – нет. Поэтому, в будущем будет развиваться персональная медицина, основанная на данных генома пациента. Однако, чтобы узнать, как работает весь организм в целом нужно смотреть протеом – совокупность всех белков организма, ткани или органа. С помощью знаний, полученных от расшифровки ДНК можно понять насколько близки разные этнические группы, как люди расселялись по планете, причины наследственных заболеваний.

**Проблема моего исследования:** Раньше секвенирование ДНК было долгим и трудоемким процессом, сейчас технология усовершенствовалась, цена упала, но она все еще остается высокой, а процесс – трудозатратным. Ученые стремятся сделать медицину более персонализированной, для чего им нужны данные, полученные от секвенирования ДНК.

**Цель моего исследования:** познакомиться с методами расшифровки ДНК и применением полученных данных на практике.

**Задачи моего исследования:**

* изучить методы расшифровки ДНК (современные и новейшие);
* узнать о свойствах и применении полученных в результате секвенирования данных.

Глава 1.

Секвенирование – это общее название методов, которые позволяют установить последовательность нуклеотидов в молекуле ДНК. Они все работают не с целой ДНК, а с ее фрагментами. Создают много копий одной ДНК или ее определенного фрагмента с помощью полимеразной цепной реакции (ПЦР). Метод ПЦР состоит из трех этапов. На первом этапе образец ДНК нагревают до 90-96 °C. Молекула ДНК денатурирует1, вследствие чего цепи раскручиваются и расходятся. Отдельные цепи обрабатываются раздельно. На втором этапе праймеры – короткие фрагменты комплементарных оснований, присоединяются по концам одноцепочечных ДНК. В ходе третьего этапа фермент ДНК-полимераза2 считывает матричную цепь и достраивает к ней комплементарную из нуклеотидов. Для этого метода нужна полимераза, которая устойчива к высоким температурам и может функционировать в условиях быстроизменяющейся температурной среды. Эту полимеразу выделяют из термофильных бактерий, живущих в горячих источниках – Thermus aquaticus. Используемую полимеразу называют Taq. Использование этого метода позволяет получить бесчисленное количество копий[2].

Существует два основных метода секвенирования: ДНК-проба и метод Сэнджера. Существуют, конечно, и другие методы.

**ДНК-проба, или ядерная проба.**
Метод состоит из двух стадий. Метод основан на стремлении одноцепочечной ДНК соединиться с цепочкой ДНК, имеющей последовательность комплементарную последовательности исходной цепи. Первая стадия - стадия гибридизации представляет собой простое смешивание одноцепочечных проб с целевой ДНК. В результате первой стадии образуется «гибридная» цепь ДНК. Она состоит из «натуральной» ДНК и «искусственной» ДНК-пробы. На второй стадии ДНК денатурируют. Двойная цепь разъединяется на цепочки, одноцепочечные пробы находят комплементарные участки и соединяются с ними.

Этот метод часто совмещают с методом анализа фрагментов ДНК, который называется Саузерн-блоттинг. Метод назван в честь его разработчика – Эдварда Саузерна[2]. Для этого метода нужно сначала разделить исследуемый участок ДНК на небольшие фрагменты. Затем фрагменты ДНК подвергают гель-электрофорезу. Электрофорез – это аналитический метод разделения молекул, основанный на различиях в их размере и/или электрическом заряде. Соединения помещают в гель и пропускают через него электрический ток. Под действием тока соединения перемещаются в геле. Скорость, с которой они двигаются, зависит от их размера и величины заряда. Соединения реагируют на электрическое поле в зависимости от того, насколько сильно они заряжены. Зависимость скорости передвижения соединений от их размера проявляется в следующем: небольшим фрагментам легче передвигаться в геле, чем большим[2].

 Различные фрагменты ДНК помещаются на одном конце геля в специально предназначенные для этого карманы. Для каждого фрагмента – свой карман. Затем включают электрический ток. Но для того чтобы увидеть результат электрофореза, надо поместить гель в этидиум бромид (EtBr) – флоуресцентный краситель, связывающийся с ДНК. Если посмотреть на гель в ультрафиолетовом свете с определенной длиной волны, можно увидеть фрагменты ДНК и их передвижение.



