

# Исследования метагенома

Выполнила : студентка 1  
курса магистратуры  
Мягкая Т.И.

- **Метагеномика** — раздел молекулярной генетики, в котором изучается генетический материал, полученный из образцов окружающей среды. Метагеномика изучает набор генов всех микроорганизмов, находящихся в образце среды, — метагеном. Метагеномный анализ позволяет определить видовое разнообразие исследуемого образца без необходимости выделения и культивирования микроорганизмов.
- Основным преимуществом использования метагеномного подхода является учёт не только культивируемых микроорганизмов, но и некультивируемых. Оказалось, что такие организмы вносят основной вклад в видовое разнообразие сообществ. Метагеномика позволяет детально изучить разнообразие сообществ, а значит и выяснить механизмы их функционирования, определить метаболические взаимосвязи.



Широкое развитие метагеномики обусловлено распространением методов секвенирования нового поколения. Они позволяют получить последовательности практически всех генов каждого микроорганизма сообщества. Так как цена на секвенирование ДНК падает с каждым днём, такой анализ становится всё более доступным.

Метагеномика позволяет изучать сообщества микроорганизмов, например, в струях кислой воды, использующихся в горной добыче



# История

- Термин «метагеномика» (англ. metagenomics) впервые был употреблён Джо Хандельсман, Джоном Кларди, Робертом Гудманом, Шоном Брэди и другими в своей публикации 1998 года. Термин «метагеном» возник из идеи, что набор генов, собранных из окружающей среды, можно анализировать подобно тому, как анализируют цельные геномы. Кевин Чен и Лайор Пэтчер (исследователи из Университета Калифорнии, Беркли) определили метагеномику как «применение современных методов геномики без необходимости изолирования и лабораторного культивирования отдельных видов».

