



## Вариация числа копий (CNV) как перспективный генетический маркер: распространение, методы валидации и гены-кандидаты в геномах сельскохозяйственных животных (обзор)

© 2020. О. А. Кошкина, Т. Е. Денискова ✉, Н. А. Зиновьева

ФГБНУ «Федеральный научный центр животноводства – ВИЖ имени академика Л. К. Эрнста», пос. Дубровицы, Московская область, Российская Федерация

Вариация числа копий (CNV) – это повторяющиеся участки генома, размером от одной тысячи до нескольких миллионов пар оснований, варьирующие между особями в популяции. Благодаря большому покрытию генома по сравнению с SNP-маркерами, CNV является важным источником генетической изменчивости и рассматривается в настоящее время как альтернативный тип ДНК-маркеров. Основное внимание уделяется идентификации регионов CNV (CNVR), перекрывающихся с генами и локусами количественных признаков (QTL) в геномах сельскохозяйственных животных. В обзоре обобщены и проанализированы результаты исследований по CNV у различных видов сельскохозяйственных животных, включая идентификацию генов-кандидатов, локусы которых перекрываются с областями CNV, а также дана краткая характеристика методических подходов для изучения вариации числа копий. У крупного рогатого скота было идентифицировано от 51 до 1265 CNVR с долей покрытия генома от 0,5 до 20 %, у свиней – 565 CNVR и 5,84 %, у коз – 978 CNVR и 8,96 %, у овец – 3488 CNVR и 2,7 %, соответственно. Локусы функциональных генов-кандидатов, связанных с экономически-значимыми признаками, перекрывались с CNVR у всех видов сельскохозяйственных животных. Были идентифицированы гены, ассоциированные с показателями роста и развития (MYH3 и GBP4 у крупного рогатого скота; ANP32B, GYS1 и CAV1 у свиней; MYLK4 у коз; SHE, BAG4, PIGY и ORMDL1 у овец), влияющие на репродуктивные признаки и плодовитость (PRP1 и PRP6 у коз, PRLR у крупного рогатого скота, PTGS1 у овец), связанные с мясной продуктивностью (KDM5B, ADAM8 и SHH у коз), ответственные за различные фенотипы окраски кожи или шерсти (KIT у свиней; ASIP, AHCU и ITCH у овец и коз) и вовлечённые в регуляцию обменных процессов (PPARA, RXRA, ADD1, FASN и PPP1CA у овец). Анализ мирового опыта продемонстрировал, что идентифицированные CNV могут быть предложены как потенциальные кандидаты для селекции по экономически значимым признакам у сельскохозяйственных животных.

**Ключевые слова:** генетический полиморфизм, организация генома, локусы количественных признаков (QTL), ДНК-чипы, хозяйственно полезные признаки

**Благодарности:** настоящий обзор подготовлен в рамках выполнения государственного задания Министерства науки и высшего образования РФ по теме № 0445-2019-0026 (AAAA-A18-118021590138-1).

Авторы благодарят рецензентов за их вклад в экспертную оценку этой работы.

**Конфликт интересов:** авторы заявили об отсутствии конфликта интересов.

**Для цитирования:** Кошкина О. А., Денискова Т. Е., Зиновьева Н. А. Вариация числа копий (CNV) как перспективный генетический маркер: распространение, методы валидации и гены-кандидаты в геномах сельскохозяйственных животных (обзор). *Аграрная наука Евро-Северо-Востока*. 2020;21(4):355-368. DOI: <https://doi.org/10.30766/2072-9081.2020.21.4.355-368>

Поступила: 23.06.2020

Принята к публикации: 23.07.2020

Опубликована онлайн: 24.08.2020

## Copy number variation (CNV) as a promising genetic marker: distribution, validation methods and candidate genes in genomes of livestock species (review)

© 2020. Olga A. Koshkina, Tatiana E. Deniskova ✉, Natalia A. Zinovieva

L. K. Ernst Federal Science Center for Animal Husbandry, Dubrovitsy, Moscow Region, Russian Federation

Copy number variations (CNVs) are repetitive genome segments, ranging from one thousand to several million base pairs and varying between individuals in a population. Due to a larger genome coverage compared to SNP markers, CNVs are important sources of genetic variation and are currently considered as an alternative type of DNA markers. The identification of CNV regions (CNVRs) which overlap with genes and quantitative trait loci (QTLs) in livestock genomes are of the greatest interest. In the review, the results of studies on CNV in various livestock species, are summarized and analyzed including the identification of candidate genes whose loci overlap with CNV regions. In addition, the methodological approaches for detection of copy number variations are briefly described. The number of identified CNVRs and a genome coverage ratio were 51-1265 and 0.5-20 % in cattle, 565 CNVRs and 5.84 % in pigs, 978 CNVR and 8.96 % in goats, 3488 CNVR and 2.7 % in sheep. Loci of functional candidate genes associated with economically significant traits overlap with CNVR in all livestock species. There were identified genes associated with growth and development indicators (MYH3 and GBP4 in cattle;

*ANP32B, GYS1 and CAV1 in pigs; MYLK4 in goats; SHE, BAG4, PIGY and ORMDL1 in sheep); affecting the reproductive traits and fertility (PRP1 and PRP6 in goats; PTGS1 in sheep); associated with meat productivity (KDM5B, ADAM8 and SHH in goats); responsible for various coat and skin colour phenotypes (KIT in pigs; ASIP, AHCY and ITCH in sheep and goats) and involved in the regulation of metabolic processes (PPARA, RXRA, ADD1, FASN and PPPICA in sheep). The analysis of international experience showed that identified CNVs could be proposed as potential candidates for selection according to economically significant traits in livestock.*

**Keywords:** genetic polymorphism, genome organization, quantitative trait loci (QTL), DNA chips, economically significant traits

**Acknowledgment:** the review was prepared within the framework of the state task of the Ministry of Science and Higher Education of the Russian Federation according to the theme No. 0445-2019-0026 (AAAA-A18-118021590138-1).

**Conflict of interest:** the authors stated that there was no conflict of interest.

**For citation:** Koshkina O. A., Deniskova T. E., Zinovieva N. A. Copy number variation (CNV) as a promising genetic marker: distribution, validation methods and candidate genes in genomes of livestock species (review). *Agrarnaya nauka Evro-Severo-Vostoka* = Agricultural Science Euro-North-East. 2020;21(4):355-368. (In Russ.). DOI: <https://doi.org/10.30766/2072-9081.2020.21.4.355-368>

Received: 23.06.2020

Accepted for publication: 23.07.2020

Published online: 24.08.2020

Благодаря активному развитию методов генотипирования с использованием ДНК-чипов в геномах сельскохозяйственных животных было обнаружено множество геномных субмикроскопических структурных вариаций. Одной из основных генетических форм таких вариаций, широко распространенных в геноме, является вариация числа копий (copy number variation, CNV) [1]. Исследования CNV способствуют лучшему пониманию эволюционного механизма одомашнивания сельскохозяйственных животных и их адаптации к различным условиям окружающей среды, обусловленной такими механизмами, как дозы генов и изменения структуры транскриптов – продуктов активности ферментов РНК-полимераз. Вариация числа копий изменяет экспрессию генов и, тем самым, оказывает влияние на фенотипические признаки особей путем делеции (выпадение участков хромосомы) или дублирования генов в регионах CNV (CNVR). Данный аспект дает основания предположить, что CNV могут оказывать значительное влияние на экономически значимые признаки сельскохозяйственных животных.

**Цель работы** – проанализировать имеющийся мировой опыт по поиску регионов с вариацией числа копий, в том числе с указанием попадающих в них генов и локусов количественных признаков, и показать применяемые методические подходы для подготовки научной базы для проведения экспериментов по поиску CNV в геномах отечественных пород сельскохозяйственных животных.

**Материал и методы.** В качестве материала для исследования были использованы литературные источники, посвященные проблеме организации генома у сельскохозяйственных животных и содержащие результаты

поиска регионов с вариацией числа копий. Поиск научных источников осуществлялся в базе данных PubMed®/The National Center for Biotechnology Information (NCBI) (<https://pubmed.ncbi.nlm.nih.gov/>). В поисковых запросах были использованы следующие основные комбинации ключевых слов: copy number variation, CNV, CNV detection, CNV livestock, CNV cattle, CNV sheep, CNV goat, CNV pig, copy number variation pig, copy number variation cattle, copy number variation sheep, copy number variation goat. Альтернативная стратегия поиска заключалась в просмотре списка цитирований в публикациях, найденных ранее в базе данных PubMed/NCBI.

Глубина поиска литературных источников, дающих общую характеристику феномена CNV и характеризующих методы идентификации регионов CNV, составляла 18 лет (с 2002 по 2020 гг.). Глубина поиска литературных источников, описывающих результаты исследований по наличию CNV в геномах сельскохозяйственных животных, не была строго определена в связи с тем, что данный подход нашел свое применение в сельскохозяйственной генетике сравнительно недавно (с 2012 г.).

