Лекция 10. Картирование генома. Стратегические подходы к картированию геномов. Методы картирования геномов млекопитающих.

1 Картирование генома.

2 Стратегические подходы к картированию геномов.

3 Методы картирования геномов млекопитающих.

1. Картирование генома. Как и любой тип карты, генетическая карта должна показывать положение отличительных особенностей. На географической карте эти [маркеры](https://www.ncbi.nlm.nih.gov/books/n/genomes/A9089/def-item/A9631/) являются узнаваемыми компонентами ландшафта, такими как реки, дороги и здания. Какие маркеры мы можем использовать в генетическом ландшафте?

Первые генетические карты, созданные в первые десятилетия 20-го века для таких организмов, как плодовая мушка, использовали гены в качестве маркеров. Это было за много лет до того, как стало понятно, что гены являются сегментами молекул ДНК. Вместо этого гены рассматривались как абстрактные сущности, ответственные за передачу наследуемых характеристик от родителя к потомству. Чтобы быть полезным в генетическом анализе, наследуемая характеристика должна существовать по крайней мере в двух альтернативных формах или [фенотипах](https://www.ncbi.nlm.nih.gov/books/n/genomes/A9089/def-item/A9771/) , примером являются высокие или короткие стебли у растений гороха, первоначально изученных Менделем.

Каждый фенотип определяется различным аллелем [соответствующего](https://www.ncbi.nlm.nih.gov/books/n/genomes/A9089/def-item/A9126/) гена. Для начала, единственными генами, которые можно было изучить, были те, которые определяли фенотипы, которые можно было различить при визуальном осмотре. Так, например, первые карты плодовой мушки показывали положение генов, отвечающих за цвет тела, цвет глаз, форму крыльев и тому подобное, причем все эти фенотипы можно было увидеть, просто глядя на мух с помощью маломощного микроскопа или невооруженным глазом. Этот подход был хорош в первые дни, но генетики вскоре поняли, что существует лишь ограниченное количество визуальных фенотипов, наследование которых можно изучать, и во многих случаях их анализ был сложным, поскольку на один фенотип могло влиять более одного гена. Например, к 1922 году более 50 генов были сопоставлены с четырьмя хромосомами плодовой мушки, но девять из них отвечали за цвет глаз; в более поздних исследованиях генетикам, изучающим плодовых мушек, пришлось научиться различать глаза мух, окрашенные в красный, светло-красный, киноварный, гранатовый, гвоздичный, киноварный, рубиновый, сепия, алый, розовый, кардинальный, бордовый, фиолетовый или коричневый цвета. Чтобы сделать генные карты более полными, необходимо было найти характеристики, которые были бы более отличительными и менее сложными, чем визуальные.

Ответом было использование биохимии для различения фенотипов. Это было особенно важно для двух типов организмов - микробов и людей. Микробы, такие как бактерии и дрожжи, имеют очень мало визуальных характеристик, поэтому картирование генов с этими организмами должно полагаться на биохимические фенотипы, такие как перечисленные. Для людей можно использовать визуальные характеристики, но с 1920-х годов исследования генетической изменчивости человека основывались в основном на биохимических фенотипах, которые можно оценить с помощью типирования крови. Эти фенотипы включают не только стандартные группы крови, такие как серия ABO, но также варианты белков сыворотки крови и иммунологических белков, таких как антигены лейкоцитов человека (система HLA).

Большим преимуществом этих маркеров является то, что многие из соответствующих генов имеют [несколько аллелей](https://www.ncbi.nlm.nih.gov/books/n/genomes/A9089/def-item/A9681/) . Например, ген *HLA-DRB1* имеет не менее 290 аллелей, а *HLA-B* — более 400. Это важно из-за способа, которым проводится картирование генов у людей. Вместо того, чтобы проводить множество экспериментов по разведению, что является процедурой с экспериментальными организмами, такими как плодовые мушки или мыши, данные о наследовании человеческих генов должны быть получены путем изучения фенотипов, демонстрируемых членами одной семьи. Если все члены семьи имеют один и тот же аллель для изучаемого гена, то никакой полезной информации получить невозможно. Поэтому необходимо, чтобы соответствующие браки произошли случайно между особями с разными аллелями. Это гораздо более вероятно, если изучаемый ген имеет 290, а не два аллеля.

