

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Вятский государственный университет»
Биологический факультет
Кафедра микробиологии

Рекомендовано к использованию
в учебном процессе
протокол заседания кафедры
№ _____ от _____
Заведующий кафедрой,
доктор медицинских наук, профессор
_____ И.В. Дармов

Е. А. Бессолицына

МОЛЕКУЛЯРНАЯ ГЕНЕТИКА

Модуль 2

Структура геномов

Конспект лекций

Учебно-методическое пособие

Киров 2011

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ
БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«ВЯТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

Биологический факультет

Кафедра микробиологии

Е. А. БЕССОЛИЦЫНА

МОЛЕКУЛЯРНАЯ ГЕНЕТИКА

Модуль 2

Структура геномов

Конспект лекций

Учебно-методическое пособие

Киров

2011

УДК

К

Допущено к изданию методическим советом биологического факультета ФГБОУ ВПО «ВятГУ» в качестве учебно-методического пособия для студентов направления 020400.62 «Биология» всех профилей подготовки, всех форм обучения

Рецензент

доцент кафедры биотехнологии ФГБОУ ВПО «ВятГУ»,

кандидат биологических наук

О. Н. Шуплецова

Бессолицына Е. А.

К

Молекулярная генетика конспект лекций модуль 2 «Структура геномов»: учебно-методическое пособие / Е. А. Бессолицына – Киров: ФГБОУ ВПО «ВятГУ», 2011. – 125 с.

УДК

Учебно-методическое пособие предназначено для студентов направления 020400.62 «Биология» всех профилей подготовки, всех форм обучения для изучения дисциплины «Молекулярная генетика».

Тех. редактор Е. В. Кайгородцева

© ФГБОУ ВПО «ВятГУ», 2011

Оглавление

Структура геномов прокариот	6
Структура Origin репликации.....	9
Структура промоторов	10
Структура терминаторов.....	12
Мобильные генетические элементы.....	17
Структура прокариотических повторов	18
Регуляция транскрипции у прокариот и структура регуляторных элементов прокариот ...	21
Регуляция транскрипции катаболитных оперонов.....	22
Репрессия катаболитных оперонов.....	24
σ -Субъединицы РНК-полимеразы	26
Аттенюаторы.....	29
Внехромосомная ДНК прокариот	32
Способы обмена генетической информацией у прокариот.....	33
Структура генома эукариот	35
Структура генома хлоропласта	35
Структура генома митохондрий	38
Организация митохондриального генома животных	38
Организация митохондриального генома грибов.....	40
Организация митохондриального генома простейших	42
Структура митохондриального генома у представителей отряда Kinetoplastida	44
Структура промоторов животных.....	48
Структура промоторов грибов	52
Структура промоторов растений.....	53
Транскрипция митохондриальной ДНК у различных групп организмов	56
Структура ядерного генома эукариот.....	57
Структура некодирующей ДНК у эукариот	63
Структура origin репликации	63
Структура элементов транскрипции.....	65
Регуляторные элементы генов, транскрибируемых РНК-полимеразой II.....	68
Энхансеры.....	75
Регуляция транскрипции, осуществляемой РНК-полимеразой III.....	80
Структура собственно некодирующей ДНК	84
Псевдогены	84
Повторы.....	85
Высокочастотные повторы	88
Сателлитная ДНК.....	89
Ферменты, участвующие в метилировании	92
Клеточные функции метилирования	93
Центромеры	98
Структурная вариабельность центромер	99
Теломеры	104
Строение теломер.....	107
Мобильные генетические элементы.....	116

Структура геномов прокариот

Геномную ДНК прокариот можно подразделить на два типа: хромосомная и внехромосомная.

Хромосомная ДНК образует нуклеоид.

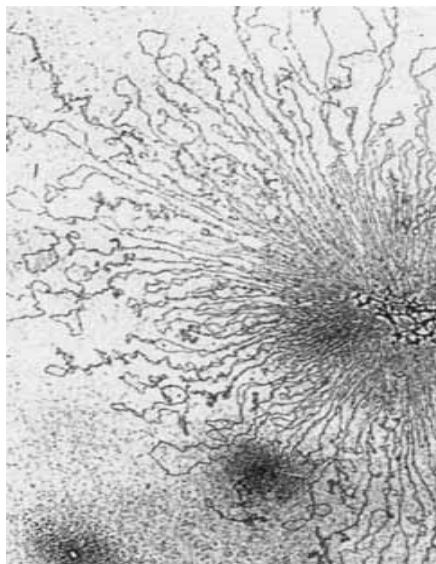


Рисунок 1 : нуклеоид *E. coli*

Полноценной хромосомой его сложно назвать, но в данной молекуле хранится основная часть генетической информации клетки необходимая для нормального ее функционирования. Хромосома прокариот кольцевая. не имеет мембранной оболочки. Хромосома бактерий суперспирализована и связана с гистоноподобными белками. Истинные нуклеосомы отсутствуют, количество белков, связанных с ДНК у прокариот значительно меньше, белок составляет 20% от общей массы нуклеоида, в связи с этим хроматин и

видимые хромосомы отсутствуют. ДНК нуклеоида так же связана с полиаминами. Идентифицированы три основных белка: белок HU конденсирует ДНК, возможно заворачивая ее в бусоподобные структуры, стимулирует репликацию ДНК; белок H1, связывается с ДНК, преимущественно с изогнутыми последовательностями, функция не известна; белок P сходе по структуре с полиаминами, однако функция не известна. Степень спирализации и «упаковки» ДНК меньше чем у эукариот, наблюдается 20 — 100 независимых суперспирализованных петель. Хромосома бактерий присоединена к клеточной мембране примерно на «экваторе» клетки.

Длина молекулы составляет 0,8 — $8 \cdot 10^6$ пар нуклеотидов. Структура самой молекулы ДНК сходна

назвать опероном.

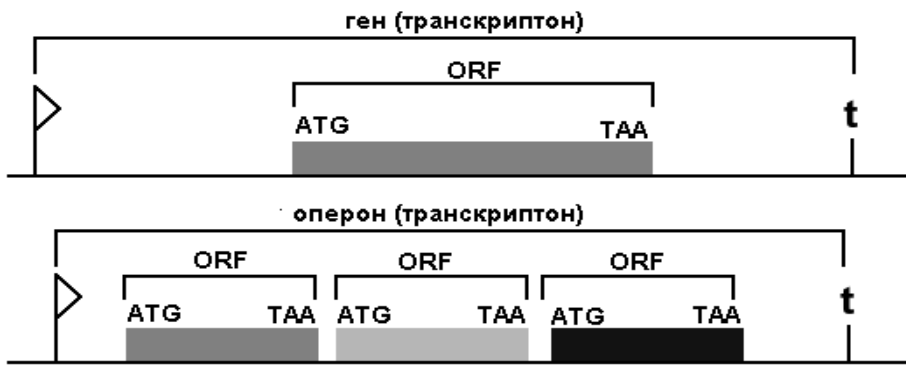


Рисунок 3: Схема организации транскриптов у прокариот и эукариот

Участки, кодирующие рРНК, объединены в единый оперон генов рРНК. В таком случае происходит транскрипция полицистронной РНК, которая затем

процессируется, это позволяет синтезировать эквимоллярные количества рРНК. ORF, кодирующие белки также объединены в опероны. В случае белок-кодирующих последовательностей объединение в транскриптоны происходит по функции. В один оперон объединяются ORF, кодирующие аминокислотные последовательности ферментов отвечающих за один метаболический путь. Например, в лактозном опероне закодированы белки расщепления лактозы, в триптофановом — ферменты биосинтеза этой аминокислоты. Такое объединение позволяет включать и выключать метаболические пути в зависимости от необходимости, но такая система менее лабильна, особенно при пересечении или перекрывании метаболических путей. Транспортные РНК так же организованы в опероны. Хотя существуют смешанные опероны, в которых закодированы как белки, так и тРНК.

Некодирующая часть ДНК прокариот включает в себя элементы отвечающие за репликацию, транскрипцию, регуляцию этих процессов, мобильные генетические элементы и собственно некодирующую ДНК.

Структура Origin репликации

Origin — это элемент, отвечающий за инициацию репликации. В

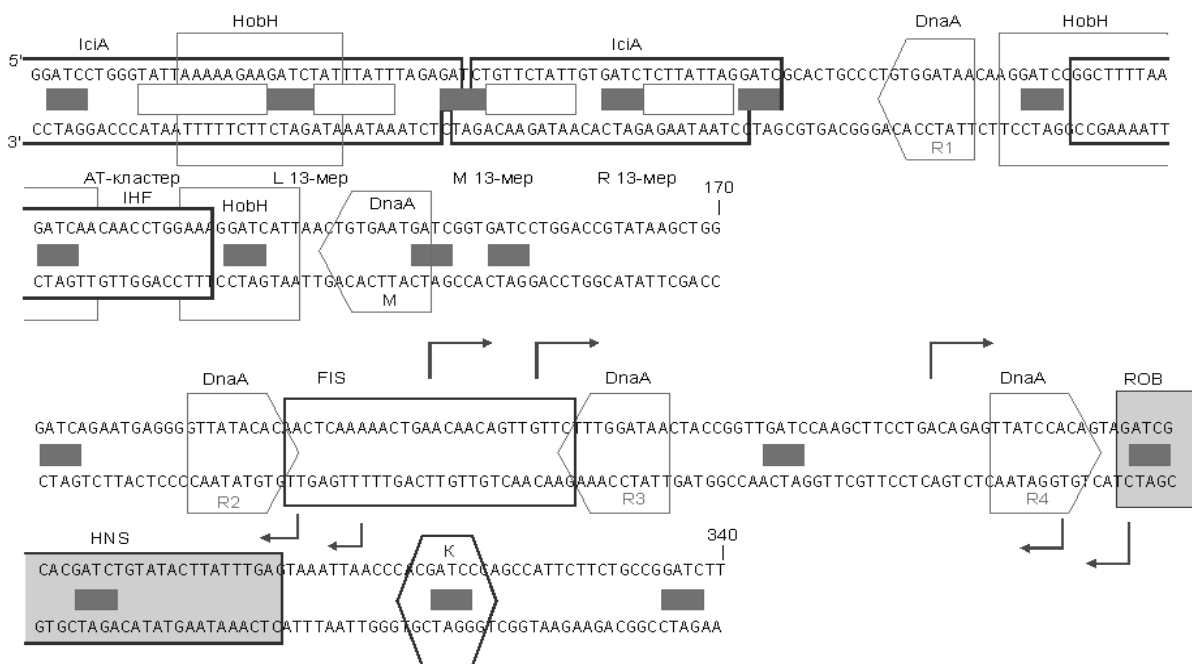


Рисунок 4: Структура области начала репликации хромосомы *E. coli* (а) Схема инициации репликации *E. coli* (б) HobH - белок, взаимодействующий с метилированной по одной цепи ДНК области начала репликации (hemimethylated origin binding)

бактериальной хромосоме располагается только один ориджин репликации. Места начала репликации ДНК (в дальнейшем origin) ДНК плазмид, фагов, бактерий и таких эукариотических репликонов, как ARS (автономно реплицирующиеся последовательности) у дрожжей или вирус SV40, имеют много общего в своей организации. Структура реплисомы представлена двумя доменами. В одном домене есть отдельная область, которая прочно связывается со специфическим для конкретного origin'a иницирующим белком, в другом домене - участок прямых повторов в количестве 2-5 штук длиной 6-16 нп, оба региона А+Т богаты. Инициация репликации начинается со связывания специфических мест с их инициаторными белками (места связывания тоже, как правило, представляют собой повторы, но не только прямые, например, как в origin'e pSC101) и расплавления в области прямых повторов. Скорее всего иницирующие белки, связывающиеся с сайтами связывания, вызывают дополнительное изгибание ДНК, во всяком случае это экспериментально обосновано для случая репликации R6K плазмиды, и, об этом свидетельствует и

сам вид ДНК-белкового "core", изначально наблюдаемый в электронный микроскоп. Как сайты связывания, так и прямые повторы погружены в 2 домена необычной структуры. Первый домен, окружающий сайты связывания, представляет изогнутую область ДНК, состоящую из А-трактов длиной 3-6 нп, которые, как известно, изгибают ось спирали для системы ДНК в растворе. Пространственное расстояние между ними близко к периоду спирали (10-11 нп) ДНК, что обеспечивает отклонение оси спирали в одном и том же направлении, приводя к макроскопическому изгибу. Второй домен, также А+Т богатый, окружающий повторы, имеет периодичность появления полиА трактов приблизительно в 6-8 пар. Этот промежуток приводит к несинхронному отклонению оси спирали от прямой по отношению к среднему периоду. Изгибы, вызванные такими повторами, в среднем уравнивают друг друга и, как следствие, приводят к отсутствию макроскопического изгиба в домене.

За транскрипцию отвечают две структуры: промоторы и терминаторы.

Структура промоторов

Основной элемент промотора — место связывания РНК-полимеразы, которое она занимает перед началом синтеза РНК. В состав промоторов могут входить также участки связывания белков-регуляторов. Размер участка связывания РНК-полимеразы соответствует ее длине и составляет примерно 70 п. н. Располагается этот участок относительно стартовой точки несимметрично: по ходу транскрипции его граница отстоит от стартовой точки на 20 п. н., а против хода — примерно на 50 п. н..

последовательностью, но искусственно сконструированный промотор с канонической последовательностью отличается очень высокой эффективностью (этот результат не был заранее очевиден: усредненная последовательность вполне могла бы обладать «средними» свойствами). О том, что каноническая последовательность является наиболее эффективной, свидетельствуют и результаты многочисленных данных по мутационным изменениям последовательности промоторов: изменения, приближающие последовательность промотора к канонической, как правило, увеличивают его силу, тогда как изменения, уменьшающие его сходство с канонической,— уменьшают его силу. Изменения нуклеотидной последовательности вне участков «—10» и «—35» обычно слабо сказываются на силе промотора. Знание этих закономерностей, однако, еще не позволяет надежно предсказывать силу промоторов и находить промоторы, рассматривая последовательность ДНК, хотя РНК-полимераза делает это очень быстро.

У разных промоторов расстояния между участками «—10» и «—35» могут быть не одинаковыми, у большинства промоторов оно составляет 16—18 п. н.; у небольшого числа промоторов вариации составляют 15—20 п. н. Оптимальное расстояние 17 п. н.: мутации, изменяющие это расстояние как в сторону увеличения, так и уменьшения, ослабляют промотор. Стартовая точка транскрипции отстоит у разных промоторов на 6—9 п. н. от последовательности Прибнова. Начальным нуклеотидом РНК чаще всего является диктуемый матрицей А или G. Предпочтение пуринов, однако, неабсолютно, и есть несколько промоторов, на которых транскрипты начинаются с С или U, несмотря на наличие на расстоянии одного-двух нуклеотидов потенциального пуринового начала. У некоторых промоторов стартовая точка не строго фиксирована: инициация происходит с двух-соседних нуклеотидов матрицы.

Структура терминаторов

Очевидная роль терминаторов транскрипции состоит в прекращении синтеза РНК в концах оперонов, что обеспечивает независимую регуляцию

экспрессии различных участков ДНК. Но терминаторы встречаются и внутри оперонов. Эффективность этих «внутренних» терминаторов может регулироваться, что позволяет клетке изменять скорость синтеза РНК на участках ДНК, расположенных за терминаторами, не меняя скорости синтеза РНК на участках, расположенных перед терминаторами. Прежде чем переходить к обсуждению функционального смысла и механизмов такой регуляции, в которой участвуют специальные белковые факторы и рибосомы, рассмотрим процесс терминации в отсутствие факторов.

В их отсутствие РНК-полимераза способна терминировать синтез РНК лишь на некоторых терминаторах, нуклеотидная последовательность в районе которых отличается двумя характерными особенностями. В них по ходу транскрипции сначала идет GC-богатый участок, обладающий центральной симметрией, а затем участок из 4—8 расположенных подряд А в значащей нити. Транскрипция заканчивается на конце олигоА последовательности или сразу за ней. Предполагается, что после прохождения РНК-полимеразой GC-богатого участка с центральной симметрией в РНК-продукте возникает шпилька, приводящая к остановке РНК-полимеразы и разрушению части РНК-ДНК гибрида транскрибирующего комплекса. Оставшаяся часть РНК-ДНК гибрида, содержащая концевую олигоU последовательность РНК» легко плавится ввиду относительной нестабильности rU-dA-пар, что приводит к освобождению РНК-продукта. Первым из комплекса освобождается РНК-продукт, а затем РНК-полимераза. В какой момент происходит схлопывание нитей ДНК. пока не известно.

Эффективность терминации зависит от прочности терминаторной шпильки в РНК. Это видно из того, что мутации, приводящие к нарушению комплементарного спаривания какого-либо основания в шпильке, ослабляют терминацию, а мутации, восстанавливающие комплементарность, усиливают терминацию. Усиление прочности РНК-ДНК гибрида в районе олигоU ослабляет терминацию. Такое усиление происходит при уменьшении длины олигоА в матрице (минимальный размер этой последовательности, при котором

терминация еще возможна, 4 нуклеотида).

Таким образом, эффективность терминации зависит, по-видимому, от баланса стабильности различных РНК: РНК и РНК : ДНК двойных спиралей. В энергетику этого баланса вносит вклад и РНК-полимераза, так как определенные мутации, затрагивающие РНК-полимеразу, влияют на эффективность терминации. Эти же мутации влияют и на продолжительность пауз транскрипции, что подтверждает представление о том, что подготовительной стадией терминации является пауза.

Терминация в отсутствие факторов, по-видимому, не полностью моделирует этот процесс в том виде, как он протекает в живой клетке. Действительно, на ряде терминаторов очищенная РНК-полимераза завершает

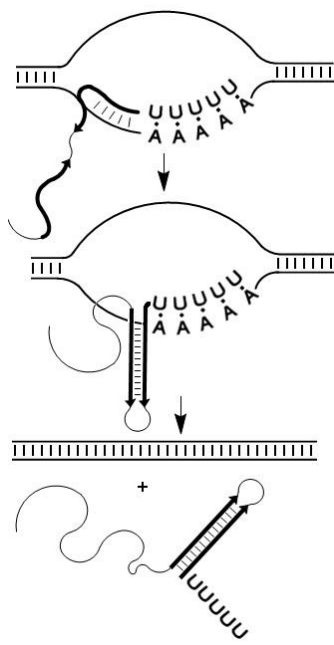


Рисунок 6: Схема ρ -независимой терминации

синтез молекул РНК, отступя на несколько нуклеотидов от того места, где она завершает его в живой клетке, и с меньшей эффективностью. Недавно обнаружен белковый фактор, названный тау(т), который улучшает эффективность и точность скорость синтеза РНК на участках ДНК, расположенных за терминаторами, не меняя скорости синтеза РНК на участках, расположенных перед терминаторами. Механизм его действия еще не изучен. Прежде чем переходить к обсуждению функционального смысла и механизмов такой регуляции, в которой участвуют специальные белковые факторы и рибосомы, рассмотрим процесс терминации в отсутствие факторов

Многие терминаторы узнаются РНК-полимеразой только с помощью фактора терминации, названного ρ . Этот белок с молекулярной массой 46 кД обладает РНК-зависимой нуклеозидтрифосфатазной активностью, обязательной для терминирующего действия. АТФазная активность фактора ρ проявляется только в комплексе с однонитевой РНК- Наибольшей АТФазной активностью ρ -фактор обладает в присутствии полицитидиловой кислоты. Фактор ρ

агрегирует с образованием гексамера, способного связываться с РНК. В комплексе с РНК гексамер защищает в ней от действия РНКазы 80—90 нуклеотидов, так что каждый мономер связывает по 12—14 нуклеотидов РНК. Фактор ρ присоединяется к РНК-продукту до того, как РНК-полимераза достигает терминатора. Присоединение происходит к определенным участкам РНК, в нуклеотидной последовательности которых пока не обнаружено каких-либо характерных особенностей. Ясно лишь, что эти участки не склонны к образованию протяженных двуспиральных структур.

В местах ρ -зависимой терминации РНК-полимераза делает паузы в отсутствие ρ -фактора, поэтому считается, что роль ρ -фактора заключается в вытеснении РНК из транскрипционного комплекса в местах пауз. Рассматриваются две модели. Согласно одной из них, фактор движется по синтезируемой РНК, а в местах пауз догоняет РНК-полимеразу и вытесняет РНК-продукт. Другая модель основана на том, что пирофосфат подавляет АТФазную активность ρ -фактора. Согласно этой модели, ρ -фактор движется за РНК-полимеразой без отставания, но при нормальной скорости элонгации ингибируется пирофосфатом, высвобождающимся при синтезе РНК. Активация ρ -фактора происходит лишь в местах пауз, где синтез цепи РНК временно останавливается, что приводит к прекращению освобождения пирофосфата.

Конкретный механизм вытеснения РНК из транскрипционного комплекса под действием ρ -фактора еще не выяснен. Это вытеснение сопряжено с гидролизом АТФ, при котором, по-видимому, происходит конформационное изменение ρ -фактора. В результате этого изменения РНК вытесняется из комплекса либо непосредственно ρ -фактором, либо за счет воздействия на РНК-полимеразу.

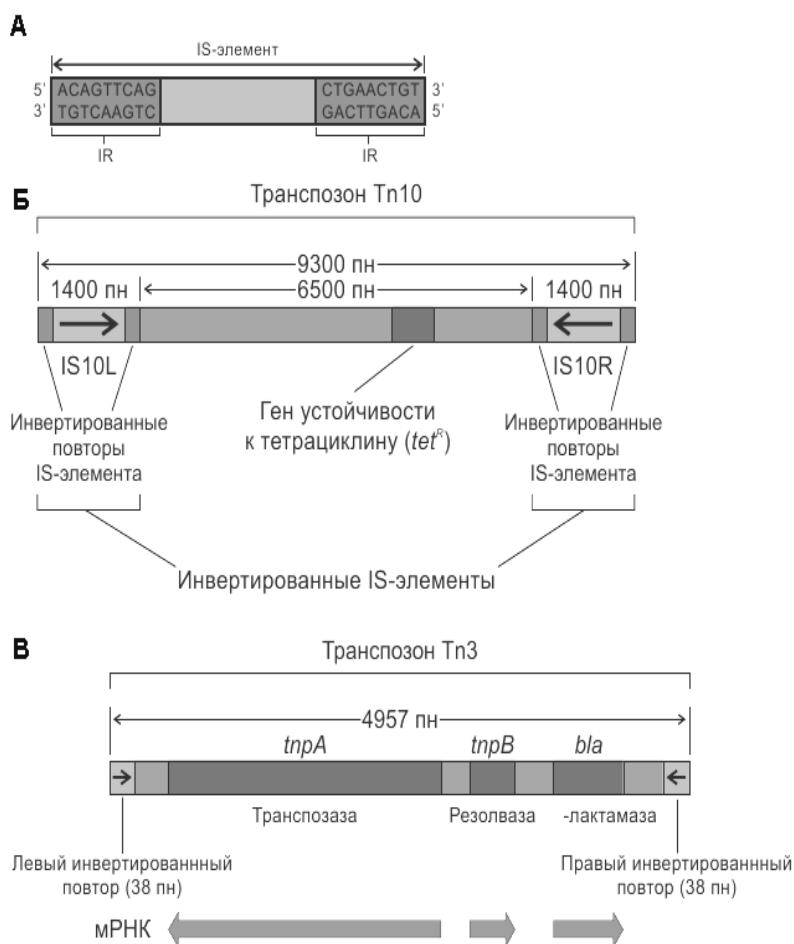
Однако ρ -фактор вызывает терминацию не во всех местах пауз. Анализ нуклеотидной последовательности показывает, что начальная часть РНК способна образовывать большое число двунитчатых структур. По-видимому, в данном случае сильно развитая вторичная структура транскрипта мешает связыванию с ним ρ -фактора, без которого терминации в местах пауз не

происходит.

Связыванию ρ -фактора с РНК мешают также рибосомы, транслирующие РНК, поэтому на ρ -зависимых терминаторах, встречающихся внутри структурных генов, ρ -фактор не обеспечивает терминации, если мРНК эффективно транслируется. Наличие таких терминаторов внутри генов, по-видимому, не случайно. Когда синтез белка по каким-либо причинам подавлен, они сигнализируют РНК-полимеразе о том, что мРНК не транслируется и синтез бессмысленен.

Регуляторные элементы также относятся к некодирующей ДНК, их структура будет рассмотрена ниже.

Мобильные генетические элементы



IS-элементы (от англ.

insertion sequences —

последовательности-

вставки) — это сегменты ДНК, способные как целое

перемещаться из одного

участка локализации в

другой. IS-элементы

содержат лишь те гены, которые необходимы для их

собственного перемещения

— *транспозиции*. Кроме

того, IS-элементы имеют

особую последовательность

на концах, как правило,

инвертированные повторы.

При встраивании в новую

последовательность ДНК

Рисунок 7: Структура мобильных генетических элементов прокариот. А - IS-элементы; Б, В -

IS-элементы вызывают небольшую дупликацию: дублированный участок с

двух сторон фланкирует встроившийся IS-элемент

Транспозонами (Tn-элементами) называют сегменты ДНК, обладающие теми же свойствами, что и IS-элементы, но содержащие, кроме того, гены, не имеющие непосредственного отношения к транспозиции. Транспозоны могут нести гены устойчивости к антибиотикам, гены токсинов или гены дополнительных ферментов клеточного метаболизма. У бактерий выделено два типа транспозонов, различающихся по типу транспозиции. Мобильные генетические элементы способны к перемещению по хромосомной и внехромосомной ДНК. Их нельзя назвать полностью некодирующими, так как в их составе присутствуют ORF, кодирующие белки транспозиции. Однако это можно сказать, что это эгоистическая ДНК, саморазмножающаяся в геноме. Хотя в составе транспозонов могут располагаться и гены, кодирующие белки, повышающие выживаемость хозяина, например, факторы устойчивости к антибиотикам.

Также к мобильным генетическим элементам относятся вирусы способные на лизогенный жизненный цикл, такие как фаги λ , μ и др.

Собственно некодирующая ДНК образована межгенными промежутками и повторами. Межгенные промежутки не содержат информации, только разделяют кодирующие фрагменты. Эти спейсеры не имеют какой-либо четкой структуры.

К некодирующей ДНК также относятся повторы.

Структура прокариотических повторов

Геномы бактерий различных видов (а иногда даже штаммов) отличаются по наличию или отсутствию повторов определённых типов, и по количеству повторов в геноме. Плотность повторов обратно пропорциональна размеру генома. Геномы бактерий содержат преимущественно короткие tandemные повторы, такие как семейство "variable number of tandem repeats" (VNTR), семейство SSR и близкие повторы - семейство CR.

Несмотря на максимальную компактность бактериальных геномов и априорную направленность селекции против сохранения прямых повторов,

практически не существует геномов, в которых вообще не было бы повторов. Достаточно сказать, что даже один из самых маленьких – геном *Nanoarchaeum equitans* содержит незначительное количество не кодирующих повторов.

Повторы могут быть диспергированными, т.е. относительно беспорядочно разбросанными по длине молекулы ДНК, или организованными в определённые структуры. К последним относятся повторы, входящие в состав мобильных элементов, сайтов интеграции геномов умеренных бактериофагов, сайтов принадлежащих интегронам или сцепленных с генами тРНК. Будем называть их «организованные повторы». Обычно общая схема строения таких повторов выглядит следующим образом: левый элемент повтора – уникальная часть – правый элемент повтора. В подавляющем большинстве случаев элементы, построенные по такой схеме, содержат повторы в обратной ориентации (IR). Исключение составляет, например, Tn9. Как диспергированные, так и организованные повторы, играют первостепенную роль в процессах рекомбинации.

Принципиально одинаковые структуры, содержащие повторы, могут находиться как в мобильном (донорном) элементе, так и в реципиентной молекуле ДНК. В зависимости от присутствия организованных повторов в реципиентном геноме или в донорном сегменте можно разделить их типы взаимодействия на классы.

Класс I. Организованные повторы присутствуют только в структуре донорного сегмента ДНК. В этом случае специфичность интеграции низкая, т.е. количество сайтов интеграции большое и выбор сайта интеграции относительно независим от структуры реципиентной молекулы ДНК. Примеры: IS-элементы, транспозоны. (Исключение составляет интеграция Tn7 в хромосому *E.coli.*).

Класс II. Организованные повторы присутствуют только в структуре реципиентного сегмента ДНК. В этом случае специфичность интеграции относительно высокая, т.е. количество сайтов интеграции небольшое и зависит от структуры реципиентной молекулы ДНК. Примеры: интегроны, сайты

интеграции, сцепленные с генами tPHK, возможно DUS-элементы.

Класс III. Организованный повтор присутствует в структуре как донорного сегмента, так и реципиентной ДНК. При этом, донорный и реципиентный повторы коэволюционировали и имеют структурное родство. В этом случае специфичность интеграции абсолютная, т.е. существует только один сайт интеграции в реципиентном геноме. Примеры: интеграция геномов умеренных бактериофагов.

Таким образом, видно, что присутствие организованных повторов в донорном сегменте или реципиентной структуре принципиально определяет результат взаимодействия этих структур. Конечно, количество сайтов интеграции внешней ДНК в геном реципиента и локализация этих сайтов имеют безусловное значение для эволюционного процесса.

Теперь рассмотрим, какое значение для геномных перестроек имеют диспергированные повторы. Эти повторы разнообразны по нуклеотидным последовательностям, размеру, количеству копий и самому наличию в геномах бактерий. Повторы являются горячими точками геномных перестроек, вовлекающих крупные сегменты генома. Такими перестройками могут быть инверсии, делеции, дупликации и транслокации. Обычно прямые повторы стимулируют делеции, дупликации и транслокации, а инвертированные повторы – инверсии.

Диспергированные повторы могут также служить и сайтами интеграции для мобильных элементов. Так, предпочтительными сайтами узнавания для транспозаз в геномах микроорганизмов различных видов оказались внегенные повторяющиеся палиндромы REP (repetitive extragenic palindromes). Тем не менее, специфичность интеграции IS- элементов и большинства транспозонов достаточно низкая.

Роль повторов в возникновении геномных перестроек сейчас доказана, в том числе в исследованиях по сравнительной геномике близкородственных микроорганизмов. Сравнивали нуклеотидные последовательности геномов пар *Mycobacterium leprae* - *Mycobacterium tuberculosis*, *Bordetella pertussis* -

Bordetella bronchiseptica, *Bartonella quintana* - *Bartonella henselae*, *Burkholderia mallei*-*Burkholderia pseudomallei*. Отличия между геномами членов пар во многих случаях были представлены делециями, транслокациями и инверсиями. На границах сегментов ДНК, претерпевших указанные перестройки, всегда обнаруживались те или иные повторы, IS-элементы, гены тРНК, части геномов профагов. Та же закономерность имеет место у *Y.pestis*: делеция 102 т.н.п. локуса *pgm* происходит с высокой частотой (10^{-4}) за счёт рекомбинации между двумя копиями фланкирующих элементов IS100.

