

УДК 577.213.6

ГЕНОМНАЯ НЕСТАБИЛЬНОСТЬ И НЕКАНОНИЧЕСКИЕ СТРУКТУРЫ ДНК

Б. Л. ЗЫБАЙЛОВ¹, В. И. ГЛАЗКО²

(¹ Медицинский университет Арканзаса (UAMS);
² РГАУ-МСХА имени К. А. Тимирязева)

Рассмотрены источники генетической нестабильности, связанные с неканоническими формами ДНК (квадруплексы, шпилечные структуры) и их распределение в секвенированных последовательностях ISSR-PCR маркеров домашней лошади и крупного рогатого скота. Показана относительно повышенная частота локализации квадруплексов в исследованных ДНК фрагментах, представленных участками гомологии к эндогенным ретровирусам. Обсуждается участие квадруплексов в процессах стабилизации/дестабилизации геномной ДНК.

Ключевые слова: ISSR-PCR маркеры, квадруплексы, шпилечные структуры, двуцепочечные разрывы ДНК, рекомбинации, эндогенные ретровирусы, геликазы.

Н.И. Вавилов был одним из первых ученых, который привлек внимание к проблеме научных основ селекционной работы, он считал, что селекция представляет собой эволюцию, направляемую волей человека, а одним из важнейших условий для управления этим процессом является выяснение источников генетической изменчивости [2]. Изучение закономерностей мутационной изменчивости на уровне фенотипических признаков привело его к выявлению гомологических рядов в наследственной изменчивости [1].

Одной из центральных задач руководимых им научно-исследовательских коллективов была разработка методов индуцированного мутагенеза с целью увеличения исходного генетического разнообразия для увеличения эффективности селекционной работы. В то же время высокий уровень генетической изменчивости необходим для отбора и подбора носителей желательных признаков, после получения планируемого сочетания которых следующей задачей становится необходимость их сохранения и, значит, снижения генетической изменчивости. С этой точки зрения селекционная работа подразделяется на два качественно разных этапа: на первом — использование методов увеличения геномной нестабильности, на конечном — поиск способов геномной стабилизации.

Источники геномной нестабильности

К настоящему времени описано множество источников геномной нестабильности на разных уровнях организации генетического аппарата [26]. Единственной общей характеристикой для большинства из них является участие в геномной дестабилизации двуцепочечных разрывов ДНК (DSBs) [15]. В последние годы накапливаются данные о тесной связи между повышенной частотой DSBs и присутствием неканонических структур ДНК. Чаще всего к таким структурам относят квадруплек-

сы. Квадруплексы исследуются особенно активно в связи с тем, что последовательности, потенциально способные к их формированию, выявлены в таких структурно-функционально важных участках ДНК, как промоторы, теломеры, участки повышенной ломкости ДНК [10]. Квадруплексы формируют четырехцепочечные структуры, в которых гуанины образуют плоскость за счет своих поперечных связей и отличаются повышенной термоустойчивостью. Разработан ряд компьютерных программ, позволяющих вести поиск в секвенированных последовательностях участки, потенциально способные к образованию таких квадруплексов [32].

Накапливаются экспериментальные данные, свидетельствующие об участии квадруплексов в различных типах геномной нестабильности. К наиболее распространенным типам относят изменчивость по копийности хромосом (нарушение прохождения митоза, работы анафаза-промотирующего комплекса), хромосомные транслокации, формирование химерных белков за счет хромосомных транслокаций, мутации в генах-супрессорах опухолей, транспозиции и рекомбинации мобильных генетических элементов SINE Alu в опухолях у человека. Оказалось, что все перечисленные типы геномной дестабилизации тесно связаны с последовательностями, способными образовывать квадруплексы. Особое внимание к квадруплексам обусловлено еще и тем, что квадруплексы могут быть стабилизированы рядом определенных молекул (например, замещенными порфиринами или теломестатином), что позволяет влиять на геномную нестабильность, связанную с нарушением работы таких структур. Некоторые примеры геномной нестабильности, связанной с формированием квадруплексов, представлены ниже.

Стабилизаторы G квадруплексов рассматриваются как новые блокаторы перехода G2/M [29]. К естественным стабилизаторам относят теломестатин, синтезируемый макроцикл, выделенный из *Streptomyces anulatus*. Искусственно синтезированы структурно похожие на теломестатин макроциклы, из которых HXDV, гексаоксазолоновый макроцикл, показал высокую специфичность к связыванию с G квадруплексами *in vitro* и проявлял антипролиферативную активность. HXDV индуцирует острый апоптоз в клетках млекопитающих через 16 ч после обработки и останавливает прохождение митоза. Таким образом, HXDV, действие которого зависит от связывания с G квадруплексами, является новым блокатором митоза.

Другой синтетический аналог теломестатина, близкий по структуре к HXDV, но не способный связываться с G квадруплексами, не обладает антипролиферативной активностью и не индуцирует апоптоз. Полученные данные свидетельствуют о том, что все наблюдаемые клеточные эффекты HXDV являются следствием его связывания с G квадруплексами.

Не исключено, что одним из факторов, принимающим участие во всех клеточных процессах, связанных с нарушением сегрегации хромосом, индукции апоптоза, являются события, обусловленные балансом изоформ известного опухолевого супрессора, p53. Ген TP53 имеет несколько вариантов сплайсинга. Полностью сплайсированный транскрипт p53 (FSp53) кодирует канонический p53 белок. Альтернативный сплайсированный p5312 транскрипт сохраняет интрон 2 и кодирует D40p53 (или DNp53), изоформа, утрачивающая первые 39 аминокислот N-участка, соответствующего трансактивации. Показано формирование G квадруплекса (G4) в G обогащенной области интрона 3, который влияет на сплайсинг интрона 2. Обработка лимфобластоидных клеток синтетическим лигандом, связывающимся с одноцепочечной G4 структурой, увеличивает формирование канонической mRNA p53 и уменьшает изоформу с альтернативным сплайсингом. Эти результаты указыва-

ют на то, что структуры G4 в интроне 3 регулируют сплайсинг интрона 2, приводя к дифференциальной экспрессии транскриптов, кодирующих различные изоформы белка p53 [22].

Известно, что хромосомные транслокации типичны для большинства онкологических заболеваний, в случае лимфом и миелом они тесно связаны с созреванием генов иммуноглобулинов и T рецепторов. Получены данные, свидетельствующие о том, что в таких транслокациях при образовании DBS непосредственное участие принимают квадруплексы [25]. Выявлена тесная ассоциация между локализацией мотивов, способных к формированию G-квадруплексов и участками повышенной ломкости в около 70% генов, участвующих в перестройках в лимфоидных опухолях [19; 24].

