

ISSN 1817-7204(Print)

ISSN 1817-7239(Online)

УДК 636.4.082.25:577.21

<https://doi.org/10.29235/1817-7204-2021-59-4-464-476>

Поступила в редакцию 24.08.2021

Received 24.08.2021

**В. Н. Кипень, М. Е. Михайлова, Е. В. Снытков, Е. Л. Романишко,
Е. В. Иванова, Р. И. Шейко**

Институт генетики и цитологии Национальной академии наук Беларуси, Минск, Беларусь

БИОИНФОРМАТИЧЕСКИЙ АНАЛИЗ ГЕНОМОВ КОММЕРЧЕСКИХ ПОРОД ДОМАШНИХ СВИНЕЙ ДЛЯ ИДЕНТИФИКАЦИИ ПОРОДОСПЕЦИФИЧНЫХ SNP

Аннотация: Определение чистопородности сельскохозяйственных животных в селекционной системе имеет ключевое значение для всей отрасли животноводства. Чистопородное разведение заводских пород призвано обеспечить производство высокоценного улучшающего племенного материала для товарного животноводства. Определение чистопородности свиней может быть проведено с использованием однонуклеотидных полиморфизмов (SNP). Технология мультиплексирования сегодня достигла уровня, который позволяет за один запуск прибора охарактеризовать десятки и сотни тысяч полиморфных вариантов одновременно для сотен животных. Впервые с использованием методов биоинформатики проведен анализ полногеномных проектов для 264 особей вида *Sus scrofa*, расположенных в базе Sequence Read Archive (NCBI-SRA). Определен *in silico* генотип для 692 SNP, из которых для 59 SNP показан значительный потенциал для дифференциации четырех коммерческих пород: крупная белая (наиболее значимые SNP – Chr.6:g.85845403T>G и Chr.16:g.74053569T>C), дюрок (Chr.4:g.55661608A>G, Chr.14:g.107689091T>C и Chr.14:g.107939105T>C), ландрас (Chr.5:g.99925204A>G, Chr.18:g.40100481A>G и Chr.18:g.7664624A>G) и пьетрен (Chr.13:g.136017764T>C и Chr.17:g.47595840A>G). Для пород свиней дюрок и пьетрен точность дифференциации была не менее 99 %, для пород свиней крупная белая и ландрас – более 80 %, однако показатель чувствительности, характеризующий процент ложноположительных результатов классификации, был немногим более 65 %. Создание моделей для молекулярно-генетических исследований данных пород позволит проводить генетическую экспертизу их чистопородности, что будет способствовать возрастанию их племенной ценности и сохранению национального генофонда. **Благодарности.** Исследование выполнено в рамках ГПНИ «Биотехнологии-2» (2021–2025 гг.), подпрограмма «Геномика, эпигеномика, биоинформатика».

Ключевые слова: *Sus scrofa scrofa*, дифференциация, однонуклеотидный полиморфизм, породоспецифичность, дюрок, ландрас, пьетрен, крупная белая, Multifactor Dimensionality Reduction, ROC-анализ

Для цитирования: Биоинформатический анализ геномов коммерческих пород домашних свиней для идентификации породоспецифичных SNP / В. Н. Кипень, М. Е. Михайлова, Е. В. Снытков, Е. Л. Романишко, Е. В. Иванова, Р. И. Шейко // Вес. Нац. акад. наук Беларуси. Сер. аграр. наук. – 2021. – Т. 59, № 4. – С. 464–476. <https://doi.org/10.29235/1817-7204-2021-59-4-464-476>

**Viachaslau N. Kipen, Mariya E. Mikhailova, Evgenij V. Snytkov, Elena L. Romanishko,
Ekaterina V. Ivanova, Ruslan I. Sheyko**

Institute of Genetics and Cytology of the National Academy of Sciences of Belarus, Minsk, Belarus

BIOINFORMATIC ANALYSIS OF GENOMES OF COMMERCIAL BREEDS OF DOMESTIC PIGS FOR IDENTIFICATION OF BREED-SPECIFIC SNPS

Abstract: Determining the purebredness of farm animals in a breeding system is of key importance for the entire livestock industry. Purebred breeding of pig breeds is designed to ensure the production of high-value improving breeding material for commercial livestock breeding. Determination of purebredness of pigs can be carried out using single nucleotide polymorphisms (SNP). The multiplexing technology today has reached a level that makes it possible to characterize tens and hundreds of thousands of polymorphic variants simultaneously for hundreds of animals in one run of the device. For the first time, using bioinformatics methods, an analysis of genome-wide projects was carried out for 264 individuals of the species *Sus scrofa* located in the Sequence Read Archive (NCBI-SRA). The *in silico* genotype was determined for 692 SNPs, of which 59 SNPs showed a significant potential for differentiation of four commercial breeds: large white (the most significant SNPs are Chr. 6: g.85845403T> G and Chr.16: g.74053569T> C), duroc (Chr. 4: g.55661608A> G, Chr. 14: g.107689091T> C and Chr. 14: g.107939105T> C), landrace (Chr. 5: g.99925204A> G, Chr. 18: g.40100481A> G and Chr. 18: g.7664624A> G) and pietrain (Chr. 13: g.136017764T> C and Chr.17: g.47595840A> G). For breeds of duroc and pietrain pigs, the accuracy of differentiation was at least 99%, for breeds of large white and landrace pigs - over 80%, however, the sensitivity indicator characterizing the percentage of false positive results of classification was slightly over 65%. Creation of models for molecular-and-genetic studies of these breeds will allow for a genetic examination of their purebredness, which will contribute to an

increase in their breeding value and preservation of the national gene pool. **Acknowledgments.** The research was carried out within the framework of the State Scientific Research Program “Biotechnology-2” (2021-2025), subprogram “Genomics, Epigenomics, Bioinformatics”.

Keywords: *Sus scrofa scrofa*, differentiation, single nucleotide polymorphism, breed-specificity, Duroc, Landrace, Pietrain, Large white, Multifactor Dimensionality Reduction, ROC-analysis

For citation: Kipen V.N., Mikhailova M.E., Snytkov E.V., Romanishko E.L., Ivanova E.V., Sheyko R.I. Bioinformatic analysis of genomes of commercial breeds of domestic pigs for identification of breed-specific SNPs. *Vestsi Natsyyanal'nay akademii navuk Belarusi. Seryya agrarnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Agrarian series*, 2021, vol. 59, no 4, pp. 464-476 (in Russian). <https://doi.org/10.29235/1817-7204-2021-59-4-464-476>

Введение. Определение чистопородности сельскохозяйственных животных в селекционной системе имеет ключевое значение для всей отрасли животноводства. Ввиду того факта, что количественные признаки, как правило, имеют полигенный характер, т.е. предетерминированы совокупностью генов и их функциональным состоянием, определение значимых корреляций между SNP (*single nucleotide polymorphism*) и фенотипическим признаком(-ми) на высоком уровне статистической значимости осуществимо только при полногеномных исследованиях – GWAS (*genome-wide association studies*).

