

Дизайн белков с заданными функциями. Принципы работы секвенаторов нового поколения. Регуляторные элементы в геноме

1. Дизайн белков с заданными функциями

Методы, которыми могут оперировать ученые на сегодняшний момент, не позволяют точно предсказывать структуру белка по его первичной структуре. Но можно идти в другом направлении: зная первичную последовательность аминокислот в белке, можно создать белок с заранее известной структурой и свойствами. Это область экспериментальных исследований.

2. Создание белков с мотивом цинковых пальцев

Одни из первых белков, которые были созданы искусственно, стали белки, содержащие мотивы цинковых пальцев (см. рис. ??). Эти мотивы чаще встречаются в транскрипционных факторах, которые узнают сайты посадки на ДНК и привлекают другие белки, в том числе РНК-полимеразу, тем самым регулируя транскрипцию. Ученые проанализировали строение транскрипционных факторов на предмет участка, распознающего ДНК. Располагая такими данными, исследователи могли точно изменять аминокислотную последовательность так, чтобы белок узнавал произвольный участок ДНК.

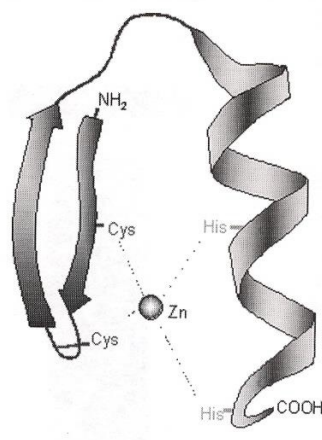


Рис. 1: Домен цинковые пальцы

Существует специальный сервис (**Zinc Finger Tools**), который конструирует наиболее удачную аминокислотную последовательность фактора, узнающего необходимую нуклеотидную последовательность. Эта область разрабатывается относительно давно и технология была довольно успешной до недавнего времени, пока не появились новые более продвинутые технологии.

3. Технология TALEN

TALEN (Transcription activator-like effector nucleases) — белки открытые у прокариот, отвечающие за «иммунный ответ». Эти белки — нуклеазы — разрезают чужеродную ДНК. TALEN-белки были выделены из *Xanthomonas bacteria*. Структура связывающегося с ДНК участка своеобразна. Все цитозины ДНК узнаются парой аминокислот HD (гистидин — аспарагиновая кислота), все тимины — NG (аспарагин — глицин), аденины — NI (аспарагин — изолейцин), гуанины — NN (аспарагин-аспарагин)(см. рис. ??). При изучении схожих белков выяснили, что это консервативный механизм для многих организмов. Ученые по тому же принципу

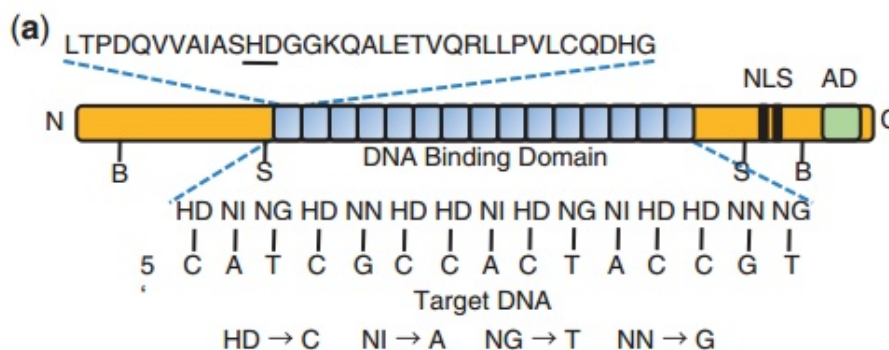


Рис. 2: Структура связывающегося с ДНК участка TALEN-белка

сконструировали искусственные нуклеазы. Сейчас существует несколько сайтов, которые по заданной нуклеотидной последовательности генерируют успешно ее узнающую аминокислотную.

На основе TALEN-белков можно создавать фьюжн-белки, одна часть которых будет распознавать ДНК по вышеописанному механизму, а вторая часть белка может нести другую функцию. Например, функцию транскрипционного фактора, привлекающего РНК-полимеразу к определенному гену. Или, например, с нуклеазной активностью. Тогда такой белок разрежет ДНК в точно определенном месте.