Картирование генов — это процесс определения местоположения генов в геноме. Ученых особенно интересуют области генома, которые напрямую кодируют белки, или белок-кодирующие гены, поэтому часто приоритетом является локализация каждого гена в геноме.

В настоящее время процедура картирования обычно включает секвенирование генома и анализ полученной последовательности с использованием цифровых методов, которые позволяют выявить нужные гены. Таким образом, большинство проектов по картированию генов начинается с секвенирования генома.

Картирование генов включает методы для определения местоположения гена и расстояния между генами. С помощью картирования также можно описать расстояния между различными участками внутри одного гена.

Основой всех карт генома является размещение нескольких молекулярных маркеров в определенных точках генома. Существует множество видов молекулярных маркеров. При создании геномных карт гены могут рассматриваться как особый класс генетических маркеров, которые картируются аналогично другим маркерам.

Генетико-сцепленные карты и физические карты — это две основные категории «карт», используемых в картировании генов.

Обе карты состоят из генетических маркеров и локусов генов. Физические карты отражают реальные физические расстояния, часто измеряемые в количестве пар оснований, тогда как расстояния на генетико-сцепленных картах основаны на данных о генетической сцепленности.

Существует множество методов картирования генов, включая сравнительное, физическое и генетико-сцепленное картирование. Однако физическое и генетико-сцепленное картирование являются наиболее распространенными.

Генетико-сцепленные карты показывают расположение каждого гена на хромосоме и их относительные расстояния друг от друга. Изначально такие карты создавались на основе анализа наследования различных признаков, таких как цвет глаз или волос. Процесс создания генетической карты начинается с анализа образцов крови, слюны или ткани как от больных, так и от здоровых членов семьи. Слюна является наиболее часто используемым образцом в генетическом картировании, особенно в личных генетических исследованиях.

Генетическое картирование становится возможным благодаря кроссинговеру — естественному биологическому процессу, происходящему во время мейоза, то есть деления клеток, приводящего к образованию сперматозоидов и яйцеклеток. В первой фазе мейоза хромосомы выстраиваются парами в центре клетки, часто «прилипая» друг к другу и обмениваясь сходными фрагментами, что и называется кроссинговером. Генетическое картирование позволяет определить, какой ген присутствует на каждой хромосоме и где он расположен. На основе расстояния между двумя генами картирование также может определить, какой из них более вероятно подвергнется рекомбинации.

***Физическое картирование***. Физические карты всегда предоставляют точные физические расстояния между маркерами ДНК, измеряемые в парах оснований. Этот подход к картированию генов характеризуется высокой степенью точности при определении последовательности оснований ДНК. Физическая карта указывает количество нуклеотидов и точное физическое расстояние между генетическими маркерами. Существует несколько методов, таких как радиационное гибридное картирование, последовательное картирование и цитогенетическое картирование, с помощью которых создается физическая карта.

Физическое картирование основывается на сборке больших фрагментов ДНК с использованием маркеров и фрагментов ДНК. Исследователи могут определить расположение оснований ДНК, анализируя перекрывающиеся участки этих фрагментов. Существуют различные методы физического картирования, которые позволяют изучать геномы разного размера и обеспечивать разную степень точности. Физическое картирование является распространенным подходом, используемым для получения полной последовательности генома и анализа взаимосвязи между определенной последовательностью ДНК и фенотипическими характеристиками.

***Методы генетического картирования.*** Методы генетического картирования используют события рекомбинации для измерения расстояния между генетическими маркерами. Одним из таких методов является полиморфизм длины случайных фрагментов (RFLP), который измеряет различия в гомологичных последовательностях ДНК для вычисления расстояния между двумя маркерами. В настоящее время анализы картирования генов, направленные на изучение моногенных заболеваний, используют полиморфизмы коротких тандемных повторов (STRP).

