

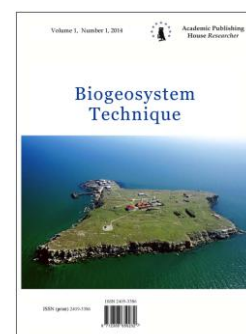
Copyright © 2015 by Academic Publishing House *Researcher*



Published in the Russian Federation  
Biogeosystem Technique  
Has been issued since 2014.  
ISSN: 2409-3386  
Vol. 4, Is. 2, pp. 138-152, 2015

DOI: 10.13187/bgt.2015.4.138

[www.ejournal19.com](http://www.ejournal19.com)



## The Science and the Problems of Development

UDC 575.174.015.3:599.723

### Genomic Scanning Using Inverted Repeats of Microsatellites (GAG)<sub>6</sub>C, (AG)<sub>9</sub>C

<sup>1</sup>Tatiana T. Glazko

<sup>2</sup>Gleb Yu. Kosovskiy

<sup>3</sup>Svetlana N. Kovaltchuk

<sup>4</sup>Boris L. Zybaïlov

<sup>5</sup>Valery I. Glazko

<sup>1, 2, 3, 5</sup> Centre of experimental embryology and reproductive biotechnology, Russian Federation

<sup>1, 5</sup> Russian State Agricultural University – Moscow Timiryazev's Agricultural Academy named after K.A. Timiryazev, Moscow, Russian Federation

<sup>4</sup> University of Arkansas for Medical Sciences (UAMS), USA  
Little Rock, AR

<sup>1</sup> Dr (Agric), Professor

<sup>2</sup> Dr (Biology), Corresponding member of Russian Academy of Natural Sciences, Director of the Centre of experimental embryology and reproductive biotechnology

<sup>3</sup> PhD (Biology), Ass. Professor

<sup>4</sup> PhD (Biology)

<sup>5</sup> Dr (Agric), Professor, Full member of Russian Academy of Natural Sciences, Full member of Russian Academy of Sciences (foreign member)

E-mail: [vigvalery@gmail.com](mailto:vigvalery@gmail.com)

#### Abstract

The sequencing of genomic DNA fragments of holsteinized cattle with a length of about 550 nucleotides, flanked by inverted repeats of microsatellites (GAG)<sub>6</sub>C and (AG)<sub>9</sub>C and search their homology to the reference bovine genome in GenBank database was first performed. In most cases, homologous (GAG)<sub>6</sub>C sequences were localized in immune genes and cellular signaling systems or in their 5' flanks in the intergenic space. The most frequently sequenced of genome regions, flanked by inverted repeat (AG)<sub>9</sub>C met areas of homology with long dispersed nuclear element LINE-1, species specific for *Bos taurus* – L1-BT, or the products of recombination between them. This leads to the conclusion that the microsatellite AG is closely linked with this product of recombination. The data obtained indicate that microsatellites, even with the very similar core motives, such as (GAG)<sub>6</sub>C and (AG)<sub>9</sub>C, differed significantly from each other in association with different genomic elements. The data obtained must be considered when using different microsatellites in order genomic scanning. A possible value for these differences of susceptibility of

the investigated microsatellites to the formation of various non-canonical DNA structures such as the G4 quadruplexes, appearing of molecular (DNA-RNA) triplexes were discussed.

**Keywords:** microsatellites, sequencing, non-canonical structure of DNA, retroviruses, the purine-pyrimidine tracks, triplexes, G4 quadruplexes.

### Введение

Генные и геномные технологии XXI века привели к возможности идентифицировать гены и геномные элементы, вовлеченные в адаптации, участвующие в процессах макро и микроэволюции, позволили перейти к анализу не отдельных генов, а генных сетей. Один из следующих ключевых шагов в эволюционной биологии будет выяснение механизмов влияния генотипической (геномной) среды на фенотипические ответы на давление искусственного и естественного отбора. Наиболее удобной моделью для таких исследований являются животные сельскохозяйственных видов. Примером этому является развитие методов геномной селекции, на которые возлагают особые надежды при работе со сложными количественными признаками, например, в молочном скотоводстве [1].

Геномные сканирования – главное направление и главный признак современной популяционной геномики XXI века. За несколько прошлых лет существенно расширился набор биологических систем, в которых выполнялись исследования геномного разнообразия в разных условиях сред обитания и давления искусственного и естественного отбора. Логика такого геномного сканирования заключается в том, что аллели в нейтральных областях геномов будут свободно перемещаться от популяции к популяции через генный поток, в то время как мишени естественного или искусственного отбора покажут более высокое геномное расхождение в разных средах обитания. Геномное сканирование может варьировать от использования пары сотен маркеров до истинного геномного сканирования путем полного секвенирования геномов. Суть этих методов довольно проста и основана на том, что среди сотен тысяч генотипов по мононуклеотидным заменам (single nucleotide polymorphism - SNP) отбираются полилокусные генотипы, общие, например, для быков с высоким селекционным индексом, в который включена оценка по молочной продуктивности их дочерей, и в дальнейшем используются для прогноза повышенной вероятности селекционной ценности молодых бычков. Такой подход, очевидно, может существенно сократить временной промежуток между поколениями. Этому способствует создание коммерческих ДНК-чипов для множественного генотипирования по SNP. В то же время, использование ДНК-чипов с высокой плотностью SNP нередко вступает в противоречие с экономической целесообразностью, в связи с чем, в частности, в мясном скотоводстве применяют ДНК-чипы с низкой плотностью SNP (3-7K). Относительно низкая экономическая эффективность применения SNP-панелей для прогноза проявления хозяйственно ценных признаков у крупного рогатого скота обусловлена разными причинами, которые могут быть подразделены на три группы. Одна из них связана с тем, что для большинства включаемых в анализ фенотипических признаков более 50 % наследственной изменчивости попадают в геномные области с небольшими фенотипическими эффектами, порядок величин которых соответствует полигенному характеру наследования. Наглядным примером является, в частности, то, что попытки картирования главных генов молочной продуктивности с использованием SNP у трех французских специализированных молочных пород привели к выявлению разных генов, общим для которых являлось только то, что их экспрессия находится под контролем гипофизарно-надпочечниковой оси.

Кроме того, такие характеристики молочной продуктивности как, например, общий удой, имеют, как правило, низкий коэффициент наследуемости и существенно зависят от факторов окружающей среды, что наглядно продемонстрировано при оценках племенной ценности дочерей одних и тех же быков по удою их дочерей, рожденных в разных эколого-географических регионах.

Другая группа источников противоречий связана собственно с методами организации SNP-панелей, а именно с неизбежными ошибками в процессе секвенирования генома крупного рогатого скота, с трудностью дифференциации между случайными, структурными и функциональными редкими аллельными вариантами SNP, с различиями между частотами встречаемости SNP в дуплицированных и одиночных последовательностях

геномной ДНК, а также с ошибками в определении дублированных геномных последовательностей.

Третья группа источников сложностей применения SNP-панелей для контроля геномной структуры групп крупного рогатого скота и выявления генотипов, ассоциированных с фенотипическим проявлением хозяйственно ценных признаков, тесно связана с проблемами математической обработки большого массива получаемых данных, моделирования таких ассоциаций и их объективного анализа.