Тип и локализация в геноме повторов различных типов определяют вероятность, характер и локализацию геномных перестроек. При этом, не только локализация, но и само присутствие IS-элементов в геномах не случайно. Например, максимальный размер повторов колеблется от 29 н.п. у *Buchnera aphidicola* до 14 317 н.п. у *Xylella fastidiosa*, а плотность повторов на 10^6 н.п. – от 12 у *Chlamydia trachomatis* до 5,069 у *Streptococcus pneumoniae*. Различная плотность повторов определяет и различную длину участка ДНК, который может быть вовлечён в рекомбинационный акт. Средняя длина такого участка варьирует от 185 н.п. у *C. trachomatis* до 1,024,894 н.п. у цианобактерий (*Nostoc* sp.). Известно, что локализация, количество и тип повторов в геномах в существенной мере видоспецифичны, соответственно, характер и частоты геномных перестроек являются относительно видоспецифичными.

Все эти элементы участвуют в реализации генома. Основной процесс реализации — транскрипция. Этот процесс регулируемый.

Регуляция транскрипции у прокариот и структура регуляторных элементов прокариот

Активность многих промоторов регулируется с помощью особых белков-регуляторов, которые присоединяются к определенным участкам ДНК и либо мешают, либо помогают РНК-полимеразе инициировать синтез РНК- В первом случае говорят о *негативной*. во втором — о *позитивной* регуляции активности промотора. Белки, осуществляющие негативную регуляцию, называются *репрессорами*. Места их связывания на ДНК называются

операторами. Способность многих репрессоров связываться со своими" операторами зависит от низкомолекулярных лигандов — *эффекторов*. Эффекторы, снижающие сродство репрессора к оператору, называются *индукторами*. В отсутствие индуктора репрессор связывается с оператором и мешает РНК-полимеразе начинать синтез РНК с промотора (промотор репрессирован). В комплексе с индуктором репрессор теряет способность связываться с оператором, в результате чего промотор активируется (индуцируется). Другие репрессоры, наоборот, могут связываться с оператором только в комплексе с эффектором (который в этом случае называется *корепрессором*). В присутствии корепрессора промотор неактивен (репрессирован), в отсутствие корепрессора активируется (дерепрессируется). Белки, осуществляющие позитивную регуляцию, называются *активаторами*. Ряд белков-регуляторов могут выступать как в роли репрессора, так и в роли активатора.

Простейший механизм репрессии заключается в стерическом блокировании репрессором присоединения РНК-полимеразы к промотору. Такой механизм имеет место в тех промоторах, в которых участок связывания репрессора перекрывается с участком связывания РНК-полимеразы. Простейший механизм активации заключается в том, что белок-активатор присоединяется к промотору рядом с РНК-полимеразой и за счет непосредственного контакта с ней облегчает образование открытого промоторного комплекса. Дискуссионными являются механизмы действия тех белков-регуляторов, которые присоединяются к ДНК на значительном расстоянии от РНК-полимеразы. Ниже рассмотрено несколько наиболее хорошо изученных, примеров, иллюстрирующих различные принципы регуляции промоторов.

Регуляция транскрипции катаболитных оперонов

Белок-активатор катаболитных оперонов (БАК) в комплексе с

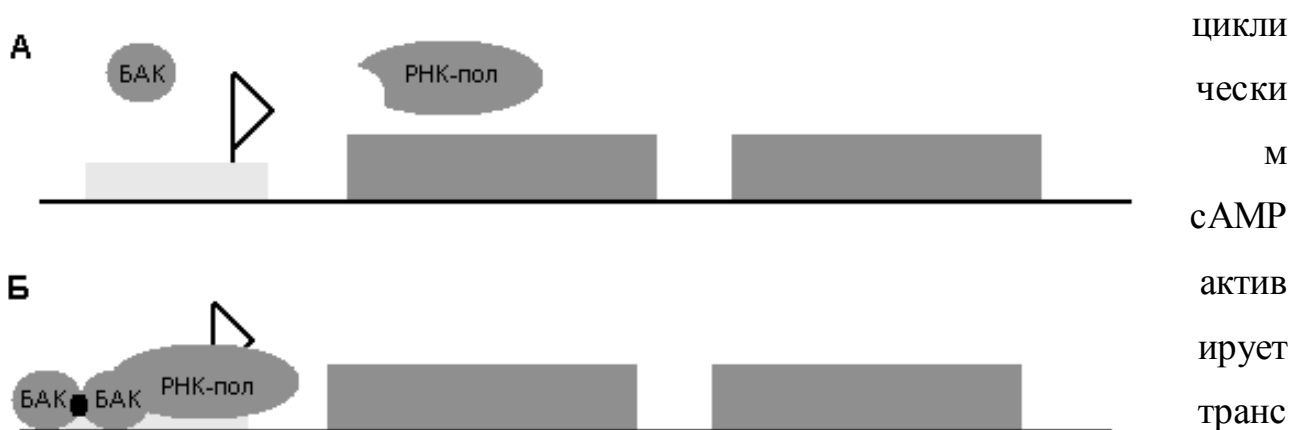


Рисунок 8: Схема позитивной регуляции катаболитных оперонов

цикли
чески
м
сАМР
актив
ирует
транс
крипц
ию

большого числа оперонов, отвечающих за расщепление различных соединений, преимущественно сахаров, используемых бактериальной клеткой в качестве источников энергии и углерода. Концентрация сАМР в клетках повышается при росте на плохо усваиваемых источниках, например ацетате или глицерине, и снижается при росте на легко усваиваемых, например глюкозе. Поэтому система регуляции с помощью БАК-сАМР позволяет клетке включать опероны катаболизма лишь по мере истощения более легко усваиваемых пищевых веществ. БАК состоит из двух идентичных субъединиц. Каждая субъединица образует два домена связывания с ДНК: С-концевой домен. N-концевой домен БАК связывается с сАМР и обеспечивает межсубъединичные контакты в димере. Связывание с ДНК происходит за счет пары биспиральных элементов, погруженных в большие бороздки В-формы ДНК. Особенностью взаимодействия БАК с ДНК является то, что он изгибает ДНК в месте своего присоединения к ней.

Участки связывания БАК в разных промоторах не идентичны по нуклеотидной последовательности, но проявляют значительное сходство друг с другом. Эти участки в разных промоторах располагаются по-разному относительно стартовых точек транскриптов, которые стимулируются БАК. В промоторе галактозного оперона участок связывания БАК непосредственно прилегает к участку связывания РНК-полимеразы. Рядом располагается второй

участок связывания БАК; с ним БАК связывается только в том случае, если на промоторе уже находится РНК-полимераза и БАК на первом участке. В промоторе лактозного оперона имеется только один участок связывания БАК, ориентированный относительно участка связывания РНК-полимеразы точно так же, как второй участок связывания БАК в галактозном промоторе. В промоторе гена, кодирующего белок-активатор мальтозных оперонов, и в промоторах мальтозных и арабинозного оперонов участок связывания отстоит от стартовой точки транскрипции еще дальше. В промоторах мальтозных и арабинозного оперонов, между участками связывания БАК и участками связывания РНК-полимеразы, располагаются участки связывания белков-активаторов этих оперонов, кодируемых генами *malT* и *agaC* соответственно. Механизм действия БАК не вполне понятен.

БАК может выступать также и в роли репрессора транскрипции. Например, в галактозном опероне кроме стимулируемого БАК промотора P_1 имеется репрессируемый БАК промотор P_2 . Эти два промотора перекрываются друг с другом, так что присоединение одной молекулы РНК-полимеразы к промотору P_2 препятствует присоединению другой молекулы РНК-полимеразы к P_1 . Присоединение БАК к ДНК мешает связыванию РНК-полимеразы с P_2 и не мешает связыванию с P_1 . Поэтому БАК оказывает не только прямое, но и опосредованное активирующее действие на промотор P_1 . Блокирование промотора P_2 приводит к усилению транскрипции с P_1 так как обеспечивает беспрепятственное связывание РНК-полимеразы с P_1 .

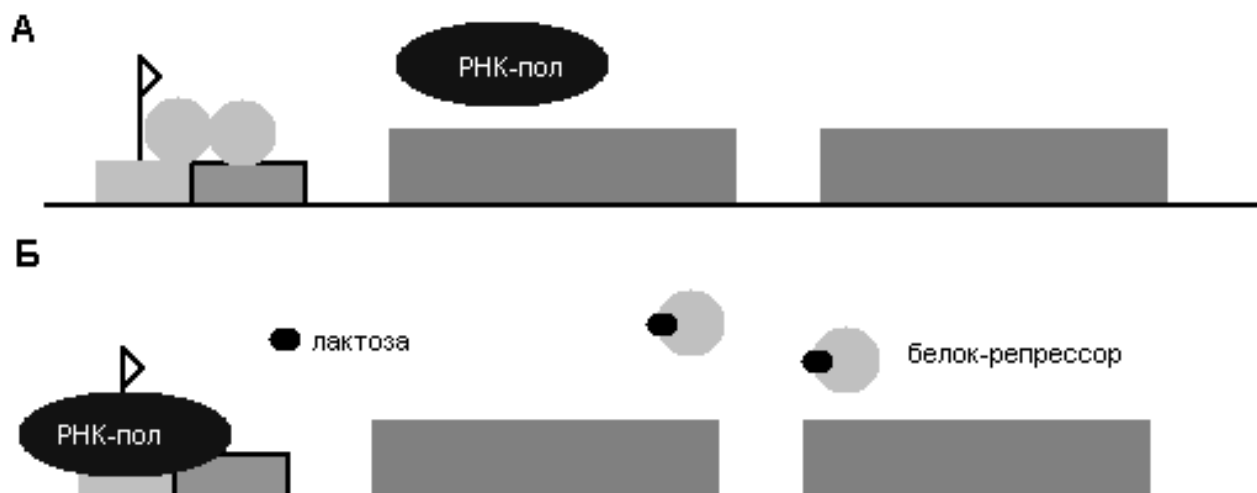


Рисунок 9: Схема негативной регуляции лактозного оперона

Репрессия катаболитных оперонов

Помимо общей регуляции с помощью БАК-сАМР существует индивидуальная регуляция катаболитных оперонов. Классическим примером является негативная регуляция лактозного оперона. В отличие от ранее рассмотренных димерных белков-регуляторов репрессор лактозного оперона представляет собой тетрамер и содержит два идентичных центра связывания ДНК- Пространственная структура этих центров формируется N-концевыми участками полипептидных цепей, которые, судя по их аминокислотной последовательности, способны образовывать биспиральные элементы, аналогичные биспиральным ДНК-узнающим элементам БАК. С -концевые домены субъединиц лактозного репрессора формирует два центра связывания индуктора лактозного оперона. Оператор лактозного оперона располагается сразу за стартовой точкой транскрипции. Долгое время считалось, что присоединение лактозного репрессора к промотору стерически мешает присоединению РНК-полимеразы. Однако недавно получены данные, свидетельствующие о том, что репрессор и РНК-полимеразы могут расположиться на промоторе рядом друг с другом. Поэтому приходится думать о более изощренных механизмах репрессии, включающих специфические контакты репрессора с РНК-полимеразой. В лактозном опероне имеется два псевдооператора, сходных по нуклеотидной последовательности с оператором, но обладающих меньшим сродством к репрессору. Один из этих

псевдооператоров располагается внутри гена, кодирующего лактозный репрессор, на расстоянии 92 п. н. от главного оператора, другой — на расстоянии 401 п. н. внутри гена β -галактозидазы. Наличие псевдооператоров несколько увеличивает сродство репрессора к оператору. Стабилизация комплекса репрессора с оператором может происходить за счет того, что одна молекула репрессора, имеющая два центра связывания ДНК может одновременно присоединить к себе и оператор, и псевдооператор. Разделяющая их ДНК при этом образует петлю.

Два оператора имеется так же и в галактозном опероне. Один из них располагается в районе —60 п. н. промотора, другой — в районе +55. Показано, что связывание репрессора с операторами не мешает связыванию БАК и РНК-полимеразы с промотором. Поскольку для эффективной репрессии нужны оба оператора, предполагается, что молекулы репрессора, расположенные на операторах, взаимодействуют друг с другом, образуя петлю ДНК- Такая конформация каким-то образом мешает инициации транскрипции.

Образование петель постулировано и при репрессии арабинозного оперона *araBAD*. Репрессором этого оперона является белок, кодируемый геном *araC*. В отсутствие арабинозы АгаС-белок, являющийся димером, репрессирует *araBAD*-оперон, а в присутствии арабинозы превращается в активатор, который активирует этот оперон. Кроме того, АгаС-белок как в присутствии, так и в отсутствие арабинозы умеренно репрессирует транскрипцию своего собственного гена, в результате чего концентрация АгаС-белка поддерживается на постоянном уровне.

Ген *araC* располагается перед началом *araBAD*-оперона, но транскрибируется в противоположном направлении.

σ -Субъединицы РНК-полимеразы

σ -Субъединицы РНК-полимеразы можно рассматривать как белки-регуляторы, включающие определенные группы генов. От других белков-активаторов σ -субъединицы отличаются лишь в деталях механизма действия: если «обычные» белки-активаторы сначала присоединяются к промотору, а

затем взаимодействуют с РНК-полимеразой, то σ -субъединица сначала присоединяется к РНК-полимеразе, а затем присоединяется к промотору как часть ДНК-связывающего центра фермента.

Впервые включение генов за счет появления в клетке новых σ -субъединиц было продемонстрировано при таком запрограммированном во времени необратимом процессе, как развитие бактериофагов. Другим запрограммированным во времени необратимым процессом, при котором происходит последовательное включение и выключение больших групп генов с участием различных σ -субъединиц, является споруляция таких бактерий, как *Bac. subtilis*.

σ -Субъединицы участвуют также в регуляции ряда промоторов, активность которых обратимо реагирует на изменения внешних условий» например на тепловой шок. Реакция теплового шока характерна для всех изученных организмов, как эу-, так и прокариот. При резком повышении температуры происходит значительное (но кратковременное) ускорение синтеза определенных белков за счет усиления транскрипции их генов. Промоторы соответствующих генов у *E. coli* отличаются по нуклеотидной последовательности от промоторов, используемых РНК-полимеразой, содержащей главную σ^{70} -субъединицу. Транскрипцию с промоторов, активируемых при тепловом шоке, осуществляет РНК-полимераза, содержащая особую σ^{32} -субъединицу с $M_2=32$ кД. Молекулы σ^{32} отличаются коротким временем жизни в клетке, т. е. расщепляются протеазами через несколько минут после своего образования. Количество σ^{32} в клетке при тепловом шоке возрастает за счет неизвестного еще механизма, что является причиной ускорения транскрипции генов белков теплового шока.

Особая σ -субъединица участвует в транскрипции ряда генов, ответственных за метаболизм азота. К ним относятся ген, кодирующий глутаминсинтетазу, и гены, контролирующие фиксацию атмосферного азота. Промоторы этих генов не содержат обычных для других промоторов последовательностей «—10» и «—35». Вместо них имеются участки гомологии,

центры которых расположены в положениях «—11» и «—21». Поэтому неудивительно, что эти промоторы не используются РНК-полимеразой, содержащей главную сигма-субъединицу, σ^{70} . Транскрипцию этих промоторов обеспечивает одна из минорных σ -субъединиц, σ^{60} , кодируемая геном *proN*. Однако для функционирования промотора гена глутаминсинтетазы белка σ^{60} недостаточно. Необходим еще ДНК-связывающийся белок, называемый NR1. Перед промотором имеется пять участков его связывания; наибольшее сродство NR1 проявляет к двум отдаленным участкам. Эти последовательности необходимы для активации промотора при низких концентрациях NR₁ и не обязательны при высоких. Если эти последовательности отодвинуть на тысячу пар нуклеотидов от промотора, они продолжают обеспечивать активность промотора. Предполагается, что белок NR1 взаимодействует с РНК-полимеразой, расположенной на промоторе. Посадка NR1 на ДНК облегчает это взаимодействие, сопровождаемое, по-видимому, образованием петли ДНК-ДНК. Промотор гена глутаминсинтетазы замечателен не только тем, что он регулируется с участием минорной сигма-субъединицы и нуклеотидных последовательностей, удаленных на большие расстояния от старта транскрипции, но и тем, что действие регуляторного белка модулируется не путем связывания лигандов-эффекторов, которыми могли бы быть глутамин или глутаминовая кислота, а путем химической модификации — фосфорилирования и дефосфорилирования NR₁ — осуществляемой несколькими ферментами, реагирующими на обеспеченность клетки источниками азота.

Аттенюаторы

Регулируемые терминаторы бактерий называют *аттенюаторами* (ослабителями). Впервые обнаружен и лучше других изучен аттенюатор триптофанового оперона *E. coli*. Этот оперон состоит из пяти генов, кодирующих ферменты биосинтеза триптофана. Регуляцию осуществляют две системы, чувствующие потребность клетки в триптофане. Первая система влияет на эффективность инициации на промоторе оперона. Репрессор триптофанового оперона в комплексе с триптофаном присоединяется к

оператору, расположенному перед стартовой точкой транскрипции в районе «— 10» и стерически препятствует РНК-полимеразе присоединиться к промотору. Таким образом, при избытке триптофана оперон репрессирован. В отсутствие триптофана репрессор теряет способность связываться с оператором, в результате чего оперон индуцируется. Эту систему дополняет регуляция в аттенуаторе, расположенном на расстоянии 180 п. н. от стартовой точки транскрипции внутри «лидерной» последовательности, предшествующей иницирующему кодону первого структурного гена. В условиях избытка триптофана лишь одна из десяти молекул РНК-полимеразы, начавших синтез РНК на триптофановом промоторе, преодолевает этот терминатор и переходит в

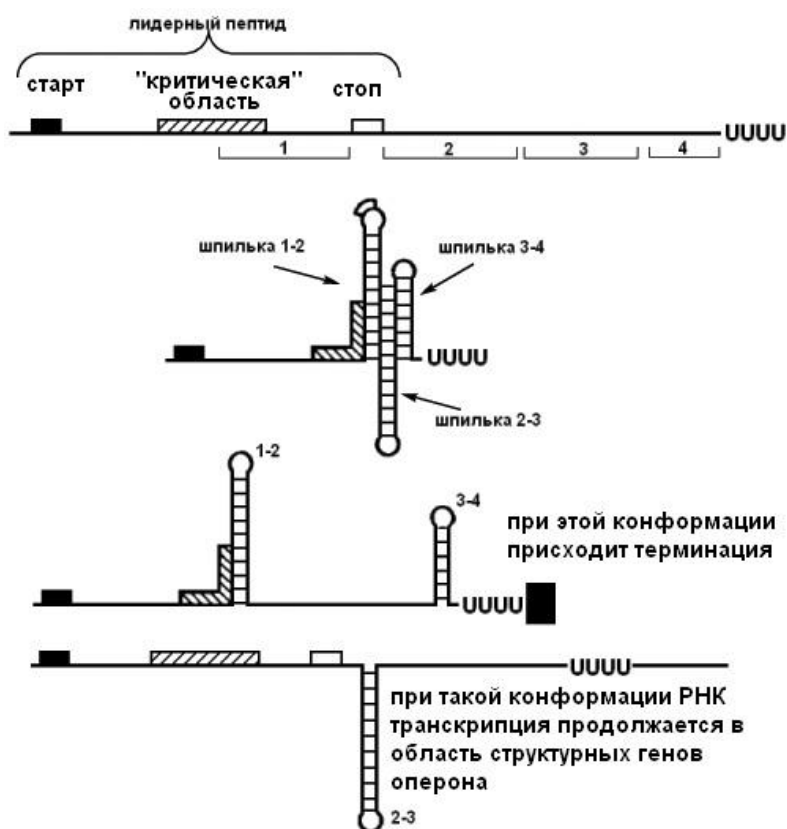


Рисунок 10: Схема регуляции транскрипции с помощью аттенуаторов

область структурных генов.

При уменьшении количества триптофана доля молекул РНК-полимеразы,

преодолевающих аттенуатор,

возрастает. Перед аттенуатором располагается несколько областей с

центральной симметрией. Образующаяся на них лидерная РНК содержит

четыре участка, способных образовывать шпильки в различных сочетаниях. Могут образоваться шпильки 1 : 2 и 3 : 4; шпилька 2 : 3 при этом

образоваться не может. Если образуется шпилька 2 : 3, то не могут образовываться шпильки 1 : 2 и 3 : 4. Шпилька 3 : 4 представляет собой терминаторную шпильку, характерную для обычных р-независимых терминаторов, так как после нее в РНК следует несколько U подряд. Поэтому

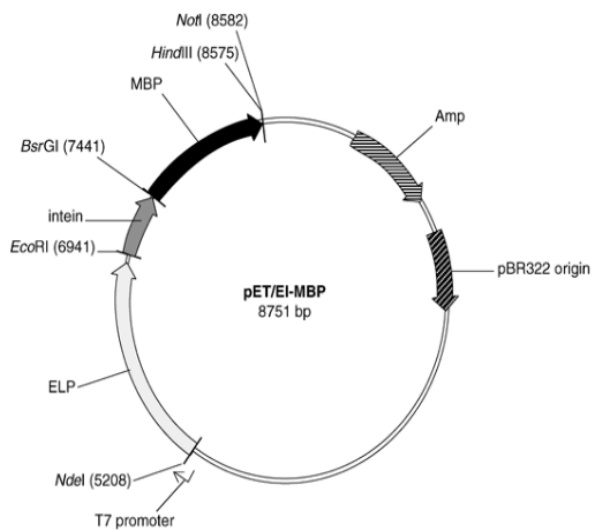
при образовании этой шпильки в аттенуаторе происходит терминация. Эффективность ее составляет около 80—90 %. Если же шпильки 3 : 4 не образуется (за счет образования шпильки 2 : 3), то терминация не происходит.

Образование альтернативных шпилек в лидерной области, от которых зависит терминация в аттенуаторе, определяется тем, как транслируется лидерная РНК. Эта РНК кодирует так называемый лидерный пептид, содержащий два триптофановых остатка подряд. Если этот пептид не транслируется, то вторичная структура РНК формируется по мере ее роста и выхода из транскрипционного комплекса: сначала формируется шпилька 1 : 2, что делает невозможным образование шпильки 2 : 3, а затем шпилька 3 : 4, что приводит к терминации синтеза РНК. Если в клетке имеется недостаток триптофана, то рибосома, транслирующая лидерный пептид, задерживается на его триптофановых кодонах. Эти кодоны расположены так, что, находясь на них, рибосома прикрывает последовательность 1 и мешает образованию шпильки 1 : 2. В результате становится возможным образование «антитерминаторной» шпильки 2 : 3. Терминация не происходит, и РНК-полимераза переходит в область структурных генов.

Сходная система регуляции в аттенуаторе используется в оперонах, отвечающих за синтез других аминокислот у *E. coli*. В лидерных РНК этих оперонов закодированы специфические пептиды, включающие по несколько остатков той аминокислоты, биосинтез которой определяется опероном. Если в лидерном пептиде триптофанового оперона содержится два tandemно расположенных остатка триптофана, то в лидерном пептиде фенилаланинового оперона — последовательно семь остатков фенилаланина, в которые вклинивается один остаток аланина и один остаток треонина, в лидерном пептиде гистидинового оперона — подряд семь остатков гистидина, в лидерном пептиде лейцинового оперона — четыре остатка лейцина подряд к т. д. При отсутствии соответствующей аминокислоты рибосома задерживается на ее кодонах и не обеспечивает образование «антитерминаторной» шпильки.

Внехромосомная ДНК прокариот

К внехромосомной ДНК относятся вирусы, чей жизненный цикл идет по литическому пути, а так же плазмиды. Данная группа вирусов называется бактериофагами, если их жизненный цикл идет по лизогенному пути, то они встраиваются в хромосомную ДНК, и входят в состав некодирующей ее части. Существует большое число групп бактериофагов, но структура их геномов подробно рассматривается в курсе вирусологии.



Плазмиды содержат от 1500 до 40000 пар нуклеотидов. Большинство плазмид состоит из трех групп генов: участка ДНК, ответственного за автономную репликацию плазмиды в клетке; системы генов, обеспечивающих возможность переноса плазмид из одной клетки в другую; генов, определяющих свойства, полезные для клетки-хозяина.

Рисунок 11: Структура плазмиды Отличительная особенность плазмид — способность к автономной репликации.

Плазмиды реплицируются автономно от хромосомной ДНК, но структура origin репликации в плазмиде и хромосомной ДНК сходны, так как удвоение ДНК осуществляется одним набором ферментов.

Плазмиды присутствуют в клетке в нескольких копиях. Количество копий зависит от плазмиды, выделяют низко-, средне- и высококопийные плазмиды.

Плазмиды содержат структурные гены, наделяющие бактериальную клетку разными, весьма важными для нее свойствами:

- R-плазмиды — лекарственной устойчивостью;
- Col-плазмиды — способностью синтезировать колицины;
- F-плазмиды — передавать генетическую информацию;
- Шу-плазмиды — синтезировать гемолизин;
- Тох-плазмиды — синтезировать токсин;

- плазмиды биодegradации — разрушать тот или иной субстрат и т. д.

Наличие F-плазмиды (фактор фертильности, половой фактор) придает бактериям функции донора, и такие клетки способны передавать свою генетическую информацию другим, F-клеткам. Наличие F-плазмиды является фенотипическим выражением (проявлением) пола у бактерий: с F-плазмидой связана не только донорская функция, но и некоторые другие фенотипические признаки — наличие F-пилей (половых ресничек) и чувствительность к L-фагам. С помощью F-ресничек устанавливается контакт между донорскими и реципиентными клетками. Через их канал и передается донорская ДНК при рекомбинации. На половых ресничках расположены рецепторы для мужских λ -фагов. F-клетки не имеют таких рецепторов и нечувствительны к таким фагам.

Геном прокариот передается от материнской клетки к дочерним без изменений, в результате репликации и деления. Такой процесс передачи генетической информации от поколения к поколению называют вертикальным переносом. У бактерий так же наблюдается и горизонтальный перенос — обмен генетической информацией между клетками в одном поколении. В основном обмен происходит для внехромосомной ДНК, но в случае мобильных генетических элементов возможен перенос и фрагментов хромосомной ДНК.

Способы обмена генетической информацией у прокариот

Трансформация — это обмен генетической информацией у бактерий путем введения в бактериальную клетку-реципиент готового препарата ДНК (специально приготовленного или непосредственно выделенного из клетки-донора). Чаще всего передача генетической информации происходит при культивировании реципиента на питательной среде, содержащей ДНК донора. Для восприятия донорской ДНК при трансформации клетка-реципиент должна находиться в определенном физиологическом состоянии (компетентности), которое достигается специальными методами обработки бактериальной популяции.

При трансформации передаются единичные (чаще 1) признаки.

Трансформация является самым объективным свидетельством связи ДНК или ее фрагментов с тем или иным фенотипическим признаком, поскольку в реципиентную клетку вводится чистый препарат ДНК.

Трансдукция — обмен генетической информацией у бактерий путем передачи ее от донора к реципиенту с помощью умеренных (трансдуцирующих) бактериофагов.

Трансдуцирующие фаги могут переносить 1 или более генов (признаков).

Трансдукция бывает:

- специфической — переносится всегда один и тот же ген;
- неспецифической — передаются разные гены. Это связано с локализацией трансдуцирующих фагов в геноме донора:

В случае специфической трансдукции они располагаются всегда в одном месте хромосомы;

При неспецифической их локализация непостоянна.

Конъюгация — обмен генетической информацией у бактерий путем передачи ее от донора к реципиенту при их прямом контакте. После образования между донором и реципиентом конъюгационного мостика одна нить ДНК-донора поступает по нему в клетку-реципиент. Чем дольше контакт, тем большая часть донорской ДНК может быть передана реципиенту.

Основываясь на прерывании конъюгации через определенные промежутки времени, можно определить порядок расположения генов на хромосоме бактерий — построить хромосомные карты бактерий (произвести картирование бактерий).

Структура генома эукариот

Как и в случае прокариот в геноме эукариот можно выявить два типа ДНК: хромосомная и внехромосомная.

Хромосомная ДНК образует хромосомы, в отличие от прокариот, у которых хромосома одна, у эукариот хромосом несколько, количество зависит от организма, например, у дрожжевой клетки хромосом 4, у человека — 46.

Хромосомы ограничены двумембранной оболочкой, и связаны с множеством белков. Это необходимо рассмотреть более подробно, что будет сделано ниже.

Внехромосомная ДНК или геномы органелл. В эукариотических клетках присутствуют органеллы, имеющие свою ДНК и аппарат для ее экспрессии, то есть часть информации о белках находится не в ядре, а в ДНК органеллы. К таким органеллам относятся митохондрии и хлоропласты. Их геном близок по структуре к хромосомам бактерий и достаточно сильно отличается от генома ядра эукариот, поэтому структуры генома органелл будут рассматриваться раньше ядерного.

Структура генома хлоропласта

ДНК хлоропластов резко отличается от ДНК ядра. Она представлена циклическими молекулами длиной до 40-60 мкм, имеющими молекулярный вес $0,8 — 1,3 \cdot 10^8$ Да. В одном хлоропласте может быть множество копий ДНК. Так, в индивидуальном хлоропласте кукурузы присутствует 20-40 копий молекул ДНК. Длительность цикла и скорость репликации ядерной и хлоропластной ДНК, как было показано на клетках зеленых водорослей, не совпадают. ДНК хлоропластов не состоит в комплексе с гистонами. Все эти характеристики ДНК хлоропластов близки к характеристикам ДНК прокариотических клеток. Более того, сходство ДНК хлоропластов и бактерий подкрепляется еще и тем, что основные регуляторные последовательности транскрипции (промоторы, терминаторы) у них одинаковы. На ДНК хлоропластов синтезируются все виды РНК (информационная, трансферная, рибосомная). ДНК хлоропластов кодирует рРНК, входящую в состав рибосом этих пластид, которые относятся к прокариотическому 70S типу (содержат 16S и 23S рРНК). Рибосомы хлоропластов чувствительны к антибиотику хлорамфениколу, подавляющему синтез белка у прокариотических клеток.

ДНК хлоропластов не состоит в комплексе с гистонами.

структурах с ограниченной автономией.

Структура генома митохондрий

Согласно эндосимбиотической теории, митохондрии произошли от риккетсиоподобных бактерий. Эти бактерии являются облигатными внутриклеточными паразитами, чей размер генома составляет 1,1 Мб. Было показано сходство организации и первичной структуры ДНК нуклеоида *Rickettsia* и митохондриальной ДНК *Reclinomonas*. Организация митохондриального генома очень вариабельна, структура генома зависит от того, к какой таксономической группе относится исследуемый организм.

Организация митохондриального генома животных

Митохондриальная ДНК животных представляет собой небольшие кольцевые молекулы, распределенные по всему матриксу митохондрий, размер которых колеблется в пределах 15-20 kb. Основную массу митохондриальной ДНК составляют единичные кольцевые молекулы, хотя встречаются димеры, представляющие собой два генома, соединенные в одно кольцо голова к хвосту, а также катенированные молекулы (подобные структуры являются результатом дубликации). Однако в митохондриях нормальных тканей димеры отсутствуют, а катенированные молекулы составляют менее 10%.

Митохондриальный геном животных содержит всего 37 генов: из них 13 кодируют белки, 2 - рибосомальные РНК и 22 - транспортные РНК, необходимые для трансляции белков. Кодировочной может быть либо только одна цепь (вторая не кодирующая), либо обе. Митохондриальные гены располагаются друг за другом без спейсерных участков. Так же они могут перекрываться, подобное явление не было обнаружено ни в ядерных, ни в бактериальных геномах, однако, встречается в некоторых вирусных ДНК.