Количественные признаки у свиней подразделяют на 4 группы: признаки, которые имеют низкий коэффициент наследуемости (0,01–0,20) – в основном они характеризуют воспроизводительную способность (хряка оценивают по оплодотворяющей способности, по живой массе потомков в возрасте 4 мес, а воспроизводительные качества свиноматок – по многоплодию, массе гнезда, количеству поросят к отъему, получаемых от матки за год); признаки, которые имеют коэффициент наследуемости от 0,2–0,4 и характеризуют откормочные качества (среднесуточный прирост, возраст достижения живой массы, затрата корма на 1 кг прироста); признаки, которые имеют сравнительно высокий коэффициент наследуемости 0,5–0,7 – это мясные качества свиней (убойный выход, длина туши, толщина шпика, площадь «мышечного глазка», содержание мяса в туше); а также относительно новая группа признаков, характеризующих мясные качества, коэффициент наследуемости которых >0,7 [1].

Определение чистопородности свиней может быть проведено с использованием маркеров двух типов – SNP (*single nucleotide polymorphism*) или STR (*short tandem repeat*). Каждый подход обладает своими преимуществами и недостатками. Одними из главных преимуществ анализа с использованием STR-локусов является возможность их мультиплексирования (10–15 локусов в одной пробирке), а также относительно низкая стоимость анализа при массовом генотипировании. Также для характеристики определенного фенотипического признака необходимо, чтобы ген, ответственный за этот признак, находился в одной группе сцепления с STR-локусом. Как правило, STR-локусы в данном контексте менее информативны, чем SNP. Технология мультиплексирования SNP сегодня достигла уровня, который позволяет за один запуск прибора (ридер чипов или секвенатор для массового параллельного секвенирования) охарактеризовать десятки и сотни тысяч SNP одновременно для сотен животных.

Так, компанией Illumina[®] разработан чип для полногеномного анализа SNP для животных вида *Sus scrofa* – PorcineSNP60 BeadChip¹, включающий в себя более 60 тыс. полиморфных вариантов. Ramos A. M. et al. (2011) применяли данную технологию для оценки дифференцирующего потенциала ряда SNP для определения чистопородности свиней пород дюрок (Duroc), ландрас (Landrace), крупная белая (Large white) и пьетрен (Pietrain) [2]. За последние 10 лет в научных журналах было опубликовано более 50 работ, также посвященных поиску видо- и породоспецифичных SNP для вида *Sus scrofa*. В основном исследователи прибегали к использованию чиповых технологий производства Illumina[®]. Также SNP были исследованы в ряде работ для оценки QTL или чистопородности для следующих пород свиней: крупная белая [2–11], ландрас [2, 5, 10–13], дюрок [2, 10–12, 14], пьетрен [2, 10–12, 15]. Оценка генетического разнообразия пород свиней в контексте решения ряда задач по криминалистике обсуждалась в исследованиях K. S. Roberts et al. [16], P. Stratz et al. [15], S. Wilkinsoet al. [10], В. Н. Кипеня [17] и др.

¹ PorcineSNP60 DNA Analysis Kit v2 [Электронный ресурс]. Режим доступа: <https://www.illumina.com/products/by-type/microarray-kits/porcine-snp60.html>. Дата доступа: 20.08.2021.

Как правило, исследователи, применявшие PorcineSNP60 BeadChip, либо не приводят информацию о дифференцирующем потенциале конкретного SNP в контексте различия пород свиней, либо предпочитают кодировать их спектр. Однако в настоящее время имеется довольно много первичных данных, находящихся в открытом доступе в Национальном центре биотехнологической информации (NCBI, National Center for Biotechnology Information), относительно полногеномных сиквенсных проектов для вида *Sus scrofa*. По состоянию на 2019 г. в открытом доступе зарегистрирована информация о более чем 200 полностью секвенированных геномах свиней пород дюрок, ландрас, крупная белая, пьетрен и др., характеризующихся достаточной глубиной прочтения и покрытия для определения генотипа с использованием методов биоинформатики.

С учетом того факта, что GWAS, как правило, проводятся с помощью микрочиповых технологий, то выявленные генетические ассоциации на уровне генома должны в дальнейших субпопуляционных исследованиях проверяться дополнительно. Результаты GWAS в первую очередь направлены на то, чтобы дать исследователю область поиска в геноме для выявления локусов (генов), обладающих высоким дифференцирующим потенциалом для решения поставленной задачи. В этой связи приоритетной задачей для исследователей биоинформатического профиля становится определение SNP, расположенных вблизи генетических маркеров, дифференцирующий потенциал которых был определен ранее. Таким образом, использование информации о хромосомной позиции SNP, включенных в PorcineSNP60 BeadChip, в совокупности с информацией о полном геноме свиньи той или иной породы позволит без существенных материальных затрат (т.е. без приобретения дорогостоящих наборов для секвенирования и приборной базы – сканеров биочипов iScan и HiScan Illumina) дополнительно определить новые SNP, потенциально информативные для решения задачи дифференциации пород свиней.

Главной задачей селекции сельскохозяйственных животных является поддержание и совершенствование их породных качеств путем подбора пар для скрещивания, умелое ведение линий в пределах породы и межлинейных кроссов наиболее удачно сочетающихся линий. В результате многообразие животных в пределах породы и ее сложная генетическая структура дают возможность добиться значительных сдвигов в ту или другую сторону и совершенствовать породу в заданном направлении. Кроме того, чистопородное разведение заводских пород призвано обеспечить производство высокоценного улучшающего племенного материала для товарного животноводства. Многочисленными исследованиями показано, что SNP являются эффективным и высокочувствительным инструментом для выявления генетической гетерогенности пород и популяций. Благодаря этой особенности они находят практическое применение не только в качестве инструмента контроля происхождения животных, но и для выявления степени генетических различий между породами, типами, стадами и генеалогическими группами животных. В настоящее время анализ SNP является практически полностью автоматизированным и высокотехнологичным процессом, позволяющим проводить исследования большого количества образцов [18]. В то же время анализ относительно небольшого количества SNP (до 100–150) может быть выполнен с использованием KASP-технологии (Kompetitive allele specific PCR), в этом случае при генотипировании даже 200 особей себестоимость анализа составит менее 0,4 евро за образец.

Цель настоящей работы – анализ целевых нуклеотидных последовательностей в геноме *Sus scrofa* на основании имеющихся в открытом доступе необработанных данных (*raw data*) полногеномных проектов и определение SNP с потенциально высоким дифференцирующим потенциалом для различения коммерческих пород домашних свиней (крупная белая, дюрок, ландрас и пьетрен) с использованием биоинформатических методов.

