

ДНК МАРКЕРЫ И «МИКРОСАТЕЛЛИТНЫЙ КОД» (обзор)

В.И. ГЛАЗКО, Г.Ю. КОСОВСКИЙ, Т.Т. ГЛАЗКО[✉], Л.М. ФЕДОРОВА

Одной из центральных проблем в современном животноводстве остается поиск генетических маркеров, которые упростили бы отбор и подбор животных для скрещиваний и повысили вероятность получения потомства с желательными хозяйственно ценными признаками. В обзоре рассмотрены примеры наиболее успешного применения различных типов ДНК маркеров для решения конкретных селекционных задач: микросателлитов — для исключения ошибок происхождения, мононуклеотидных полиморфизмов (SNP) — для создания карт геномных районов, полиморфизм в которых ассоциирован с модификацией фенотипических характеристик (D.J. Rigden, X.M. Fernández, 2023) и локализацией ключевых генов адаптации к давлению естественного отбора на краях ареалов и в зонах рискованного животноводства (Е.К. Cheruiyot с соавт., 2022; L. Buggiotti с соавт., 2021, 2022), областей с повышенным варьированием копийности геномных участков (CNV) — для анализа механизмов ответа полигенных систем (сенсорных, иммунных, транспортных) на воздействие факторов естественного и искусственного отбора (Y. Huang с соавт., 2021; P. Davoudi с соавт., 2022). Обсуждается преимущественное вовлечение диспергированных и тандемных повторов (в частности, микросателлитных) как элементов регуляторных сетей в эпигенетическую и фенотипическую изменчивость (R.P. Kumar с соавт., 2010). Рассматривается сложная структурно-функциональная организация микросателлитов, индивидуальная изменчивость некоторых локусов, их участие в процессах эволюции, рекомбинации, транскрипции, в формировании вторичной структуры ДНК, модификации архитектоники интерфазного ядра, в регуляции профилей геной экспрессии (R.P. Kumar с соавт., 2010; X. Tang с соавт., 2022). Исследования регуляторных сетей приобретают особое значение в связи с накоплением информации о существенных различиях по размерам генома у животных разных таксонов, распространенности и составу мобильных генетических элементов (источников компонентов регуляторных сетей) при сходстве числа генов, кодирующих белки (В.И. Глазко с соавт., 2022). Эти данные свидетельствуют о перспективности полилокусного генотипирования на основе микросателлитных и диспергированных повторов для выявления особенностей популяционно-генетической структуры, консолидированности и различий между группами животных.

Ключевые слова: ДНК маркеры, микросателлиты, short tandem repeat, STR, мононуклеотидные полиморфизмы, single nucleotide polymorphism, SNP, изменчивость числа копий, copy number variations, CNV, полногеномные исследования ассоциаций, genome-wide association study, GWAS.

Успешность селекционной работы определяется качеством отбора и подбора животных для скрещиваний, что непосредственно зависит от возможностей прогнозировать проявление желательных хозяйственно ценных признаков в конкретных условиях окружающей среды. Идея отбора с помощью маркеров впервые была выдвинута советским генетиком А.С. Серебровским, который ввел понятие сигнального гена. Согласно этим представлениям, так называемые сигналии — это удобные для менделистических наблюдений альтернативные гены с более или менее известной хромосомной локализацией, которые, не оказывая прямого воздействия на изучаемый трансгрессирующий признак и влияя достаточно определенным образом, облегчают генетический анализ этого признака, позволяя следить за наследованием того участка хромосомы, в котором эти сигналии расположены (1). В дальнейшем в качестве маркеров у животных рассматривали антитела к белковым антигенам биообъектов, затем и у растений, и у животных — полиморфизм белковых продуктов одних и тех же генов, выявляемый по электрическому заряду молекул (биохимические маркеры) (2).

Следующим этапом стала разработка ДНК маркеров различных типов. По общему определению ДНК маркеры — это маркеры полиморфизма

геномных участков, изменчивость которых потенциально может быть ассоциирована с ее проявлениями на фенотипическом уровне. Однако по мере углубленного изучения некоторых ДНК маркеров различных геномных элементов становится очевидной их вовлеченность в сложные сети взаимосвязей между ними, в обеспечение устойчивости, изменчивости и эволюции биологических систем.

В настоящем обзоре мы представили примеры эффективности ДНК-маркеров разных поколений как инструментов в прикладных исследованиях и проанализировали данные о микросателлитных (short tandem repeat, STR) ДНК-маркерах. Рассмотрено разнообразие биологических функций STR, включая, в частности, участие в формировании вторичных структур ДНК, архитектоники интерфазного ядра, элементов регуляторных сетей генной экспрессии, что, по нашему мнению, изменяет представления об их биологической роли и возможностях практического применения,

Прикладное применение ДНК маркеров. *Типы ДНК маркеров и их значение для генетических и геномных исследований.* С 1980-х годов для генетического маркирования начали использовать полиморфизм геномных участков. В 1980-х годах наиболее распространенными были ДНК маркеры, основанные на полиморфизме длин фрагментов рестрикции геномной ДНК (restriction fragment length polymorphism, RFLP), в 1990-х — маркеры, выявляемые с помощью полимеразной цепной реакции (селекция с помощью маркеров, marker-assisted selection, MAS), после 2000 года широкое распространение получили ДНК-чипы для обнаружения однонуклеотидных полиморфизмов (single nucleotide polymorphism, SNP) при полногеномном секвенировании.

Полилокусное генотипирование по ДНК маркерам (геномное сканирование) у сельскохозяйственных видов широко применяется для определения параметров изменчивости внутри и между породами; при идентификации популяционно-генетических особенностей у географически разобщенных групп и/или при смешивании в группах животных разного происхождения; для изучения эволюционных взаимоотношений, поиска центров происхождения и маршрутов миграции; для картирования главных генов, полиморфизм которых ассоциирован с изменчивостью фенотипических характеристик (включая идентификацию известных аллелей в случае генетически детерминированных заболеваний и выявление их носителей, а также обнаружение аллелей, связанных с повышенной устойчивостью к инфекционным и неинфекционным болезням). Для полилокусного генотипирования широко используются микросателлитные маркеры — тандемные повторы, элементарная (повторяющаяся) единица которых может быть длиной от 2 до 6 п.н. (простые, или короткие тандемные, повторы, short tandem repeat, STR) (3).

По согласованию с Организацией по продовольствию и сельскому хозяйству при ООН (Food and Agricultural Organization, FAO) и Международному сообществу по генетике животных (International Society of Animal Genetics, ISAG) для сельскохозяйственных животных первоначально были разработаны панели генотипирования по группам крови, затем по биохимическим маркерам, в настоящее время — по полиморфизму микросателлитов. Эти панели специфичны для каждого вида, включают пару десятков локусов и разработаны таким образом, что позволяют проводить ПЦР одновременно по нескольким маркерам и решать ряд текущих проблем (идентификация происхождения, выявление породных популяционно-генетических особенностей животных). На протяжении более чем двух десятилетий генетическая паспортизация животных, необходимая для исключения оши-

бок происхождения, осуществлялась с помощью таких панелей высокополимерных маркеров. Степень их гетерозиготности, несмотря на частые близкородственные скрещивания у сельскохозяйственных животных, нередко составляет 75 % и более. Обычно для паспортизации используют динуклеотидные и тринуклеотидные повторы. Наряду с ними, благодаря расширению полногеномного секвенирования представителей разных видов появляются новые поколения ДНК маркеров, как правило, основанные на SNP.

К настоящему времени более 30000 быков голштинской породы генотипированы с использованием ДНК-микроматриц BovineSNP50 BeadChip («Illumina», США), позволяющих анализировать одновременно 54001 SNP (примерно один SNP на 50000 п.н.). С помощью таких микроматриц для разных видов построены карты распределения SNP по геномам (4) и созданы полногеномные карты ассоциаций сайтов локализации SNP с изменчивостью фенотипических характеристик (genome-wide association study, GWAS) (5, 6).

ДНК-микроматрицы (чипы) позволяют выявить не только сцепленные с SNP маркеры, но и изменчивость участков их локализации по числу копий (copy number variations, CNV), включая делеции, дупликации, транслокации и инверсии (7). CNV привлекает все большее внимание, поскольку достаточно часто оказывается связанной с вариабельностью фенотипических признаков у сельскохозяйственных видов животных (8-13) и с неблагоприятными фенотипическими проявлениями у человека (7). Созданы подробные хромосомные карты распределения CNV маркеров. В частности, у человека локусы с CNV покрывают 12 % генома (Database of Genomic Variants, <http://projects.tcag.ca/variation/>), то есть изменчивость CNVs вовлекает больше нуклеотидов в расчете на геном, чем SNPs (14). Предполагается, что спонтанные CNV возникают в среднем с частотой 10^{-4} п.н. (15). Высокий уровень изменчивости дает основание полагать, что применение CNV маркеров полиморфизма участков геномной ДНК повысит точность картирования главных генов таких хозяйственно значимых характеристик животных, как устойчивость к действию биотических и абиотических факторов среды и продуктивные качества (16-18).

Таким образом, следует отметить успешное развитие методов секвенирования биомолекул, особенно с помощью технологий четвертого поколения (nanopore-based sequencing) (19), а также полное прочтение геномов у большинства сельскохозяйственных видов животных, в том числе у млекопитающих (20). Однако направления практического применения огромного объема результатов этих генных и геномных исследований до сих пор недостаточно разработаны. Требуется соотнести накопленные за десятилетия данные о генотипах представителей разных пород по микросателлитам (STR) и результаты генотипирования с помощью SNP панелей, которое в настоящее время (как ранее STR анализ) проводят для контроля происхождения, породной идентификации, выявления связей между генотипами и изменчивостью фенотипических признаков (21-24). Выяснилось, что, наряду с другими сложностями, STR панель, рекомендованная ISAG для генотипирования сельскохозяйственных животных, включает микросателлиты, которые существенно отличаются друг от друга по полиморфизму и эффективности дифференциации животных внутри и между породами (21). Изучение колокализации STR и SNP в геноме крупного рогатого скота показало, что только 57,1 % STR находятся в неравновесии по сцеплению с SNP (linkage disequilibrium, LD), тогда как остальные 42,9 % расположены

вне таких блоков (25). Иными словами, в геноме крупного рогатого скота значительное число STR не сцеплено с SNP (вероятно, из-за особенностей механизмов мутаций и геномного распространения STR, а также их повышенного полиморфизма). Из этого следует, что каждый микросателлит имеет свои особенности полиморфизма, мутабельности, и объединение разных микросателлитов в общую панель может привести к ошибочным выводам.

Распространенная замена полногеномного секвенирования (whole genome sequencing, WGS) сравнительно ограниченным числом SNP на ДНК-чипах в целях выявления SNP, ассоциированных с изменчивостью фенотипических характеристик, также достаточно часто несет в себе источники существенных ошибок (26). Как и в случае STR, возможность использования ограниченного числа SNP для прогноза изменчивости фенотипических признаков будет зависеть от локализации SNP в различных геномных элементах, их структурно-функциональных особенностей (26).

Развитие нового направления — пангеномики (в частности, пангеномики крупного рогатого скота, КРС) благодаря накоплению данных WGS по многим геномам позволяет добавить к референсному геному КРС, представленному в GenBank (*Bos taurus*, <https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/9913/>), около 4 % нуклеотидных последовательностей (27). Это может влиять на смещение позиций анализируемых аллельных вариантов SNP. Так, у крупного рогатого скота дополнительно к эталонному геному выявлено 83250 SNP, по которым отмечается полиморфизм как внутри пород, так и между породами (27).

Поиск ДНК маркеров хозяйственно ценных признаков. В последнее десятилетие развитие GWAS позволило продвинуться в понимании генетической основы сложных признаков и болезней как у человека, так и у домашнего скота (28). При этом интересен тот факт, что большинство SNP, ассоциированных с фенотипической изменчивостью свойств, локализуется в некодирующих последовательностях генома (28). Во многих случаях такие последовательности тесно связаны с многочисленными регуляторными элементами (RE), влияющими на профили генной экспрессии (29). Однако их эволюционная сохранность, изменчивость, вовлеченность в проявление сложных полигенных признаков до сих пор остаются недостаточно изученными.

Полиморфизм STR или SNP может быть использован для картирования геномных участков локализации геномных элементов, изменчивость которых вносит существенный вклад в проявление количественных хозяйственно ценных признаков у животных сельскохозяйственных видов, и обнаружения ключевых генов и/или RE таких признаков. Как правило, эти геномные элементы выявляются у животных, находящихся в экстремальных условиях обитания. Ярким примером успешных поисков служат работы по изучению устойчивости молочной продуктивности австралийских коров голштинской породы к высоким температурам (30, 31), а холмогорского и якутского крупного рогатого скота — к низким (32, 33).