В митохондриальном геноме обнаружены некодирующие области, размер которых варьирует от 100 bp до 2 kb. У некоторых животных показано наличие только одной некодирующей области (например, аллигатор, в которой располагаются элементы, отвечающие за регуляцию транскрипции и

репликации, тогда как у других - две некодирующие области, различающиеся по размеру (например, гаттерия).

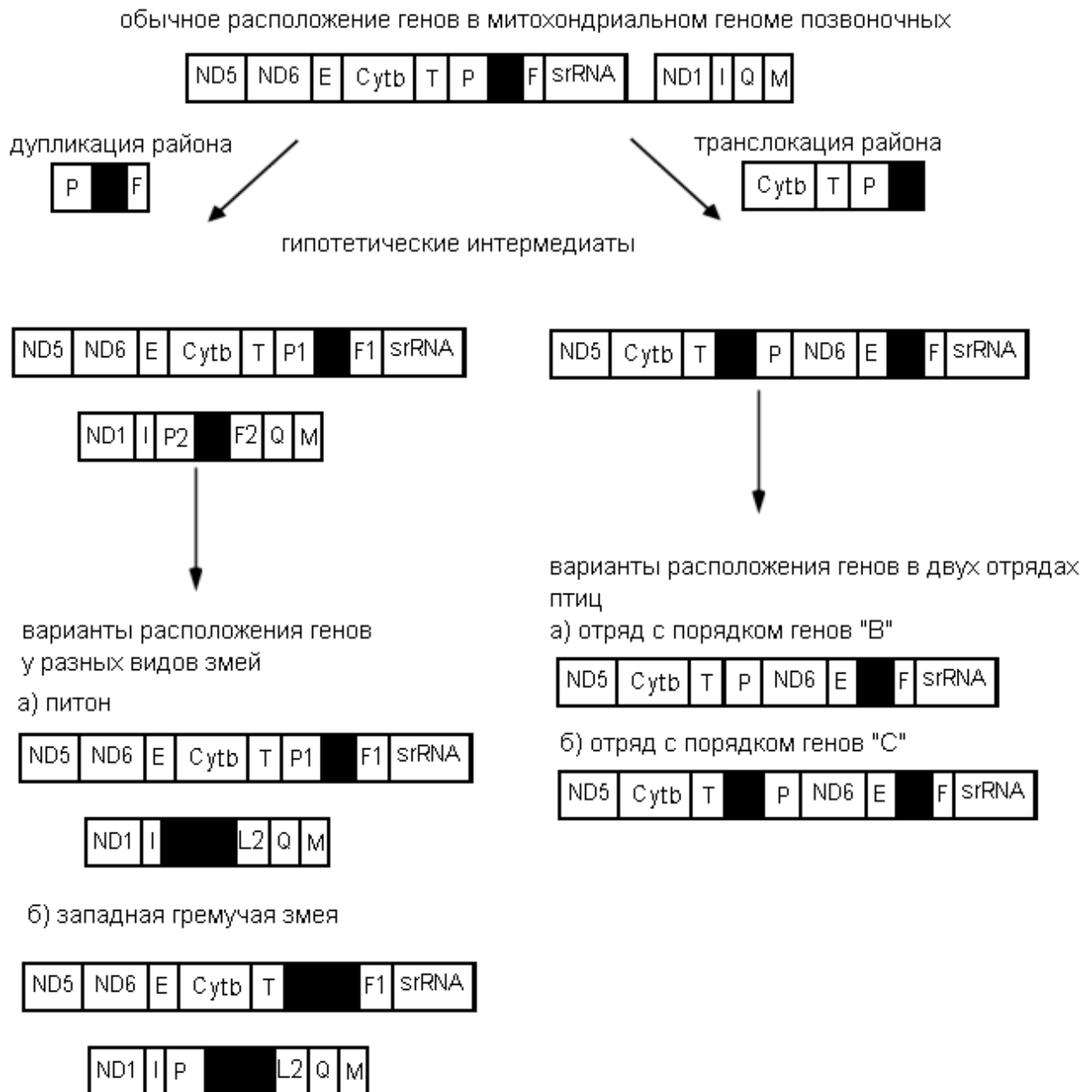


Рисунок 13: Схема перестроек митохондриальной ДНК у различных представителей позвоночных на примере птиц и пресмыкающихся (змей) путем транслокаций, дупликаций и делеций отдельных фрагментов. (Черные прямоугольники обозначают не кодирующие области, не заштрихованные прямоугольники - кодирующие последовательности; буквами обозначены гены тРНК по однобуквенному коду).

Для большинства митохондриальных геномов животных было показано отсутствие интронов. Однако, у актинии *Metridium senile* в гене 5-ой субъединицы NADH-дегидрогеназы был обнаружен интрон первого класса с

необычными свойствами. Этот интрон после правильного сплайсинга представляет собой полиаденилированную бицистронную мРНК, содержащую гены 1-ой и 3-ей субъединиц NADH-дегидрогеназы.

Также открытые рамки считывания в митохондриальном геноме могут быть не полными, в стоп кодоне (TAA) может отсутствовать либо один, либо оба аденозинфосфата.

Набор генов в митохондриальном геноме одинаков практически для всех животных, хотя их расположение может достаточно сильно варьировать за счет перестроек ДНК в ходе эволюции. В основном, эти изменения происходят за счет дупликаций и/или транслокаций некоторых сегментов. При анализе расположения митохондриальных генов у большого числа видов животных было установлено, что в различных группах перегруппировкам подвергались не одни и те же фрагменты митохондриальной ДНК.

Так у рептилий (змеи) изменению подвергся фрагмент, содержащий последовательности генов тРНК ($tRNA^{Pro}$ и $tRNA^{Phe}$) и некодирующий район, тогда как у птиц - участок цитохрома b, двух тРНК ($tRNA^{Pro}$ и $tRNA^{Thr}$) и некодирующей области (см. рисунок 2). По этим изменения можно отслеживать эволюцию организмов.

Таким образом, митохондриальная ДНК животных представляет собой небольшой, компактный геном, достаточно консервативный по своему набору генов и вариабельный по их расположению.

Организация митохондриального генома грибов

Митохондриальная ДНК большинства грибов, также как и у животных, представляет собой замкнутую кольцевую молекулу. Однако ее размер больше и составляет от 16 kb у *Saccharomyces pombe* до 148,5 kb у *Agaricus bitorquis*. По сравнению с животными, у грибов митохондриальный геном более вариабелен по длине, структуре и набору генов, например, описаны «дополнительные гены», кодирующие некоторые рибосомальные белки, мтРНК-полимеразу, а также различное количество и расположение не кодирующих районов.

По структуре митохондриального генома у грибов можно выделить несколько групп: Oomycota (*Phytophthora infestans*), низшие грибы (Chytridiomycota), Ascomycota (дрожжи) и высшие грибы (Basidiomycota).

Наиболее хорошо охарактеризованным представителем Oomycota является *Phytophthora infestans*. Митохондриальный геном *Phytophthora infestans* представляет собой кольцевую ДНК размером около 38kb, которая по своей организации имеет черты сходства с аналогичными молекулами простейших и высших растений. Кодирующими являются обе цепи ДНК, гены могут перекрываться. Также у *Phytophthora infestans* обнаружены гены, которые не были найдены в митохондриальном геноме животных и других грибов, например, 7-я, 9-я, и 11-я субъединицы NADH-дегидрогеназного комплекса, некоторые рибосомальные белки. Однако, с другой стороны, присутствуют черты сходства с мтДНК животных, нехарактерные для других грибов: расположение тРНК по всему геному, отсутствие интронов и очень короткие межгенные промежутки.

Для низших грибов (Chytridiomycota) характерна кольцевая ДНК размером примерно 57 kb. Митохондриальный геном содержит обычный для других групп организмов набор генов, разделенных длинными (до 800 bp) межгенными промежутками. Для Chytridiomycota описаны три особенности, нехарактерные для других грибов: во-первых, наличие редактирования тРНК, во-вторых, мобильные генетические элементы, один из которых кодирует сайт-специфическую эндонуклеазу, в-третьих, очень большое количество интронов первого класса.

Для дрожжей (*Saccharomyces cerevisiae*), относящихся к группе Ascomycota, характерна мтДНК меньшего, чем у других грибов размера, всего около 20 kb. Гены закодированы только на одной цепи ДНК и разделены длинными G/C богатыми спейсерами. В отличие от низших грибов (Chytridiomycota) в митохондриальном геноме дрожжей обнаружены интроны обоих классов. Интроны второго класса также являются мобильными генетическими элементами, которые кодируют обратную транскриптазу и

эндонуклеазу.

Митохондриальная ДНК высших грибов (Basidiomycota) наиболее крупная из всех изученных мтДНК грибов, ее размер составляет 50-70 kb. Геном содержит обычный набор митохондриальных генов, а также 4 дополнительные рамки считывания, одна из которых кодирует мтРНК-полимеразу, для трех других не было показано достоверных гомологий с уже известными последовательностями. Кодирующими являются обе цепи. Однако, основная часть генов располагается на одной, и только три структурных гена (1-я субъединица цитохромоксидазы и 8-я, 9-я субъединицы АТФ-азы) и 11 генов тРНК закодированы по другой цепи.

Таким образом, геномы митохондрий всех грибов обладают следующими особенностями: большее количество некодирующей ДНК, некодирующая ДНК не сконцентрирована в определенном участке кольцевой ДНК, а распределена по всему геному, равномерно по геному распределены промоторы митохондриального генома грибов, кроме того некоторые открытые рамки считывания содержат интроны первого и второго типов.

Организация митохондриального генома простейших

О структуре митохондриального генома простейших в настоящее время известно очень мало, в базах данных присутствует только 23 полных последовательности мтДНК представителей этой группы организмов. Эти данные не позволяют достаточно достоверно обобщать имеющуюся информацию для всех простейших, так как по другим параметрам показана большая гетерогенность внутри этого таксона. Но даже имеющаяся информация указывает на большую вариабельность их митохондриального генома по размеру, структуре и набору генов.

МтДНК простейших чаще всего кольцевая, хотя встречаются и линейные формы (*Paramecium aurelia*, *Tetrahymena pyriformis*), которые распределены в матриксе митохондрии, хотя у некоторых видов (представители отряда Kinetoplastida) они катенированы и образуют единую сеть. Размер этих молекул колеблется от 6 kb (*Plasmodium*) до 76,5 kb (*Malawimonas jakobiformis*).

Мтгеном основной части простейших компактен и размер не кодирующих областей обычно не больше 10%. Часто гены мтгенома простейших перекрываются, например 8-ой и 9-ой субъединиц NADH-дегидрогеназы у *Leishmania tarentolae*.

Для некоторых видов характерно наличие большого количества не кодирующих областей, в которых располагаются повторы, различающиеся по размеру. Простейшие с линейной мтДНК имеют концевые (*Paramecium aurelia*) или теломероподобные повторы различной длины (*Tetrahymena pyriformis*).

Так же, как и в случае большинства других организмов, в мтгеноме простейших закодированы две рРНК, несколько белков дыхательного комплекса и транспортные РНК. Однако, по набору генов, кодирующих белки, Protista более сходны с высшими растениями, чем с животными и грибами. Например, для них показано наличие генов 9-ой субъединицы NADH-дегидрогеназы, 1-ой и 9-ой субъединиц АТР-азы, РНК-полимеразы (*Reclinomonas*), а также большое количество повторов (все эти элементы отсутствуют в мтДНК грибов и животных). Кроме того, в отличие от других групп организмов у простейших обнаружено большое количество неидентифицированных рамок считывания.

У большинства простейших в митохондриальном геноме закодированы все тРНК. Однако, у некоторых представителей этой группы имеется только минимальный набор тРНК, то есть одна тРНК взаимодействует со всеми кодонами, кодирующими данную аминокислоту. У небольшой группы простейших (*Acanthamoeba castellanii*, *Paramecium aurelia*, *Tetrahymena pyriformis*) наблюдается явление импорта некоторых цитоплазматических тРНК в митохондрии. У представителей трипаносоматид и у Apicomplexa гены тРНК отсутствуют в митохондриальном геноме и все тРНК импортируются в митохондрию из цитоплазмы, то есть кодируются в ядре.

Митохондриальные гены могут содержать интроны первого и второго классов, этот факт является еще одним доказательством большого сходства митохондриальных геномов простейших, грибов и высших растений.

Представители отряда Kinetoplastida содержат уникальную по своей

организации митохондриальную ДНК. Их мтгеном образован двумя типами колец (мини и максикольца), различающихся по своей структуре, размеру, функциям, а также представленности. Максикольца являются аналогом мтДНК других организмов. В нем выделяют две области: дивергентную и консервативную. В консервативной области располагаются митохондриальные гены, разделенные небольшими спейсерами, в дивергентной области, по мнению некоторых авторов, располагаются регуляторные элементы. В организации максикольца присутствуют черты сходства с мтДНК как животных, так и грибов. С одной стороны, четкое разделение на кодирующую и регуляторную области указывают на сходство с животными. С другой, большой размер дивергентной области и наличие в ней повторов указывают на сходство с грибами и высшими растениями.

Структура митохондриального генома у представителей отряда Kinetoplastida

Отряд Kinetoplastida относится к жгутиковым простейшим. Одной из таксономических особенностей этой группы является наличие одной гигантской митохондрии, расположенной в основании базального тельца жгутика. К этому отряду относят два семейства: Bodonidae (свободноживущие и паразитические формы) и Trypanosomatidae (паразиты животных и растений). Для представителей данной группы характерна уникальная организация митохондриального генома. Вся митохондриальная ДНК сконцентрирована в одной области митохондрии (кинетопласте) и образует единую сеть из катенированных молекул. Хотя у некоторых представителей семейства Bodonidae, например, *Cryptobia helioides*, молекулы не образуют крупной сети и могут встречаться в виде отдельных релаксированных колец.

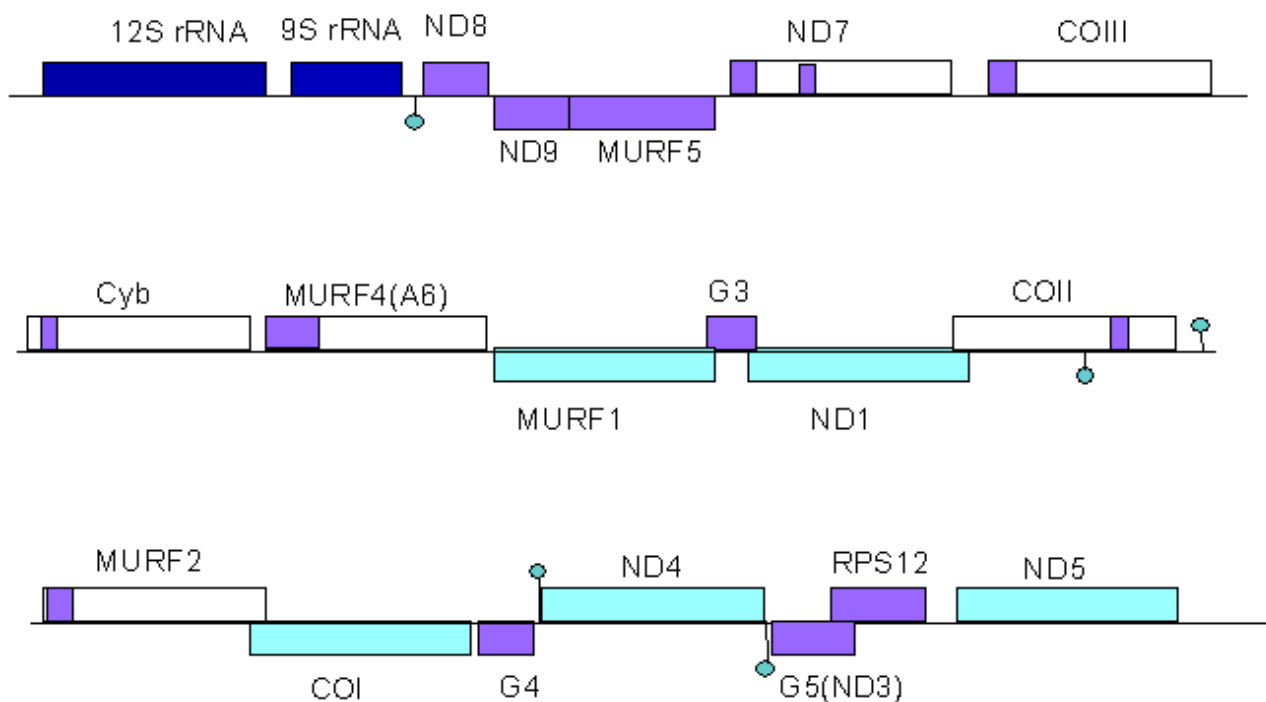


Рисунок 14 : Организация максикольца *Leishmania tarentolae*. Темно синими прямоугольниками обозначены рРНК, голубыми прямоугольниками - нередактируемые гены, синими - полностью и домены частично редактируемых генов, белыми - нередактируемые домены. Кружками обозначены гены гидовых РНК.

Кольцевые ДНК, входящие в митохондриальный геном трипаносоматид, могут быть разделены на два класса, различающиеся по размеру, структуре, представленности в геноме и выполняемой функции. Показано, что все выше указанные признаки могут варьировать у разных видов трипаносоматид. Основную массу кпДНК составляют так называемые миникольца. Это небольшие кольцевые ДНК, размер которых колеблется в пределах от 0,46 до 10 kb. Функции миниколец до сих пор полностью не изучены. Однако, для некоторых видов (*Trypanosoma brucei*, *Leishmania tarentolae*) показано, что в них закодированы гидовые РНК (о функциях гидовых РНК см. ниже). Происхождение этих молекул точно не выяснено. Хотя для представителей бодонид было высказано предположение о том, что аналогичные миниколецам трипаносоматид молекулы - это плазмиды, которые были захвачены митохондрией древней бодониды, которая уже имела собственную ДНК, и сохранились в процессе эволюции как носители генов гидовых РНК .

Второй класс молекул ДНК составляют максикольца. Это крупные молекулы размером 20-38 kb и представленные в количестве всего 25-50 копий на геном. Максикольцевые ДНК представляют собой аналог митохондриального генома других организмов. По структуре максикольцо напоминает мтДНК. В нем можно четко выделить две области: дивергентную и консервативную. Дивергентная область сильно варьирует по размеру и первичной структуре даже у представителей одного рода трипаносоматид. Размер максикольца у разных видов трипаносоматид различается за счет изменений длины дивергентной области. По данным некоторых исследователей, в дивергентной области располагаются элементы, обеспечивающие репликацию и транскрипцию максикольца, а также регуляторные элементы. Тогда как, консервативная область является кодирующей. В максикольце располагаются гены рРНК, белков дыхательной цепи, субъединицы АТФ-азы, а также несколько неидентифицированных рамок считывания (MURF).

Так же, как и в мтгеноме некоторых животных, гены в максикольце закодированы по обеим цепям и разделены небольшими спейсерными участками, в которых могут находиться последовательности гидовых РНК. Однако, в отличие от животных, у трипаносоматид в максикольце не обнаружено ни одной тРНК. Несмотря на достаточно большое сходство структуры консервативной области максикольцевой ДНК у разных трипаносоматид, были обнаружены виды, у которых структура максикольца сильно отличается. Например, *Trypanosoma equiperdum*, у этого вида отсутствуют гены 9S рРНК; 1-ой, 7-ой, 8-ой, 9-ой субъединиц NADH-дегидрогеназы; 2-ой, 3-ей субъединиц цитохромоксидазы; цитохрома b и 6-ой субъединицы АТФ-азы. Тогда как район, включающий гены 3-ей, 4-ой, 5-ой субъединиц NADH-дегидрогеназы и RPS12 (гомологичный одному из рибосомальных белков *E. coli*), наоборот дублирован.

Митохондриальный геном также содержит не кодирующую ДНК, к которой относятся origin репликации, промоторы, терминаторы, интроны и межгенные спейсеры. ДНК митохондрий достаточно компактна, поэтому

интроны и межгенные спейсеры достаточно короткие вплоть до отсутствия, длина зависит от организма.

Origin репликации в митохондриальной ДНК один, структура также зависит от генома.

Структура промоторов очень многообразна. Структура промоторов сильно зависит от таксона организма. Поэтому необходимо рассмотреть структуру промоторов митохондриального генома в зависимости от таксона.

Структура промоторов животных

Как было сказано ранее, структура митохондриального генома животных сходна для представителей этого таксона. Соответственно и инициация транскрипции мт-генома осуществляется по сходной схеме: все элементы, отвечающие за этот процесс, располагаются в контрольной области (D-петле). Различия наблюдаются только в структуре промоторных элементов.

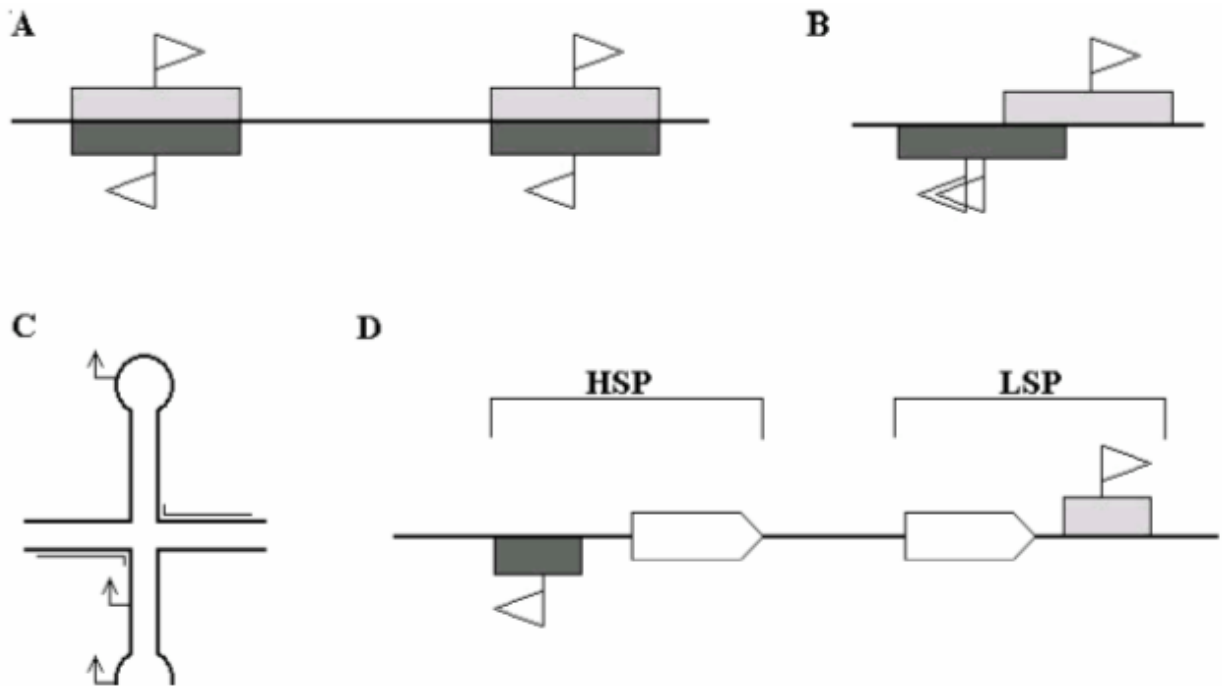
Амфибии

Инициация митохондриальной транскрипции амфибий была достаточно подробно изучена на примере *Xenopus laevis*.

У *Xenopus laevis* в контрольной области локализованы два промотора, отстоящие друг от друга на расстоянии около 65 нуклеотидов. Каждый из них состоит из двух, расположенных на разных цепях ДНК, перекрывающихся восьминуклеотидных элементов (5'-ACRTTATA-3'), которые обеспечивают инициацию транскрипции, каждый на одной из цепей.

Птицы

У птиц структура промоторов изучена на примере кур. У этого вида контрольная область содержит один промотор, локализованный в контрольной области и обеспечивающий считывание информации с обеих цепей ДНК (см. рис. 4 В). Для легкой цепи ДНК существует один сайт инициации, а для тяжелой - два, находящиеся на расстоянии двух нуклеотидов друг от друга. Как



Р

исунок 15: Структура митохондриальных промоторных элементов позвоночных животных. Прямоугольниками различных оттенков серого показаны участки, взаимодействующие с РНК-полимеразой. Флажками - участки инициации транскрипции. А - схема D-петли амфибий (*Xenopus laevis*); В - схема D-петли птиц; С - схема крестообразной структуры образуемой промоторными элементами птиц; стрелками отмечены сайты инициации транскрипции, тонкими линиями - положения консервативного 8-нуклеотидного элемента промотора. D - схема организации D-петли позвоночных, белыми пятиугольниками отмечены участки ДНК, взаимодействующие с фактором транскрипции mtTF.

и у *Xenopus*, митохондриальный промотор кур состоит из двух восьминуклеотидных элементов, 5'-ACRTTATA-3', которые находятся на разных цепях ДНК, однако, в данном случае, они не перекрываются. Перед каждым восьминуклеотидным элементом находится АТ богатый участок, без которого промотор не функционирует. Таким образом, по первичной структуре митохондриальный промотор курицы представляет собой два инвертированных относительно друг друга 16-18 нуклеотидных повтора. Эта последовательность

способна образовать крестообразную структуру, причем сайты инициации для обеих цепей окажутся в петлях (см. рис. 4 С).

Для других птиц, структура митохондриального генома и механизмы его реализации изучены только у уток. В их митохондриальной ДНК содержится тот же набор генов, что и у кур, и они расположены в той же последовательности. Контрольная область содержит ориджин репликации и промотор для инициации транскрипции на обеих цепях ДНК с той же консенсусной последовательностью, что и у кур.

Млекопитающие

Митохондриальный геном млекопитающих обладает практически одинаковой структурой: в области D-петли располагаются два промотора: HSP - для тяжелой цепи и LSP - для легкой. Промоторы расположены очень близко друг к другу, но не перекрываются, и инициация транскрипции с них осуществляется независимо. В митохондриальном промоторе человека выделяют два функционально важных элемента (см. рис. 4.D). Первый прилегает к сайту инициации и является участком связывания РНК полимеразы. Второй элемент расположен за 12 - 40 нуклеотидов перед стартовой точкой транскрипции и представляет собой ряд повторов. Он определяет специфичность инициации. С ним взаимодействует транскрипционный фактор h-mtTF - 29 kDa белок, содержащий два HMG домена (high mobility group, ДНК-связывающий домен, имеющий высокое сродство к АТ-богатым областям), соединенных короткой линкерной последовательностью и С-концевой хвост.

Расположенные upstream элементы HSP и LSP очень схожи по первичной структуре. Они представляют собой ряд повторов и отстоят друг от друга на расстоянии 80 нуклеотидов. За счет работы этих элементов, каждый митохондриальный промотор может вызывать инициацию транскрипции в двух направлениях (т.е. с любой цепи ДНК), однако преобладает синтез одного транскрипта (эффективность инициации на разных цепях различается приблизительно в 50 раз). Стартовые точки транскрипции в обоих направлениях для каждого промотора соответствуют одной и той же паре

нуклеотидов.

В регуляции транскрипции митохондриальных генов, помимо РНК-полимеразы и h-mtTF, принимает участие LON протеаза. Она сайт-специфически связывается с одноцепочечными участками ДНК промоторной области. Кроме того, известно, что LON протеаза взаимодействует с транскрипционными факторами, вызывая их деградацию. Таким образом, этот фермент осуществляет репрессию транскрипции и репликации за счет протеолиза белковых факторов, которые взаимодействуют с ДНК вблизи участка D-петли, с которым связана LON.

За пределами D-петли, со стороны HSP, находятся еще два участка, с которыми может взаимодействовать h-mtTF. Функциональное значение этих областей неизвестно. Однако, их положение на ДНК относительно HSP аналогично положению промоторных upstream элементов относительно LSP. В связи с этим, существует гипотеза о том, что два промотора в митохондриальном геноме человека образовались за счет дубликации одного промотора, причем этот промотор обеспечивал транскрипцию обеих цепей ДНК. Сходная организация промоторных элементов характерна для других видов позвоночных. Однако, при изучении контрольной области мтДНК мыши было обнаружено, что транскрипция тяжелой цепи осуществляется с двух независимых промоторных элементов.

Беспозвоночные

Структура митохондриального генома беспозвоночных животных сходна с таковой у позвоночных. Для представителей насекомых рода *Drosophila* было показано наличие A/T богатого района, который сильно варьирует по длине и нуклеотидной последовательности у различных видов этого рода. В этом районе были идентифицированы ориджин репликации и сайты инициации транскрипции для обеих цепей. Также для этих насекомых было показано наличие участков ДНК, с которыми связываются митохондриальные факторы транскрипции - это последовательность ТТАТС/Г несколько раз повторяющаяся в A/T богатой области.

Для других беспозвоночных информации значительно меньше. Только для *Artemia franciscana* (представитель отряда ракообразных) для тяжелой цепи мтДНК были определены два сайта инициации транскрипции, содержащие последовательность ТССА(Г)СА(Н)(Н)ТААА. Однако данные о структуре промотора для легкой цепи отсутствуют.

Структура промоторов грибов

Наиболее изучена транскрипция митохондриального генома Ascomycota (*Saccharomyces cerevisiae* и *Neurospora crassa*).

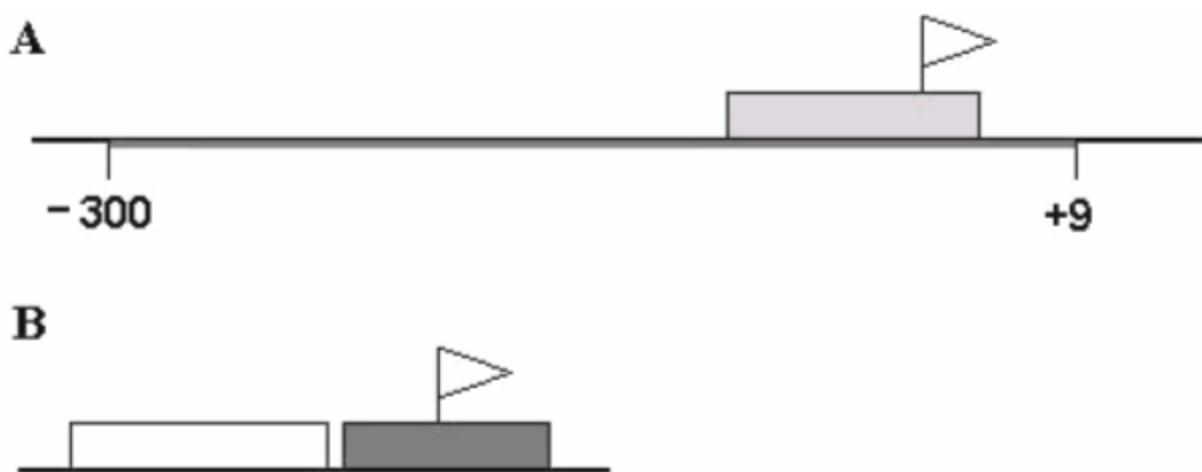


Рисунок 16: Схемы организации митохондриальных промоторов грибов. Флажками обозначены участки инициации транскрипции, серыми прямоугольниками - консервативные последовательности с которыми взаимодействует РНК-полимераза. Белым - А/Т богатый участок. А - промоторный элемент дрожжей. В - промотор *Neurospora crassa*.