Материалы и методы исследования. Определение генотипа *in silico*. Генотипы *in silico* для животных, геномы которых были представлены в открытом доступе в NCBI, определяли с использованием online алгоритма SRA-Blast (файлы в формате *.sam) и программы Unipro UGENE v.1.31.1 [19]. Генотип для пород дюрок, крупная белая и ландрас был определен по 692 SNP, из них 193 – с потенциально высоким дифференцирующим потенциалом [2], 499 – локусы, ассоциированные с QTL по таким характеристикам, как экстерьер и продуктивность [20–34]. В био-

информатическом анализе были задействованы 194 особи вида *Sus scrofa* (дюрок – 79, ландрас – 39, крупная белая – 68, пьетрен – 8), «сырые» данные полных геномов которых расположены в базе Sequence Read Archive² (SRA,) – проекты PRJNA41185 (год регистрации проекта в BioProject – 2009, University of Aarhus, Дания), PRJNA176478 (2012, Uppsala University, Швеция), PRJNA186497 (2013, Novogene, США), PRJEB1683 (2013, Wageningen University & Research, Нидерланды), PRJNA239399 (2014, Agricultural Biotechnology Center, Венгрия), PRJNA260763 (2014, Seoul National University и NICEM – Южная Корея, BGI – КНР), PRJEB9922 (2015, Wageningen University & Research, Нидерланды), PRJNA309108 (2016, Novogene, США), PRJNA322309 (2016, China Agricultural University, КНР), PRJNA343658 (2016, USDA-ARS-USMARC, США), PRJNA369600 (2017, Universitat Autònoma de Barcelona, Испания), PRJNA378496 (2017, China Agricultural University, КНР), PRJNA393920 (2017, INRA, Франция), PRJNA487172 (2018, Huazhong Agricultural University, КНР), PRJNA506339 (2018, University of Edinburgh, Шотландия), PRJNA507853 (2018, Sichuan Agricultural University, Китай).

Процесс глобального выравнивания в среде SRA-Blast и сохранение SAM файлов были автоматизированы с использованием скрипта на языке программирования Python³.

Статистический анализ данных. Дифференцирующий потенциал SNP для определения чистопородности определяли с использованием ROC-анализа в SPSS v.20.0. При наличии нижней границы асимптотического 95%-ного доверительного интервала более 0,5 для параметра AUC (площадь под кривой) SNP позиционировался как генетический маркер со значительным дифференцирующим потенциалом.

Оценку дифференцирующего потенциала для совокупности SNP проводили с использованием программы MDR v.3.0.2⁴. В процессе моделирования нами были использованы высоко консервативные настройки поиска конфигурации модели, которые позволили однозначно дифференцировать наличие/отсутствие статистически значимых эффектов: количество атрибутов (*attribute count range*) – от 1 до n (где n – количество переменных в модели); воспроизводимость модели (*cross-validation count*) – 10; анализ топ-моделей (*track top models*) – 1000; поиск конфигурации модели (*search method configuration*) – exhaustive; классификация ячеек (*ambiguous cell assignment*) – unclassified. Вклад конкретного генотипа определялся величиной энтропии H (%). При $H = 100$ % полиморфизм способен дифференцировать к какой группе (породе) относится неизвестный образец. В программе MDR в результате множественных перестановок (пермутаций) первичных данных определяется наиболее оптимальная модель дифференциации. Корректность модели оценивали по значению сбалансированной точности (Balanced Accuracy), которая зависит от чувствительности и специфичности модели [35].

Результаты и их обсуждение. Проведенный биоинформатический анализ, направленный на определение генотипа по 692 SNP для 194 животных вида *Sus scrofa*, позволил рассчитать частоты встречаемости аллелей и генотипов. Полученные результаты легли в основу математического анализа классификаций с применением ROC-кривых (*receiver operating characteristic*), также известных как кривые ошибок. Количественную интерпретацию ROC дает показатель AUC (*area under ROC curve*) – площадь, ограниченная ROC-кривой и осью доли ложных положительных классификаций. Чем выше показатель AUC, тем качественнее осуществляется классификация, при этом значение 0,5 демонстрирует непригодность выбранного метода классификации.

Для свиней породы крупная белая нами выявлены два строгоспецифичных SNP с высоким дифференцирующим потенциалом, т.е. породоспецифичный аллель встречался только у данной породы и не был выявлен у свиней других породы (табл. 1). Для данных SNP нижний предел дифференциации (нижняя граница 95 % ДИ) составил 0,564, частота породоспецифичного аллеля – 22,2 % (g.85845403T>G, Chr.6) и 23,7 % (g.34974276A>G, Chr.9). Дополнительно нами были определены SNP, для которых частота породоспецифичного аллеля составила менее 15 % – g.90473255A>G (Chr.3), g.48649643A>G (Chr.5), g.97244415T>C (Chr.5), g.74053569T>C (Chr.16)

² SRA database [Electronic resource]. Mode of access: <https://www.ncbi.nlm.nih.gov/sra/> Date of access: 20.08.2021.

³ Python™ [Electronic resource]. Mode of access: <https://www.python.org/> Date of access: 20.08.2021.

⁴ Multifactor Dimensionality Reduction [Electronic resource]. Mode of access: <https://sourceforge.net/projects/mdr/> Date of access: 20.08.2021.

и g.40047092T>C (Chr.X) – дифференцирующий потенциал данных SNP невысокий, однако дополнительный биоинформатический анализ прилегающих к ним геномных регионов позволит выявить новые SNP, не представленные в научной литературе.

Т а б л и ц а 1. SNP с наибольшим потенциалом для дифференциации свиней породы крупная белая на основании ROC-анализа

Table 1. SNP with the greatest potential for differentiation of large white pigs based on ROC analysis

SNP*	Площадь (AUC)	Стд. ошибка (sd)	Асимптотический 95%-ный доверительный интервал	
			нижняя граница	верхняя граница
Chr.6: 85845403T>G	0,750	0,095	0,564	0,936
Chr.9: 34974276A>G	0,750	0,095	0,564	0,936

* Версия сборки генома Sscrofa10.2 (GCF_000003025.5).

