

На правах рукописи

ШИБАЛЁВ ДМИТРИЙ ВАЛЕРЬЕВИЧ

**КОНСЕРВАТИВНОСТЬ И ВАРИАБЕЛЬНОСТЬ ДНК
ЯДРЫШКОВЫХ ОРГАНИЗАТОРОВ ЧЕЛОВЕКА**

Специальность 03.01.03 –молекулярная биология

АВТОРЕФЕРАТ
диссертации на соискание учёной степени
кандидата биологических наук

Москва – 2012 г.

Работа выполнена в Федеральном государственном бюджетном учреждении науки Институте биологии гена Российской академии наук, лаборатории организации генома.

Научный руководитель:

кандидат биологических наук
Куприянова Наталья Сергеевна

Официальные оппоненты:

Носиков Валерий Вячеславович
доктор биологических наук, профессор,
заведующий лабораторией, Государственный
научный центр Российской Федерации
ФГУП Государственный научно-исследовательский
институт генетики и селекции
промышленных микроорганизмов

Машкова Тамара Дмитриевна
кандидат химических наук, старший научный
сотрудник ИМБ РАН

Ведущая организация:

Федеральное государственное бюджетное
учреждение науки Институт биохимии
и генетики Уфимского научного центра
Российской академии наук

Защита состоится « 16 » апреля 2012 г. в 12 часов на заседании диссертационного совета Д 002.037.01 при ИБГ РАН по адресу: 119334, Москва, ул. Вавилова 34/5.

С диссертацией можно ознакомиться в библиотеке Федерального государственного бюджетного учреждения науки Институте молекулярной биологии им. В. А. Энгельгардта Российской академии наук по адресу: 119991, Москва, В-334, ул. Вавилова 32.

Автореферат разослан « 14 » марта 2012 года.

Учёный секретарь Диссертационного совета
канд. фарм. наук

Грабовская Л.С.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Рибосома является одной из самых древних и важных органелл клетки, сохранившей общие черты организации у всех ныне живущих организмов. Гены, ответственные за синтез нуклеиновых кислот и белков, формирующих рибосому, а также гены, обслуживающие процесс работы этих генов, созревание продуктов транскрипции и переход зрелых продуктов в активное состояние, образуют крупнейший полигенный комплекс, от согласованной работы которого зависит жизнеспособность отдельных клеток и всего организма в целом. Центральную роль в этом комплексе занимают гены рибосомной ДНК (рДНК), организованные в кластеры тандемных повторов. Мономер состоит из консервативной транскрибируемой области и нетранскрибируемого межгенного спейсера (рМГС), строение которого имеет видоспецифический характер.

Тот факт, что определенные количества рибосомной РНК (рРНК) необходимы практически во всех клетках организма, делает актуальным изучение механизмов транскрипции рибосомных генов и выявление особенностей организации тех участков рДНК, которые отвечают за регуляцию этого процесса. О получении черновой версии полноразмерной нуклеотидной последовательности генома человека было заявлено в 2000, а завершенной – в 2003 году. Однако неисследованными остались гетерохроматиновые области (около 8% всего генома), к которым относятся центромеры, теломеры и кластеры рибосомных ДНК человека. Это объясняется сложной организацией данных областей и высокой концентрацией повторов в них. По всей вероятности, центромеры, теломеры и кластеры рибосомных ДНК человека будут просеквенированы не раньше разработки новых технологий для изучения первичной структуры ДНК.

Синтез рРНК является этапом, лимитирующим скорость биогенеза рибосом. Скорость работы Pol I подвергается интенсивной регуляции на нескольких уровнях. Известно, что эффективность экспрессии генов рРНК у эукариот непосредственно определяется взаимодействием ряда цис- и транс- факторов в предпромоторной области. Скорость пролиферации раковых клеток в опухолях прямо пропорциональна размеру ядрышка и активности Pol I. Сверхэкспрессия пре-рРНК коррелирует с неблагоприятным прогнозом при раке [White, 2008; Montarano et al., 2008].

За последние годы обнаружен ряд эпигенетических факторов, участвующих в регуляции активности рибосомных генов. Обнаружен комплекс, ремоделирующий структуру хроматина в области промотора рРНК (NORC), который не только участвует в формировании гетерохроматина и подавлении транскрипционной активности, но и связывает эти процессы с модификацией гистонов и метилированием de novo [Santoro and Grummt, 2005; Bernstein and Allis, 2005]. Большую роль в процессе регуляции играют

также некодирующие РНК, считывающиеся с рибосомного межгенного спейсера (рМГС) [Mayer et al., 2008; Shiao et al., 2009].

Хотя регуляция транскрипции рРНК представляет собой весьма сложную систему, последние достижения позволяют подойти к пониманию множественных связей между биогенезом рРНК и развитием серьезных заболеваний. Отсюда следует, что изучение организации и полиморфизма рДНК является важной задачей, так как даже небольшие изменения в первичной структуре могут подавлять гены супрессоры, а также активировать онкогены и протеинкиназы, стимулирующие Pol I, и ускоряющие клеточную пролиферацию. Все выше изложенное показывает актуальность изучения структурно-функциональной организации и полиморфизма генов рРНК, играющих важную роль в регуляции их экспрессии.

Цели и задачи исследования. Целью данной работы является поиск и изучение вариаций в первичной структуре рМГС человека и прилежащих к кластеру рДНК областей генома. Для решения поставленной цели были сформулированы следующие задачи:

С помощью ПЦР-амплификации провести анализ участков предпромоторной рДНК, содержащих повторы различных типов.

Определить уровень и изучить молекулярную природу вариабельности участка LR2 рибосомного межгенного спейсера.

Из космидной клонотеки хромосомы 13 человека выделить клоны, содержащие рДНК и определить первичную структуру участков, предшествующих кластеру рДНК.

Определить вероятные «горячие точки рекомбинации» фрагментов области LR1-LR2, с наибольшей вероятностью реализующиеся при рекомбинации в рДНК человека.

Научная новизна и практическая ценность работы.

При изучении полиморфизма предпромоторной области рМГС выявлен феномен образования укороченных рекомбинантных (химерных) молекул ДНК, как следствие преждевременной терминации полимеразной цепной реакции в сложно организованных локусах рМГС человека, содержащих однонаправленные *Alu* повторы, а также тандемные повторы между ними. Установлены сайты терминации Taq полимеразы на *Alu* повторах в рМГС человека.

Впервые показан высокий уровень полиморфизма участка LR2_{var} в рМГС человека; определена природа изменчивости этих участков, которая связана с дупликациями, делециями и конверсионными переносами блоков повторов между аллелями.

Показано, что область, предшествующая кластеру рДНК в хромосоме 13 (не менее 7 т.п.н.), гомологична на 84% ранее определенному ортологичному участку из хромосомы 21. Обнаружение высокой консервативности последовательностей, прилежащих к кластеру рДНК со стороны теломеры, в различных ЯОР⁺ хромосомах указывает на важное значение этих последовательностей для согласованной эволюции генов рРНК и их принадлежность к ядрышковому организатору в совокупности с кластерами рДНК.

Обнаружен новый случай сегментных дупликаций в геноме человека. Показана высокая гомология прителомерного участка (5 т.п.н.) хромосомы 13 с прицентромерным участком хромосомы 19.

Картированы *in silico* сегменты, высоко гомологичные LR1-LR2 участкам рДНК, на ЯОР⁻ хромосомах, путем их сравнения с полноразмерной последовательностью ДНК генома человека; данные предполагают существование горячих точек рекомбинации в повторах рДНК.

Результаты работы имеют ценность для понимания структурно-функциональной организации генов рРНК человека, природы внутри геномного полиморфизма рДНК и регуляции работы мультигенного семейства рДНК. Полученные рекомбинантные клоны могут быть использованы для решения научных и практических задач по геномике человека, в том числе, при определении роли рДНК в различных патологиях человека.