Относительно этого процесса у других грибов известно мало, но имеющиеся данные сходны с полученными для аскомицетов. В отличие от животных, в мтДНК грибов обнаружены несколько точек инициации транскрипции, которых у *Saccharomyces cerevisiae* тринадцать. Большая часть этих промоторов не связана с сайтами инициации репликации, однако, в участках ori-репликации обнаружены консенсусные последовательности, которые узнаются митохондриальной РНК-полимеразой. Это позволяет предположить, что в транскрипции митохондриальных генов и в образовании РНК-затравок для репликации участвует один и тот же фермент. Промоторный участок дрожжевой мтДНК представляет собой 9-нуклеотидный консервативный

элемент: (-7)TATAAGTAA(+2), который узнается митохондриальной РНК-полимеразой. Этот элемент абсолютно необходим для инициации транскрипции. В некоторых промоторах он может быть частично изменен, что оказывает влияние на их силу. Кроме того, на эффективность инициации транскрипции могут оказывать влияние последовательности, находящиеся в районе от -300 п.н. до консервативного элемента и первые девять, следующих за ним. Так, например, замечено, что наличие Т(тимина) в участке, прилегающем к промотору с 3' стороны, ослабляет его работу. Неактивные промотор-подобные последовательности содержат в этом участке целые кластеры Т (тиминов), а самые сильные промоторы вообще не имеют Т (тиминов) в данной области. Кластеры других нуклеотидов такого влияния не оказывают.

В митохондриальном геноме *Neurospora crassa* инициация транскрипции осуществляется с помощью нескольких промоторов следующего строения: АТ-богатый участок, длиной порядка 15-27 нуклеотидов, и следующий за ним одинадцатинуклеотидный консервативный элемент TTAG(A/T)RR(G/T)(G/C)N(A/T). Начало транскрипта приходится на 5-6 положение консервативного элемента (см. рис. 5 В). В отсутствие АТ-богатого участка промотор неактивен.

Структура промоторов растений Водоросли

По структуре митохондриального генома красная водоросль *Chondrus crispus* ближе к животным, чем к высшим растениям. Ее митохондриальная ДНК небольшого размера (26 kb), и практически не содержит некодирующих областей. Так же как и в мтДНК животных регуляторная область, обеспечивающая инициацию транскрипции обеих пре-РНК, локализована в самом крупном некодирующем участке ДНК (282 п.н.) и представляет собой четыре шестидесятинуклеотидных повтора, причем 1-й, 2-й и 4-й направлены в одну сторону, а 3-й - в противоположную. Последовательность нуклеотидов регуляторной области теоретически способна образовать крестообразную структуру, как и регуляторная область мтДНК кур.

Для обеих пре-РНК детектировано по несколько сайтов инициации транскрипции, причем всем им предшествует консенсус АСГТТТ(Т/А)У. В регуляторной области есть три участка с палиндромом АААСГТТТ, который, видимо, участвует в инициации транскрипции на обеих цепях ДНК, то есть промоторы у *C. crispus* двунаправлены. Такой палиндром не обнаруживается ни в одном из прочих межгенных участков, что согласуется с отсутствием в них сайтов инициации транскрипции.

Pylaiella littoralis

Митохондриальный геном бурой водоросли *Pylaiella littoralis* транскрибируется с помощью двух РНК-полимераз. Первая - многосубъединичная полимеразы бактериального типа, закодированная в ядре. Вторая - полимеразы фагового типа, гомологичная митохондриальным РНК-полимеразам других организмов. Интересной особенностью этого фермента является то, что его ген находится в митохондриальном геноме. Таким образом, *Pylaiella littoralis* является на данный момент единственным из исследованных организмов, использующая для транскрипции митохондриальных генов две РНК-полимеразы, причем одна из них закодирована в геноме митохондрии.

Высшие растения

Механизм инициации транскрипции у высших растений напоминает таковой у грибов. В мтДНК располагается множество промоторов, в некоторых случаях несколько на один ген (например, ген *cox2* кукурузы). Во всех митохондриальных промоторах растений присутствует core последовательность: СRТА. Однако, по организации промоторных элементов растения можно разделить на две группы, соответствующие выделенным ранее порядкам: однодольные и двудольные.

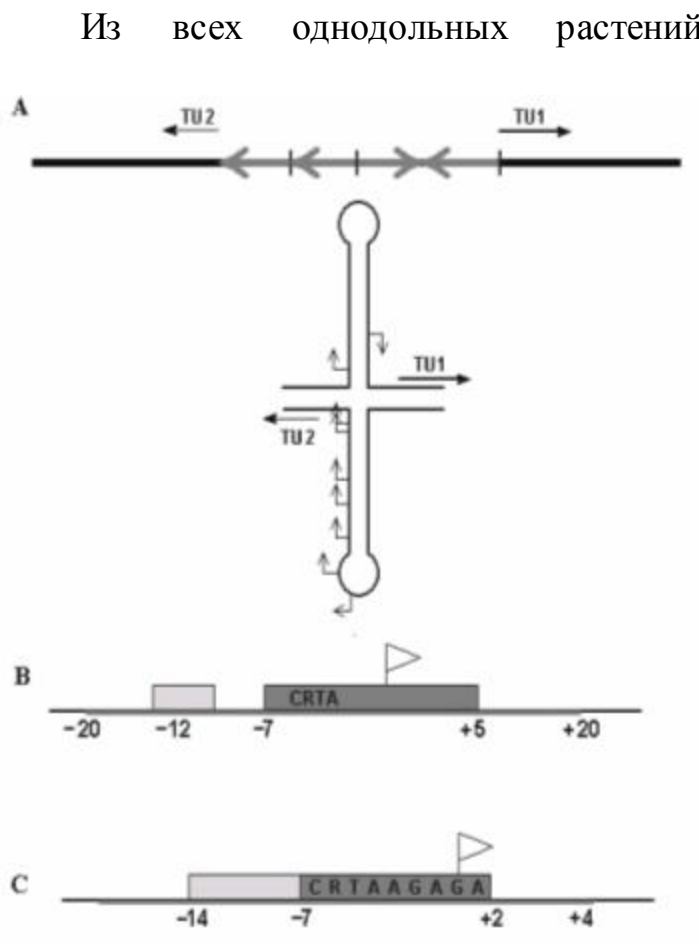


Рисунок 17: Схемы организации митохондриальных промоторов растений. А - схема организации промотора водоросли *Chondrus crispus*; серыми стрелками обозначены повторы, тонкими стрелками сайты инициации и направление транскрипции. В - схема промотора однодольных растений; С - схема промотора двудольных растений. Флажками обозначены участки инициации транскрипции, темно-серыми прямоугольниками - консервативные последовательности, с которыми взаимодействует РНК-полимераза; светло-серыми прямоугольниками - А/Т богатый участок.

Из всех однодольных растений наиболее детально изучены митохондриальные промоторы кукурузы. Они состоят из двух доменов: центральный домен от -7 до +5 п.н., в состав которого входит консервативная для большинства промоторов последовательность CRTA, и пурин-богатый домен, локализованный возле -12 п.н. (см. рис. 6 В). В редких случаях консервативный участок промотора может быть изменен. Встречаются

промоторы, содержащие в этой области TRTA, YYTA или RRTA последовательности. Расстояние от консервативного элемента до сайта инициации транскрипции может варьировать от 0 до 7 п.н., чаще всего составляет 2-5 п.н.. Расстояние между доменами промотора так же может быть

разным. Показано, что оно влияет на силу промотора, однако, регуляторные области некоторых генов вовсе лишены домена -10 - -12, например промотор гена *cox2* (II субъединица цитохром c оксидазы) кукурузы. На силу промотора влияет его окружение. Важна область от -20 до +20 п.н. С этой областью взаимодействуют белки, необходимые для инициации транскрипции.

Промоторы двудольных проявляют большую консервативность первичной структуры, чем промоторы однодольных. Они состоят из центрального домена, в состав которого входит последовательность (-7)CRTAAGAGA(+2), и предшествующего ему пятинуклеотидного АТ-богатого участка (см. рис. 6 С). Следует отметить, что, в отличие от однодольных, положение консервативного ядра промотора двудольных строго фиксировано относительно положения сайта инициации транскрипции. Известно, что на эффективность инициации оказывает влияние последовательность нуклеотидов от -14 до +4. АТ-богатый участок, скорее всего, необходим для плавления ДНК в ходе инициации транскрипции. Показано, что в этой области важно не количество нуклеотидов каждого типа, а их расположение, а также расстояние между АТ-богатым участком и центральным доменом.

АТ-богатый участок присутствует не во всех промоторах. Для митохондрий картофеля известно несколько сайтов инициации транскрипции, окружение которых практически не проявляет сходства со стандартной структурой промоторов двудольных. Существует гипотеза, что для взаимодействия мтРНК-полимеразы с такими промоторами существуют специальные транскрипционные факторы, однако, на данный момент, они неизвестны.

Транскрипция митохондриальной ДНК у различных групп организмов

Из всего выше перечисленного можно заключить, что существуют два принципиально различных механизма митохондриальной транскрипции. Эти механизмы подразумевают различия не только в инициации, но и по форме пре-мРНК (поли или моноцистронные), а также по терминации транскрипции.

Инициация осуществляется в одном участке мтДНК (контрольной области), в котором располагаются промоторы и элементы, отвечающие за регуляцию транскрипции. Количество промоторов ограничено 1-2 для каждой цепи. В результате транскрибируется крупная полицистронная пре-мРНК, в некоторых случаях (например, *Plasmodium falciparum*) совпадающая по размеру

с длиной мтДНК. Затем эта пре-мРНК подвергается нарезанию и дальнейшему процессингу с образованием зрелых моноцистронных мРНК. Элементов, отвечающих за терминацию транскрипции, в мтДНК два. Располагаются они следующим образом: один за генами рРНК, второй в контрольном регионе.

Сайтов инициации несколько (около 10) и распределены они по всему митохондриальному геному. В результате считывается несколько моноцистронных или небольших полицистронных пре-мРНК, содержащих две или три открытых рамки считывания. Элементов, отвечающих за терминацию транскрипции, в мтДНК несколько и располагаются они в 3' UTR. Однако, для растений было показано, что эти элементы чаще являются сигналами процессинга, а не терминации.

Для этого случая завершения транскрипции необходимо и присутствие тридекамерной последовательности в tRNA LeuUUR ген и определенный белковый фактор, mTERF, является необходимым и достаточным. Продвигающая завершение деятельность mTERF двунаправлена и, так как никакие дополнительные гены не кодируются в этой цепи upstream и downstream. mTERF обязательный участок {сайт}, это могло также действовать в конце транскрипции L-цепи.

Структура ядерного генома эукариот

Основной частью генетической информации у эукариот является именно ядерная ДНК, в отличие от прокариот она организована в несколько хромосом, в связи с этим существует специальная органелла — ядро. Хромосомы ограничены двумембранной оболочкой и за счет ядерного матрикса осуществляется организация генетического материала в ядре. Хромосомная ДНК связана с большим количеством белков и этот комплекс имеет два типа организации: собственно хромосома (в ходе деления клетки) и хроматин (в интерфазном ядре). Хроматин также бывает двух видов эухроматин и гетерохроматин. Низший уровень организации хромосом и хроматина — нуклеосомный. Нуклеосома состоит из октамера, образованного четырьмя

парами гистоновых белков (H2A, H2B, H3, H4) и фрагмента ДНК длиной 146 п.н., накрученного на этот октамер. Дополнительно в организации нуклеосом участвует гистоновый белок H1. На более высоких уровнях организации хроматина и хромосом участвуют негистоновые белки.

Главной особенностью генетического материала эукариот в сравнении с прокариотами является наличие избыточной ДНК. Если средний размер гена бактерий - 1500 п.н., а длина кольцевой молекулы ДНК хромосомы *E. coli* и *Bacillus subtilis* составляет около 1,1 мкм, то в такой хромосоме могут разместиться около 3000 генов. Примерно такое число генов было экспериментально определено у бактерий по числу типов мРНК. Если это число умножить на средний размер гена, то получится, что около 95% генома бактерий состоит из кодирующих (генных) последовательностей.

Иная картина наблюдается у эукариотических организмов. Например, у человека насчитывают приблизительно $5 \cdot 10^4$ генов. В то же время размер генома человека - $3 \cdot 10^9$ п.н. Это означает, что кодирующая часть его генома составляет всего 15-20% от всей ДНК. Существуют виды, геном которых в десятки раз больше генома человека, например некоторые рыбы, хвостатые амфибии, лилейные растения. Избыточная ДНК характерна для всех эукариот. В связи с этим необходимо разграничить понятия геном и генотип.

Генотип - это совокупность генов, имеющих фенотипическое проявление.

Геном — это количество ДНК в гаплоидном наборе хромосом данного вида.

Также как и у прокариот хромосомная ДНК эукариот подразделяется на кодирующую и не кодирующую. Кодирующая включает в себя последовательности рРНК, тРНК и открытые рамки считывания. Транспортные и рибосомные РНК сходны по строению с аналогичными у прокариот, а вот строение ORF, кодирующих белки несколько отличается.

При изучении первичной структуры, т.е. последовательности нуклеотидов

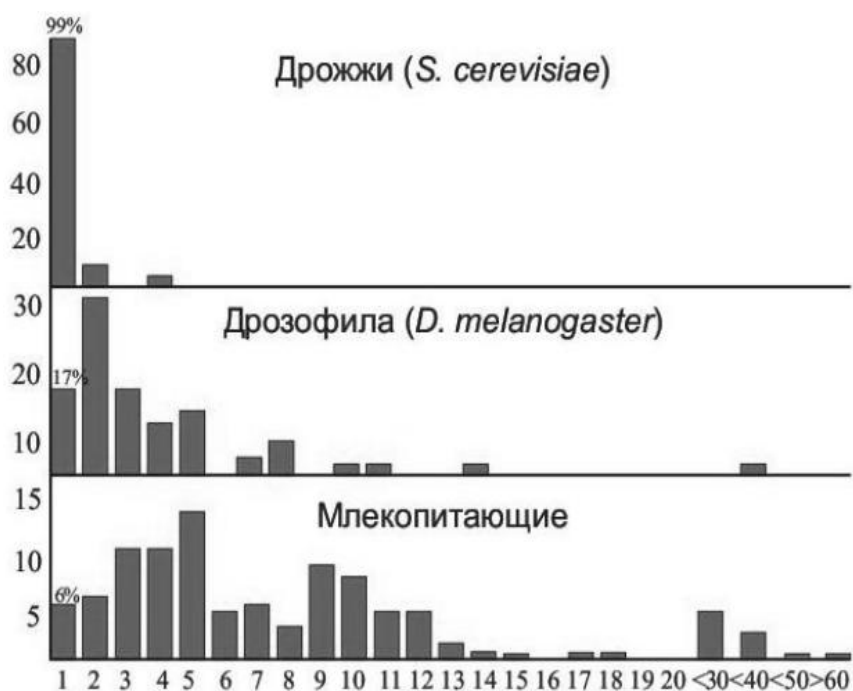


Рисунок 18: Распределение частот генов, имеющих различное количество интронов. Ось ординат - процент генов, ось абсцисс - количество интронов.

ряда генов выяснилось, что в них, наряду с участками, кодирующими специфичный для этого гена продукт (полипептид, рРНК, тРНК и т.д.) имеются участки, которые ничего не кодируют, т.е. они подобно межгенным спейсерам (участкам

между генами) не содержат генетической информации. Группы учёных, возглавляемых Р. Робертсом и П.А. Шарпом обнаружили такие расщеплённые гены у аденовируса 2 в 1977 году. Некодирующие участки получили название интронов, кодирующие - экзонов. Такой тип структурной организации обнаружен для множества генов, локализованных в хромосомах, некоторых генов внутриклеточных органелл эукариот - пластид и митохондрий, а также для генов нескольких РНК-содержащих и ДНК-содержащих вирусов, поражающих эукариот. У бактерий интронов в генах нет. Нет интронов и в генах вирусов, поражающих бактерии. Число и внутригенная локализация интронов характерны для каждого гена, что становится очевидным в результате сравнения организации гомологичных генов у разных видов. Некоторые гены содержат только один-два интрона, но часто их значительно больше. Так, например в гене овальбумина курицы 7 интронов, в гене сывороточного альбумина крысы их 13, а один из генов коллагена курицы имеет даже 51 интрон. Видно, что у низших эукариот, таких как дрожжи, 95% генов содержат

только один экзон, значит такие гены не прерываются нитронами. У дрозофилы таких генов только 17%, а у млекопитающих - только 6%. Экзоны имеют, как правило, небольшую длину. Длина интрона может быть очень различной - от нескольких десятков пар нуклеотидов до многих тысяч. Общая длина всех интронов часто значительно превышает суммарную длину экзонов. К примеру, из приблизительно 7000 пар нуклеотидов, образующих ген овальбумина, на долю экзонов приходится всего 1872 п.н., т.е. почти 3/4 длины составляют интроны. Интроны транскрибируются наравне с экзонами, так что про-мРНК содержит участки, транскрибированные как с экзонов, так и с интронов. В дальнейшем в ходе процессинга, происходящего в ядре, участки про-мРНК, транскрибированные с интронов, вырезаются, а бывшие разобщенными участки, считанные с экзонов, "сращиваются", и зрелая мРНК содержит только транскрипты экзонов. Эти прежде разобщенные участки соединяются в нужном порядке. Воссоединение участков, транскрибированных с экзонов при образовании зрелой мРНК, называют сплайсингом (splicing - сращивание морских канатов - англ.). Длина гена существенно больше длины мРНК. Интроны всегда (установлено для генов, кодирующих белки) имеют на 5'-конце двойку последовательностей гуанин-тимин, а на 3'конце — аденин-гуанин. Последовательности нуклеотидов в экзонах консервативны, а в интронах сильно варьируют. Иногда экзон одного гена может быть гомологичным экзону даже другого гена. Например, два β -глобиновых гена мыши имеют по три гомологичных экзона в каждом гене. Между нитронами этих генов гомология не найдена. Связано это с тем, что интроны эволюционируют значительно быстрее, чем экзоны. При сравнениях последовательностей нуклеотидов в одних и тех же генах у разных видов, находят большую гомологию в интронах. В интронах некоторых генов располагаются другие гены. Удивительным по сложности организации и величине интронов является ген *dnc* у *D.melanogaster*. Ген занимает минимум 130 т.п.н. (от -90 до +40 т.п.н.) и содержит 13 экзонов. Между первым экзонам и экзонам, расположенным еще более дистально (около 40 т.п.н.), расположено несколько генов из семейства. Между вторым и третьим

экзонами (около 70 т.п.н.) располагается еще 4 гена.

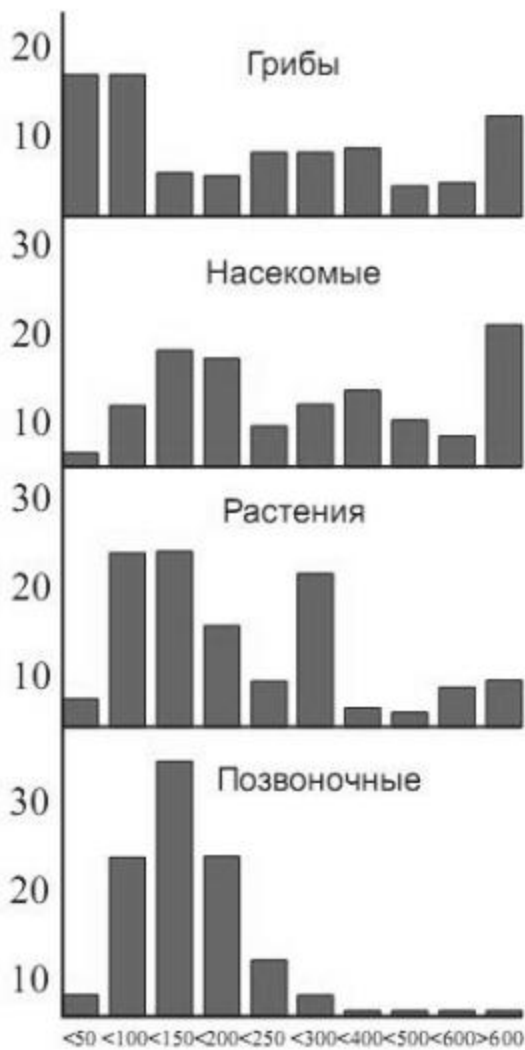


Рисунок 19: Длины интронов у различных организмов. Ось абсцисс — длина интрона, ось ординат — процент встречаемости.

Так же отличается и расположение генов в геноме. В отличие от прокариот у эукариот не распространен оперонный тип расположения генов, т.е. объединение в блоки генов, находящихся под общим контролем. Гены, контролирующие даже последовательные биохимические реакции, расположены в различных районах хромосомы и даже в различных хромосомах. Например, у дрозофилы гены, кодирующие ферменты, под контролем которых происходит превращение триптофана в бурый глазной пигмент, разбросаны во множестве участков генома. Вместе с тем, известны некоторые примеры кластерной организации генов. У человека известно несколько типов гемоглобинов. Каждый из них синтезируется в определенном органе и на определенной стадии развития. Например, гемоглобины ζ и ϵ синтезируются в клетках эмбрионального желточного мешка на ранних этапах развития.

В это время молекула белка состоит из двух цепей ζ -подобный α -гемоглобин и двух цепей ϵ (β -подобный). К концу третьего месяца развития синтез гемоглобина этого типа заканчивается и начинается синтез фетального (утробного) гемоглобина (Hb-F) в клетках печени и селезенки. Фетальный гемоглобин состоит из двух α -полипептидов и двух γ (β -подобных): двух γ A или двух γ G. Во время постэмбрионального развития гемоглобин синтезируется в клетках костного мозга, и состоит из α , β полипептидов и некоторого количества β -подобного δ полипептидов. Большая часть

гемоглобина в крови взрослого человека состоит из тетрамера $\alpha 2\beta 2$ (Hb-A). В геноме человека гены гемоглобина расположены двумя кластерами: в хромосоме 16 расположены все α -подобные гены, в то время как все β -подобные гены расположены в хромосоме 11. В каждом кластере есть псевдогены. Интересно, что гены расположены в том порядке, вдоль по хромосоме, в какой очередности они включаются в работу в ходе онтогенеза. Однако, каких-либо данных, свидетельствующих об их функциональной сцепленности или общем контроле по принципу оперонной организации, не получено. Тот факт, что эти гены функционируют в разных тканях и на разных этапах онтогенеза, скорее свидетельствует о независимом контроле экспрессии этих генов. Гены транспортной РНК у дрозофилы расположены во многих районах по 1-2 гена в каждом. Однако, в одном из районов 16 генов занимают небольшой участок длиной 50 т.п.н., хотя их транскрипция не контролируется одним промотором. Организация генов 18S и 28S рРНК у всех эукариот в общих чертах одинаковая. Последовательности генов 18S и 28S рРНК, лидерной последовательности, а также транскрибируемого и нетранскрибируемого спейсера составляют единицу длиной около 11 т.п.н., которая повторена несколько сот раз. Как правило число копий варьирует в пределах от 100 до 1000: у дрожжей - 140 единиц повторенности, у дрозофилы - 200-250, у человека - 1250. Лидерная последовательность расположена перед геном 18S, и с неё начинается транскрипция. Каждая единица повтора генов рРНК считывается РНК-полимеразой I от лидерной последовательности до конца гена 28S рРНК в виде одного цистрона. Участок ДНК между генами 18S и 28S транскрибируется вместе с этими генами и называется транскрибируемым спейсером. В нём расположены короткие последовательности, которые выделяются из общего транскрипта в ходе процессинга РНК. У млекопитающих и амфибий в коротком спейсере формируется 5.8S РНК - небольшая молекула, образующая водородную связь с 28S рРНК в рибосоме. Транскрипционная единица рРНК самая короткая у бактерий, у которых при общей длине транскрипта 6 т.п.н. 80% приходится на

кодирующую часть. У эукариот длина транскрипта варьирует от 7 до 8 т.п.н. и кодирующая часть оставляет 70-80%. В ходе созревания рРНК лидирующая последовательность и часть транскрибируемого спейсера, не кодирующая 5.8S РНК, деградируют до нуклеотидов. Транскрибируемые единицы разделяются участком ДНК, называемым нетранскрибируемым спейсером (НТС). Его длина варьирует в широких пределах: от 1750 п.н. у дрожжей до 30000 п.н. у мыши, у дрозофилы он имеет длину 3750-6450 п.н. В самом НТС обнаружили внутренние повторы, которые в частности у дрозофилы, играют существенную роль в мейотическом спаривании X и Y хромосом. Существенная часть генов 28S рРНК содержит инсерции. Каждая из единиц повтора может функционировать независимо от других. Иногда в кластеры объединяются и неповторенные гены. Семь генов, кодирующих белки теплового шока, располагаются во фрагменте ДНК длиной 12 т.п.н., картируемом с помощью гибридизации *in situ* в пуфе теплового шока 67В. Для генов этого района также характерна разная направленность транскрипции.

Структура некодирующей ДНК у эукариот

Некодирующая ДНК эукариот составляет основную массу генома, это составляет 80 — 85% всего генома. Эта ДНК сложно организована, в ней можно выделить следующие элементы: участки инициации репликации, элементы инициации, терминации и регуляции, транскрипции, мобильные генетические элементы, а также собственно некодирующая ДНК.

Структура *origin* репликации

В отличие от прокариот каждая эукариотическая хромосома содержит несколько *origin* репликации, это обеспечивает высокую скорость репликации крупного генома.

Инициация репликации у эукариот происходит на специфических множественных последовательностях нуклеотидов - репликаторах. Наиболее изученными являются репликаторы дрожжей *S. cerevisiae*, впервые идентифицированные как автономно реплицирующиеся последовательности

(ARS - autonomously replicating sequence), способные поддерживать внехромосомную репликацию плазмид в дрожжевых клетках.

У млекопитающих области начала репликации располагаются на расстоянии около 100 т.п.н. друг от друга; часть их удалось клонировать и изучить на молекулярном уровне. Установлено, что синтез ДНК в отдельных репликациях происходит по двум направлениям, причем перемещение репликативной вилки осуществляется предпочтительно в одном направлении, которое может изменяться в зависимости от стадии развития организма и уровня экспрессии генов, содержащих репликаторы. Частота использования отдельных репликаторов изменяется в онтогенезе, уменьшаясь в клетках взрослого организма.

Сравнение первичных структур шести отдельных репликаторов эукариот показало, что все они содержат DUE-элементы, участки прикрепления к ядерному матриксу (SAR/MAR), канонические ARS- последовательности дрожжей, пиримидиновые тракты, а также каноническую последовательность

WAWTTDDWWWDHWGWHMAWTT,

где $W = A/T$, $D = A/C/T$, $H = A/C/T$, а $M = A/C$.

Имеются сообщения о том, что в репликаторах животных присутствуют пуриновые тракты, канонические последовательности, взаимодействующие с факторами транскрипции и белками репликативного комплекса, энхансерный октамерный мотив, сайты связывания продуктов онкогенов, АТ-богатые последовательности и участки перегибов (bent) ДНК. Предполагается, что многие из этих последовательностей участвуют в регуляции транскрипции (а следовательно, и регуляции экспрессии генов).

Ориджины репликации могут располагаться внутри кодирующей части генов, например в гене (beta-глобине цыпленка находится по крайней мере четыре ориджина репликации, причем было показано что активность генов не связана со степенью репликации. На выбор ориджинов репликации влияет ацетилирование гистонов. Например, репликация в бетта-глобиновом локусе цыпленка не чувствительна ни к ацетилированию ни к деацетилированию

гистонов. В мышинном *NoxV* локусе, ацетилирование гистонов имеет зависимость с молчанием ориджинов скорее чем их активация.

Структура элементов транскрипции

К элементам обеспечивающим транскрипцию относятся: промоторы, терминаторы и цис-элементы регуляции транскрипции.

Как уже было сказано выше у эукариот РНК полимераз три. РНК-полимераза I (или А по другой, номенклатуре) осуществляет в ядрышке синтез 18 S и 23 S рРНК- РНК-полимераза II (или В) — синтез информационных РНК- РНК-полимераза III (или С) — синтез тРНК и некоторых других низкомолекулярных РНК. Рассмотрим транскрипцию этих групп генов по отдельности.

Транскрипция рибосомных генов с помощью РНК-полимеразы I

Более половины валового синтеза РНК эукариотической клетки приходится на образование 18S, 28S и 5,8S РНК рибосом. Рибосомные гены представлены сотнями повторяющихся копий, сгруппированных в одном или



Рисунок 20: Типичный межгенный спейсер рДНК *X. laevis* (IGS область, расположенная между 3-концом последовательности, кодирующей 28S-рРНК, и сайтом, в котором начинается транскрипция про-рРНК в начале 5-ETS) Овалами со стрелками обозначены спейсерные промоторы; участки, выделенные черным цветом, гомологичные повторы из 60 81 п.н. Символами T1 и T2 отмечены неэффективный и эффективный терминаторы транскрипции соответственно. Функции участков, обозначенных цифрами 0 и 1, неизвестны.

нескольких участках генома. Многократная повторяемость генов рРНК, а также их амплификация на определенных стадиях развития отражают потребность клетки в больших количествах транскриптов этих генов. Блок активно работающих рибосомных генов формирует особую структуру — ядрышко, где осуществляется синтез РНК- Гены, кодирующие рибосомные РНК, разделены

промежутками — спейсерами (англ. spacer), длина и структура которых может достаточно сильно различаться даже у близких видов. Ранее считали, что районы спейсера не транскрибируются. Действительно, электронно-микроскопические картины транскрипции рибосомных генов, казалось бы, подтверждали этот вывод. Однако оказалось, что транскрипты спейсеров подвергаются особенно быстрому процессингу и деградации. В конце спейсера находится сайт терминации T, включающий консервативную последовательность из нескольких нуклеотидов. Функциональная особенность этого сайта заключается в том, что он же отвечает и за реинициацию нового транскрипта. Одно и то же мутационное изменение в этом районе спейсера снижает как эффективность терминации, так и реинициации транскрипции. При освобождении транскрипта в сайте T полимеразы сохраняется связанной с ДНК. Взаимодействуя с дополнительным белковым фактором, полимеразы тут же настраиваются на следующий акт реинициации транскрипции. Благодаря сопряжению актов терминации и инициации в одном сайте обеспечивается высокая эффективность процесса транскрипции, поскольку каждый новый акт инициации не требует поиска промотора. В спейсере обнаружены повторяющиеся последовательности, которые резко, в десятки раз, активируют транскрипцию. Подобно энхансерам генов, транскрибируемых РНК-полимеразой II, такие повторы, включающие по 40—60 п. н., активируют транскрипцию вне зависимости от их ориентации по отношению к направлению транскрипции. Эффекты таких повторов также мало зависят от расстояния между ними и стартом транскрипции. Интенсивность транскрипции пропорциональна числу таких усилителей в составе спейсера. Однако механизмы, с помощью которых энхансеры, локализованные в районе транскрибируемого спейсера, оказывают влияние на транскрипцию рибосомных генов, могут быть совсем иными по сравнению с механизмом действия энхансеров генов, транскрибируемых РНК-полимеразой II. Повторяющиеся по длине спейсера элементы способствуют связыванию молекул РНК-полимеразы I с ДНК-матрицей в районе спейсера.