В популяциях опухолевых клеток человека относительно часто наблюдаются рекомбинации Alu-повторов в генах супрессорах опухолевого роста [9]. Сиквенс этих генов в семействах с повышенным риском опухолей молочной железы позволил обнаружить потенциальную роль геликазы Pif1 в предрасположенности к развитию опухолей. Особенностью этой геликазы является ее высокое сродство к последовательностям G4 квадруплексов [9]. Предполагается, что Pif1 участвует в рекомбинациях, опосредуемых Alu повторами, и полиморфизм по этому гену может вносить свой вклад в предрасположенность к онкологическим заболеваниям.

Появление новых химерных генов, образующихся в результате межгенных обменов и часто используемых как биомаркеры опухолевого процесса, в ряде случаев тесно связаны с формированием G4 квадруплексов. Такой квадруплекс обнаруживается в последовательности протоонкогена Ret в том участке, где наиболее часто наблюдаются разрывы при слиянии с фрагментом гена Kif5B и образовании нуклеотидной последовательности химерного белка [16].

Показана важная регуляторная роль G квадруплекса, присутствующего в области гиперчувствительности к нуклеазам (nuclease hypersensitive element—NHE III) в промоторе гена человека c-мус (h c-мус), в регуляции транскрипции этого гена. Обнаружено, что ядерная поли(АДФ-рибоза) полимеразы 1 человека (h PARP-1) участвует в регуляции экспрессии h c-мус гена через взаимодействие этого белка с G квадруплексом [12]. Важно подчеркнуть, что сам фермент h PARP-1 является одним из ключевых ферментов репарации одноцепочечных разрывов ДНК [23], т. е. участие во влиянии на транскрипцию G квадруплекса в промоторе гена c-мус реализуется через ключевой фермент репарации одноцепочечных разрывов.

Исследуется роль G-квадруплексов в развитии других болезней человека, таких, например, как экспансия триплета CGG. Такие повторы предрасположены к формированию G-квадруплексов, их размножение в гене FMR-1 ассоциировано с нейродегенеративными болезнями, связано с повышенной ломкостью хромосомы X и рядом других заболеваний [24].

В популяционно-генетических исследованиях человека и набора клеточных линий опухолей человека обнаружено, что мононуклеотидные замены (SNP), нарушающие структуру G4 квадруплекса в промоторах ряда структурных генов, сопровождаются изменением их экспрессии. Авторы приходят к заключению о том, что G4 квадруплексы и нарушение таких структур могут быть прямыми механическими причинами популяционной изменчивости в экспрессии генов [7].

Таким образом, накопленные к настоящему времени данные свидетельствуют о широком участии G квадруплексов в процессах поддержания стабильности/нестабильности генома.

Высокополиморфные геномные участки сельскохозяйственных видов животных

Современное развитие генных и геномных методов привело к возможности перехода от контроля аллельных вариантов немногочисленного количества генов или других геномных элементов к полилокусному геномному сканированию. Геномное сканирование (полилокусное генотипирование) может варьировать от использования нескольких десятков или сотен маркеров до истинного геномного сканирования путем полного секвенирования геномов [21]. Особое значение геномное сканирование приобретает в исследованиях генофондов сельскохозяйственных видов, поскольку только наличие достаточного количества геномных молекулярно-генетических маркеров позволяет сохранять, контролировать и усовершенствовать необходимое биоразнообразие таких видов. В последние несколько десятилетий с этой целью используют генотипирование «анонимных» фрагментов ДНК, фланкированных инвертированным повтором либо одного из случайно выбранных декануклеотидов (Random Amplified Polymorphic DNA — RAPD), фрагмента микросателлита (Inter-Simple Sequence Repeat — ISSR), или фланга ретротранспозона (Inter-Retrotransposon Amplified Polymorphism — IRAP) [5]. Каждый из перечисленных методов позволяет получать в полимеразной цепной реакции (PCR) спектр продуктов амплификации, совокупность которых отражает полиморфизм нескольких десятков или сотен геномных участков, в которых локализованы соответствующие инвертированные повторы, и достаточно надежно устанавливать особенности генетической структуры разных групп животных [14]. В то же время общим недостатком методов такого геномного сканирования является то, что остается недостаточно исследованным информационное содержание таких «анонимных» фрагментов ДНК и источники их высокой изменчивости, позволяющие надежно отличать при их использовании даже близкородственные группы животных. Накопление данных об источниках изменчивости фрагментов ДНК, фланкированных инвертированными повторами, может позволить разрабатывать методы увеличения/уменьшения геномной стабильности сельскохозяйственных видов животных.

Ранее нами были получены данные, свидетельствующие о том, что одним из источников высокого полиморфизма ISSR-PCR маркеров может быть их формирование в результате рекомбинационных событий между мобильными генетическими элементами [3]. Следует отметить, что в секвенированных последовательностях геномов крупного рогатого скота и лошадей мобильные генетические элементы, в частности с длинными концевыми повторами (LTR), встречаются на порядок чаще, чем микросателлитные локусы [11;31], их полиморфизм является видоспецифичным [13].

Для того чтобы оценить возможность вовлечения в формирование высокополиморфных спектров ISSR-PCR маркеров неканонических структур ДНК, в настоящей работе выполнены исследования локализации в секвенированных фрагментах генома лошади и крупного рогатого скота продуктов амплификации, полученных в результате полимеразной цепной реакции (PCR) с использованием в качестве праймера микросателлитного повтора (AG)_nC.

Исследования распределения квадруплексов и шпилечных структур в секвенированных последовательностях ISSR-PCR маркеров

Для исследований использовали результаты секвенирования продукта амплификации спектра ISSR-PCR маркеров генома лошади длиной в 416 нуклеотидов [3] и фрагмента длиной в 1920 нуклеотидов спектра ISSR-PCR маркеров крупного рогатого скота [4].