Полиморфизм однонуклеотидов (SNP) применяется в генетических исследованиях, ориентированных на анализ ассоциаций и сцепления генов на уровне всего генома. Для анализа сцепления используются такие признаки и генетические маркеры, как SNP и микросателлиты. Исследования ассоциации на уровне всего генома (GWA) изучают связи между признаками и маркерами, такими как SNP и микросателлиты, рассматривая популяцию как одну семью. Этот метод применяется для картирования генов, связанных с распространенными заболеваниями.



***Рисунок 1.*** *Иллюстрация, показывающая разницу между двумя основными способами картирования генома: генетическое картирование и физическое картирование (Image credit: Laura Olivares Boldú / Wellcome Connecting Science).*

1. **Стратегические подходы к картированию геномов.** Картирование геномов — это основополагающий процесс в генетике и геномике, который позволяет определить расположение генов и других функциональных элементов в пределах генома. Этот процесс включает в себя использование различных методов и стратегий, направленных на точное и детальное изучение генетической структуры и связи между элементами генома. В последние годы картирование геномов достигло значительных успехов благодаря развитию секвенирования нового поколения и внедрению усовершенствованных методов анализа данных.

Существует несколько ключевых подходов к картированию геномов, среди которых генетическое картирование, физическое картирование и сравнительное картирование. Генетическое картирование строится на данных сцепленности генов, позволяя определить их относительное положение на хромосоме. Этот метод был разработан в начале 20-го века и основан на анализе наследственных признаков. Генетическое картирование долгое время служило основой генетических исследований и позволило создать базовые генетические карты для ряда организмов, таких как плодовая мушка и кукуруза. Сегодня этот подход используется для исследования полигенных признаков и наследственных заболеваний, где требуется анализ нескольких генов, влияющих на развитие фенотипа.

В отличие от генетического, физическое картирование направлено на точное определение расстояний между генетическими маркерами на уровне пар оснований. Физическое картирование обеспечивается методами, такими как оптическое картирование, радиационно-гибридное картирование и последовательное картирование. Оптическое картирование стало важной частью современных стратегий картирования, так как оно позволяет с высокой точностью определять крупные структурные изменения в геноме, такие как инверсии и дупликации. Этот метод особенно полезен при работе с организмами, геномы которых содержат многочисленные повторяющиеся последовательности, что затрудняет использование других методов. Радио-гибридное картирование, основанное на анализе гибридных клеток, также позволяет получить высокую точность при построении карт геномов, особенно при работе с крупными геномами.

Сравнительное картирование — это ещё один значимый подход, который нашёл широкое применение в эволюционной биологии и медицине. Сравнительное картирование используется для сопоставления геномов различных видов, что помогает выявить консервативные и изменчивые элементы в геноме, а также понять эволюционные процессы, которые привели к их изменению. Благодаря этому подходу стало возможным анализировать гены, участвующие в развитии определённых заболеваний, что позволяет создавать эффективные диагностические тесты и новые методы терапии. Например, сравнение геномов людей и других млекопитающих позволяет исследователям определять консервативные гены, связанные с метаболизмом, и выявлять мутации, которые могут быть причиной генетических заболеваний.

Развитие секвенирования нового поколения и массового параллельного секвенирования вывело картирование геномов на новый уровень, особенно в области физического картирования. Современные секвенаторы, такие как Illumina и Oxford Nanopore, обеспечивают возможность быстро получать большие объёмы данных, что делает физическое картирование более доступным и точным. Например, использование технологии длинных ридов Oxford Nanopore позволило значительно повысить точность картирования сложных геномов, таких как геномы растений и многоклеточных организмов. Эта технология позволяет детально изучать структурные вариации, включая крупные делеции и инверсии, которые ранее было трудно идентифицировать с помощью традиционных методов.