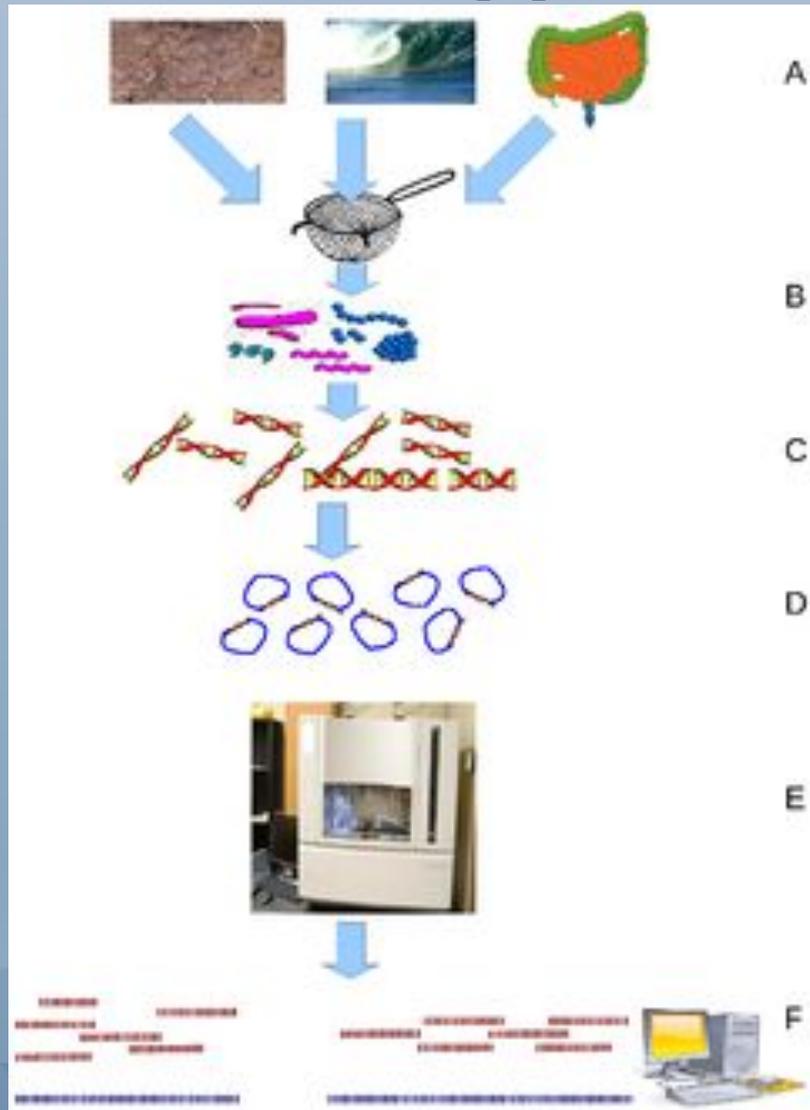


- Долгое время при секвенировании геномов микроорганизмов в качестве источников ДНК использовались, как правило, культуры одинаковых клеток. Однако ранние исследования, посвящённые метагеномике, показали, что во многих средах обитания присутствуют большие группы микроорганизмов, которые нельзя вырастить в лабораторной культуре и, следовательно, секвенировать их геномы. В этих ранних работах изучались последовательности 16S рРНК, которые довольно коротки, часто консервативны в пределах одного вида и, как правило, различаются от вида к виду.
- Начало молекулярным исследованиям в этой сфере положил Норман Пейс и коллеги, которые использовали ПЦР для изучения разнообразия последовательностей рРНК. Благодаря этим исследованиям Пейс выдвинул идею клонирования ДНК непосредственно из образцов из окружающей среды в 1985 году. В 1991 году Пейс и коллеги опубликовали первое сообщение об изоляции и клонировании ДНК из образца из окружающей среды.

- В 1995 году Хили сообщил о метагеномной изоляции функциональных генов из сложной лабораторной культуры микроорганизмов из окружающей среды, выращенной на сухой траве.
- В 2002 году Мия Брейтбард, Форест Роуэр и коллеги при помощи секвенирования образцов окружающей среды по методу дробовика показали, что 200 литров морской воды содержат более 5000 видов вирусов. Дальнейшие исследования показали, что в человеческих фекалиях находится более тысячи видов вирусов, а в килограмме морских донных отложений, возможно, обитает более миллиона видов вирусов, в том числе бактериофагов. Практически все эти вирусы были новыми видами.
- В 2004 году была полностью секвенирована ДНК из кислых вод рудников. Благодаря этому исследованию удалось получить полные или почти полные геномы видов бактерий и архей, которых до этого не удавалось культивировать в лаборатории.

- В начале 2003 года Крейг Вентер, глава проекта, параллельного проекту «Геном человека», организовал экспедицию по сбору образцов океанской воды со всей Земли. Все образцы были секвенированы по методу дробовика с целью идентифицировать геномы новых организмов. В Саргассовом море была определена ДНК 2000 различных видов, в том числе 148 новых видов бактерий.
- В 2005 году Стефан Шустер и коллеги из университета штата Пенсильвания опубликовали первую последовательность из образца окружающей среды, полученную с помощью пиросеквенирования.
- Несколько проектов метагеномики человека находятся на разных стадиях выполнения или уже завершены, в том числе анализ микрофлоры кожи и кишечника.
- В 2007—2008 годах был запущен глобальный проект, получившего название «Микробиом человека». В 2011 году были представлены некоторые результаты. С 2010 года масштабное исследование метагенома человека наметилось и в России. Консорциум ведущих Российских институтов в области гастроэнтерологии и молекулярной биологии в качестве инициативного проекта начал проводить первые эксперименты по широкомасштабному секвенированию образцов ДНК из кишечника человека.

# Секвенирование с использованием случайного фрагментирования



- (A) Получение образца окружающей среды;
- (B) Разделение составляющих образца (например, по размеру);
- (C) Лизис клеток и выделение ДНК;
- (D) Клонирование и создание библиотеки клонов;
- (E) Секвенирование;
- (F) Сборка в контиги и скаффолды.