Совокупность перечисленных проблем приводит к возвращению к относительно более простым методам геномного сканирования (полилокусного генотипирования) крупного рогатого скота с целью разработок более простых, эффективных и менее затратных методов решения классических задач генетики животных сельскохозяйственных видов, таких как исключение ошибок происхождения, разработка методов прогноза количества и качества конечной продукции, устойчивости к условиям содержания и к инфекционным агентам, контроль наличия наследственных заболеваний и инфицированности патогенами, реконструкция истории и генеалогии пород и выяснение специфики их генофонда, разработка генетически обоснованных программ устойчивого использования местных пород и их сохранения, картирование главных генов количественных признаков в целях геномной селекции.

Для решения перечисленных задач в последние годы широко использовалось генотипирование по микросателлитным локусам. По-видимому, сложности использования генотипирования по микросателлитам обусловлены, прежде всего тем, что они выполняются по небольшому количеству локусов (не более двух десятков.). Поэтому их используют, в основном для генетической паспортизации. В этом отношении особый интерес представляют микросателлиты, формирующие инвертированные повторы в пределах расстояний до 2 тысяч пар нуклеотидов, что позволяет использовать их в качестве праймеров в полимеразной цепной реакции (ПЦР или PCR) для получения полилокусных спектров фрагментов ДНК, удобных для геномного сканирования. Этот метод получил название ISSR-PCR (ISSR-Inter-Simple Sequence Repeat, внутренние простые повторяющиеся последовательности).

ISSR-маркеры выгодно отличаются наличием информации о структурных особенностях их флангов (предрасположенность к петлеобразованию) от таких, например, как AFLP (Amplified Fragment Length Polymorphism – полиморфизм длины амплифицированных фрагментов), применяемых для полилокусного генотипирования и являющихся также «анонимными» по нуклеотидному содержанию амплифицируемых фрагментов ДНК.

Ранее нами были получены данные, свидетельствующие о том, что по ISSR-PCR маркерам группы черно-пестрого голштинизированного скота, отличающиеся по инфицированию вирусом бычьего лейкоза (Bovine Leukose Virus – BLV) и характеристиками молочной продуктивности, существенно дифференцируются по генетической структуре. Причем, генетические взаимоотношения между группами зависят от микросателлитов, используемых в PCR в качестве праймеров. Полиморфизм спектров продуктов амплификации (ампликонов), полученных в результате применения в качестве праймеров последовательностей  $(AGC)_6C$  и  $(GAG)_6C$  оказался сходным у всех групп коров, инфицированных BLV, вне зависимости от их молочной продуктивности, а спектры праймера  $(AG)_9C$  дифференцировали коров с высокой молочной продуктивностью от животных с низкой продуктивностью, независимо от их инфицированности BLV [2]. Секвенирование фрагмента ДНК, фланкированного инвертированным повтором  $(AGC)_6C$ , позволило обнаружить присутствие последовательностей, гомология к которым отсутствовала в секвенированном геноме герефорда (мясная порода) GenBank и относительно повышенную частоту встречаемости участков ретротранспозонов и их рекомбинантов, видоспецифичных для *Bos taurus* [3]. Полученные данные соответствуют результатам геномного секвенирования, свидетельствующим о тесной связи микросателлита с коровым мотивом AGC с видоспецифичным ретротранспозоном и их ассоциированным распространением в геноме крупного рогатого скота [4].

В целях выяснения нуклеотидного содержания фрагментов ДНК, фланкированных инвертированным повторами участков микросателлитов  $(GAG)_6C$  и  $(AG)_9C$  и их

изменчивости, в настоящей работе выполнено секвенирование фрагментов геномной ДНК коров голштинизированной породы хозяйства «Можайское» длиной около 500 пар нуклеотидов, фланкированных инвертированными повторами (GAG)<sub>6</sub>C и (AG)<sub>9</sub>C. Сравнение секвенированных последовательностей, фланкированных микросателлитом (GAG)<sub>6</sub>C и (AG)<sub>9</sub>C, представляет особый интерес потому, что оба этих микросателлита являются пурин/пиримидиновыми треками, предрасположенными к формированию таких неканонических структур ДНК, как триплексы, выполняющих существенную роль в формировании вторичных структур ДНК и РНК, влияющих на процессы репликации, транскрипции, частоту мутационных событий [5, 6]. Причем, не смотря на близость коровых мотивов и принадлежности к пурин-пиримидиновым трекам, эти микросателлиты отличаются тем, что (GAG)<sub>6</sub>C, в отличие от (AG)<sub>9</sub>C, предрасположены к формированию других неканонических мотивов ДНК, G4 квадруплексов.

### Материалы и методы

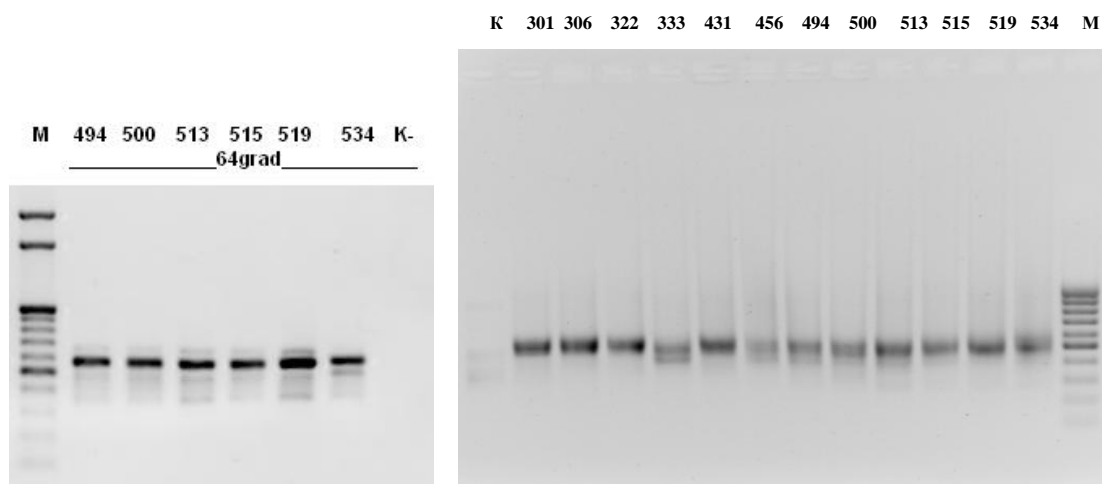
В работе использовали образцы тотальной ДНК, выделенной из цельной крови 12-ти коров черно-пестрой голштинизированной породы. Пробы крови были отобраны в декабре 2011 г. из хвостовой вены стерильными катетерами с использованием ЭДТА в качестве антикоагулянта. По результатам предыдущих исследований [7], три коровы были инфицированы вирусом бычьего лейкоза (№№ 494, 513, 519) и три коровы оказались свободными от инфекции (№№ 500, 515, 534).