Биоинформатические алгоритмы и методы машинного обучения стали неотъемлемой частью картирования геномов, так как позволяют обрабатывать огромные объемы данных, полученные при секвенировании. С помощью современных методов глубокого обучения можно предсказывать функциональные элементы генома, а также обнаруживать сложные генетические паттерны. Одним из важных инструментов, используемых в геномном картировании, является граф де Брейна, который позволяет строить геномные последовательности даже при наличии многочисленных повторяющихся элементов. Также активное развитие получили методы интеграции данных, когда данные секвенирования дополняются данными протеомики и эпигеномики, что позволяет создать более полные и информативные геномные карты.

Одним из значительных достижений последних лет стало создание глобальных референсных геномов для различных популяций и этнических групп. Этот подход позволяет выявлять уникальные генетические особенности, характерные для разных популяций, и вносит огромный вклад в развитие персонализированной медицины. Исследования показали, что включение данных о разнообразии геномов из разных частей мира позволяет улучшить точность диагностики и повысить эффективность лечения за счёт адаптации к индивидуальным генетическим особенностям пациента.

Картирование геномов играет важную роль в разработке методов ранней диагностики и таргетной терапии. Например, в онкологии картирование позволяет выявлять специфические мутации, связанные с агрессивными формами рака, что позволяет разрабатывать индивидуализированные схемы лечения. В сфере инфекционных заболеваний картирование геномов патогенов помогает отслеживать мутации, отвечающие за лекарственную устойчивость, и прогнозировать эпидемии, что особенно актуально в условиях пандемий.

Таким образом, стратегические подходы к картированию геномов включают комбинацию методов генетического, физического и сравнительного картирования, а также новейшие технологии секвенирования и биоинформатические методы. Эти подходы расширяют наши знания о структуре и функции геномов, создавая возможности для новых открытий в медицине, биотехнологии и сельском хозяйстве.

**3 Методы картирования геномов млекопитающих.** Методы картирования геномов млекопитающих играют решающую роль в изучении генетической структуры и функциональной организации генома, а также в понимании эволюции и генетической изменчивости этих видов. Эти исследования имеют важное значение для биомедицинских наук, так как млекопитающие, включая человека и модельные организмы, такие как мыши, крысы и крупные животные, служат объектами изучения генетических основ заболеваний, разработки лекарств и методов терапии. Картирование геномов млекопитающих включает использование нескольких подходов, которые дополняют друг друга и позволяют получить детальное представление о генетической архитектуре и функциональной организации генов.

Один из основных методов картирования геномов млекопитающих — это **генетическое картирование**, основанное на изучении наследственных признаков и данных сцепленности генов. Этот метод используется для построения карт сцепленности, которые показывают относительное расположение генов на хромосомах. Генетическое картирование требует анализа полиморфизмов, таких как SNP (однонуклеотидные полиморфизмы) и STR (кратные тандемные повторы), что позволяет детально изучить сцепленность признаков. Этот метод особенно полезен для исследования полигенных признаков, таких как предрасположенность к определённым заболеваниям и агрессивность опухолевых процессов. В модельных организмах, таких как лабораторные мыши, генетическое картирование используется для поиска генов, связанных с различными заболеваниями, такими как диабет и сердечно-сосудистые заболевания. Эти карты служат основой для дальнейшего картирования и анализа мутаций, влияющих на развитие фенотипических признаков.

**Физическое картирование** является еще одним ключевым методом, который используется для точного определения расстояний между генетическими маркерами на уровне пар оснований. Этот метод включает такие подходы, как оптическое картирование и радиационно-гибридное картирование, которые позволяют получить высокую разрешающую способность и точность при картировании геномов. Оптическое картирование стало особенно важным в исследованиях крупных геномов млекопитающих, так как позволяет визуализировать крупные хромосомные перестройки и структурные вариации. Этот метод часто используется в комбинации с полногеномным секвенированием для уточнения расположения сложных повторяющихся элементов и детекции структурных аномалий. В последние годы физическое картирование применялось для анализа геномов крупных животных, таких как крупный рогатый скот и свиньи, что способствует развитию генетических методов селекции и повышению продуктивности в сельском хозяйстве.