Литературные источники были включены в обзор по следующим критериям: актуальность; публикация в рецензируемых изданиях; наличие секции обсуждения полученных результатов с другими авторами; детальное описание методологии работы; описание генов-кандидатов, перекрывающихся в регионах CNV (предпочтение отдавалось тем источникам, где функция генов-кандидатов была четко обозначена).

**Основная часть.** Что такое CNV? Краткая характеристика и потенциальные аспекты прикладного использования.

Вариация числа копий – это полиморфные участки генома, представляющие собой сегменты ДНК размером от одной тысячи до нескольких миллионов пар оснований, варьирующих по сравнению с эталонным или референсным геномом [1].

CNV оказывают влияние на экспрессию генов, фенотипическое разнообразие и эволюционную адаптационную способность животных посредством включения широкого спектра механизмов, таких как дозировка генов и изменение структуры транскриптов [2]. Частота мутаций CNV варьирует от  $1,1 \times 10^{-2}$  [3] до  $1 \times 10^{-8}$  на locus в расчете на поколение [4, 5], что отражает разнообразные процессы, с помощью которых создаются CNV. Существует четыре вида механизмов образования CNV: неаллельные гомологичные рекомбинации (Nonallelic Homologous Recombination, NaHR), также известны под названием неравного кроссинговера; негомологичное соединение концов (Non-homologous End-Joining, NHEJ); остановка вилки репликации и переключение матрицы (Fork Stalling and Template Switching, FoSTeS); L1-опосредованная ретротранспозиция [6].

Вариация числа копий охватывает больше нуклеотидных полиморфизмов ( $\approx 1\%$  генома индивидуума), чем однонуклеотидные полиморфизмы SNP ( $\approx 0,1\%$ ) [7, 8], внося тем самым значительный вклад в разнообразие генетических вариаций в геноме [9].

CNV отличаются от таких типов геномных вариаций, как инсерции (встраивание фрагментов хромосом в новый locus) и делеции, тем, что последние, как правило, имеют размеры менее 1 Kb [1]. CNV также отличаются и от сегментных дупликаций (SD), имеющих длину более 1 Kb с идентичностью последовательности не менее 90% между дублированными фрагментами, которые часто не являются полиморфными в популяции [10]. Вероятно, что SD когда-то были CNV, которые впоследствии закрепились в популяции.

CNV могут присутствовать в одних и тех же или перекрывающихся областях генома у нескольких индивидуумов. Такие области называют регионами или областями вариации числа копий (CNVR). В литературе имеются данные о том, что CNVR преимущественно расположены вне генных областей [11, 12, 13], и что те CNV, которые перекрывают гены, с большей вероятностью, будут представлены дупликациями (удвоение участка хромосомы), чем делециями [7, 14]. Вероятно, это связано

с тем, что делеции оказывают более разрушительный эффект на функции генов, чем дупликации, и поэтому подвергаются большему избирательному отбору. G. M. Cooper и соавторы [15] предполагают, что CNV, которые перекрывают SD, с большей вероятностью будут представлены в генных регионах, в то время как CNV, которые не перекрывают SD, представлены в бедных генами участках генома.

В последние годы вариация числа копий все чаще рассматривается как важный источник генетического и фенотипического разнообразия [16, 17, 18]. В геноме человека был обнаружен ряд CNV, связанных с заболеваниями, включающими аутоиммунные нарушения, шизофрению, аутизм, а также инфекционные и сердечно-сосудистые заболевания [19, 20]. Кроме этого, доказана роль CNV в фармакокинетике лекарственной эффективности и токсичности [21]. Таким образом, системный анализ и характеристика CNV способствуют лучшему пониманию генетической вариабельности и являются важным инструментом для расшифровки роли CNV в наследуемости сложных признаков. В последние годы был проведен ряд исследований CNV у различных видов млекопитающих, включая человека [22], крупный рогатый скот [23, 24], коз [25], овец [26, 27], собак [28], кроликов [29] и свиней [30, 31].

*Методы поиска и подтверждения присутствия CNV в геноме.* В животноводстве, согласно имеющимся данным, CNV могут влиять на экономически значимые признаки, поэтому могут быть интегрированы в программы селекции по ДНК-маркерам у свиней [32], коз [33], овец [34], крупного рогатого скота [35]. В связи с ростом потребительских требований к качеству продукции животноводства и развитием маркерной селекции, в современном мире возникает потребность использовать CNV в качестве молекулярно-генетических маркеров для проведения эффективного отбора по экономически значимым признакам и профилактики заболеваний сельскохозяйственных животных.

С развитием высокоплотных платформ для генотипирования и секвенирования нового поколения (NGS) идентификация CNV становится результативным подходом в геномной селекции животных, создавая огромное количество возможностей для повышения генетического прогресса в животноводстве.

В настоящее время разработан широкий спектр молекулярно-генетических методов для

установления вариации числа копий, включая традиционное кариотипирование, флуоресцентный гибридизационный анализ *in situ* (FISH), скрининг количества копий на основе микрочипов, мультиплексная амплификация лигированных зондов (MLPA), секвенирование нового поколения (NGS), сравнительная геномная гибридизация (CGH) и количественная полимеразная цепная реакция (qPCR). Каждый метод имеет свои преимущества, и выбор конкретного подхода в значительной степени зависит от области применения, используемого оборудования, его разрешающей способности, рабочей нагрузки и стоимости анализа [36].

Традиционное кариотипирование позволяет обнаружить структурные вариации по всему геному, но его разрешение ограничено ( $> 5\text{--}10\text{ Mb}$ ).

FISH-анализ для целевых регионов рутинно используется в течение многих лет и требует либо метафазных хромосом (по аналогии с кариотипированием), либо межфазных ядер (разрешение около  $100\text{ Kb}$ ).

Сравнительная геномная гибридизация является наиболее широко используемой платформой благодаря её низкой стоимости и возможности быстрого, точного и одновременного анализа многих генов и мутаций [37, 38]. В дополнении к CGH была разработана платформа SNP-CGH, в основе которой лежит высокопроизводительное SNP-генотипирование, и значения плотности SNP, полученные из каждого образца, используются для обнаружения CNV в каждом отдельном случае. Данный подход отличается высокой производительностью и пригоден для проведения ассоциативных исследований [39].

Для целенаправленного скрининга соответствующих генов и окружающих их регуляторных областей отдается предпочтение методам с фиксированным количеством копий, таким как MLPA, микрочипы и qPCR [40, 41]. Технология MLPA позволяет выявлять широкий диапазон генетических нарушений – от точечных мутаций единичных нуклеотидов до делеций/дупликаций обширных хромосомных регионов. Большинство современных биочипов являются SNP-микрочипами, с помощью которых можно исследовать более миллиона однонуклеотидных замен одновременно. С развитием ДНК-чипов высокой плотности было достигнуто более высокое разрешение [42], что способствует выявлению значительно меньших (до нескольких пар

оснований) CNV в геномах сельскохозяйственных животных. Недостатком ДНК-чипов является их высокая стоимость. Несмотря на преимущества биочипов высокой плотности, сегодня их частично замещают технологией секвенирования нового поколения (NGS). NGS – наиболее высокопродуктивная технология исследования генома [43], однако также имеющая недостатки, главным из которых является высокая частота так называемых «indel error rate» (степень ошибок при выявлении инсерций / делеций) [44]. Количественная ПЦР в реальном времени обладает рядом преимуществ перед альтернативными методами: низкая стоимость расходных материалов; менее высокие требования к приборной базе; быстрота анализа; высокая чувствительность и открытый формат (не зависящий от одного поставщика). Количественная ПЦР – это современный «золотой» стандарт анализа экспрессии генов [36], который находит все большее применение для подтверждения присутствия регионов CNV в геноме.

*Исследования вариации числа копий у крупного рогатого скота.* С использованием целого ряда платформ были построены и доработаны тонкие карты CNV для человека и других видов млекопитающих [45, 46, 47, 48]. Среди сельскохозяйственных животных наибольшее число исследовательских работ по идентификации CNV было проведено в геноме крупного рогатого скота. По данным различных исследований, у крупного рогатого скота идентифицировано от 51 до 1265 CNVR [46, 48], при этом оценки доли генома, содержащей CNV регионы, варьируют от 0,5 до 20 % [48, 49]. Широкий диапазон оценок, вероятно, обусловлен рядом факторов, среди которых можно выделить различия в технологиях обнаружения CNV, используемых критериях идентификации CNV, породной принадлежности и количестве исследуемых животных.

J. S. Вае и соавторы [45] представили CNV-карту для крупного рогатого скота. Применение Illumina BovineSNP50 BeadChip позволило идентифицировать 855 CNV, имеющих средний размер  $149,8\text{ Kb}$ , из которых 368 областей CNV были уникальными.

J. W. Kijas и соавторы [48] провели сравнительную геномную гибридизацию тауринного и зебувидного скота с использованием олигонуклеотидной платформы на основе 385 000 зондов. В результате был обнаружен 51 CNV со средним индивидуальным разме-

ром от 213 до 335 Kb, которые в совокупности покрывали примерно 0,5 % генома крупного рогатого скота. Большая часть CNV (82 %) охватывала не менее одного гена, среди которых были идентифицированы функциональные кандидаты, в том числе ответственные за устойчивость к паразитарным болезням и влияющие на фертильность у молочных пород.