Апробация работы. Основные результаты диссертационной работы были представлены на конференции «Фундаментальные науки – медицине» (Москва 3-4 декабря 2007года),

Публикации. По материалам диссертации опубликовано 5 печатных работ.

Структура и объём работы. Диссертация изложена на 96 страницах, включает 18 рисунков, 2 таблицы и состоит из введения, обзора литературы, материалов и методов, результатов, обсуждения результатов, выводов, списка литературы, включающего 224 источника, благодарности, а также приложения.

Результаты и их обсуждение

1. Феномен образования рекомбинантных (химерных) молекул ДНК, как следствие преждевременной терминации полимеразной цепной реакции в локусах содержащих однонаправленные *Alu* элементы и микросателлитные последовательности.

Рибосомная ДНК содержит практически все типы последовательностей генома, как эволюционно консервативные, кодирующие, так и переменные, некодирующие спейсерные области с высоким содержанием повторов различного типа. При рестриктно-гибризационном анализе повторов рДНК, клонированных в составе космидной клонотеки хромосомы 13 человека, стандартные характеристики получены лишь для части мономеров, значительная же часть клонов рДНК содержала большое число нуклеотидных замен, дополнительные микросателлитные кластеры и делеции различной протяженности [Рысков и др., 1993; Куприянова и др., 1996].

Полученные результаты указывали на то, что геномная рДНК также может обладать значительно большей вариабельностью, чем это представлялось ранее. Структурная организация рДНК человека представлена на (рис.1).

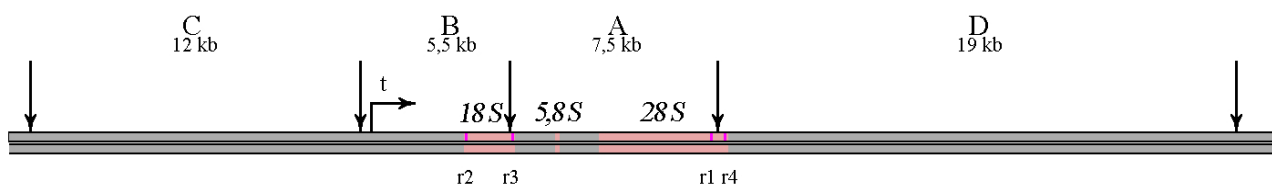


Рис. 1 Схема строения мономера рДНК. Точками обозначены высоко-консервативные области r1-r4.

Известно, что наиболее переменные участки рДНК находятся в районах рМГС, связанных с инициацией и терминацией транскрипции. Описанный полиморфизм обычно проявляется в вариациях числа повторяющихся звеньев в микросателлитных кластерах и в точковых заменах оснований, которые могут приводить к инактивации рестриктных сайтов, ведущей к полиморфизму длин рестриктных фрагментов (ПДФ).

Для поиска внутригеномного полиморфизма в области, предшествующей точке инициации транскрипции, сначала нами была избрана область между *EcoRI* и *BamHI* сайтами, содержащая четыре *Alu* элемента в различных ориентациях (рис.2). Проводили ПЦР – амплификацию участка, размером около 1,8 т.п.н., ограниченного праймерами P3 (5'-aatgaaatgaaatgaacgca-3') и P4 (5'-aaagcgaatgagtcaacag-3'), содержащего однонаправленные повторы *Alu3* и *Alu4*. Эти повторы разделены блоком 90 п.н. повторов, которые образованы микросателлитными кластерами (TCCC)_n и (TTGC)_n (рис. 3).

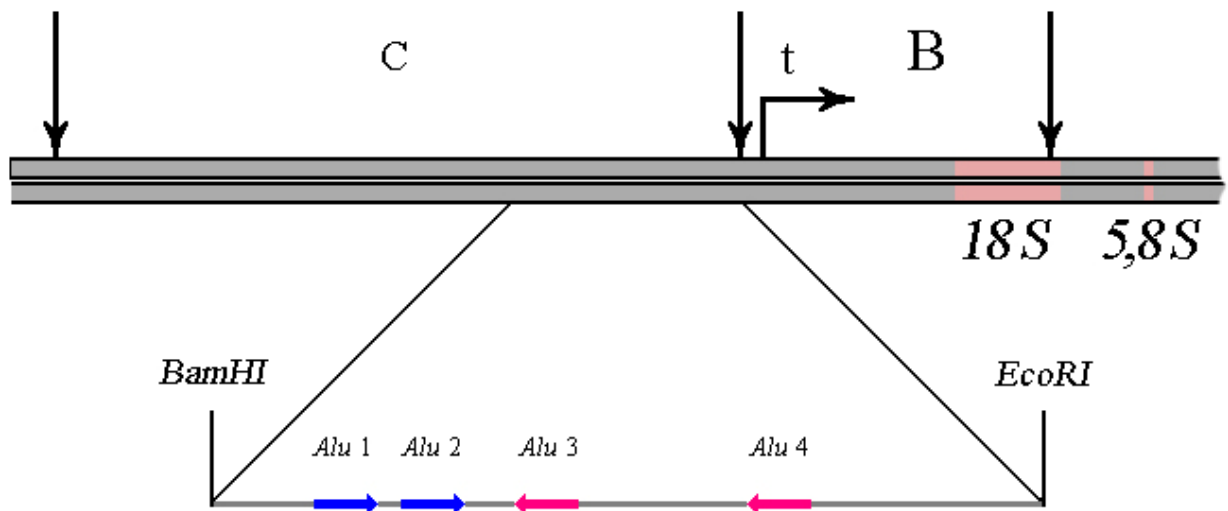


Рис.2. Развернутое схематичное изображение изучаемой области рМГС. Вынесен *BamHI*-*EcoRI* фрагмент размером $\sim 4,5$ т.п.н. Стрелками обозначено расположение и направление *Alu*-повторов.

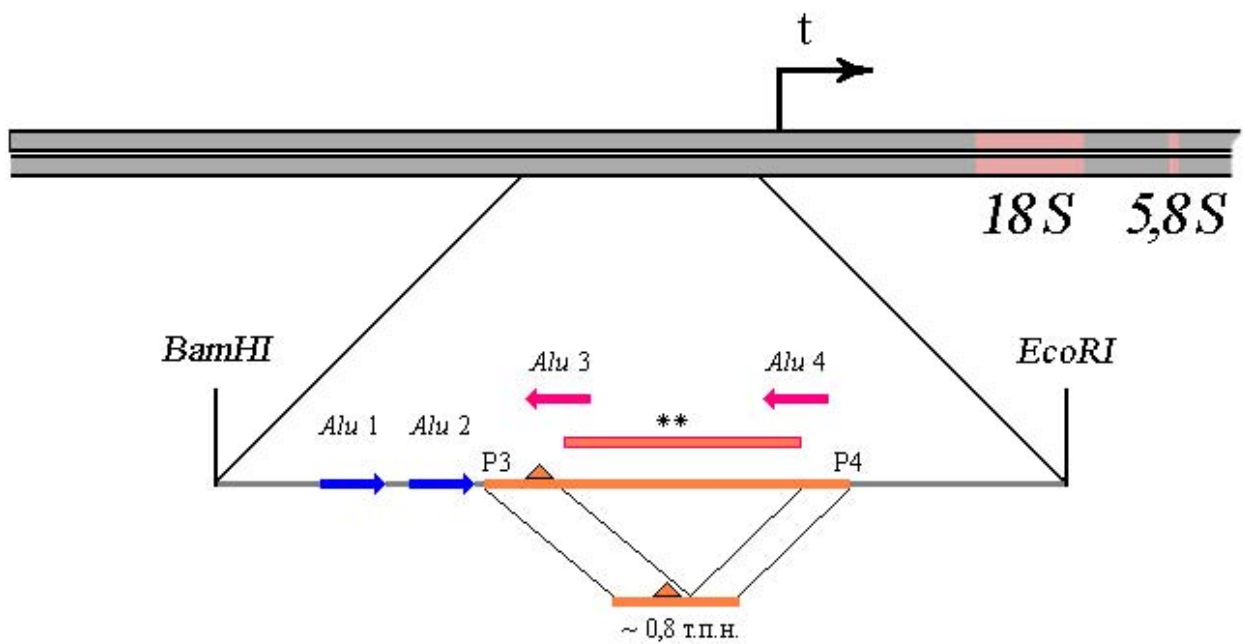


Рис. 3 Схема строения и ПЦР-анализа области рДНК, содержащей *Alu3* и *Alu4*. Положение олигонуклеотидной пробы ОР обозначено треугольником, микросателлитных мотивов звездочкой.