**Сравнительное картирование** — это метод, который позволяет сопоставлять геномы различных млекопитающих, что важно для понимания эволюционных процессов и выявления генетически консервативных элементов. Этот подход применяется для выявления гомологичных генов и общих генетических путей, что делает его незаменимым в эволюционной биологии и медицине. Сравнительное картирование используется для создания генетических карт на основе данных о других видах, что значительно ускоряет картирование новых геномов и облегчает поиск ключевых генов, связанных с важными физиологическими функциями. Например, сравнение генома человека с геномами шимпанзе и мышей позволило определить гены, важные для развития мозга и когнитивных функций, что даёт возможность глубже понять уникальные черты человеческого генома.

Современные технологии секвенирования нового поколения (NGS) и методы длинных ридов, такие как **PacBio** и **Oxford Nanopore**, стали важной частью картирования геномов млекопитающих, так как они позволяют получать непрерывные и длинные последовательности, что облегчает сборку сложных регионов генома. Эти методы широко используются в полногеномном секвенировании, так как они позволяют точно картировать геномы даже в присутствии многочисленных повторов. С использованием длинночитающих технологий стало возможным создание высококачественных референсных геномов, которые открывают новые горизонты в изучении генетических основ сложных признаков. Для млекопитающих, таких как мыши и крысы, высококачественные референсные геномы способствуют созданию моделей для изучения заболеваний, связанных с раковыми мутациями, и исследования механизмов наследования.

**Генетические и физические карты** используются в комбинации для построения полных геномных карт, которые включают как последовательности ДНК, так и функциональные элементы, такие как регуляторные элементы и энхансеры. Эти карты помогают исследовать регуляцию экспрессии генов, геномные перестройки и идентифицировать ключевые области, влияющие на фенотипические проявления. В сельском хозяйстве и ветеринарии картирование геномов млекопитающих, таких как крупный рогатый скот и свиньи, позволяет выявлять генетические маркеры, которые можно использовать для улучшения селекции и повышения устойчивости к болезням.

На уровне биоинформатики стратегическое картирование геномов млекопитающих сопровождается использованием алгоритмов машинного обучения и глубокого обучения для обработки больших объемов данных, полученных в ходе секвенирования. Эти алгоритмы позволяют анализировать геномные данные и идентифицировать уникальные генетические паттерны, а также использовать их для прогнозирования функциональных характеристик генов. Сочетание биоинформатических инструментов с методами картирования облегчает создание детализированных генетических карт, что важно для точной диагностики и разработки персонализированных подходов в медицине.



***Рисунок 2.*** *Два класса якорных локусов для построения карт генов животных.*

Результаты сравнительного генетического анализа геномов млекопитающих выявили несколько важных закономерностей, заслуживающих внимания. Во-первых, хромосомная эволюция у млекопитающих отличается сравнительной консервативностью. Хромосомные перестройки случаются настолько редко, что предковые кариотипы некоторых родов, семейств и даже отрядов можно восстановить на основе анализа G-бандинговых паттернов у современных видов. Эта цитологическая консервация часто, но не всегда, подтверждается сохранением синтении (сходства генетического порядка на хромосомах). Сравнительное картирование генов также выявило случаи гомологии генов даже при отсутствии очевидной хромосомной гомологии.

Несмотря на общую тенденцию к хромосомной консервации, существуют значительные исключения. Например, у гиббонов и совиных обезьян наблюдаются значительно перестроенные кариотипы по сравнению с предковыми формами приматов. Их синтенийные карты также сильно отличаются от карты гомологичных локусов у человека. У хищников, таких как медведи (семейство *Ursidae*) и собаки (семейство *Canidae*), кариотипы отличаются от прототипического кариотипа хищников, но остаются на уровне, который позволяет восстановить эволюционные события хромосомной радиации.

Сравнение синтенийных групп, сохранившихся между по крайней мере двумя отрядами млекопитающих (например, приматами и хищниками), показывает высокий уровень цитологической гомологии (около 30-40% по длине). Эти гомологии по G-бандированию, вероятно, были унаследованы от общего предка линий хищников и приматов и могли присутствовать у других предков плацентарных млекопитающих. По мере разработки карт геномов и кариотипов с гомологичными генами для других отрядов млекопитающих появится возможность использовать эти данные для реконструкции филогенетических связей между отрядами млекопитающих на основе принципа парсимонии.