Состав и функциональная организация геномных элементов, вовлекаемых в генетически детерминированный контроль молочной продуктивности у крупного рогатого скота, могут быть примером сложной генетической основы количественных хозяйственно ценных признаков. К настоящему времени создана хромосомная карта генов, участвующих в формировании вымени и продукции молока у крупного рогатого скота (так называемая карта лактома), которая включает 197 генов белков молока и более 6000 генов, участвующих в развитии и функционировании молочной железы (34). Оказалось, что эти гены рассеяны по всем 30 хромосомам круп-

ного рогатого скота. Сравнение геномов утконоса, опоссума, плацентарных млекопитающих (крупный рогатый скот, собака, человек, мышь, крыса) (34) по генам белков молока и формирования молочной железы выявило потери и дубликации, филогенетические связи, консерватизм последовательностей этих генов и их эволюцию. Получены данные о том, что у крупного рогатого скота гены белков молока и молочной железы эволюционируют медленнее, чем у других изученных видов плацентарных. Обнаружено, что у крупного рогатого скота по сравнению с другими перечисленными видами наиболее дивергировали гены белков молока, определяющие его питательные и иммунные свойства, наиболее консервативными оказались гены, ассоциированные с процессами секреции молока.

Анализ транскриптома на разных стадиях лактации показал, что 16892 гена экспрессируются в промежуточный период лактационного цикла, 19094 — на пике лактации и 18070 — при затухании лактации. Уровень экспрессии генов, кодирующих казеины, сывороточные белки и ферменты метаболического пути синтеза лактозы, был повышен в начале лактации, большинства генов метаболических путей липидного обмена — в промежуточный период и на пике лактации (35).

Очевидно, что на молочную продуктивность влияют генетические особенности особи, эпигенетические процессы, питание, патогены, климатические условия и другие внешние факторы. Это особенно наглядно обнаруживается на примере различий селекционных индексов, рассчитанных для одних и тех же голштинских быков по молочной продуктивности их дочерей, рожденных в разных эколого-географических регионах — в Люксембурге и Тунисе (36).

За последнее десятилетие существенно возросло число работ по оценке влияния эпигенетической изменчивости, связанной с регуляторными сетями, которые представлены, в частности, микроРНК (miRNAs, небольшие некодирующие РНК с транскрипционными и посттранскрипционными эффектами). Все чаще выявляется прямое участие микроРНК в контроле профилей генной экспрессии, в регуляции развития и функционирования молочной железы (37, 38). Обнаружено, что микроРНК играют важную роль во многих процессах, связанных с развитием и болезнями молочной железы, а также в секреции молока. В молочной железе идентифицированы сотни микроРНК, однако число микроРНК, функции которых полностью известны, очень мало. Проблема в том, что одна микроРНК может участвовать в контроле сотни генов, поэтому функциональная валидация каждого гена-мишени для этой микроРНК затруднительна. Ситуация усложняется еще и тем, что ответ на воздействие одних и тех же факторов окружающей среды могут обеспечивать разные микроРНК, причем не только у близкородственных видов (39), но и у пород (40).

Разнообразие спектров микроРНК очень широко, как и их связь с регуляцией экспрессии генов разных метаболических путей и узлов их пересечения. Из 19994 пар ортологичных генов *Bos taurus* и вымершего вида *B. primigenius*, кодирующих белок, 1620 генов различаются по полиморфизму сайтов связывания микроРНК в 3'-UTR (41). В основном эти 1620 генов участвуют в контроле пигментации, репродукции, нейронных функций, общего метаболизма, иммунных реакций и в изменчивости признаков продуктивности животных, включая качество молока и эффективность использования корма.

Применимость ДНК маркеров для практической селекции. Одно из первых заключений, которое можно сделать на основании данных генети-

ческих и геномных исследований сельскохозяйственных животных, по-видимому, следующее. Не нужно искать универсального метода для решения одновременно всех задач селекционной работы. Если панели STR показали успешность для исключения ошибок происхождения, нет смысла заменять их на более сложные и дорогостоящие тест-системы, основанные на SNP. Тем более что уже сейчас ведется поиск универсальных панелей STR, ортогональных у разных видов, с желательным уровнем полиморфизма (42), которые позволят дифференцировать видовую и внутривидовую изменчивость биологических объектов (43).

Очевидно, что карты SNP незаменимы при поиске геномных районов, в которых локализованы гены устойчивости к действию относительно критических биотических и абиотических внешних факторов. На основе их уточнений и расширения объемов картирования будет формироваться экологическая геномика как продолжение экологической генетики — научного направления, заложенного еще А.С. Серебровским (1).

SNP-анализ важен для поиска генов-кандидатов хозяйственно ценных признаков, когда их проявление существенно различается у исследуемых животных, поскольку в этом случае можно вести поиск небольшого числа генов, обуславливающих такие различия. Трудно переоценить успешность применения SNP для реконструкции генетической динамики популяций на основе анализа протяженных гомозиготных областей (runs of homozygosity, ROH) (44) или при поиске мутаций генов, критичных для воспроизводства животных (loss-of-function или гаплотипы фертильности) (45).

Значительный объем данных о распределении SNP позволил выполнить полногеномный анализ их ассоциаций (GWAS) с различными признаками у видов растений и животных (46). Результаты представлены на ресурсах Китайского национального центра данных по геномике (National Genomics Data Center, China National Center for Bioinformatics) (47) и в коллекции баз данных журнала *Nucleic Acids Research* (47). В то же время, если речь идет о таком комплексном признаке, как молочная продуктивность, сложно ожидать очевидного и надежного успеха (49). L. Flori с соавт. (49) сообщают об обнаруженных ассоциациях между гаплотипами SNP (при неравновесном сцеплении между ними) с изменчивостью параметров молочной продуктивности у трех основных молочных пород Франции — голштинов, нормандов и монтбельярдов. В районах повышенной плотности таких гаплотипов у трех пород суммарно выявили 40 генов, в основном различающихся у исследованных пород. Возможно, наблюдаемые противоречия обусловлены эпистатическими взаимодействиями между генными ансамблями, вовлеченными в проявление таких хозяйственно ценных признаков, как величина общего удоя, процент и количество молочного жира и белка. При GWAS-анализе с использованием 76109 SNP у 294079 голштинских коров первой лактации оценили влияние попарного эпистаза на показатели молочной продуктивности и воспроизводства (общий удой, выход молочного жира, белка, процент молочного жира и белка, коэффициент стельности дочерей) (50). Из 50000 лучших выявленных эффектов парного эпистаза по каждому признаку пять связаны с большими хромосомными областями с внутривнутрихромосомным эпистазом (50). По сути этим можно объяснить хорошо известные нежелательные корреляции между общим удоем и белковостью молока, жирномолочностью молока и воспроизводством у коров.

Наглядной демонстрацией сложности выявления генных элементов, полиморфизм которых ассоциирован с изменчивостью хозяйственно цен-

ного признака, служит локус количественного признака на 18-й хромосоме (BTA18), связанный с легкостью отела и мертворождением у голштино-фризского крупного рогатого скота и его помесей (51). Этот факт известен более 20 лет, но его генетическая основа до сих пор не идентифицирована. Для ее выявления на основе генотипирования 2697 голштино-фризов выполнили подробный анализ соответствующего участка BTA18 и оценку неравновесия по сцеплению в этом районе. В результате была подтверждена связь полиморфизма с описанной патологией, идентифицированы 4 SNP с почти идеальным неравновесием по сцеплению с ней, однако не выявлен ни один ген-кандидат, ассоциированный с указанной патологией. При этом обнаружено обилие сегментного дуплицирования внутри региона и вокруг него (51).

Метод генотипирования по CNV маркерам, по-видимому, очень эффективен при выявлении физиологических систем, вовлекаемых в направление отбора, прежде всего на межвидовом уровне. Косвенным подтверждением этого предположения служит тот факт, что у сельскохозяйственных животных CNV чаще всего выявляют в случае генов, участвующих в различных функциях иммунной системы (52-55).

МикроРНК как элементы регуляторных сетей привлекают внимание в связи с их вовлеченностью в контроль и модуляцию функциональной активности генов (37-40). Однако нужно учитывать, что степень изменчивости уровня экспрессии у разных генов неодинакова. Выполненный нами анализ профилей генной экспрессии в печени свиней (56) выявил две группы генов — с наличием и отсутствием индивидуальных различий между исследованными животными. У свиней с использованием базы данных о метаболических путях KEGG Metabolic pathway (<https://www.genome.jp/kegg/pathway.html>) мы оценили вовлеченность в разные метаболические пути для 17 генов с близкими значениями экспрессии у особей (эти гены условно обозначили как группу с конститутивной экспрессией) и 18 генов, составляющих вариабельную часть профилей генной экспрессии с выраженными индивидуальными различиями между животными. В нашем эксперименте продукты 17 генов с конститутивной экспрессией участвовали в 25 метаболических путях, или в среднем 1,5 пути на один ген, а на каждый из 18 генов с вариабельным уровнем экспрессии приходилось по 3 метаболических пути (56). То есть чем больше число метаболических путей, в которые вовлечен продукт гена, чем сложнее объединяющие и контролируемые их потенциальные регуляторные элементы, тем выше индивидуальная изменчивость экспрессии гена. Таким образом, хотя роль микроРНК в регуляторных сетях бесспорна (как и роль самих регуляторных сетей в формировании признаков и изменении их проявлений под действием факторов влияния), перспективность микроРНК в качестве молекулярных маркеров неоднозначна. Как уже отмечалось выше, одна микроРНК может влиять на транскрипты нескольких десятков генов, разные микроРНК — на активность одного гена (37-40). При этом тысячи факторов регуляции транскрипции также изменяют ее, взаимодействуя с некодирующими нуклеотидными последовательностями гена (40, 41). Кроме того, взаимодействие продуктов генов одного метаболической пути и/или разных метаболических путей тоже могут активировать процессы, приводящие к изменению профилей генной активности (40, 41).

Микросателлиты, структурное и функциональное разнообразие. *Структурно-функциональные особенности STR маркеров.* Микросателлиты относятся к первому поколению ДНК маркеров, которые уже более 30 лет широко используются в генетике и геномике человека, сель-

скохозяйственных животных и растений (42). Накопленные за этот срок экспериментальные данные позволили составить достаточно полное представление о сложной структурно-функциональной организации ДНК маркеров этого типа. К настоящему времени уже исследованы видоспецифичные особенности числа и геномного распределения микросателлитов. Созданы карты распределения микросателлитов в геномах разных видов (57).

На примере распределения STR в геноме домашнего кролика (*Oryctolagus cuniculus*) можно отметить выраженное различие по частоте встречаемости STR в связи с длиной элементарного повтора: 579097 моно-, 927755 ди-, 122482 три-, 767458 тетра-, 614173 пента- и 1739144 гексануклеотидных повторов (57). Большое число гексануклеотидных повторов объясняется тем, что у млекопитающих консервативный теломерный повтор TTAGGG продублирован несколько тысяч раз (58). Интересно, что суммарно в геноме и кролика, и ряда других видов млекопитающих, включая человека, тринуклеотидных повторов существенно меньше, чем ди-, тетра- и пентануклеотидных, то есть это различие не связано с длиной элементарной единицы повтора (57). Вероятно, потому, что структура тринуклеотидных повторов соответствует принципу триплетности генетического кода, они испытывают давление отбора.

Есть выраженные различия между частотой встречаемости микросателлитов с одинаковой длиной элементарной единицы, но с разным коровым мотивом, например AG и AC. У кролика AG мотив встречается примерно в 5 раз чаще, чем AC, и чаще, чем все другие микросателлиты, тогда как у человека, наоборот, микросателлитов с коровым мотивом AC больше, чем с AG (57, 59).

У прокариот тоже встречаются STR, но с низкой частотой (56). У нематод и насекомых распространены тринуклеотидные STR, у рыб — динуклеотидные STR, относительный дефицит которых наблюдается у птиц (60). Близкородственные виды могут существенно различаться по числу микросателлитных локусов (60). Мы уже отмечали, что в целом у многих видов млекопитающих ди- и тетрануклеотидные мотивы STR более часты, чем тринуклеотидные, но частота последних существенно варьирует в зависимости от коровых мотивов микросателлитов, причем эти различия могут быть видоспецифичными (61).

Разные STR в целом чаще всего встречаются в межгенных пространствах, затем в интронах, в промоторных областях и реже всего в экзонах (56). Судя по базе данных микросателлитов у разных видов (56), в экзонах тринуклеотидные STR обнаруживаются чаще, чем ди- и тетрануклеотидные. Причем частота встречаемости тринуклеотидных STR в экзонах может быть достаточно высокой.

Несмотря на относительно пониженную частоту встречаемости STR в экзонах, их вклад в полиморфизм кодируемых белков может быть существенным. Так, недавно при сравнении результатов полногеномного секвенирования у лабораторных мышей (всего 71 линия) были идентифицированы аллели STR, присутствующие в кодирующих областях 562 генов, и представлены свидетельства того, что эти аллели могут менять фолдинг кодируемого белка и таким образом оказывать существенное влияние на его функцию (62).

Экспансии микросателлитов при патологиях и адаптациях. Известно, что многие патологии человека (чаще всего нейрокогнитивные и нейродегенеративные) связаны с полиморфизмом по числу копий (длинам) тринуклеотидных микросателлитов (63-66). Во многих исследованиях генетических основ заболеваний у человека в кодирующих участках генома обнаружена

амплификация триплетов (CCG)_n, (CGG)_n, (GCC)_n, (GCG)_n и (CAG)_n, которые обуславливают синтез соответственно полипролина, полиаргинина, полиаланина и полиглутамина (67). С развитием ряда заболеваний также связывают изменение длин динуклеотидных микросателлитов с коровым мотивом GA при их определенной геномной локализацией (67).