Для того чтобы определить фрагмент, связавшийся с пробой, используют способ обнаружения нужных молекул – система соединенных антител и флуорохромов4. Этот метод получил название флуоресцентной in situ гибридизации. Он позволяет синтезировать пробы со встроенными флуоресцентными молекулами, определяющимися с помощью флуоресцентных антител. Для дальнейшей работы гель высушивается на системе фильтров и у образцов получается твердая основа. Если фрагменты ДНК пометили с помощью радиоактивных атомов, например, трития, то на основу кладут фоточувствительную пленку. Тогда элемент проявит пленку так же, как при взаимодействии света на нее. Если же фрагмент пометили флуоресцентной меткой, то необходимо подвергать основу воздействию ультрафиолетовым светом определенной длины волны. В таком случае понадобится метка, требующая для своего выявления иную длину волны, чем этидиум бромид[2]. Таким образом, можно узнать о локализации специфичных последовательностей ДНК благодаря использованию определенных рестриктаз3, а также о локализации последовательности оснований, комплементарной пробе во фрагменте ДНК.

**Метод Сэнджера, или метод обрыва цепи.**

Его открыл английский биофизик Фредерик Сэнджер (1918-2013гг) – единственный ученый, получивший сразу две Нобелевские премии по химии. Первую премию в 1958 году ему присудили за установление структур белков, в особенности инсулина. Вторую премию в 1980 году он получил за разработку методов определения первичной последовательности нуклеиновых кислот. Эту методику Сэнджер и его коллеги предложили в 1977 году, с тех пор она прошла несколько модификаций, но до сих пор используется[4]. У метода есть второе название – «метод обрыва цепи». Второе название объясняется тем, что метод основан на остановке синтеза новой цепи ДНК[2]. Он зависит от следующих факторов:

1. Синтез фрагмента двухцепочечной ДНК из одноцепочечной происходит при наличии ДНК-полимеразы;
2. Синтез ДНК останавливается, если включенное в цепь основание находится в форме дидезоксинулеотида вместо дезоксинуклеотида. В дидезокси-форме у нуклеотида отсутствует гидроксильная группа в определяющей позиции (3’ – позиция).



В реакционную смесь добавляют дидезоксинуклеозидтрифосфаты (аналоги оснований). Включение дидезоксинуклеозидтрифосфатов в синтезируемую цепь приводит к тому, что цепь не может дальше синтезироваться. Образуется «обрывок» цепи, по которому можно установить последнюю букву секвенируемого фрагмента ДНК[2].

Метод Сэнджера позволяет считывать последовательности до 1000 пар нуклеотидов и применяется для секвенирования небольших фрагментов генома или генов. Например, он используется для:

* секвенирования отдельных участков генома в целях анализа мутаций и полиморфизмов;
* идентификации вирусов и организмов;
* перестроек (потерь или вставок участков хромосомы, в следствии различных процессов) хромосом
* и т.д. [4]

**Метод Максама-Гилберта.**

Развивались другие методы секвенирования, в частности, метод химической деградации Максама и Гилберта, он тоже используется в настоящее время и иногда незаменим. Участки с сильной вторичной структурой не всегда бывает возможно секвенировать методом Сэнджера и тогда используют данный метод. В основе этого метода лежит ограниченное расщепление меченого фрагмента ДНК под действием специфических реагентов. Для проведения этого метода необходимо наличие фрагмента ДНК, меченного только по одному концу[5]. Разделение продуктов деградации по размеру проводят также с помощью электрофореза. Таким образом, можно разделять фрагменты ДНК, различающиеся между собой по длине всего на один нуклеотид в достаточно широком диапазоне. После разделения также применяют радиоавтографию для определения нуклеотидной последовательности.