ПЦР-амплификация данной области, проведенная на тотальной ДНК из крови нескольких неродственных индивидов, выявила помимо ожидаемого фрагмента ($\sim 1,8$ т.п.н.) также дополнительный укороченный фрагмент, размером ~ 800 п.н. (рис.4а).

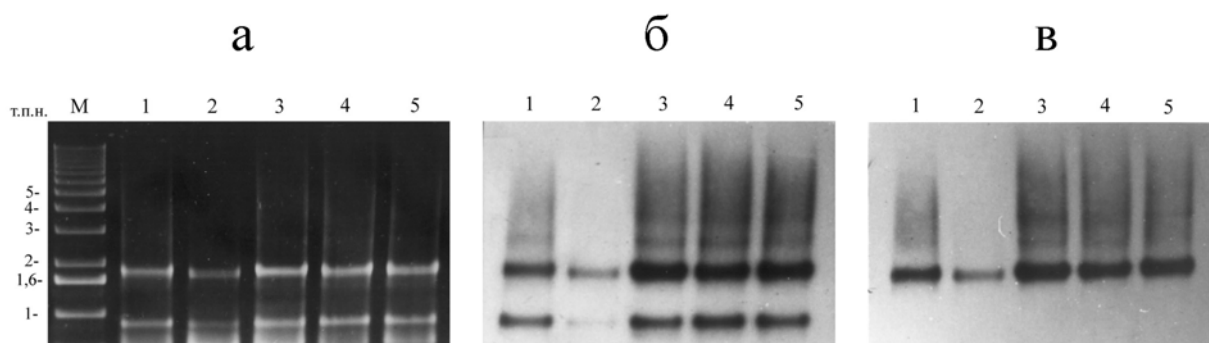


Рис 4. Саузерн-блот анализ ПЦР-продуктов, а-негативное изображение 1,5 % агарозного геля в УФ-свете, б-гибридизация с пробой 0P и в – с пробой (ttgc)₄.

Для выяснения природы укороченного фрагмента был проведен блот-гибридизационный анализ полученного ПЦР продукта. Оказалось, что оба фрагмента гибридизуются с олигонуклеотидной пробой 0P (5'-ctaggaaatcgccacttga-3'), соответствующей уникальному участку, вблизи праймера P1 (рис.4 б).

Однако только фрагмент размером 1,8 т.п.н. гибридизуется со специфической пробой, содержащей микросателлитный мотив (ttgc)₄, кластер которого разделяет *Alu3* и *Alu4* (рис. 4в). Эти результаты позволяют предположить, что фрагмент, размером 0,8 т.п.н. представляет собой делетированный вариант фрагмента размером 1,8 т.п.н., утратившего участок между двумя *Alu* элементами. Вариация условий ПЦР-амплификации (вариация концентрации Mg²⁺ от 1,5 до 4 мМ и использование разных препаратов Taq-полимеразы) приводила к изменению относительного содержания этих продуктов, но не устраняло полностью укороченный вариант (данные не приведены). Блот-гибридизационный анализ геномной ДНК, гидролизованной *EcoRI*, с олигонуклеотидными пробами, использованными для идентификации этих ПЦР продуктов, не выявил укороченных рестриктных вариантов рДНК. Было высказано предположение, что укороченные ПЦР-продукты образуются *in vitro* в результате преждевременной терминации полимеразной реакции.

Для проверки этого предположения, были проведены контрольные эксперименты, в которых в качестве матрицы использовали ДНК космидного клона 84A3, содержащего рДНК человека и ранее выделенного из клонотеки хромосомы 13 человека (LA13NCO1, Los Alamos). Оказалось, что и в этом случае амплифицируются два фрагмента ДНК –1,8 т.п.н. и 0,8 т.п.н. Это означает, что укороченный фрагмент ДНК действительно образуется как побочный продукт ПЦР-амплификации, и причиной этого может быть преждевременная терминация реакции полимеризации.

Для определения структуры укороченного фрагмента проводили его клонирование в pGEM T-Easy вектор, отбирали клоны, дающие положительные сигналы со специфической олигонуклеотидной пробой 0P, и секвенировали вставки рекомбинантных клонов. Оказалось, что в укороченном амплификанте отсутствует один *Alu* элемент и последовательность в 700 п.н., разделяющая *Alu3* от *Alu4*.

Известно, что степень отличия *Alu3* и *Alu4* из рМГС человека составляет 25% [Gonzalez et al., 1989]. Наличие в *Alu3* и *Alu4* специфических нуклеотидных вариаций позволила нам выявить химерную (рекомбинантную) структуру *Alu*-повтора в 0,8 т.п.н. фрагменте. (рис 5)

TGTAATCCCAGCACTTTGGGAGG

1) GG**ACGAGGCT**GGGGCG**GGTTG**CACGCCTGTTCATCCCAGCACTTTGGGAGGCTAAGGCAGGCGGAC-65
 2) GGCCGGGCACGGGCGC-----ACGCCCGTCATCCCAGCACTTCGAGAGGCTAAGGCAGGCGGAC
 3) GGCCGGGCACGGT**CGT**-----CACGCCCGTCATCCCAGCACTTCGAGAGG**CGAGGTGGGCGCAT**

1) CACCTGAGGTTGGGAGTGGGAGACCAGCCTGACCAACATGGAAAAACACC--GTCTCTACTAAAA-128
 2) CACCTGAGGTTGGGAGTGGGAGACCAGCCTGAACAACATGGAAAAACACC--GTCTCTACTAAAA
 3) CAC**AGGAGGT**CGGGAGT**TGGGAGACCAGCCTGAGCAACATGGAGAGACACGGCGTGCCTACTAAAA**

1) GTACAAACATCAGCCA-----GCACGGTGGCCCATGCCT----GT-AATCCCAGCTAATCAGGAG-183
 2) GTACAAACATCAGCCAAGCCAGCACGGTGGCCCATGCCT----GT-AATCCCAGCTAATCAGGAG
 3) **ACACAAACATCAGCCAAGCCAGG--CGTGGGGG-TGCCTCCCTGTCAATCCCCGCTAATCGGGAG**

1) GCT-GAGGCAGGAGAATCGCTTGAACCTGGGAGGCGGAGGGTGCGGTGAGCCGAGATCGCGCCAT-247
 2) GCT-GAGGCAGGAGAATCGCTTGAACCTGGGAGGCGGAGGGTGCGGTGAGCCGAGATCGC**ACCAT**
 3) GCT**GGAGGCAGGAGAAGCGCTCGAACC**CGGGAGGCGGA**AGGTGCGGTGAGCC****AAGATCGCGCCAT**

1) TGCCCTCTAGCCTGGGCAACAAGAGTGAAACTCTGTCTCAAAAAAAAAAG-295
 2) TGCCCTCTAGCCTGGGCAACAAGAGTGAAACTCTGTCTCAAAAAAAAA
 3) **TGCACTCTAGCCGTGGAAACAAGAGTGAAACTCTGTCTCAAAA**

Рис.5 Сравнение рекомбинантного *Alu* 0,8 т.п.н. фрагмента с известными вариантами *Alu3* и *Alu4* из рМГС человека. Строки 1-3: *Alu3*, химерный *Alu*, *Alu4*. Номера справа возле 1 строки соответствует позициям нуклеотидов в *Alu3*. Жирным шрифтом выделены нуклеотиды характерные для *Alu3* и *Alu4*, черточками - делеции. Светло-серым цветом отмечен *Alu3* темно-серым *Alu4*. Положение последовательности *Alu* - “core” показано над верхней строкой.