4. Технология CRISPR

Еще одна интересная технология CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats). Ключевым ферментом этой технологии является нуклеаза Cas9. Нуклеаза Cas9 взаимодействует с таргетной РНК. Эта РНК состоит из двух частей: первая узнается нуклеазой, а вторая часть вариабельна и комплементарна какому-либо участку ДНК. Эта РНК связывается с ДНК, их дуплекс узнается нуклеазой, которая разрезает ДНК по месту связи с РНК.

Все эти технологии (конструирование белков с цинковыми пальцами, TALEN и CRISPR) позволяют ученым сделать специфические разрезы в ДНК.

5. Рекомбинационная репарация двуниевых разрывов

Для чего это может понадобиться? Известно, что если в клетке возникнет разрыв ДНК, то запустится **система репарации**, которая восстановит разрыв. Восстановление может происходить несколькими путями:

- Негомологичная рекомбинация — способ репарации, при котором особые белки узнают разрыв и напрямую соединяют фрагменты между собой. При таком способе в восстановленную ДНК включаются лишние нуклеотиды, либо из нее несколько нуклеотидов выпадает (инсерции или делеции). В итоге с этого гена белка вообще не будет, или будет, но нефункциональный.
- Гомологичная рекомбинация — способ репарации, при котором в клетке есть фрагмент ДНК, комплементарный участкам двух фрагментов. Такую донорную ДНК ДНК-полимераза распознает как матрицу для восстановления бреши. Таким образом, в ДНК можно вставить фрагмент любой длины. То есть можно делать надрезы в генах с вредными мутациями, вставлять в них фрагмент, содержащий нормальный ген и, таким образом, исправлять мутацию.

6. Применение CRISPR и TALEN технологий

Известный исследователь George Church в статье 2013 года (Yang et al., 2013) пытался оценить, какая из технологий работает лучше. В результате при применении технологии TALEN успешная негомологичная рекомбинация происходила в 0,4% случаев, а при применении Cas-9-gRNA (CRISPR) — 3%. Гомологичная

рекомбинация происходит в обеих клеточных культурах примерно одинаково (TALEN — 0,6%, Cas-9-gRNA — 1%). То есть технология CRISPR не намного лучше TALEN.

В июле 2014 года был создан сайт, который делает эти технологии доступными для пользования (Montague et al., 2014). На сайте можно указать место генома, куда нужно вставить фрагмент, и технологию, с помощью которой это будет осуществляться. На выходе получают данные о необходимых для этой процедуры белках, праймерах, фрагментах ДНК, которые синтезируют в лаборатории и доставляют исследователю для эксперимента.

Вышеупомянутая работа 2013 года по оценке эффективности технологий CRISPR и TALEN была сделана на гене **CCR5**. Продукт этого гена является одной из мишеней для ВИЧ, а также рецептором Т-лимфоцитов на хемокины. ВИЧ очень мутабелен, этим объясняется устойчивость вируса к различным вакцинам. Но не все люди могут заболеть СПИДом, потому что у некоторых просто нет рецептора CCR5. При этом его ген в геноме таких людей все же есть, но в нем делеция в 32 нуклеотида, которая приводит к тому, что рецептор не синтезируется. Около 2% людей имеют такую мутацию и не подвержены СПИДу.

Работа по оценке эффективности технологий проводилась на плюрипотентных стволовых клетках. В этой работе авторы предложили схему использования технологий редактирования генома. По этой схеме для начала проводят биопсию кожи (забор фибробластов) и репрограммирование клеток (получение индуцированных стволовых клеток). Далее в этих клетках можно провести коррекцию геномного материала — включение в геном интересующих генов с помощью технологии TALEN. Затем их подсаживают обратно в организм донора и заставляют дифференцироваться в ткани. Эту схему теоретически можно применять к больным СПИДом.