Реакции проводятся в несколько этапов. На первом этапе проводится ограниченная модификация определенных нуклеотидов под действием различных веществ. Концентрация вещества и продолжительность его воздействия на молекулы ДНК подбирается с таким расчетом, чтобы в каждой молекуле произошла модификация только одного нуклеотида. Тогда, по теории вероятности все основания данного типа в секвенируемом фрагменте ДНК окажутся модифицированными. На следующих этапах происходит удаление модифицированных оснований, разрыв цепи. Для каждого типа нуклеотидов или их комбинации проводят отдельные реакции ограниченной модификации и количественного расщепления. Таким образом, в результате нескольких реакций образуется смесь из олигонуклеотидных5 молекул, которые различаются по размеру на один нуклеотид и несут на одном из концов метку. Однако кроме меченых молекул в реакционной смеси будут и олигонуклеотидные фрагменты без меток, поэтому на этапе радиоавтографии они окажутся невидимыми и «не будут существовать». После разделения продуктов реакции и этапа радиоавтографии на рентгеновской пленке будет видна лестница из полос ДНК на соседних дорожках, "чтение" которой позволяет восстановить последовательность нуклеотидов секвенируемого фрагмента ДНК[5].

Словарь определений:

1- Денатурация - процесс нарушения природной структуры белков, за счет разрушения водородных связей.

2-ДHK-полимераза - фермент, катализирующий процесс синтеза полинуклеотидной цепи ДНК из отдельных нуклеотидов при использовании другой цепи в качестве матрицы.

3-Рестрикция - процесс расщепления чужеродной молекулы ДНК под действием специфических бактериальных ферментов – рестриктаз. (см рис.1.1)

4-Флуорохром - краситель, обладающий люминисцентными свойствами и используемый при флуоресцентной микроскопии.

5- Олигонуклеотид - олигомерная форма нуклеиновой кислоты, содержащая относительно небольшое количество нуклеотидов (до 20).

Глава 2.

В 1990 году министерство энергетики США, а также страны - Великобритания, Франция, Япония, Китай и Германия – поддержали запуск международного проекта «Геном человека». Проект возглавил доктор Фрэнсис Коллинз.
В 1998 году аналогичный проект запустил доктор Крейг Вентер1 и его фирма «Celera Genomics» с целью быстрее и более дешево секвенировать геном человека. Бюджет международного проекта составлял более 3 миллиардов долларов, а бюджет проекта от «Celera Genomics» всего 300 миллионов. Кроме того, в отличии от международного проекта, доктор Крейг Вентер и «Celera Genomics» не собирались открывать доступ к полученным данным своего исследования[3].

Цель проекта заключалась в выяснении последовательности оснований во всех молекулах ДНК человека и установлении локализации всех генов. Это помогло бы выяснить причины наследственных заболеваний и открыть способы их лечения. Кроме этой цели были сформулированы другие:

* идентифицировать 20 000–25 000 генов ДНК;
* завершить составление детальной генетической карты, на которой были бы помечены гены, отстоящие друг от друга на расстоянии, не превышающем в среднем 2 млн оснований;
* составить физические карты каждой хромосомы (разрешение 0,1 млн оснований);
* завершить к 2004 году полное секвенирование ДНК (разрешение 1 основание);
* нанести на полностью завершенную секвенсовую карту все гены человека (к 2005 году);
* усовершенствовать приборы для анализа данных и внедрить новейшие технологии в область частного использования;
* исследовать этические, правовые и социальные вопросов, возникающие при расшифровке генома[3].

Ожидалось, что при достижении поставленных целей исследователи определят все функции генов и разработают методы применения полученных данных.

**Типы карт.**

В ходе работы над проектом последовательно создавались три типа карт хромосом: генетические, физические и секвенсовые.

Локализовать каждый ген в хромосомах позволяют выявление всех генов и установление примерного расстояния между ними. Генетические карты перечисляют гены, входящие в геном, и указывают места их расположения, отвечают на вопрос о вовлеченности генов в образование отдельных признаков организма. Ведь многие признаки формируются под контролем нескольких генов, часто расположенных в разных хромосомах, и знание локализации каждого из них способствует лучшему распознаванию законов дифференцировки клеток, органов и тканей, а также лучшему лечению болезней[3].

Физическую карту хромосомы можно увидеть на рисунке 2.2. В 60-е годы цитогенетики использовали методы окрашивания хромосом для выявления бэндов - поперечных полосок на хромосомах, которые можно заметить в микроскоп. Были разработаны методы, позволившие физически следить за присоединением коротких отрезков радиоактивно-меченых или флуоресцентно-меченых ДНК к хромосомной ДНК.