# Секвенирование эволюционно консервативных генов

- Задача определения видового состава сообщества решается с помощью секвенирования определенных генов, которые должны быть у всех организмов в сообществе. Некоторые участки таких последовательностей геномной ДНК, как например ген, кодирующий 16S рРНК, состоят из высококонсервативных последовательностей и гипервариабельных участков. Эта особенность позволяет использовать праймеры для секвенирования, которые комплементарны консервативным участкам для получения последовательностей гипервариабельных участков. Полученные последовательности позволяют отнести организм к тому или иному виду

# Биоинформатический анализ

Данные, полученные в результате метагеномного эксперимента, содержат огромное количество информации и шума, так как они представляют из себя фрагменты последовательностей ДНК, принадлежащих тысячам и десяткам тысяч разных видов. Сбор, курирование и извлечение полезной биологической информации из наборов данных такого размера представляют из себя сложные вычислительные задачи, которые могут быть решены с помощью биоинформатики

- **Предварительная обработка данных**

Первый этап метагеномного анализа заключается в предварительной фильтрации данных. Он включает удаление избыточных и некачественных последовательностей. Для метагеномов, полученных из организмов животных, важно удаление последовательностей эукариотического происхождения. Удаление загрязнений эукариотической геномной ДНК производится с помощью алгоритмов Eu-Detect и DeConseq.

# Биоинформатический анализ

## Сборка последовательностей

По своей сути последовательности ДНК из геномных и метагеномных экспериментов одинаковые. Тем не менее, метагеномные эксперименты обеспечивают более низкое покрытие, а использование для анализа методов секвенирования нового поколения приводит к ограничению на длину секвенируемой последовательности. Также задача усложняется из-за разной представленности видов в сообществе. Эти особенности приводят к тому, что сборка участков генома из данных метагеномного эксперимента становится сложной задачей, она требует высоких вычислительных мощностей и может приводить к ошибочному результату. Например, могут получиться химерные последовательности, которые представляют из себя комбинацию участков последовательностей ДНК разных организмов

# Биоинформатический анализ

Существует несколько программ, которые производят сборку с учетом парно-концевых чтений, этот метод позволяет уменьшить число ошибок. Такие программы, как Phrap или Celera Assembler, изначально были созданы для сборки единичных геномов, однако они дают неплохие результаты при обработке метагеномных данных. Другие программы, например, Velvet assembler, используют графы де Брейна для того, чтобы справляться с короткими последовательностями (ридами), которые получаются в результате работы методов секвенирования нового поколения. Сборку геномов наиболее распространенных видов облегчает использование референсных геномов. После сборки возникает следующая проблема: необходимо определить, каким видам принадлежат полученные последовательности .

# Биоинформатический анализ

## Поиск кодирующих последовательностей

В метагеномном анализе для аннотации кодирующих последовательностей после сборки, используются два основных подхода. В основе первого метода лежит поиск гомологичных аннотированных генов обычно с помощью BLAST. Такой подход реализован в программе MEGAN4. Второй подход (*ab initio*) использует внутренние особенности последовательности для предсказания кодирующих областей, для его реализации используются обучающие выборки генов родственных организмов. Этот подход используют программы GeneMark и GLIMMER. Основное преимущество подхода *ab initio* состоит в том, что он может определить кодирующие последовательности, для которых неизвестны гомологи.

## Определение видового состава

В то время как аннотация метагенома указывает на то, какие функции реализуются в сообществе, определение видового состава позволяет определить, какие организмы ответственны за их выполнение.

# Биоинформатический анализ

Процесс соотнесения определенных генов, а значит и функций, которые они могут выполнять, с определенными видами организмов называется биннинг. Он реализуется с помощью метода BLAST через поиск похожих генов, для которых известно, какому организму они принадлежат. Этот подход реализован в программе MEGAN (MEta Genome ANalyzer).

## Интеграция данных

Анализ большого экспоненциально растущего количества доступных последовательностей ДНК является сложной задачей. Кроме того анализ затрудняют, сложные метаданные, связанные с метагеномными проектами. Они включают в себя информацию о географии исследуемого образца, особенностях окружающей среды, физические данные, а также методику отбора проб. Эта информация необходима для обеспечения воспроизводимости экспериментов и для дальнейшего анализа.