ДНК выделяли с помощью набора Magna™ DNA Prep 200 («Лаборатория Изоген», Россия). Инкубационная смесь объемом 20 мкл содержала 2 мкл 10-кратного буфера и 1 мкл (5 ед.) Taq-полимеразы («Синтол», Россия), 2 мкл раствора dNTP (10 mM каждого), 1 мкл праймера (AGC)<sub>6</sub>G (20 пкмоль), 2 мкл (0,5-1 мкг) ДНК, 12 мкл деионизированной воды. ПЦР проводили при следующих условиях: начальная денатурация в течение 1 мин при 94°C; 35 циклов (30 с при 94°C, 30 с при 55°C, 2 мин при 72°C); финальная элонгация при 72°C в течение 10 мин. Электрофорез продуктов амплификации проводили в 1 % агарозном геле. В качестве маркеров использовали M25 DNA Ladder и M11 DNA Ladder («СибЭнзим», Россия). На секвенирование были взяты фрагменты ДНК длиной 450 - 550 пар нуклеотидов (п.н.), выделенные из агарозных гелей. Получение библиотеки фрагментов ДНК проводили по протоколу подготовки быстрых библиотек, клональную эмульсионную ПЦР и секвенирование проводили с использованием наборов реактивов согласно рекомендациям фирмы-производителя (Roche). Определение нуклеотидных последовательностей ампликонов проводилось на геномном анализаторе GS Junior (Roche). Определено около 22 млн. п.н., 70 223 прочтений были идентифицированы как пригодные для дальнейшего анализа, из них были отобраны только последовательности длиной не менее 380 нп.

Секвенированные данные организовывали с учетом разработанных методов такого анализа [8, 9, 10]. Последовательности группировались в отдельные кластеры («риды») на основании их идентичности.

Выравнивание секвенированных последовательностей каждого кластера с хромосомными последовательностями *Bos taurus* выполняли с использованием алгоритмов программы BLASTn (ресурс <https://blast.ncbi.nlm.nih.gov/>). Для анализа присутствия в секвенированных последовательностях участков, идентичных микросателлитам и диспергированным повторам, использовали компьютерные программы RepeatMasker (<http://www.repeatmasker.org/>) и Giri (<http://www.girinst.org>). Принадлежность генов, в которых присутствовали участки гомологии к секвенированным последовательностям, к различным функциональным группам и метаболическим путям, определяли с использованием базы данных GeneCards.

## Результаты и их обсуждение



*Рис. 1.* Амплифицированные и секвенированные последовательности геномных ДНК чернопестрых коров, фланкированных инвертированным повтором микросателлитов GAG (слева) и AG (справа).

М – маркер молекулярных масс, k – контроль без геномной матрицы с полным составом реакционной смеси, с праймером.

Над каждой дорожкой указаны номера исследованных коров

В результате выполненных процедур амплификации, выделения и секвенирования фрагмента геномной ДНК черно-пестрых голштинизированных коров, получен ряд нуклеотидных последовательностей, анализ которых выполнялся с использованием алгоритмов программ BLASTn (ресурс <https://blast.ncbi.nlm.nih.gov/>), RepeatMasker (<http://www.repeatmasker.org/>) и Giri (<http://www.girinst.org/>), а также базы данных GeneCards.

С использованием алгоритмов программы BLSTn, выполнен поиск участков гомологии к последовательности праймера gaggaggaggaggaggagc в плюс- и минус цепях секвенированного генома коровы мясной породы герефорд, представленных в базе GenBank. Оказалось, что таких последовательностей в геноме существенно меньше по сравнению, например, с таким микросателлитом, как agcagcagcagcagcagcagcagcagc [3]. Если учитывать только полную гомологию, суммарно таких участков 79, с одной или двумя заменами – 762. В связи с низкой частотой встречаемости микросателлит с коровым мотивом gag даже не вошел в результаты геномного анализа частот встречаемости три-нуклеотидных микросателлитов у крупного рогатого скота, в отличие от микросателлита с мотивом agc [4].

Учитывая нуклеотидные ошибки, возникающие при проведении полимеразной цепной реакции, процедуры самого секвенирования, наличия фильтров для удаления из результатов недостаточно качественно секвенированных участков, традиционно рекомендуется в анализ секвенированных последовательностей включать только те, которые образуют кластеры («риды»), содержащие не менее 12 идентичных копий нуклеотидных последовательностей [11].

В результате секвенирования получены следующие данные. В последовательностях, несущих участки гомологии к праймеру (GAG)<sub>6</sub>C в «ридах», содержащих более 12-ти идентичных копий, среди межгенных областей только 4 встречаются у всех без исключения коров (хромосома 3, на 3' конце участка гомологии к gaggaggaggaggaggag располагается ген предшественника макрофаг-колонии стимулирующего фактора; хромосома 20 - ген фактора регуляции транскрипции IRX-1, опухоль-супрессирующий фактор; хромосома 22 – ген белка RFT1, участвующего в гликолизировании белков для их транспорта в каналцы эндоплазматического ретикулума; хромосома 25 – ген рибосомального белка 39S L28, митохондриального предшественника) и 8 в структурных генах (хромосома 5 – cytohesin-4 –

член семейства белков PSCD, участвующих в белковом сортинге и связывании белков с плазматическими мембранами; хромосома 8 – palladin isoform X2, AKT1 – член семейства серин/треонин киназ, ингибирующих апоптоз; хромосома 13 – potassium voltage-gated channel subfamily B member 1, участвующий в контроле клеточного объема и ионной проводимости; хромосома 15 – arf-GAP с доменом Rho-GAP – ADP-рибозилирующий фактор (ARF) GTPаза активирующий белок, участвующий в перестройках цитоскелета в процессах клеточной дифференцировки; хромосома 16 – MAP kinase-activated protein kinase 2 – член семейства серин/треонин киназ MAP метаболического пути, который является интегральной точкой пересечения множества внутриклеточных процессов, таких как пролиферация, цитодифференцировка, регуляция транскрипции; хромосома 18 – phosphotyrase b kinase регуляторная субъединица, взаимодействует с кальмодулином, участвует в фосфорилировании серина в ряде субстратов, в частности, тропонина; хромосома 19 – gas-related protein Rab-40B, участвует в формировании внутриклеточных секреторных везикул; хромосома 27 – deoxycytidylate deaminase, специфичная ДНК-одноцепочечная цитидин-деаминаза, связана с соматическим гипермутационным процессом, генной конверсией и класс-переключающей рекомбинацией В-лимфоцитов.

В общем, из 12-ти разных участков полной гомологии к последовательности gagaggaggaggaggagc, два связаны с регуляцией транскрипции, два с функцией иммунной системы, все остальные – со структурными элементами клетки, плазматическими и внутриклеточными мембранами, элементами цитоскелета.

Сама последовательность микросателлита с коровым мотивом GAG предрасположена не только к формированию триплексов, в том числе и межмолекулярных (ДНК-РНК), но и к образованию G4 квадруплексов – сложной вторичной структуре ДНК, непосредственно вовлекаемой как в регуляцию генной экспрессии, так и в повышенную частоту мутационных событий [12]. Можно ожидать, что ее локализация в перечисленных геномных участках в определенной степени ассоциирована с регуляторными событиями, вовлеченными в функцию иммунной системы, а также перестройки клеточной архитектоники.

В связи с тем, что фланги секвенированного участка включали инвертированный повтор микросателлита, что само по себе предполагает высокую вероятность возникновения спонтанных мутаций и ошибок при получении конечных данных, далее в анализ мы включали все без исключения последовательности, даже если в кластере («риде») их количество не превышало двух копий.