Используя аналогичную платформу с целью выявления CNV в геноме домашней козы, L. Fontanesi и соавторы [50] провели эксперимент по сравнительной межвидовой гибридизации генома крупного рогатого скота и козы (aCGH). Всего был выявлен 161 CNV и идентифицировано 127 CNVR, охватывающих около 11,47 Mb виртуального генома козы, относящегося к геному крупного рогатого скота (0,435 % в соответствии с последней сборкой генома). Выявленные перекрытия между регионами CNV козы и крупного рогатого скота были статистически значимыми ( $P < 0,0001$ ), поэтому можно предположить, что некоторые хромосомные области содержат повторяющиеся межвидовые CNVR.

Получены данные о влиянии CNV на хозяйственно полезные признаки у крупного рогатого скота и других видов животных. Например, CNV гена *GBP4* крупного рогатого скота является частью локуса количественных признаков (QTL), влияющего на рост взрослых животных [51]. Дублирование участка размером 133 Kb, содержащего четыре гена *FGF3*, *FGF4*, *FGF19* и *ORAOV1*, приводит к появлению волосяного покрова у родезианских и тайских собак [52].

Y. Xu и соавторы [35] провели анализ ассоциации вариации числа копий гена *MYH3* с уровнем транскрипции и признаками роста у четырех китайских пород крупного рогатого скота. Ген *MYH3*, критический регуляторный фактор развития скелетных мышц, был обнаружен в области CNV путем сравнительной геномной гибридизации. Было установлено, что увеличенное число копий гена *MYH3* ( $\geq 3$  копий) положительно коррелирует с уровнем транскрипции в скелетных мышцах как в эмбрионах, так и у взрослых животных ( $P < 0,05$ ).

*Исследования вариации числа копий у свиней.* В 2008 году J. Fadista и соавторы [30] представили первые данные по наличию регионов CNV в геноме свиньи, локализованных на хромосомах 4, 7, 14 и 17. С. Chen и соавторы [53], исследуя 1693 свиней из 18 различных популяций с использованием мик-

рочипа SNP60 BeadChip и программного обеспечения PennCNV [54], идентифицировали 1315 предполагаемых CNV, принадлежащих к 565 регионам CNV (CNVR). Несколько областей CNV были обнаружены на хромосомах 6, 11, 13, 14 и 17. Семь CNV перекрывались с QTL, ассоциированных с длиной туши, толщиной шпика, массой нутряного сала, длиной лопаточной части, содержанием внутримышечного жира в длиннейшей мышце спины, живой массой в возрасте 240 дней и гликолитическим потенциалом длиннейшей мышцы спины. Кроме того, был идентифицирован один CNV, который ранее ассоциировался с цветом кожи у свиней [55].

*Исследования вариации числа копий у коз.* В 2009 году L. Fontanesi и соавторы [56] провели сравнительную гибридизацию с использованием платформы aCGH на основе сиквенсов кодирующих последовательностей гена *ASIP* в 6 породах коз с различным окрасом, в том числе у зааненской и мурсиана градина. Были обнаружены CNV, покрывающие область генома размером менее 100 Kb и включающие гены *ASIP* и *AHCY*. Был сделан вывод о том, что разница в количестве копий в локусе *Agouti* у коз вносит определенный вклад в вариабельность окраски шерсти. В 2010 году тот же научный коллектив идентифицировал 127 CNVR, охватывающих около 11,47 Mb генома козы, и представил первую карту CNVR в геноме козы [50].

В 2019 году R.Q. Zhang и соавторы [57] на основе данных полногеномного генотипирования провели скрининг CNV в популяции лаошанских молочных коз, представленных группами с высокой и низкой плодовитостью. Выявленная дупликация генов *PRP1* и *PRP6*, ответственных за репродуктивные процессы, в группе коз с высокой плодовитостью позволила предположить, что CNV могут способствовать увеличению размера помета у лаошанских молочных коз.

S. Y. Shi и соавторы [33] обнаружили, что ген *MYLK4* перекрывался с регионом CNV. Было показано, что вариации числа копий гена *MYLK4* существенно влияют на живую массу, длину туловища и высоту в холке у коз.

Первой системной попыткой провести анализ распределения CNV в популяциях коз, разводимых в различных странах мира, стало исследование M. Liu и соавторов, проведенное в 2019 году [25]. Анализ полногеномных SNP-профилей, сгенерированных в рамках проекта

ADAPTmap, выявил 6286 CNV у 1023 коз и 978 CNVR, покрывающих 262 Mb, или 8,96 % генома коз. Кроме того, было идентифицировано 154 функциональных гена, перекрывающихся в CNVR, в том числе *EDNRA*, *NR3C2*, *ADAMTS20*, *ASIP*, *AHCY*, *ITCH*, *KDM5B*, *ADAM8* и *SHH*. Уровень амплификации гена *EDNRA* положительно коррелирует со степенью покрытия белыми пятнами в масти у коз. Ген *NR3C2* кодирует минералокортикоидный рецептор, участвующий в регуляции артериального давления. Сообщалось, что дополнительные копии гена *NR3C2* могут быть связаны со специфическими признаками, определяющими адаптации к суровым и засушливым климатическим условиям [50]. Ген *ADAMTS20* сыграл значительную роль в вариации цвета шерсти в процессе одомашнивания коз. Гены *ASIP*, *AHCY* и *ITCH* ответственны за изменение окраса шерсти у коз, что подтверждается проведенными ранее исследованиями L. Fontanesi с соавторами [56]. За мясную продуктивность отвечают гены *KDM5B*, *ADAM8* и *SHH*. Полученные результаты показали, что идентифицированные CNV могут рассматриваться как потенциальные кандидаты для улучшения экономически-значимых признаков, в том числе мясной и молочной продуктивности у коз.

*Исследования вариации числа копий у овец.* В последние годы был достигнут определенный прогресс в идентификации вариации числа копий у овец. В 2011 году L. Fontanesi и соавторы [58] представили первую карту CNV генома овцы, полученную с использованием метода межвидовой сравнительной гибридизации с геномом крупного рогатого скота. Однако данный подход не был оптимальным из-за низкой гомологии между зондами крупного рогатого скота и ДНК овец для некоторых регионов и не отражал реального распределения CNV регионов в геноме овцы.

В 2013 году с использованием ДНК-чипа Ovine SNP50 впервые была построена карта CNV овец [59]. В трех породах овец идентифицировали 238 регионов CNV, покрывающих 60,35 Mb генома, что соответствовало 2,27 % аутомсомной геномной последовательности и 2,17 % полного генома овец. Полученные данные отличались от результатов, полученных ранее L. Fontanesi и соавторами [58].

J. Liu и соавторы [59] попытались идентифицировать гены, расположенные внутри

или частично перекрывающиеся с регионами CNV. Геном овцы не так хорошо изучен по сравнению с геномом других сельскохозяйственных животных [60, 61, 62], поэтому только 4 региона CNV частично совпали с пятью известными генами. Тем не менее, было выявлено 563 ортологичных гена человека в CNVR овец. Также были определены области CNV, которые могут перекрываться с генами, влияющими на устойчивость к заболеваниям и собранными в базе данных OMIM (Online Mendelian Inheritance in Man).

В связи с тем, что CNV трудно идентифицируемы независимо от используемых платформ наилучшим подходом, по-видимому, будет являться консервативное использование нескольких методов для выявления вариации числа копий. Учитывая отсутствие комплексного исследования CNV у овец, в 2016 году G. M. Jenkins и соавторы [27] с помощью специально разработанной платформы Roche-NimbleGen 2.1 M CGH и целого ряда биоинформационных методов обнаружили 3488 CNV регионов на аутосомах овец со средней длиной 19 Kb. Обнаруженные CNVR были сопоставимы с выявленными в предыдущих работах. G. M. Jenkins с соавторами [27] удалось обнаружить 31 % CNVR, идентифицированных в первом исследовании CNV у овец [58], и 16-62 % CNVR, обнаруженных ранее в геноме крупного рогатого скота [45, 46, 47, 48, 63, 64]. Идентифицированные CNV покрывали 2,7 % генома овец, что примерно в семь раз больше, чем было установлено ранее [58], что, вероятно, связано с большим размером исследуемой выборки животных. Из 3335 аутомсомных регионов CNV 1335 (40 %) перекрывали кодирующую tandemную последовательность, охватывающую гены *ASIP* и *AHCY*, а также и область промотора гена *ITCH*, что подтверждается наличием дупликации, приводящей к появлению белой шерсти у овец масти агути, о котором ранее также сообщали B. J. Norris и V. A. Whan [65]. Результаты исследования, в целом, были сопоставимы с диапазоном оценок для крупного рогатого скота [45, 46, 47, 48, 63, 64].