Предполагается, что спейсер наряду с молекулами РНК-полимеразы, осуществляющими транскрипцию, может содержать связанный пакет молекул полимеразы, тесно расположенных по длине спейсера и продвигающихся к промотору. В районе промотора осуществляется инициация транскрипции этими молекулами РНК-полимеразы. В результате тандемная организация рибосомных генов, разделенных длинными спейсерами, обеспечивающими связывание полимеразы и содержащими регуляторные сайты транскрипции, предназначена для обеспечения высокой эффективности функционирования рибосомных генов. Эффективность процесса, по-видимому, также значительно увеличивается благодаря компартментализации (концентрированию) всех компонентов системы в одной структуре — ядрышке.

Различия в структуре спейсеров у близких видов очень велики, поэтому спейсеры рассматривались как функционально нейтральные. Обнаружение важных регуляторных элементов в составе спейсера позволило поставить под сомнение общее представление о том, что в процессе эволюции быстро меняются те районы ДНК, которые функционально нейтральны. Напротив, можно предполагать, что здесь осуществлялась сравнительно быстрая совместная эволюция («коэволюция») функционально важных участков ДНК спейсера и белков, взаимодействующих с ними. Транс-действующие факторы транскрипции, связывающиеся с элементами промотора РНК-полимеразы I, не изучены.

Регуляторные элементы генов, транскрибируемых РНК-полимеразой
II

Промоторные элементы генов одноклеточных эукариот — дрожжей — содержат сайты инициации (И), нуклеотидную последовательность ТАТА (обычно ТАТААА1, а также другие элементы — активирующие

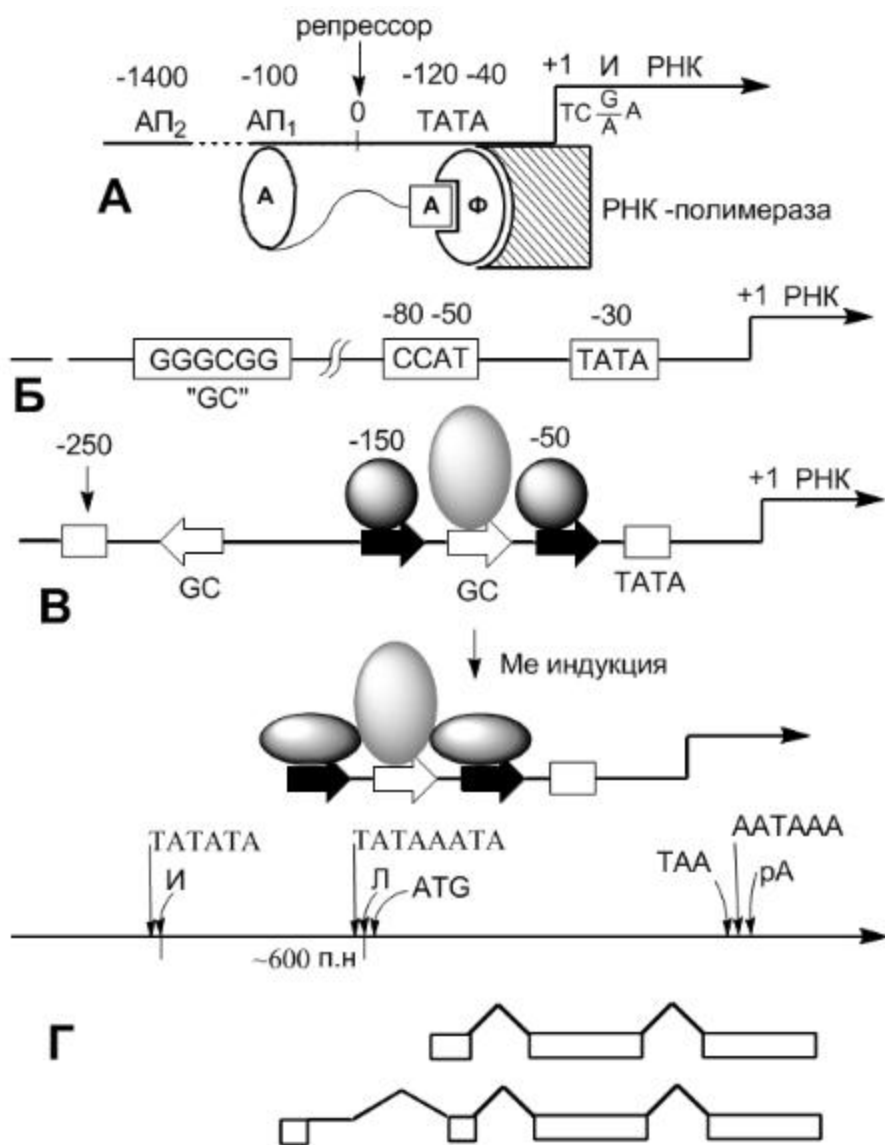


Рисунок 21: Структура промотора и регуляторных элементов генов транскрибируемых РНК-полимеразой II.

последовательности (АП, UAS, англ. upstream activating sequences), находящиеся перед сайтом инициации транскрипции. Кроме того, промотор может содержать элементы оператора О, участвующего в репрессии транскрипции. Расстояние между ТАТА-элементом и сайтом инициации может варьировать от 40 до 120 п. н., и в отличие, например, от промоторов позвоночных в

промоторах дрожжей правильная точная инициация транскрипции сохраняется при изменении расстояния между сайтом инициации и ТАТА-элементом. Инициаторный элемент представляет собой особый участок, включающий нуклеотидную последовательность ТС(G/A)А. Участок узнается РНК-полимеразой или связанными с ней белковыми факторами. Элементы АП являются сайтами связывания белков, регулирующих транскрипцию. Ряд их

свойств напоминает свойства усилителей транскрипции (энхансеров) в генах позвоночных, действующих вне зависимости от ориентации и на значительных расстояниях (около 1000 п. н.) от сайта инициации. Благодаря использованию большого набора мутаций по промоторам и генам активирующих белков дрожжей удалось выяснить некоторые особенности взаимодействия белков-активаторов с АП, а также характерные свойства этих белков. Белок GAL4 активирует гены, необходимые для катаболизма галактозы. GAL4 связывается с АП, представленной повторяющимися элементами по 17 п. н. Степень активирующего действия пропорциональна числу этих элементов в промоторе. Функция связывания ДНК и активации транскрипции принадлежит разным участкам белка GAL4, который содержит 881 аминокислоту. 73 остатка с N-конца молекулы белка достаточны для обеспечения специфического связывания с ДНК. Этот участок связывает ионы цинка и содержит характерную структуру — «цинковые пальцы», обнаруженные в целом ряде белков, активирующих транскрипцию. Два других дискретных участка белка, включающих аминокислоты 149—196 и 768—881, достаточны для обеспечения активации транскрипции. Эти участки содержат кислые аминокислотные остатки. По-видимому, в разных активаторных белках эти районы обладают сходными структурными особенностями, но не содержат гомологичных последовательностей. Предполагается, что эти белки стимулируют транскрипцию благодаря взаимодействию с другими факторами транскрипции, например с ТАТА-связывающимся белком, который в свою очередь взаимодействует с РНК-полимеразой. Белки — активаторы транскрипции, обладают, по крайней мере, двумя дискретными доменами, один из которых предназначен для узнавания ДНК, а другой для осуществления белок-белковых взаимодействий, играющих очень большую роль в образовании активного транскрипционного комплекса.

У высших многоклеточных эукариот (насекомых, позвоночных) в промоторная область варьирует в пределах 100—200 п. н. Перед стартом

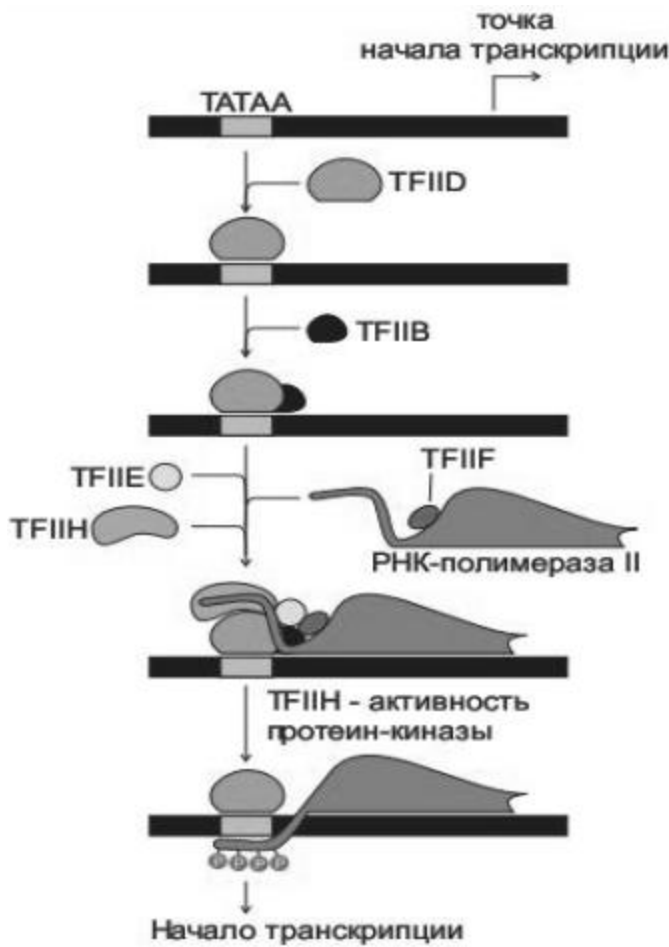


Рисунок 22: Схема инициации транскрипции РНК-полимеразой II.

транскрипции выявлена более сложная мозаика промоторных элементов, представленных короткими нуклеотидными последовательностями (“мотивами”). На расстоянии 27—30 п. н. от кэп-сайта расположен ТАТА-мотив, усредненный вариант которого (так называемый consensus) можно представить как ТАТА(Т/А)А(Т/А). Положение ТАТА-элемента строго определяет сайт инициации транскрипции, т. е. 5'-ко-нец транскрипта. При повреждении или удалении ТАТА-элемента образуется набор молекул РНК с разными 5'-концами.

Отдельные нуклеотидные замены в ТАТА-элементе могут приводить к резкому снижению эффективности транскрипции. Промоторная зона некоторых генов (например, генов лизоцима яйца птиц или гидроксиметилглутарилКоА-редуктазы — ключевого фермента биосинтеза холестерина у млекопитающих) не содержит ТАТА-элемента, и транскрипция начинается с нескольких разных сайтов. Образующиеся РНК различаются по 5'-концам в районе нетранслируемой лидерной последовательности. Возможно, различные лидерные зоны определяют характер регуляции экспрессии генов на уровне трансляции. То же самое наблюдается и для генов дрожжей, лишенных ТАТА-элемента в районе промотора.

В пределах 100—200 п. н. перед стартом транскрипции многих генов

находятся по крайней мере еще два коротких нуклеотидных «мотива», усредненные варианты которых можно представить как GGGCGG («GC-мотив») и ССААТ. Предполагается, что «GC-мотив», который может встречаться несколько раз по длине промоторной зоны, включающей 200—300 п. н., характерен для промоторных районов генов, работающих конститутивно и обеспечивающих общеклеточные функции. Действительно, такие повторы обнаруживаются перед генами, кодирующими белки, необходимые для жизнедеятельности самых разных клеток организма. Перед геном дигидрофолат-редуктазы мыши вкраплены четыре таких мотива, перед генами глюкозо-6-фосфатдегидрогеназы и фосфоглицераткиназы человека соответственно девять и два. Нельзя исключить, что «GC-мотивы» принимают непосредственное участие и в индукции генной активности, поскольку они обнаружены и в промоторах ряда индуцируемых генов. Последовательность ССААТ (СААТ* встречается в промоторной зоне разных тканеспецифических генов: в положении — 80—50 разных генов глобинов, в гене тиреоглобулина млекопитающих, в гене α -актина скелетных мышц цыпленка и других генах. Роль «мотива ССААТ» может быть достаточно важной в регуляции активности глобиновых генов, активируемых или репрессируемых на определенных стадиях развития. Репрессия синтеза γ -глобина плода у взрослого организма снимается в случае мутационного замещения одного нуклеотида в последовательности ССААТ. Мутация приводит к так называемой наследственной персистенции (сохранению синтеза γ -глобина плода у взрослых). В промоторах глобиновых генов отсутствует «GC-мотив». Молекулярные механизмы, с помощью которых описанные элементы промотора регулируют транскрипцию, еще не выяснены, но несомненно, что активность промоторных элементов обусловлена связыванием с определенными белковыми факторами, обеспечивающими точную и эффективную транскрипцию генов РНК-полимеразой II. Выделены разные белки, взаимодействующие с разными участками промотора, содержащими ТАТА, ССААТ или «GC-мотив». По-видимому, существует несколько белков, способных связываться с «мотивом

ССААТ», среди них — гетеродимер, состоящий из разных субъединиц. Белок, узнающий «GC-мотив», связывается с участком ДНК, включающим 18—20 п. н., в центре которого находится GC-элемент. Эффективность промотора, по крайней мере частично, определяется эффективностью отдельного элемента («мотива») в составе промотора, числом этих элементов и их взаимным расположением. Эти элементы, вероятно, функционируют в зависимости от ближайшего нуклеотидного окружения. Замены близлежащих нуклеотидов могут сильно сказываться на эффективности действия элемента. Так, например, замены выделенных жирным шрифтом нуклеотидов в окружении «GC-мотива» (GGGGCGGGGC) могут снижать активность промотора, тогда как замена первого G на T вполне допустима. Если область промотора содержит как GC, так и СААТ-элементы, то разные белковые факторы транскрипции, взаимодействующие с ними, могут согласованно активировать транскрипцию. Итак, область эукариотического промотора рассматривается как специфический ДНК-остов, на котором собираются белки транскрипции, узнающие свои сайты связывания и взаимодействующие как друг с другом, так и с РНК-полимеразой. Нельзя исключить, что факторы транскрипции являются ферментами и в процессе этих взаимодействий осуществляются ферментативные модификации как белковых факторов, так и ДНК. Появление нового фактора транскрипции в дифференцированных клетках можно рассматривать как способ включения гена на нужной стадии развития.

Формирование белкового комплекса на промоторной последовательности начинается с того, что фактор транскрипции TFIID связывается с ТАТА-доменом, который расположен выше точки инициации транскрипции примерно на 25 нуклеотидов. TFIID состоит в свою очередь из многих субъединиц. Те из них, которые ответственны за распознавание ТАТА-последовательности, называются ТВР (ТАТА-binding proteins). Теперь для начала транскрипции РНК-полимераза II должна освободиться от комплекса факторов транскрипции. Ключевым в процессе инициации транскрипции является присоединение фактора транскрипции TFIIF. Одна из его субъединиц, обладающая протеин-

киназной активностью, фосфорилирует молекулу РНК-полимеразы II. Установлено на примере по крайней мере нескольких генов, что это фосфорилирование освобождает РНК-полимеразу II и позволяет начать транскрипцию. В состав комплексов общих факторов транскрипции и РНК-полимеразы II входит до 50 белков, иногда эти комплексы называют транскриптосомами. Две другие РНК-полимеразы, I и III, найденные у эукариот, также требуют для активирования набора общих факторов транскрипции. С ними не все ясно. Установлено лишь, что белки ТВР требуются для всех трех полимераз. Другие факторы отличаются от тех, что были описаны в комплексах с РНК-полимеразой II. Кроме промотора в контрольной зоне находятся регуляторные последовательности, с которыми связываются регуляторные белки, необходимые для контроля процесса образования белкового комплекса на промоторе. Ниже приведены примеры организации регуляторных участков, обнаруженных у генов эукариот. Гены, отвечающие на тепловой шок, имеют в составе регуляторной зоны от трех до шести "элементов теплового шока"(HSE), которые представлены 14-членными последовательностями нуклеотидов. Известно, что гены теплового шока очень быстро реагируют на индуцирующее воздействие. В соответствии с таким типом индукции у них устроены регуляторные районы. В неиндуцированной тепловой шок клетке находятся белковые факторы HSF, они неактивны и не связываются с промоторными HSE-элементами. РНК-полимераза II занимает участок гена в районе точки инициации транскрипции. Под воздействием теплового шока белок HSF модифицируется и связывается с промотором гена hsp. Это активирует РНК-полимеразу II, либо воздействуя на ТАТА-домен, либо освобождая гипотетический белок "X", блокирующий РНК-полимеразу II. Некоторые из генов hsp индуцируются также гормонами. Соответственно, промоторы этих генов организованы сложнее. В них перемежаются контролируемые элементы обоих типов.

Энхансеры

Опыты с искусственными генными конструкциями, составленными из отрезков ДНК разного происхождения, выявили существование особого cis-действующего элемента регуляции генов эукариот, получившего название

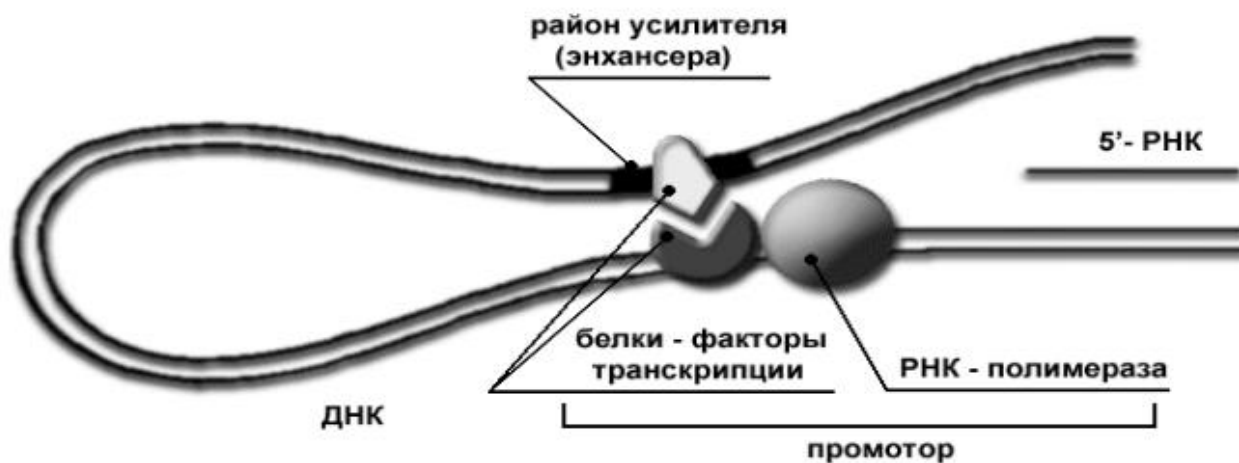


Рисунок 23: Схема взаимодействия энхансерного элемента с РНК-полимеразой II.

усилителя (энхансера) или активатора транскрипции. Энхансеры представлены короткими последовательностями ДНК, состоящими из отдельных элементов (модулей), включающих десятки нуклеотидных пар. Модули могут представлять собой повторяющиеся единицы. Энхансер увеличивает эффективность транскрипции гена в десятки и сотни раз. Впервые энхансеры были обнаружены в составе геномов животных ДНК-содержащих вирусов (SV40 и полиомы), где они обеспечивают активную транскрипцию вирусных генов. Извлеченные из вирусных геномов и включенные в состав искусственных генетических конструкций, они резко усиливали экспрессию ряда клеточных генов. Позднее были обнаружены собственные энхансеры генов эукариотической клетки. Особенность энхансеров состоит в том, что они способны действовать на больших расстояниях (более чем 1000 п. н.) и вне зависимости от ориентации по отношению к направлению транскрипции гена. Оказалось, что энхансеры могут располагаться как на 5'-, так и на 3'-конце фрагмента ДНК, включающего ген, а также в составе интронов. Например, энхансеры были выявлены в районе 400 п. н. перед стартом транскрипции генов инсулина и химотрипсина крысы. В случае гена алкогольдегидрогеназы

дрозофилы энхансер был локализован за 2000 п. н. Перед промотором. Энхансеры обнаружены на 3'-фланге гена, кодирующего полипептидный гормон-плацентарный лактоген человека, а также в составе интронов генов иммуноглобулинов и коллагена. Энхансеры могут содержать разные нуклеотидные последовательности, составленные из нескольких «нуклеотидных мотивов», каждый из которых обладает указанными особыми свойствами. Такие «мотивы» (модули) могут быть повторены в одном сайте или чередоваться друг с другом. В составе определенного семейства энхансеров (например, в наиболее изученных энхансерах геномов вирусов) можно выделить отдельные общие «мотивы», нуклеотидные замены в которых приводят к резкому снижению их биологической активности. Например, резко падает активность энхансера с последовательностью TGGAAAC в результате замены гуанилового нуклеотида, отмеченного звездочкой.

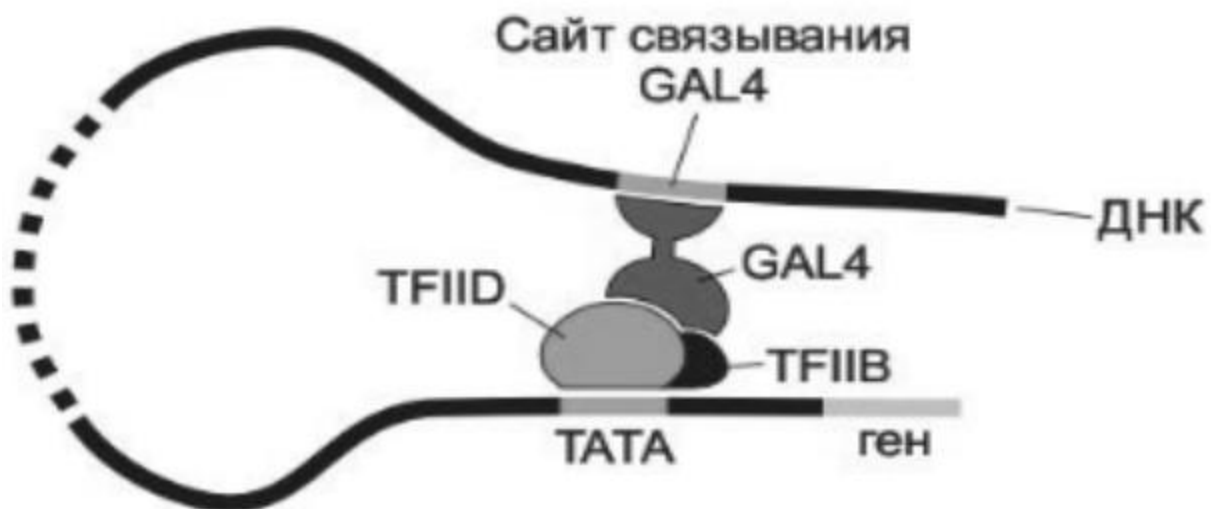


Рисунок 24: Взаимодействие белка GAL4 с участком UAS и транскрипционным комплексом

Принципы действия энхансеров, способных оказывать свое влияние на значительном расстоянии (более чем тысячи нуклеотидных пар) и вне зависимости от ориентации по отношению к старту транскрипции, не выяснены. Короткие нуклеотидные блоки могут служить центрами связывания специфических ядерных белков, выступающих как транс-действующие факторы. Сила энхансера, вероятно, может зависеть от числа таких блоков (модулей). Обсуждаются следующие два основных механизма действия

энхансеров. Считается, что функциональные участки генома, содержащие один или несколько генов, образуют длинные петли, включающие десятки тысяч нуклеотидных пар ДНК. Высказано представление, что петли закреплены в матриксе клеточного ядра и могут быть сверхспирализованы. В состав матрикса входит топоизомераза II, по-видимому, определяющая топологию петли ДНК. В таком случае взаимодействие энхансера с белками может менять конформацию всей петли, включая и удаленный от энхансера участок ДНК, в результате чего в составе петли изменяется локальная структура хроматина и облегчается транскрипция гена. Более вероятно, что влияние энхансера, связанного с белком, определяется его непосредственным взаимодействием с РНК-полимеразой и другими факторами транскрипции в процессе инициации. Такое взаимодействие может осуществляться благодаря сгибанию молекулы ДНК, что создает возможность непосредственного контакта районов промотора и удаленного от него энхансера, связанных со специфическими белками. Способность ряда энхансеров взаимодействовать со специфическими белками дифференцированной клетки, вероятно, обеспечивает их важное свойство — тканевую специфичность. Тканеспецифический энхансер впервые был выявлен в генах, кодирующих тяжелую полипептидную цепь иммуноглобулинов. При образовании функционирующего гена иммуноглобулина происходит программированная в развитии рекомбинация генетического материала. Один из нескольких сотен геномных сегментов, кодирующих варьирующую часть молекулы антитела (V-гены), в результате последовательных рекомбинационных процессов соединяется с D- и -J-элементами и, наконец, с районом, кодирующим константные C-районы иммуноглобулина и содержащим энхансер. В процессе транслокаций каждый V-ген сохраняет свой промотор. Однако только промотор, перенесенный к энхансеру и находящийся в составе функционально перестроенного гена, активируется в В-лимфоцитах, где ген должен работать. Энхансер, свойства и функции которого были доказаны с помощью искусственных генетических конструкций, располагается в районе интрона. Заметим, что приближение усилителя к V-генам — это лишь один из

многих этапов в сложном и многоступенчатом процессе активации экспрессии генов иммуноглобулинов. Эnhансер обнаруживает выраженную тканевую специфичность: его активность проявляется в лимфоидных клетках, но не в фибробластах или клетках легочной ткани. Эnhансеры с тканеспецифическим действием описаны и для других генов, например для генов инсулина и химотрипсина, которые не подвергаются перестройке в процессе клеточной дифференцировки. По мере исследования регуляторных элементов генома эукариот оказалось, что не удается провести строгих функциональных различий между эnhансерами и вышеописанными элементами промотора. Прежде всего ориентация таких активных элементов промотора, как «GC-мотивы», может быть любой относительно направления транскрипции гена. Более того, промоторные элементы генов металлотионеина и генов теплового шока, отделенные от ТАТА-последовательности, определяющей сайт инициации транскрипции, также сохраняют специфические регуляторные свойства вне зависимости от ориентации. Промоторный участок гена металлотионеина, включающий ряд регуляторных мотивов, способен сильно стимулировать экспрессию гена β -глобина и в том случае, если промотор расположить на 3'-фланге относительно направления транскрипции, причем на значительном расстоянии от β -глобина. Примечательно, что экспрессия β -глобина теперь будет зависеть от ионов металлов. В результате регуляторный участок гена металлотионеина можно рассматривать не только как промотор, но и как индуцируемый - эnhансер активности гена. Подобным образом специфический промотор генов теплового шока, отделенный от мотивов ТАТА и ССААТ и находящийся на значительном расстоянии от старта транскрипции, вне зависимости от ориентации по отношению к нему активирует ген β -глобина при повышении температуры. Таким образом, регуляторные элементы генов, которые первоначально относили либо к промоторным, либо к эnhансерам, обладают рядом общих функциональных характеристик. На физической карте регуляторной области гена они могут располагаться в одном районе, создавая сложную мозаику регуляторных сигналов.

У дрожжей известны элементы, функционально похожие на энхансеры, которые называются upstream activator sequences (UAS). Подобно энхансерам UAS могут функционировать в любой ориентации и на различных расстояниях до промотора. Однако, в отличие от энхансера, UAS не могут функционировать когда локализованы ниже промотора. UAS-последовательности активируются, когда с ними связываются, белки GAL4. В последние несколько лет появился новый метод активирования направленной экспрессии генов у дрозофилы. Суть его заключается в том, что исследователи создают две трансформированные линии мух, в одну из которых вводится GAL4-транспозон, а в другую UAS-транспозон. Для создания первой линии в Р-элемент, наряду с обычными маркерами вводится ген, кодирующий белок GAL4. Такой транспозон встраивается в случайные районы хромосом дрозофилы, в том числе может встроиться под какой-нибудь энхансер, например, под энхансер гена, работающего в клетках формирующегося крыла или ноги. Поэтому GAL4 будет экспрессироваться в клетках крыла или ноги, однако никаких последствий для этих клеток не будет, поскольку белок GAL4 может активировать транскрипцию другого гена лишь в том случае, если он свяжется с промотором UAS, в нормальном развитии функционирующий также только в геноме дрожжей. Поэтому и создают вторую линию дрозофилы, которую трансформируют транспозоном, содержащим испытуемый ген (X - на рис. 7.50.), сшитый с промотором UAS. Когда в результате скрещивания двух линий оба транспозона объединяются в генотипе потомка, ген GAL4 начинает функционировать в тех клетках, в которых активен его энхансер. Белок GAL4 связывается с промотором UAS, который и активирует ген "X". При этом активация "X" будет совсем не в той ткани, в которой он активируется своим собственным энхансером, а в той, в которой активен энхансер гена GAL4. Такая экспрессия называется эктопической, т.е. "не на своем месте". Рассмотрим один из наиболее интересных примеров, связанный с использованием GAL4-UAS транспозонов.

Регуляция транскрипции, осуществляемой РНК-полимеразой III

Транскрипция генов, считываемых РНК-полимеразой III определяется промотором, который лежит внутри этих генов и состоит, в случае генов тРНК, из двух блоков, разделенных отрезком ДНК длиной около 20 п.н. Первый блок начинается через 8-30 п.н. после точки инициации транскрипции и представлен в усредненном виде последовательностью TGGCNNAGTGG. Усредненная последовательность для второго блока - (+51 -+72) GGTTCGANCC.

Гены, транскрибируемые РНК-полимеразой III, разделяют на три класса.

К 1-му классу относят гены 5S рРНК.

Гены тРНК относят ко второму классу.

Гены третьего класса кодируют малые ядерные РНК, например, U6-РНК у высших эукариот, и содержат только регуляторные элементы, расположенные перед точкой инициации транскрипции. У большинства генов этого класса регуляторные элементы состоят из ТАТА-последовательности с центром в положении -30 и из проксимального элемента PSE (proximal sequence element) с центром в положении -60. Несмотря на наличие ТАТА-элемента, обычно определяющего положение точки инициации транскрипции, именно PSE фиксирует положение этой точки. Аналогичную функцию у промоторов РНК-полимеразы II выполняет ТАТА-последовательность. Состав белковых компонентов, образующих комплекс с PSE, до конца не выяснен. Обнаружены два различных PSE-связывающих белковых комплекса, названные SNAP и PTF. Активность промоторов третьего класса зависит от наличия TBP (ТАТА-связывающего белка). Белок TBP, входящий в состав фактора транскрипции TFIIIB, взаимодействует непосредственно с ТАТА-элементом промотора. Поскольку для инициации транскрипции на этих промоторах требуется наличие PSE-элемента, предполагают, что PSE-связывающий фактор взаимодействует с TBP-компонентом фактора TFIIIB.

Интересно, что промотор гена U6-РНК животных, узнаваемый РНК-полимеразой III, отличается от промотора гена U2-РНК, с которым взаимодействует РНК-полимераза II, только отсутствием ТАТА-элемента U2-

промотора. У растений оба этих промотора содержат такие регуляторные элементы, а изменение расстояния между ТАТА-элементом и расположенным перед ним регуляторным элементом превращает промотор U6 из Pol III-зависимого в Pol II-зависимый. Несмотря на наличие у промоторов этих РНК-полимераз характерных черт, механизмы, по которым РНК-полимеразы II и III различают свои промоторы, остаются невыясненными.