Изменчивость длин тринуклеотидных STR обнаруживается и при некоторых примерах адаптаций. Так, связь между увеличением копийности микросателлитов и повышением адаптивного потенциала обнаружена у гигантской панды при адаптации к пище, богатой углеводами, у белого и бурого медведей — к температуре местообитания (68). Изменчивость по STR, ассоциированную с указанными адаптациями, в основном выявили в регуляторных генах (например, гены факторов регуляции транскрипции, сигнальный путь, включающий рецептор инсулиноподобного фактора роста). Эти гены в основном вовлечены в два метаболических пути, участвующих в ключевых физиологических процессах (сердечно-сосудистая функция и регуляция энергетического обмена).

Наличие STR в кодирующих последовательностях может влиять на упаковку белка (фолдинг) и изменять его гибкость, что позволяет связываться с различными субстратами, будь то нуклеотиды, липиды или белки. Ожидается, что такие белки, содержащие аминокислотные повторы, кодируемые STR, участвуют в регуляции экспрессии генов, часто многофункциональны и обладают плейотропными эффектами, повышая устойчивость клеток и многоклеточных организмов к изменчивости факторов окружающей среды, что, по-видимому, оправдывает сложности, связанные с потенциальной высокой мутабельностью микросателлитов (69, 70).

Некодирующие STR тоже способны существенно влиять на фенотипическую изменчивость. Сообщалось (71), что у человека с присутствием STR связаны 10-15 % наследуемой изменчивости в экспрессии генов. Выявлены STR, аллели которых значительно различались между этническими группами, обнаружены 15 STR, у которых длина повтора коррелирует с уровнем экспрессии генов, два из которых (*Glutathione Peroxidase 7* и *Glutathione S-Transferase Mu 3*) участвуют в метаболизме глутатиона (72).

Обсуждается сложность механизмов влияния на транскрипцию и трансляцию при увеличении длин STR. Увеличение копийности тринуклеотидных STR в кодирующих последовательностях может приводить к патологиям из-за появления полиглутаминовых или полиаланиновых участков, неправильно локализованных в белках, что приводит, в частности, к нарушению межбелковых взаимодействий, многие из которых участвуют в регуляции транскрипции, репарации ДНК и/или препятствуют образованию молекулярных конденсатов (73). Образование конденсата определенного транскрипта с факторами регуляции транскрипции влияет на экспрессию ряда генов, а накопление аминокислотных повторов может приводить к нарушению такой конденсации и изменению регуляции транскрипции. Например, полиглутаминовый белок ТВР (TATA-box Binding Protein) связывается с TATA-боксом промоторов генов, чтобы инициировать транскрипцию, но когда длина полиглутаминового повтора увеличивается, изменяется его способность к конденсации с активаторами транскрипции, что приводит к нарушению регуляции транскрипции, наблюдаемому при многих полиглутаминовых заболеваниях (73). В других случаях, например с Ataxin-2, РНК-связывающий белок (RNA Binding Protein, RBP), участвующий в сборке конденсатов и стрессовых гранул и в процессинге РНК, несет полиглутаминовую последовательность. Увеличение ее копийности приводит к нейродегенеративным заболеваниям (Spinocerebellar ataxia type 2,

SCA2) (74, 75). Напомним, что РНК-связывающие белки играют существенную роль и в ядерно-цитоплазматическом транспорте, нарушение которого также ведет к нейродегенеративным заболеваниям (76).

Если для исследованных случаев, как отмечалось выше (57), в экзонах типична экспансия тринуклеотидных повторов, в интронах, промоторах, 3'- и 5'-UTRs обнаруживается изменчивость длин не только тринуклеотидных, но и тетра-, пента-, гекса- и декануклеотидных повторов (77). При амплификации STR возможно изменение метилирования соответствующего участка ДНК, само удлинение STR может изменять расстояние между регуляторными мотивами в промоторах, что будет существенно влиять на экспрессию (78). Транскрипция длинных STR приводит к формированию РНК-агрегатов, служащих ловушкой для белков, и к мультимолекулярным взаимодействиям с другими транскриптами, что, в свою очередь, способно повлиять на сплайсинг, профили генной экспрессии, РНК интерференцию (79-81).

STR могут транскрибироваться в смысловом и антисмысловом направлениях. Это приводит к появлению токсичных пептидов из-за трансляции с не-ATG иницируемых триплетов (repeat-associated non-ATG translation, RAN), транскрибируемых с STR (82).

В некоторых исследованиях отмечается, что многие патологии, связанные с амплификацией STR, выявляются в генах, кодирующих факторы регуляции транскрипции, что сопровождается нарушением формирования мультимолекулярных конденсатов и препятствует их взаимодействию с РНК-полимеразой II (77).

Особенности возникновения мутаций в STR. Обнаружено (83), что локализация STR и горячих точек рекомбинации в мейозе, которые обычно находятся в промоторах генов, часто совпадает. STR могут влиять на процессы рекомбинации в таких точках. Это показано для STR с коровыми мотивами GA, CA, GT, CT из-за их высокого сродства к ферментам рекомбинации (84). Частота мутаций по STR имеет широкий диапазон изменчивости — от 10^{-2} до 10^{-8} на локус на поколение, но она широко варьируется от локуса к локусу (85-87). Зависимость частоты мутаций STR от действия факторов окружающей среды была описана на модельных объектах (*Caenorhabditis elegans*) при сравнении мутабельности в лабораторных и естественных условиях (88).

В некоторых случаях обнаруживаются механизмы компенсации неблагоприятных последствий мутаций STR. Так, отмечается тесная связь между мутабельностью STR и полиморфизмом упаковывающих их белков хроматина, которые смягчают неблагоприятные влияния изменений в длине и организации STR на процессы, участвующие в подавлении экспрессии транспозонов, точной передаче хромосом и обеспечении их целостности (89).

Особое структурно-функциональное значение, в том числе для мутабельности, имеет способность STR формировать вторичные структуры ДНК. Формирование G4 квадруплексов в треках, обогащенных G/C, триплексов в пурин-пиримидиновых треках, R-петель (ДНК-РНК дуплексы с вытеснением второй цепи ДНК, которая некомплементарна РНК), других петлевых и шпилечных структур влияет на профили генной экспрессии, функцию ферментов репарации, работу ДНК-полимеразы, нестабильность STR (90-92).

STR могут приводить к появлению неканонических структур ДНК, отличающихся от классической правовращающей В-спирали, что определяется первичной нуклеотидной последовательностью. Например, левовращающая Z-ДНК содержит чередующиеся обогащенные GC пурин-пиримидиновые последовательности, при этом для образования и стабилизации Z-

формы ДНК решающее значение имеет суперспирализация. Считается, что Z-ДНК регулирует уровень суперспирализации и таким образом играет важную роль в транскрипции, экспрессии генов, рекомбинациях, транслокациях и делециях (93). Так, формирование Z-ДНК индуцирует нестабильность в области тринуклеотидных повторов (CAG, CGG и GAC), которые связаны с различными нейродегенеративными заболеваниями (93).

В появлении неканонических многоцепочечных структуры в областях пурин-пиримидиновых треков, состоящих из микросателлитов, например AG/TC, GAG/CTC, включают ДНК-ДНК взаимодействия с выделением одной нити ДНК, ДНК-РНК — с таким же эффектом и двуцепочечной РНК с РНК (94-96). Взаимодействие третьей нити с дуплексной ДНК или РНК специфичным для двуцепочечной последовательности образом, приводящим к образованию межмолекулярного триплекса, оказывает существенное влияние на транскрипцию, посттранскрипционные модификации, мутагенез (95).

Следует подчеркнуть, что все вторичные структуры по своей природе динамичны, при определенных условиях возникают, а при их изменении исчезают (93).

Особое внимание привлекают последовательности STR, предрасположенные к образованию крестообразных структур в районах локализации инвертированных повторов, так как многие белки, участвующие в контроле клеточного деления, например топоизомеразы, p53, Rif1 (Replication Timing Regulatory Factor 1), могут индуцировать формирование таких структур (97). Крестообразные формы ДНК играют важную роль в регуляции репликации, экспрессии генов, вовлечены в формирование структуры нуклеосом и рекомбинацию (98), служат мишенями для многих архитектурных и регуляторных белков, таких как гистоны H1 и H5, топоизомераза II β , белки HMG, HU, p53, протоонкогенный белок DEK (97, 98). Ряд ДНК-связывающих белков (например, члены семейства HMG-box, Rad54, белок BRCA1, а также полимераза PARP-1) предпочтительно связываются с крестообразными структурами (97). Предполагается, что по своей функции белки, взаимодействующие с крестообразными структурами, в основном подразделяются на четыре основные группы: топоизомеразы; белки репарации ДНК и факторы регуляции транскрипции; белки, участвующие в репликации; белки, ассоциированные с хроматином (98). Распространенность крестообразных структур, формируемых инвертированными повторами STR, их роль в эпигенетической регуляции и поддержании клеточного гомеостаза позволяют считать инвертированные повторы существенным компонентом регуляторных систем (98).

У млекопитающих описаны половые отличия по частоте мутаций STR. Поскольку ооциты в отличие от сперматозоидов, в частности у млекопитающих, находятся в состоянии покоя длительное время, мутации, возникающие в STR при гомологичных рекомбинациях, неравном кроссинговере, двуцепочечных разрывах, могут накапливаться и иметь более выраженные проявления (99). Обычно у мутаций STR, унаследованных по материнской зародышевой линии, частота немного выше, чем у унаследованных от отца (100). Однако с возрастом число мутаций STR в ооцитах практически не изменяется, тогда как в сперматозоидах увеличивается 2-кратно (исследования проведены в группе мужчин от 20 до 58 лет) (101).

Полное секвенирование геномов 544 человек из 29 семей в трех поколениях (база данных Центра изучения полиморфизма человека — Centre d'Etude du Polymorphisme Humain, CEPH, <https://uofuhealth.utah.edu/center-genomic-medicine/research/ceph-resources>) показало высокое разнообразие

новых типов мутаций STR, возникающих в разных семьях и в разных локусах STR (102). Обнаружена связь между длиной повтора и частотой мутаций (за исключением мононуклеотидных STR). В экзонах выявлена наименьшая частота мутаций (всего две в тринуклеотидных STR), их наибольшая часть (53,38 %) приходится на межгенные области, чуть менее половины (44,87 %) находится в интронах, около 1,6 % — в 5'- и 3'-UTR. В среднем рассчитанная частота мутаций ($5,24 \times 10^{-5}$) соответствует таковой для других типов мутаций, включая усредненные значения для SNP, однако в этом семейном анализе по данным из CEPH не были выявлены корреляции между появлением новых аллелей по STR и SNP. Оказалось, что приблизительно 30 % мутаций STR происходят в элементах Alu (short interspersed element, SINE), составляющих 11 % генома, тогда как во вставках LINE-1 (long interspersed element-1, LINE-1, или L1), охватывающих 17 % генома человека, находится только 10 % мутаций STR (102). То есть при семейном анализе трех поколений выяснилось, что пятая часть всех новых мутаций STR возникает в последовательностях ретротранспозонов.

Вовлеченность микросателлитов в геномную изменчивость и организацию интерфазного ядра. *Связь между микросателлитами и мобильными генетическими элементами.* Тесные взаимоотношения между STR и мобильными генетическими элементами — транспозонами (TE) выявлены достаточно давно (103). Многие микросателлиты возникли в результате геномных встроек TE. Предполагается, что это может быть результатом ряда событий: tandemных вставок TE в определенные участки генома хозяина (дуплицированные сайты-мишени, target site duplication, TSD), наличия в TE прямых и инвертированных повторов и взаимодействий между ними, захватов последовательностей генома хозяина на флангах TE, рекомбинаций между различными TE. Широкие взаимодействия TE друг с другом приводят к заключению о том, что они могут быть реконструированы в повторяющиеся некодирующие или кодирующие последовательности. Это предполагает эволюционную связь между такими последовательностями ДНК и то, что эволюция геномов включала частые перетасовки повторяющихся последовательностей — процесс, который обозначают как ремоделирование ДНК (103-105). Множественные вставки TE обусловлены наличием в геномах-мишенях предпочтительных мест для такой интеграции, благодаря чему появляются новые продукты рекомбинаций TE, образующиеся с высокой скоростью в периоды активной транспозиции. Иными словами, сам процесс транспозиций TE регулярно генерирует последовательности, из которых могут возникать новые микросателлиты (58, 105).

Так, в геноме человека широко распространены неавтономные короткие диспергированные ретротранспозоны Alu (SINE), содержащие поли(A)-хвост и центральную линкерную область, богатую аденинами. Показаны значительные связи между 3'-концами последовательностей Alu не только с мононуклеотидными повторами (A)_n, но также с (AAC)_n, (AAT)_n и тетра- и гексануклеотидными повторами, обогащенными А-нуклеотидами, тогда как для динуклеотидных повторов (AT)_n такая связь существенно слабее (106). С элементами Alu также предпочтительно ассоциирована локализация динуклеотидных повторов (AC)_n, причем 75 % из них выявлены на 3'-конце элемента, в то время как остальные — в центральной области. Интересно, что (GAA)_n — тринуклеотидный повтор, амплификация которого наблюдается при атаксии Фридрайха, возможно, возник при участии элемента Alu. Из 788 локусов генома человека, содержащих повторы (GAA)_n, 63 % (501 локус), имеют гомологию с Alu длиной не менее 25 п.н. Среди

них 94 % связаны с поли(А)-хвостом, а остальные — либо с 5'-концом элемента, либо с центральной областью (106). У исследованных видов *Carnivore* выявлено несколько сотен возникших из tRNALys SINE, которые содержат микросателлитные повторы, преимущественно обогащенные AG и A (107). У ряда видов, включая морского котика (*Phoca vitulina concolour*), крупный рогатый скот (*Bos taurus*), на флангах наиболее распространенного автономного TE — L1, обнаруживаются микросателлитные последовательности CA/GT разной длины (108).