Локализация флуоресцентных или радиоактивных меток повысила разрешение карт. Использование метода флуоресцентной in situ гибридизации дало возможность достичь разрешения от 2 до 5 миллионов оснований, а потом повысить его (при изучении хромосом делящихся клеток) до 100 тысяч оснований. Используя рестриктазы построили рестрикционные физические карты[3].

Разработка методов секвенирования привела к созданию секвенсовых карт, на которых указано положение всех нуклеотидов в ДНК.

**Два подхода к картированию геномов.**

Поскольку у всех видов число хромосом и их длина различны и для разных видов нельзя применить одни и те же методы картирования, были предложены два различных подхода, которые условно назвали «сверху вниз» и «снизу вверх». Первый метод использовали для небольших последовательностей, а второй для длинных геномов. При использовании первого метода шли «сверху вниз» - разрезали геномную ДНК на небольшое число кусков и анализировали их по отдельности, а затем воссоздавали всю начальную структуру. При использовании второго метода шли «снизу вверх» - сначала секвенировали более длинные, произвольно выбранные куски, затем прикладывали их друг к другу, чтобы найти перекрывающиеся концы. Когда находили перекрытие, куски объединяли в один, так называемый контиг. С применением компьютерных и математических методов искали перекрытие все более крупных контигов и постепенно шли вверх выше[3].

Важная составляющая проекта – это применение новых методов исследований. Развитые еще до начала выполнения проекта методы или методы первого поколения включали применение рестрикционных ферментов; создание гибридных молекул, их клонирование и перенос участков ДНК с помощью векторов в клетки-доноры; синтез ДНК на матрицах информационной РНК (сДНК-синтез); методы секвенирования генов; получение практически неограниченного количества копий генов с помощью полимеразной цепной реакции; методы, предназначенные для разделения молекул ДНК по плотности, массе, различной вторичной структуре.

Благодаря проекту "Геном человека", были развиты новые методы или методы второго поколения, в которых некоторые процессы автоматизированы.

**Банки данных.**

Результаты работы находятся в общем доступе. Были созданы международные банки данных о последовательностях нуклеотидов в ДНК разных организмов (примеры - GenBank, EMBL и DDBJ) и о последовательностях аминокислот в белках (примеры - PIR и SwissPot)[3].

**Краткая характеристика результатов.**

Было решено, что сначала надо отработать методы на более простых моделях. Сначала изучили 8 разных представителей мира микроорганизмов, затем список расширяли. Изучили представителей многих родов бактерий: архебактерии, хламидобактерии, кишечная палочка, возбудители пневмоний, сифилиса, гемофилии, метанобразующие бактерии, микоплазмы, риккетсии, цианобактерии, первого эукариотического одноклеточного организма - дрожжей Saccharomyces cerevisae и первого многоклеточного животного организма - нематоды C. elegans.
В 1995 году был полностью картирован геном бактерии Hemophilus influenzae, в 1996 году было закончено картирование ДНК дрожжевой клетки (12,5 миллионов оснований, 6 тыс. генов), в середине декабря 1998 года был полностью картирован геном круглого червя Caenorhabditis elegans (97 миллионов оснований и 19 099 генов, что составляет от 1/3 до 1/4 общего числа генов человека).

С января 1995 до января 1996 года длина участков ДНК человека с установленной полной последовательностью оснований увеличилась почти в 10 раз. К июлю 1998 года было секвенировано почти 9% всего генома. К 23 октября 1998 года были установлены последовательности 30 181 гена человека. К 11 ноября того же года число секвенированных генов достигло 30 261, примерно половины всех генов человека.

6 июня 2000 г. президент США и премьер-министр Великобритании объявили о расшифровке человеческого генома. Был опубликован рабочий черновик человеческого генома, и лишь к 2003 г. он был расшифрован практически полностью, хотя и сегодня все еще проводят дополнительный анализ некоторых участков генома.

Была получена информация о вовлеченности генов в образование и функционирование отдельных органов и тканей человеческого тела. Например, оказалось, что самое большое число генов необходимо для формирования мозга и поддержания его активности, а самое маленькое для создания эритроцитов - всего 8 генов.