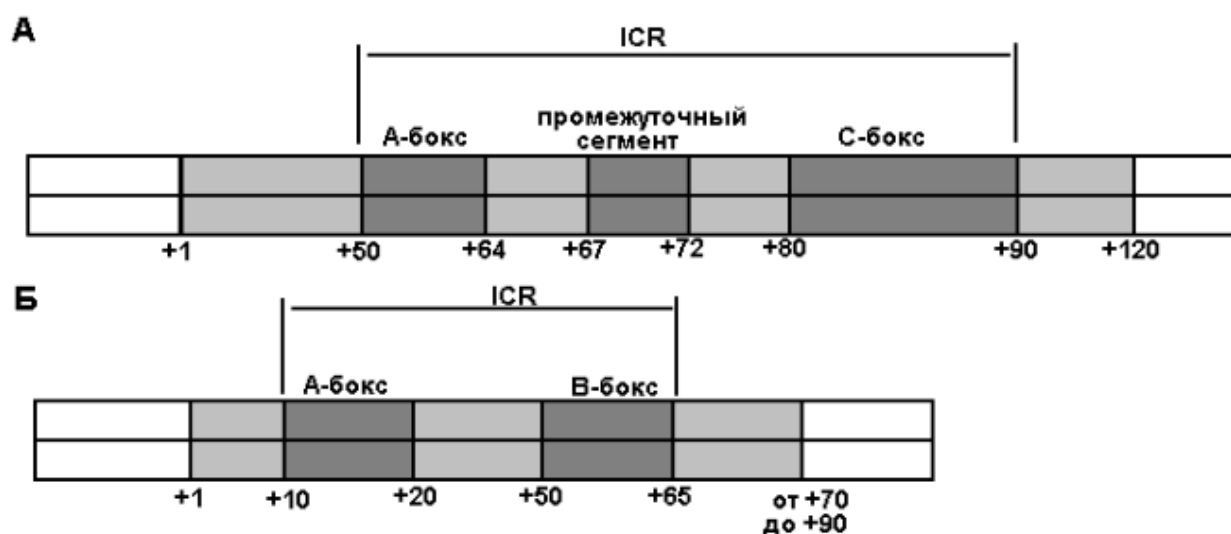


Рисунок 25: А. Структура внутренней регуляторной области (ICR) гена 5S-рРНК *X. laevis*. Эта область содержит А-бокс, промежуточный сегмент и С-бокс. Они выделены более темным цветом, чем кодирующая часть гена 5S-рРНК. Б. Внутренняя регуляторная область (ICR) гена тРНК *K*. Показано положение А- и В-боксов, которое в разных тРНК слегка различается. + 1 — сайт инициации транскрипции. Такое же попарное расположение А- и В-боксов характерно для генов 7SL-РНК. Канонические последовательности А- и В-боксов имеют вид 5'-TGGCNNAGTGG-3' и 5'-GGTTCGANNC-3' соответственно (N — любой нуклеотид).

РНК-полимераза III транскрибирует гены 5SРНК, тРНК и 7SL РНК. Гены 5S РНК, как и гены других рибосомных РНК, обычно представлены сотнями или даже тысячами тандемно повторяющихся копий, разделенных спенсерами. Эти гены собраны в одном или нескольких районах генома. Гены разных тРНК (в эукариотической клетке насчитывается 40—60 основных типов тРНК) также часто сгруппированы в кластеры, расположенные в разных хромосомах. Характер организации генов тРНК в составе таких блоков сильно варьирует у разных организмов. В целом организация генов тРНК в кластерах не столь

регулярна, как в случае рибосомальных РНК: направление транскрипции близлежащих генов бывает противоположным, а гены могут быть разделены нерегулярными промежутками, включающими 100—150, а иногда и несколько тысяч нуклеотидных пар. Природа и функции участков ДНК, разделяющих гены тРНК, не выяснены. Особенность участков генов 5S РНК и тРНК, узнаваемых факторами транскрипции и РНК-полимеразой и обеспечивающих инициацию транскрипции, состоит в том, что они локализованы непосредственно в транскрибируемом районе (или районах) гена. Так, район внутри генов 5S РНК лягушки, кодирующий нуклеотиды с 55-го по 80-й, необходим для инициации транскрипции. Размеры транскрипционного комплекса, включающего субъединицы полимеразы, достаточно велики, чтобы взаимодействовать с внутренними районами небольшого по размерам гена и одновременно инициировать транскрипцию. Сигналом терминации транскрипции служит последовательность из нескольких следующих друг за другом тимидиловых нуклеотидов. Роль спейсера в транскрипции, по крайней мере в опытах *in vitro*, пока не была обнаружена. Новообразованная 5S РНК, по-видимому, не подвергается модификации и процессингу. Внутри генов тРНК обнаружены два района, необходимые для того, чтобы осуществилась правильная и эффективная транскрипция. Например, для одного из исследованных генов тРНК это области, кодирующие нуклеотиды 8—25 и 50—58 молекулы тРНК. Два участка такого «расщепленного промотора» совпадают с районами гена, которые соответствуют наиболее консервативным районам молекул тРНК, содержащих наибольшее количество инвариантных нуклеотидов. Для активной транскрипции требуется сохранность расстояния между районами промотора, которое составляет около 50 п. н.. В то же время эту область гена можно заменить без ущерба на чужеродную. Внутри районов генов 5S РНК и тРНК, определяющих точную и активную транскрипцию, можно найти гомологии по нуклеотидной последовательности. Транскрипция генов 5S РНК и тРНК осуществляется с участием выделенных и очищенных белков — факторов транскрипции. Особенно хорошо изучен специфический

фактор транскрипции TF III A (англ. transcription factor) 5S-генов. Фактор представляет собой полипептид с $M_r=40\ 000$, он связывается с внутренним контролирующим элементом 5S-гена. Вслед за ним связываются два других белка и присоединяется РНК-полимераза. Одна из особенностей белка TF III A состоит в том, что он специфически связывается не только с ДНК, но и с 5S РНК. Поэтому при большой концентрации продукта транскрипции — 5S РНК — фактор связывается, а экспрессия 5S-генов подавляется по принципу обратной связи. Были рассмотрены три группы эукариотических генов, транскрипция которых осуществляется разными РНК-полимеразами при участии белковых факторов, взаимодействующих с характерными для каждой группы регуляторными элементами. Однако кроме них существуют еще «гибридные» системы транскрипции, в которых, по-видимому, одновременно могут использоваться способы регуляции, представленные в каждом из рассмотренных типов транскрипции. Так, РНК-полимераза III транскрибирует гены малых ядерных РНК типа U6, а также гены 7SK РНК «неизвестной функции» хотя те и другие не содержат «внутренних» промоторов и, напротив, на 5'-конце несут ряд элементов, характерных для систем транскрипции с помощью РНК-полимеразы II.

Все эти элементы аналогичны таковым у прокариот, и являются можно сказать функциональной ДНК, остальная часть генома представляется собой так называемую избыточную ДНК, которая, как считали ранее, не является функциональной. В настоящее время функции некоторых участков выявлены.

Структура собственно некодирующей ДНК

Некодирующая часть генома эукариот не только занимает основную часть генома, но и очень многообразна по составу. Можно выделить несколько видов:

- псевдогены;
- повторы;
- мобильные генетические элементы.

Все эти структуры по мнению большинства исследователей не имели определенных функций кроме ограниченного их числа, в частности теломерные

повторы участвуют в организации хромосом и их защите.

Но необходимо рассмотреть эти данные более подробно, так как вся некодирующая часть ДНК имеет определенные функции.

Псевдогены

Иногда в геномах встречаются т.н. псевдогены. Они имеют все необходимые черты генов, т. е. полный набор экзонов, характерных для этого гена, поли-А/Т хвост и короткие прямые повторы (как у мобильных элементов) в ДНК мишени, однако они не содержат элементов необходимых для транскрипции, таким образом данные рамки считывания не реализуются. В результате в данных фрагментах ДНК наблюдается высокий уровень мутагенеза. Псевдоген начинается с 5' точки, эквивалентной 5' точке мРНК, а заканчивается трактом поли-А нуклеотидов, что вероятно имеет происхождение от поли-А-конца в молекуле мРНК. У дрозофилы псевдогены встречаются редко, чаще это гены транспортных или малых ядерных РНК (тРНК, snRNA). Еще один известный пример - псевдоген личиночного кутикулярного белка и псевдоген *Adh* у *D. mulleri*, они сохраняют свои интроны. Пример процессированного псевдогена (т.е. без интронов) описан для гена *Adh* у *D. teissieri* и *D. Yakuba*.

Считалось, что псевдогены являются «осколками генома», и образуются в результате или рекомбинации, или обратной транскрипции мРНК, под действием ревертазы ретровирусов или ретротранспозонов, таким образом, функции для псевдогенов до конца не определены, хотя в некоторых случаях псевдогены могут рекомбинировать с истинными генами обуславливая многообразие продуктов, такое характерно при генной конверсии.

Повторы

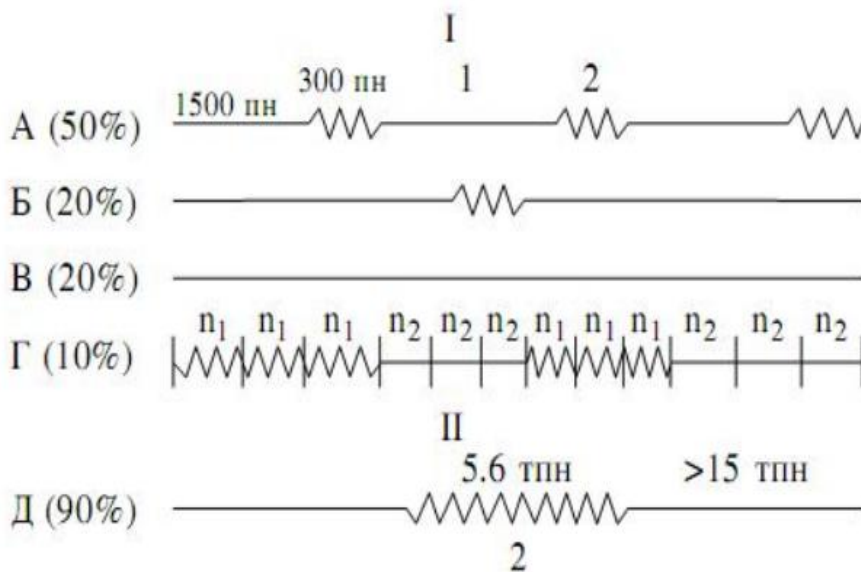
В конце 60-х годов американские ученые Р. Бриттен и Э. Дэвидсон открыли фундаментальную особенность молекулярной структуры генома эукариот - последовательности нуклеотидов разной степени повторяемости. Это открытие было сделано с помощью молекулярно биологического метода

изучения кинетики ренатурации денатурированной ДНК. Различают следующие фракции в геноме эукариот:

1. Уникальные, т.е. представленные в одном экземпляре.
2. Промежуточные (или среднечастотные) повторы. Это последовательности, повторяющиеся десятки и сотни раз.
3. Высокочастотные повторы, число которых в геноме достигает 10^6 копий.

Уникальные последовательности чаще всего представлены генами. Число генов у эукариот определяют одним из двух способов. Первый способ прямой, т.е. экспериментально определяют последовательности нуклеотидов во всем геноме, число последовательностей, содержащих длинные рамки считывания или кДНК-клоны. Понятно, что такой анализ можно провести пока на очень ограниченном числе видов, и эти исследования исключительно интенсивно проводятся в настоящее время (программы "Геном человека", "Геном дрозофилы", "Геном дрожжей" - всего около 30 таких программ). Выполнение подобных программ требует огромных финансовых затрат и скоординированных усилий большого числа ученых из всех развитых стран мира. Например, для того, чтобы расшифровать последовательность нуклеотидов в геноме дрожжей *Saccharomyces cerevisiae* потребовались усилия более чем 600 ученых из 96 лабораторий мира. Другой подход используют уже 2-3 десятка лет. С помощью довольно простых процедур рассчитывают возможное число генов у того или иного вида. Сначала определяют общий размер генома этого вида, затем, зная средний размер гена у этого вида и добавив к этому значению половину размера собственно гена (межгенный промежуток), делят значение размера генома на значение размера гена + межгенного промежутка и получают число генов. Все эти оценки в какой-то степени субъективны, поэтому варьируют в довольно широких пределах. Варьирование связано с тем, что разные авторы берут в расчет несколько различающиеся значения длин генов, межгенных промежутков, да и общий объем генома.

Если у прокариот уникальные повторы (гены) располагаются более или менее равномерно, то у эукариот распределение



менее равномерно, то у эукариот распределение назвать равномерным нельзя. Повторы образуют семейства - совокупность последовательностей, полностью или по большей части гомологичных друг

Рисунок 26: Схематическое изображение характера чередования последовательностей нуклеотидов разной степени повторяемости в геномах типа *Xenopus* (I) и типа *Drosophila* (II). 1 - уникальная последовательность, 2 диспергированная повторяющаяся последовательность, n₁ и n₂ - короткие повторы. А-Д - части генома и доля (%) генома, имеющие такой тип чередования последовательностей

другу. Нередко из-за существенных различий в нуклеотидном составе высокочастотных повторов и остальной

ДНК первые образуют при центрифугировании в градиенте плотности хлористого цезия так называемые сателлитные пики, которые имеют большую или меньшую плавучую плотность, чем остальная ДНК. Эта фракция генома представлена небольшим (10-15) числом семейств коротких (5-12 п.н.) повторов, образующих протяженные блоки. У огромного большинства видов эта фракция занимает не более 10% генома. Близкие виды, например мышь и крыса имеют совершенно различные высокочастотные последовательности: у крысы их нуклеотидный состав не отличается от основной ДНК, тогда как геном мыши содержит четкий АТ-богатый сателлит. Это означает, что высокочастотные повторы способны к быстрым изменениям в ходе видообразования. Остальные 90% генома эукариот построены по принципу чередования (интерсперсии) уникальных и повторяющихся последовательностей. Условно выделяют два основных типа интерсперсии,

получивших названия по тем видам, у которых они впервые были описаны: интерсперсия типа "ксенопус" (обнаружена у *Xenopus laevis*) и типа "дрозофила" (впервые описана у *D. Melanogaster*). Примерно в 50% генома *Xenopus laevis* уникальные последовательности из 800-1200 п.н. чередуются с повторяющимися, средний размер которых 300 п.н. В остальной части геномов типа "ксенопус" расстояния между соседними повторами значительно превышают 1-2 т.п.н. Структура генома типа "ксенопус" широко распространена, особенно среди животных. Млекопитающие и человек также относятся к этому типу организации генома. Особенность генома человека и других приматов составляют интерсперсные высокочастотные повторы длиной около 300 п.н. У человека эти повторы содержат сайт, разрезаемый ферментом рестрикции Alu I. Число Alu- подобных повторов достигает 5×10^5 - 10^6 копий. У дрозофилы параметры интерсперсии резко отличаются от видов с типом генома "ксенопус". Повторяющиеся последовательности длиной 5600 п.н. чередуются с уникальными, длина которых не менее 13000 п.н. Интересно отметить, что у *Musca domestica*, - вида, близкого *D. melanogaster*, геном устроен по типу "ксенопус". Этот факт прямо указывает на то, что в ходе эволюции возможны очень быстрые преобразования характера чередования последовательностей. Птицы по параметрам интерсперсии занимают промежуточное положение между типом "ксенопус" и типом "дрозофилы". Многие виды не могут быть отнесены ни к какому типу.

Высокочастотные повторы

У эукариот существенная часть генома состоит из коротких многократно повторенных последовательностей, состав оснований которых отличается от среднестатистического состава нуклеотидов основной части генома. Высокочастотные повторы составляют основную часть некодирующей ДНК, они достаточно многообразны и вариабельны, к этим повторам относятся: сателлитная ДНК, центромерная и теломерная ДНК.

Сателлитная ДНК

Повторенные последовательности ДНК могут быть охарактеризованы с

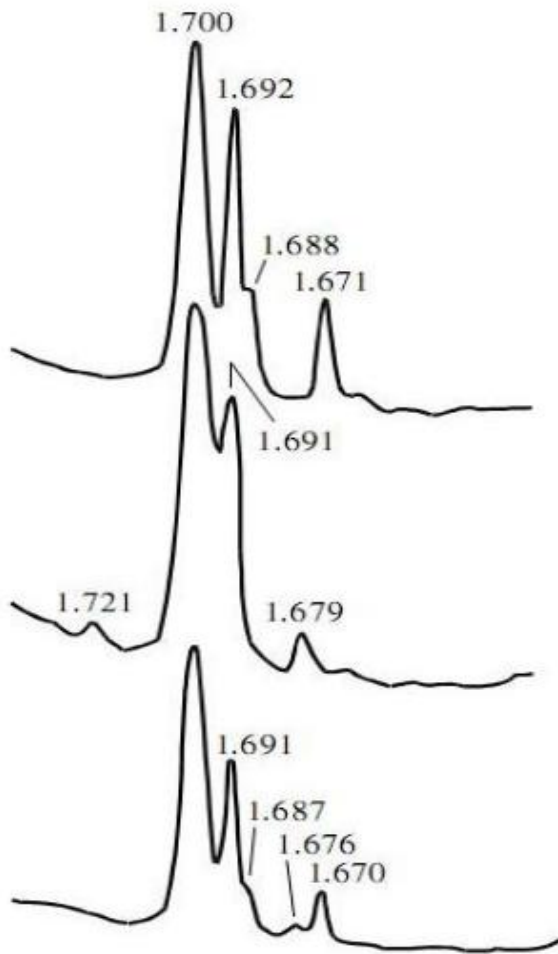


Рисунок 27: Распределение фракций ДНК из генома *Drosophila virilis* (вверху), *D. americana* (в центре) и их гибрида (внизу) в градиенте плотности нейтрального CsCl

помощью двух подходов: по их исключительно высокой скорости ренатурации или градиентном центрифугировании в хлористом цезии. В последнем случае основная часть ДНК составляет главную полосу осаждения (главный пик), а повторенная фракция в силу обогащенности определенным набором нуклеотидов и, следовательно, другой молекулярной массы, дает одну или несколько дополнительных (сателлитных) полос. Все данные, полученные до сих пор с помощью гибридизации *in situ*, свидетельствуют о том, что сателлитная ДНК располагается в участках прицентромерного гетерохроматина в метафазных хромосомах или в хромоцентрах интерфазных ядер.

Иногда, очень редко, некоторое количество сателлитных последовательностей локализовано в эухроматине. Кроме сателлитов в прицентромерном гетерохроматине располагаются умеренные повторы генов рибосомной ДНК. Анализ высоко-повторенных ДНК дрозофилы позволяет разбить их на три группы в соответствии с составом нуклеотидов и степенью повторенности:

1. Фракция, осаждающаяся в зоне 1,679 г/см³ и состоящая из последовательностей рибосомной ДНК;
2. Сателлиты, состоящие из повторенных 5-10 п.н.;

3. Сателлит, осаждающийся в зоне 1,688 г/см³. Это тандемный повтор длиной 359 п.н.

Клонирование коротких (300-600 п.н.) фрагментов, выделенных из различных трактов сателлитных ДНК с последующим определением последовательностей нуклеотидов показал следующее:

Всего найдено 11 типов сателлитных ДНК, четыре из них мажорные и семь - минорные.

Как правило, разные сателлиты есть в каждой хромосоме, однако, есть и сателлиты, специфичные для определенных хромосом, например, сателлит 359 п.н.

Как правило нуклеотидный состав повторов гомогенен, единичные замены нуклеотидов встречаются на 1 т.п.н., хотя в некоторых случаях частоты замен могут быть в 100 раз выше. Каждая горизонтальная полоса представляет копию повтора. Если есть замены, то указан новый нуклеотид в данном положении.

Среди сателлитных последовательностей часто располагаются мобильные элементы (т. е. умеренные повторы).

Высоко повторенные последовательности, как полагают, являются наиболее быстро эволюционирующими частями генома. Сравнение сателлитных профилей даже у очень близких видов выявляет большие различия. *D. virilis* 40% ДНК состоит из сателлитных последовательностей, у *D. texana* - 35%, но с другими константами седиментации, совсем не найдено сателлитов у третьего близкого вида - *D. ezoana*. 7-хромосома у многих животных, в том числе, дрозофилы, млекопитающих и человека состоит главным образом из повторенных последовательностей. Эта хромосома наследуется только от отца к сыну и не претерпевает рекомбинации в мейозе. Поэтому все индивидуальные изменения структуры уникальных и повторенных последовательностей как бы запечатлевают историю происхождения именно этой 7-хромосомы, присутствующей именно у конкретного мужчины. Очевидно, что все современные Y-хромосомы имеют единственного предка мужского пола. Это

неизбежно, хотя неясно каким был этот предок: совсем недавно существовавший и похожий на современного человека или же древний, доисторический. 7-хромосомы в разных популяциях могут иметь различные отклонения в молекулярной структуре: инсерции, деления, дубликации, инверсии, изменения в последовательностях крупных и мелких сателлитов, изменения длин рестрикционных фрагментов.

Комбинация этих различий (полиморфизмов) дает гаплотип. Наиболее переменные полиморфизмы позволяют делать различия между всеми независимыми Y-хромосомами. Менее переменные признаки позволяют находить что-то общее для объединения хромосом в группы. На сегодняшний день известно 9 гаплотипов. Они по-разному представлены в разных популяциях. Данные свидетельствуют о том, что концентрация тех или иных гаплотипов существенно варьирует в разных популяциях.

Долгое время считалось, что сателлитная ДНК не выполняет никаких функций и просто является избыточной, но сейчас существуют данные что именно сателлитная ДНК входит в состав гетерохроматина и отвечает за гетерохроматизацию геномной ДНК, кодирующие участки, прилегающие к сателлитной ДНК также могут подвергаться гетерохроматизации, таким образом можно менять транскрипционную активность генома в отдельно взятой клетке.

Активность хроматина регулируется ковалентными модификациями ядерных белков и ДНК. Метилирование цитозинового остатка в пятом положении пиримидинового кольца (m5C) является наиболее характерной модификацией ДНК у высших эукариот. Наследуемость молекулярных модификаций в ДНК создает основу для эпигенетической памяти. Важность метилирования ДНК в процессе развития убедительно доказана для животных и растений. Распределение метилцитозина в геноме далеко не равномерно. Большинство метилцитозиновых остатков сконцентрировано в гетерохроматиновых областях, сателлитных повторах, в неактивных транспозонах и эндогенных вирусах. С другой стороны, промоторные области некоторых генов, ассоциированных с CpG островками, и внеядерные геномы

(ДНК пластид и митохондрий) обычно обходятся без метилирования. У животных метилирование обнаруживается в симметричных CG динуклеотидах, в то время как в растениях почти каждый цитозин может быть метилирован, хотя преобладают симметричные m5CG и m5CNG мотивы. У нитевидных грибов *Neurospora* и *Ascobolus* большая часть генома свободна от метилирования за исключением некоторых сильно метилированных повторяющихся последовательностей. ДНК некоторых организмов с малым размером генома, например *S. Cerevisia*, *D. melanogaster*, *C. elegans*, не метилированы и программы их развития осуществляются с использованием других эпигенетических модификаций хроматина. Общий уровень метилирования у растений выше, чем у животных, что возможно отражает наличие большего количества сателлитных последовательностей (и гетерохроматина) в их геномах. Наблюдаются значительные различия в уровне метилирования у разных видов растений. К примеру, в то время как у *Arabidopsis* метилировано 6% цитозина, у большинства других покрытосеменных эта цифра достигает 20–30%. Эти различия можно объяснить как количеством повторяющихся последовательностей, так и разным уровнем метилирования CNG последовательностей.

Ферменты, участвующие в метилировании

Наблюдается рост числа изолированных генов (белков), связанных с процессом метилирования ДНК. Перенос метильной группы на цитозинный остаток является пострепликативным ферментативным процессом, катализируемым ДНК+метилтрансферазами. Эукариотические метилтрансферазы по сравнению с прокариотическими ферментами являются большими молекулами, часто размером более 100 кДа, и содержат разные функциональные домены. Хромодомены, обнаруженные недавно в малом семействе хромометилаз (СМТ+семейство), указывают на существование связи между структурой хроматина и метилированием ДНК. Высказано предположение, что хромодомены могут направить фермент к определенному геномному участку. В самом деле, у *Arabidopsis* и кукурузы разрушение

хромометилазных генов приводит к специфическому снижению метилирования сателлитных последовательностей, указывая на сродство ферментов этого типа с гетерохроматиновыми участками. Обе мутации влияют больше на CNG, чем на CG метилирование. Механизм узнавания гетерохроматина возможно включает в себя взаимодействие фермента с гетерохроматин-специфическим HP1 белком у *Arabidopsis*. Интересно отметить, что взаимодействие HP1 с нуклеосомами происходит через метилированный гистон H3, связывающий при этом две эпигенетические модификации, основанные на метилировании. Функции CNG метилирования в повторах остаются неясными, так как ни один из хромометилазных мутантов не обнаруживает фенотипических отличий. Обнаружение *ddm1* мутации у *Arabidopsis* в 90-ые годы показало, что метилирование эукариотического генома является сложным процессом, требующим помимо участия метилтрансфераз других факторов. В первых генерациях *ddm1* мутантных растений наблюдается резкое снижение уровня метилирования сателлитных гетерохроматиновых участков – до 70%, в то время как гипометилирование в уникальных последовательностях обнаруживается лишь после нескольких поколений инбридинга. Оставалось загадкой, почему это происходит, когда у мутантных растений сохраняется нормальный уровень активности ДНК+метилтрансферазы и нормальный уровень кофакторов метилирования. Недавно мутации были увязаны с геном, кодирующим член семейства SNF2/SWI2 ДНК+зависимой АТФ+азы, реконструирующим хроматин. Возможно в будущем будут открыты белки (гены), которые либо прямо (путем связывания метилтрансфераз через хромодомены), либо косвенно, например, через факторы, регулирующие доступность к хроматину, участвуют в процессе метилирования.

Клеточные функции метилирования

Становится очевидным, что метилирование цитозина может выполнять и плейотропные функции в зависимости от организма и локализации генома. У высших эукариот метилирование может быть вовлечено в число важных клеточных функций, таких как выключение генов, геномный импринтинг,

инактивация транспозонов, тканеспецифическая экспрессия и эпигенетическая инактивация хромосомных доменов. Из этого следует, что обобщенная формулировка общей функции метилирования ДНК у эукариот является неправильной. В этом обзоре мы рассмотрим вопрос о возможной роли метилирования ДНК в сателлитах. Имеется ряд работ, показывающих высокий уровень метилирования цитозина в сателлитных повторах в разных участках генома как у животных, так и у растений. Считается, что сателлиты, как и все повторяющиеся последовательности, являются исключительно хорошими мишенями для метилирования *de novo*. Доказательством этого являются полученные в трансгенных экспериментах данные, показавшие, что многие tandemно расположенные трансгенные копии подвергаются метилированию и отключению их экспрессии в процессе, именуемом умолчанием генов под воздействием повторов. Сигнал метилирования часто включает гомологозависимое взаимодействие между повторяющимися единицами (как tandemно расположенными, так и дисперсными) на уровне ДНК или РНК. Очевидно, что эндогенные сателлиты могут быть естественной мишенью для выключения генов под воздействием повторов и метилирования. Эндогенные вирусы являются хорошим примером того, как одна или несколько копий чужой последовательности, однажды интегрированные в геном, могут расширяться, превратиться в сателлиты и претерпеть новые эпигенетические модификации. В недавней эволюционной истории *Nicotiana* имела место интеграция растительной вирусной ДНК. В настоящее время ряд видов содержит около сотни копий амплифицированной и высокометирированной вирусной последовательности в геноме. Несомненно, не все сателлиты являются равномерно метилированными, к тому же не существует точной корреляции между числом копий и уровнем метилирования. Локусы рДНК являются хорошим примером повторяющейся последовательности, содержащей как гипо-, так и гиперметирированные единицы. Гипометирирование рДНК связывают с транскрипционной активностью и понижением метилирования, вызываемое 5-азациитидином, сопровождается повышенной транскрипцией исходно молчащих

единиц у аллополиплоидов Brassica. Метилирование может способствовать компартиментализации рДНК на активные и молчащие участки в клеточных ядрах. Возникает вопрос о роли метилирования цитозина в сателлитах, которые никогда не транскрибируются и образуют конститутивный гетерохроматин.

1. Берд высказал предположение, согласно которому метилирование предотвращает нежелательный транскрипционный «шум» у видов с большим геномом. Частота сайтов, связывающих факторы транскрипции может быть высокой в геноме, значительную часть которого составляет гетерохроматин. В результате этого, следовые транскрипции, инициированные в гетерохроматине, могут достичь чрезвычайно высокого уровня. Метилирование цитозина может непосредственно влиять на транскрипцию путем предотвращения связывания транскрипционных факторов с сайтом узнавания или же непрямым путем – образованием конденсированной структуры хроматина, недоступной для транскрипционных комплексов. Выделены белки с высоким сродством к метилцитозину у животных и показана их высококооперативная связуемость с ДНК. Возможно, некоторые из этих белков способствуют конденсации хроматина.

2. Хотя транспозоны не рассматриваются как классические сателлиты, они могут составлять до 50% повторяющейся ДНК у некоторых видов растений, например кукурузы. У *Arabidopsis* мутации, влияющие на метилирование, могут увеличить активность транспозонов. И наоборот, активные транспозоны часто содержат меньше метилцитозина, чем их неактивные копии. Таким образом, метилирование может контролировать транспозиционную активность.

3. Другая роль метилирования цитозина выяснилась в ходе недавних исследований в разных экспериментальных системах. У *Ascobolus* частота кроссинговера между двумя генетическими маркерами снижалась в сотни раз, в тех случаях, когда взаимодействующие ДНК были метилированы. Этот результат показывает, что метилирование может влиять на ДНК-ДНК взаимодействие и подавлять генетическую рекомбинацию в организме. Было

показано, что редкий у людей ISN-синдром может быть связан с мутацией в метилтрансферазном гене DDM3. Классическая сателлитная ДНК обычно сверхметилирована, но при ISN-синдроме почти полностью неметилирована. Цитогенетический анализ ISN-пациентов показал наличие хромосомных aberrаций в центромерном гетерохроматине, в основном его удлинение в метафазе. В интерфазе наблюдалась устойчивая самоассоциация околоцентромерных сателлитов. Экспериментальное снижение метилирования вызывает aberrантную конденсацию и разделение сестринских хроматид в субтеломерном гетерохроматине у пшеницы. У растений с большим геномом сателлитные повторы, состоящие из сотен тысяч единиц, часто обнаруживаются в разных хромосомах. Возможно метилирование способствует предотвращению нежелательных рекомбинаций между гомологичными сателлитами в сдвинутых позициях.

Таким образом, можно предположить, что метилирование сателлитных последовательностей может выполнять различные функции – понижать транскрипционный «шум», инактивировать транспозоны и стабилизировать тандемно расположенные единицы в клеточном ядре. Несмотря на значительные достижения в понимании метилирования цитозина, все еще остается много вопросов.

1. Необходимо ли метилирование для сателлитов? Известно, что сателлиты одноклеточных дрожжей и *Drosophila* выполняют свои функции без модификации цитозина.