Недавние исследования показали, что различное число копий генов в геноме оказывают различное влияние на рост и развитие овец. В 2016 г. C. Zhu и соавторы [66] установили, что такие гены, как *PPARA*, *RXRA*, *KLF11*, *ADD1*, *FASN*, *PPP1CA* влияют на процессы отложения жира и формирование



жирного хвоста у овец. Эти области CNV содержат большое количество генов, связанных с метаболизмом жира и активностью ГТфазных ферментов. L. Yang и соавторы [26] установили, что у овец области CNV, которые были связаны с развитием мышечной ткани плода, синтезом простагландина и размерами костей, перекрывались с генами *BTG3*, *PTGS1* и *PSPH* соответственно.

В 2017 году Q. Ma и соавторы [34], применив ДНК-чип высокой плотности *Ovine SNP 600K BeadChip* (Illumina Inc., США), провели генотипирование 48 овец из китайской породы тан и выявили 1296 областей CNV, что составляет 4,7 % генома овец. Результатом исследований стало построение CNV-карты генома китайских местных овец [34].

В ряде недавних исследований была установлена корреляция экономически значимых признаков животных с изменением числа копий. R. Jiang и соавторы [67] обнаружили, что ген *SHE* находится в QTL, связанном с процентным содержанием молочного жира и плотностью костей, и идентифицировали вариацию числа копий гена *SHE* длиной 2000 п.о. Ген *SHE*, расположенный на хромосоме 1, содержит 6 экзонов и кодирует белок с 495 аминокислотами. Ген *SHE* является важным членом доменного семейства SH2, структурно консервативного белкового домена, встречающегося в онкопротеине Src [68] и многих других белках внутриклеточной сигнальной трансдукции [69]. Функция этого домена состоит в том, чтобы распознавать фосфорилированное состояние остатков тирозина и инициировать различные реакции, что в итоге приводит к изменению уровня экспрессии генов или других клеточных реакций. Ассоциативный анализ числа копий гена *SHE* с признаками роста и развития в четырех китайских породах овец показал, что носители нормального и увеличенного числа копий демонстрировали лучшие фенотипы. Полученные результаты позволяют предположить, что CNV гена *SHE* влияют на развитие хозяйственно полезных признаков и могут быть предложены в качестве перспективного кандидата в овцеводстве.

X. Wang и соавторы [70] обнаружили, что ген *ORMDL1* характеризуется вариацией числа копий 2800 п. о. на хромосоме 2 у четырех китайских пород овец, и что этот CNV содержится в области с высокой плотностью QTL различных признаков экономической

важности, в том числе ассоциированных с формированием мышц и жира. Ген *ORMDL1*, член семейства *ORMDL*, является белок-кодирующим геном, который связан с метаболизмом сфинголипидов, а также регулирует биосинтез и гомеостаз церамидов в клетках млекопитающих [71, 72, 73]. Анализ показал, что CNV гена *ORMDL1* коррелирует с показателями роста туловища, обхватом груди и окружностью берцовой кости ( $P < 0,05$ ), а также оказывает значительное влияние на массу тела, высоту в холке, длину туловища, глубину груди и высоту в крестце ( $P < 0,05$ ). На основании полученных данных CNV гена *ORMDL1* можно рекомендовать в качестве перспективного молекулярного маркера для использования в селекционных программах овец.

Z. Feng и соавторы [74] обнаружили, что ген *PIGY*, расположенный на хромосоме 6, содержит область CNV в экзоне 2 длиной 3600 п. о., перекрывающуюся с 28 QTL, связанных с такими признаками, как плотность мышц и масса тела. Исследователи провели скрининг CNV гена *PIGY* у трех китайских пород овец. Ген *PIGY* инициирует биосинтез гликозилфосфатидилинозитол-ацетилглюкозаминилтрансфераз (GPI) [75] и играет важную роль в межклеточном взаимодействии. Проведенный анализ ассоциаций показал значительное влияние CNV гена *PIGY* на живую массу, обхват груди и окружность берцовой кости овец. В связи с тем, что овцы с увеличенным количеством копий характеризовались лучшими показателями роста, чем овцы с другими типами CNV, то можно diskutieren о возможности применения CNV гена *PIGY* в качестве нового молекулярно-генетического маркера у овец.

Z. Yang и соавторы [76] изучали потенциальные ассоциации CNV гена *BAG4* с показателями роста и развития у трех китайских пород овец. Ген *BAG4* локализован в области QTL признаков роста и формирования вымени у овец. CNV гена *BAG4* был статистически связан с ростом туловища ( $P < 0,05$ ), с косой длиной туловища ( $P < 0,05$ ) и высотой в крестце ( $P < 0,05$ ). CNV гена *BAG4* может рассматриваться как перспективный кандидат для внедрения в маркерную селекцию овец.

Информация о наиболее значимых функциональных генах-кандидатах, частично или полностью локализованных в регионах CNV у разных видов сельскохозяйственных животных, обобщена в таблице.

Таблица – Наиболее значимые функциональные гены-кандидаты, частично или полностью локализованные в регионах CNV, у разных видов сельскохозяйственных животных /

Table – The most significant functional candidate genes, partially or fully overlapped in the CNV regions, in different livestock species

Ген / Gene	Xp / Chr	Функции / Function	Источник / Reference
1	2	3	4
<i>Идентифицированные гены в регионах CNV у свиней / Genes overlapped in CNV regions in pigs</i>			
<i>ANP32B</i> (кислотный лейцин - богатый ядерный фосфопротеин 32 член семьи B) / <i>ANP32B</i> (acidic leucine-rich nuclear phosphoprotein 32 family member B)	1	Фактор прогрессирования клеточного цикла и выживания клетки. Требуется для перехода от G1 к S фазе / Factor of cell cycle progression and cell survival. It is required for the transition from G1 to S phase	C. Chen et al. (2012) [53]
<i>BSCL2</i> (сейпин, связанный с биогенезом липидных капель) / <i>BSCL2</i> (seipin lipid droplet biogenesis associated)	2	Особенно активен в моторных нейронах и в адипоцитах. Участвует в формировании липидных капель, хранящих молекулы жира / High activity in motor neurons and adipocytes. Formation of lipid droplets storing fat molecules	C. Chen et al. (2012) [53]
<i>LTBP3</i> (скрытый трансформирующий фактор роста бета-связывающий белок 3) / <i>LTBP3</i> (latent transforming growth factor beta binding protein 3)	2	Участвует в субклеточной локализации, играет структурную роль во внеклеточном матриксе / A role in subcellular localization and in the extracellular matrix	C. Chen et al. (2012) [53]
<i>KIT</i> (KIT протоонкоген рецептор тирозинкиназы) / <i>KIT</i> (KIT proto-oncogene, receptor tyrosine kinase)	5	Контролируют пролиферацию клеток. Развитие и функционирование половых клеток, гемопоэтических стволовых клеток, лейкоцитов и меланоцитов / Cell proliferation control. Involved in development and functioning of germ cells, hematopoietic stem cells, white blood cells and melanocytes	C. Chen et al. (2012) [53]
<i>GDF3</i> (фактор дифференциации роста 3) / <i>GDF3</i> (3growth differentiation factor 3)	5	Регулирует рост и дифференцировки клеток в эмбриональный и постнатальный периоды / Regulates cell growth and cell differentiation in the embryonic and postnatal periods	C. Chen et al. (2012) [53]
<i>GYS1</i> (гликоген синтаза 1) / <i>GYS1</i> (glycogen synthase 1)	6	Синтез фермента мышечной гликогенсинтазы, необходимой для связывания молекул глюкозы с образованием гликогена / Synthesis of the muscle glycogen synthase required for binding glucose molecules to form glycogen	C. Chen et al. (2012) [53]
<i>CAVI</i> (кавеолин-1) / <i>CAVI</i> (caveolin-1)	18	Обеспечивает межклеточный и внутриклеточный транспорт молекул. Поддержание структуры клетки. Регуляция пролиферации, дифференцировки и апоптоза клеток / Intercellular and intracellular transport of molecules. Maintaining the cell structure. Regulation of cell proliferation, differentiation, and apoptosis	C. Chen et al. (2012) [53]
<i>Идентифицированные гены в регионах CNV у крупного рогатого скота / Genes overlapped in CNV regions in cattle</i>			
<i>GBP4</i> (связывающий гуанила белок 4) / <i>GBP4</i> (guanylate-binding protein 4)	3	Влияет на рост взрослого животного / Affects the growth of an adult animal	X. K.Cao et al. (2018) [51]
<i>MYH3</i> (ген тяжелой цепи 3 миозина) / <i>MYH3</i> (myosin heavy chain 3)	19	Регуляция экспрессии генов, участвующих в формировании скелетных мышц и сердца / Regulation of expression of genes involved in the formation of skeletal muscles and heart	Y. Xu et al. (2014) [35]