В 2003г. ученые обнаружили, что существует меньше генов, чем они ожидали, но впоследствии убедились в противоположном. Традиционно ген определяли, как участок ДНК, который кодирует белок. Однако, расшифровывая геном, ученые выяснили, что 98,5% участков ДНК не кодируют белки, а отвечают за ее функционирование. Таким образом, если данные участки ДНК также считать генами, то их количество удвоится. В итоге исследования изменилось представление о генах, и сейчас ученые считают, что ген — это единица наследственности, которую нельзя понимать как просто участок ДНК, кодирующий белки.

**Будущее проекта.**

Главная будущая задача сформулирована следующим образом: изучить однонуклеотидные вариации ДНК в разных органах и клетках отдельных индивидуумов и выявить различия между индивидуумами. За этой задачей стоят и положительные, и отрицательные последствия. С одной стороны, это даст возможность персонализированной медицине, с другой – такие сведения можно использовать против индивидуума.

Словарь терминов:

1 - Марк Адамс - ведущий сотрудник Института геномных исследований в штате Мэриленд (США), частной исследовательской компании, занимающейся исключительно работами в области картирования генома человека, Крэйг Вентер - директор этого института.

Глава 3.

**Новые методы секвенирования. Секвенирование «нового поколения».**

Технология методов секвенирования «нового поколения» позволяет «прочитать» единовременно сразу несколько участков генома. Появление этих методов позволило ускорить и удешевить определение полной последовательности миллионов геномов организмов. Появилась возможность единовременно оценивать работу тысяч генов в организмах, тканях и клетках (секвенирование транскриптомов1) и анализировать регуляцию их активности. Сейчас на рынке представлены секвенаторы2, позволяющих определять последовательность полных геномов организмов, проводить анализ экспрессии3 генов и многое другое[4]. Снижается себестоимость анализа генома[6]. Появляются высокопроизводительные технологии секвенирования, и одновременно происходит прогресс программного обеспечения – создаются алгоритмы с открытым программным кодом, появляются открытые источники данных и платформы для вычислений. В геномике4 используются более сложные алгоритмы, которые позволяют работать с очень большим объемом данных. Секвенирование «нового поколения» применяется и для анализа геномов организмов, у которых он уже секвенирован и известна последовательность оснований (ресеквенирование), и для того, чтобы впервые расшифровать геном организма (секвенирование de novo). Для ресеквенирования используются платформы, которые генерируют большое количество коротких секвенируемых фрагментов. Даже короткие фрагменты успешно картируются на референсный5 геном при анализе данных. Это используется для выявления изменений в геноме. В случае секвенирования de novo используют комбинированный подход — сочетание платформ, генерирующих как короткие, так и длинные чтения[4]. Новые методы секвенирования НК позволяют оценивать уровень метилирования ДНК6, проводить анализ дифференциальной экспрессии генов, в том числе и генов-регуляторов, например, таких, как микроРНК7. Оценка уровня метилирования генома, позволяет определить, что реагирует в ответ на меняющиеся факторы окружающей среды, такие иссследования зачастую проводят для изучения эволюционных механизмов в живых системах.

**Пиросеквенирование.**

Одним из методов секвенирования нового поколения является пиросеквенирование. Основной смысл этого метода заключается в последовательном синтезе ДНК на ДНК-фрагментах изучаемого организма в специальных пиколитровых «реакторах». В ходе синтеза дочерней цепочки ДНК детектируют пирофосфаты, высвобождающиеся при включении нуклеотида в синтезируемую на матрице комплементарную цепь[4].

Эту технологию предложил Пол Нирен с коллегами из Королевского технологического института в Стокгольме в 1996 году. Этот метод позволяет определять нуклеотидную последовательность фрагментов геномной ДНК размером 300–500 пар оснований. Методы секвенирования нового поколения требуют предварительной фрагментации ДНК. К обоим концам фрагмента ДНК «пришивают» ДНК-адаптеры, необходимые для последующего секвенирования. Конструкции фрагмент ДНК+ДНК-адаптер называются ДНК-библиотеки. Готовые ДНК-библиотеки фиксируют на магнитных сферах. Затем магнитные сферы с нанесенной на них клональной библиотекой доставляют на проточную ячейку, где в присутствии праймера8, дезоксинуклеотидтрифосфатов и ферментов — ДНК-полимеразы, люциферазы, АТФ-сульфурилазы — происходит циклический синтез новой цепи. Во время цикла пиросеквенирования при образовании фосфодиэфирной связи между матричной цепочкой ДНК и нуклеотидом синтезируемой цепи выделяется пирофосфат. Он запускает каскад химических реакций, приводящих к выделению АТФ, необходимой для реакции окисления люциферина9 с выделением кванта света, который фиксируют с помощью микросхемы, состоящей из светочувствительных фотодиодов. Нуклеотиды, которые не участвуют в синтезе новой цепочки, удаляют из проточной ячейки. После начинается следующий цикл, в ходе которого добавляют дезоксинуклеотидтрифосфат другого типа. В общем, при включении нуклеотида в синтезируемую цепочку ДНК происходит регистрация высвобождающихся пирофосфатов - побочных продуктов реакции полимеризации нуклеотидов в ДНК[4].