Для свиней породы дюрок нами выявлены 38 строгоспецифичных SNP с высоким дифференцирующим потенциалом (табл. 2). Для следующих 10 SNP – g.130814016T>G (Chr.1), g.249596526T>G (Chr.1), g.78754003A>C (Chr.5), g.59161671T>C (Chr.13), g.107689091T>C (Chr.14), g.107939105T>C (Chr.14), g.12577892A>G (Chr.14), g.84388841T>G (Chr.15) и g.86214333T>G (Chr.15) – нижний предел дифференциации (нижняя граница 95 % ДИ) был максимальным и находился в диапазоне 0,842–1,000. При этом частота породоспецифичного аллеля для данных SNP находилась в диапазоне 68,3–96,4 %. Стоит отметить, что на хромосоме 14 выявлен фрагмент длиной около 250 тыс. п.н., на котором определены несколько SNP с высоким дифференцирующим потенциалом для породы свиней дюрок. На данном участке хромосомы или непосредственно рядом с ним расположены гены: *AICF* (APOBEC1 complementation factor, NCBI Gene ID – 100155074) – гомологичный ген у человека выполняет роль редактирования РНК или процессинга РНК; *PRKG1* (protein kinase cGMP-dependent 1, NCBI Gene ID – 733640) – действует как ключевой медиатор в пути передачи сигналов оксида азота; *ASAH2* (N-acylsphingosine amidohydrolase 2, NCBI Gene ID – 100157065) и *SGMS1* (sphingomyelin synthase 1, NCBI Gene ID – 100037303) – участвуют в сфинголипидном метаболизме; и LOC110256620 (NCBI Gene ID – 110256620), функциональная роль которого неизвестна. Наличие ряда SNP, которые расположены в одной группе сцепления (в случае с SNP g.107689091T>C и g.107939105T>C на Chr.14 вероятность кроссинговера составляет не более 0,25 %), дает исследователю возможность в перспективе выявить компактный кластер (менее 200 н.п.) с несколькими SNP с высоким потенциалом и разработать модель дифференциации свиней породы дюрок по упрощенной схеме, например, с использованием технологии HRM (high resolution melting).

Для свиней породы ландрас нами выявлены три строгоспецифичных SNP с высоким дифференцирующим потенциалом (табл. 3). Для данных SNP нижний предел дифференциации (нижняя граница 95 % ДИ) был невысоким относительно установленного при статистическом анализе – 0,511–0,610, при этом частота породоспецифичного аллеля составила 19,6 % (g.67905693T>C, Chr.3), 23,9% (g.99925204A>G, Chr.5) и 20,5 % (g.40100481A>G, Chr.18). Дополнительно нами выявлены SNP, для которых частота породоспецифичного аллеля составила менее 15 % – g.48069819A>G (Chr.4), g.55120557T>C (Chr.9), g.7664624A>G (Chr.18) и g.8496831T>C (Chr.18), дифференцирующий потенциал данных SNP также, как и для свиней породы крупная белая, оказался невысоким.

Однако биоинформатический анализ, который мы планируем провести, позволит выявить новые SNP со значимым дифференцирующим потенциалом. Стоит отметить, что на хромосоме 18 выявлен фрагмент длиной около 800 тыс. п.н., на котором определены несколько SNP с высоким дифференцирующим потенциалом для породы свиней ландрас. На данном участке хромосомы или непосредственно рядом с ним расположено довольно значительное количество генов, например: *EPHB6* (EPH receptor B6, NCBI Gene ID – 100511727) – участвует в формировании аксонов (нейронной сети); *TRPV5* (transient receptor potential cation channel subfamily V member 5, NCBI Gene ID – 100622162) – задействован в процессах регулируемой эндокринной системой реаб-

сорбции кальция; *SSBPI* (single stranded DNA binding protein 1, NCBI Gene ID – 100512994) – принимает участие в механизмах репликации ДНК; *MGAM* (maltase-glucoamylase, NCBI Gene ID – 102160115) – задействован в процессах пищеварения.

Таблица 2. SNP с наибольшим потенциалом для дифференциации свиней породы дюрок на основании ROC-анализа

Table 2. SNP with the greatest potential for differentiation of Duroc pigs based on ROC analysis

SNP*	Площадь (AUC)	Стд. ошибка (sd)	Асимптотический 95%-ный доверительный интервал	
			нижняя граница	верхняя граница
Chr.1: 130814016T>G	0,932	0,044	0,847	1,000
Chr.1: 132217476A>G	0,773	0,070	0,636	0,909
Chr.1: 204595633A>G	0,926	0,044	0,840	1,000
Chr.1: 249596526T>G	0,998	0,003	0,993	1,000
Chr.1: 304963553A>G	0,886	0,054	0,780	0,993
Chr.2: 146082066A>G	0,737	0,071	0,597	0,877
Chr.2: 153706211A>G	0,773	0,070	0,636	0,909
Chr.3: 28483632G>A**	0,736	0,071	0,596	0,876
Chr.3: 61093246T>C	0,886	0,054	0,780	0,993
Chr.4: 23873948G>A**	0,720	0,073	0,577	0,863
Chr.4: 55661608A>G	0,909	0,049	0,812	1,000
Chr.5: 48861940A>C	0,805	0,065	0,677	0,933
Chr.5: 78754003A>C	0,977	0,026	0,926	1,000
Chr.7: 112783948T>C	0,864	0,058	0,749	0,978
Chr.8: 33753357G>A**	0,750	0,071	0,610	0,890
Chr.8: 109044778T>C	0,886	0,054	0,780	0,993
Chr.10: 56611984A>G	0,909	0,049	0,812	1,000
Chr.11: 20538983A>C	0,750	0,071	0,610	0,890
Chr.11: 20679297A>G	0,773	0,070	0,636	0,909
Chr.11: 24063007A>C	0,853	0,059	0,737	0,968
Chr.12: 44570609A>G	0,682	0,075	0,535	0,829
Chr.12: 47729442A>G	0,886	0,054	0,780	0,993
Chr.13: 191444960A>T	0,841	0,062	0,720	0,962
Chr.13: 51703949A>G	0,750	0,071	0,610	0,890
Chr.13: 59161671T>C	0,928	0,044	0,842	1,000
Chr.14: 102343810A>G	1,000	0,000	1,000	1,000
Chr.14: 107689091T>C	0,977	0,026	0,926	1,000
Chr.14: 107939105T>C	0,977	0,026	0,926	1,000
Chr.14: 12577892A>G	0,972	0,027	0,920	1,000
Chr.14: 79763861A>C	0,795	0,067	0,663	0,928
Chr.15: 55290487A>G	0,886	0,054	0,780	0,993
Chr.15: 60719311T>C	0,905	0,050	0,807	1,000
Chr.15: 84388841T>G	0,944	0,037	0,872	1,000
Chr.15: 86214333T>G	0,950	0,037	0,878	1,000
Chr.17: 11901977T>C	0,727	0,073	0,585	0,870
Chr.18: 27274159T>G	0,773	0,070	0,636	0,909
Chr.18: 34861012T>C	0,705	0,074	0,560	0,850
Chr.18: 3614625A>G	0,909	0,049	0,812	1,000

*Версия сборки генома Sscrofa10.2 (GCF_000003025.5); **версия сборки генома Sscrofa11.1 (GCF_000003025.6).