1	2	3	4
Идентифицированные гены в регионах CNV у овец/ Genes overlapped in CNV regions in sheep			
<i>BTG3</i> (BTG антипролиферативный фактор 3) / <i>BTG3</i> (BTG anti-proliferation factor 3)	1	Развитие мышечной ткани плода. Потенциально влияет на процесс нейрогенеза в центральной нервной системе / The development of the fetal muscular tissue. Potentially affects the neurogenesis in the central nervous system	L. Yang et al. (2018) [26]
<i>SHE</i> (Src гомологическая домен - содержащая E фосфатаза 2) / <i>SHE</i> (Src homology 2 domain containing E)	1	Ген коррелирует с QTL ассоциированным с процентным содержанием жира в молоке и плотностью костей / The gene correlates with QTL associated with milk fat percentage and bone density	R. Jiang et al. (2019) [67]
<i>ORMDL1</i> (ORMDL регулятор биосинтеза сфинголипидов 1) / <i>ORMDL1</i> (ORMDL sphingolipid biosynthesis regulator 1)	2	Регулирование метаболизма и гомеостаза церамида / Regulation of ceramide metabolism and homeostasis	X. Wang et al. (2019) [70]
<i>PTGS1</i> (простагландин-эндопероксид-синтаза 1) / <i>PTGS1</i> (prostaglandin-endoperoxide synthase 1)	3	Катализирует превращение арахидоната в простагландин, влияющего на плодовитость. Регуляция ангиогенеза в эндотелиальных клетках / Catalyzes the conversion of arachidonic acid to prostaglandin, which affects fertility. Regulation of angiogenesis in endothelial cells	L. Yang et al. (2018) [26]
<i>PPARA</i> (альфа-рецептор, активируемый пероксисомным пролифератором) / <i>PPARA</i> (peroxisome proliferator activated receptor alpha)	3	Связывает ненасыщенные жирные кислоты в печени, почках, скелетных мышцах, сердце для активации генов-регуляторов метаболизма жирных кислот / Binds unsaturated fatty acids in the liver, kidneys, skeletal muscles, and heart to activate of genes regulating fatty acid metabolism	C. Zhu et al. (2016) [66]
<i>RXRA</i> (альфа-рецептор ретиноида X) / <i>RXRA</i> (retinoid X receptor alpha)	3	Играет роль в процессах поддержания липидного гомеостаза и развитии плода / Involved in lipid homeostasis maintaining and fetal development	C. Zhu et al. (2016) [66]
<i>KLF11</i> (Крупнель-подобный фактор 11) / <i>KLF11</i> (Kruppel like factor 11)	3	Регулирует транскрипцию бурого жира в адипоцитах / Regulates the transcription of brown fat in adipocytes	C. Zhu et al. (2016) [66]
<i>ADD1</i> (аддуцин 1) / <i>ADD1</i> (adducin 1)	6	Важная роль в процессах дифференцировки жировых клеток и поддержании гомеостаза холестерина / An important role in fat cell differentiation and in maintaining of cholesterol homeostasis	C. Zhu et al. (2016) [66]
<i>PIGY</i> (фосфатидилинозитол гликан-анкорный биосинтез класса Y) / <i>PIGY</i> (phosphatidylinositol glycan anchor biosynthesis class Y)	6	Иницирует биосинтез GPI и играет важную роль в межклеточном взаимодействии / Initiates GPI biosynthesis and plays an important role in intercellular interaction	Z. Feng et al. (2020) [74]
<i>FASN</i> (синтаза жирных кислот) / <i>FASN</i> (fatty acid synthase)	11	Ключевая роль в процессах биосинтеза и анаболизма жирных кислот. Регулирует депонирование жира / A key role in fatty acid biosynthesis and anabolism. Regulates fat deposition	C. Zhu et al. (2016) [66]
<i>ASIP</i> (сигнальный белок агути) / <i>ASIP</i> (agouti signaling protein), <i>AHCY</i> (аденозилгомоцистеиназа) / <i>AHCY</i> (adenosylhomocysteinase), <i>ITCH</i> (E3 убиквитинная протеиновая лигаза) / <i>ITCH</i> (itchy E3 ubiquitin protein ligase)	13	Антагонисты альфа-меланоцит-стимулирующего гормона. Участвуют в нейроэндокринной системе действия меланокортина. Регуляция липидного обмена в адипоцитах / Antagonists of alpha-melanocyte-stimulating hormone. Involved in the neuroendocrine system in regards with melanocortin. Regulation of lipid metabolism in adipocytes	B. J. Norris и V. A. Whan (2008) [65] *
<i>PPP1CA</i> (протеинфосфатаза 1 каталитическая субъединица альфа) / <i>PPP1CA</i> (protein phosphatase 1 catalytic subunit alpha)	21	Дефосфорилирование и инактивация гликогенсинтазы в скелетных мышцах / Dephosphorylation and inactivation of glycogen synthase in skeletal muscles	C. Zhu et al. (2016) [66]

1	2	3	4
<i>PSPH</i> (фосфосеринфосфатаза) / <i>PSPH</i> (phosphoserine phosphatase)	24	Отвечает за последний этап биосинтеза L-серина. Потенциально влияет на цвет костной ткани / Responsible for the last stage of L-serine biosynthesis. Potentially affects a color of bones	L. Yang et al. (2018) [26]
<i>BAG4</i> (Bcl-2-ассоциированный с атаногеном 4) / <i>BAG4</i> (BAG co-chaperone)	26	Активация различных клеточных процессов, в том числе реакции на стресс, пролиферации, миграции и апоптоза. Расположен внутри QTL, ассоциированного с признаками развития туловища и вымени / Involved in the activation of various cellular processes, including responses to stress, proliferation, migration, and apoptosis. Located within the QTL associated with traits for body stature and udder development	Z. Yang et al. (2020) [76]

*Идентифицированные гены, перекрывающиеся в CNV у коз /  
Genes overlapped in CNV regions in goats*

<i>SHH</i> (молекула сигнального пути Sonic Hedgehog) / <i>SHH</i> (sonic hedgehog signaling molecule)	4	Ингибирует образование жира и ускоряет миогенную и остеогенную дифференцировку. Запускает химический сигнал, необходимый для развития эмбриона. Важен для развития головного и спинного мозга / Inhibits fat formation and accelerates myogenic and osteogenic differentiation. Triggers a chemical signal required for embryo development. Crucial for brain and spinal cord development	M. Liu et al. (2019) [25]
<i>KDM5B</i> (лизиндеметиلاза 5B) / <i>KDM5B</i> (lysine demethylase 5B)	16	Деметилюет «Lys-4» гистона H3 / Demethylation «Lys-4» of histone H3	M. Liu et al. (2019) [25]
<i>EDNRA</i> (рецептор эндотелина тип A) / <i>EDNRA</i> (endothelin receptor type A)	17	Опосредует действие эндотелина-1 путем ассоциации с G-белками, активирующими кальциевую систему второго мессенджера. Играет роль в развитии мышц плода / Mediates endothelin-1 by association with G-proteins that activate the calcium system of the second messenger. Involved in fetal muscle development	M. Liu et al. (2019) [25]
<i>NR3C2</i> (член группы C2 подсемейства ядерных рецепторов 3) / <i>NR3C2</i> (nuclear receptor subfamily 3 group C member 2)	17	Регуляция количества натрия и кровяного давления. Потенциально ассоциирован с развитием адаптаций к засушливому климату / Regulates sodium level blood pressure. Potentially associated with adaptation traits to arid climate	M. Liu et al. (2019) [25]
<i>PRP1</i> (связанный с пролактином белок 1) / <i>PRP1</i> (prolactin-related protein-1) и <i>PRP6</i> (связанный с пролактином белок 6) / <i>PRP6</i> (prolactin-related protein-6)	18	Кодируемый белок связывает андрогенный рецептор, обеспечивая связь между активацией транскрипции и сплайсингом. Играют важную регуляторную роль в процессах воспроизводства и беременности / Encoded protein binds an androgen receptor, providing a link between transcription activation and splicing. An important regulatory role in reproduction and pregnancy	R. Q. Zhang et al. (2019) [57]
<i>MYLK4</i> (член семейства киназ легкой цепи миозина 4) / <i>MYLK4</i> (myosin light chain kinase family member 4)	23	Влияет на сокращение мышц. Распространен в гладких мышцах, сердечной мышце, скелетных мышцах / An important role in muscle contraction. Found in smooth muscle, heart muscle, and skeletal muscle	S. Y. Shi et al. (2019) [33]

\* L. Fontanesi и соавторы (2009) [56] выявили ту же группу генов, перекрывающихся в CNV у коз /

\* L. Fontanesi et al (2009) [56] identified the same group of genes, which have been overlapped in the CNV in goats

**Выводы.** Таким образом, суммируя результаты проведенного анализа мирового опыта в области поиска регионов вариации числа копий и методов их идентификации, можно сделать несколько выводов.

1. Прежде всего, следует отметить, что локусы значительного количества функциональных генов-кандидатов, отвечающих за разнообразные экономически значимые признаки, полностью или частично перекрываются.

ся с регионами CNV. В частности, каждый вид сельскохозяйственных животных, рассматриваемый в настоящем обзоре, характеризовался наличием генов, перекрывающихся с регионами CNV и локализованных внутри или рядом с QTL показателей роста или мясной продуктивности. В связи с этим, полученные результаты так же, как и большее покрытие генома по сравнению с SNP-маркерами, свидетельствует об огромном потенциале CNV выявленных генов в качестве ДНК-маркеров для использования в мясном животноводстве.