Анализ нуклеотидной последовательности рекомбинантного *Alu*-повтора показывает ее идентичность с *Alu4* от 5'-конца до нуклеотида C52, и последующее ее совпадение с *Alu3*, начиная от нуклеотида T53 (рис. 5). Это означает, что *Alu* в составе укороченного ПЦР фрагмента состоит из половинок *Alu3* и *Alu4*. По-видимому, досрочная терминация происходит на матрице *Alu3*, и синтез останавливается на нуклеotide 47-52 при элонгации праймера P1. С уверенностью можно сказать, что с нуклеотида C53 наблюдается гомология между рекомбинантным (химерным) продуктом и *Alu3*, более точно сказать сложно так как от нуклеотида G47 до нуклеотида C52 *Alu3* и

Alu4 не отличаются по первичной структуре. В случае отжига любого из вариантов преждевременно терминировавших последовательностей с гомологичными участками *Alu3* при последующей достройке может образовываться наблюдаемый нами рекомбинантный (химерный) продукт. Это может происходить по схеме на (рис.6).

Оставалось неясным, является ли обнаруженный феномен спецификой для данного участка рМГС. Поэтому эксперимент, аналогичный выше описанному, был проведен нами на соседней области рМГС, ограниченной праймерами P1 (5'-gttactctgaaaacggaggc -3'), P2 (5'-gttgcgtttcattttcattttca -3'), содержащей однонаправленные повторы *Alu1* и *Alu2* эти повторы разделены 93 п.н., которые образованы АТ обогащенными последовательностями и микросателлитным кластером (TACA)_n. Повторы *Alu1* и *Alu2* в этом участке рМГС имеют направление, противоположное направлению повторов *Alu3* и *Alu4* (рис.3). При ПЦР-амплификации этой области также было получено два продукта 1 и 0,6 т.п.н. Секвенирование этих продуктов показало химерную природу укороченного варианта (данные не переведены).

Таким образом, феномен преждевременной терминации ПЦР, в определенных участках *Alu*, с образованием делетированного (укороченного) продукта показан для двух зон рМГС, содержащих как однонаправленные *Alu*, так и микросателлитные повторы между ними.

Полученные данные находятся в соответствии с литературными, в которых отмечалось, что амплифицированный продукт не всегда может точно отражать структуру нативной ДНК. В частности, при амплификации сегментов, образованных мультигенными семействами, авторы наблюдали появление, кроме ожидаемых, дополнительных продуктов амплификации [Bradley and Hillis, 1997], [Scharf et al., 1987], [Bradley et al., 1993] Эти авторы сделали предположение, что подобные продукты могут возникать за счет того, что укороченные копии амплифицируемой ДНК гибридизуются с альтернативной гомологичной матрицей и выступают в качестве праймеров в последующих циклах амплификации. Однако природа укороченных продуктов ими не была определена.

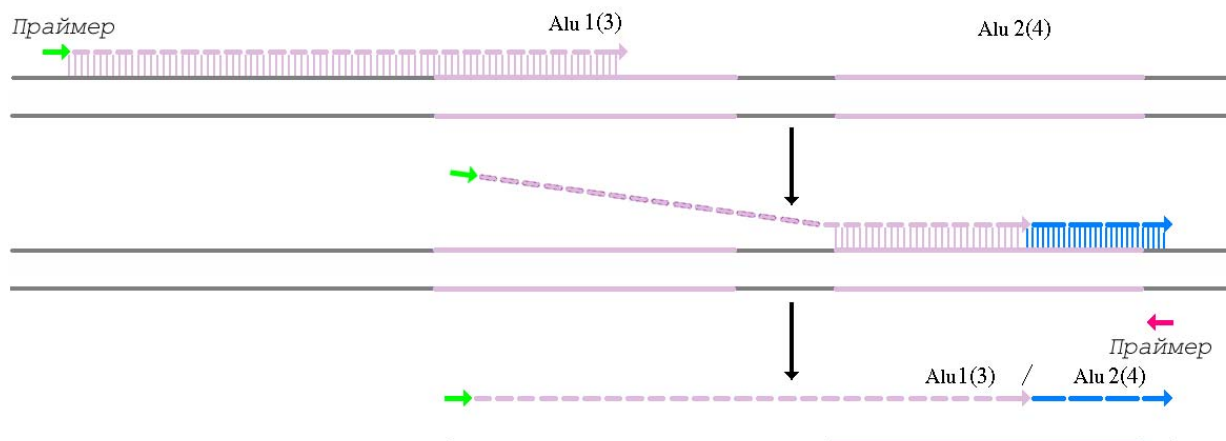


Рис 6. Предполагаемая схема образования укороченного ПЦР-продукта.

В наших экспериментах было впервые показано, преждевременная терминация происходит в определенной точке *Alu* элементов. Этот феномен был показан при ПЦР-амплификации двух независимых участков рМГС, имеющих сходную структурную организацию и был также воспроизведен при амплификации аналогичного участка рМГС в составе космидного клона.

Остается неясной причина преждевременной терминации ПЦР внутри *Alu* повторов, ведущей к появлению рекомбинантных продуктов. Вполне вероятно, что наиболее консервативный участок *Alu* (*Alu* - “core”), будучи частью *PoIII* промотора, может иметь повышенное сродство к *Taq* полимеразе, тормозить ее движение вдоль цепи ДНК и тем самым вызывать явление преждевременной терминации *in vitro*, которое ведет к появлению неполноразмерных амплифицированных продуктов.

Очевидно, что описанное нами явление необходимо учитывать в исследованиях по сравнительной геномике, когда используется ПЦР-амплификация участков, содержащих сложно организованные повторяющиеся элементы генома, и требуются дополнительные доказательства существования аллельных вариантов при образовании более одного амплификанта.

2. Природа гетерогенности сегмента LR2_{var} рМГС человека.

рМГС человека содержит в своей центральной части два крупных повтора (~2 т.п.н.) LR1 и LR2, гомологичных между собой на 90% (рис. 7), каждый из которых включает большого числа микросателлитов.

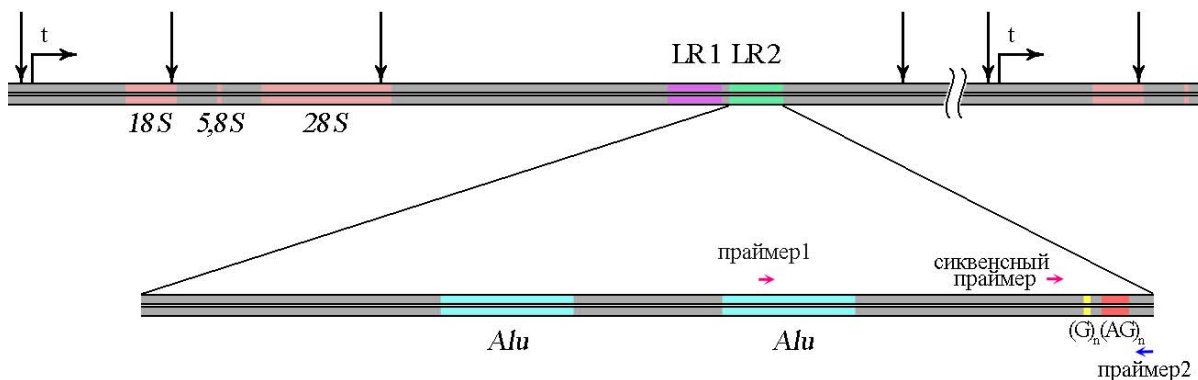


Рис.7 Расположение LR1 и LR2 в повторе рДНК. Вынесено подробное строение LR2.