Таблица 3. SNP с наибольшим потенциалом для дифференциации свиней породы ландрас на основании ROC-анализа

Table 3. SNP with the greatest potential for differentiation of Landrace pigs based on ROC analysis

SNP*	Площадь (AUC)	Стд. ошибка (sd)	Асимптотический 95 %-ный доверительный интервал	
			нижняя граница	верхняя граница
Chr.3: 67905693T>C	0,682	0,075	0,535	0,829
Chr.5: 99925204A>G	0,750	0,071	0,610	0,890
Chr.18: 40100481A>G	0,659	0,075	0,511	0,807

*Версия сборки генома Sscrofa10.2 (GCF_000003025.5).

Для свиней породы пьетрен нами выявлены 4 строгоспецифичных SNP с высоким дифференцирующим потенциалом (табл. 4). Для данных SNP нижний предел дифференциации (нижняя граница 95 % ДИ) находился в диапазоне 0,601–1,0, частота породоспецифичного аллеля составила 66,7 % (g.142179174A>G, Chr.6), 50,0 % (g.136017764T>C, Chr.13), 33,3 % (g.48426806T>C, Chr.17) и 58,3 % (g.47595840A>G, Chr.17). Дополнительно нами были определены SNP – g.149172524A>G (Chr.6), g.49934017A>G (Chr.12) и g.14724810A>G (Chr.X), дифференцирующий потенциал которых оказался относительно невысоким, несмотря на достаточно высокие значения распространенности в популяции пороспецифичных аллелей (16,7–50,0 %). Как нам представляется, данная проблема связана с небольшим количеством полногеномных прочтений для свиней породы пьетрен. Также на хромосоме 17 выявлен фрагмент длиной около 830 тыс. п.н., на котором определены несколько SNP с высоким дифференцирующим потенциалом для данной породы свиней. На этом участке хромосомы или непосредственно рядом с ним расположены гены: *DHX35* (DEAH-box helicase 35, NCBI Gene ID – 100154906) – вовлечен в процессы сплайсинга; и два локуса с длинными некодирующими РНК – LOC110257441 (NCBI Gene ID – 110257441) и LOC110257425 (NCBI Gene ID – 110257425), функция которых неизвестна.

Таблица 4. SNP с наибольшим потенциалом для дифференциации свиней породы пьетрен на основании ROC-анализа

Table 4. SNP with the highest potential for differentiation of Pietrain pigs based on ROC analysis

SNP*	Площадь (AUC)	Стд. ошибка (sd)	Асимптотический 95%-ный доверительный интервал	
			нижняя граница	верхняя граница
Chr.6: 142179174A>G	0,917	0,090	0,739	1,000
Chr.13: 136017764T>C	1,000	0,000	1,000	1,000
Chr.17: 48426806T>C	0,833	0,119	0,601	1,000
Chr.17: 47595840A>G	0,917	0,090	0,739	1,000

*Версия сборки генома Sscrofa10.2 (GCF_000003025.5).

Таким образом, нами определены 59 SNP со средним или высоким потенциалом (по данным ROC-анализа) для дифференциации свиней пород крупная белая (7 SNP), дюрок (38 SNP), ландрас (7 SNP) и пьетрен (7 SNP). Наибольшее количество информативных SNP выявлено для свиней породы дюрок. Данная порода широко используется в промышленном свиноводстве при создании новых линий в трех- и четырехпородных типах скрещивания. Такой племенной молодняк характеризуется высоким (до 65 %) содержанием мяса отличного качества в тушах [36]. В настоящее время порода дюрок широко распространена в европейских странах, однако в Республике Беларусь данная порода используется слабо – доли процента от общей численности поголовья *Sus scrofa domesticus*.

По нашему мнению, для создания модели определения чистопородности свиней с использованием молекулярно-генетического анализа должно быть включено небольшое количество генетических маркеров – оптимальное количество SNP для решения данной задачи не более 5–7. Поэтому для достижения данной цели необходимо с использованием статистических методов

редуцировать общее количество выявленных нами SNP до логически и экономически приемлемого количества, оставив в моделях наиболее значимые комбинации SNP с наивысшим дифференцирующим потенциалом.

Построение модели взаимодействий SNP (определение минимального и достаточного количества генетических маркеров для решения поставленной задачи) проводили с использованием биоинформатического метода многомерного сокращения размерности (MDR v.3.0.2). В результате проведенного моделирования были определены модели, включающие такие сочетания SNP для свиней пород крупная белая, дюрок, ландрас и пьетрен, которые позволили наилучшим образом отличить данные породы между собой (табл. 5).

Таблица 5. Модели для дифференциации пород домашних свиней с использованием SNP (на основании MDR-анализа)

Table 5. Models for differentiating breeds of domestic pigs using SNP (based on MDR analysis)

Порода	SNP* в модели	Сбалансированная точность, %	Чувствительность, %	Специфичность, %	Воспроизводимость модели
Крупная белая	Chr.6: 85845403T>G Chr.16: 74053569T>C	83,2	66,7	100	100/100
Дюрок	Chr.4: 55661608A>G Chr.14: 107689091T>C Chr.14: 107939105T>C	99,1	100	100	100/100
Ландрас	Chr.5: 99925204A>G Chr.18: 40100481A>G Chr.18: 7664624A>G	81,7	66,7	100	95/100
Пьетрен	Chr.13: 136017764T>C Chr.17: 47595840A>G	100	100	100	100/100

*Версия сборки генома Sscrofa10.2 (GCF_000003025.5).

Вклад конкретного SNP в дифференцирующий потенциал модели характеризует такой параметр, как энтропия. В математической статистике энтропия – мера неопределенности распределения вероятностей: чем выше значение энтропии, тем больше вклад SNP в дифференцирующий потенциал модели. Наибольшим совокупным дифференцирующим потенциалом из числа выявленных для определения чистопородности свиней породы обладают следующие SNP:

1) для свиней породы крупная белая – g.85845403T>G (Chr.6) и g.74053569T>C (Chr.16) (SNP g.85845403T>G имеет величину энтропии, равную 20,20 %, SNP g.74053569T>C имеет величину энтропии, равную 12,06 %, совокупный вклад двух SNP составил 36,04 %) (рис. 1, a);

2) для свиней породы дюрок – g.55661608A>G (Chr.4), g.107689091T>C (Chr.14) и g.107939105T>C (Chr.14) (SNP g.55661608A>G имеет величину энтропии, равную 52,17 %, SNP g.107689091T>C имеет величину энтропии равную 73,21 %, SNP g.107939105T>C имеет величину энтропии, равную 73,21 %, совокупный вклад трех SNP составил 80,08 %) (рис. 1, b);

3) для свиней породы ландрас – g.99925204A>G (Chr.5), g.40100481A>G (Chr.18) и g.7664624A>G (Chr.18) (SNP g.99925204A>G имеет величину энтропии, равную 25,53 %, SNP g.40100481A>G имеет величину энтропии, равную 15,36 %, SNP g.7664624A>G имеет величину энтропии, равную 10,69 %, совокупный вклад трех SNP составил 51,33 %) (рис. 1, c);

4) для свиней породы пьетрен – g.136017764T>C (Chr.13) и g.47595840A>G (Chr.17) (SNP g.136017764T>C имеет величину энтропии, равную 29,56 %, SNP g.47595840A>G имеет величину энтропии, равную 22,41 %, совокупный вклад двух SNP составил 29,56 %) (рис. 1, d).