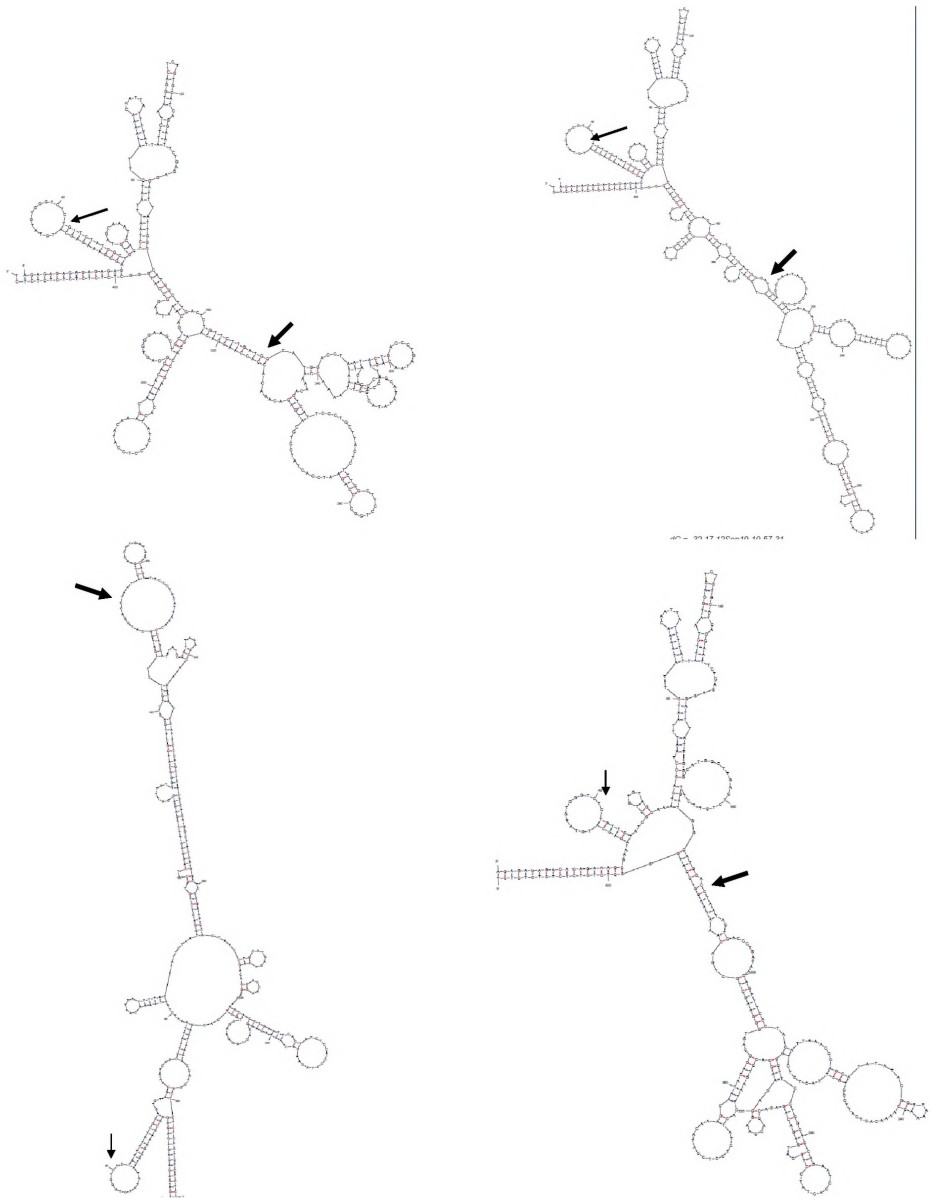


Рис. 3. Примеры (из 11 возможных) предрасположенности к формированию шпичечных структур в секвенированном фрагменте генома лошади длиной в 416 нуклеотидов: стрелками указаны нуклеотидные последовательности, локализованные между участками гомологии к фрагментам ДНК эндогенных ретровирусов (A — последовательность 39-46; —последовательность 166-182)

(см. рис. 3), локализована в том участке, разнообразие упаковки которого вносит существенный вклад в наличие 11 возможных вариантов фолдинга этого участка, т. е. обладает повышенной вероятностью лабильности вторичной структуры ДНК.

5'AGAGAGAGAGAGCGACATGGGGAGACCAGTATTTTCGAGAAACTGATCCCAATTCT
TTATTTTCCAGAGTCTGTTTTATACACTAAGATGTTATACAAAAGTCACGTGGGGA
CAGCAGTCTGCTGATTTTAAAGTCAGGTGCTTCACACAAAATGTATACAGAGGTCT
TAGG**GGTGTACATCATCTTCTGGCCAGGGGGC**CCTGCTGACAATTTACGACCCTC
TCCTTGTGACAGCGGTCAAGTCAATCAGGACTTATTTCTCCAGGCTGATTTATCTCA
AAACAGACGCCACCAAGTAAAGTTACACTCCTACA**GGTGAAGGGTGTAGTGGGT**
TTAGTTAAGGAAAAGAATTAAGTACCTAAGTCTAACGTGATTAATATCAAAGGTC
AATCTATATATTCAATAATATGTGTAAG**GGCAGGGGATATGG**AGACTTAGCAACA
AASACTGGCTCAGCAAATGAAAACCCCTCACCAACACAATTTCTAAATAGCCCAT
ATACTTATACTTAATGTTTCTAAACCTTTCTAAGGAACCTGTTTTAGAGGTTTTAAA
GCATCTCGTGCCTCTCATGGTTGGGAGGCTGTGAGCAATCACATGT**GGCCGACGA**
GCCTGTCAAGCAGGCCAGAGAACCTTCAGAGGAGTTTGTAGGTTAAAACACTCTTG
TCACACAAAGAGTTTTATTGACTGAAGCTCTAGGTTAACTCCTTCTCCGAAAG**AG**
CTGGTGGGGACAGCCCCCGTAAAGTCAGAGGTGTAGGTAAGAGCACAAAGTAG
TAAAGTC**GGCAGGCTCTGCTTATGG**GGGTAGACGCTCGA**GGATTCCAGGGGGAG**
TCTGAGGCTCGATCCCGCTTTGCGTATGTCGAGCCTCCTTCTCATGACCTTTGCC
ATGGGCAGAGTACCTCACTCTGCCCCAACAGGTGTGGTCTCTTAAATGTTGAAAA
ACSSAAATAATCGCCCTCAATTCAGCTCTGGGCCGATGTCTCTCAGTTACCTGA
ACATGGGATTTTCCATCAATAAATGACTGTGACTGCTTACCATTAGAGGAACCATCTGTG
AAAACAAAATTCGCTCTTCCAGAGGTGAATAGAAAATTTCTTTGGAAAATCACT
GGATGTGTTTTCAAAAAATGCAGTAGGGGGCTTGAGGGCAAATGAAAACAAAATTT
GCCCTGAAATCTACAGGGCAACTGGCAATCATCACTATTTTGTAAAAGAAAATTCG
AGTTGATCCTTATGATTGGAATCACTATATTTGCTACTCTTTTCCAAAAATTTCCA
CACTCTTTTCTATCTTTATAATTAATTTGTGCCACCAAGCTGGGATAAGATAAAAAT
ATTTTTCTT**GGGTCACTGGTAGATGG**AGCCATTCCAATGGGCCTTCTTCCACAGC
ATCCTATAGGTGTATGTTCTGAGGCAAGAATAATTAATCCCATCCTCTGGTAATAT
TAACSTAATTAAGTATCAATTAAGCCCTCATCCACTTAAACAAGAGCCAACTCTG
CCTCACTAACCAACTTTCTCTTGAAGTGGGATAGGATCGCTTTTAAAGCAATCAA
ACAAAAGCTTTAAATCTGCAGTAGTCACTTTTAAATGTGGGTGATCCAATTAATAT
CCCTAATAAATTTCTGGAATCTTTTGTTA**GGTAGGTAAGTGAAGGAAAAGG**AGT
CCAAAATGGTGGTGGCTAAAAGACAA**GGAGGTCAAAGGAAGG**TCCGAGGACCA
GAGTGAAGACTTCAGGCAGAACAAACAGCAGCCCTGGCTAAGCCCAATTTGCATAG
GGCAGGCTAGGTGGAAAGAAAACATATAAAAGGAGGAGCCAAAACACTTTCTC
TTGGACTCTCTCTCCATGTGCGTGGCTCTTTCTCTCACTCTCTCTCTC 3'