При анализе всех секвенированных последовательностей получены следующие данные. В отличие от фрагментов ДНК, фланкированных инвертированным повтором (AGC)<sub>6</sub>C [3], все, без исключения, кластеры имели участки гомологии в полностью секвенированном геноме *Bos taurus*, представленном в GenBank, полученном на основании генома коровы мясной породы герефорд [4].

Участки гомологии к 49-ти фрагментам обнаруживались в межгеномном пространстве на разных хромосомах. Анализ функциональных особенностей структурных генов, локализованных по данным GenBank на 3' концах таких участков гомологии с использованием базы данных GeneCards, позволил их подразделить на 13 функциональных групп. Чаще всего, в 8-ми из 49 случаев, их 3' концы находились вблизи с 5' флангами генов, участвующими в функционировании иммунной системы, такими как макрофаг-стимулирующий фактор, семейства иммуноглобулин-подобных рецепторов лейкоцитов, тяжелая цепь иммуноглобулинов эpsilon, поверхностные рецепторы Т-клеток. Вторая по частоте встречаемости функциональная группа таких последовательностей включала системы передачи сигнала (в 6-ти из 49 случаев), такие как ольфакторные рецепторы, содержащие домен SH3 киназа – связывающие белки, тирозин фосфатазы, гены цАМФ метаболического пути. Остальные фрагменты гомологии локализовались рядом с началами генов, которые относились к регуляторам пролиферации (3 участка), ДНК фолдинга (4 участков), геометрии клетки (4 участка), адгезии (1 участок), к факторам регуляции транскрипции (5 участков), рибосомальным белкам (4 участка), к регуляторам транспорта в клетку (3 участка), секреции (2 участка), убиквитин\_зависимых путей (4 участка), липидного синтеза (2 участка), углеводного синтеза (3 участка). В общем, участки гомологии к секвенированным последовательностям, локализованные в межгеномных пространствах и обладающие предрасположенностью к формированию триплексов (как содержащие

пурин/пиримидиновые треки), располагались преимущественно перед структурными генами, продукты которых участвуют в функциях иммунной системы и клеточных каскадах передачи сигнала.

Собственно в структурных генах обнаружено 57 участков гомологии, причем, в отличие от 13-ти функциональных групп генов на 3' флангах таких участков, в выборке генов с внутренними гомологичными последовательностями обнаруживается 11 функциональных групп, без генов, продукты которых участвуют в регуляции секреции и липидного синтеза. Ведущими группами оказались гены иммунной системы (7 генов), клеточных систем передачи сигналов (8 генов) и регуляции транскрипции (8 генов). Несколько иное распределение обнаруживается и по участкам гомологии в генах, продукты которых участвуют в регуляции пролиферации (6 генов), ДНК фолдинга (5 генов), геометрии клетки (3 гена), адгезии (5 генов), рибосомальным белкам (3 гена), к регуляторам транспорта в клетку (4 гена), убиквитин-зависимых путей (4 гена), углеводного синтеза (4 гена).

Таким образом, полученные данные свидетельствуют о том, что среди секвенированных последовательностей чаще всего встречаются ассоциированные с генами иммунной и сигнальной систем. Причем из 8-ми генов, расположенных на 3' фланге участков гомологии к микросателлиту (GAG)<sub>6</sub>C и из 7 структурных генов, в которых есть участки гомологии к этой последовательности, продукты которых вовлечены в функции иммунной системы, четыре гена одинаковы (табл.1).

Таблица 1. Гены иммунной системы, в которых обнаруживаются участки гомологии к нуклеотидной последовательности (GAG)<sub>6</sub>C, или расположенные рядом с 3' концом этой последовательности

Межгенное пространство с участками гомологии к (GAG) <sub>6</sub> C на 3' флангах которой расположены следующие гены:								
Гены на 3' флангах	antimicrobial peptide NK-lysin-like isoform X2	leukocyte immunoglobulin-like receptor subfamily B member	cancer/testis antigen 47A-like	immunoglobulin heavy constant epsilon	receptor-type tyrosine-phosphatase U precursor	T-cell surface glycoprotein CD4 isoform X	B-cell lymphoma/leukemia 11B isoform X	macrophage colony-stimulating factor
	Хромосома 11	Хромосома 18	Хромосома X	Хромосома 21	Хромосома 2	Хромосома 5	Хромосома 21	Хромосома 3
Структурные гены с участками гомологии к (GAG) <sub>6</sub> C								
Структурные гены	antimicrobial peptide NK-lysin-like isoform X2	leukocyte immunoglobulin-like receptor, subfamily B	cancer/testis antigen 47A-like	immunoglobulin heavy constant epsilon	C-C motif chemokine 25 precursor	rho guanine nucleotide exchange factor 40	paired box protein Pax-5 isoform X1	
	Хромосома 11	Хромосома 18	Хромосома X	Хромосома 21	Хромосома 7	Хромосома 10	Хромосома 8	

То есть, такие гены как antimicrobial peptide NK-lysin-like isoform X2 (хромосома 11); leukocyte immunoglobulin-like receptor subfamily B member (хромосома 18); immunoglobulin heavy constant epsilon (хромосома 21) и cancer/testis antigen 47A-like (хромосома X) могут нести участки гомологии к последовательности gagaggaggaggaggagc как в межгенном пространстве, на своем 5' конце, так и внутри гена. Следует отметить, что на хромосоме 21 ген immunoglobulin heavy constant epsilon (присутствующий в результатах секвенирования у всех исследованных коров) расположен вблизи гена B-cell lymphoma/leukemia 11B isoform X и между ними у некоторых животных также обнаруживается участок гомологии к последовательности (GAG)<sub>6</sub>C. Из этого

следует, что участок гомологии может располагаться как рядом с 5' флангом этого гена, внутри него и за его 3' концом. Причем ген B-cell lymphoma/leukemia 11B isoform X хорошо известен предпочтительной интеграцией в нем провирусной ДНК при миелоидной лейкемии у мышей, индуцируемой встройкой ретровируса.

Известно также, что длинные концевые повторы (LTR) провирусной ДНК вируса BLV содержат последовательности, предрасположенные к формированию таких неканонических структур ДНК, как G4 квадруплексы [13], триплексы (in vivo и in vitro) [6]. В наших собственных исследованиях ранее нами было показано, что сравнительно длинный пурин-пиримидиновый трек длиной в 13 п.о., предрасположенный к формированию межмолекулярных триплексов ДНК-РНК, и отличающийся относительно повышенным консерватизмом как в секвенированных последовательностях GenBank, так и лимфомах, индуцированных BLV, начинается в LTR BLV с 90 нуклеотида (СТТССССТТССС) и локализован в области второго участка TxRE [13]. Нами также были выявлены короткие пурин-пиримидиновые треки, которые могут создавать как шпильчатые, так и триплексные структуры в LTR BLV - два участка ТСТСТССТ, от 364 нуклеотида и от 458 нуклеотида, и одна инвертированная по отношению к ним последовательность AGGAGAGA (с позиции 485 нуклеотид), первая на расстоянии в 113 нуклеотидов от инвертированной последовательности, вторая – на расстоянии 19 нуклеотидов. Можно ожидать, что встройки провирусной ДНК с такими флангами, предрасположенными к формированию неканонических структур ДНК, оказывают существенное влияние на экспрессию рядом расположенных генов.