Заключение. В ходе проведенного биоинформатического исследования были выявлены кандидатные полиморфные локусы, для которых показан средний или высокий потенциал дифференциации. Для определения чистопородности свиней породы крупная белая были отобраны SNP с наибольшим совокупным дифференцирующим потенциалом: g.85845403T>G (Chr.6) и g.74053569T>C (Chr.16); для породы дюрок – g.55661608A>G (Chr.4), g.107689091T>C (Chr.14) и g.107939105T>C (Chr.14); для породы ландрас – g.99925204A>G (Chr.5), g.40100481A>G (Chr.18) и g.7664624A>G (Chr.18); для породы пьетрен – g.136017764T>C (Chr.13) и g.47595840A>G (Chr.17).

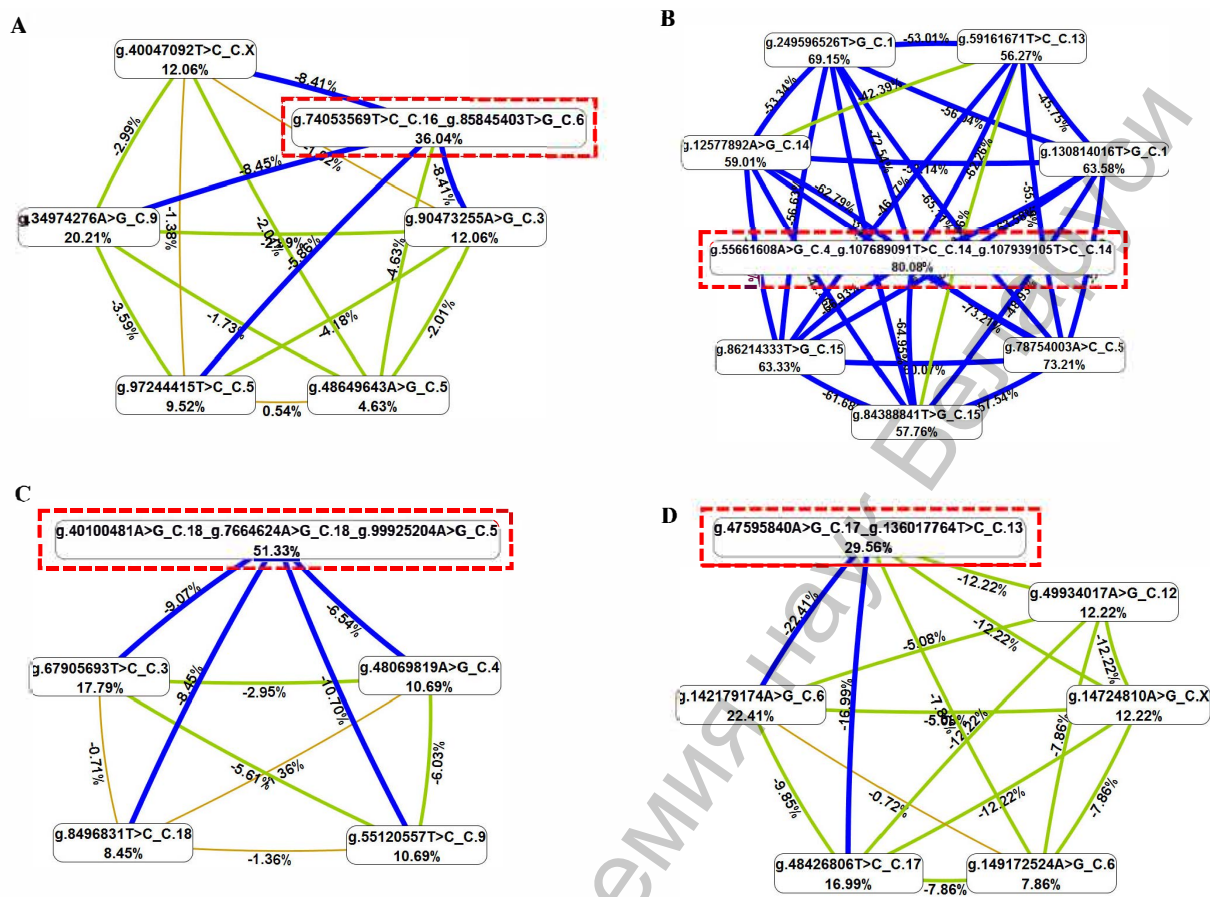


Рис. 1. Графическое представление взаимодействий для SNP с высоким дифференцирующим потенциалом, адаптировано из ПО MDR v.3.0.2 (абсолютные значения (%) – показатель энтропии, характеризующий вклад конкретного SNP в дифференцирующий потенциал модели): А – для свиней породы крупная белая; В – для свиней породы дюрок; С – для свиней породы ландрас; D – для свиней породы пьетрен

Fig. 1. A graphical presentation of interactions for SNPs with high differentiating potential, adapted from MDR v.3.0.2 software (absolute values (%) - entropy index characterizing the contribution of a specific SNP to the differentiating potential of the model); A - for Large White pigs; B - for Duroc pigs; C - for Landrace pigs; D - for Pietrain pigs

Для пород свиней дюрок и пьетрен точность дифференциации была не менее 99 %, для пород свиней крупная белая и ландрас – более 80 %, однако показатель чувствительности, характеризующий процент ложноположительных результатов классификации был немногим более 65 %. Поэтому приоритетной задачей в ближайшей перспективе нами видится в более углубленном биоинформатическом анализе хромосомных локусов Chr.3: g.60000000-g.70000000, Chr.5: g.95000000-g.105000000 и Chr.18: g.35000000-g.45000000 для свиней породы ландрас, а также Chr.6: g.80000000-g.90000000 и Chr.9: g.30000000-g.40000000 для свиней породы крупная белая. В целом дополнительный анализ выявил более 9 тыс. потенциальных SNP, дифференцирующих потенциал которых будет оценен в последующих работах.

Подобный биоинформатический анализ проведен в Республике Беларусь впервые. Полученные результаты позволят определить значимые SNP и разработать тест-модели, с использованием которых появится возможность определить чистопородность свиней и, как следствие, поддерживать на оптимальном уровне такие параметры популяции (пород, линии), как уровень гетерозиготности, инбредности, генетической дифференциации и др. В итоге появится возможность контролировать генетические особенности пород и определять соответствие отдельных особей генофонду породы, осуществлять подбор родительских пар с оптимальной степенью гетерогенности, обуславливающей повышение продуктивности животных и определять тактику и стратегию разведения и дальнейшего совершенствования пород.