2. Несмотря на то, что разработан широкий спектр лабораторных методов для выявления CNV они требуют дальнейшего совершенствования. Чувствительность обнаружения CNV сильно зависит от выбранной технологии.

Вероятно, наиболее адекватным и информативным является комплексный подход, основанный на использовании ДНК-чипов высокой плотности (или NGS данных) с последующим подтверждением присутствия обнаруженных CNV с помощью количественной ПЦР.

3. Идентификация CNV была проведена только у китайских пород овец, поэтому полученные результаты будет методически неверно экстраполировать на российские породы овец. В связи с этим, на основе полученных ранее данных полногеномного SNP генотипирования [77, 78] были начаты исследования распространения CNV в геномах отечественных пород овец с целью внедрения полученной информации в маркерную селекцию для улучшения экономически значимых показателей овец.

### References

1. Feuk L., Carson A. R., Scherer S. W. Structural variation in the human genome. *Nat Rev Genet.* 2006;7(2):85-97. DOI: <https://doi.org/10.1038/nrg1767>
2. Cloup A., Vidal O., Amills M. Copy number variation in the genomes of domestic animals. *Anim Genet.* 2012;43(5):503-517. DOI: <https://doi.org/10.1111/j.1365-2052.2012.02317.x>
3. Egan C. M., Sridhar S., Wigler M., Hall I. M. Recurrent DNA copy number variation in the laboratory mouse. *Nat Genet.* 2007;39(11):1384-1389. DOI: <https://doi.org/10.1038/ng.2007.19>
4. 1000 Genomes Project Consortium, Abecasis G. R., Altshuler D., Auton A., Brooks L. D., Durbin R. M., et al. A map of human genome variation from population-scale sequencing. *Nature.* 2010;467:1061-1073. DOI: <https://doi.org/10.1038/nature09534>
5. Michaelson J. J., Shi Y., Gujral M., Zheng H., Malhotra D., Jin X., et al. Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell.* 2012;151(7):1431-1442. DOI: <https://doi.org/10.1016/j.cell.2012.11.019>
6. Gu W., Zhang F., Lupski J. R. Mechanisms for human genomic rearrangements. *Pathogenetics.* 2008;1(1):4. DOI: <https://doi.org/10.1186/1755-8417-1-4>
7. Conrad D. F., Pinto D., Redon R., Feuk L., Gokcumen O., Zhang Y., et al. Origins and functional impact of copy number variation in the human genome. *Nature.* 2010;464(7289):704-712. DOI: <https://doi.org/10.1038/nature08516>
8. Pang A. W., MacDonald J. R., Pinto D., Wei J., Rafiq M. A., Conrad D. F., et al. Towards a comprehensive structural variation map of an individual human genome. *Genome Biol.* 2010;11(5):R52. DOI: <https://doi.org/10.1186/gb-2010-11-5-r52>
9. Iafrate A. J., Feuk L., Rivera M. N., Listewnik M. L., Donahoe P. K., Ying Y., et al. Detection of large-scale variation in the human genome. *Nat Genet.* 2004;36(9):949-951. DOI: <https://doi.org/10.1038/ng1416>
10. Kim P. M., Lam H. Y., Urban A. E., Korbel J. O., Affourtit J., Grubert F., et al. Analysis of copy number variants and segmental duplications in the human genome: Evidence for a change in the process of formation in recent evolutionary history. *Genome Res.* 2008;18(12):1865-1874. DOI: <https://doi.org/10.1101/gr.081422.108>
11. Berglund J., Nevalainen E. M., Molin A. M., Perloski M., The LUPA Consortium, André C., et al. Novel origins of copy number variation in the dog genome. *Genome Biology.* 2012;13(8):R73. DOI: <https://doi.org/10.1186/gb-2012-13-8-r73>
12. Zhang F., Gu W., Hurler M. E., Lupski J. R. Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet.* 2009;10:451-481. DOI: <https://doi.org/10.1146/annurev.genom.9.081307.164217>
13. Conrad D. F., Andrews T. D., Carter N. P., Hurler M. E., Pritchard J. K. A high-resolution survey of deletion polymorphism in the human genome. *Nat Genet.* 2006;38(1):75-81. DOI: <https://doi.org/10.1038/ng1697>
14. Emerson J. J., Cardoso-Moreira M., Borevitz J. O., Long M. Natural selection shapes genome-wide patterns of copy-number polymorphism in *Drosophila melanogaster*. *Science.* 2008;320(5883):1629-1631. DOI: <https://doi.org/10.1126/science.1158078>
15. Cooper G. M., Nickerson D. A., Eichler E. E. Mutational and selective effects on copy-number variants in the human genome. *Nat Genet.* 2007;39(7 Suppl):S22-S29. DOI: <https://doi.org/10.1038/ng2054>
16. Zhang L., Jia S., Yang M., Xu Y., Li C., Sun J., et al. Detection of copy number variations and their effects in Chinese bulls. *BMC Genomics.* 2014;15(1):480. DOI: <https://doi.org/10.1186/1471-2164-15-480>
17. Stankiewicz P., Lupski J. R. Structural variation in the human genome and its role in disease. *Annu Rev Med.* 2010;61:437-455. DOI: <https://doi.org/10.1146/annurev-med-100708-204735>
18. McCarroll S. A., Kuruvilla F. G., Korn J. M., Cawley S., Nemesh J., Wysoker A., et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet.* 2008;40(10):1166-1174. DOI: <https://doi.org/10.1038/ng.238>