Участок TxRE (Tax-responsive elements), связывающий активатор транскрипции провирусной ДНК (Tax), в котором выявлен пурин-пиримидиновый трек, представляет особый интерес, поскольку локализован в энхансерном участке LTR. Показано, что Tax непосредственно участвует в регуляции транскрипции провирусной ДНК BLV и оказывает существенное влияние на экспрессию генов хозяина, связанных с клеточной пролиферацией и дифференцировкой, иммортализацией первичных клеточных популяций в системе in vitro, увеличения скорости мутирования путем подавления эксцизионной репарации ДНК [14]. Показано, что Tax активирует белки, принадлежащие сигнальному пути активаторов белков 1 (activator protein 1 – AP-1), в который входят, в частности, FBJ остеосаркома онкоген (FOS), jun proto-oncogene – JUN) и ряд других, через взаимодействие с другими транскрипционными метаболическими путями (G-protein, GTP-binding proteins и т.д.) [15]. Не смотря на широкий спектр мишеней регуляторных эффектов Tax белка, принято считать, что в общем, одной из ведущих мишеней регуляторных эффектов этого белка являются гены, кодирующие белки, связанные с функцией иммунной системы [14]. Мотивы TxRE, предрасположенные к формированию неканонических вторичных структур ДНК, как мишени опознавания Tax, могут быть вовлечены в эти процессы. В наших исследованиях последовательности gaggaggaggaggaggagc, предрасположенные к формированию G4 квадруплексов и межмолекулярных триплексов (ДНК-РНК), также обнаруживаются преимущественно в связи с генами, продукты которых участвуют в функциях иммунной системы, а также в передаче сигналов.

Среди полученных нами результатов секвенирования геномных фрагментов инфицированных BLV и свободных от инфекции коров был только один структурный ген antimicrobial peptide NK-lysin, участок которого всегда входил в секвенированные последовательности продукта амплификации геномной ДНК, фланкированной инвертированным повтором gaggaggaggaggaggagc, у всех трех инфицированных BLV коров и отсутствовал у 2-х коров, свободных от инфекции. То есть, у всех инфицированных коров высока вероятность подавления экспрессии этого гена за счет образования межмолекулярного триплекса (ДНК-РНК) и, соответственно, подавления транскрипции [5]. У двух коров, свободных от BLV, участки гомологии к (GAG)<sub>6</sub>C в этом гене отсутствовали, что приводило к его отсутствию в секвенированных продуктах амплификации при использовании (GAG)<sub>6</sub>C в полимеразной цепной реакции в качестве праймера.

У третьей коровы, свободной от инфекции, присутствие этого участка в структурном гене совпадало с присутствием его же в межгенном пространстве на 5' конце гена, что позволяет предполагать сложный механизм взаимодействия между ними.

Очевидно, что связь между присутствием в гене NK-lysin (сапозин) последовательности, предрасположенной к формированию неканонических ДНК структур и сниженной



чувствительностью коров к инфицированию BLV, требует дальнейших экспериментальных исследований. Однако факт совпадения присутствия в нем инвертированной последовательности (GAG)<sub>6</sub>C, предрасположенной к формированию сложных неканонических вторичных структур ДНК и снижению эффективности транскрипции, у инфицированных BLV коров, по нашему мнению, заслуживает особого внимания. Сапозин – член семейства сапозин-подобных белков (SAPLIP), локализованных в цитотоксических гранулах Т клеток, освобождающихся при антиген стимуляции. Этот белок присутствует в цитотоксических гранулах цитотоксических Т лимфоцитов и нуллеров-киллеров (NK клетки), описана их антимикробная активность по отношению к туберкулезной бактерии *M. tuberculosis*, других инфекционных агентов, а также участие в защите от некоторых ретровирусных инфекций, таких как вирус иммунодефицита человека (HIV) [16, 17, 18, 19, 20].

Следует подчеркнуть, что HIV по ряду своих характеристик очень близок к BLV и исследования организации последнего, взаимодействия его с многоклеточными организмами часто используются как модель для получения новой информации об особенностях HIV [20].

С использованием программ RepeatMasker (<http://www.repeatmasker.org/>) и Gini (<http://www.girinst.org>) был выполнен в секвенированных последовательностях поиск участков гомологии среди диспергированных повторов. В отличие от результатов такого поиска в секвенированных последовательностях ДНК, фланкированных инвертированным повтором (AGC)<sub>6</sub>G, выполненного нами ранее [3], в данном случае диспергированных повторов оказалось всего 12, причем участки гомологии были относительно короткими (табл. 2).

Таблица 2. Мобильные генетические элементы, гомология к которым обнаружена в секвенированных фрагментах, фланкированных инвертированным повтором (GAG)<sub>6</sub>C, в геномной ДНК коров, инфицированных вирусом бычьего лейкоза (+BLV), и свободных от инфекции (-BLV)

№№ Коровы	494 +BLV	500 - BLV	513+BLV	515-BLV	519+BLV	534-BLV
Транспозоны	Координаты участков гомологии в п.о. (гомология не менее 80%)					
ДНК транспозон Sola1-1_ACas	1 - 53	нет	1 - 53	1 - 53	1 - 53	нет
ДНК транспозон DNA/hAT	нет	нет	нет	нет	нет	2031-2116
ДНК транспозон CHARLIE2A	нет	нет	нет	нет	нет	12-96
Ретротранспозон Copia-22_PIT-I	257-305	257-305	160-208	161-209	нет	нет
Ретротранспозон Copia-8_TA-I	нет	349-461	нет	нет	нет	нет
Ретротранспозон Gypsy-15_CFI-I	нет	нет	358-437	357-436	358-437	нет
Ретротранспозон Gypsy-18_SCH-I	нет	нет	366-421	367-422	1-56	нет
Ретротранспозон Gypsy-32_Mad-I	нет	нет	нет	нет	нет	126-222
Ретротранспозон LTR76_EC	нет	401-484	400-484	нет	нет	нет
Ретротранспозон LTR-14B_Crp	429-457	нет	нет	нет	нет	нет
Ретротранспозон RMER17C2 (ERV/ERV2)	нет	нет	нет	нет	нет	472-503
Ретротранспозон REP-19_CPВ	нет	нет	322-433	нет	нет	нет



Такой продукт рекомбинации, как представленный на рис. 2, присутствующий в межгенном пространстве между генами томорегулин и предшественником транспортера цинка, обнаруживается в секвенированных последовательностях всех исследованных коров без исключения, а также встречается в ряде секвенированных последовательностях, имеющих гомологию с межгенными и генными последовательностями референсного генома.

Оценена распространенность L1\_BT с использованием алгоритма BLAT (в открытом доступе <http://genome.ucsc.edu/cgi-bin/hgBlat> <http://emb-bioinfo.fsaa.ulaval.ca/bioinfo/html/cgi-bin/hgBlat?hgsid=2400&command=start>). Обнаружено, что на всех без исключения хромосомах референсного генома *Bos taurus* с высокой частотой встречается этот вариант рекомбинации между двумя L1-BT. Это позволяет сделать вывод о том, что микросателлит AG тесно связан именно с этим продуктом рекомбинации.