Известно, что в геномах TE формируют участки преимущественной локализации — так называемые гнезда (109, 110), причем и сами TE, и сайты их интеграции содержат микросателлитные последовательности, благодаря чему здесь происходят рекомбинации и возникают новые варианты TE.

Важно подчеркнуть, что способность к горизонтальному переносу между разными таксонами описана для некоторых TE (например, для L1 и BovB). Особенно часто они встречаются у млекопитающих, причем их представленность в геноме может существенно различаться даже у относительно близких по происхождению групп животных (111).

Для некоторых STR показана прямая связь между их последовательностью и TE. Например, у крупного рогатого скота и овец микросателлит (AGC)_n встречается чаще, чем у других видов млекопитающих. В частности, у крупного рогатого скота представленность микросателлитных локусов с кором AGC соответственно в 90 и 142 раза выше, чем у человека и собаки (112). При этом в геноме крупного рогатого скота 39 % таких микросателлитных локусов непосредственно связаны с ретротранспозоном Bov-A2 (часть BovB) (112, 113) — эволюционно молодым и видоспецифичным для крупного рогатого скота. Интересно, что у крупного рогатого скота Bov-A2 функционирует как энхансер экспрессии гена интерферона II типа (113). Тесная связь (AGC)_n с Bov-A2 оказалась специфичной для этого STR и TE, примерно у 60 % других STR генетического сцепления с TE не выявили (112).

Таким образом, и сами TE, и их фланги содержат последовательности STR, в ряде случаев они тесно связаны друг с другом. Изменчивость STR и TE определяют различные события и механизмы (процессы рекомбинации, распространенность гнездового принципа локализации TE друг в друге), но каждый раз локусы STR и TE имеют выраженную индивидуальность по скорости и особенностям эволюции, что обусловлено, по-видимому, структурно-функциональными характеристиками и действием факторов отбора.

Микросателлиты, сайты повышенной ломкости хромосом — эволюция кариотипа. У человека описано около 230 участков с потенциально повышенной хрупкостью, содержащих STR (114). Выявлены 10 хрупких участков хромосом человека, для которых показана экспансиями геноспецифичных tandemных повторов с коровыми мотивами CGG и CCG. Авторы этого исследования полагают, что увеличение длин STR может приводить к появлению новых сайтов повышенной ломкости (114).

Род туко-туко (*Ctenomys*) подземных грызунов (*Rodentia: Ctenomyidae*) насчитывает около 65 видов, у которых обнаруживаются наиболее значительные хромосомные вариации среди млекопитающих (от $2n = 10$ до $2n = 70$). Причем кариотипическая изменчивость возможна даже в пределах вида, например $2n$ у *C. minutus* — от 42 до 50, у *C. talarum* — от 44 до 48 и у *C. lami* — от 54 до 58 (115). Среди них выделяется *C. minutus* с уже идентифицированными 45 различными цитотипами, из которых исходных предположительно семь (они распространены на прибрежных равнинах Южной

Бразилии). У туко-туко выполнено картирование повторяющихся участков ДНК, включая микросателлиты и LINE-1, и выявлена прямая связь между локализацией STR с разными коровыми мотивами и LINE-1, с одной стороны, и кариотипической изменчивостью (формированием цитотипов) — с другой в разных популяциях в пределах вида (115). Важно подчеркнуть, что описанные цитотипы включали не только робертсоновские транслокации, слияния и расщепления хромосом, но и тандемные повторы, парацентрические и перичентрические инверсии. Вовлеченность центромерных и теломерных повторов во внутривидовую изменчивость и формирование внутривидовых хромосомных рас хорошо известна у ряда видов землероек (116), три варианта внутривидовых хромосомных рас («робертсоновские веера») описаны Н.Н. Воронцовым (117) у мышей рода *Leggada*, домовых мышей *Mus domesticus* надвида *Mus musculus* и слепушонка группы *Ellobius tancrei*, относящейся к надвиду *Ellobius talpinus*. Описана вовлеченность STR и TE в кариотипическую изменчивость у рыб (118). Так, сравнительный анализ секвенированных геномов трех видов рыб показал, что коммерческий вид *Solea senegalensis* претерпел обширную хромосомную эволюцию, связанную с локализацией STR и TE в районах с повышенной плотностью хромосомных перестроек.

Вспышки стремительной эволюции кариотипа, часто называемые кариотипической мегаэволюцией, или хромосомной тахителией, обнаружены у разных таксонов (119). По-видимому, наиболее наглядным примером служат два вида оленей, кариотипы которых резко различаются, — индийский мунтжак *Muntiacus muntjak* ($2n = 6$) и китайский мунтжак *M. reevesi* ($2n = 46$). Сравнительный анализ секвенированных геномов мунтжаков, благородных оленей (*Cervus elaphus*) и крупного рогатого скота (*Bos taurus*) подтверждает эволюционную последовательность делений и слияний хромосом, описанную цитогенетически. Обнаружено, что с момента расхождения видов оленей и крупного рогатого скота (около 20 млн лет назад) стремительная эволюция кариотипа индийского мунтжака не сопровождалась крупными инверсиями или другими внутренними перестройками (за исключением дискретных событий расщепления и слияния хромосом) (119).

Сравнение геномов на уровне хромосом, выполненное для *Hydropotes inermis* (водяной олень, $2n = 70$), *Muntiacus reevesi* ($2n = 46$), самок и самцов *M. crinifrons* (черный мунтжак, $2n$ 8 или 9) и *M. gongshanensis* (олень Гонгшанских гор, $2n$ 8 или 9) (120), позволило авторам прийти к выводу, что уникальные центромерные сателлитные повторы, включая STR, теломерные STR и палиндромные повторы могли быть причиной повторяющихся слияний хромосом у оленей этих видов (120).

В целях реконструкции кариотипов 16 филогенетических узлов млекопитающих, в том числе кариотипа их общего предка, выполнены масштабные исследования геномных последовательностей у 32 видов, относящихся к эвтериевым (19 отрядов), сумчатым и однопроходным (по 3 отряда) как представителей трех надотрядов млекопитающих — соответственно *Euarchonta*, *Laurasiatheria* и *Xenarthra* (121). Современными видами, у которых оценили число хромосомных перестроек по сравнению с предполагаемым общим предком млекопитающих, были человек, ленивец (*Choloepus didactylus*) и крупный рогатый скот. Полученные данные дают основание предположить, что у общего предка млекопитающих, вероятно, было 19 пар аутосом. При этом девять из самых маленьких хромосом были общими у общего предка млекопитающих и общего предка всех амниот (из них три хромосомы все еще сохраняются у существующих млекопитающих) (121).

Определено число и типы хромосомных перестроек для переходов между кариотипом предков млекопитающих, потомственными предками и существующими видами. Выделены общие районы повышенной скорости эволюционных преобразований хромосом (evolutionary breakpoint regions, EBRs) и эволюционно консервативные блоки (homologous synteny blocks, HSBs) (121). Оказалось, что EBR участки отличаются от HSB повышенной плотностью активно транскрибируемых генов и повторяющихся элементов. Отмечается неслучайное распределение EBR в геномах и их связь с фрагильными сайтами при онкогенезе (121). Высокая плотность экспрессирующихся генов, выявленная в EBR, как предполагают авторы, может объяснять повышенную склонность к разрывам двойных цепей ДНК (ДНК в открытых транскрипционно активных участках хроматина более чувствительна к повреждениям). Анализ показал, что районы EBR обогащены генами, продукты которых участвуют в сенсорном восприятии и в регуляции транскрипции, тогда как в блоках HSB повышена плотность генов, вовлеченных в формирование анатомических характеристик и развитие центральной нервной системы (121). В EBR значительно выше плотность повторов всех типов, сегментных дупликаций, SINE (SINE; все SINE и Alu), LINE (LINE; L1) и длинные концевые повторы (LTR; все LTR и эндогенный ретровирус 1 — ERV1), чем в HSB (121).

По данным этих же авторов (121), эволюционная история хромосом предков млекопитающих (mammalian ancestor chromosomes, MAM) значительно варьировала в зависимости от размера хромосом. Большие длинные MAM чаще участвовали в хромосомных перестройках, чем короткие, у некоторых существующих видов (например, у *Mus musculus*, *Equus caballus*, *Canis familiaris*, *Bos taurus* и *Capra hircus*) сохранялась синтения генов на уровне хромосом. Девять из 14 маленьких MAM у перечисленных видов млекопитающих оказались ортологичными по синтении генов с хромосомами *Gallus gallus* и реконструированными хромосомами предков птиц и амниот. Некоторые MAM сохранялись как отдельные хромосомы или как замкнутые единицы (то есть целые хромосомы, слитые с одной или несколькими хромосомами без разрыва синтении) в геномах млекопитающих. Например, MAM7 сохранялся в виде целой хромосомы у *Oryctolagus cuniculus*, *Rhinolophus luctus* и *Procapra capensis*, которые представляют три отряда млекопитающих (122). MAM13 и MAM14 присутствуют как отдельные хромосомные единицы у более чем 15 сохранившихся видов млекопитающих (121).

В совокупности эти результаты демонстрируют поразительное сохранение синтении в течение примерно 320 млн лет эволюции позвоночных со времен общего предка всех амниот. Реконструированный геном предков млекопитающих показал, что существующие геномы млекопитающих представляют собой мозаику, полученную в результате эволюционной перетасовки 2557 синтенных сегментов, что составляет от 69 до 94 % от размера генома у проанализированных видов, причем в качестве связок между такими сегментами выступают EBR последовательности, обогащенные TE и STR (121). Анализируя полученные данные, авторы высказали предположение (121), что эволюционно консервативные синтенные сегменты — HSB служат основными строительными блоками (геномные элементы, аналогичные элементам периодической системы химических элементов) для геномов всех млекопитающих, сохраняющимися, наряду с синтенией, биологические функции (121). Подобное предположение сформулировано и в исследованиях особенностей организации и распространения микрохромосом у реп-

тилий, птиц и млекопитающих (122). Показано, что геномы птиц и рептилий, но не млекопитающих, состоят из нескольких крупных макрохромосом длиной от 3 до 6 мкм и множества крошечных (меньше 0,5 мкм) микрохромосом. Микрохромосомы имеют центромерные и теломерные участки, несут большое число генов, обогащены GC нуклеотидами, высококонсервативны среди птиц и рептилий, имеют гомологию с одной или несколькими крошечными хромосомами беспозвоночных, разошедшихся с позвоночными более 680 млн лет назад. Микрохромосомы ассоциируют друг с другом, бедны TE и собираются вместе в центре интерфазных ядер, что, по мнению авторов, указывает на функциональную когерентность. У черепах, змей и ящериц многие микрохромосомы исчезли вследствие слияния в макрохромосомы, у большинства млекопитающих микрохромосомы исчезли полностью, однако некоторые хромосомы утконоса совпадают с несколькими описанными ранее микрохромосомами рептилий и птиц. Это позволяет предположить, что такие хромосомы представляют собой строительные блоки хромосом млекопитающих, связь между которыми формируется с участием TE и STR (121, 122).

STR и TE принимают непосредственное участие в архитектонике интерфазного ядра. В интерфазном ядре хроматин организован в виде иерархии от нуклеосом до доменов хроматина (chromatin domains, CD), далее к формированию топологически ассоциированных доменов (topologically associated domains, TADs) и до компартментов более высокого уровня; вершина иерархии — так называемые территории хромосом (chromosome territories, CT) (123). Согласно современным представлениям, организация хроматина — критический фактор, регулирующий экспрессию генов (123-125). Так, энхансеры взаимодействуют с целевыми генами почти исключительно в пределах TAD, дистально расположенные гены, которые экспрессируются совместно, вовлекаются в общие белковые скопления при активации, компактные домены демонстрируют движение и конфигурационные изменения *in vivo* (124, 125).

Неслучайное радиальное позиционирование СТ в ядре указывает на возможность предпочтительных паттернов взаимодействия между хромосомными территориями. Обнаружена их способность формировать специфические межхромосомные сети, которые изменяются в течение клеточного цикла, при клеточной дифференцировке, при неопластической трансформации (125). Предполагается, что динамика этих сетей коррелирует с глобальным контролем структурных изменений и регуляцией функциональной активности генома. Склонность к различным транслокациям при патологиях может объясняться близким и в норме демонстрирующим специфичность расположением определенных хромосомных областей при коэкспрессии локализованных в них генов. Так, возможно, что для областей генома из районов EBR в основном типичен открытый эухроматин, что способствует эпигенетическим модификациям из-за доступности ДНК для регуляции и активной работе генов и/или их взаимодействиям (124).