Рис. 4. Локализация квадруплекса в фрагменте ДНК длиной в 1921 нуклеотид спектра продуктов амплификации, полученных с использованием в рСг в качестве праймера последовательности (AG)₅C с геномной ДНК крупного рогатого скота: последовательность квадруплекса выделена цветом, пары гуанинов — подчеркиванием

Gig1 в секвенированной последовательности генома крупного рогатого скота длиной в 1921 нуклеотид выполнен поиск повторов, в результате получены следующие данные. С 5 нуклеотида по 938 выделенный участок имеет гомологию в 96% с последовательностью видо специфично го для крупного рогатого скота эндогенного ретровируса BTLTR1B (ERV2, [18]; с 976 по 1674 нуклеотид — с более чем 70% гомологией к последовательности эндогенного ретровируса ERV2-1_{TSy-I}, впервые описанного у вида обезьян *Tarsius tarsier* [17], а с позиции 1675 по 1921 в обратном направлении, с 1 по 256 нуклеотид, с 92% гомологией — к последовательности эндогенного ретровируса крупного рогатого скота ERV1-2-LTRBT [20] (ERV1, рис. 5).

Сравнение распределения последовательностей, потенциально способных к формированию квадруплексов (см. рис. 4), и участков гомологии к мобильным генетическим элементам (см. рис. 5) свидетельствует о предпочтительной локализации потенциальных квадруплексов (с высоким уровнем перекрытия) близко к флангам рекомбинирующих фрагментов эндогенных ретровирусов.

Выполненный анализ вариантов упаковки нуклеотидной последовательности длиной в 1921 нуклеотид генома крупного рогатого скота показал наличие большого количества инвертированных повторов, что приводит к потенциальным возможностям формирования 22 вариантов упаковки этого участка. Причем существенный

В секвенированной последовательности фрагмента ДНК генома крупного рогатого скота длиной в 1921 нуклеотид выявлено 11 участков локализации потенциальных квадруплексов (без перекрытия) (рис. 4).

Если учитывать варианты перекрытия друг с другом последовательностей, потенциально способных формировать квадруплексы, их оказывается существенно больше (см. таблицу).

По данным, представленным в таблице, видно, что наиболее высокая плотность перекрывающихся последовательностей, потенциально способных к формированию квадруплексов, локализована в участках между 168 и 206, 801 и 857 и между 1680 и 1755 нуклеотидами.

С использованием

программ RepeatMasker и

**Локализация последовательностей, потенциально способных
к формированию квадруплексов с учетом их перекрытия, в секвенированном
участке ДНК длиной в 1921 нуклеотид генома крупного рогатого скота**

Позиция	Длина	Последовательность потенциальных квадруплексов	G-Score
168	29	<u>GGTCTTAGGGGTGTTACATCATCTTCTGG</u>	5
175	27	<u>GGGGTGTTACATCATCTTCTGGCCAGG</u>	5
175	28	<u>GGGGTGTTACATCATCTTCTGGCCAGGG</u>	5
175	29	<u>GGGGTGTTACATCATCTTCTGGCCAGGGG</u>	5
175	29	<u>GGGGTGTTACATCATCTTCTGGCCAGGGG</u>	3
175	30	<u>GGGGTGTTACATCATCTTCTGGCCAGGGGG</u>	5
175	30	<u>GGGGTGTTACATCATCTTCTGGCCAGGGGG</u>	0
175	30	<u>GGGGTGTTACATCATCTTCTGGCCAGGGGG</u>	4
175	30	<u>GGGGTGTTACATCATCTTCTGGCCAGGGGG</u>	3
176	28	<u>GGGTGTTACATCATCTTCTGGCCAGGGG</u>	4
176	29	<u>GGGTGTTACATCATCTTCTGGCCAGGGGG</u>	5
176	29	<u>GGGTGTTACATCATCTTCTGGCCAGGGGG</u>	4
176	30	<u>GGGTGTTACATCATCTTCTGGCCAGGGGGG</u>	6
176	30	<u>GGGTGTTACATCATCTTCTGGCCAGGGGGG</u>	5
176	30	<u>GGGTGTTACATCATCTTCTGGCCAGGGGGG</u>	4
177	27	<u>GGTGTTACATCATCTTCTGGCCAGGGG</u>	5
177	28	<u>GGTGTTACATCATCTTCTGGCCAGGGGG</u>	6
177	28	<u>GGTGTTACATCATCTTCTGGCCAGGGGG</u>	5
177	29	<u>GGTGTTACATCATCTTCTGGCCAGGGGGG</u>	7
177	29	<u>GGTGTTACATCATCTTCTGGCCAGGGGGG</u>	6
177	29	<u>GGTGTTACATCATCTTCTGGCCAGGGGGG</u>	5
<i>21 перекрывающийся участок</i>			
322	30	<u>GGGTGAGGGTGTAGTGGGIIIAGTTAAGG</u>	14
322	30	<u>GGGTGAGGGTGTAGTGGGIIIAGTTAAGG</u>	15
322	30	<u>GGGTGAGGGTGTAGTGGGMMAGTTAAGG</u>	15
322	30	<u>GGGTGAGGGTGTAGTGGG IIIAGTTAAGG</u>	16
323	29	<u>GGTGAGGGTGTAGTGGGIIIAGTTAAGG</u>	13
323	29	<u>GGTGAGGGTGTAGTGGGIIIAGTTAAGG</u>	14
323	29	<u>GGTGAGGGTGTAGTGGGIIIAGTTAAGG</u>	14
323	29	<u>GGTGAGGGTGTAGTGGGIIIAGTTAAGG</u>	15
427	15	<u>GGGCAGGGGATAT GG</u>	17
428	14	<u>GGCAGGGGATATGG</u>	17
588	30	GGTTGGGAGGCTGTGAGCAATCACATGTGG	5
588	30	GGTTGGGAGGCTGTGAGCAATCACATGTGG	4
592	30	GGGAGGCTGTGAGCAATCACATGTGGCCGG	5
593	29	GGAGGCTGTGAGCAATCACATGTGGCCGG	4
616	24	GGCCGGACGAGCCTGTCAAGCAGG	11
716	30	GGTTAACTCCTTCTCCGAAAGAGGTGGTGG	2
738	10	<u>GGTGGTGGGG</u>	20
738	11	<u>GGTGGTGGGGG</u>	21
738	11	<u>GGTGGTGGGGG</u>	19
801	18	GGCAGGCTCTGGTTAT GG	19
801	19	GGCAGGCTCTGGTTATGGG	18
801	20	GGCAGGCTCTGGTTATGGGG	17
801	20	<u>GGCAGGCTCTGGTTATGGGG</u>	11
801	20	<u>GGCAGGCTCTGGTTATGGGG</u>	13
801	21	GGCAGGCTCTGGTTATGGGGG	16