Не наблюдается выраженных отличий между геномами коров по представленности в секвенированных последовательностях ретротранспозонов или ДНК транспозонов, вне зависимости от инфицированности коров вирусом бычьего лейкоза. У каждой коровы участки, фланкированные инвертированными повторами (AG)<sub>n</sub>C, оказывались ассоциированы с простыми и сложными микросателлитами (ACTC, TC, GA).

Суммарно у 12-ти исследованных коров секвенированные последовательности размером около 500 пар оснований, фланкированные инвертированным повтором (AG)<sub>n</sub>C, локализируются с уровнем гомологии выше 95 % в 156 межгенных пространствах и 80-ти структурных генах референсного генома.

Во всех группах секвенированных последовательностей, полученных на геномной матрице разных коров, присутствовали с 3'конца фрагмента такие гены как zinc transporter ZIP10 precursor (хромосома 2), solute carrier family 22 member 16 (хромосома 9), Attractin и attractin isoform X1 (хромосома 13), proteasome assembly chaperone 3 (хромосома 25), zinc finger C4H2 domain-containing protein isoform X2 (хромосома X).

Участки гомологии присутствовали в следующих структурных генах: SET and MYND domain-containing protein 1 (хромосома 11), neuronal PAS domain-containing protein 3 (хромосома 21), inactive carboxypeptidase-like protein X2 precursor (хромосома 26), disrupted in schizophrenia 1 protein isoform 1 (хромосома 28). В большинстве фрагментов присутствовали участки гомологии к мобильным генетическим элементам, как правило, образующие рекомбинантные продукты.

### **Заключение**

В популяционной геномике в последние годы широкое распространение получили методы полилокусного генотипирования (геномного сканирования) с использованием разных молекулярно-генетических маркеров, начиная с мононуклеотидных замен (Single Nucleotide Polymorphisms – SNP), до полного секвенирования. К одному из таких методов относится оценки полиморфизма фрагментов ДНК, фланкированных инвертированными повторами участков микросателлитных локусов (Inter-Simple Sequence Repeat – ISSR-PCR) [1]. Фирма «Прикладные биосистемы» (Applied Biosystems – [www.appliedbiosystems.com](http://www.appliedbiosystems.com)) по согласованию с Продовольственной и сельскохозяйственной организацией ООН (ФАО) (Food and Agriculture Organization (FAO)) и Международным обществом генетики животных (International Society of Animal Genetics (ISAG)) разработала тест-систему для генотипирования для крупного рогатого скота по 11-ти микросателлитам. Однако, по-видимому, сложности использования генотипирования по микросателлитам обусловлены прежде всего тем, что они выполняются по небольшому количеству локусов (не более двух десятков), полиморфизм которых (аллельное разнообразие, частота спонтанного мутирования, скорость фиксации аллелей) существенно отличается от одного локуса к другому. В этом отношении особый интерес представляют микросателлиты, формирующие инвертированные повторы в пределах расстояний до 2 тысяч пар нуклеотидов, использование которых получило название ISSR-PCR (ISSR-Inter-Simple Sequence Repeat, внутренние простые повторяющиеся последовательности). К достоинствам этого метода относится то, что в полимеразной цепной реакции (PCR) происходит множественная адресная амплификация фрагментов геномной ДНК, длиной, как правило, от 200 до 2000 пар оснований (п.о.), у которых есть выраженная структурная особенность – наличие

на флангах инвертированного повтора, маркирующего предрасположенность к формированию неканонических структур – ДНК петель. В то же время, существенным его недостатком, кроме доминантного характера присутствия амплифицируемого фрагмента ДНК, является «анонимность» нуклеотидных последовательностей, заключенных между известными флангами. Этот вопрос представляет особый интерес, поскольку количество амплифицируемых фрагментов ДНК, их полиморфизм у одних и тех же животных существенно отличается в зависимости от микросателлита, участок которого используется в качестве праймера в PCR [1].

Существенным недостатком ISSR-маркеров является доминантный характер их проявления, поскольку присутствие фрагмента ДНК определенной длины в спектре продуктов амплификации, полученных в ПЦР с использованием фрагмента микросателлита в качестве праймера, не позволяет отличить гомозиготу от гетерозиготы. Кроме того, «анонимное» нуклеотидное содержание амплифицируемых фрагментов ДНК одинаковой длины не исключает возможности их внутренней нуклеотидной гетерогенности при амплификации с разных геномных участков. С другой стороны объективным достоинством ISSR-маркеров является их «адресность» – принадлежность амплифицируемых участков ДНК только к тем участкам геномов, где на расстоянии до 2 тысячи пар нуклеотидов (т.п.н.) локализован инвертированный повтор микросателлита, что позволяет исследовать геномное распределение таких повторов.

Полученные результаты наших исследований по геномным секвенированным нуклеотидным последовательностям позволяют сделать следующее заключение. В результате впервые выполненного секвенирования нуклеотидных последовательностей, длиной около 550 п.о., фланкированных инвертированными повторами (GAG)<sub>6</sub>C и (AG)<sub>9</sub>C, обнаружены существенные отличия между ними. Инвертированный повтор первого микросателлита встречается с примерно одинаковой частотой в межгенных пространствах и в структурных генах, участки гомологии ко второму микросателлиту обнаруживаются в два раза чаще в межгенных пространствах, чем в структурных генах. Участки гомологии к последовательности (GAG)<sub>6</sub>C фланкируют геномные районы, в которых не наблюдается повышенная частота встречаемости видоспецифичных для *Bos taurus* ретротранспозонов или продуктов их рекомбинации.

Не обнаружено выраженных отличий в секвенированных фрагментах, фланкированных инвертированным повтором (AG)<sub>9</sub>C, полученных на геномной ДНК инфицированных вирусом бычьего лейкоза и свободных от инфекции. Это хорошо согласуется с полученными нами ранее данными о том, что полиморфизм в спектрах продуктов амплификации, полученных с праймером (AG)<sub>9</sub>C, не дифференцирует коров с интеграцией провирусного генома этого вируса в геном коров, и свободных от такой интеграции.

Полученные данные свидетельствуют о том, что микросателлит (AG)<sub>9</sub>C тесно ассоциирован с другими ди- и тетромикросателлитами, с ретротранспозонами, включая видоспецифичные для *Bos taurus* LINE и SINE и эндогенные ретровирусы, с их фрагментами (в частности, с LTR) и продуктами их рекомбинации. Все это свидетельствует в пользу предположений, имеющих в литературных источниках, о том, что некоторые микросателлиты являются «меткой» или «реперной точкой» геномных участков, предпочтительных для встраивания и выщепления ретротранспозонов. Отсюда следует их активная возможность участия в микро и макроэволюционных событиях. В то же время, полученные нами данные свидетельствуют о выраженных отличиях между содержанием фрагментов геномной ДНК крупного рогатого скота, фланкированных инвертированными повторами микросателлитов, даже имеющих такие выраженные сходства по коровым мотивам, как микросателлиты (GAG)<sub>6</sub>C и (AG)<sub>9</sub>C.