Благодарности. Исследование выполнено в рамках ГПНИ «Биотехнологии-2» (2021–2025 гг.), подпрограмма «Геномика, эпигеномика, биоинформатика», НИР «Разработка системы генетического анализа для определения чистопородности свиней на основе изучения SNP-локусов».

Список использованных источников

1. Корневская, П. А. Продуктивность и биологические особенности свиней французской селекции и их помесей : автореф. дис. ... канд. биол. наук : 06.02.10 / П. А. Корневская ; Рос. гос. аграр. ун-т – МСХА им. К. А. Тимирязева. – М., 2018. – 24 с.
2. Identification of high utility SNPs for population assignment and traceability purposes in the pig using high-throughput sequencing / A. M. Ramos [et al.] // *Animal Genetics*. – 2011. – Vol. 42, N 6. – P. 613–620. <https://doi.org/10.1111/j.1365-2052.2011.02198.x>
3. A genome-wide association study for a proxy of intermuscular fat level in the Italian Large White breed identifies genomic regions affecting an important quality parameter for dry-cured hams / L. Fontanesi [et al.] // *Animal Genetics*. – 2017. – Vol. 48, N 4. – P. 459–465. <https://doi.org/10.1111/age.12542>
4. Genome-wide association study for ham weight loss at first salting in Italian Large White pigs: towards the genetic dissection of a key trait for dry-cured ham production / L. Fontanesi [et al.] // *Animal Genetics*. – 2017. – Vol. 48, N 1. – P. 103–107. <https://doi.org/10.1111/age.12491>
5. Genetic diversity analysis of two commercial breeds of pigs using genomic and pedigree data / R. Zanella [et al.] // *Genetics Selection Evolution*. – 2016. – Vol. 48. – Art. 24. <https://doi.org/10.1186/s12711-016-0203-3>
6. Genome-wide study on intramuscular fat in Italian Large White pig breed using the PorcineSNP60 BeadChip / R. Davoli [et al.] // *J. of Animal Breeding a. Genetics*. – 2016. – Vol. 133, N 4. – P. 277–282. <https://doi.org/10.1111/jbg.12189>
7. A genome-wide association study in large white and landrace pig populations for number piglets born alive / S. Bergfelder-Drüing [et al.] // *PLoS ONE*. – 2015. – Vol. 10, N 3. – P. e0117468. <https://doi.org/10.1371/journal.pone.0117468>
8. A genomewide association study for average daily gain in Italian Large White pigs / L. Fontanesi [et al.] // *J. of Animal Science*. – 2014. – Vol. 92, N 4. – P. 1385–1394. <https://doi.org/10.2527/jas.2013-7059>
9. A genome-wide association study to detect QTL for commercially important traits in Swiss Large White boars / D. Becker [et al.] // *PLoS ONE*. – 2013. – Vol. 8, N 2. – P. e55951. <https://doi.org/10.1371/journal.pone.0055951>
10. Development of a genetic tool for product regulation in the diverse British pig breed market / S. Wilkinson [et al.] // *BMC Genomics*. – 2012. – Vol. 13. – Art. 580. <https://doi.org/10.1186/1471-2164-13-580>
11. Using SNP array data to test for host genetic and breed effects on Porcine Reproductive and Respiratory Syndrome Viremia / S. Biffani [et al.] // *BMC Proc*. – 2011. – Vol. 5, suppl. 4. – Art. S28. <https://doi.org/10.1186/1753-6561-5-S4-S28>
12. Estimation of U.S. Yorkshire breed composition using genomic data / Y. Huang [et al.] // *J. of Animal Science*. – 2014. – Vol. 92, N 4. – P. 1395–1404. <https://doi.org/10.2527/jas.2013-6907>
13. Uimari, P. Whole-genome SNP association analysis of reproduction traits in the Finnish Landrace pig breed / P. Uimari, A. Sironen, M. L. Sevón-Aimonen // *Genetics Selection Evolution*. – 2011. – Vol. 43. – Art. 42. <https://doi.org/10.1186/1297-9686-43-42>
14. Multi-breed genome-wide association study reveals novel loci associated with the weight of internal organs / Y. He [et al.] // *Genetics Selection Evolution*. – 2015. – Vol. 47. – Art. 87. <https://doi.org/10.1186/s12711-015-0168-7>
15. Investigations on the pattern of linkage disequilibrium and selection signatures in the genomes of German Piétrain pigs / P. Stratz [et al.] // *J. of Animal Breeding a. Genetics*. – 2014. – Vol. 131, N 6. – P. 473–482. <https://doi.org/10.1111/jbg.12107>
16. Roberts, K. S. Relationships among and variation within rare breeds of swine / K. S. Roberts, W. R. Lamberson // *J. of Animal Science*. – 2015. – Vol. 93, N 8. – P. 3810–3813. <https://doi.org/10.2527/jas.2015-9001>
17. Оценка интрогрессии генов свиньи домашней (*Sus scrofa domestica*) в генофонд дикого кабана (*Sus scrofa scrofa*) на основе исследования полиморфизма генов MC1R и NR6A1 / В. Н. Кипень [и др.] // Молекулярная и прикладная генетика : сб. науч. тр. / Ин-т генетики и цитологии Нац. акад. наук Беларуси. – Минск, 2019. – Т. 26. – С. 83–95.
18. Kolb, A. Accurate SNP analysis using the IntelliQube® and duplex BHQplus® genotyping assays with a fast PCR protocol [Electronic resource] / A. Kolb, L. Linz // Annual SLAS conference and exhibition, Washington, Feb. 4–8, 2017 / Walter E. Washington Convention Center. – Mode of access: <http://info.biosearchtech.com/hubfs/docs/SLAS%20Poster%20Abstract.pdf>. – Date of access: 31.10.2019.
19. Okonechnikov, K. Unipro UGENE: a unified bioinformatics toolkit / K. Okonechnikov, O. Golosova, M. Fursov // *Bioinformatics*. – 2012. – Vol. 28, N 8. – P. 1166–1167. <https://doi.org/10.1093/bioinformatics/bts091>
20. Analysis of MC4R polymorphism in Italian Large White and Italian Duroc pigs: association with carcass traits / R. Davoli [et al.] // *Meat Science*. – 2012. – Vol. 90, N 4. – P. 887–892. <https://doi.org/10.1016/j.meatsci.2011.11.025>
21. Evaluation of QTL for carcass merit and meat quality traits in a US commercial Duroc population / I. Choi [et al.] // *Meat Science*. – 2012. – Vol. 92, N 2. – P. 132–138. <https://doi.org/10.1016/j.meatsci.2012.04.023>
22. A genome-wide association scan in pig identifies novel regions associated with feed efficiency trait / G. Sahana [et al.] // *J. of Animal Science*. – 2013. – Vol. 91, N 3. – P. 1041–1050. <https://doi.org/10.2527/jas.2012-5643>
23. Genomic regions affecting backfat thickness and cannon bone circumference identified by genome-wide association study in a Duroc pig population / N. Okumura [et al.] // *Animal Genetics*. – 2013. – Vol. 44, N 4. – P. 454–457. <https://doi.org/10.1111/age.12018>