19. Wain L. V., Armour J. A., Tobin M. D. Genomic copy number variation, human health, and disease. *Lancet*. 2009;374(9686):340-350. DOI: [https://doi.org/10.1016/S0140-6736\(09\)60249-X](https://doi.org/10.1016/S0140-6736(09)60249-X)
20. Schoumans J., Ruivenkamp C., Holmberg E., Kyllerman M., Anderlid B. M., Nordenskjöld M. Detection of chromosomal imbalances in children with idiopathic mental retardation by array based comparative genomic hybridisation (array-CGH). *J Med Genet*. 2005;42(9):699-705. DOI: <https://doi.org/10.1136/jmg.2004.029637>
21. He Y., Hoskins J. M., McLeod H. L. Copy number variants in pharmacogenetic genes. *Trends Mol Med*. 2011;17(5):244-251. DOI: <https://doi.org/10.1016/j.molmed.2011.01.007>
22. Li X., Tan L., Liu X., Lei S., Yan T., Chen X., et al. A genome wide association study between copy number variation (CNV) and human height in Chinese population. *J Genet Genomics*. 2010;37(12):779-785. DOI: [https://doi.org/10.1016/S1673-8527\(09\)60095-3](https://doi.org/10.1016/S1673-8527(09)60095-3)
23. Ma Y. L., Wen Y. F., Cao X. K., Cheng J., Huang Y. Z., Ma Y., et al. Copy number variation (CNV) in the IGF1R gene across four cattle breeds and its association with economic traits. *Arch Anim Breed*. 2019;62(1):171-179. DOI: <https://doi.org/10.5194/aab-62-171-2019>
24. Da Silva J. M., Giachetto P. F., da Silva L. O., Cintra L. C., Paiva S. R., Yamagishi M. E., et al. Genome-wide copy number variation (CNV) detection in Nelore cattle reveals highly frequent variants in genome regions harboring QTLs affecting production traits. *BMC Genomics*. 2016;17:454. DOI: <https://doi.org/10.1186/s12864-016-2752-9>
25. Liu M., Zhou Y., Rosen B. D., Van Tassell C. P., Stella A., Tosser-Klopp G., et al. Diversity of copy number variation in the worldwide goat population. *Heredity (Edinb)*. 2019;122(5):636-646. DOI: <https://doi.org/10.1038/s41437-018-0150-6>
26. Yang L., Xu L., Zhou Y., Liu M., Wang L., Kijas J. W., Zhang H., Li L., Liu G. E. Diversity of copy number variation in a worldwide population of sheep. *Genomics*. 2018;110(3):143-148. DOI: <https://doi.org/10.1016/j.ygeno.2017.09.005>
27. Jenkins G. M., Goddard M. E., Black M. A., Brauning R., Auvray B., Dodds K. G., Kijas J. W., Cockett N., McEwan J. C. Copy number variants in the sheep genome detected using multiple approaches. *BMC Genomics*. 2016;17:441. DOI: <https://doi.org/10.1186/s12864-016-2754-7>
28. Alvarez C. E., Akey J. M. Copy number variation in the domestic dog. *Mamm Genome*. 2012;23(1-2):144-163. DOI: <https://doi.org/10.1007/s00335-011-9369-8>
29. 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: <https://doi.org/10.1016/j.ygeno.2012.07.001>
30. Fadista J., Nygaard M., Holm L. E., Thomsen B., Bendixen C. A snapshot of CNVs in the pig genome. *PLoS One*. 2008;3(12):e3916. DOI: <https://doi.org/10.1371/journal.pone.0003916>
31. Ramayo-Caldas Y., Castelló A., Pena R. N., Alves E., Mercadé A., Souza C. A., Fernández A. I., Perez-Enciso M., Folch J. M. Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip. *BMC Genomics*. 2010;11:593. DOI: <https://doi.org/10.1186/1471-2164-11-593>
32. Fowler K. E., Pong-Wong R., Bauer J., Clemente E. J., Reitter C. P., Affara N. A., Waite S., Walling G. A., Griffin D. K. Genome wide analysis reveals single nucleotide polymorphisms associated with fatness and putative novel copy number variants in three pig breeds. *BMC Genomics*. 2013;14:784. DOI: <https://doi.org/10.1186/1471-2164-14-784>
33. Shi S. Y., Li L. J., Zhang Z. J., Wang E.-Y., Wang J., Xu J.-W., Liu H.-B., Wen Y.-F., He H., Lei C.-Z., Chen H., Huang Y.-Z. Copy number variation of MYLK4 gene and its growth traits of *Capra hircus* (goat). *Anim Biotechnol*. 2019;1-6. DOI: <https://doi.org/10.1080/10495398.2019.1635137>
34. Ma Q., Liu X., Pan J., Ma L., Ma Y., He X., Zhao Q., Pu Y., Li Y., Jiang L. Genome-wide detection of copy number variation in Chinese indigenous sheep using an ovine high-density 600 K SNP array. *Sci Rep*. 2017;7(1):912. DOI: <https://doi.org/10.1038/s41598-017-00847-9>
35. Xu Y., Shi T., Cai H., Zhou Y., Lan X., Zhang C., Lei C., Qi X., Chen H. Associations of MYH3 gene copy number variations with transcriptional expression and growth traits in Chinese cattle. *Gene*. 2014;535(2):106-111. DOI: <https://doi.org/10.1016/j.gene.2013.11.057>
36. D'haene B., Vandesompele J., Hellemans J. Accurate and objective copy number profiling using real-time quantitative PCR. *Methods*. 2010;50(4):262-270. DOI: <https://doi.org/10.1016/j.ymeth.2009.12.007>
37. Gryadunov D., Dementieva E., Mikhailovich V., Nasedkina T., Rubina A., Savvateeva E., Fesenko E., Chudinov A., Zimenkov D., Kolchinsky A., Zasedatelev A. Gel-based microarrays in clinical diagnostics in Russia. *Expert Rev Mol Diagn*. 2011;11(8):839-853. DOI: <https://doi.org/10.1586/erm.11.73>
38. Marasso S. L., Mombello D., Cocuzza M., Casalena D., Ferrante I., Nesca A., Poiklik P., Rekker K., Aaspolu A., Ferrero S., Pirri C. F. A polymer lab-on-a-chip for genetic analysis using the arrayed primer extension on microarray chips. *Biomed Microdevices*. 2014;16(5):661-670. DOI: <https://doi.org/10.1007/s10544-014-9869-x>
39. Peiffer D. A., Le J. M., Steemers F. J., Chang W., Jenniges T., Garcia F. et al. High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome Res*. 2006;16(9):1136-1148. DOI: <https://doi.org/10.1101/gr.5402306>
40. Schouten J. P., McElgunn C. J., Waaijer R., Zwijnenburg D., Diepvens F., Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res*. 2002;30(12):e57. DOI: <https://doi.org/10.1093/nar/gnf056>
41. Carter N. P. Methods and strategies for analyzing copy number variation using DNA microarrays. *Nat Genet*. 2007;39(7 Suppl):S16-S21. DOI: <https://doi.org/10.1038/ng2028>
42. Rincon G., Weber K. L., Eenennaam A. L., Golden B. L., Medrano J. F. Hot topic: performance of bovine high-density genotyping platforms in Holsteins and Jerseys. *J Dairy Sci*. 2011;94(12):6116-6121. DOI: <https://doi.org/10.3168/jds.2011-4764>

43. Ong F. S., Lin J. C., Das K., Grosu D. S., Fan J. B. Translational utility of next-generation sequencing. *Genomics*. 2013;102(3):137-139. DOI: <https://doi.org/10.1016/j.ygeno.2013.04.012>
44. Robasky K., Lewis N. E., Church G. M. The role of replicates for error mitigation in next-generation sequencing. *Nat Rev Genet*. 2014;15(1):56-62. DOI: <https://doi.org/10.1038/nrg3655>
45. Bae J. S., Cheong H. S., Kim L. H., NamGung S., Park T. J., Chun J.-Y., Kim J. Y., Pasaje C. F. A., Lee J. S., Shin H. D. Identification of copy number variations and common deletion polymorphisms in cattle. *BMC Genomics*. 2010;11:232. DOI: <https://doi.org/10.1186/1471-2164-11-232>
46. Bickhart D. M., Hou Y., Schroeder S. G., Alkan C., Cardone M. F., Matukumalli L. K., et al. Copy number variation of individual cattle genomes using next-generation sequencing. *Genome Res*. 2012;22(4):778-790. DOI: <https://doi.org/10.1101/gr.133967.111>
47. Hou Y., Liu G. E., Bickhart D. M., Cardone M. F., Wang K., Kim E., Matukumalli L. K., Ventura M., Song J., VanRaden P. M., Sonstegard T. S., Van Tassell C. P. Genomic characteristics of cattle copy number variations. *BMC Genomics*. 2011;12:127. DOI: <https://doi.org/10.1186/1471-2164-12-127>
48. Kijas J. W., Barendse W., Barris W., Harrison B., McCulloch R., McWilliam S., Whan V. Analysis of copy number variants in the cattle genome. *Gene*. 2011;482(1-2):73-77. DOI: <https://doi.org/10.1016/j.gene.2011.04.011>
49. Cicconardi F., Chillemi G., Tramontano A., Marchitelli C., Valentini A., Ajmone-Marsan P., Nardone A. Massive screening of copy number population-scale variation in *Bos taurus* genome. *BMC Genomics*. 2013;14:124. DOI: <https://doi.org/10.1186/1471-2164-14-124>
50. Fontanesi L., Martelli P. L., Beretti F., Riggio V., Dall'Olio S., Colombo M., Casadio R., Russo V., Portolano B. An initial comparative map of copy number variations in the goat (*Capra hircus*) genome. *BMC Genomics*. 2010;11:639. DOI: <https://doi.org/10.1186/1471-2164-11-639>
51. Cao X. K., Huang Y. Z., Ma Y. L., Cheng J., Qu Z. X., Ma Y., Bai Y. Y., Tian F., Lin F. P., Ma Y. L., Chen H. Integrating CNVs into meta-QTL identified GBP4 as positional candidate for adult cattle stature. *Funct Integr Genomics*. 2018;18(5):559-567. DOI: <https://doi.org/10.1007/s10142-018-0613-0>
52. Salmon Hillbertz N. H., Isaksson M., Karlsson E. K., Hellmén E., Pielberg G. R., Savolainen P., et al. Duplication of FGF3, FGF4, FGF19 and ORAOV1 causes hair ridge and predisposition to dermoid sinus in Ridgeback dogs. *Nat Genet*. 2007;39(11):1318-1320. DOI: <https://doi.org/10.1038/ng.2007.4>
53. Chen C., Qiao R., Wei R., Yuanmei G., Ai H., Ma J., Ren J., Huang L. A comprehensive survey of copy number variation in 18 diverse pig populations and identification of candidate copy number variable genes associated with complex traits. *BMC Genomics*. 2012;13:733. DOI: <https://doi.org/10.1186/1471-2164-13-733>
54. Wang K., Li M., Hadley D., Liu R., Glessner J., Grant S. F. A., Hakonarson H., Bucan M. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res*. 2007;17(11):1665-1674. DOI: <https://doi.org/10.1101/gr.6861907>
55. Johansson Moller M., Chaudhary R., Hellmén E., Höyheim B., Chowdhary B., Andersson L. Pigs with the dominant white coat color phenotype carry a duplication of the KIT gene encoding the mast/stem cell growth factor receptor. *Mamm Genome*. 1996;7(11):822-830. DOI: <https://doi.org/10.1007/s003359900244>
56. Fontanesi L., Beretti F., Riggio V., Gómez González E., Dall'Olio S., Davoli R., Russo V., Portolano B. Copy number variation and missense mutations of the agouti signaling protein (ASIP) gene in goat breeds with different coat colors. *Cytogenet Genome Res*. 2009;126(4):333-347. DOI: <http://dx.doi.org/10.1159/000268089>
57. Zhang R. Q., Wang J. J., Zhang T., Zhai H. L., Shen W. Copy-number variation in goat genome sequence: A comparative analysis of the different litter size trait groups. *Gene*. 2019;696:40-46. DOI: <https://doi.org/10.1016/j.gene.2019.02.027>
58. Fontanesi L., Beretti F., Martelli P. L., Colombo M., Dall'Olio S., Occidente M., Portolano B., Casadio R., Matassino D., Russo V. A first comparative map of copy number variations in the sheep genome. *Genomics*. 2011;97(3):158-165. DOI: <https://doi.org/10.1016/j.ygeno.2010.11.005>
59. Liu J., Zhang L., Xu L., Ren H., Lu J., Zhang X., Zhang S., Zhou X., Wei C., Zhao F., Du L. Analysis of copy number variations in the sheep genome using 50K SNP BeadChip array. *BMC Genomics*. 2013;14:229. DOI: <https://doi.org/10.1186/1471-2164-14-229>
60. International Chicken Genome Sequencing Consortium. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*. 2004;432(7018):695-716. DOI: <https://doi.org/10.1038/nature03154>
61. Bovine Genome Sequencing and Analysis Consortium, Elsik C. G., Tellam R. L., Worley K. C. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*. 2009;324(5926):522-528. DOI: <https://doi.org/10.1126/science.1169588>
62. Humphray S. J., Scott C. E., Clark R., Marron B., Bender C., Camm N., et al. A high utility integrated map of the pig genome. *Genome Biol*. 2007;8(7):R139. DOI: <https://doi.org/10.1186/gb-2007-8-7-r139>
63. Fadista J., Thomsen B., Holm L. E., Bendixen C. Copy number variation in the bovine genome. *BMC Genomics*. 2010;11:284. DOI: <https://doi.org/10.1186/1471-2164-11-284>
64. Jiang L., Jiang J., Yang J., Liu X., Wang J., Wang H., Ding X., Liu J., Zhang Q. Genome-wide detection of copy number variations using high-density SNP genotyping platforms in Holsteins. *BMC Genomics*. 2013;14:131. DOI: <https://doi.org/10.1186/1471-2164-14-131>
65. Norris B. J., Whan V. A. A gene duplication affecting expression of the ovine ASIP gene is responsible for white and black sheep. *Genome Res*. 2008;18(8):1282-1293. DOI: <https://doi.org/10.1101/gr.072090.107>
66. Zhu C., Fan H., Yuan Z., Hu S., Ma X., Xuan J., Wang H., Zhang L., Wei C., Zhang Q., Zhao F., Du L. Genome-wide detection of CNVs in Chinese indigenous sheep with different types of tails using ovine high-density 600K SNP arrays. *Sci Rep*. 2016;6:27822. DOI: <https://doi.org/10.1038/srep27822>