Нами изучался внутригеномный полиморфизм участка LR2, обогащенного поли (G)_n- и (AG)_m- кластерами, обозначенного нами как LR2_{var}. В первой серии экспериментов проводили ПЦР-амплификацию этого участка с праймерами П1 (5'-ttctgtcagtctgtcagacac-3') 22764-22786 и П2 (5'- gaaacccccctgactcaggtcaag -3') 23523-23500 на суммарной клеточной ДНК одного индивида. Продукт амплификации, размером ~ 750 п.н., клонировали в векторе рGEM-T Easy. В результате получили 35 рекомбинантных клонов. Секвенирование этих клонов показало высокую гетерогенность клонированных участков LR2_{var}.

В последующих экспериментах проводили амплификацию LR2_{var}, используя ДНК 10 неродственных индивидов. Всего было клонировано и секвенировано 689 копий сегмента LR2_{var}, с координатами 22763-23523, полученных из геномов десяти индивидов. Сравнение последовательностей также выявило их высокую гетерогенность и специфические особенности строения.

Центральную часть LR2_{var} занимают протяженные кластеры (G)_n(AG)_m. Значения 'n' и 'm' существенно варьируют между копиями LR2_{var} в интервале 'n'- от 5 до 12 и 'm' от 13 до 27, соответственно, при случайных комбинациях 'n' и 'm'. Варианты (G)₈₋₁₁ и (AG)₁₈₋₂₀ наиболее часто встречаются в тотальной выборке (таблица 1).

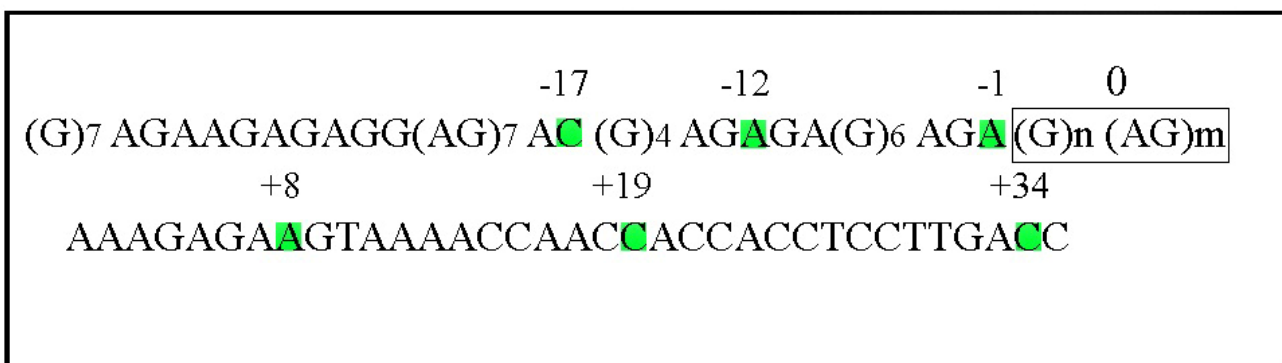


Рис. 8 Схема строения LR2_{var}. Цифрами обозначены точковые замены на основании которых происходило деление на группы.

Для систематизации последовательностей, фланкирующих центральный кластер, их сравнивали без учета варьирующих значений 'n' и 'm', в результате чего сегменты были подразделены на 4 группы на основании нуклеотидной замены в положении -17, где возможны C или T (Рис.8). Группа А наиболее многочисленна, представлена 13 вариантами и включает 567 из 689 проанализированных последовательностей, что составляет 82,29%. Наиболее представлен вариант 1 включающий 334 последовательности. В группу В попали 14 вариантов последовательности LR2_{var} (12,33%). Наиболее распространенный вариант представлен 15-ю копиями. Также следует отметить, что группа В отличается от группы А, в большинстве случаев делециями во фланкирующих областях кластера (G)_n(AG)_m в положении от -12 и до +7-9.

К группе А+В были отнесены последовательности, содержащие в своем составе участки, характерные как для А, так и для В (в таблице 1 выделены различной окраской). Следует отметить, что их число невелико. Наиболее представленный вариант состоит из 5'-части последовательности А и 3'-от В, и встречается 5 раз.

Группа С наименьшая - 12 вариантов (1,75%). В нее попали варианты в большинстве своем содержащие участки гомологии с LR1 см. диаграмму 1.

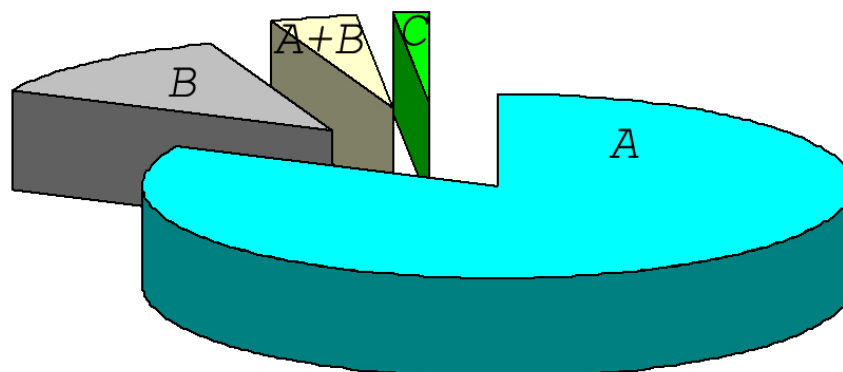


Диаграмма 1 Представленность вариантов групп А, В, А+В и С в выборке из 689 последовательностей.

Как видно из диаграммы, наиболее представлены нуклеотидные последовательности группы А, которые практически идентичны (без учета варьирующих значений 'n' и 'm' в центральном кластере) ортологичному участку из последовательности GeneBank U13369.

В группе В (12,3%, от всех LR2_{var}) обнаруживается гетерогенность до и после центрального кластера (G)_n (AG)_m (Табл.1). Фланкирующие участки с 5'-конца у последовательностей группы В включают переменные кластеры (G)₆₋₈ и (AG)₆₋₁₀, одну замену С→Т в фиксированном положении -17 и короткие делеции перед кластером (G)_n (AG)_m. Для 3'- фланкирующей области характерны делеции, примыкающие к центральному кластеру (G)_n (AG)_m и замещения оснований в фиксированных позициях +8, +19 и +34.

Последовательности группы С (1,75% от всех LR2_{var}) сильно отличаются от таковых из групп А и В. Три варианта не содержат центрального кластера (G)_n (AG)_m. Вместо этого, первый кластер (AG)₆₋₁₀, характерный для аллелей В, расширен до значений (AG)₁₈₋₃₂ и часто содержит замены G → С. (обозначены звездочкой в таблице 1). Четыре последовательности в своем составе имеют группу нуклеотидов отличающихся от канонического состава LR2. 3'- фланкирующая последовательность содержит замещения оснований, делеции и инсерции. Сравнение последовательностей LR2_{var} с ортологами из LR1_{var} (GeneBank, U13369) часто обнаруживает идентичные замещения оснований. Это с высокой вероятностью позволяет предположить фрагментный обмен LR1 с LR2 и/или вставку сегментов LR1 в LR2.

Различия между нуклеотидными последовательностями, фланкирующими центральный блок микросателлитов (G)_n (AG)_m в группах А-С могли возникнуть в результате кроссинговера и мозаичных межаллельных переносов. В группах А и В, были выделены варианты LR2_{var} объединяющие в себе характерные особенности обеих групп. Их число невелико, что позволяет предположить, что рекомбинации в центральной части LR2_{var} являются весьма редким событием.