У пациента уже невозможно изъять вирус из крови. Можно провести биопсию, репрограммирование клеток, корректирование их генома. Затем подвергнуть пациента жесткой химиотерапии, которая убьет все зараженные клетки; человек останется без иммунитета. Затем измененные клетки посадить пациенту обратно и заставить их дифференцироваться в клетки крови. Эти клетки будут устойчивы к ВИЧ.

Корейская компания в 2014 году претворила эту схему в жизнь (Ye et al., 2014). Она испытала этот метод на 3 людях и было показано, что действительно, отредактированные клетки дифференцировались в моноциты и макрофаги, которые были устойчивы к ВИЧ. Но эти клетки пока не подсаживались донору.

Используя ту же технологию можно лечить и другие генетические заболевания. В частности бета-талассемию, которая проявляется в результате нарушения синтеза гемоглобина из-за мутаций в геноме.

До недавнего времени получение ГМО было примитивным: создавали вирусные частицы, которые встраивали интересующий ген случайно в геном. В результате могут возникать вредные побочные эффекты, которые необходимо отсеивать. Потом нужно изучать метаболизм ГМО и постоянно следить за ним. Технологии типа TALEN и CRISPR позволяют направленно и точно изменять свойства организма, что экономит много времени.

7. Принципы работы секвенаторов нового поколения

Пиросеквенирование Этот метод секвенирования заключается в том, что в реакционную смесь добавляют матрицу ДНК, праймеры и ДНК-полимеразу, которая осуществляет синтез второй цепи ДНК. В реакции расходуются нуклеотидтрифосфаты, которые встраиваются в ДНК с образованием фосфодиэфирной связи и выделением дифосфата. Этот дифосфат регистрируют с помощью добавленной в реакционную смесь **люциферазы** (фермента, который, используя энергию люциферина и АТФ, дает вспышку света). В реакцию добавляют один определенный нуклеотид и, если он встроился, наблюдают вспышку света (или более яркую вспышку света при присоединении двух одинаковых нуклеотидов). Если же нуклеотид не встроился — света нет, как нет и комплементарного нуклеотида в последовательности. Таким образом, по свету от встраиваемых известных нуклеотидов можно определить всю последовательность нуклеотидов фрагмента ДНК. В секвенаторе в разных лунках происходят тысячи таких реакций, вспышки от которых фиксирует камера.

Технология секвенирования SOLiD Особенность технологии в использовании меченых олигонуклеотидов. В реакционную смесь с праймером и матричной ДНК добавляют специальные олигонуклеотиды (по 6–7 нуклеотидов) и лигазу, которая создает фосфодиэфирную связь между праймером и олигонуклеотидами, если они комплементарно отожделились (гибридизовались) на ДНК. Эти олигонуклеотиды имеют флуоресцентную метку, которая испускает свет определенной волны при комплементарном присоединении к ДНК. Это свечение от каждой ячейки регистрируется приборами. Таким образом, исходная последовательность устанавливается не по отдельным нуклеотидам, а по блокам олигонуклеотидов с помощью лигазы (а не с помощью полимеразы).

Технология секвенирования Illumina Эта технология основана на использовании ДНК-полимеразы и флуоресцентно меченых нуклеотидов. Принцип работы похож на пиросеквенирование: в лунках происходит

репликация ДНК и камера фиксирует вспышки света определенной длины волны от каждого типа включившегося в цепь нуклеотида.

Технология секвенирования IonTorrent При присоединении нуклеотида к растущей цепи ДНК во время репликации кроме пирофосфата выделяется еще и протон. В технологии IonTorrent его выделение регистрируют чувствительные рН-метры. Процедура секвенирования схожа с предыдущими: добавляют аденозинтрифосфат и наблюдают подкисление среды — аденозинтрифосфат включился в состав ДНК; в следующем цикле добавляют тимидинтрифосфат и не наблюдают изменения рН — в матричной цепи в конкретном месте нет аденина. Миниатюрные рН-метры располагаются в каждой реакционной ячейке на площадке.

Все эти технологии имеют ограничения по длине матрицы, поэтому перед запуском секвенатора исходную ДНК режут на мелкие фрагменты. В среднем длина таких фрагментов порядка 100 – 300 нуклеотидов, в зависимости от технологии число может варьировать.