Домашняя мышь, помимо человека, является млекопитающим, для которого составлено наибольшее количество генетических карт: к настоящему моменту картировано более 2800 локусов. Сравнение геномов человека и мыши положило начало полю сравнительного картирования генов млекопитающих, и сейчас можно сопоставить 560 локусов мыши с их гомологами у человека. Эти исследования показывают, что геном мыши состоит примерно из 90 коротких сегментов с гомологичными последовательностями в геноме человека, что позволяет точно предсказывать расположение генов у человека, основываясь на данных мыши, и наоборот.

На сегодняшний день у млекопитающих было создано два типа генетических карт: физические карты и карты сцепления (рекомбинационные карты). Физические карты основываются на локализации генов на хромосомах с использованием гибридных панелей, гибридизации in situ одноразовых молекулярных клонов, анализа транслокаций и делеции, а также упорядочивания молекулярных клонов и континов. Карты сцепления, напротив, зависят от частоты мейотических обменов между сцепленными полиморфными маркерами в результате полового скрещивания. Анализ сцепления требует наличия полиморфных маркеров. В геномной карте человека описано более 2000 полиморфных локусов, однако фактическая карта сцепления, составленная в основном Центром изучения полиморфизма человека (CEPH) на основе анализа 80 многопоколенных семей, включает менее 1000 маркеров. Цель исследований — объединить физические и карты сцепления для видов, однако этот процесс требует значительных трудозатрат.

Физические карты млекопитающих разрабатывались методами, применяемыми для человека, такими как панели гибридных клеток и гибридизация in situ. Карты сцепления у мышей создавались методом возвратного скрещивания гибридов F1, полученных путем скрещивания инбредных линий, различающихся по многим локусам. Этот метод сложен, так как требует наличия большого числа инбредных линий и генетических различий между ними по многим локусам. Недавно возвратное скрещивание гибридов F1, полученных между генетически далекими видами мышей, позволило создать многоточечные карты сцепления хромосом мыши. Этот метод межвидовых скрещиваний эффективен, поскольку эволюционное различие в последовательностях ДНК вокруг гомологичных генов почти всегда приводит к вариациям полиморфизма длины рестрикционных фрагментов (RFLP), которые можно выявить с помощью кДНК-зондов.

Работа Коупленда, Дженкинса и их коллег позволила создать карты сцепления для 15 из 20 хромосом мыши, основываясь на межвидовых скрещиваниях между лабораторной мышью (Mus musculus domesticus) и Mus spretus. Подход межвидовых скрещиваний также используется в проектах по картированию геномов крупного рогатого скота, овец и домашних кошек. В исследованиях картирования генов у человека и мышей ведутся работы по выделению группы якорных локусов, которые могут служить в качестве индексных маркеров для разных лабораторий. Важнейшими критериями для этих маркеров у человека являются их однокопийность, точное упорядочение, свободная доступность (в виде клона или последовательности, полученной методом полимеразной цепной реакции) и высокий информационный индекс полиморфизма.

Вопросы для самоконтроля:

1. В чем заключается основное отличие между физическим и генетическим картированием геномов?
2. Как межвидовые скрещивания помогают повысить точность карт сцепления в картировании геномов млекопитающих?
3. Какие основные задачи решает сравнительное картирование геномов в изучении эволюции млекопитающих?
4. Какова роль якорных локусов в обеспечении стандартизации генетического картирования?
5. Почему объединение физического и генетического картирования геномов считается трудоемким процессом?
6. Какие преимущества дают современные методы длинных ридов, такие как PacBio и Oxford Nanopore, для работы с повторяющимися последовательностями в геномах млекопитающих?
7. Какие биоинформатические подходы используются для реконструкции структуры геномов млекопитающих на уровне целых хромосом?