В настоящее время для секвенирования используют коммерческие платформы: HiSeq, MiSeq и NextSeq 500 (Illumina), Roche 454 GS, Ion torrent (Thermo Fisher Scientific) и SOLiD (Applied Biosystems) (Cronin, Ross, 2011; Meldrum et al., 2011; Ross, Cronin, 2011; Desai, Jere,2012; Ku et al.,2012; Rizzo, Buck, 2012).

Платформы для секвенирования работают по следующим этапам: подготовка библиотек, секвенирование и детекция. Технологии секвенирования различаются у разных платформ.

**Практическое применение знаний расшифрованного генома человека.**

Одна из коммерческих возможностей применения, полученных данных – это разработка и продажа микрочипов. Микрочипы – большие матрицы для генных фрагментов. Микрочипы являются основой при определении генов, экспрессируемых в данных клетках. Их используют, например, для определения различий между нормальным и аномальным функционированием клеток.



Чтобы воспользоваться микрочипом надо сначала выделить мРНК из исследуемых клеток – и из мутантных, и из нормальных. мРНК отвечают за синтез определенных белков, и таким образом, могут рассказать, какие гены экспрессируются. мРНК используются для того, чтобы на ее основе с помощью обратной транскриптазы синтезировать ДНК, от которой она произошла. Новая ДНК называется кДНК. Обе кДНК помечаются флоуресцентными метками. Получается ДНК-чип с денатурировавшими фрагментами генома исследуемых клеток. Фрагменты ДНК располагаются на предметном стекле определенным образом, известны локализация и соотношение каждого фрагмента. Микрочипы инкубируют с кДНК здоровых и больных клеток. кДНК связываются согласно принципу комплементарности. Сканируя микрочип лазерами цвета меток можно узнать какие гены экспрессируются в обоих типах клеток, только в нормальных и только в аномальных.

Когда изучается ген, но нет возможности проверить, экспрессируется он или нет, используются репортерные гены. Эти гены активны, если активен исследуемый. Исследуемые системы включают вирусную ДНК и мутантные гены. Некоторые репортерные гены отвечают за синтез белка, легко выявляющегося. Один из популярных репортерных генов, является ген, кодирующий зеленый флоуресцентный белок (GFP)[2]. Репортерный ген связывают с определенным исследуемым событием в клетке, например, влиянием человеческого гормона роста – соматотропина – на клетку. Его соединяют с промотером10, чувствительным к этому гормону. Когда соматотропин действует на клетку, она светится зеленым светом.
Репортерные гены используются и в отслеживании передвижения белков[2]. В исследованиях по изучению нормального и аномального метаболизма рецептора рилизинг-фактора гонадотропного гормона11 присоединили ген GFP к гену рецептора. По зеленой флоуресценции можно наблюдать где локализуется данный белок.

Были разработаны методы, которые позволяют по аминокислотной последовательности белка разработать и синтезировать ген, который кодирует данный белок. Этот способ работает даже если такой ген никогда не существовал. Таким образом можно синтезировать ДНК вне живой клетки. Чтобы начать синтез новой цепи необходима основа для удержания молекул на месте и обеспечения их взаимодействия. Синтезаторы ДНК используют силиконовую основу в небольших колонках, связывающую первый нуклеотид в последовательности. Второй нуклеотид пропускают через колонку, он связывается с первым посредством связи между фосфатом и рибозой. Такие связи называются фосфодиэфирными. Благодаря им формируется сахарофосфатный остов ДНК. Молекулы второго нуклеотида должны связываться только с первым нуклеотидом, а не между собой, поэтому их они должны быть химически блокированы определенным образом. Избыток второго нуклеотида вымывается из колонки, химический блок снимается с нуклеотида, связавшегося с основой. Затем добавляют другие нуклеотиды, повторяя процедуру. Сейчас этот процесс автоматизирован.