Об участии STR и TE в регуляции смены программ генной экспрессии посредством динамических изменений в архитектонике интерфазного ядра сообщается во многих работах (123-125). По нашему мнению, наиболее наглядный пример — исследование, выполненное на интерфазных ядрах клеток цилиндрических фоторецепторов у ночных млекопитающих (125). Авторы выявили инверсию в локализации гетерохроматина и эухроматина по сравнению с ядрами клеток ганглиев: гетерохроматин размещался внутри, эухроматин — на периферии ядра под ламиной (125). Обычно SINE

ассоциированы с активно транскрибируемыми районами, LINE — с гетерохроматизированными, локализованными на периферии интерфазного ядра под ядерной оболочкой, и только в случае клеток цилиндрических фоторецепторов ночных млекопитающих LINE вместе с гетерохроматином концентрируется в центре интерфазного ядра, SINE — на периферии. По мнению авторов исследования (125), такое расположение гетеро- и эухроматина в значительной степени обусловлено контактами между гомологичными диспегирированными повторами, которые локализованы в разных районах хроматина.

Сформулировано предположение о том, что STR можно рассматривать в качестве компонентов так называемого «кода упаковки генома», определявшего особенности его конденсации в зависимости от типа клеток (126). Отмечено, что у некоторых организмов STR сгруппированы в определенных областях хромосом, например в перичентромерных или субтеломерных, в сложных геномах STR распределены по всей его протяженности неслучайным образом, располагаясь преимущественно в межгенных пространствах (57, 126). Известны несколько белков, которые специфически связываются с STR (125, 126). Таким образом, STR могут быть «якорями» для вовлечения групп локусов, с которыми они сцеплены, в процессы внутри- и межхромосомных взаимодействий, формирования TAD и CD. Участие STR в межклеточных взаимодействиях в сложных органах через предполагаемое влияние на архитектуру упаковки хроматина и, как следствие, на программы генной экспрессии, обсуждалась в ряде работ, в частности в случае микросателлита с коровым мотивом GATA у животных и растений (126, 128). Показано, что повторы GAGA, которые в избытке присутствуют в геноме эукариот, опознаются GAGA-ассоциированным фактором GAF, влияющим на упаковку хроматина (GAGA pioneer factor, GAF) у дрозофил, белками BBR (barley B recombinant protein) у ячменя, GBP (GAGA-binding protein) у сои и риса (126, 128).

Гипотеза о существовании такого «микросателлитного кода» программ генной экспрессии (126) также важна для понимания молекулярно-генетических механизмов формирования конвергентных признаков у эволюционно удаленных таксонов. Попытки найти общие структурные гены для таких признаков предпринимались многократно (129). В частности, проведено сравнение транскриптомов у восьми видов позвоночных (ящериц, млекопитающих, акул), которые вынашивают эмбрионы в матке (129). Оказалось, что во всех живородящих группах основной набор физиологических функций матки не различается, но при этом в одинаковом наборе генов ни один не экспрессируется специфично для всех живородящих линий или даже во всех линиях живородящих амниот, формирующих зародышевые оболочки (129). Таким образом, морфологические и физиологические признаки, необходимые для успешного течения беременности, у отдаленно родственных позвоночных, как оказалось, контролируются неодинаковыми генами. По всей видимости, эволюционные изменения живорождения как способа воспроизведения потомства происходили многократно с вовлечением разных генов благодаря кооперации и коллатеральности метаболических путей, однако набор таких генов все же исходно ограничен их составом у предков каждой линии (129).

Пример с живорождением может наглядно объяснить относительно низкую эффективность селекции с помощью маркеров (MAS), когда отбор и подбор животных для скрещиваний основывается на генотипах по небольшому числу генов, кодирующих белки (130). Именно поэтому особое

значение приобретает поиск ДНК маркеров регуляторных сетей, изменчивость которых лежит в основе организации профилей генной экспрессии. Накопленные данные, свидетельствующие о том, что, несмотря на хромосомные перестройки и структурные преобразования вследствие взаимодействия между макро- и микрохромосомами, сохраняется достаточно высокая эволюционная консервативность по генному составу хромосомных районов в TAD A (активно транскрибируемые домены интерфазного хроматина) и TAD B (гетерохроматизированные домены) (121, 123-125), что позволяет предполагать наличие спектра регуляторных элементов, вовлекаемых в такое подразделение. Одни из этих элементов — STR, участвующие в структурных взаимодействиях макро- и микрохромосом и в формировании архитектуры интерфазного ядра, что, в свою очередь, тесно связано с модуляцией профилей генной экспрессии.

Подводя итог обсуждению применения ДНК маркеров, важно отметить, что многие значимые для селекции вопросы уже решаются с помощью современных молекулярно-генетических методов. Ключевыми этапами селекционного процесса служат отбор и подбор животных для скрещиваний и оценка племенной ценности родителей по характеристикам потомства. ДНК маркеры позволяют исключить ошибки происхождения, облегчают выявление мутаций, связанных с фенотипическими и репродуктивными дефектами, устойчивостью к действию биотических факторов и экологических стрессов. Если желательное развитие хозяйственно ценных признаков контролируется небольшим числом ключевых генов, ДНК-маркирование применимо для поиска соответствующих аллельных вариантов, как правило, связанных с изменчивостью качества конечной продукции. Попытки обнаружить ассоциации между наборами SNP генотипов и изменчивостью фенотипических признаков не всегда успешны из-за сложной конструкции генных сетей, конкурентных взаимоотношений между молекулярно-генетическими структурами, которые служат мишенями для факторов естественного и искусственного отбора, вариабельности вклада элементов регуляторных сетей во взаимодействия между метаболическими путями в зависимости от генотипической среды и влияния внешних факторов. «Неуловимые главные гены» количественных признаков и различия в генетическом контроле сходных фенотипических характеристик — также примеры того, что далеко не всегда для успешной селекции достаточно ДНК-маркирования генов и локусов, ассоциированных с желательным признаком.

Отметим, что у видов животных число генов, кодирующих белки, сравнительно мало различается, тогда как вариации в размере геномов значительны и обусловлены различиями в распространенности в них диспергированных и тандемных повторов (110, 126). На тандемные повторы (в частности, на микросателлиты) в геноме млекопитающих приходится больше нуклеотидов, чем на гены, кодирующие белки (126). Подразделенность геномов различных таксонов животных на эволюционно консервативные (HSB) и эволюционно нестабильные (EBR) блоки генов, обогащенность последних диспергированными повторами (121) позволяют предполагать прямую вовлеченность диспергированных повторов и их производных, таких как STR, в регуляцию генных сетей. Широкая представленность микросателлитных повторов в геномах и многообразие их биологических эффектов выделяют эти ДНК маркеры в качестве элементов регуляторных сетей и, возможно, самостоятельных мишеней изменчивости и селекционного отбора. Полилокусное генотипирование именно этих геномных элементов позволит с высокой разрешающей способностью анализировать популя-

ционно-генетическую структуру групп животных, оценивать степень их консолидированности и выявлять отличия от близкородственных групп.

Итак, одной из центральных проблем в современном животноводстве остается поиск генетических маркеров, которые упростили бы отбор и подбор животных для скрещиваний и повысили вероятность получения потомства с желательными хозяйственно ценными признаками. Разработанные ДНК маркеры разных типов и поколений успешно применяются при решении ряда значимых для селекции вопросов: микросателлиты (STR) — для исключения ошибок происхождения, мононуклеотидные полиморфизмы (SNP) — для создания карт геномных районов, ассоциированных с фенотипическими характеристикам и адаптацией к давлению естественного отбора на краях ареалов и в зонах рискованного животноводства, области с повышенным варьированием копийности геномных участков (CNV) — для анализа механизмов ответа полигенных систем на действие факторов естественного и искусственного отбора. При небольшом числе основных генов, определяющих проявление признака, ДНК-маркирование применимо для поиска аллельных вариантов структурных генов, анализа изменчивости элементов регуляторных сетей и взаимосвязей между метаболическими путями. Многолетние исследования выявили множественную вовлеченность STR в базовые процессы (репликация, репарация, транскрипция, трансляция, адаптация и формообразование, эпигенетические эффекты), определяющие стабильность, изменчивость и эволюцию биологических систем. В этой связи STR могут рассматриваться как элементы регуляторных сетей — основных мишеней естественного и искусственного отбора. Полилокусное генотипирование на основе микросателлитных и диспергированных повторов представляется перспективным для анализа популяционно-генетической структуры и консолидированности групп животных и их отличий от близкородственных групп.

ФГБНУ НИИ пушного звероводства и кролиководства
им. В.А. Афанасьева,
140143 Россия, Московская обл., Раменский р-н, пос. Родники,
ул. Трудовая, 6,
e-mail: vigvalery@gmail.com, gkosovsky@mail.ru, tglazko@rambler.ru ✉,
felami@mail.ru

Поступила в редакцию
10 января 2023 года

Sel'skokhozyaistvennaya biologiya [Agricultural Biology], 2023, V. 58, № 2, pp. 223-248

DNA MARKERS AND MICROSATELLITE CODE (review)

V.I. Glazko, G.Yu. Kosovsky, T.T. Glazko ✉, L.M. Fedorova

Afanas'ev Research Institute of Fur-Bearing Animal Breeding and Rabbit Breeding, 6, ul. Trudovaya, pos. Rodniki, Ramenskii Region, Moscow Province, 140143 Russia, e-mail vigvalery@gmail.com, gkosovsky@mail.ru, tglazko@rambler.ru (✉ corresponding author), felami@mail.ru

ORCID:

Glazko V.I. orcid.org/0000-0003-4242-2239

Kosovsky G.Yu. orcid.org/0000-0003-1889-6063

The authors declare no conflict of interests

Final revision received January 10, 2023

Accepted February 6, 2023

Glazko T.T. orcid.org/0000-0002-0520-7022

Fedorova L.M. orcid.org/0000-0002-1514-3050

doi: 10.15389/agrobiology.2023.2.223eng

Abstract

The search for genetic markers that simplify the selection of animals for crosses, increasing the likelihood of offspring obtaining with the desired manifestation of economically valuable traits is a central problem in modern animal husbandry. Here, we discuss the most successful applications of various types of DNA markers of genomic element polymorphisms for solving specific breeding problems. Microsatellites are used to exclude errors of origin, single nucleotide polymorphisms (SNPs) to create maps of genomic regions in which polymorphism is associated with the variability of phenotypic

characteristics (D.J. Rigden, X.M. Fernández, 2023) and to identify the localization of key genes of adaptation to natural selection factors at the natural habitat edges and in areas of animal husbandry risky (E.K. Cheruiyot et al., 2022; L. Buggiotti et al., 2021, 2022). The loci of increased variability in the copyicity of genome regions (CNV) are used to assess their involvement in responses to natural and artificial selection factors of such polygenic systems as sensory, immune, and transporter (Y. Huang et al., 2021; P. Davoudi et al., 2022). The predominant involvement of regulatory networks including dispersed and tandem repeats, in particular microsatellite repeats, in epigenetic and phenotypic variability is discussed (R.P. Kumar et al., 2010). The structural and functional complexity of microsatellite loci, individual features of variability of specific loci and their participation in evolutionary, recombination, transcription processes are considered. Their involvement in the organization of secondary DNA structures, participation in the formation and variability of the architectonics of the interphase nucleus and regulation of gene expression profiles is noted (R.P. Kumar et al., 2010; X. Tang et al., 2022). The study of regulatory networks is of particular importance, since there is evidence that the size of the genome in animals of different taxa, as well as the distribution and composition of mobile genetic elements (sources of components of regulatory networks) differ significantly, in contrast to the similarity in the number of genes encoding proteins (V.I. Glazko et al., 2022). Accumulating evidence suggests that polylocus genotyping of individual microsatellites and dispersed repeats can contribute to solving practical problems, such as information on the specific features of the population-genetic structure, consolidation and differences between closely related groups of animals.

Keywords: DNA markers, microsatellites, short tandem repeat, STR, single nucleotide polymorphism, SNP, copy number variations, CNV, genome-wide association studies, GWAS.