Позиция	Длина	Последовательность потенциальных квадруплексов	G-Score
801	21	GGCAGGCTCTGGTTATGGGGG	12
801	21	GGCAGGCTCTGGTTATGGGGG	10
801	21	GGCAGGCTCTGGTTATGGGGG	14
801	21	GGCAGGCTCTGGTTATGGGGG	13
805	16	GGCTCTGGTTATGGGGG	17
805	17	GGCTCTGGTTATGGGGG	18
805	17	GGCTCTGGTTATGGGGG	16
805	30	GGCTCTGGTTATGGGGGTAGACGCTCGAGG	11
805	30	GGCTCTGGTTATGGGGGTAGACGCTCGAGG	12
805	30	GGCTCTGGTTATGGGGGTAGACGCTCGAGG	13
805	30	GGCTCTGGTTATGGGGGTAGACGCTCGAGG	14
805	30	GGCTCTGGTTATGGGGGTAGACGCTCGAGG	9
805	30	GGCTCTGGTTATGGGGGTAGACGCTCGAGG	11
805	30	GGCTCTGGTTATGGGGGTAGACGCTCGAGG	10
811	24	GGTTATGGGGGTAGACGCTCGAGG	9
811	24	GGTTATGGGGGTAGACGCTCGAGG	11
811	24	GGTTATGGGGGTAGACGCTCGAGG	10
817	27	GGGGGTAGACGCTCGAGGATTCCAGG	9
817	27	GGGGGTAGACGCTCGAGGATTCCAGG	11
817	28	GGGGGTAGACGCTCGAGGATTCCAGGG	9
817	28	GGGGGTAGACGCTCGAGGATTCCAGGG	11
817	29	GGGGGTAGACGCTCGAGGATTCCAGGGG	9
817	29	GGGGGTAGACGCTCGAGGATTCCAGGGG	11
817	29	GGGGGTAGACGCTCGAGGATTCCAGGGG	1
817	29	GGGGGTAGACGCTCGAGGATTCCAGGGG	7
817	30	GGGGGTAGACGCTCGAGGATTCCAGGGGG	9
817	30	GGGGGTAGACGCTCGAGGATTCCAGGGGG	0
817	30	GGGGGTAGACGCTCGAGGATTCCAGGGGG	11
817	30	GGGGGTAGACGCTCGAGGATTCCAGGGGG	2
817	30	GGGGGTAGACGCTCGAGGATTCCAGGGGG	0
817	30	GGGGGTAGACGCTCGAGGATTCCAGGGGG	8
817	30	GGGGGTAGACGCTCGAGGATTCCAGGGGG	7
818	26	GGGGTAGACGCTCGAGGATTCCAGG	10
818	27	GGGGTAGACGCTCGAGGATTCCAGGG	10
818	28	GGGGTAGACGCTCGAGGATTCCAGGGG	10
818	28	GGGGTAGACGCTCGAGGATTCCAGGGG	8
818	29	GGGGTAGACGCTCGAGGATTCCAGGGGG	10
818	29	GGGGTAGACGCTCGAGGATTCCAGGGGG	1
818	29	GGGGTAGACGCTCGAGGATTCCAGGGGG	9
818	29	GGGGTAGACGCTCGAGGATTCCAGGGGG	8
819	27	GGGTAGACGCTCGAGGATTCCAGGGG	9
819	28	GGGTAGACGCTCGAGGATTCCAGGGGG	10
819	28	GGGTAGACGCTCGAGGATTCCAGGGGG	9
820	26	GGTAGACGCTCGAGGATTCCAGGGG	10
820	27	GGTAGACGCTCGAGGATTCCAGGGGG	11
820	27	GGTAGACGCTCGAGGATTCCAGGGGG	10
833	24	GGATTCCAGGGGGACTCCTGAGG	12
833	24	GGATTCCAGGGGGACTCCTGAGG	14
833	24	GGATTCCAGGGGGACTCCTGAGG	13

55 перекрывающихся
участков

1373	18	GGGGTCACTGGTAGATGG	16
1375	30	GGTCACTGGTAGATGGAGCCATTCCAATGG	14

Позиция	Длина	Последовательность потенциальных квадруплексов	G-Score
1680	22	GGTAGGTAGAATAGGGAAAAGG	16
1680	22	GGTAGGTAGAATAGGGAAAAGG	15
1684	30	GGTAGAATA GG GAAAA GGAGTCCAAAAT GG	16
1684	30	GGTAGAATAGGGAAAAGGAGTCCAAAATGG	15
1693	24	GGGAAAAGGAGTCCAAAATGGTGG	12
1693	27	GGGAAAAGGAGTCCAAAATGGTGGTGG	15
1693	27	GGGAAAAGGAGTCCAAAATGGTGGTGG	9
1693	27	GGGAAAAGGAGTCCAAAATGGTGGTGG	5
1694	23	GGAAAAGGAGTCCAAAATGGTGG	12
1694	26	GGAAAAGGAGTCCAAAATGGTGGTGG	15
1694	26	GGAAAAGGAGTCCAAAATGGTGGTGG	9
1694	26	GGAAAAGGAGTCCAAAATGGTGGTGG	6
1700	20	GGAGTCCAAAATGGTGGTGG	12
1712	21	GGTGGTGGCTAAAAGACAAGG	11
1712	25	GGTGGTGGCTAAAAGACAAGGAAGG	7
1712	25	GGTGGTGGCTAAAAGACAAGGAAGG	8
1712	25	GGTGGTGGCTAAAAGACAAGGAAGG	12
1715	22	GGTGGCTAAAAGACAAGGAAGG	11
1715	29	GGTGGCTAAAAGACAAGGAAGGTCAAAGG	11
1715	29	GGTGGCTAAAAGACAAGGAAGGTCAAAGG	7
1715	29	GGTGGCTAAAAGACAAGGAAGGTCAAAGG	9
1718	26	GGCTAAAAGACAAGGAAGGTCAAAGG	12
1718	30	GGCTAAAAGACAAGGAAGGTCAAAGGAAGG	12
1718	30	GGCTAAAAGACAAGGAAGGTCAAAGGAAGG	12
1718	30	GGCTAAAAGACAAGGAAGGTCAAAGGAAGG	8
1731	17	GGAAGGTCAAAGGAAGG	18
1731	24	GGAAGGTCAAAGGAAGGTCCGAGG	14
1731	24	GGAAGGTCAAAGGAAGGTCCGAGG	14
1731	24	GGAAGGTCAAAGGAAGGTCCGAGG	14
1735	20	GGTCAAAGGAAGGTCCGAGG	18
<i>30 перекрывающихся</i>			
<i>участков</i>			
1814	16	GGGCAGGCCTAGGTGG	18
1815	15	GGCAGGCCTAGGTGG	18
1825	28	GGTGGAAGAAAAACATATAAAAAGGAGG	4