Болезни, вызванные образованием дефектных белков, можно излечить предотвратив транскрипцию гена, кодирующего данный белок. Кроме этого можно предотвратить некоторые вирусные инфекции, блокировав транскрипцию вирусных генов. Антисмысловые молекулы – это последовательности оснований, связывающих одноцепочечную мРНК. Связывание с такими молекулами означает, что ген не будет экспрессироваться. Антисмысловая молекула – это одноцепочечная нуклеиновая кислота с последовательностью оснований идентичной кодирующей последовательности гена. Эта молекула будет связываться с мРНК, транскрибирующийся с данного гена, потому что последовательность оснований мРНК комплементарна последовательности в гене. мРНК функциональна только в состоянии одноцепочечной молекулы. Антисмысловая цепочка связывается с мРНК, и мРНК уже не может служить матрицей для синтеза белка. Антисмысловые последовательности используются при изучении функций различных генов[2].

Есть и другие способы применить полученные данные в разных областях биологии, медицины и других наук.

Все данные об эволюции взяты из археологических находок, а расшифровка генома не только дала возможность подтвердить теории археологов, но в будущем позволит точнее узнать историю эволюции человека. Появление рибосом и других органелл, развитие эмбриона, иммунной системы позвоночных, можно будет проследить на молекулярном уровне. Это позволит пролить свет на многие вопросы о сходстве и различиях между людьми и нашими ближайшими сородичами: приматами, неандертальцем (чей геном недавно был реконструирован из 1,3 млрд фрагментов, подвергавшихся тысячелетнему разложению и загрязненных генетическими следами археологов, державших в руках останки этого существа), а также и всеми млекопитающими. Мы сможем узнать, как гены влияют на физические и умственные характеристики и наше поведение[6]. Возможно, в будущем, посмотрев на генетический код, можно будет не только предсказать, как будет выглядеть человек, но и, к примеру, будет ли у него актерский талант. Хотя никогда нельзя будет это определить со 100%-ной точностью.

Зная расшифрованный геном можно будет проследить историческое развитие биоты12. Межвидовое сравнение покажет, чем отличается один вид от другого, как они разошлись на эволюционном древе. Межпопуляционное сравнение покажет, как этот вид эволюционирует. Сравнение ДНК отдельных особей внутри популяции покажет, чем объясняется различие особей одного вида, одной популяции. Наконец, сравнение ДНК различных клеток внутри одного организма поможет понять, как происходит дифференцирование13 тканей, как они развиваются и что идет не так в случае заболеваний, таких, например, как рак.

Если соотнести данные о анатомии и физиологии человека с данными генетики, то лечить человека можно будет персонально. Однако, нужно сказать, что для того, чтобы понять, как работает организм конкретного человека информации о геноме недостаточно, нужно посмотреть протеом14 этого человека, он даст более полную картину о работе организма этого человека. Наука, изучающая протеомы, называется протеомика. Она находится в начальной стадии развития. Однако, в 2010 году Россия, США, Южная Корея, Швеция, Канада и Иран запустили проект «Протеом человека». Его цель – систематизировать все белки, которые производит наш организм. Масштабы этого проекта сравнимы с проектом «Геном человека» [6]. Кроме того, необходимо обратить внимание на то, что для того, чтобы понять, как ткани индивидуума дифференцируются, нужно посмотреть транскриптом этого человека. Геном одинаковый у всех клеток организма, но все клетки разные. Чтобы понять, какие гены в клетке экспрессируются и нужны данные о транскриптоме. Обычно исследуют транскриптом органа или клеток определенного типа. Транскриптом меняется со временем: например, гены больных и иммунных клеток работают по-другому[6].

Словарь терминов:

1 - Транскриптом – набор присутствующих в органе или ткани РНК. Именно от РНК будет зависеть прочтение гена, а значит и то как клетки в этом органе или ткани будут дифференцироваться.