Полученные после секвенирования последовательности фрагментов с помощью биоинформатических методов выравниваются на последовательность тотальной ДНК (для человека она уже известна). Какие-то участки останутся невыравненными относительно референсной последовательности, или совпадут не полностью — это значит, что в этих участках произошла мутация.

Технология секвенирования MinION Основным элементом секвенатора является нанопора с чувствительными сенсорами, через которую проходит ДНК. В зависимости от проходящего через нее нуклеотида, размер нанопоры изменяется. Эти изменения регистрируют сенсоры. Преимущество этого метода в том, что для его работы не нужны реактивы и ДНК перед секвенированием не нужно фрагментировать. Образцы этого секвенатора очень компактные.

РНК тоже можно отсеквенировать, переведя предварительно РНК в ДНК с помощью вирусного фермента обратной транскриптазы. Также у большинства РНК есть полиА-хвост. Это обстоятельство значительно облегчает подбор праймеров.

8. Регуляторные элементы

Как регуляторные элементы — **промоторы, энхансеры** и пр. — обнаружить в геноме? Промотор может находиться непосредственно перед геном, энхансеры и **сайленсеры** — перед, после и внутри гена. За счет гибкости ДНК регуляторные элементы могут взаимодействовать с промотором и регулировать транскрипцию гена. Регуляторные элементы можно обнаружить через взаимодействия с транскрипционными факторами с помощью метода иммунопреципитации.

Имунопреципитация — это метод выделения белка из смеси с использованием специфичных антител к нему. Для этого клетку обрабатывают фиксирующими агентами (например, формальдегидом). При этом фиксируются все молекулярные взаимодействия. Мембрану клетки разрушают, и, с помощью антител выделяют белки, связанные с ДНК. Антитела предварительно синтезируют специально для интересующего белка и закрепляют на магнитный носитель, за который их удобно доставать из раствора обратно. Далее определяют последовательность, с которой этой белок связался, и секвенируют ее.

При изучении регуляторных последовательностей важно помнить, что они специфичны для каждой ткани. Поэтому, чтобы изучить все сайты связывания одного транскрипционного фактора, необходимо рассмотреть их во всех клеточных типах, которых у человека больше 200. Методом иммунопреципитации можно определить, где сидел белок, но нельзя определить, какой ген он регулирует.

Как определить какой же ген он регулирует? Здесь помогут методы функциональной геномики. Можно, например, изменять экспрессию транскрипционного фактора в клетке и посмотреть, изменится ли экспрессия гена, который находился рядом или нет. Этот метод может быть неточен, так как один транскрипционный фактор может регулировать много генов. Тогда изменения в его экспрессии могут привести к драматическим трансформациям в клетке и колебания экспрессии интересующего гена может быть артефактом, не имеющим ничего общего с нормально функционирующей клеткой.

Перед интересующим геном почти всегда есть промоторная область, с которой связывается РНК-полимераза. На промотор оказывают воздействие (стимулирующее или супрессирующее) последовательности энхансеров и сайленсеров, которые могут находиться на большом расстоянии от промотора. За счет гибкости ДНК эти регуляторные элементы встречаются в пространстве и определенным образом воздействуют на экспрессию гена.

Недавно вышла интересная работа сингапурской лаборатории. В ней комплекс регуляторных элементов зафиксировали с помощью формальдегида. Далее его обработали ДНК-азой так, чтобы остались только белки, сидящие на обрывках ДНК. Обрывки остаются «торчащими» из белков, эти концы ДНК-аза не может расщепить из-за своей пространственной конфигурации. Далее комплекс обрабатывается лигазой. Это небольшой

фермент, который случайным образом сшивает между собой обрывки ДНК с образованием петель. Далее разрушают белки комплекса, а оставшуюся лигированную ДНК отдают на секвенирование. По полученной последовательности определяют какие регуляторные элементы клетки взаимодействовали между собой при жизни.

Современные методы секвенирования позволяют определить не просто структуру ДНК и мутации, но и ее регуляцию, или, в более широком смысле, эпигенетику. То есть секвенирование в исследованиях применяется как один из методов познания ДНК.