вклад в лабильность такой упаковки вносили последовательности нуклеотидов, локализованные между участками гомологии к фрагментам рекомбинирующих ретровирусов (см. рис. 5, нуклеотидные последовательности между 938 — 976 и между 1674 и 1675 позициями нуклеотидов). Примеры вариантов упаковок в этих участках представлены на рисунке 6, а, б.

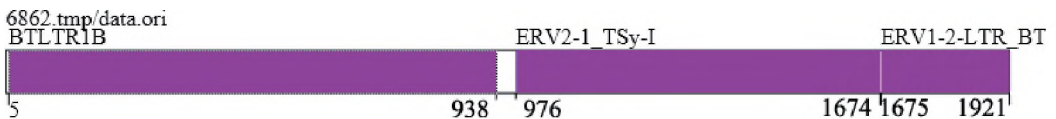


Рис. 5. Распределение участков гомологии к фрагментам эндогенных ретровирусов в полном размере фрагменте нуклеотидных последовательностей участка хромосомы 15 крупного рогатого скота (GenBank: NC 007313.5)

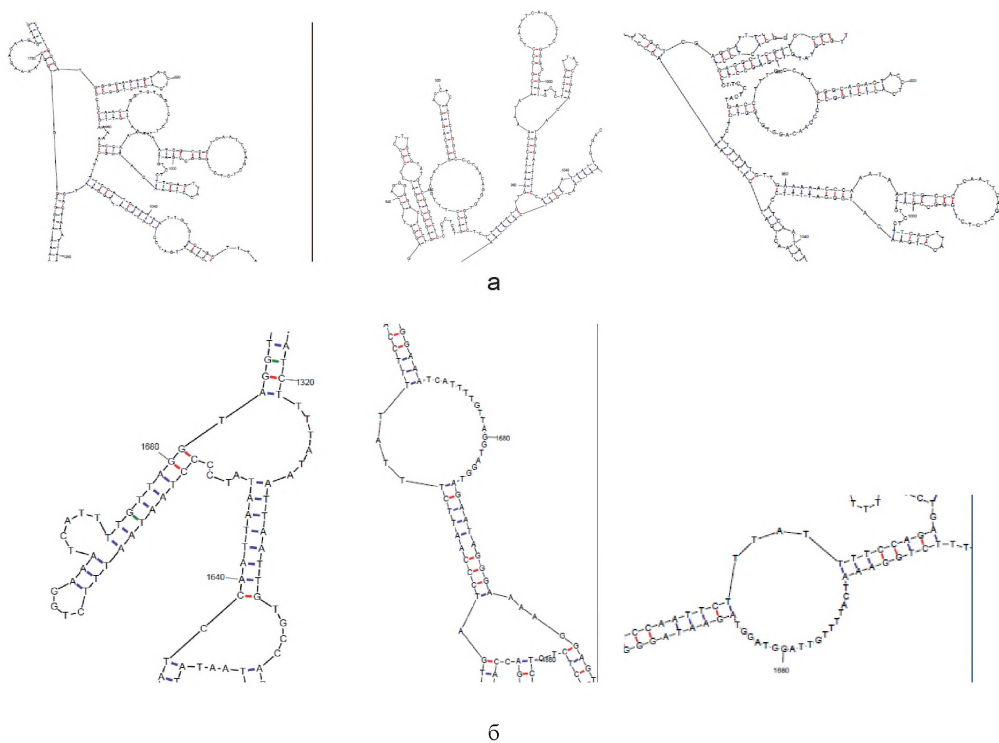


Рис. 6. Примеры вариантов упаковки последовательностей нуклеотидов, локализованных между участками гомологии к различным эндогенным ретровирусам в фрагменте длиной в 1921 нуклеотид генома крупного рогатого скота: а — варианты упаковки ДНК, включающие последовательность от 938 до 976 нуклеотида; б — варианты упаковки ДНК, включающие последовательность от 1674 до 1675 нуклеотида

Следует отметить, что, по литературным данным, встречаемость последовательностей, потенциально способных к образованию квадруплексов, обнаруживается в ядерном и митохондриальном геноме, в частности дрожжей, с существенно меньшей частотой, чем в исследованных нами секвенированных фрагментах геномной ДНК домашней лошади и крупного рогатого скота, представленными участками, гомологичными комбинациям фрагментов эндогенных ретровирусов. Так, у *Saccharomyces cerevisiae* мтДНК содержит на порядок больше G4 DNA на одну тысячу пар нуклеотидов (0,373), чем любая ядерная хромосома этого же вида (0,034-0,067) [8]. У человека в ядерном геноме на одну тысячу нуклеотидов встречается примерно 0,1 потенциальной последовательности G4 (-376000 G4 на 3×10^9 нуклеотидов гаплоидного генома человека [27]). В то же время в секвенированном фрагменте ДНК генома лошади последовательность нуклеотидов, потенциально способная к формированию квадруплекса, встречается 1 раз на 416 нуклеотидов (см. рис. 1), т. е. условно около трех последовательностей на тысячу пар нуклеотидов. В случае секвенированной последовательности фрагмента генома крупного рогатого скота даже без учета перекрытия потенциальных G4 участков (см. таблица), такие последовательности встречаются приблизительно с частотой 5 потенциальных G4 на

одну тысячу нуклеотидов (рис. 4), что на порядок больше, чем в мтДНК, и на два порядка — в ядерной ДНК *S. cerevisiae*. Следует отметить также, что потенциальные G4 в наших исследованиях локализованы близко к тем участкам, которые отличаются повышенной вероятностью формирования различных вариантов шпилечных структур (см. рис. 3,6).