Мини-группа С LR2_{var} лишена центрального (G)_n (AG)_m кластера, тогда как предшествующий кластер (AG)₆₋₁₀ расширен до значений (AG)₁₈₋₃₂. В 3'- части последовательностей С-группы содержатся замены, характерные для LR1. Соответствующее выравнивание позволяет предположить, что варианты С возникли в результате генной конверсии между LR1_{var} и LR2_{var}. В кластере (AG)₇ могли возникнуть двуцепочечные разрывы, аналогичные тем, которые описаны для минисателлитов [Buard J., et al., 2000; Jeffreys A.J., May C., 2004], с последующей достройкой выступающих одностранных концов, приводящей к внедрению LR1 подобных участков.

Предполагаемое внедрение участка LR1 в аллели C2 и C5 начинается с нуклеотида -17. Конечную точку внедрения определить сложно в виду идентичности двенадцати 3'-концевых нуклеотидов в LR1_{var} и LR2_{var} (таблица 1). Таким образом, можно предположить, что выраженная гетерогенность LR2 может возникать по нескольким механизмам, включающим дубликации, делеции и конверсионный перенос блоков повторов между аллелями.

Интересно отметить, что вся область LR2_{var} вплоть до позиции +11 на Рис. 8 состоит из пуринов, за исключением C/T в позиции -17. Поскольку во всех случаях за C/T следует G, есть основания полагать, что «Т» возникают в результате деметилирования 5-MeC в CpG.

Таким образом, нами обнаружен высокий уровень внутри- и межгеномного полиморфизма участка LR2_{var} в составе рМГС человека. Его вариации связаны с микросателлитной изменчивостью, инсерциями/ делециями, с точковыми мутациями (нуклеотидными заменами) и фрагментными обмeнами между LR1 и LR2.

3. Анализ *in silico* горячих точек рекомбинации, связанных с рДНК.

Описанная выше высокая гетерогенность LR2_{var} может быть связана с повышенной частотой рекомбинационных событий в области LR1-LR2 рМГС. В этом случае фрагменты данной области можно было бы обнаружить в различных ЯОР - хромосомах. Мы провели поиск наличия LR1-LR2 фрагментов рДНК *in silico* в различных хромосомах человека. Такие фрагменты различного размера со средней степенью гомологии ~80% были обнаружены на нескольких хромосомах. Распределение концевых участков фрагментов рДНК может соответствовать положению горячих точек рекомбинации в рДНК.

Было обнаружено, что участки, гомологичные области LR1-LR2, наиболее часто встречаются на ЯОР⁺ хромосомах **2** – 10 раз, на хромосоме **12** – 5 раз, а также на хромосомах **3**, **6**, **19** и на половых хромосомах среди гипотетических транскрипционных факторов.

При изучении последовательностей, окружающих горячие точки рекомбинации, оказалось, что наиболее часто они представлены микросателлитами типа (TC)_n, (TG)_n, поли AT – последовательностями, а также вблизи них часто встречаются последовательности, потенциально способные образовывать квадруплексы согласно формуле GxNaGxNbGxNcGx, где x – число G оснований в каждом G-тракте, из четырех, формирующих тетраду.

4. Сегментная дупликация прилегающих к рДНК участков генома человека.

Следующей задачей нашей работы являлась характеристика ранее не изученных участков, дистальных по отношению к кластеру рДНК (рДО) хромосомы **13** человека. Для экспериментов использовали клонотекку рекомбинантных космид LA13NCO1, полученную из хромосомы **13** человека, очищенной на хромосомном сортере. Клонотекка состояла из 16896 отдельных клонов, что соответствует четырем эквивалентам ДНК хромосомы **13** человека. Использовали фильтры с упорядоченной панелью клонов, ранее изготовленные в лаборатории Н.К. Янковского (ИОГен) и любезно предоставленные нам для работы. Отбор соответствующих клонов проводился с помощью гибридизации с пробой DJU-R [Gonzalez and Sylvester, 1997], специфичной для рДО акроцентрических хромосом человека и высших приматов. Из 1500 клонов семь дали положительный сигнал, причем у двух (92E7 и 82F6) ярко выраженный, и значительно слабее у клонов 84A3, 91D10, 86G5, 76G2 и 123C5. Блот-гибридизационный анализ ДНК этих клонов с пробой DJU-R подтвердил специфичность гибридизации колоний для 92E7 и 82F6 (рис.9).

Последовательная гибридизация EcoRI-фрагментов ДНК с пробами, специфичными для значащих областей рДНК, показала, что 3'-конец вставки клона 82F6 локализован в 18S рДНК, а 3'-конец вставки клона 92E7 находится в области 28S рДНК (рис.10). Это позволило предположить, что протяженность фрагмента рДО в клоне 82F6 больше, чем в 92E7. Таким образом, клон 82F6 оказался наиболее перспективным для изучения прителомерной зоны, соседствующей с кластером рДНК.

Отобранные клоны были охарактеризованы также с помощью FISH на метафазных хромосомах. Работа производилась О.В. Муравенко из Института молекулярной биологии им. В.А. Энгельгардта Российской академии наук.

ДНК клонов 82F6 и 92E7, как и следовало ожидать, хорошо гибридизовалась с рДНК в прителомерных областях всех акроцентрических хромосом (данные не приведены).

Поскольку размер космидной вставки 82F6 составляет ~30 т.п.н., а один из ее концов картирован нами в зоне 18S рДНК, можно было ожидать, что эта вставка простирается вглубь прителомерной зоны. Секвенирование ДНК клона 82F6 проводили, используя в качестве праймеров DJU-R и произвольно выбранные олигонуклеотиды по ходу секвенирования. Ранее был секвенирован участок 8.3 т.п.н., дистальный относительно кластера рДНК в хромосоме 21 (acc.no.U67616) [Gonzalez and Sylvester, 1997]. Нами была показана высокая гомология (98 %) между участком клона 82F6 из хромосомы 13 и 1700 п.н. 5'-концевого участка 8.3 т.п.н. хромосомы **21** человека и

выйти за пределы этого участка в направлении теломеры. Было просеквенировано еще 3500 п.н. (Асс. Nu.AF478540) в ранее не изученной прителомерной зоне.

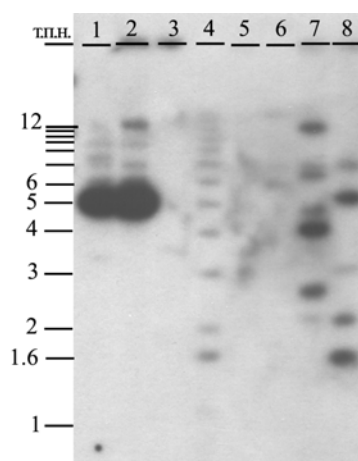


Рис. 9 Блот-гибризационный анализ ДНК клонов с пробой DJU-R. 1-92E7, 2-82F6, 3-84A3, 4- маркер, 5-91D10, 6-86G5, 7-76G2 и 8-123C5.

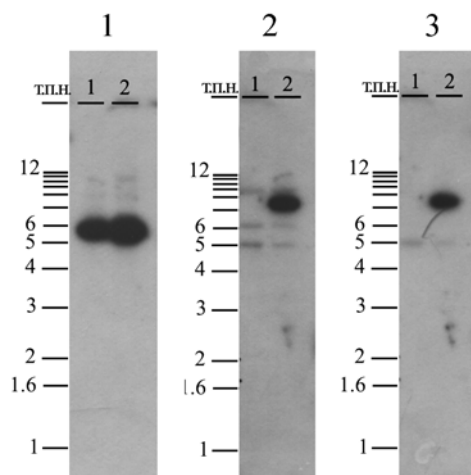


Рис.10 Картирование концов вставок клонов 82F6 и 92E7, гидролизованных эндонуклеазой *EcoRI*. 1-гибридизация с пробой r2, 2-с пробой r3 и 3- с пробой r1 (см. рис.1).