REFERENCES

1. Serebrovskiy A.S. *Geneticheskiiy analiz* [Genetic analysis]. Moscow, 1970 (in Russ.).
2. **Glazko V.I.** Gene and genomic levels of domestication signature (review). *Sel'skokhozyaistvennaya biologiya [Agricultural Biology]*, 2018, 53(4): 659-672 (doi: 10.15389/agrobiology.2018.4.659eng).
3. Thakur J., Packiaraj J., Henikoff S. Sequence, chromatin and evolution of satellite DNA. *Int. J. Mol. Sci.*, 2021, 22: 4309 (doi: 10.3390/ijms22094309).
4. Li C., Tian D., Tang B., Liu X., Teng X., Zhao W., Zhang Z., Song S. Genome variation map: a worldwide collection of genome variations across multiple species. *Nucleic Acids Res.*, 2021, 49(D1): D1186-D1191 (doi: 10.1093/nar/gkaa1005).
5. Wang Z.H., Zhu Q.H., Li X., Zhu J.W., Tian D.M., Zhang S.S., Kang H.L., Li C.P., Dong L.L., Zhao W.M., Li M.H. iSheep: an integrated resource for sheep genome, variant and phenotype. *Front. Genet.*, 2021, 12: 714852 (doi: 10.3389/fgene.2021.714852).
6. Kirichenko A.V., Zlobin A.S., Shashkova T.I., Volkova N.A., Iolchiev B.S., Bagirov V.A., Borodin P.M., Karszen L.S., Tsepilov Y.A., Aulchenko Y.S. The GWAS-MAP|ovis platform for aggregation and analysis of genome-wide association study results in sheep. *Vavilovskii Zhurnal Genetiki i Seleksii*, 2022, 26(4): 378-384 (doi: 10.18699/VJGB-22-46).
7. Pös O., Radvanszky J., Buglyó G., Pös Z., Rusnakova D., Nagy B., Szemes T. DNA copy number variation: main characteristics, evolutionary significance, and pathological aspects. *Biomed. J.*, 2021, 44(5): 548-559 (doi: 10.1016/j.bj.2021.02.003).
8. Strillacci M.G., Gorla E., Cozzi M.C., Vevey M., Genova F., Scienski K., Longeri M., Bagnato A. A copy number variant scan in the autochthonous Valdostana Red Pied cattle breed and comparison with specialized dairy populations. *PLoS ONE*, 2018, 3(9): e0204669 (doi: 10.1371/journal.pone.0204669).
9. Lee Y.L., Bosse M., Mullaart E., Groenen M.A.M., Veerkamp R.F., Bouwman A.C. Functional and population genetic features of copy number variations in two dairy cattle populations. *BMC Genomics*, 2020, 21(1): 89 (doi: 10.1186/s12864-020-6496-1).
10. Qiu Y., Ding R., Zhuang Z., Wu J., Yang M., Zhou S., Ye Y., Geng Q., Xu Z., Huang S., Cai G., Wu Z., Yang J. Genome-wide detection of CNV regions and their potential association with growth and fatness traits in Duroc pigs. *BMC Genomics*, 2021, 22(1): 332 (doi: 10.1186/s12864-021-07654-7).
11. Davoudi P., Do D.N., Rathgeber B., Colombo S.M., Sargolzaei M., Plastow G., Wang Z., Karimi K., Hu G., Valipour S., Miar Y. Genome-wide detection of copy number variation in American mink using whole-genome sequencing. *BMC Genomics*, 2022, 23(1): 649 (doi: 10.1186/s12864-022-08874-1).
12. Yuan C., Lu Z., Guo T., Yue Y., Wang X., Wang T., Zhang Y., Hou F., Niu C., Sun X., Zhao H., Zhu S., Liu J., Yang B. A global analysis of CNVs in Chinese indigenous fine-wool sheep populations using whole-genome resequencing. *BMC Genomics*, 2021, 22(1): 78 (doi: 10.1186/s12864-021-07387-7).
13. Laseca N., Molina A., Valera M., Antonini A., Demyda-Peyrás S. Copy number variation (CNV): a new genomic insight in horses. *Animals (Basel)*, 2022, 12(11): 1435 (doi: 10.3390/ani12111435).

14. Bhanuprakash V., Chhotaray S., Pruthviraj D.R., Rawat C., Karthikeyan A., Panigrahi M. Copy number variation in livestock: a mini review. *Vet. World*, 2018, 11(4): 535-541 (doi: 10.14202/vet-world.2018.535-541).
15. Seroussi E., Glick G., Shirak A., Yakobson E., Weller J.I., Ezra E., Zeron Y. Analysis of copy loss and gain variations in Holstein cattle autosomes using BeadChip SNPs. *BMC Genomics*, 2010, 11: 673 (doi: 10.1186/1471-2164-11-673).
16. Wang Z., Guo J., Guo Y., Yang Y., Teng T., Yu Q., Wang T., Zhou M., Zhu Q., Wang W., Zhang Q., Yang H. Genome-wide detection of CNVs and association with body weight in sheep based on 600K SNP Arrays. *Front. Genet.*, 2020, 11: 558 (doi: 10.3389/fgene.2020.00558).
17. Taghizadeh S., Gholizadeh M., Rahimi-Mianji G., Moradi M.H., Costilla R., Moore S., Di Gerlando R. Genome-wide identification of copy number variation and association with fat deposition in thin and fat-tailed sheep breeds. *Sci. Rep.*, 2022, 12(1): 8834 (doi: 10.1038/s41598-022-12778-1).
18. Kalds P., Zhou S., Gao Y., Cai B., Huang S., Chen Y., Wang X. Genetics of the phenotypic evolution in sheep: a molecular look at diversity-driving genes. *Genet. Sel. Evol.*, 2022, 54(1): 61 (doi: 10.1186/s12711-022-00753-3).
19. Feng Y., Zhang Y., Ying C., Wang D., Du C. Nanopore-based fourth-generation DNA sequencing technology. *Genom. Proteom. Bioinform.*, 2015, 13: 4-16 (doi: 10.1016/j.gpb.2015.01.009).
20. Suminda G.G.D., Ghosh M., Son Y.O. The innovative informatics approaches of high-throughput technologies in livestock: spearheading the sustainability and resiliency of agrigenomics research. *Life (Basel)*, 2022, 12(11):1893 (doi: 10.3390/life12111893).
21. Nolte W., Alkhoder H., Wobbe M., Stock K.F., Kalm E., Vosgerau S., Krattenmacher N., Thaller G., Tetens J., Kühn C. Replacement of microsatellite markers by imputed medium-density SNP arrays for parentage control in German warmblood horses. *J. Appl. Genet.*, 2022, 63(4): 783-792 (doi: 10.1007/s13353-022-00725-9).
22. Marina H., Suarez-Vega A., Pelayo R., Gutiérrez-Gil B., Reverter A., Esteban-Blanco C., Arranz J.J. Accuracy of imputation of microsatellite markers from a 50K SNP chip in Spanish Assaf sheep. *Animals (Basel)*, 2021, 11(1): 86 (doi: 10.3390/ani11010086).
23. McClure M.C., Sonstegard T.S., Wiggins G.R., Van Eenennaam A.L., Weber K.L., Penedo C.T., Berry D.P., Flynn J., Garcia J.F., Carmo A.S., Regitano L.C., Albuquerque M., Silva M.V., Machado M.A., Coffey M., Moore K., Boscher M.Y., Genestout L., Mazza R., Taylor J.F., Schnabel R.D., Simpson B., Marques E., McEwan J.C., Cromie A., Coutinho L.L., Kuehn L.A., Keele J.W., Piper E.K., Cook J., Williams R., Bovine HapMap Consortium, Van Tassell C.P. Imputation of microsatellite alleles from dense SNP genotypes for parentage verification across multiple *Bos taurus* and *Bos indicus* breeds. *Front Genet.*, 2013, 4: 176 (doi: 10.3389/fgene.2013.00176).
24. McClure M., Sonstegard T., Wiggins G., Van Tassell C. Imputation of microsatellite alleles from dense SNP genotypes for parental verification. *Front. Genet.*, 2012, 3: 140 (doi: 10.3389/fgene.2012.00140).
25. Xu L., Haasl R.J., Sun J., Zhou Y., Bickhart D.M., Li J., Song J., Sonstegard T.S., Van Tassell C.P., Lewin H.A., Liu G.E. Systematic profiling of short tandem repeats in the cattle genome. *Genome Biol. Evol.*, 2017, 9(1): 20-31 (doi: 10.1093/gbe/evw256).
26. Jiang Y., Song H., Gao H., Zhang Q., Ding X. Exploring the optimal strategy of imputation from SNP array to whole-genome sequencing data in farm animals. *Front. Genet.*, 2022, 13: 963654 (doi: 10.3389/fgene.2022.963654).
27. Crysanto D., Leonard A.S., Fang Z.H., Pausch H. Novel functional sequences uncovered through a bovine multiassembly graph. *Proc. Natl. Acad. Sci. USA*, 2021, 118(20): e2101056118 (doi: 10.1073/pnas.2101056118).
28. Chen S., Liu S., Shi S., Jiang Y., Cao M., Tang Y., Li W., Liu J., Fang L., Yu Y., Zhang S. Comparative epigenomics reveals the impact of ruminant-specific regulatory elements on complex traits. *BMC Biol.*, 2022, 20(1): 273 (doi: 10.1186/s12915-022-01459-0).
29. Li J., Xiang Y., Zhang L., Qi X., Zheng Z., Zhou P., Tang Z., Jin Y., Zhao Q., Fu Y., Zhao Y., Li X., Fu L., Zhao S. Enhancer-promoter interaction maps provide insights into skeletal muscle-related traits in pig genome. *BMC Biol.*, 2022, 20(1): 136 (doi: 10.1186/s12915-022-01322-2).
30. Hayes B.J., Bowman P.J., Chamberlain A.J., Savin K., van Tassell C.P., Sonstegard T.S., Goddard M.E. A validated genome wide association study to breed cattle adapted to an environment altered by climate change. *PLoS ONE*, 2009, 4(8): e6676 (doi: 10.1371/journal.pone.0006676).
31. Cheruyiot E.K., Haile-Mariam M., Cocks B.G., Pryce J.E. Improving genomic selection for heat tolerance in dairy cattle: current opportunities and future directions. *Front. Genet.*, 2022, 13: 894067 (doi: 10.3389/fgene.2022.894067).
32. Buggiotti L., Yurchenko A.A., Yudin N.S., Vander Jagt C.J., Vorobieva N.V., Kusliy M.A., Vasiliev S.K., Rodionov A.N., Boronetskaya O.I., Zinovieva N.A., Graphodatsky A.S., Daetwyler H.D., Larkin D.M. Demographic history, adaptation, and NRAP convergent evolution at amino acid residue 100 in the world northernmost cattle from Siberia. *Mol. Biol. Evol.*, 2021, 38(8): 3093-3110 (doi: 10.1093/molbev/msab078).
33. Buggiotti L., Yudin N.S., Larkin D.M. Copy number variants in two northernmost cattle breeds are related to their adaptive phenotypes. *Genes*, 2022, 13: 1595 (doi: 10.3390/genes13091595).