2. Каким образом некоторые сателлитные единицы не метилируются, тогда как другие гиперметилированы?

3. Как «машина» метилирования узнает повторы? Определяется это структурными свойствами ДНК, например изгибами, факторами хроматина и/или же другими эпигенетическими модификациями, например метилированием, ацетилированием, и/или фосфорилированием гистонов?

4. Влияет ли метилирование на мутабельность и эволюцию сателлитов?

Давно известно и широко обсуждается в литературе то обстоятельство, что количество ДНК в гаплоидном наборе хромосом эукариот значительно превышает количество ДНК, содержащееся в структурных и регуляторных генах. Этот парадокс объясняют тем, что эукариоты содержат избыточную ДНК, функциями которой не являются кодирование белков или регуляция экспрессии структурных генов. По мнению этих авторов, избыточная ДНК является ненужной, как бы случайной. Она не влияет на фенотип, и ее единственная функция заключается в выживании внутри генома. В роли этой ДНК, названной «эгоистичной», может выступать ДНК с простой последовательностью, а также умеренно повторяющиеся ДНК. К эгоистичной ДНК относят и часть уникальных ДНК, которые не кодируют какие-либо белки и не участвуют в регуляции экспрессии генетической активности. Эгоистичная ДНК не вносит никакого вклада в определение фенотипа. Она лишь несколько отягощает в энергетическом отношении клетку, в которой она содержится. Появление эгоистичной ДНК в геноме сравнивают с распространением не особенно вредного паразита внутри хозяина. В то же время предполагается, что иногда клетка использует некоторую часть эгоистичной ДНК. Одна из возможных областей применения – это контрольные механизмы генной активности. По мнению авторов гипотезы, амплификация или делеция эгоистичной ДНК может происходить путем неравного кроссинговера в мейозе или митозе. Процессы образования и элиминации последовательностей эгоистичной ДНК должны быть сбалансированными, хотя ситуация может постоянно меняться. Вышеуказанной гипотезе более двадцати лет, однако до сих пор нельзя четко сказать, является ли сателлитная ДНК действительно ненужной (junk DNA – мусор) или же обладает какими-либо биологическими функциями. Наличие сателлитных ДНК почти у всех эукариот (если не у всех) может служить аргументом в пользу существования у них какой-либо общей функции, но точных доказательств этому почти нет.

В то же время появились работы, которые приписывают сателлитным ДНК ряд функций. Например, считается, что они необходимы для образования

центромер. Высказано предположение, что сателлитные ДНК Y-хромосомы обеспечивают хранение белков семенников у *Drosophila*. Обнаружена рибозимная активность сателлитной ДНК pDo500 у сверчка *Dolichopoda*. Недавно Ксинк и Хеникофф предложили более общую модель функционирования сателлитных ДНК. По их наблюдениям, определенные белки в интерфазе связываются с эухроматическими участками, а в метафазе – с сателлитными ДНК. Согласно этой модели сателлитные кластеры ДНК, связываясь с этими белками, обеспечивают стабилизацию митотического процесса. Хотя эта модель кажется привлекательной, мы все еще далеки от понимания биологических функций этих молекул. Более того, все еще нельзя считать доказанным существование таких функций вообще.

Центромеры

Центромеры представляют собой очень сложную специализированную структуру хромосом. Цитологически они обнаруживаются в виде первичной перетяжки в метафазных хромосомах и участвуют в точном распределении хроматид в митозе и мейозе. Центромеры выполняют ряд функций: сцепление сестринских хроматид перед анафазой, предотвращение начала анафазы при отсутствии точного биполярного прикрепления веретена, связывание хромосом с аппаратом веретена и содействие движению хромосом. У большинства многоклеточных эукариот микротрубочки аппарата веретена не взаимодействуют непосредственно с ДНК /хроматином центромерной области а контактируют с белковой структурой, именуемой кинетохором, который со своей стороны тесно ассоциирован с центромерным хроматином. Иногда термин центромера используется для обозначения ассоциации центромерной ДНК (CEN ДНК) со связывающими ее белками (хроматин) и кинетохором. Интенсивные исследования по выяснению строения и функций этих структур отражены во многих статьях и обзорах.

Структурная варибельность центромер

К настоящему времени интерес исследователей сфокусирован на

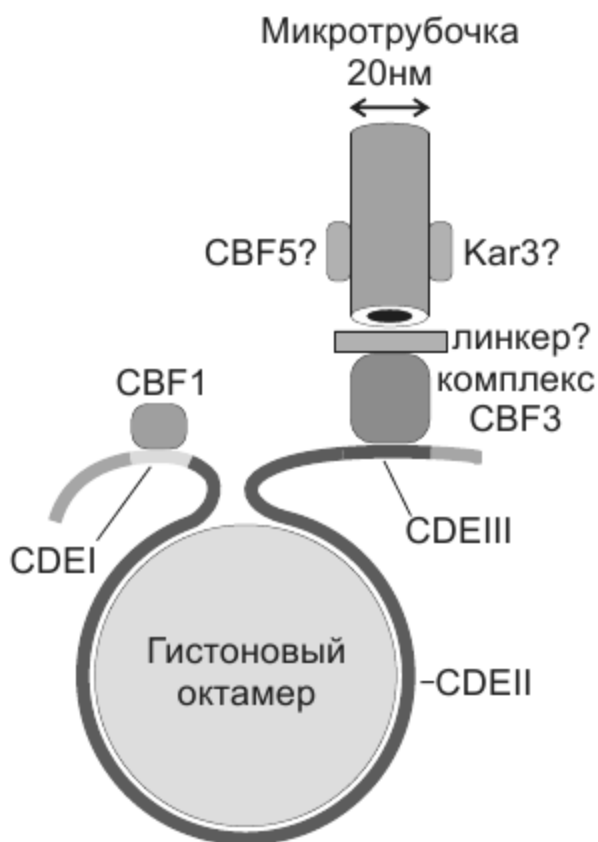


Рисунок 28: Функциональная центромера *S. cerevisiae*, которая охватывает только 125 п.н. и содержит три консервативных элемента ДНК – CDEI, CDEII и CDEIII, третий из которых (длиной 25 п.н.) абсолютно необходим для функционирования центромеры.

хромосомах многоклеточных эукариот с так называемыми локализованными центромерами. Следует однако отметить, что другие организмы показывают значительную вариабельность структуры центромер. К примеру, у некоторых простейших и *S. cerevisiae* микротрубочки не вставлены в сложный аппарат кинетохора. В голоцентричных хромосомах, обнаруживаемых у хорошо изученных *Caenorhabditis elegans* и других организмах (например, некоторых членистоногих), (голо)-кинетохор расположен вдоль всей или значительной части хромосом. Полицентрические хромосомы характеризуются многочисленными кинетохорами, которые распределены вдоль хромосомы, например, у некоторых покрытосеменных (*Luzula*, *Cyperus*) и членистоногих. Организация митотических и мейотических центромер/кинетохор может быть разной даже у одного и того же организма. В отличие от них, локализованные центромеры имеют одну область, которая образует перетяжку (первичная перетяжка) по отношению к хромосомным плечам. Имеется два класса локализованных центромер – точечные центромеры и региональные. Точечные центромеры упакованы в малые, компактные локусы и обнаруживаются у некоторых одноклеточных эукариот, таких как *S. cerevisiae*, *Schizosaccharomyces uvarum*, *Kluuyveromyces lactii*, *Candida maltosa*, *Candida*

glabrata и *Yarrowia lipolytica*. Наиболее хорошо изучена функциональная центромера *S. cerevisiae*, которая охватывает только 125 п.н. и содержит три консервативных элемента ДНК –CDEI, CDII и CDEIII, третий из которых (длиной 25 п.н.) абсолютно необходим для функционирования центромеры. Эти элементы связывают разные белковые факторы, Cbf1p, Cse4p и CBF3, которые, со своей стороны, контактируют с другими белками и микротельцами. Такая простая структура точечных центромер, по-видимому, является результатом эволюции, так как другие грибы содержат большие региональные центромеры. Вместе с тем, обнаруживается эволюционная консервативность на разном уровне. Некоторые белки центромер *S. Cerevisiae* проявляют сходство с белками кинетохор других эукариот. Например, Cse4p представляет собой гистон H3-подобную молекулу, которая связывается с CDEII. Анализ человеческого центромерного полипептида CENP-A показывает, что этот белок также произошел от гистона H3. Другой пример –это ограниченная гомология центромерного белка дрожжей Mif2p с человеческим центромерным белком CENP-C и белком HCP-4 *C. elegans*. Региональные центромеры имеют более сложное строение ДНК. Их размер варьирует от 40 т.п.н. до нескольких мегабаз. Большинство этих ДНК представлено повторяющимися элементами. Центромера *S. pombe* содержит разные повторяющиеся последовательности в виде длинных обращенных повторов, которые фланкируют 4–7 т.п.н. уникальную центральную область. Центромерные последовательности разных хромосом *S. Pombe* умеренно консервативны. Центромерные ДНК *S. cerevisiae* и *S. pombe* довольно сильно различаются по строению. Необходимо заметить, что структура центромеры *S. pombe* сходна с центромерами некоторых других эукариот. У *Drosophila* размер минимальной функциональной центромеры составляет 420 т.п.н. Эту центромеру можно идентифицировать в экспериментах с использованием производных минихромосомы Dp1187, полученных облучением. Они содержат участки многократно повторяющихся tandemно расположенных последовательностей, перемешанных со сложной ДНК, состоящей из уникальных и среднеповторяющихся последовательностей.

Около 350 т.п.н.

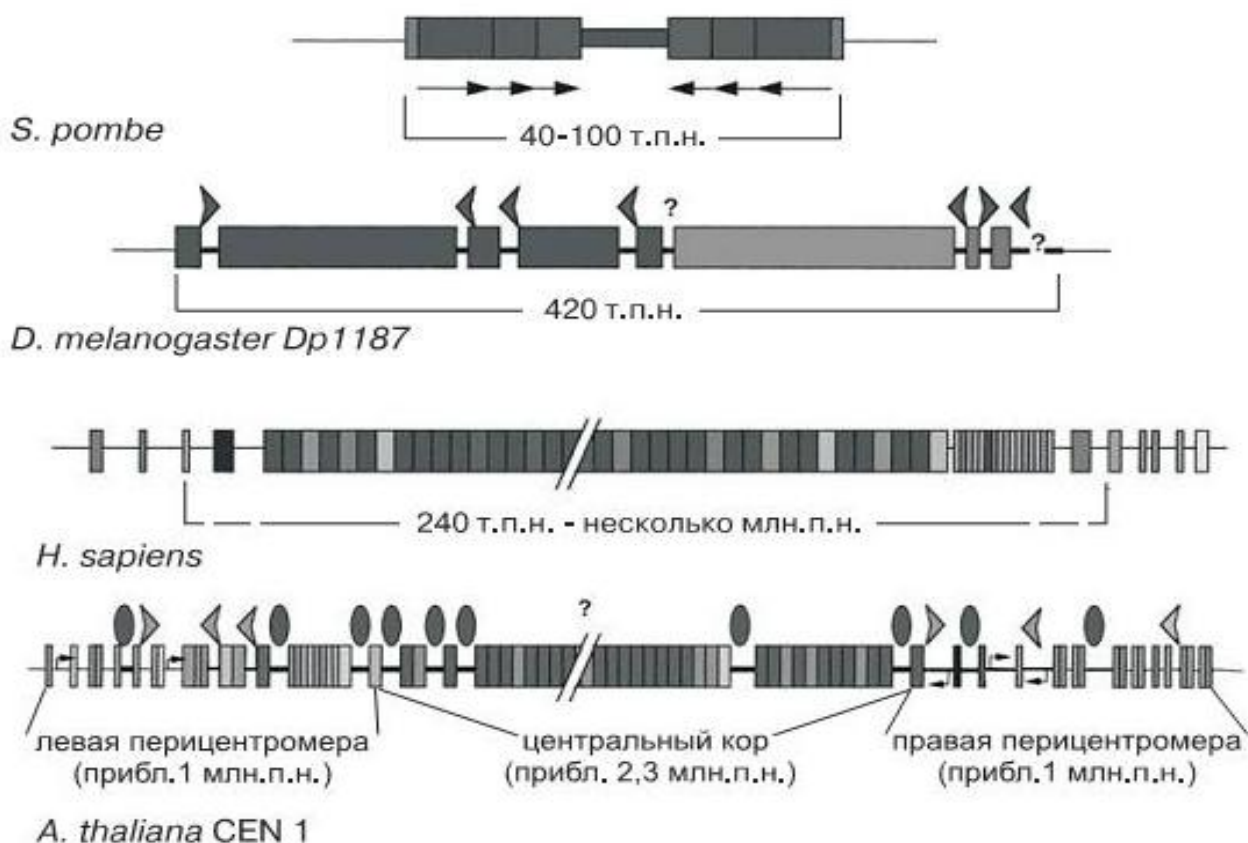


Рисунок 29: Структура центромер в хромосомах различных животных.

Пояснения в тексте

этого участка состоит из двух простых повторов. В остальной части перемешаны целый транспозон и новая АТ-богатая последовательность. Центромерный транспозон обнаруживает значительную консервативность по сравнению с участками вне центромер. Центромеры животных расположены около или внутри длинной повторяющейся ДНК, которые охватывают участок длиной от 500 т.п.н. до нескольких мегабаз. У приматов основная часть этих центромер представлена тандемной АТ+ богатой повторяющейся последовательностью длиной 171 п.н., именуемой альфоидной ДНК. Иногда встречаются другие элементы типа Alu или ретротранспозонных элементов L1 (например, центромеры человека). Бросается в глаза, что за исключением *Drosophila*, длина сателлитного повтора довольно близка у разных организмов, включая растения, и соответствует нуклеосомной длине. Знания относительно центромер многоклеточных эукариот довольно ограничены. Однако некоторые публикации, появившиеся за последние годы, значительно улучшили

понимание структуры и функции центромер. Сделано несколько важных наблюдений, касающихся центромер дрожжей, животных и человека. Хотя структурный анализ этих центромер не закончен, имеются значительные сомнения в существовании какой-то еще не обнаруженной «магической последовательности», способной быть «ядром» сворачивания большой области в функциональную центромеру. Этот взгляд поддерживается рядом исследований о наличии нецентромер. Полученные данные свидетельствуют об отсутствии какой-то определенной последовательности ДНК и подтверждают точку зрения, согласно которой центромеры являются эпигенетической конструкцией ДНК-белкового комплекса, т.е. изменяющейся структурой, которая наследуется без соответствующего изменения последовательности ДНК. Другие интересные исследования касаются состава центромерных последовательностей. Несмотря на дивергенцию последовательностей, региональные центромеры дрожжей и животных имеют сходную «архитектуру». Центральная центромерная область состоит из нескольких однообразных сателлитных мотивов, фланкированных перицентромерными участками более сложной ДНК, включая разные повторяющиеся элементы, которые отличаются от основного сателлитного повтора. Эти фланкирующие участки всегда образуют гетерохроматин. Возможно гетерохроматиновые перицентромеры представляют собой границу, которая отделяет гетерохроматиновое центромерное ядро от эукариотических хромосомных плеч. Ряд исследований подтверждает, что упомянутые структурные домены одновременно являются и функциональными доменами. Не вызывает сомнений тот факт, что внутренняя область центромерного ядра является областью связывания кинетохор. Однако эта область недостаточна для образования всей функционирующей центромеры. Возможно перицентромеры имеют важное значение для сцепления сестринских хроматид или сестринских центромер. Следует заметить, что област перицентромерного гетерохроматина является поздно реплицирующейся. Ее репликации предшествует репликация центромерного кора, которой, в свою очередь, предшествует репликация

эухроматиновых областей. По содержанию белков эти области четко различаются по ассоциации и/или модификации специфических белков. Наиболее заметна разница между эухроматином и центромерным гетерохроматином, которая заключается в наличии варианта гистона H3 (CENP+A) в перицентромерной области, замещающего обычный H3 гистон в нуклеосомах. Обнаруженный поначалу в человеческих клетках гомолог H3 гистона позднее был найден в таких разных организмах, как дрожжи, *C. elegans* и *Drosophila*. Эти центромерные H3 гистоны соответствуют нецентромерному H3 гистону своим C-концевым доменом, однако NH₂-концевые домены у них различны. В моделях, объясняющих хроматин-зависимое эпигенетическое воспроизводство идентичности центромер, CENP-A и их гомологи играют существенную роль. При образовании новых центромерных нуклеосом «нормальные» H3-содержащие нуклеосомы должны исчезнуть, чтобы избежать постепенную потерю идентичности центромер. Это может быть достигнуто образованием новых центромерных нуклеосом при помощи старых («модель узнавания»). Медиаторами этого процесса могут служить CENP-A гистон и специфические хроматиновые факторы. Альтернативным (или дополнительным) вариантом к этому процессу может быть доступность центромерных гистонов только в ограниченное время и в ограниченном ядерном участке в течение клеточного цикла. Предполагается, что поздно реплицирующийся (пери-) центромерный гетерохроматин осуществляет этот процесс при помощи пространственного исключения нецентромерного H3 гистона из этих доменов с одновременным накоплением центромерного H3 гистона в то же время или в течение поздней репликационной стадии («секвестровая модель»). При помощи этой модели можно объяснить образование нецентромер, что подчеркивает важную роль гетерохроматина для точного функционирования центромер. Рассмотрение накопившихся генетических, цитологических и молекулярных данных подтверждает, что региональные центромеры растений могут иметь сходную с животными центромерами структурную и функциональную организацию.

Теломеры

Три группы фактов, полученных к началу 1980-х годов, свидетельствовали о наличии на концах хромосом особых структур:

Цитологи, изучавшие кариотипы растений или животных, отмечали, что морфология отдельных хромосом и наборы хромосом в целом являются постоянными. Если бы была возможность хромосомам склеиваться "конец - в конец", такого постоянства бы не было. Значит существует некая структура, запрещающая объединение теломерных концов.

1. Г. Дж. Меллер в 1927 году открыл возможность получения мутаций в результате облучения клеток дрозофилы рентгеновскими лучами. При этом мутации могут быть точковыми или сопровождаться перестройками хромосом (инверсиями, транслокациями, делениями и дупликациями). Перестройки образуются в результате разрывов в различных участках хромосом под действием излучения и последующего воссоединения фрагментов в новом порядке. Исследуя закономерности образования хромосомных перестроек Г. Меллер в 1932 году пришел к двум важным выводам о поведении наиболее дистальных частей нормальных хромосом:

2. Искусственно образовавшиеся концы фрагментов хромосом объединяются только с такими же искусственно образовавшимися и никогда - с концевыми фрагментами, существующими в нормальных хромосомах.

3. В свою очередь концевые участки нормальных хромосом в результате облучения никогда не перемещаются во внутренние районы хромосом. Для объяснения результатов этих наблюдений Меллер предположил, что участки хромосом (или "гены", как он их называл), расположенные во внутренних районах хромосом, обладают свойством "биполярности", т.е. они могут соединяться с другими фрагментами каждым из двух своих свободных концов. Тот "ген" (или фрагмент хромосомы), который в нормальной хромосоме расположен на самом ее конце, по мнению Меллера, "униполярный", т.е. способен к объединению только с той стороны, которая обращена к внутренней части хромосомы. Меллер предположил, что этот концевой "униполярный" ген

необходим для особой функции, "запечатывания конца хромосомы". Такой "ген" был назван специальным термином - "теломерой" (по-гречески "телос" - означает конец, "мерос" — часть). Барбара Макклиток в 1941 году на основании полученных ею результатов по индукции хромосомных перестроек у кукурузы пришла к тем же выводам, что и Меллер. Более того, она постулировала, что пассивность теломер в объединении с другими участками хромосом играет исключительно важную роль, поскольку предотвращает склеивание хромосом конец - в конец, а это в свою очередь, способствует сохранению их индивидуальности и всего кариотипа в целом.

4. После открытия структуры ДНК и механизмов ее репликации А.М. Оловникова в России, заинтересовала загадка организации репликации на конце молекулы ДНК. Дело в том, что каждая хромосома содержит единственную непрерывную молекулу ДНК, и концы этой молекулы соответствуют концам хромосомы. В соответствии со стандартной моделью репликации ДНК процесс удвоения отстающей цепи ДНК начинается с синтеза коротких РНК-праймеров, или затравок, с 3'-концов которых синтезируются отрезки ДНК -фрагменты Оказаки. Затем отрезок РНК удаляется, образовавшиеся бреши (гэпы) заполняются фрагментами ДНК. Последние синтезируются, используя в качестве праймеров 3'-концы фрагментов Оказаки. Поскольку для крайнего гэпа нет праймера, вновь синтезированная цепь оказывается на 8-12 нуклеотидов (длина РНК-праймера) короче исходной. В результате, после каждого цикла репликации молекула ДНК должна становиться все короче: одна из четырех цепей, образовавшихся в результате раунда репликации стала короче на 8-12 п.н., т.е. на одном из концов хромосомы средняя длина четырех цепей ДНК уменьшилась на 2-3 п.н. Поэтому А.М. Оловников пришел к выводу, что если в клетке нет особых механизмов, компенсирующих эти потери нуклеотидов с каждого конца нити ДНК, то хромосома начнет сокращаться: сначала должны исчезнуть теломерные районы, затем ближайшие к теломерам гены, потом более удаленные гены и т.д. Очевидно, что этот процесс должен, в конце концов, привести к гибели клетки. Действительно, у клеток, растущих в

культуре (*in vitro*), есть лимит на число делений. Американским ученым Л. Хейфликом в 1965 году было показано, что если для культивирования взять клетки у новорожденных, они могут пройти 80-90 делений; клетки, взятые у 70-летних, делятся только 20-30 раз. Ограничение на число клеточных делений называют барьером Хейфлика. Обычно клетки не преодолевают барьер из 2-90 делений, по мнению Хейфлика - 50 ± 10 .

А.М. Оловников в 1971 году предложил следующую формулу для расчета продолжительности жизни любого клона клеток *in vitro*:

$$T = k(lt/lm - M),$$

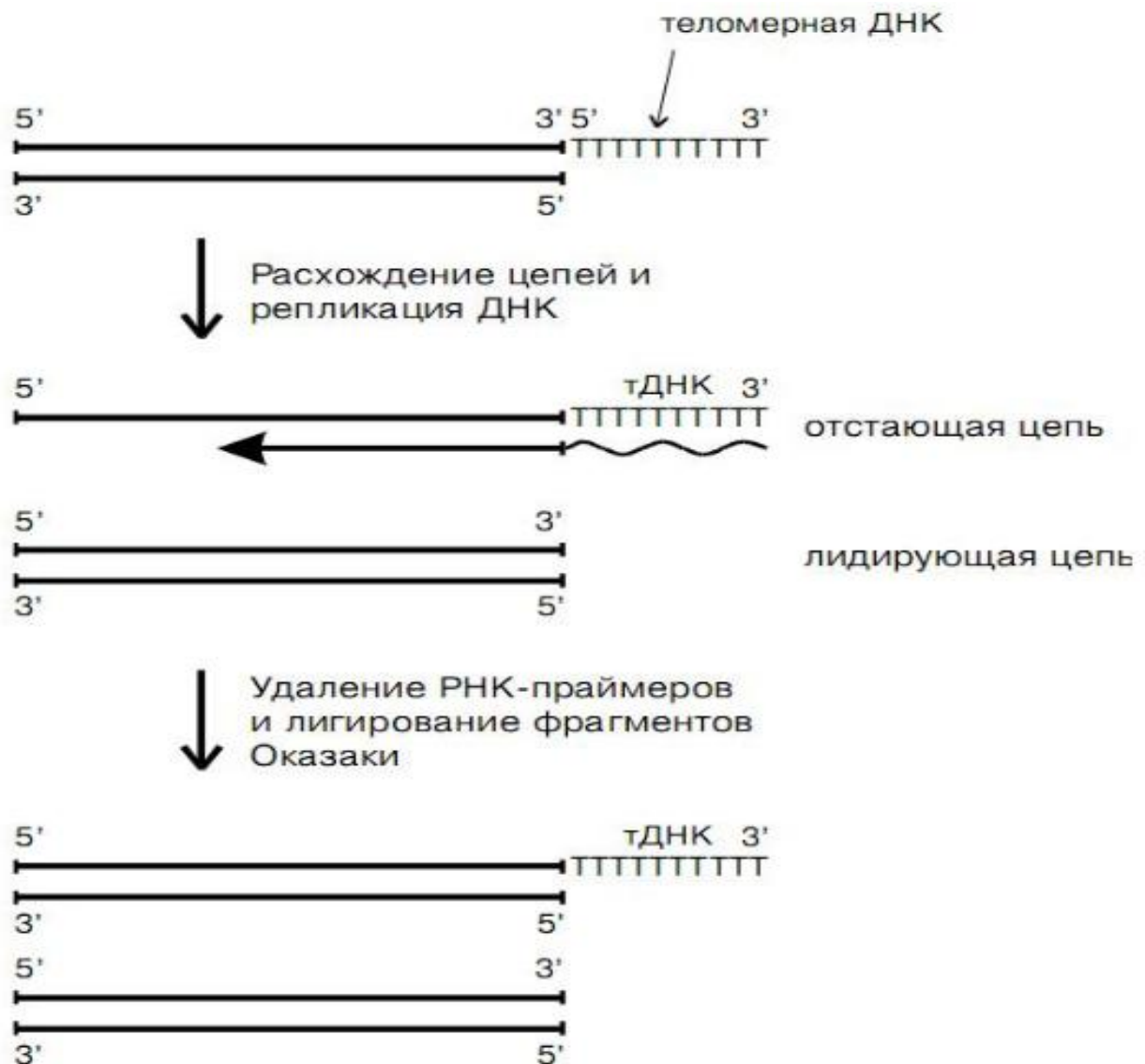


Рисунок 30: Схема репликации концевого участка отстающей цепи ДНК при наличии теломерного повтора на 3'-конце ДНК

где T - срок предстоящей жизни клеток, K — коэффициент корреляции

между сроком жизни клона клеток и числом репликаций ДНК, lt — длина теломерного участка, lm — длина фрагмента ДНК, утрачиваемого в ходе каждого цикла репликации M — число уже прошедших репликаций. Таким образом, Оловников напрямую связал срок жизни клеточных клонов с длиной теломерной ДНК. О том, что укорочение концевых фрагментов ДНК реально, свидетельствуют результаты следующего опыта. При мобилизации перемещений транспозона из теломерного района X-хромосомы дрозофилы в последнем обнаружены терминальные делеции, удаляющие наиболее дистальные последовательности ДНК. В последующих поколениях в линиях с такими делециями монотонно уменьшается длина концевых фрагментов ДНК (от теломеры к центромере) со скоростью 50-100 (в среднем 75) пар нуклеотидов за одно поколение. Полагают, что размер утрачиваемого фрагмента коррелирует с размером октонуклеотида РНК-праймера, инициирующего синтез фрагментов Оказаки при репликации ДНК. Если учесть, что клетки полового пути у этого вида в течение одного поколения делятся примерно 35 раз, совпадение расчетной длины фрагментов, утерянных за одно поколение, и экспериментальных данных оказалось поразительным. В нормальных же хромосомах, в которых теломеры присутствуют, укорочение длины молекулы ДНК не происходит.

Строение теломер

Молекулярно-генетическое изучение теломер осложняется тем, что

Таблица 1

Последовательности теломерных повторов у разных организмов

Теломерные повторы в хромосомах некоторых видов		
Вид		Последовательность нуклеотидов (5'-3')
Простейшие	<i>Euplotes</i>	TTTTGGGG
Слизневые грибы	<i>Phusarum</i>	TTTAGGG
Жгутиковые	<i>Trypanosoma</i>	TTAGGG
Споровики	<i>Plasmodium</i>	TT(T/C)AGGG
Грибы	<i>Neurospora</i>	TTAGGG
	<i>Candida maltosa</i>	ACGGATGCAGACTCGCTTGGTGT
Нематоды	<i>Ascaris</i>	TTAGGC
Насекомые	<i>Bombyx mori</i>	TTAGG
Водоросли	<i>Chlamidomonas</i>	TTTTAGGG
Высшие растения	<i>Arabidopsis</i>	TTTAGGG
Позвоночные животные	<i>Homo sapiens</i>	TTAGGG

концентрация теломерной ДНК в ядрах эукариот очень низка (8 теломер в $1,7 \times 10^8$ п.н. генома дрозофилы или 46 теломер в 3×10^9 п.н. генома человека). Подход к решению проблемы выделения теломерной ДНК был найден неожиданно. Цитологи, изучавшие развитие простейших одноклеточных животных из класса Ciliata, таких как например всем известная инфузория туфелька, обнаружили у каждого организма по два ядра - микронуклеус и макронуклеус. Микронуклеус является покоящимся ядром, в котором хромосомы находятся в компактном состоянии, и служащим для передачи наследственного материала от одного поколения другому. Другое дело - макронуклеус. Это трофическое ядро. Хромосомы там находятся в активном состоянии - на них синтезируется РНК, кодирующая белки, необходимые для роста и жизнедеятельности этого одноклеточного организма. На начальных этапах развития инфузории хромосомный материал макронуклеуса испытывает серию сложных превращений. Сначала хромосомы политенизируются, т.е. каждая хромосома реплицируется примерно десять раз и все вновь

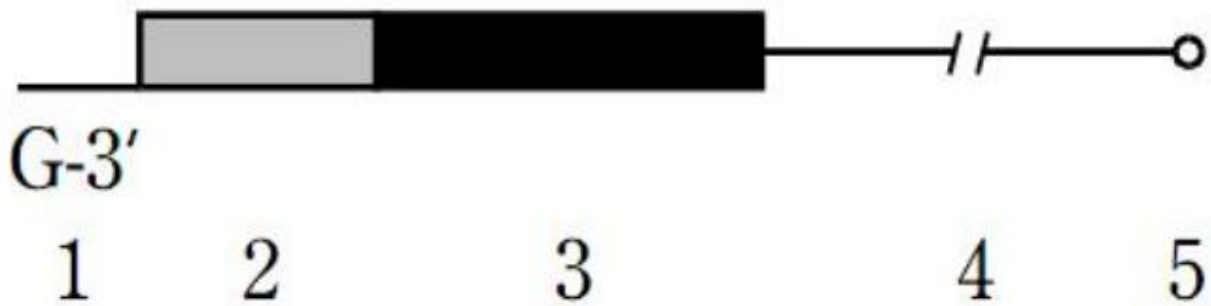


Рисунок 31: Схема организации теломерного конца хромосомы. 1 - повтор из гуаниновых нуклеотидов; 2 - теломерный концевой повтор (TR); 3 - последовательности, связанные с теломерой (TAS); 4 - собственно хромосома; 5 - центромера

образованные хромосомы остаются тесно связанными друг с другом, образуя пучек хромосом или, как в таких случаях говорят, политенную хромосому с характерным рисунком темных хромомеров и светлых межхромомерных участков. Затем политенная хромосома как бы разрезается поперек на тысячи долек, в каждой из которых находится один или несколько хромомеров. Каждая долька обтянута особой белковой оболочкой, формируя пузырек. Таким образом, на этой стадии развития в созревающем макронуклеусе находятся тысячи пузырьков, в каждом из которых находится один или несколько хромомеров. В пузырьках происходит "созревание" наследственного материала: из фрагмента хромосомы удаляется и переваривается вся ДНК, не имеющая прямого отношения к кодированию наследственной информации, например, межгенные фрагменты, внутренние участки генов, не кодирующие белки (интроны), а так же мобильные элементы генома. Затем пузырьки разрываются, и их содержимое выходит в полость макронуклеуса, который превращается по образному выражению Д.М. Прескотта исследователя этих процессов - в "мешок с генами". И наконец, самое важное: к каждому из этих генов присоединяется по одной теломерной последовательности с каждого конца. В итоге всех этих процессов в макронуклеусе в общей сложности присутствует несколько миллионов теломер. Наличие "мешка с генами", в котором находятся только индивидуальные гены, предоставляет уникальные возможности для их выделения, последующего клонирования (т.е. размножения многих миллионов

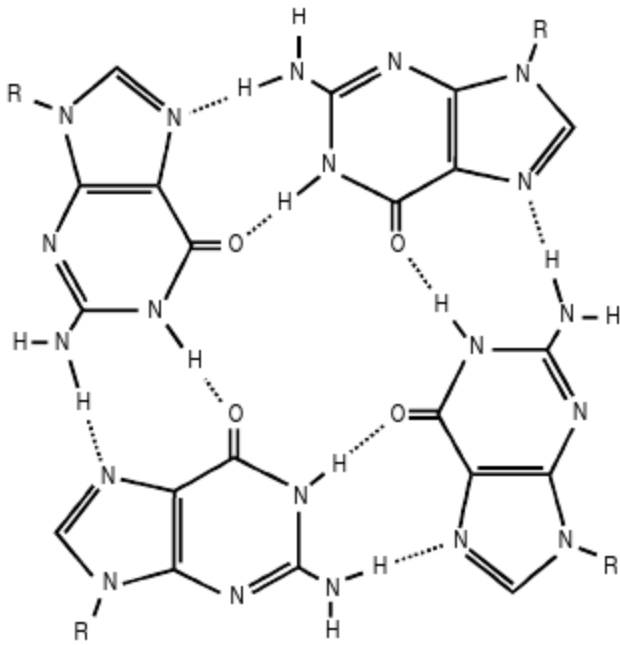


Рисунок 32: Плоскостное расположение молекул гуанина в теломерном повторе. Каждая соединена с двумя соседними.