2. – Секвенатор – машина, секвенирующая ДНК.

3 – Экспрессия генов - проявление данного гена в организме в форме какого-либо специфического для него признака.

4 – Геномика – наука, изучающая геномы.

5 - Референсный геном — последовательность ДНК в цифровом виде, составленная учеными как общий репрезентативный пример генетического кода того или иного вида.

6 – Метилирование ДНК - это химическая модификация, реакция добавления метильных групп (CH3) в ДНК. При этом последовательность нуклеотидов не меняется.

7 – МикроРНК - это семейство маленьких одноцепочечных РНК, длиной в 21-23 нуклеотида, негативно регулирующих экспрессию генов на посттранскрипционном уровне.

8 - **Праймер** — это короткий фрагмент нуклеиновой кислоты (олигонуклеотид), комплементарный ДНК или РНК мишени, служит затравкой для синтеза комплементарной цепи с помощью ДНК полимеразы, а также при репликации ДНК.

9 – Люциферин - класс светоизлучающих биологических пигментов.

10 - Промотор— последовательность нуклеотидов ДНК, узнаваемая РНК-полимеразой как стартовая площадка для начала транскрипции.

11 – Гонадотропные гормоны - гормоны передней доли гипофиза и плаценты, физиологической функцией которых является регуляция работы половых желёз.

12 – Биота - исторически сложившаяся совокупность флоры, фауны и микроорганизмов (не всегда экологически взаимосвязанных, в отличие от биоценоза) населяющих какую-либо определенную территорию. Термин предложил биолог Э. Раковицэ в 1907 году.

13 – Дифференцирование - возникновение в организме или отдельном его участке в процессе онтогенеза генетически детерминированных необратимых морфологических и функциональных различий, приводящих к формированию специализированных клеток, органов и тканей.

14 – Протеом - совокупность белков организма, производимых клеткой, тканью или организмом в определённый период времени.

 Заключение

Итак, чтобы расшифровать последовательность нуклеотидов в ДНК ученые пользовались и пользуются разнообразными методами. Общее название этих методов – методы секвенирования ДНК. Первая расшифровка человеческого генома продолжалась в течении 10 лет, в результате совместной работы ученых из многих стран в 2003 году эта задача была решена. В итоге исследования изменилось представление о генах. Тогда секвенирование ДНК было долгим и трудоемким процессом, весь проект стоил более 3 миллиардов долларов. Сейчас технологии секвенирования усовершенствовались, процессы были автоматизированы. Были разработаны аппараты, определяющее последовательность нуклеотидов в ДНК - секвенаторы. Благодаря их появлению себестоимость исследования упала, а скорость возросла. Одновременно с развитием техники для секвенирования развиваются и программные коды для них. У ученых появилась возможность единовременно оценивать работу тысяч генов в организмах, тканях и клетках и анализировать регуляцию их активности. Зачем все это? Когда материал исследования подготовлен ученые сравнивают частоту мутаций в генах, выясняют наличие однонуклеотидных полиморфизмов, которые влияют и не влияют на работу генов. По этим данным можно выяснить предрасположенность к заболеваниям в популяциях и родство этих популяций, пути миграций их предков. Данные, полученные в результате секвенирования ДНК позволят лечить заболевания, вызванные дефектными белками. Если же соотнести данные о анатомии и физиологии человека с данными секвенированного ДНК, то лечить человека можно будет персонально. Это значит, что врачи будут выписывать рецепты пациентам «не вслепую», а «лично». Пациент будет принимать лекарства, которые помогут вылечиться именно ему. Однако, на мой взгляд, генетика и геномика только начинают развиваться. За ними светлое будущее человечества.

Список литературных источников:

1. Марк Хендерсон **«**Генетика: 50 идей, о которых нужно знать»
2. Шарон Уолкер "Биотехнология без тайн"
3. <https://elementy.ru/nauchno-populyarnaya_biblioteka/431280/Proekt_Genom_cheloveka_desyat_let_spustya>
4. <https://biomolecula.ru/articles/metody-v-kartinkakh-cekvenirovanie-nukleinovykh-kislot>
5. <http://enc.sci-lib.com/article0001457.html>
6. <http://kot.sh/statya/3785/sekveniruy-menya-polnostyu>