34. Lemay D.G., Lynn D.J., Martin W.F., Neville M.C., Casey T.M., Rincon G., Kriventseva E.V., Barris W.C., Hinrichs A.S., Molenaar A.J., Pollard K.S., Maqbool N.J., Singh K., Murney R., Zdobnov E.M., Tellam R.L., Medrano J.F., German J.B., Rijkels M. The bovine lactation genome: insights into the evolution of mammalian milk. *Genome Biol.*, 2009, 10(4): R43 (doi: 10.1186/gb-2009-10-4-r43).
35. Wickramasinghe S., Rincon G., Islas-Trejo A., Medrano J.F. Transcriptional profiling of bovine milk using RNA sequencing. *BMC Genomics*, 2012, 13: 45 (doi: 10.1186/1471-2164-13-45).
36. Hammami H., Rekik B., Bastin C., Soyeurt H., Bormann J., Stoll J., Gengler N. Environmental sensitivity for milk yield in Luxembourg and Tunisian Holsteins by herd management level. *J. Dairy Sci.*, 2009, 92(9): 4604-4612 (doi: 10.3168/jds.2008-1513).
37. Cai W., Li C., Li J., Song J., Zhang S. Integrated small RNA sequencing, transcriptome and GWAS data reveal microRNA regulation in response to milk protein traits in Chinese Holstein cattle. *Front. Genet.*, 2021, 12: 726706 (doi: 10.3389/fgene.2021.726706).
38. Dysin A.P., Barkova O.Y., Pozovnikova M.V. The role of microRNAs in the mammary gland development, health, and function of cattle, goats, and sheep. *Noncoding RNA*, 2021, 7(4): 78 (doi: 10.3390/ncrna7040078).
39. Ren W., Zhang Y., Dingkao R., Huang C., Ma X., Wu X., La Y., Chu M., Bao P., Guo X., Pei J., Yan P., Liang C. Comparative study of the expression profiles of miRNAs of milk-derived exosomes of yak and Jeryak. *Animals (Basel)*, 2022, 12(22): 3189 (doi: 10.3390/ani12223189).
40. Deb R., Sengar G.S. Comparative miRNA signatures among Sahiwal and Frieswal cattle breeds during summer stress. *Biotech.*, 2021, 11(2): 79 (doi: 10.1007/s13205-020-02608-4).
41. Braud M., Magee D.A., Park S.D., Sonstegard T.S., Waters S.M., MacHugh D.E., Spillane C. Genome-wide microRNA binding site variation between extinct wild aurochs and modern cattle identifies candidate microRNA-regulated domestication genes. *Front. Genet.*, 2017, 8: 3 (doi: 10.3389/fgene.2017.00003).
42. Liu Y., Xu J., Chen M., Wang C., Li S. A unified STR profiling system across multiple species with whole genome sequencing data. *BMC Bioinformatics*, 2019, 20(Suppl. 24): 671 (doi: 10.1186/s12859-019-3246-y).
43. Cui W., Jin X., Guo Y., Chen C., Zhang W., Wang Y., Lan J., Zhu B. Development and validation of a novel five-dye short tandem repeat panel for forensic identification of 11 species. *Front. Genet.*, 2020, 11: 1005 (doi: 10.3389/fgene.2020.01005).
44. Forutan M., Ansari Mahyari S., Baes C., Melzer N., Schenkel F.S., Sargolzaei M. Inbreeding and runs of homozygosity before and after genomic selection in North American Holstein cattle. *BMC Genomics*, 2018, 19(1): 98 (doi: 10.1186/s12864-018-4453-z).
45. Häfliger I.M., Spengeler M., Seefried F.R., Drügemüller C. Four novel candidate causal variants for deficient homozygous haplotypes in Holstein cattle. *Sci. Rep.*, 2022, 12(1): 5435 (doi: 10.1038/s41598-022-09403-6).
46. Liu X., Tian D., Li C., Tang B., Wang Z., Zhang R., Pan Y., Wang Y., Zou D., Zhang Z., Song S. GWAS Atlas: an updated knowledgebase integrating more curated associations in plants and animals. *Nucleic Acids Res.*, 2023, 51(D1): D969-D976 (doi: 10.1093/nar/gkac924).
47. CNGB-NGDC Members and Partners. Database resources of the National Genomics Data Center, China National Center for Bioinformation in 2023. *Nucleic Acids Res.*, 2023, 51(D1): D18-D28 (doi: 10.1093/nar/gkac1073).
48. Rigden D.J., Fernández X.M. The 2023 Nucleic Acids Research Database Issue and the online molecular biology database collection. *Nucleic Acids Res.*, 2023, 51(D1): D1-D8 (doi: 10.1093/nar/gkac1186).
49. Flori L., Fritz S., Jaffrézic F., Boussaha M., Gut I., Heath S., Foulley J.L., Gautier M. The genome response to artificial selection: a case study in dairy cattle. *PLoS ONE*, 2009, 4(8): e6595 (doi: 10.1371/journal.pone.0006595).
50. Prakash D., Liang Z., Jiang J., Ma L., Da Y. A large-scale genome-wide association study of epistasis effects of production traits and daughter pregnancy rate in U.S. Holstein cattle. *Genes (Basel)*, 2021, 12(7): 1089 (doi: 10.3390/genes12071089).
51. Dachs N., Upadhyay M., Hannemann E., Hauser A., Krebs S., Seichter D., Russ I., Gehrke L.J., Thaller G., Medugorac I. Quantitative trait locus for calving traits on *Bos taurus* autosome 18 in Holstein cattle is embedded in a complex genomic region. *J. Dairy Sci.*, 2023, 106(3): 1925-1941 (doi: 10.3168/jds.2021-21625).
52. Huang Y., Li Y., Wang X., Yu J., Cai Y., Zheng Z., Li R., Zhang S., Chen N., Asadollahpour Nanaei H., Hanif Q., Chen Q., Fu W., Li C., Cao X., Zhou G., Liu S., He S., Li W., Chen Y., Chen H., Lei C., Liu M., Jiang Y. An atlas of CNV maps in cattle, goat and sheep. *Sci. China Life Sci.*, 2021, 64(10): 1747-1764 (doi: 10.1007/s11427-020-1850-x).
53. Revay T., Quach A.T., Maignel L., Sullivan B., King W.A. Copy number variations in high and low fertility breeding boars. *BMC Genomics*, 2015, 16(1): 280 (doi: 10.1186/s12864-015-1473-9).
54. Fontanesi L., Martelli P.L., Scotti E., Russo V., Rogel-Gaillard C., Casadio R., Vernesi C. Exploring copy number variation in the rabbit (*Oryctolagus cuniculus*) genome by array comparative genome hybridization. *Genomics*, 2012, 100(4): 245-251 (doi: 10.1016/j.ygeno.2012.07.001).
55. Davoudi P., Do D.N., Rathgeber B., Colombo S.M., Sargolzaei M., Plastow G., Wang Z.,

- Karimi K., Hu G., Valipour S., Miar Y. Genome-wide detection of copy number variation in American mink using whole-genome sequencing. *BMC Genomics*, 2022, 23(1): 649 (doi: 10.1186/s12864-022-08874-1).
56. Khlopova N.S., Glazko T.T., Glazko V.I. Constitutive and variable components of gene expression profiles in pig liver. *Russ. J. Genet. Appl. Res.*, 2011, 1: 302-307 (doi: 10.1134/S2079059711040046).
 57. Avvaru A.K., Sharma D., Verma A., Mishra R.K., Sowpati D.T. MSDB: a comprehensive, annotated database of microsatellites. *Nucleic Acids Res.*, 2020, 48(D1): D155-D159 (doi: 10.1093/nar/gkz886).
 58. Zattera M.L., Bruschi D.P. Transposable elements as a source of novel repetitive DNA in the eukaryote genome. *Cells*, 2022, 11(21): 3373 (doi: 10.3390/cells11213373).
 59. Gharesouran J., Hosseinzadeh H., Ghafouri-Fard S., Taheri M., Rezazadeh M. STRs: ancient architectures of the genome beyond the sequence. *J. Mol. Neurosci.*, 2021, 71(12): 2441-2455 (doi: 10.1007/s12031-021-01850-6).
 60. Verbiest M., Maksimov M., Jin Y., Anisimova M., Gymrek M., Bilgin Sonay T. Mutation and selection processes regulating short tandem repeats give rise to genetic and phenotypic diversity across species. *J. Evol. Biol.*, 2023, 36(2): 321-336 (doi: 10.1111/jeb.14106).
 61. Astolfi P., Bellizzi D., Sgaramella V. Frequency and coverage of trinucleotide repeats in eukaryotes. *Gene*, 2003, 317(1-2): 117-125 (doi: 10.1016/s0378-1119(03)00659-0).
 62. Arslan A. Compendious survey of protein tandem repeats in inbred mouse strains. *BMC Genom. Data*, 2022, 23(1): 62 (doi: 10.1186/s12863-022-01079-1).
 63. Wang X., Budowle B., Ge J. USAT: a bioinformatic toolkit to facilitate interpretation and comparative visualization of tandem repeat sequences. *BMC Bioinformatics*, 2022, 23(1): 497 (doi: 10.1186/s12859-022-05021-1).
 64. Jafarian Z., Khamse S., Afshar H., Khorshid H.R.K., Delbari A., Ohadi M. Natural selection at the RASGEF1C (GGC) repeat in human and divergent genotypes in late-onset neurocognitive disorder. *Sci. Rep.*, 2021, 11(1): 19235 (doi: 10.1038/s41598-021-98725-y).
 65. Khamse S., Alizadeh S., Bernhart S.H., Afshar H., Delbari A., Ohadi M. A (GCC) repeat in *SBFI* reveals a novel biological phenomenon in human and links to late onset neurocognitive disorder. *Sci. Rep.*, 2022, 12(1): 15480 (doi: 10.1038/s41598-022-19878-y).
 66. Khamse S., Jafarian Z., Bozorgmehr A., Tavakoli M., Afshar H., Keshavarz M., Moayedi R., Ohadi M. Novel implications of a strictly monomorphic (GCC) repeat in the human *PRKACB* gene. *Sci. Rep.*, 2021, 11(1): 20629 (doi: 10.1038/s41598-021-99932-3).
 67. Khamse S., Arabfard M., Salehi M., Behmard E., Jafarian Z., Afshar H., Khazaei M., Ohadi M. Predominant monomorphism of the RIT2 and GPM6B exceptionally long GA blocks in human and enriched divergent alleles in the disease compartment. *Genetica*, 2022, 150(1): 27-40 (doi: 10.1007/s10709-021-00143-5).
 68. Cheng M., Xie D., Price M., Zhou C., Zhang X.. Comparative analysis of microsatellites in coding regions provides insights into the adaptability of the giant panda, polar bear and brown bear. *Genetica*, 2022, 150(6): 355-366 (doi: 10.1007/s10709-022-00173-7).
 69. Song X., Yang T., Zhang X., Yuan Y., Yan X., Wei Y., Zhang J., Zhou C.. Comparison of the microsatellite distribution patterns in the genomes of *Euarctonogilres* at the taxonomic level. *Front. Genet.*, 2021, 12: 622724 (doi: 10.3389/fgene.2021.622724).
 70. Verbiest M.A., Delucchi M., Bilgin Sonay T., Anisimova M. Beyond microsatellite instability: intrinsic disorder as a potential link between protein short tandem repeats and cancer. *Front. Bioinform.*, 2021, 1: 685844 (doi: 10.3389/fbinf.2021.685844).
 71. Gymrek M., Willems T., Guilmatre A., Zeng H., Markus B., Georgiev S., Daly M.J., Price A.L., Pritchard J.K., Sharp A.J., Erlich Y. Abundant contribution of short tandem repeats to gene expression variation in humans. *Nat. Genet.*, 2016, 48(1): 22-29 (doi: 10.1038/ng.3461).
 72. Kinney N., Kang L., Bains H., Lawson E., Husain M., Husain K., Sandhu I., Shin Y., Carter J.K., Anandkrishnan R., Michalak P., Garner H. Ethnically biased microsatellites contribute to differential gene expression and glutathione metabolism in Africans and Europeans. *PLoS ONE*, 2021, 16(3): e0249148 (doi: 10.1371/journal.pone.0249148).
 73. Basu S., Mackowiak S.D., Niskanen H., Knezevic D., Asimi V., Grosswendt S., Geertsema H., Ali S., Jerkovi I., Ewers H., Mundlos S., Meissner A., Ibrahim D.M., Hnisz D. Unblending of transcriptional condensates in human repeat expansion disease. *Cell*, 2020, 181(5): 1062-1079.e30 (doi: 10.1016/j.cell.2020.04.018).
 74. Paul S., Dansithong W., Figueroa K.P., Scoles D.R., Pulst S.M. Staufen1 links RNA stress granules and autophagy in a model of neurodegeneration. *Nat. Commun.*, 2018, 9(1): 3648 (doi: 10.1038/s41467-018-06041-3).
 75. Johnson S.L., Tsou W.L., Prifti M.V., Harris A.L., Todi S.V. A survey of protein interactions and posttranslational modifications that influence the polyglutamine diseases. *Front. Mol. Neurosci.*, 2022, 15: 974167 (doi: 10.3389/fnmol.2022.974167).
 76. Loureiro J.R., Oliveira C.L., Silveira I. Unstable repeat expansions in neurodegenerative diseases: nucleocytoplasmic transport emerges on the scene. *Neurobiol. Aging*, 2016, 39: 174-183 (doi: 10.1016/j.neurobiolaging.2015.12.007).