Поиск в базе данных показал, что определенная нами нуклеотидная последовательность обладает значительной гомологией (83-84%) с ВАС-клоном из хромосомы **19** (Асс.пн. АС006504), расположенным наиболее близко к центромере из всех секвенированных на сегодня клонов.

Кластеры рДНК образуют мультигенные семейства и при этом располагаются на ЯОР+ хромосомах человека перед теломерами. Полноразмерный повтор рДНК человека был секвенирован сравнительно давно [Gonzalez and Sylvester, 1995]. Нами выявлена высокая гомология (98%) рДО хромосом **13** и **21** на протяжении более 1700 п.н. и определена последовательность еще ~ 3300 п.н. в направлении теломеры. Обнаружена протяженная область гомологии (81-84%) между прителомерными областями хромосом

13 и прицентромерным участком хромосомы **19**, а также «перебегающая» гомология между прителомерным и прицентромерным участками хромосомы **13**.

Проведенное нами сравнение последовательности рДНК человека с полноразмерной последовательностью генома человека выявляет на ЯОР⁻ хромосомах большое число участков с высокой (в среднем, 80% гомологией) средней длины 1-1,5 т.п.н. Особенно часто на ЯОР⁻ хромосомах присутствуют фрагменты последовательностей, гомологичные тем, которые предшествуют кластерам рДНК как это показано на (рис.11). Возможно, это говорит о частых перестройках генома, в которых участвуют области ЯОР⁺ хромосом, дистальных по отношению к кластерам рДНК.

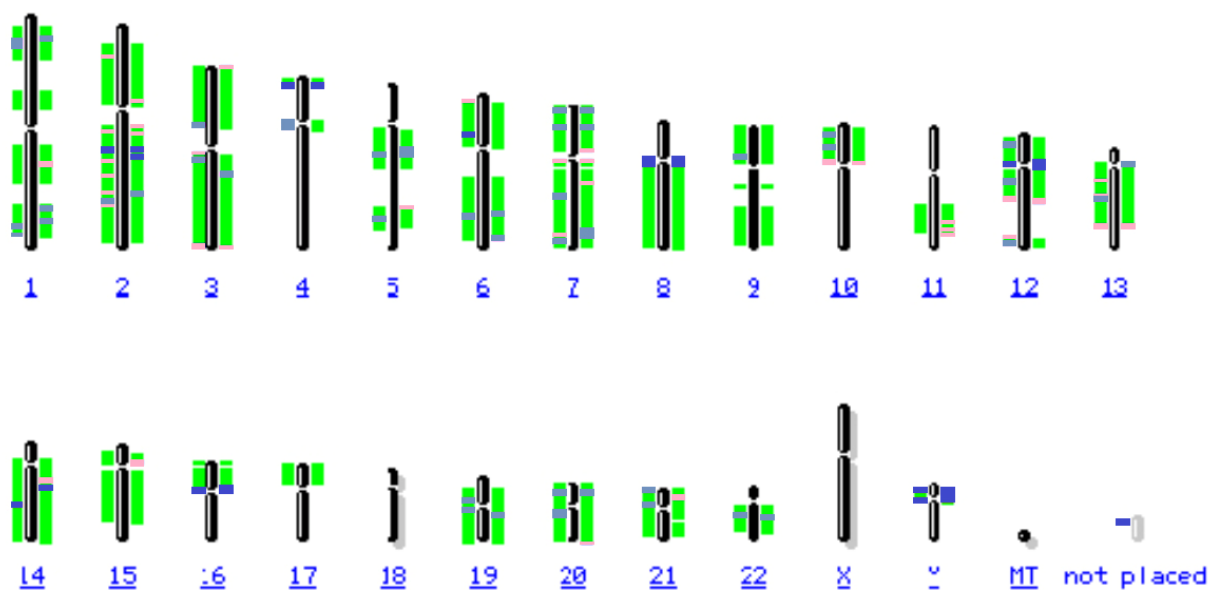


Рис. 11. Фрагменты, предшествующие рДНК человека, повторяющиеся на ЯОР⁻ хромосомах. Штрихи, перпендикулярные «телу» хромосомы, обозначают участки, гомологичные тем, которые встречаются перед кластерами рДНК на хромосоме **13**. Цвет штрихов отражает размер сегментов. Светло серый – от 50 до 80 н., серый – от 80 до 200 н., и темно-серый – больше 200 н. Чрезвычайно высокая плотность фрагментов с высокой степенью гомологии на большинстве хромосом свидетельствует в пользу ранее сделанного заключения о том, что участки, предшествующие кластерам рДНК, часто вступают в рекомбинации с различными хромосомами.

Можно было ожидать, что рДО, представляющие собой уникальные нуклеотидные последовательности на акроцентрических хромосомах, эволюционируют независимо друг от друга. Вместо этого обнаружена высокая межхромосомная гомогенность в районе рДО, а небольшие вариации свойственны в равной степени как гомологичным, так и негомологичным хромосомам, что напоминает ситуацию, характерную для наиболее консервативных значащих областей рДНК, в эволюции которых преобладают частые внутри- и межхромосомные рекомбинации. По всей видимости, истинной причиной консервативного характера рДО акроцентрических хромосом является активная

независимая их гомогенизация, аналогичная процессу для значащей и регуляторной областей генов рРНК.

В настоящее время намечается тенденция к пересмотру основных концепций эволюции геномной ДНК. Общепринятый взгляд на эволюцию генома до сих пор базировался на постулатах о неизменности его общего строения, на фоне которого происходило лишь не более одной крупной реорганизации за 10 млн. лет [Bailey J.A., et al., 2002], [Melford H.C., Trask B.J. 2002].

Теперь же в результате реализации программы “Геном человека” появились свидетельства беспрецедентно высокой частоты сегментных дупликаций в ходе эволюции приматов, т.е. за последние 35 млн. лет [Bailey J.A., et al., 2002]. Сегментные дупликации представляют собой дупликативную транспозицию геномной ДНК, размером от одной до сотен т.п.н. при среднем значении гомологии 95.5%. Дупликации могут существовать в виде тандемов, [van Geel Mat., et al., 2002] но чаще они рассеяны по геному и бывают как внутри-, так и межхромосомными. Межхромосомные дупликации предпочтительно локализуются в прицентромерных и прителомерных областях хромосом. Анализ черновой версии генома показал, что прицентромерные области обогащены межхромосомными дупликациями не менее чем в 4.5 раза. а прителомерные - не менее чем в 2.7 раза [Murphy W.J., et al., 2001]. Структура и эволюция наиболее значительных по размеру (>1 т.п.н) и степени гомологии (>90%) сегментных дупликаций хромосомы **22** подробно рассмотрены в [Bailey J.A., et al., 2002]. Один из самых интересных результатов этой работы - обнаружение на хромосоме **22** человека участка, для которого не найдено дубликонов на хромосомах высших приматов, и именно этим участком заканчивается расшифрованная на сегодняшний день нуклеотидная последовательность в прицентромерной области хромосомы **22**. Это позволяет думать, что самые молодые дубликоны расположены наиболее близко к центромерам.