### Примечания

1. Глазко В.И., Косовский Г.Ю., Глазко Т.Т. Введение в геномную селекцию животных. М.: ООО "Приятная компания", 2012. 258 с.
2. Косовский Г.Ю., Глазко В.И., Архипов А.В., Петрова И.О., Глазко Т.Т. Популяционно-генетическая дифференциация молочного скота по ISSR-PCR маркерам. // Доклады Российской академии сельскохозяйственных наук. 2014. №5. С.53-56.

3. Глазко В.И., Косовский Г.Ю., Ковальчук С.Н., Архипов А.В., Петрова И.О., Дедович Г.О., Глазко Т.Т. Инвертированный повтор микросателлита (AGC)<sub>6</sub> фланкирует районы ДНК с участками гомологии к ретротранспозонам в геноме крупного рогатого скота. // *Инновационные технологии в медицине*. 2014. 2(03). С. 63-79.
4. Tellam R.L., Worley K.C. The genome sequence of taurine cattle: a window to ruminant biology and evolution. // *Science*. 2009. 324: 522-528
5. Buske F.A., Mattick J.S., Bailey T.L. Potential in vivo roles of nucleic acid triple-helices// *RNA Biology*. 2011. V. 8. №3. P. 427-439
6. Limanskaya O.Yu. Polypurine/polypyrimidine sequences with potential of forming triplexes in the proviral DNA of bovine retroviruses // *Cytology and Genetics*. 2010 V. 44. №1: 10-18.
7. Косовский Г.Ю., Сотникова Е.А., Мудрик Н.Н., Cuong V.C., Тоан Т.Х., Ноан Т.Х., Глазко В.И. Диагностика лейкоза КРС с помощью праймеров к генам gag и pol. // *Ветеринария*. 2013. №8: 58-61.
8. Ewing B., Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. // *Genome Res*. 1998. 8: 186-194.
9. Ewing B., Hillier L., Wendl M. C., Green, P. Basecalling of automated sequencer traces using phred. I. Accuracy assessment. // *Genome Res*. 1998. 8: 175-185.
10. Gordon D., Abajian C., Green, P. Consed: a graphical tool for sequence finishing. // *Genome Res*. 1998. №8: 195-202
11. Ребриков, Д.В., Коростин, Д.О., Шубина, Е.С., Ильинский, В.В. NGS: высокопроизводительное секвенирование/ Д. В. Ребриков. М., 2014. 232 с.
12. Zybailov B.L., Sherra M.D., Glazko G.V., Raney K.D., Glazko V.I. G4-quadruplexes and genome instability // *Molecular Biology*. 2013. V. 47. № 2. P. 197-204.
13. Самуйленко А.Я., Косовский Г.Ю., Гринь С.А., Синковец С.М., Глазко Т.Т., Глазко В.И. Полиморфизм и потенциальные неканонические структуры в LTR вируса бычьего лейкоза / В сб: *Научные основы производства и обеспечения качества биологических препаратов для АПК. Материалы международной научно-практической конференции, посвященной 45-летию института ВНИТИБП, 27-28 ноября 2014 г.* / Под ред. акад. РАН А.Я. Самуйленко. М., 2014. 525 с. С. 106-117.
14. Arainga M., Takeda E., Aida Y. Identification of bovine leukemia virus tax function associated with host cell transcription, signaling, stress response and immune response pathway by microarray-based gene expression analysis // *BMC Genomics* 2012. 13:12, <http://www.biomedcentral.com/1471-2164/13/121>
15. Aida Y, Murakami H, Takahashi M., Takeshima S-N. Mechanisms of pathogenesis induced by bovine leukemia virus as a model for human T-cell leukemia virus// *Front. Microbiol*. 2013. V. 4. Article 328, doi:10.3389/fmicb.2013. 00328
16. Lee M.O., Kim E-H, Jang H-J., Park M-N, Woo H-J., Han J.Y., Womack J.E. Effects of a single nucleotide polymorphism in the chicken NK-lysin gene on antimicrobial activity and cytotoxicity of cancer cells // *PNAS*. 2012. Vol. 109, №30. P. 12087-12092.
17. Zhang M, Li M-f, Sun L. NKLP27: A Teleost NK-Lysin Peptide that Modulates Immune Response, Induces Degradation of Bacterial DNA, and Inhibits Bacterial and Viral Infection. // *PLoS ONE*. 2014. V. 9. Is. 9. e106543. doi:10.1371/journal.pone.0106543
18. Pitabut N., Sakurada S., Tanaka T., Ridruechai C., Tanuma J., Aoki T., Kantipong P., Piyaworawong S., Kobayashi N., Dhepakson P., Yanai H., Yamada N., Oka S., Okada M., Khusmith S., Keicho N. Potential Function of Granulysin, Other Related Effector Molecules and Lymphocyte Subsets in Patients with TB and HIV/TB Coinfection. // *Int. J. Med. Sci*. 2013. Vol. 10. №8. P. 1003-1014.
19. Endsley J. J., Furrer J. L., Endsley M.A., McIntosh M.A., Maue A.C., Waters W. R., Lee D.R., Estes D.M. Characterization of Bovine Homologues of Granulysin and NK-lysin // *The Journal of Immunology*. 2004. V. 173. P. 2607-2614.
20. Gillet N.A., Gutierrez G., Rodriguez S.M., de Brogniez A., Renotte N., Alvarez I., Trono K., Willems L. Massive Depletion of Bovine Leukemia Virus Proviral Clones Located in Genomic Transcriptionally Active Sites during Primary Infection. // *PLoS Pathog*. 2013. V. 9. №10: e1003687. doi:10.1371/journal.ppat.1003687

### References:

1. Glazko VI, Kosovskiy GYu, Glazko TT. Introduction to genomic selection of animals. M.: OOO "Good company", 2012. 258 pp. (in russian)
2. Kosovskiy GYu, Glazkov VI, Arkhipov AV, Petrova IO, Glazko TT. Population-genetic differentiation of dairy cattle by ISSR-PCR markers. // *Reports of the Russian Academy of Agricultural Sciences*. 2014. №5. pp. 53-56. (in russian)
3. Glazko VI, Kosovskiy GYu, Kovalchuk SN, Arkhipov AV, Petrova IO, Dedovich GO, Glazko TT. Inverted repeat of microsatellite (AGC)<sub>6</sub> flanks the regions of DNA with homology to retrotransposons in the genome of cattle. // *Innovative technologies in medicine*. 2014 2 (03). pp. 63-79. (in russian)