77. Loureiro J.R., Castro A.F., Figueiredo A.S., Silveira I. Molecular mechanisms in pentanucleotide repeat diseases. *Cells*, 2022, 11(2): 205 (doi: 10.3390/cells11020205).
78. Li Y., Lu Y., Polak U., Lin K., Shen J., Farmer J., Seyer L., Bhalla A.D., Rozwadowska N., Lynch D.R., Butler J.S., Napierala M. Expanded GAA repeats impede transcription elongation through the *FXN* gene and induce transcriptional silencing that is restricted to the *FXN* locus. *Hum. Mol. Genet.*, 2015, 24(24): 6932-6943 (doi: 10.1093/hmg/ddv397).
79. Jain A., Vale R.D. RNA phase transitions in repeat expansion disorders. *Nature*, 2017, 546(7657): 243-247 (doi: 10.1038/nature22386).
80. Isiktas A.U., Eshov A., Yang S., Guo J.U. Systematic generation and imaging of tandem repeats reveal base-pairing properties that promote RNA aggregation. *Cell Rep. Methods*, 2022, 2(11): 100334 (doi: 10.1016/j.crmeth.2022.100334).
81. Murmann A.E., Patel M., Jeong S.Y., Bartom E.T., Jennifer Morton A., Peter M.E. The length of uninterrupted CAG repeats in stem regions of repeat disease associated hairpins determines the amount of short CAG oligonucleotides that are toxic to cells through RNA interference. *Cell Death & Disease*, 2022, 13(12): 1078 (doi: 10.1038/s41419-022-05494-1).
82. Nguyen L., Cleary J.D., Ranum L.P.W. Repeat-associated non-ATG translation: molecular mechanisms and contribution to neurological disease. *Annu. Rev. Neurosci.*, 2019, 42: 227-247 (doi: 10.1146/annurev-neuro-070918-050405).
83. Li Y.C., Korol A.B., Fahima T., Beiles A., Nevo E. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Mol. Ecol.*, 2002, 11(12): 2453-2465 (doi: 10.1046/j.1365-294x.2002.01643.x).
84. Chistiakov D.A., Hellemans B., Volckaert F.A. Microsatellites and their genomic distribution, evolution, function and applications: a review with special reference to fish genetics. *Aquaculture*, 2006, 255(1-4): 1-29 (doi: 10.1016/j.aquaculture.2005.11.031).
85. Sun J.X., Helgason A., Masson G., Ebenesersdyttir S.S., Li H., Mallick S., Gnerre S., Patterson N., Kong A., Reich D., Stefansson K. A direct characterization of human mutation based on microsatellites. *Nat. Genet.*, 2012, 44(10): 1161-1165 (doi: 10.1038/ng.2398).
86. Eckert K.A., Hile S.E. Every microsatellite is different: intrinsic DNA features dictate mutagenesis of common microsatellites present in the human genome. *Mol. Carcinog.*, 2009, 48(4): 379-388 (doi: 10.1002/mc.20499).
87. Lynch M. Rate, molecular spectrum, and consequences of human mutation. *Proc. Natl. Acad. Sci. USA*, 2010, 107(3): 961-968 (doi: 10.1073/pnas.0912629107).
88. Rajaei M., Saxena A.S., Johnson L.M., Snyder M.C., Crombie T.A., Tanny R.E., Andersen E.C., Joyner-Matos J., Baer C.F. Mutability of mononucleotide repeats, not oxidative stress, explains the discrepancy between laboratory-accumulated mutations and the natural allele-frequency spectrum in *C. elegans*. *Genome Res.*, 2021, 31(9): 1602-1613 (doi: 10.1101/gr.275372.121).
89. Brand C.L., Levine M.T. Functional diversification of chromatin on rapid evolutionary time-scales. *Annu. Rev. Genet.*, 2021, 55: 401-425 (doi: 10.1146/annurev-genet-071719-020301).
90. Murat P., Guilbaud G., Sale J.E. DNA polymerase stalling at structured DNA constrains the expansion of short tandem repeats. *Genome Biol.*, 2020, 21(1): 209 (doi: 10.1186/s13059-020-02124-x).
91. Stein M., Hile S.E., Weissensteiner M.H., Lee M., Zhang S., Kejnovský E., Kejnovská I., Makova K.D., Eckert K.A. Variation in G-quadruplex sequence and topology differentially impacts human DNA polymerase fidelity. *DNA Repair (Amst)*, 2022, 119: 103402 (doi: 10.1016/j.dnarep.2022.103402).
92. Wheeler V.C., Dion V. Modifiers of CAG/CTG repeat instability: insights from mammalian models. *Journal of Huntington's Disease*, 2021, 10(1): 123-148 (doi: 10.3233/JHD-200426).
93. Bansal A., Kaushik S., Kukreti S. Non-canonical DNA structures: diversity and disease association. *Front. Genet.*, 2022, 13: 959258 (doi: 10.3389/fgene.2022.959258).
94. Georgakopoulos-Soares I., Chan C.S.Y., Ahituv N., Hemberg M. High-throughput techniques enable advances in the roles of DNA and RNA secondary structures in transcriptional and post-transcriptional gene regulation. *Genome Biol.*, 2022, 23(1): 159 (doi: 10.1186/s13059-022-02727-6).
95. Kunkler C.N., Schiefelbein G.E., O'Leary N.J., McCown P.J., Brown J.A. A single natural RNA modification can destabilize a U•A-T-rich RNA•DNA-DNA triple helix. *RNA*, 2022, 28(9): 1172-1184 (doi: 10.1261/rna.079244.122).
96. Brown J.A. Unraveling the structure and biological functions of RNA triple helices. *Wiley Interdiscip. Rev. RNA*, 2020, 11(6): e1598 (doi: 10.1002/wrna.1598).
97. Bowater R.P., Bohálová N., Brázda V. Interaction of proteins with inverted repeats and cruciform structures in nucleic acids. *Int. J. Mol. Sci.*, 2022, 23(11): 6171 (doi: 10.3390/ijms23116171).
98. Brázda V., Laister R.C., Jagelská E.B., Arrowsmith C. Cruciform structures are a common DNA feature important for regulating biological processes. *BMC Mol. Biol.*, 2011, 12: 33 (doi: 10.1186/1471-2199-12-33).
99. Gao Z., Moorjani P., Sasaki T.A., Pedersen B.S., Quinlan A.R., Jorde L.B., Amster G., Przeworski M. Overlooked roles of DNA damage and maternal age in generating human germline mutations. *Proc. Natl. Acad. Sci. USA*, 2019, 116(19): 9491-9500 (doi: 10.1073/pnas.1901259116).

100. Mitra I., Huang B., Mousavi N., Ma N., Lamkin M., Yanicky R., Shleizer-Burko S., Lohmueller K.E., Gymrek M. Patterns of de novo tandem repeat mutations and their role in autism. *Nature*, 2021, 589(7841): 246-250 (doi: 10.1038/s41586-020-03078-7).
101. Hannan A.J. Tandem repeats mediating genetic plasticity in health and disease. *Nat. Rev. Genet.*, 2018, 19(5): 286-298 (doi: 10.1038/nrg.2017.115).
102. Steely C.J., Watkins W.S., Baird L., Jorde L.B. The mutational dynamics of short tandem repeats in large, multigenerational families. *Genome Biol.*, 2022, 23(1): 253 (doi: 10.1186/s13059-022-02818-4).
103. McGurk M.P., Barbash D.A. Double insertion of transposable elements provides a substrate for the evolution of satellite DNA. *Genome Res.*, 2018, 28(5): 714-725 (doi: 10.1101/gr.231472.117).
104. Nishiyama E., Ohshima K. Cross-kingdom commonality of a novel insertion signature of RTE-related short retroposons. *Genome Biol Evol.*, 2018, 10(6): 1471-1483 (doi: 10.1093/gbe/evy098).
105. Paço A., Freitas R., Vieira-da-Silva A. Conversion of DNA sequences: from a transposable element to a tandem repeat or to a gene. *Genes (Basel)*, 2019, 10(12): 1014 (doi: 10.3390/genes10121014).
106. Richard G.F., Kerrest A., Dujon B. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiol. Mol. Biol. Rev.*, 2008, 72(4): 686-727 (doi: 10.1128/MMBR.00011-08).
107. López-Giráldez F., Andrés O., Domingo-Roura X., Bosch M. Analyses of carnivore microsatellites and their intimate association with tRNA-derived SINES. *BMC Genomics*, 2006, 7: 269 (doi: 10.1186/1471-2164-7-269).
108. Duffy A.J., Coltman D.W., Wright J.M. Microsatellites at a common site in the second ORF of L1 elements in mammalian genomes. *Mamm. Genome*, 1996, 7(5): 386-387 (doi: 10.1007/s003359900111).
109. Lexa M., Jedlicka P., Vanat I., Cervenansky M., Kejnovsky E. TE-greedy-nester: structure-based detection of LTR retrotransposons and their nesting. *Bioinformatics*, 2020, 36(20): 4991-4999 (doi: 10.1093/bioinformatics/btaa632).
110. Glazko V.I., Kosovskiy G.Yu., Glazko T.T. **The sources of genome variability as domestication drivers (review).** *Sel'skokhozyaistvennaya biologiya [Agricultural Biology]*, 2022, 57(5): 832-851 (doi: 10.15389/agrobiology.2022.5.832eng).
111. Ivancevic A.M., Kortschak R.D., Bertozzi T., Adelson D.L. Horizontal transfer of BovB and L1 retrotransposons in eukaryotes. *Genome Biol.*, 2018, 19(1): 85 (doi: 10.1186/s13059-018-1456-7).
112. Tellam R.L., Lemay D.G., Van Tassel C.P., Lewin H.A., Worley K.C., Elsik C.G. Unlocking the bovine genome. *BMC Genomics*, 2009, 10: 193 (doi: 10.1186/1471-2164-10-193).
113. Kelly C.J., Chitko-McKown C.G., Chuong E.B. Ruminant-specific retrotransposons shape regulatory evolution of bovine immunity. *Genome Res.*, 2022, 32(8): 1474-1486 (doi: 10.1101/gr.276241.121).
114. Mirceta M., Shum N., Schmidt M.H.M., Pearson C.E. Fragile sites, chromosomal lesions, tandem repeats, and disease. *Front. Genet.*, 2022, 13: 985975 (doi: 10.3389/fgene.2022.985975).
115. de Oliveira T.D., Bertocchi N.A., Kretschmer R., de Oliveira E.H.C., Cioffi M.D.B., Liehr T., de Freitas T.R.O. Genomic organization of microsatellites and LINE-1-like retrotransposons: evolutionary implications for *Ctenomys minutus* (Rodentia: Ctenomyidae) cytotypes. *Animals*, 2022, 12(16): 2091 (doi: 10.3390/ani12162091).
116. Kryshchuk I.A., Orlov V.N., Cherepanova E.V., Borisov Y.M. Unusual chromosomal polymorphism of the common shrew, *Sorex araneus* L., in southern Belarus. *Comparative Cytogenetics*, 2021, 15(2): 159-169 (doi: 10.3897/CompCytogen.v15.i2.63084).
117. Vorontsov N.N. *Razvitie evolyutsionnykh idey v Mezhdunbiologii*. Moscow, 2004 (in Russ.).
118. Ramírez D., Rodríguez M.E., Cross I., Arias-Pérez A., Merlo M.A., Anaya M., Portela-Bens S., Martínez P., Robles F., Ruiz-Rejón C., Rebordinos L. Integration of maps enables a cytogenomics analysis of the complete karyotype in *Solea senegalensis*. *Int. J. Mol. Sci.*, 2022, 23(10): 5353 (doi: 10.3390/ijms23105353).
119. Mudd A.B., Bredeson J.V., Baum R., Hockemeyer D., Rokhsar D.S. Analysis of muntjac deer genome and chromatin architecture reveals rapid karyotype evolution. *Commun. Biol.*, 2020, 3(1): 480 (doi: 10.1038/s42003-020-1096-9).
120. Yin Y., Fan H., Zhou B., Hu Y., Fan G., Wang J., Zhou F., Nie W., Zhang C., Liu L., Zhong Z., Zhu W., Liu G., Lin Z., Liu C., Zhou J., Huang G., Li Z., Yu J., Zhang Y., Yang Y., Zhuo B., Zhang B., Chang J., Qian H., Peng Y., Chen X., Chen L., Li Z., Zhou Q., Wang W., Wei F. Molecular mechanisms and topological consequences of drastic chromosomal rearrangements of muntjac deer. *Nat. Commun.*, 2021, 12(1): 6858 (doi: 10.1038/s41467-021-27091-0).
121. Damas J., Corbo M., Kim J., Turner-Maier J., Farré M., Larkin D.M., Ryder O.A., Steiner C., Houck M.L., Hall S., Shiue L., Thomas S., Swale T., Daly M., Korlach J., Uliano-Silva M., Mazzoni C.J., Birren B.W., Genereux D.P., Johnson J., Lindblad-Toh K., Karlsson E.K., Nweeia M.T., Johnson R.N., Zoonomia Consortium, Lewin H.A. Evolution of the ancestral mammalian karyotype and syntenic regions. *Proc. Natl. Acad. Sci. USA*, 2022, 119(40): e2209139119 (doi: 10.1073/pnas.2209139119).
122. Waters P.D., Patel H.R., Ruiz-Herrera A., Álvarez-González L., Lister N.C., Simakov O., Ezaz T., Kaur P., Frere C., Grützner F., Georges A., Graves J.A.M. Microchromosomes are building blocks of bird, reptile, and mammal chromosomes. *Proc. Natl. Acad. Sci. USA*, 2021,

- 118(45): e2112494118 (doi: 10.1073/pnas.2112494118).
123. Fritz A.J., Sehgal N., Pliss A., Xu J., Berezney R. Chromosome territories and the global regulation of the genome. *Genes Chromosomes Cancer*, 2019, 58(7): 407-426 (doi: 10.1002/gcc.22732).
 124. Deakin J.E., Potter S., O'Neill R., Ruiz-Herrera A., Cioffi M.B., Eldridge M.D.B., Fukui K., Marshall Graves J.A., Griffin D., Grutzner F., Kratochvíl L., Miura I., Rovatsos M., Srikulnath K., Wapstra E., Ezaz T. Chromosomics: bridging the gap between genomes and chromosomes. *Genes (Basel)*, 2019, 10(8): 627 (doi: 10.3390/genes10080627).
 125. Falk M., Feodorova Y., Naumova N., Imakaev M., Lajoie B.R., Leonhardt H., Joffe B., Dekker J., Fudenberg G., Solovei I., Mirny L.A. Heterochromatin drives compartmentalization of inverted and conventional nuclei. *Nature*, 2019, 570(7761): 395-399 (doi: 10.1038/s41586-019-1275-3).
 126. Kumar R.P., Senthilkumar R., Singh V., Mishra R.K. Repeat performance: how do genome packaging and regulation depend on simple sequence repeats? *Bioessays*, 2010, 32(2): 165-174 (doi: 10.1002/bies.200900111).
 127. van Kruistum H., Nijland R., Reznick D.N., Groenen M.A.M., Megens H.J., Pollux B.J.A. Parallel genomic changes drive repeated evolution of placentas in live-bearing fish. *Mol. Biol. Evol.*, 2021, 38(6): 2627-2638 (doi: 10.1093/molbev/msab057).
 128. Tang X., Li T., Liu S., Wisniewski J., Zheng Q., Rong Y., Lavis L.D., Wu C. Kinetic principles underlying pioneer function of GAGA transcription factor in live cells. *Nat. Struct. Mol. Biol.*, 2022, 29(7): 665-676 (doi: 10.1038/s41594-022-00800-z).
 129. Foster C.S.P., Van Dyke J.U., Thompson M.B., Smith N.M.A., Simpfendorfer C.A., Murphy C.R., Whittington C.M. Different genes are recruited during convergent evolution of pregnancy and the placenta. *Mol. Biol. Evol.*, 2022, 39(4): msac077 (doi: 10.1093/molbev/msac077).
 130. Glazko V.I., Andreichenko I.N., Kovalchuk S.N., Glazko T.T., Kosovsky G.Yu. Candidate genes for control of cattle milk production traits. *Russ. Agricult. Sci.*, 2016, 42: 458-464 (doi: 10.3103/S1068367416060082).