Полученные данные свидетельствуют о том, что в высокополиморфных участках ДНК, фланкируемых инвертированными повторами и состоящими из последовательностей, гомологичных рекомбинантам между разными эндогенными ретровирусами, обнаруживается относительно повышенная частота встречаемости нуклеотидных последовательностей, способных формировать квадруплексы. Причем наблюдается смещение их локализации к флангам участков гомологии к эндогенным ретровирусам (местам рекомбинаций между ними), для которых характерно наличие нуклеотидных последовательностей, предрасположенных к формированию лабильных вариантов шпилечных структур.

Направления поисков геномных стабилизаторов

Последовательности, потенциально способные к формированию квадруплетов, часто предрасположены к формированию шпилечных структур [6]. Районы, вовлеченные в образование шпилек и одновременно несущие последовательности, способные формировать G4 структуры, можно обнаружить и в участках на флангах фрагментов эндогенных ретровирусов, входящих в исследованные нами нуклеотидные последовательности длиной в 416 нуклеотидов генома домашней лошади и длиной в 1921 нуклеотид генома крупного рогатого скота (см. рис. 3, 6), образующие разные варианты фолдинга.

Для раскручивания структур G4 и шпилек необходимо участие разных геликаз. Геликаза 5-3' Pif1 способствует лиенизации G4 структур, а геликаза 3-5' Sgs1 дрожжей и ее гомолог у человека BLM раскручивает шпилечные структуры [6]. В то же время в некоторых исследованиях отмечается их определенная дополняющая функция [27], что, по-видимому, может быть обусловлено возможностями перехода из одного варианта вторичной структуры ДНК в другой.

В работах [8] анализ распределения G4 DNA по ядерному геному показал прочную ассоциацию мотивов ДНК G4 с сайтами частых двуцепочечных разрывов ДНК (DSBs) в митотических и мейотических клетках. У человека выявлено около 25 618 «горячих» сайтов мейотических рекомбинаций, которые существенно обогащены потенциальными участками формирования ДНК G4 квадруплетов [30]. Предполагается, что существенное перекрытие «горячих» сайтов DSB в митозе и мейозе обусловлено присутствием структур G4 с общими функциями в обоих типах клеток.

Присутствие структур ДНК G4 в сайтах DSB позволяет предполагать наличие двух функций, не исключających друг друга. Во-первых, формирование структур G4 во время репликации ДНК может увеличивать вероятность DSBs. Образование их во время репликации, когда ДНК временно одноцепочечная, может замедлять или даже останавливать прогрессию вилки репликации, приводя к образованию DSBs. Известно, что от дрожжей до человека теломерная ДНК препятствует прогрессии вилки репликации во время стандартного полуконсервативного синтеза ДНК. К тому же мутации в ДНК геликазах, которые раскручивают G4, приводят к хромосомным делениям в местах локализации G4.

Во-вторых, мотивы ДНК G4 могут быть связаны с DSBs потому, что они выполняют особую роль в формировании DSB. Один из начальных шагов в репариро-

вании DSB разрыва заключается в появлении одиночных хвостов с 3' концов. Формирование структуры ДНК G4 в этих хвостах могут способствовать образованию повреждений ДНК и влиять на расположенные вниз по течению события. Действительно, некоторые репарирующие и рекомбинирующие белки *S. cerevisiae* имеют активность, меняющую ДНК G4 структуры.

Важно подчеркнуть, что нуклеотидная последовательность геликазы Pif1 *Bos taurus* содержит большое количество последовательностей, способных к образованию G4 структур — на 2461 нуклеотидов этого гена выявляется 28 потенциальных неперекрывающихся G4 структур (814 перекрывающихся, собственные данные). Эта GC богатая последовательность предрасположена также к формированию множества коротких шпилек. Таким образом, факторы, способствующие стабилизации участков потенциального образования квадруплексов, препятствуют не только накоплению DSBs в геномной ДНК в общем, но и увеличению частот мутаций в гене Pif1, ключевой геликазе разворачивания G4. Такой комплексный эффект может вносить существенный вклад в стабилизацию генома. Следует отметить, что в последние годы появилось много работ, в которых подбираются и подробно исследуются способы стабилизации потенциальных G4 последовательностей (например, [10]).

Таким образом, накопленные данные свидетельствуют о том, что именно квадруплексы могут быть одной из ключевых мишеней воздействия в целях разработки методов увеличения/уменьшения геномной изменчивости в зависимости от целей и стадиями селекционной работы с различными группами организмов, поскольку потенциальные G4 структуры не только оказываются тесно связанными с DSBs — общей характеристикой дестабилизации генетического аппарата, но и могут находиться в конкурентных взаимоотношениях с другими неканоническими структурами ДНК (шпилечными, например), а также с кодирующими последовательностями геликазы Pif1 — ключевого фермента стабилизации репликации.

Библиографический список

1. Вавилов Н.И. Закон гомологических рядов в наследственной изменчивости. Л: Наука, 1987. 256 с.
2. Вавилов Н.И. Теоретические основы селекции. М: Наука, 1987. 510 с.
3. Глазко В.И., Феофилов А.В., Бардуков Н.В., Глазко Т. Т. Видоспецифичные ISSR-PCR маркеры и пути их формирования // Известия ТСХА. 2012. №1. С. 118-125.
4. Глазко В.И., Феофилов А.В., Бардуков Н.В., Глазко Т.Т. Микросателлиты и эндогенные ретровирусы в геномном «ландшафте» высших млекопитающих // Российские нанотехнологии. — В печати.
5. Календарь Р.Н., Глазко В.И. Типы молекулярно-генетических маркеров и их применение // Физиология и биохимия культурных растений. 2002. Т. 34. № 4. С. 279-296.
6. Anand R.P., Shah K.A., Niu P., Sung P., Mirkin S.M., Freudenreich C.H. Overcoming natural replication barriers: differential helicase requirements // Nucleic Acids Research. 2012. Vol. 40. No. 3. P. 1091-1105.
7. Baral A., Kumar P., Haider R., Mani P. Quadruplex-single nucleotide polymorphisms (Quad-SNP) influence gene expression difference among individuals // Nucleic Acids Research. 2012. Vol. 40. No. 9. P. 3800-3811.
8. Capra J.A., Paeschke K, Singh M., Zakian V.A. G-Quadruplex DNA Sequences Are Evolutionarily Conserved and Associated with Distinct Genomic Features in *Saccharomyces cerevisiae* // PLoS Comput Biol. 2010. No. 6(7). P. 1-13.
9. Chisholm K.M., Aubert S.D., Freese K.P., Zakian V.A., King M.-C. A Genomewide Screen for Suppressors of Alu-Mediated Rearrangements Reveals a Role for PIF1 // PLoS ONE. 2012. Vol. 7. No. 2. e30748. 10 p.