# Области применения метагеномики

## Медицина

Микробные сообщества играют ключевую роль в поддержании здоровья человека, но их состав и механизмы функционирования до сих пор остаются неразрешенными. Метагеномное секвенирование использовалось для описания микробных сообществ у сотен индивидуумов. Это часть так называемого проекта Микробиома Человека, главными целями которого является: определить базовый набор микробов человека, понять, как изменение микрофлоры человека коррелирует с изменением здоровья, а также разработать технологическую и биоинформатическую базу для достижения этих целей. Основные центры, получившие финансирование на секвенирование геномов микроорганизмов кишечника, пригодных к культивированию в лабораторных условиях:

- Бэйлорский университет (г. Уэйко, Техас, США);
- Broad Institute (объединенный институт, в состав которого входят Массачусетский технологический институт, Гарвардский университет и институт Уайтхеда);
- университет Вашингтона в Сент-Луисе, (штат Миссури).

# Области применения метагеномики

## Биотопливо

Биотопливо получается за счет конверсии биомассы, например, при превращении целлюлозы, полученной из кукурузы и проса, в гидролизный спирт. Этот процесс зависит от микробных консорциумов, которые преобразуют целлюлозу в сахара, с последующим сбраживанием сахара в этанол. Микроорганизмы также производят различные источники биоэнергии, включая метан и водород.

Для эффективного промышленного получения из биомассы новых соединений требуются новые ферменты с более высокой производительностью и меньшими затратами на производстве. Метагеномные подходы к анализу сложных микробных сообществ позволяют проводить целевой скрининг ферментов промышленного применения в производстве биотоплива, таких как гликозил-гидролаз. Кроме того, знания о том, как микробные сообщества функционируют, необходимы для управления этими сообществами, а метагеномика является ключевым инструментом в их понимании. Метагеномные подходы позволяют проводить сравнительный анализ между конвергентными системами микроорганизмов.

# Области применения метагеномики

## Биоремедиация

С помощью метагеномики можно улучшить стратегии для мониторинга воздействия загрязняющих веществ на экосистему, а также разработать новые методы очистки загрязненных сред. Более глубокое понимание того, как микробные сообщества справляются с загрязняющими агентами дает надежду на то, что этот процесс можно в будущем использовать для борьбы с техногенными загрязнениями.

## Биотехнология

Микробные сообщества могут продуцировать широкий спектр биологически активных веществ, которые далее используются другими организмами. Многие лекарственные препараты, используемые сегодня, первоначально были обнаружены у микроорганизмов. Недавний успех в получении разнообразного генетического материала из некультивируемых микроорганизмов привел к открытию новых генов, ферментов и других активных соединений. Применение метагеномики позволило развить новые отрасли химической и фармацевтической промышленности.

# Области применения метагеномики

## Сельское хозяйство

В одном грамме почвы, используемой для выращивания растений, содержится от  $10^9$  –  $10^{10}$  микробных клеток. Состав микробных сообществ, обитающих в почве, давно привлекал внимание ученых, однако до сих пор остается плохо изученным, несмотря на их экономическое значение. Микробные сообщества выполняют широкий спектр экосистемных функций, необходимых для роста растений (например, фиксация азота), защиты растений от болезней, участие в круговороте железа и других металлов. Метагеномика помогает изучать взаимодействия микробов в этом сообществе, а также взаимодействие между растениями и микробами. На основе данных, полученных с помощью метагеномного анализа, можно выявить свойства микроорганизмов, относящихся к некультивируемым таксонам, понять их роль в круговороте веществ, а также их взаимоотношения с растениями. Всё это необходимо для улучшения здоровья сельскохозяйственных культур.

# Области применения метагеномики

## Экология

Метагеномика может дать ценную информацию о функциональной экологии сообществ окружающей среды. Например, метагеномный анализ бактериальных сообществ, обнаруженных в дефекациях австралийских морских львов, показывает, что фекалии морских львов богаты питательными веществами и могут быть важным источником питания для прибрежных экосистем. Это происходит потому, что бактерии, которые выбрасываются одновременно с фекалиями, могут превращать трудноусваиваемые соединения в биологически доступные формы, которые далее могут быть вовлечены в пищевую цепь.

Секвенирование ДНК также может быть использовано для идентификации видов, присутствующих в толще воды. Это может помочь установить диапазон инвазионных видов и исчезающих видов, а также отслеживать сезонные популяции.

**Спасибо за внимание!**