67. Jiang R., Cheng J., Cao X. K., Ma Y. L., Chaogetu B., Huang Y. Z., Lan X. Y., Lei C. Z., Hu L. Y., Chen H. Copy Number Variation of the SHE Gene in Sheep and Its Association with Economic Traits. *Animals (Basel)*. 2019;9(8):531. DOI: <https://doi.org/10.3390/ani9080531>
68. Russell R. B., Breed J., Barton G. J. Conservation analysis and structure prediction of the SH2 family of phosphotyrosine binding domains. *FEBS Lett.* 1992;304(1):15-20. DOI: [https://doi.org/10.1016/0014-5793\(92\)80579-6](https://doi.org/10.1016/0014-5793(92)80579-6)
69. Hosoe Y., Numoto N., Inaba S., Ogawa S., Morii H., Abe R., Ito N., Oda M. Structural and functional properties of Grb2 SH2 dimer in CD28 binding. *Biophysics and Physicobiology*. 2019;16:80-88. DOI: [https://doi.org/10.2142/biophysico.16.0\\_80](https://doi.org/10.2142/biophysico.16.0_80)
70. Wang X., Cao X., Wen Y., Ma Y., Elnour I. E., Huang Y., Lan X., Chaogetu B., Hu L., Chen H. Associations of ORMDL1 gene copy number variations with growth traits in four Chinese sheep breeds. *Arch Anim Breed*. 2019;62(2):571-578. DOI: <https://doi.org/10.5194/aab-62-571-2019>
71. Siow D. L., Wattenberg B. W. Mammalian ORMDL proteins mediate the feedback response in ceramide biosynthesis. *J Biol Chem*. 2012;287(48):40198-40204. DOI: <https://doi.org/10.1074/jbc.C112.404012>
72. Cai L., Oyeniran C., Biswas D. D., Allegood J., Milstien S., Kordula T., Maceyka M., Spiegel S. ORMDL proteins regulate ceramide levels during sterile inflammation. *J Lipid Res*. 2016;57(8):1412-1422. DOI: <https://doi.org/10.1194/jlr.M065920>
73. Hjelmqvist L., Tuson M., Marfany G., Herrero E., Balcells S., González-Duarte R. ORMDL proteins are a conserved new family of endoplasmic reticulum membrane proteins. *Genome Biol*. 2002;3(6):research0027. DOI: <https://doi.org/10.1186/gb-2002-3-6-research0027>
74. Feng Z., Li X., Cheng J., Huang R., Wang D., Huang Y., Pi L., Hu L., Chen H. Copy Number Variation of the PIGY Gene in Sheep and Its Association Analysis with Growth Traits. *Animals (Basel)*. 2020;10(4):688. DOI: <https://doi.org/10.3390/ani10040688>
75. Ilkovski B., Pagnamenta A. T., O'Grady G. L., Kinoshita T., Howard M. F., Lek M. Mutations in PIGY: expanding the phenotype of inherited glycosylphosphatidylinositol deficiencies. *Hum Mol Genet*. 2015;24(21):6146-6159. DOI: <https://doi.org/10.1093/hmg/ddv331>
76. Yang Z., Cao X., Ma Y., Cheng J., Song C., Jiang R., Wang X., Huang Y., Buren C., Lan X., Ibrahim E. E., Hu L., Chen H. Novel copy number variation of the BAG4 gene is associated with growth traits in three Chinese sheep populations. *Anim Biotechnol*. 2020;1-9. DOI: <https://doi.org/10.1080/10495398.2020.1719124>
77. Deniskova T. E., Dotsev A. V., Selionova M. I., Kunz E., Medugorac I., Reyer H., Wimmers K., Barbato M., Traspov A. A., Brem G., Zinovieva N. A. Population structure and genetic diversity of 25 Russian sheep breeds based on whole-genome genotyping. *Genet Sel Evol*. 2018;50(1):29. DOI: <https://doi.org/10.1186/s12711-018-0399-5>
78. Yurchenko A. A., Deniskova T. E., Yudin N. S., Dotsev A. V., Khamiruev T. N., Selionova M. I., Egorov S. V., Reyer H., Wimmers K., Brem G., Zinovieva N. A., Larkin D. M. High-density genotyping reveals signatures of selection related to acclimation and economically important traits in 15 local sheep breeds from Russia. *BMC Genomics*. 2019;20(Suppl 3):294. DOI: <https://doi.org/10.1186/s12864-019-5537-0>

#### Сведения об авторах

**Кошкина Ольга Андреевна**, аспирант, младший научный сотрудник группы генетики и геномики мелкого рогатого скота, ФГБНУ «Федеральный научный центр животноводства – ВИЖ имени академика Л. К. Эрнста», п. Дубровицы, д. 60, Городской округ Подольск, Московская область, Российская Федерация, 142132, ORCID: <https://orcid.org/0000-0003-4830-6626>, e-mail: olechka1808@list.ru

✉ **Денисова Татьяна Евгеньевна**, кандидат биол. наук, старший научный сотрудник группы генетики и геномики мелкого рогатого скота, ФГБНУ «Федеральный научный центр животноводства – ВИЖ имени академика Л. К. Эрнста», п. Дубровицы, д. 60, Городской округ Подольск, Московская область, Российская Федерация, 142132, ORCID: <http://orcid.org/0000-0002-5809-1262>, e-mail: horarka@yandex.ru

**Зинovieva Наталия Анатольевна**, доктор биол. наук, профессор, академик Российской академии наук, директор ФГБНУ «Федеральный научный центр животноводства – ВИЖ имени академика Л. К. Эрнста», п. Дубровицы, д. 60, Городской округ Подольск, Московская область, Российская Федерация, 142132, ORCID: <http://orcid.org/0000-0003-4017-6863>, e-mail: n\_zinovieva@mail.ru

#### Information about the authors

**Olga A. Koshkina**, graduate student, junior researcher of the Group for Genetics and Genomics of Small Ruminants of L. K. Ernst Federal Science Center for Animal Husbandry, Dubrovitsy village, 60, Podolsk City District, Moscow Region, Russian Federation, 142132, ORCID: <https://orcid.org/0000-0003-4830-6626>, e-mail: olechka1808@list.ru

✉ **Tatiana E. Deniskova**, PhD in Biology, senior researcher of the Group for Genetics and Genomics of Small Ruminants of L. K. Ernst Federal Science Center for Animal Husbandry, Dubrovitsy village, 60, Podolsk City District, Moscow Region, Russian Federation, 142132, ORCID: <http://orcid.org/0000-0002-5809-1262>, e-mail: horarka@yandex.ru

**Natalia A. Zinovieva**, DSc in Biology, Professor, Academician of the Russian Academy of Sciences, Director of L. K. Ernst Federal Science Center for Animal Husbandry, Dubrovitsy village, 60, Podolsk City District, Moscow Region, Russian Federation, 142132, ORCID: <http://orcid.org/0000-0003-4017-6863>, e-mail: n\_zinovieva@mail.ru

✉ – Для контактов / Corresponding author