Наши данные о гомологии между рДО хромосомы **13** и областью **19q12** и возрастанию этой гомологии в хромосоме **13** в направлении теломеры вносят вклад в изучение дупликативной транспозиции ДНК и, тем самым, в развитие общей картины эволюции генома. Представляет интерес то, что протяженный участок гомологии (~2 т.п.н.) обнаружен именно между прителомерным и прицентромерным участками негомологичных хромосом (рис. 12)

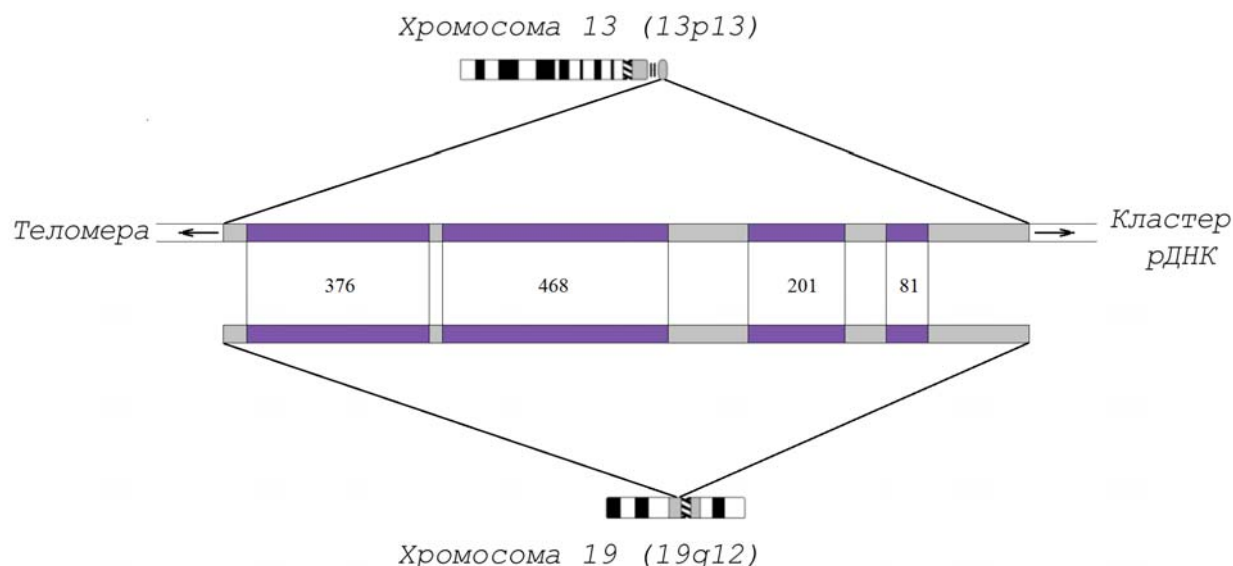


Рис.12 Схематическое сравнение фрагмента рДО хромосомы 13 человека с гомологичным участком ВАС-клона, содержащего прицентромерную область хромосомы 19 человека.

Таких случаев описано немного. Наиболее известный из них - участки, специфичные для хромосомы **22** и других акроцентрических хромосом, в локусе **2p11** [Horvath J., et al., 2000]. Общеизвестно, что эта гомология обусловлена слиянием концов двух предковых форм (синтетических хромосом **12** и **13** шимпанзе) около 5 млн. лет назад, что и привело к возникновению хромосомы **2** человека [Ijdo J., et al., 1991], [Luke S., Verma R.S. 1992]. Высокоповторяющиеся и паралогичные участки **4q24** и **4q35** в виде мозаичных вкраплений присутствуют на коротких плечах акроцентриков и в прицентромерных и прителомерных областях различных хромосом [Piccini I., et al. 2001]. Недавно при изучении прителомерных дупликаций у гоминоидов обнаружено, что копия, филогенетически ортологичная прителомерному участку хромосомы **1** человека (HSA1q42.3), у бабуина локализована в прицентромерном пространстве (PNA 1q) [van Geel Mat et al., 2002]. Можно предположить, что снижение порога степени сходства до 80-90% при поиске гомологичных нуклеотидных последовательностей расширит число выявленных дупликаций за счет более древних.

Степень гомологии между рДО хромосом **13** и **21** и прителомерной областью хромосомы **19** относительно невысока (от 81 до 84%), что, по-видимому, указывает на древний ("доприматный") возраст дупликации. Обе нуклеотидные последовательности гомологичны последовательности *Alu* ретропозона, что говорит о его присутствии в данной области еще до дупликации. Интересно, что пока участок прицентромерной области хромосомы **19** подвергался независимой дивергенции, гомологичные участки в рДО почти всех акроцентрических хромосом эволюционировали согласованно, что

предполагает эволюционное давление, а, следовательно, и важное функциональное значение поддержания консервативности рДО.

За последние несколько лет получено чрезвычайно много новых данных не только о механизмах регуляции экспрессии рРНК, но и о строении как ядрышка в целом, так и его многочисленных субструктур. Полученные результаты позволили назвать ядрышко клеточным сенсором, т.к. многие реакции в ответ на внутренние и внешние стрессы затрагивают его в первую очередь. Поскольку модуляция синтеза рРНК является ведущим фактором в скорости образования зрелых рибосом, любой, даже незначительные вариации в структуре рДНК (например, SNP) могут оказывать серьезное влияние на активность рибосомного синтеза и, следовательно, нуждаются в специальном изучении.

ВЫВОДЫ

1. Показана консервативность двух сложно-организованных локусов рМГС человека вблизи промотора генов рРНК, содержащих однонаправленные *Alu*-повторы и тандемные повторы между ними. Обнаружен феномен образования укороченных рекомбинантных (химерных) молекул ДНК при ПЦР-амплификации этих локусов в результате преждевременной терминации Taq-полимеразы в определенных участках *Alu*-повторов.
2. Обнаружен гипервариабельный участок LR2_{var} в центральной части рМГС человека. Установлено, что гетерогенность LR2_{var} определяется микросателлитной изменчивостью, как следствие репликативных ошибок, однонуклеотидными заменами, инсерциями/делециями небольших участков и конверсионными обменами между LR2 и LR1.
3. Показана высокая гомология прителомерного участка (5 т.п.н.) хромосомы **13** с прицентромерным участком хромосомы **19**.
4. С помощью анализа *in silico* установлены возможные «горячие точки» рекомбинации в рДНК и оценена представленность фрагментов области LR1-LR2 на различных хромосомах человека.

Список работ, опубликованных по теме диссертации.

Шибалёв Д.В., Воронов А.С., Башкиров В.Н., Куприянова Н.С., Рысков А.П. Обнаружение рекомбинантных продуктов при ПЦР-амплификации ДНК, содержащей однонаправленные Alu-повторы. Доклады академии наук. 2003. Т. 388. С. 55-59.

Куприянова Н.С., Шибалёв Д.В., Воронов А.С., Муравенко О.В., Зеленин А.В., Рысков А.П. Сегментные дубликации прителомерных участков хромосомы 13 человека. Молекулярная биология. 2003. Т. 37. С. 221-227.

Kupriyanova N.S., Shibalev D.V., Voronov A.S., Ryskov A.P. PCR-generated artificial ribosomal DNAs from premature termination at Alu sequences. Biomolecular Engineering. 2004. V.21. P. 21-25.

Шибалёв Д.В., Воронов А.С., Фирсов С.Ю., Рысков А.П., Куприянова Н.С. Обнаружение внутригеномного полиморфизма участка LR2 межгенного рибосомного спейсера человека. Молекулярная биология. 2004. Т. 38. С. 835-838.

Kupriyanova N.S., Shibalev D.V., Voronov A.S., Ryskov A.P. Enhanced heterogeneity of the LR2 segment in the human ribosomal intergenic spacer. Gene. 2008. V. 425. P. 44-47.

Материалы конференций.

А.П. Рысков, Д.В. Шибалёв, К.К. Нечволодов, Н.С. Куприянова. Структурная организация и полиморфизм рибосомного межгенного спейсера у человека и высших обезьян. Конференция по научным направлениям Программы фундаментальных исследований Президиума РАН «Фундаментальные науки – медицине», Москва, 2007 г. С. 156-157.

А.П. Рысков, Д.В. Шибалёв, К.К. Нечволодов, Н.С. Куприянова. Изучение внутри и межгеномного полиморфизма нуклеотидных последовательностей ДНК области ядрышкового организатора у человека и высших обезьян. Конференция и семинары по научным направлениям Программы фундаментальных исследований Президиума РАН «Фундаментальные науки – медицине», Москва, 2008 г. С. 212-213.