10. Collie G.W., Parkinson G.N. The application of DNA and RNA G-quadruplexes to therapeutic medicines // *Chem. Soc. Rev.* 2011. №. 40. P. 5867-5892.
11. Elsik C.G., Tellam R.L., Worley K.C. The Genome Sequence of Taurine Cattle: A Window to Ruminant Biology and Evolution // *Science.* 2009. Vol. 324. №. 5926. P. 522-528.
12. Fekete A., Kenesi E., Hunyadi-Gulyas E., Durgo H., Berko B. The Guanine-Quadruplex Structure in the Human c-myc Gene's Promoter Is Converted into B-DNA Form by the Human Poly(ADP-Ribose) Polymerase-1 // *PLoS ONE.* 2012. Vol. 7. №. 8. e42690.
13. Garcia-Etxebarria K., Jugo B.M. Genome-Wide Detection and Characterization of Endogenous Retroviruses in Bos Taurus Hi. *Virology.* 2010. Vol. 84. №. 20. P. 10852-10862.
14. Glazko V.I., Bardukov N.V., Pheophilov A.V., Sipko T.P., Elkina M.A., Glazko T.T. Polymorphism of ISSR and IRAP markers in genomes of musk-oxen (*Ovibos moschatus*) and horse (*Equus caballus*) of Altai breed // *Izvestia of Timiryazev Agricultural Academy.* 2012. Special Issue P. 16-26.
15. Greaves M., Maley C.C. Clonal evolution in cancer // *Nature.* 2012. Vol. 481. №. 7381. P. 287-294.
16. Ji Y.S., Lee W.-C., Shin J.-Y., Lee S., Bleazard T. A transforming KIF5B and RET gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing // *Genome Research.* 2012. Vol. 22. P. 436-445.
17. Jurka J. Endogenous retroviruses from tarsier // *Rebase Reports.* 2010. №. 10(9). P. 1202-1202.
18. Jurka J. Long terminal repeats from domestic cow // *Rebase Reports.* 2008. №. 8(8). P. 847-847.
19. Kadapadi V.K., Nambiar M., Raghavan S.C. Potential G-quadruplex formation at breakpoint regions of chromosomal translocations in cancer may explain their fragility // *Genomics.* 2012. Vol. 100. №. 2. P. 72-80.
20. Kohany O., Jurka J. LTR retrotransposons from cow // *Rebase Reports.* 2008. №. 8(6). P. 674-674.
21. Lowry D.B. Landscape evolutionary genomics // *Biol. Lett.* 2010. Vol. 6. P. 502-504.
22. Marcel V., Tran P.L.T., Sagne C., Martel-Planche G., Vaslin L., Teulade-Fichou M.-P., Hall J., Mergny J.-L., Hainaut P., Van Dyck E. G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms // *Carcinogenesis.* 2011. Vol. 32. №. 3. P. 271-278.
23. Masaoka A., Gassman N.R., Kedar P.S., Prasad R., Hou E.W., Horton J.K., Bustin M., Wilson S.H. HMG1 Protein Regulates Poly(ADP-ribose) Polymerase-1 (PARP-1) Self-PARylation in Mouse Fibroblasts // *J. Biol. Chem.* 2012. Vol. 287. P. 27648-27658.
24. Nambiar M., Goldsmith G., Moorthy B. T., Lieber M.R., Joshi M.V., Choudhary B., Hosur R.V., Raghavan S.C. Formation of a G-quadruplex at the BCL2 major breakpoint region of the t(14;18) translocation in follicular lymphoma // *Nucleic Acids Research.* 2011. Vol. 39. №. 3. P. 936-948.
25. Nambiar M., Raghavan S.C. How does DNA break during chromosomal translocations? // *Nucleic Acids Research.* 2011. Vol. 39. №. 14. P. 5813-5825.
26. Schuster-Bockler B., Lehner B. Chromatin organization is a major influence on regional mutation rates in human cancer cells // *Nature.* 2012. Vol. 488. N.7412. P. 504-507.
27. Sharma S. Non-B DNA Secondary Structures and Their Resolution by RecQ Helicases // *Journal of Nucleic Acids.* 2011. Vol. 2011. Art. ID 724215. 15 p.
28. Smit A.F. Identification of a new, abundant superfamily of mammalian LTR-transposons // *Nucleic Acids Res.* 1993. Vol. 21, №. 8. P. 1863-1872.
29. Tsai Y.-C., Qi H., Lin C.-P., Lin R.-K., Kerrigan J.E., Rzuczek S.G., LaVoie E.J., Rice J.E., Pilch D.S., Lyu Y.L., Liu L.F. A G-quadruplex Stabilizer Induces M-phase Cell Cycle Arrest Hi. *Biological Chemistry.* 2009. Vol. 284. №. 34. P. 22535-22543.
30. Uringa E.-J., Youds J.L., Lisaingo K., Lansdorp P.M., Boulton S.J. RTEL1: an essential helicase for telomere maintenance and the regulation of homologous recombination // *Nucleic Acids Research.* 2011. Vol. 39. No. 5 P. 1647-1655.

31. Wade CM., Giulotto E., Sigurdsson S. **Genome Sequence, Comparative Analysis, and Population Genetics of the Domestic Horse** // *Science*. 2009. Vol. 326. №. 5954. P. 865-867.
32. Wong H.M., Stegle O., Rodgers S., Huppert J.L. **A Toolbox for Predicting G-Quadruplex Formation and Stability** // *Journal of Nucleic Acids*. 2010. Vol. 2010. P. 6.

GENOME INSTABILITY AND NON CANONIC DNA STRUCTURES

B.L. ZYBAILOV, VI. GLAZKO

(¹ University of Arkansas for Medical Sciences (UAMS), Little Rock, AR, USA,
² RTSAU named after K.A. Timiryazev, Moscow, Russia)

The sources of genetic instability connected with non canonical DNA structure of DNA (quadruplex, hairpins) and their distribution in sequences of horse and cattle ISSR-PCR markers have been considered. The increased frequency of quadruplex localization in studied DNA fragments presented by homologous fragments to endogen retroviruses is shown. Participation of quadruplexes in processes of stabilization/destabilization of genomic DNA is discussed.

Keywords: ISSR-PCR markers, quadruplex, hairpins, double-stranded break of DNA, recombination, endogen retroviruses, helicase.

Зыбайлов Борис Леонидович — и. с. Медицинского университета Арканзаса (4301 W Markham Little Rock, AR 72205).

Глазко Валерий Иванович — д. с.-х. н., проф., акад. РАСХН (инстр. член), руководитель Центра нанобиотехнологий РГАУ-МСХА имени КА. Тимирязева (125550, г. Москва, ул. Тимирязевская, 49. Тел.: (499) 976-03-75; e-mail: rector@timacad.ru).