идентичных копий гена в клетках живых бактерий или дрожжей) и последующего биохимического анализа . В 1978 г. Е. Блэкберн и Дж. Голл, выделив из макронуклеуса инфузории тетрахимены ("мешка с генами") фрагмент ДНК длиной 22 т.п.н., содержащий два сцепленных друг с другом гена рибосомной РНК и применив только еще набирающую популярность методику определения последовательностей нуклеотидов в молекулах ДНК, установили, что на

обоих концах этой пары генов находится относительно простая последовательность из цитидина (С) и аденозина (А), а во второй цепи ДНК, соответственно гуанидина (G) и тимидина (Т): т.е. на концах генов была последовательность CCCCAA/GGGGTT, повторенная несколько раз подряд. У двух других видов, *Stylonychia* и *Oxytricha* как было показано несколько позже, они похожи: CCCCAAAA/GGGGTTTT. В последние годы с помощью различных биохимических методик теломеры были выделены из ДНК многих организмов. Оказалось, что у большинства видов животных и растений теломерные районы имеют в целом очень похожий тип строения. Как правило, в состав теломерного района входят два типа фрагментов: собственно конечная часть хромосомы (или теломерный концевой повтор - TR) и "последовательность, связанная с теломерой" (TAS), которые располагаются вглубь хромосомы. Как уже было отмечено выше, у инфузории тетрахимены самая концевая часть ДНК содержит многократно повторенную последовательность гексануклеотида CCCCAA/TTGGGG, при этом последовательность ориентирована так, что TTGGGG-нить находится на 3' конце ДНК, в то время как CCCCAA-нить формирует 5' конец. Для удобства

первую из вышеупомянутых нитей называют G-нитью, вторую - C-нитью. К настоящему времени известно, что теломерные фрагменты у большинства живых существ очень похожи друг на друга, они обогащены нуклеотидами G и C, которые располагаются в определенной, похожей у всех организмов последовательности. Несколько отличается структура TR у некоторых видов грибов и особенно сильно - у дрозофилы. Кроме инфузории тетрахимены, теломерные повторы изучены у двух других представителей простейших, стилонихии и окситрихи. У них структура теломерного повтора несколько отличается, в частности единицей повторенности является октамер CCCC AAAA/GGGG TTTT. Этот октануклеотид расположен на концах хромосом особым образом, так что часть G-цепи ДНК остается одиночной из-за отсутствия второй цепи, в результате чего теломерные концы хромосомы выглядят следующим образом. Видно, что на нити ДНК, 3'-конец которой является концом хромосомы, основной повторяющейся единицей является октонуклеотид T4G4, если считать от 5'- к 3'-концу. Этот октонуклеотид, выступающий за обычную двойную Уотсон-Криковскую двойную спираль, называют "однонитчатым свободным G-концом". Длина простого повторенного дуплекса, расположенного проксимальнее, может варьировать от <50 п.н. (у *Euplotes*) до > 100 т.п.н. (у мыши). Интересно, что теломерные однонитчатые свободные G-концы фактически одинаковы у представителей совершенно различных филогенетических таксонов. Анализ показал что четыре молекулы гуанина, не имеющие второй цепи ДНК и таким образом - возможности контактировать с молекулами цитозина, как это бывает в обычной молекуле ДНК, все же спариваются. Они располагаются в одной плоскости и образуют водородные связи между собой. Особенность формирования этой структуры заключается в том, что соединение нуклеотидов происходит не по принципу комплементарности оснований Уотсона и Крика, а между молекулами одного и того же основания гуанина. С этой структурой связываются особые белки, образуя ДНК-белковый комплекс - теломеру. Она действительно как бы запечатывает хромосому с каждого конца. Как уже отмечалось выше, рядом с

теломерным повтором располагаются участки ДНК большой протяженности, называемые субтеломерными повторами-TAS. У дрожжей описано два субтеломерных повтора, X и Y'. X - имеет размеры 0,3-3,75 т.п.н. и менее консервативен, чем Y' (5,2-6,7 т.п.н.). Они чередуются или могут состоять только из Y'-элементов. Последние имеют открытые рамки считывания. На основе слабой гомологии в первичных последовательностях с мобильными элементами, предположили, что Y'-элементы возникли из одиночной инсерции в предковый геном дрожжей. Субтеломерные повторы обнаружены у многих видов. Общая длина тракта может быть очень большой - у хирономуса до 300 т.п.н. в каждой теломере. По иронии судьбы, термин "теломера" возник в результате исследований на дрозофиле. В то же время, до сих пор нет никаких доказательств того, что в геноме дрозофилы есть короткие теломерные G-богатые последовательности, характерные для теломер большинства других видов. У дрозофилы на концах хромосом расположены теломеро-специфичные ретротранспозоны, HeT-A и TART. Обнаруженные два типа ретротранспозонов, He-T и TART, имеют в общих чертах похожее строение. У обоих на 5'-конце расположены небольшие повторенные последовательности, а на 3'-конце - более длинный повтор, завершающийся длинным фрагментом, содержащим только нуклеотиды с аденином: An. Различаются они тем, что He-T имеет две частично наложенные одна на другую последовательности (ORF), кодирующие белки, в то время как TART - две, ORF1 и ORF2, расположенные тандемно. Оба транспозона присутствуют во многих копиях, главным образом в прицентромерных районах. Они могут перемещаться по геному и в случае утраты теломерного района, встраиваются на самый конец хромосомы, восстанавливая теломерную структуру. Несмотря на то, что молекулярная структура теломеры была к началу 1990-х годов, в основном расшифрована, проблема неполной репликации на конце линейной молекулы ДНК осталась. В последние несколько лет выяснилось, что природа выработала механизм удлинения (или элонгации) самого конца хромосомы - т.е. теломерного концевого повтора. Это связано с активностью особого фермента - теломеразы

рибонуклеопропротеида, т.е. белка, содержащего в своем составе короткую молекулу РНК (примерно 150 нуклеотидов, среди которых 2 копии теломерного повтора). Теломераза выполняет две функции:

1. Перед началом цикла репликации ДНК теломераза добавляет несколько копий теломерных повторов на 3' конец ДНК. После этого репликация идет в обычном порядке. На отстающей цепи синтезируются РНК-праймеры, при этом наиболее важно то, что самый концевой праймер синтезируется на теломерном повторе. После заполнения гэпов и завершения репликации остается незаполненным только участок РНК-праймера, синтезированного на теломерной последовательности. В результате дочерние цепи ДНК получаются той же длины, что и родительские цепи.

2. Нетрудно заметить, что часть нуклеотидов на 3' конце отстающей цепи все же не будет реплицироваться. Но эта неполная репликация произойдет в зоне теломерного повтора, что не причинит никакого вреда генам, расположенным рядом. Однако, через несколько циклов репликации теломерный повтор "растет" и недорепликация начнет затрагивать гены. Поэтому второй функцией теломеразы является постоянное наращивание G-нити. Теломераза имеет свою матричную молекулу РНК, с помощью которой фермент распознает теломерный повтор. Последовательность 5'-СААССССАА-3' в молекуле теломеразы спаривается с последовательностью теломерного повтора 5'-ТТГГГГ-3'. Нуклеотиды ААС в РНК теломеразы остаются неспаренными, и на них достраиваются нуклеотиды ТТГ. Фермент перемещается на самый конец теломерной последовательности т.е. на всю длину ТТГГГГТТГ, и нуклеотиды ААС из молекулы теломеразы спариваются с ТТГ теломеры, после чего достраивается вся последовательность повтора. Таким образом, очевидно, что конец хромосомы как бы "запечатан" особой ДНК-белковой структурой, которая позволяет нормально реплицироваться ДНК всей хромосомы.

В последнее время накапливаются данные о том, что нарушения в

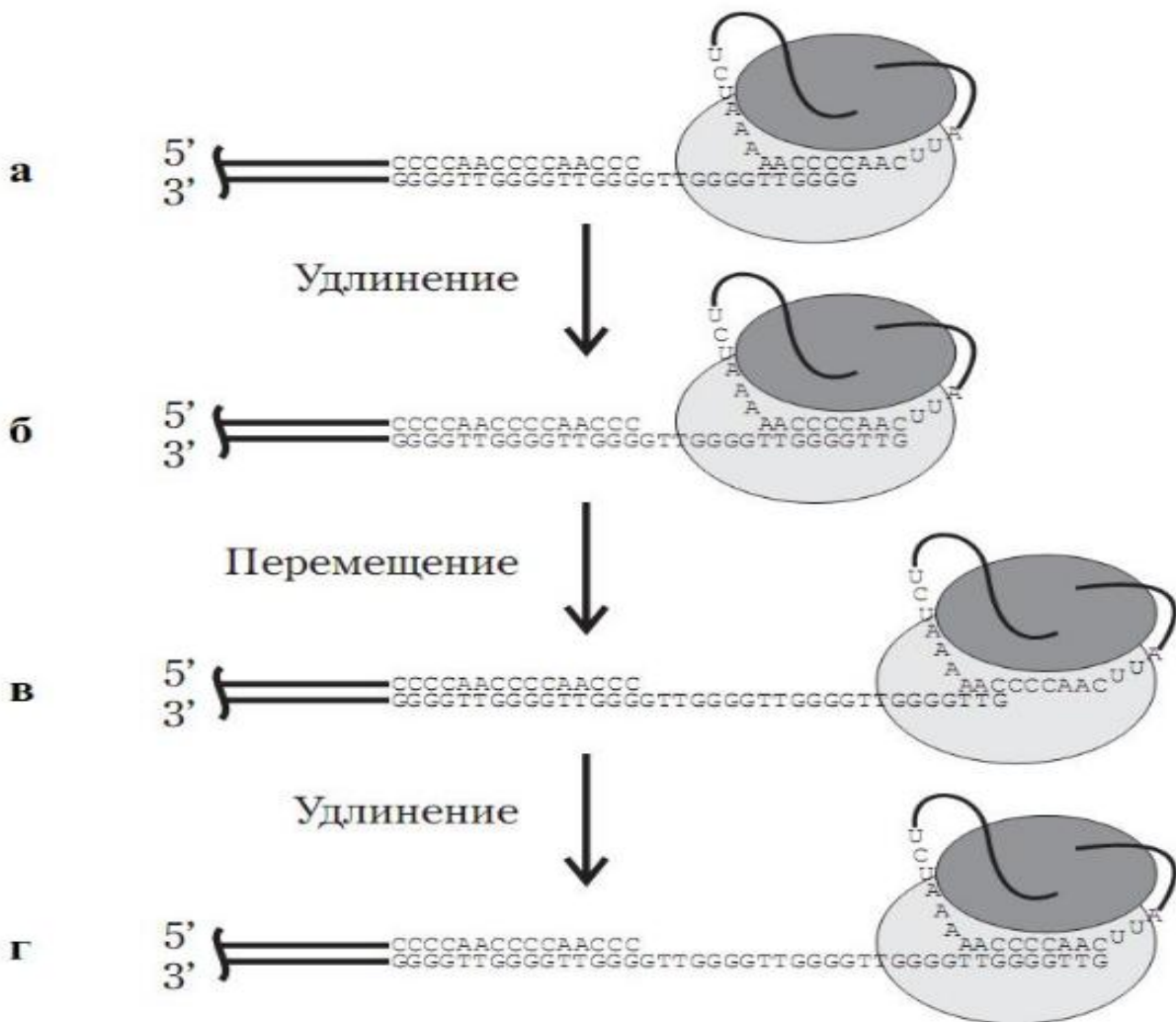


Рисунок 33: Удлинение теломерного повтора с помощью фермента теломеразы

механизме удлинения теломерного повтора непосредственно связаны с формированием злокачественных новообразований, а также играют важную роль в процессе старения. Теломеры в клетках зародышевого пути, благодаря постоянно высокой активности теломеразы сохраняют нормальную длину. Однако, в соматических клетках, культивируемых *in vitro*, теломераза неактивна, и теломеры постоянно укорачиваются. Это и объясняет существование барьера Хейфлика. В раковых клетках, которые также являются соматическими, клеточные деления не прекращаются, и теломеры у них не укорачиваются. Оказалось, что почти во всех образцах опухолевых клеток,

взятых как из культуры, так и из целого организма, активность теломеразы сохраняется на высоком уровне. Это обстоятельство позволило ученым пофантазировать на следующую тему: если бы удалось найти химический реагент, избирательно инактивирующий теломеразу, то при его применении опухолевые клетки быстро достигали бы барьера Хейфлика и погибали, в то время как в соматических клетках действие этого агента не ощущалось бы, т.к. в них теломеразы нет. И наконец, в середине января 1998 года в американском журнале Science (N5349) появилась статья, взбудоражившая общественное мнение. Ее авторы сообщили об успешном эксперименте, в ходе которого удалось преодолеть Хейфлика в культуре клеток человека. Группе американского исследователя Дж. Шей с помощью генно-инженерных методов удалось ввести в геном соматических клеток человека ген теломеразы, снабженный регулируемыми фрагментами ДНК, которые и заставили этот ген активно работать в тех клетках, в которых он обычно не работает. Авторы обнаружили, что длина теломер в этих клетках начала увеличиваться, так же как и продолжительность жизни клеточных культур: сверх обычных 50 делений клетки прошли 20 дополнительных. Результаты, приведенные выше, показывают, какую огромную роль в жизни организма играют маленькие фрагменты ДНК, расположенные на концах хромосом.

Согласно данным цитогенетиков теломеры образуют гетерохроматин, который связан с ядерной оболочкой.

Теломерные районы хромосом имеют следующие свойства, сближающие их с гетерохроматином:

1. В состав теломер входят повторенные последовательности, число копий в кластере варьирует.
2. Теломеры обычно расположены на ядерной обложке.
3. Теломерные концы хромосом чаще всего создают ассоциации, т.е. теломеры разных хромосом конъюгируют.
4. Теломеры вступают в ассоциации с многими другими районами генома, в первую очередь сприцентромерным гетерохроматином и

интеркалярным гетерохроматином.

5. В некоторых случаях (очень редко) в теломерных районах выявляется слабая С-окраска.

6. В политенных хромосомах теломерные районы представлены неполно т.е. теломерная ДНК недореплицирована.

7. Белок HP1 выявляется в теломерах.

8. Гены могут инактивироваться если перенесены в окрестности прицентромерного гетерохроматина.

9. Есть данные о действии модификаторов эффекта положения на проявления свойств теломер.

Мобильные генетические элементы

В начале 1940-х годов американская исследовательница Барбара МакКлинток открыла существование гена или локуса, который вызывал повышенные частоты хромосомных перестроек у кукурузы. Среди потомков от скрещивания, в котором оба родителя несли такие перестройки, появлялись нестабильные мутации с неожиданно высокой частотой. В 1948 году она опубликовала результаты исследований этого локуса, вызывающего разрывы хромосом, сделав вывод, что он был совершенно необычным поскольку мог перемещаться из одного участка хромосомы в другой. Б. МакКлинток назвала феномен перемещения транспозицией, а сами локусы — "контролирующими элементами" (КЭ). Эти элементы характеризуются следующими свойствами:

1. Они могут перемещаться из одного сайта в другой.

2. Их встраивание в данный район влияет на активность генов, расположенных рядом.

3. Утрата КЭ в данном локусе превращает прежде мутабельный локус в стабильный.

4. В сайтах, в которых присутствуют КЭ, могут возникать делеции, транслокации, транспозиции инверсии, а также разрывы хромосом.

Геном кукурузы содержит несколько семейств КЭ. Члены каждого семейства могут быть подразделены на два класса:

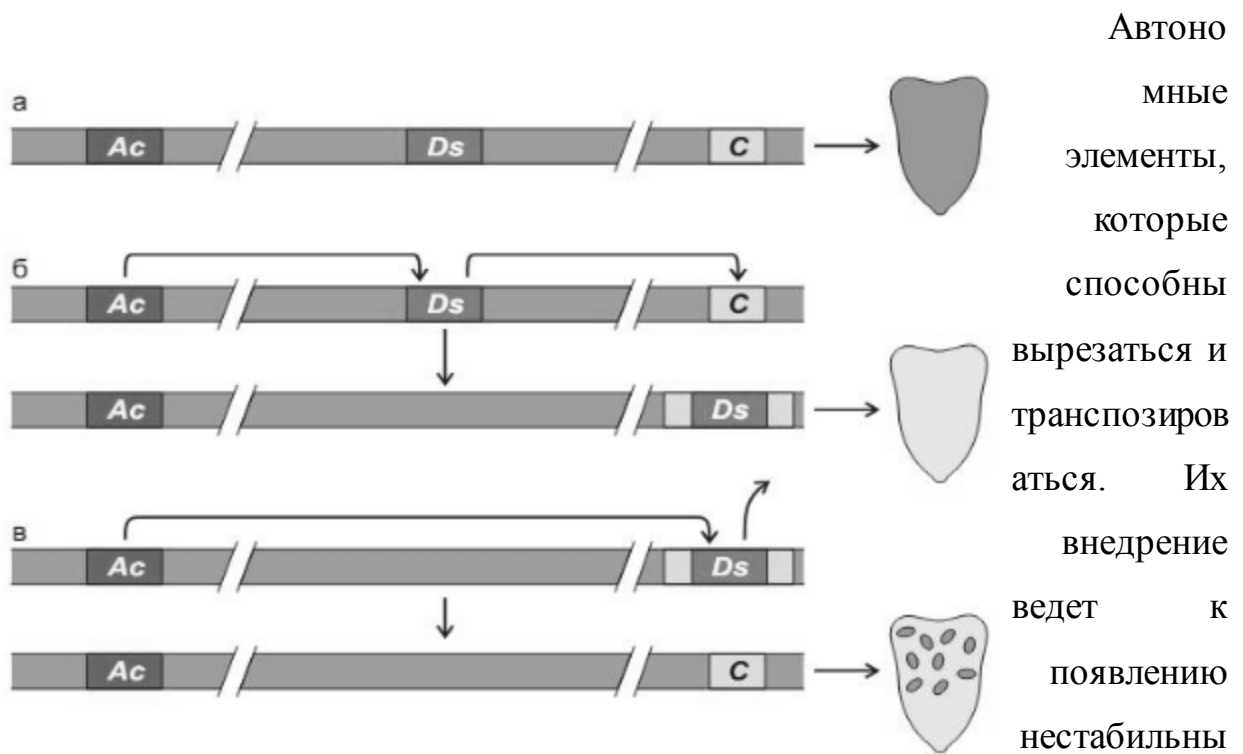


Рисунок 34: Изменение окраски кукурузного зерна под влиянием перемещений элементов Ac-Ds

Неавтономные элементы теряют свою стабильность только в том случае, если в какой-то области генома присутствует автономный член того же семейства.

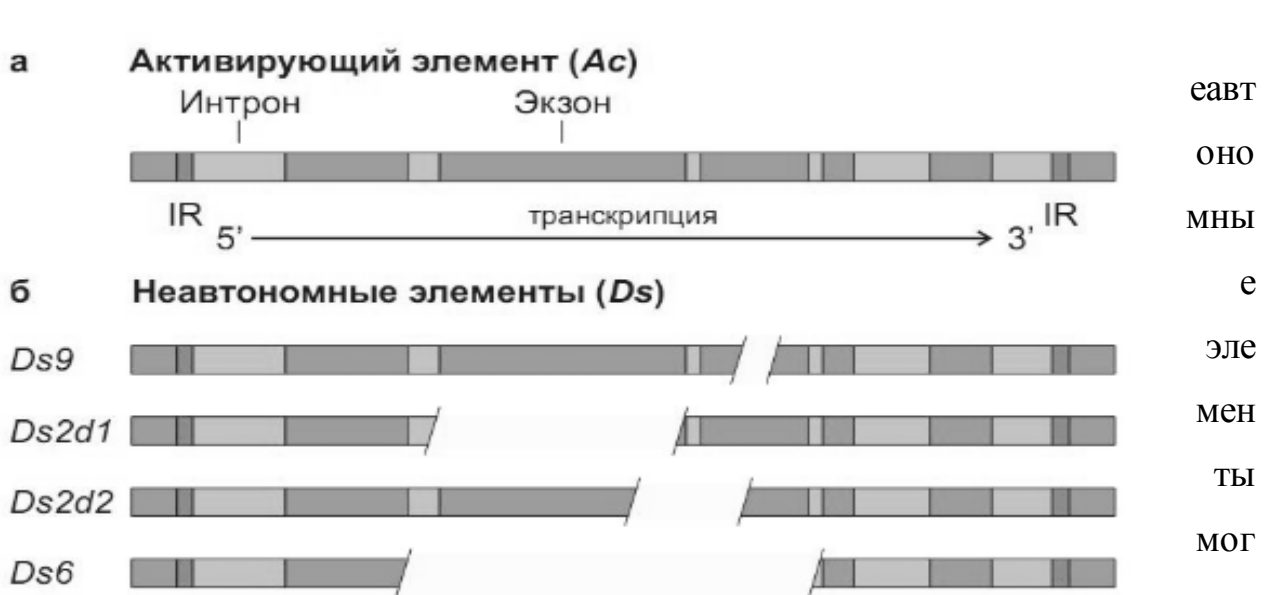


Рисунок 35: Структура автономного Ac-элемента (а) и неавтономных Ds-элементов (б) у кукурузы.

Активированы только определенными автономными элементами (членами того же семейства). У кукурузы изучены лучше всего Ac-Ds, Spm (супрессор-мутатор) и Dt семейства.

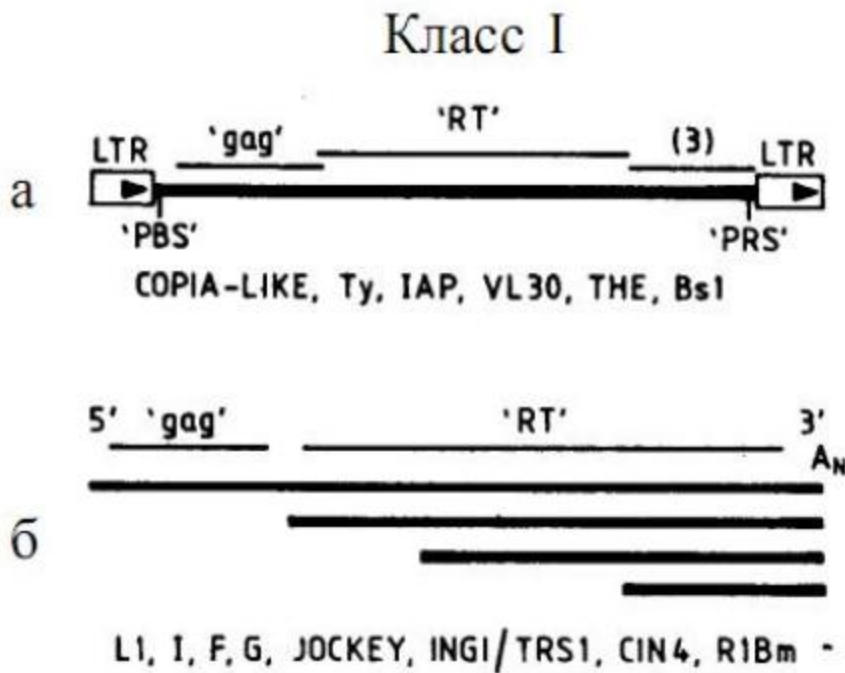
При встраивании автономного элемента в ген, последний мутирует, однако мутация эта будет нестабильной, поскольку элемент может выйти из данного гена и переместиться в другой участок генома. Поскольку частота транспозиции значительно выше, чем частота обратного спонтанного мутирования, то аллель, индуцированный автономным элементом, называется мутабельным. Система Ac-Ds у кукурузы была изучена в деталях. Ac-элемент имеет длину 4563 п.н. с инвертированными повторами (IR) на концах. Он содержит единственную единицу транскрипции (5 экзонов и 6 интронов), кодирующего фермент транспозазу. Элементы Ds происходят из Ac в результате делеций внутренних участков Ac-элемента. Если растение имеет аллель гена C дикого типа, зерно будет иметь пурпурную окраску, если Ac элемент индуцировал инсерцию Ds в ген C, возникает мутантный аллель c, и окраска зерна будет бесцветной. В ходе развития Ds элемент может в некоторых клетках выйти из гена C, в результате чего зерна вновь приобретут пурпурную окраску. Таким образом, возникает мозаичность. К настоящему времени мобильные элементы открыты у множества видов растений, животных и микроорганизмов. Ниже дается их классификация и подробное описание. По механизмам транспозиции мобильные элементы делятся на две большие группы. У млекопитающих в составе генома обнаружено несколько классов умеренно повторенных последовательностей: SINE (short interspersed nuclear sequences) и LINE (long interspersed nuclear sequences). Элементы SINE - это фрагменты длиной 100-300 п.н., чередующиеся с уникальными последовательностями от 1000 до 2000 п.н. длиной. Элементы LINE имеют длину более 5 т.п.н., они чередуются с уникальными последовательностями до 35 т.п.н. длиной. Как SINE, так и LINE представлены семействами, состоящими из одинаковых элементов. В геноме человека SINE широко представлено семейством элементов Alu. Члены этого семейства имеют длину 300 п.н. и повторены от 300000 до 500000 раз в геноме. Около 3% генома человека приходится на долю этих повторов. Наименование Alu этот элемент получил поскольку содержит сайт рестрикции рестриктазы AluI. Каждая Alu-последовательность

фланкирована прямыми повторами длиной от 7 до 20 п.н. По этой причине полагают, что Alu-повторы являются мобильными элементами, скорее всего ретротранспозонами. Одно из семейств, принадлежащих LINE, это LINE-1 (или L1-элемент). Полагают, что в геноме человека присутствует от 50 до 100 тыс. копий L1-элемента, т.е. он представляет около 5% генома. Максимальная длина этих элементов составляет 6500 п.н., хотя таких элементов в геноме не более 3500. Остальные же копии по аналогии с Ds элементами кукурузы имеют внутренние делеции различной длины. Полноразмерные L1 элементы содержат большие открытые рамки считывания, имеющие гомологию с известными обратными транскриптазами.

Наличие мобильных элементов в геномах имеет разнообразные генетические последствия.

1. Перемещения и внедрение мобильных элементов в гены может вызывать мутации. Так, большинство "спонтанных" мутаций в локусе white у дрозофилы индуцированы инсерциями мобильных генетических элементов. Около 80% спонтанных мутаций, изученных в разных локусах дрозофилы, вызвано инсерциями мобильных элементов. Внедряясь в ген, мобильный элемент может повредить экзон, разорвав его. В таком случае ген будет лишен возможности кодировать белок. Попадая в район промоторов или энхансеров, мобильный элемент может повредить регуляторную зону гена. Наконец, инсерция в район интрона может оказаться безвредной, поскольку вся последовательность интрона вместе с мобильным элементом будет вырезана во время процессинга мРНК, а соседние экзоны беспрепятственно сплайсируются.

2. Изменение состояния активности генов. LTR (длинный концевой участок) мобильного элемента являются промоторами



ретротранспозона, причем как LTR, так и сам ретротранспозон содержит нуклеотидные последовательности, являющиеся энхансерами транскрипции.

Поэтому перемещение этих сигналов в геноме может изменять регуляцию активности генов. Например, если мобильный элемент оказался около протоонкогена, то результатом может быть сверхпродукция белка и злокачественное перерождение клетки. В

Рисунок 36: Структура мобильных элементов 4х различных типов у эукариот: а, б - мобильные элементы класса I, которые, как полагают, транспозируют через посредничество РНК; в, г - мобильные элементы, транспозирующие непосредственно через ДНК

случае ретротранспозонов особые возможности для перенесения и преобразования регуляторных сигналов возникают в том случае, когда элемент удаляется за счет кроссинговера между LTR с идентичными

последовательностями, в результате сохраняется лишь один ДКП на месте внедрения ретротранспозона. Это явление широко распространено в клетках дрожжей. Установлено, что такие одинокие LTR оказывают серьезное влияние на изменчивость регуляторных систем дрожжевой клетки.

3. Формирование хромосомных перестроек. В результате кроссинговера между одинаково ориентированными элементами возникает делеция и дупликация материала, расположенного между инсерциями. Если инсерции ориентированы в противоположном направлении, возникает инверсия.

4. У дрозофилы отсутствует теломеразная машина, но концы ДНК удлиняются за счет перемещений ретротранспозонов.

5. Участие в горизонтальном переносе генов. Инфекционные ретровирусы способны заражать организмы, принадлежащие к разным видам и переносить собственный генетический материал, образуя копии ДНК, встраивающиеся в геном. Таким образом могли распространяться ретротранспозоны. Подобный способ передачи генов получил название горизонтального в отличие от вертикального наследования генов из поколения в поколение. Широкое распространение транспозона *mariner* среди филогенетически отдаленных групп организмов может свидетельствовать о повторных переходах данного элемента от вида к виду. Так, один из ретротранспозонов дрозофилы (*gypsy*), как оказалось, является настоящим ретровирусом: путем инъекции или скармливанием вирусных частиц удается заразить мух, не несущих этих ретротранспозонов.

6. Транспозоны на основе Р-элемента используют для трансформации у эукариот, клонирования генов, поиска энхансеров.