4. Tellam R.L., Worley K.C. The genome sequence of taurine cattle: a window to ruminant biology and evolution. // *Science*. 2009. 324: 522-528
5. Buske F.A., Mattick J.S., Bailey T.L. Potential in vivo roles of nucleic acid triple-helices// *RNA Biology*. 2011. V. 8. №3. P. 427-439
6. Limanskaya O.Yu. Polypurine/polypyrimidine sequences with potential of forming triplexes in the proviral DNA of bovine retroviruses // *Cytology and Genetics*. 2010 V. 44. №1: 10-18.
7. Kosovskiy GYu, Sotnikova EA, Mudrik NN, Cuong VC, Toan TX, Hoan TX, Glazko VI. Diagnosis of bovine leukemia by primers for genes gag and pol. // *Veterinary Medicine*. 2013. №8: 58-61. (in russian)
8. Ewing B., Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. // *Genome Res*. 1998. 8: 186-194.
9. Ewing B., Hillier L., Wendl M. C., Green, P. Basecalling of automated sequencer traces using phred. I. Accuracy assessment. // *Genome Res*. 1998. 8: 175-185.
10. Gordon D., Abajian C., Green, P. Consed: a graphical tool for sequence finishing. // *Genome Res*. 1998. №8: 195-202
11. Rebriko, DV, Korostin DO, Shubin ES, Ilyinsky VV. NGS: high-throughput sequencing / DV Rebrikov. M.: BINOM. Knowledge Laboratory, 2014. 232 pp. (in russian)
12. Zybailov B.L., Sherpa M.D., Glazko G.V., Raney K.D., Glazko V.I. G4-quadruplexes and genome instability // *Molecular Biology*. 2013. V. 47. № 2. P. 197-204.
13. Samujlenko AYa, Kosovskiy GYu, Grin' SA, Sinkovets SM, Glazko TT, Glazko VI. Polymorphism and potential non-canonical structure of bovine leukemia virus LTR / In: Scientific bases of production and quality of biological products for the agricultural sector. Proceedings of the international scientific-practical conference dedicated to the 45th anniversary of the Institute VNITIBP, 27-28 November 2014 / Ed. Acad. RAS AYa Samujlenko. Moscow, 2014. 525 pp. Pp. 106-117. (in russian)
14. Arainga M., Takeda E., Aida Y. Identification of bovine leukemia virus tax function associated with host cell transcription, signaling, stress response and immune response pathway by microarray-based gene expression analysis // *BMC Genomics* 2012. 13:12, <http://www.biomedcentral.com/1471-2164/13/121>
15. AidaY, Murakami H, Takahashi M., Takeshima S-N. Mechanisms of pathogenesis induced by bovine leukemia virus as a model for human T-cell leukemia virus// *Front. Microbiol*. 2013. V. 4. Article 328, doi:10.3389/fmicb.2013. 00328
16. Lee M.O., Kim E-H, Jang H-J., Park M-N, Woo H-J., Han J.Y., Womack J.E. Effects of a single nucleotide polymorphism in the chicken NK-lysin gene on antimicrobial activity and cytotoxicity of cancer cells // *PNAS*. 2012. Vol. 109, №30. P. 12087-12092.
17. Zhang M, Li M-f, Sun L. NKLP27: A Teleost NK-Lysin Peptide that Modulates Immune Response, Induces Degradation of Bacterial DNA, and Inhibits Bacterial and Viral Infection. // *PLoS ONE*. 2014. V. 9. Is. 9. e106543. doi:10.1371/journal.pone.0106543
18. Pitabut N., Sakurada S., Tanaka T., Ridruechai C., Tanuma J., Aoki T., Kantipong P., Piyaworawong S., Kobayashi N., Dhepakson P., Yanai H., Yamada N., Oka S., Okada M., Khusmith S., Keicho N. Potential Function of Granulysin, Other Related Effector Molecules and Lymphocyte Subsets in Patients with TB and HIV/TB Coinfection. // *Int. J. Med. Sci*. 2013. Vol. 10. №8. P. 1003-1014.
19. Endsley J. J., Furrer J. L., Endsley M.A., McIntosh M.A., Maue A.C., Waters W. R., Lee D.R., Estes D.M. Characterization of Bovine Homologues of Granulysin and NK-lysin // *The Journal of Immunology*. 2004. V. 173. P. 2607-2614.
20. Gillet N.A., Gutierrez G., Rodriguez S.M., de Brogniez A., Renotte N., Alvarez I., Trono K., Willems L. Massive Depletion of Bovine Leukemia Virus Proviral Clones Located in Genomic Transcriptionally Active Sites during Primary Infection. // *PLoS Pathog*. 2013. V. 9. №10: e1003687. doi:10.1371/journal.ppat.1003687

УДК 575.174.015.3:599.723

### **Геномное сканирование с использованием инвертированных повторов микросателлитов (GAG)<sub>6</sub>C, (AG)<sub>9</sub>C**

- <sup>1</sup>Татьяна Теодоровна Глазко
- <sup>2</sup>Глеб Юрьевич Косовский
- <sup>3</sup>Светлана Николаевна Ковальчук
- <sup>4</sup>Борис Леонтьевич Зыбайлов
- <sup>5</sup>Валерий Иванович Глазко

<sup>1, 2, 3, 5</sup> Центр экспериментальной эмбриологии и репродуктивных биотехнологий, Российская Федерация

127422, Москва

<sup>1, 5</sup> РГАУ – МСХА имени К.А. Тимирязева, Российская Федерация

127550, Москва, ул. Тимирязевская, 49

E-mail: vigvalery@gmail.com

<sup>4</sup> Университет Арканзаса по медицинским наукам (UAMS), США

4301 W Markham St, Little Rock, AR 72205

<sup>1</sup> Доктор биологических наук, профессор

<sup>2</sup> Доктор биологических наук, член-корреспондент Российской академии естественных наук, директор центра экспериментальной эмбриологии и репродуктивных биотехнологий

<sup>3</sup> Кандидат биологических наук, доцент

<sup>4</sup> Кандидат биологических наук

<sup>5</sup> Доктор биологических наук, профессор, академик Российской академии естественных наук, академик Российской академии наук (иностраный член)

E-mail: vigvalery@gmail.com

**Аннотация.** Впервые выполнено секвенирование фрагментов геномной ДНК крупного рогатого скота черно-пестрой голштиinizированной породы длиной около 550 пар нуклеотидов, фланкированных инвертированными повторами микросателлитами (GAG)<sub>6</sub>C и (AG)<sub>9</sub>C, произведен поиск их гомологии в референсном геноме крупного рогатого скота в базе GenBank. В большинстве случаев гомологичные (GAG)<sub>6</sub>C последовательности локализуются в генах иммунной и клеточной сигнальной систем или на их 5' флангах в межгенном пространстве. Наиболее часто в исследованных секвенированных последовательностях фланкированных инвертированным повтором (AG)<sub>9</sub>C встречались участки гомологии с длинным диспергированным ядерным элементом LINE-1, видоспецифичным для *Bos taurus* – L1-BT, или продукты рекомбинации между ними. Это позволяет сделать вывод о том, что микросателлит AG тесно связан именно с этим продуктом рекомбинации. Полученные данные свидетельствуют о том, что микросателлиты, даже с близкими значениями коровых мотивов, такие как (GAG)<sub>6</sub>C и (AG)<sub>9</sub>C, существенно отличаются друг от друга по ассоциациям с разными геномными элементами. Полученные данные необходимо учитывать при использовании разных микросателлитов в целях геномного сканирования. Обсуждается возможное значение для таких отличий предрасположенности исследованных микросателлитов к формированию различных неканонических структур ДНК, таких как G4 квадруплексы, межмолекулярные (ДНК-РНК) триплексы.

**Ключевые слова:** микросателлиты, секвенирование, неканонические структуры ДНК, ретровирусы, пурин-пиримидиновые треки, триплексы, G4 квадруплексы.