24. Genome-wide association study reveals genetic architecture of eating behavior in pigs and its implications for humans obesity by comparative mapping / D.N. Do [et al.] // *PLoS One*. – 2013. – Vol. 8, N8. – P. e71509. <https://doi.org/10.1371/journal.pone.0071509>
25. Genome-wide association analysis identifies quantitative trait loci for growth in a Landrace purebred population / E. J. Jung [et al.] // *Animal Genetics*. – 2014. – Vol. 45, N3. – P. 442–444. <https://doi.org/10.1111/age.12117>
26. A genome-wide association study of production traits in a commercial population of Large White pigs: evidence of haplotypes affecting meat quality / M.P. Sanchez [et al.] // *Genetics Selection Evolution*. – 2014. – Vol. 46, N12. – P. 1–12. <https://doi.org/10.1186/1297-9686-46-12>
27. Genome-wide association and systems genetic analyses of residual feed intake, daily feed consumption, backfat and weight gain in pigs / D.N. Do [et al.] // *BMC Genetics*. – 2014. – Vol. 15. – Art. 27. <https://doi.org/10.1186/1471-2156-15-27>
28. Genome-wide association analysis for growth, muscularity and meat quality in Piétrain pigs / P. Stratz [et al.] // *Animal Genetics*. – 2014. – Vol. 45, N3. – P. 350–356. <https://doi.org/10.1111/age.12133>
29. Feed intake, average daily gain, feed efficiency, and real-time ultrasound traits in Duroc pigs: II. Genome-wide association / S. Jiao [et al.] // *J. of Animal Science*. – 2014. – Vol. 92, N7. – P. 2846–2860. <https://doi.org/10.2527/jas.2014-7337>
30. Genome-wide association study using deregressed breeding values for cryptorchidism and scrotal/inguinal hernia in two pig lines / C.A. Sevillano [et al.] // *Genetics Selection Evolution*. – 2015. – Vol. 47. – Art. 18. <https://doi.org/10.1186/s12711-015-0096-6>
31. Genome-wide association study on legendre random regression coefficients for the growth and feed intake trajectory on Duroc Boars / J. T. Howard [et al.] // *BMC Genetics*. – 2015. – Vol. 16. – Art. 59. <https://doi.org/10.1186/s12863-015-0218-8>
32. Genome wide association analysis reveals new production trait genes in a male Duroc population / K. Wang [et al.] // *PLoS One*. – 2015. – Vol. 10, N9. – P. e0139207. <https://doi.org/10.1371/journal.pone.0139207>
33. A WUR SNP is associated with European Porcine Reproductive and Respiratory Virus Syndrome resistance and growth performance in pigs / G. Abella [et al.] // *Research in Veterinary Science*. – 2016. – Vol. 104. – P. 117–122. <https://doi.org/10.1016/j.rvsc.2015.12.014>
34. SNP- and haplotype-based genome-wide association studies for growth, carcass, and meat quality traits in a Duroc multigenerational population / S. Sato [et al.] // *BMC Genetics*. – 2016. – Vol. 17. – Art. 60. <https://doi.org/10.1186/s12863-016-0368-3>
35. *Motsinger, A. A.* Multifactor dimensionality reduction: An analysis strategy for modelling and detecting gene – gene interactions in human genetics and pharmacogenomics studies / A. A. Motsinger, M. D. Ritchie // *Human Genomics*. – 2006. – Vol. 2, N5. – P. 318–328. <https://doi.org/10.1186/1479-7364-2-5-318>
36. *Шейко, И. П.* Белорусский внутривидовой тип свиней в породе дюрок / И. П. Шейко, Р. И. Шейко, Т. Н. Тимошенко // *Вест. Нац. акад. наук Беларусі. Сер. аграр. навук.* – 2016. – № 2. – С. 92–97.

References

1. Korenevskaya P. A. *Productivity and biological characteristics of French selection pigs and their crossbreeds*. Abstract of Ph.D. diss. M., 2018. 24 p. (in Russian).
2. Ramos A. M., Megens H. J., Crooijmans R. P. M. A., Schook L. B., Groenen M. A. M. Identification of high utility SNPs for population assignment and traceability purposes in the pig using high-throughput sequencing. *Animal Genetics*, 2011, vol. 42, no. 6, pp. 613–620. <https://doi.org/10.1111/j.1365-2052.2011.02198.x>
3. Fontanesi L., Schiavo G., Galimberti G., Bovo S., Russo V., Gallo M., Buttazzoni L. A genome-wide association study for a proxy of intermuscular fat level in the Italian Large White breed identifies genomic regions affecting an important quality parameter for dry-cured hams. *Animal Genetics*, 2017, vol. 48, no. 4, pp. 459–465. <https://doi.org/10.1111/age.12542>
4. Fontanesi L., Schiavo G., Gallo M., Baiocco C., Galimberti G., Bovo S., Russo V., Buttazzoni L. Genome-wide association study for ham weight loss at first salting in Italian Large White pigs: towards the genetic dissection of a key trait for dry-cured ham production. *Animal Genetics*, 2017, vol. 48, no. 1, pp. 103–107. <https://doi.org/10.1111/age.12491>
5. Zanella R., Peixoto J.O., Cardoso F.F., Cardoso L.L., Biegelmeyer P., Cantão M. E., Otaviano A., Freitas M.S., Caetano A.R., Ledur M. C. Genetic diversity analysis of two commercial breeds of pigs using genomic and pedigree data. *Genetics Selection Evolution*, 2016, vol. 48, art. 24. <https://doi.org/10.1186/s12711-016-0203-3>
6. Davoli R., Luise D., Mingazzini V., Zambonelli P., Braglia S., Serra A., Russo V. Genome-wide study on intramuscular fat in Italian Large White pig breed using the PorcineSNP60 BeadChip. *Journal of Animal Breeding and Genetics*, 2016, vol. 133, no. 4, pp. 277–282. <https://doi.org/10.1111/jbg.12189>
7. Bergfelder-Drüing S., Grosse-Brinkhaus C., Lind B., Erbe M., Schellander K., Simianer H., Tholen E. A genome-wide association study in large white and landrace pig populations for number piglets born alive. *PLoS ONE*, 2015, vol. 10, no. 3, p. e0117468. <https://doi.org/10.1371/journal.pone.0117468>
8. Fontanesi L., Schiavo G., Galimberti G., Calò D.G., Russo V. A genomewide association study for average daily gain in Italian Large White pigs. *Journal of Animal Science*, 2014, vol. 92, no. 4, pp. 1385–1394. <https://doi.org/10.2527/jas.2013-7059>
9. Becker D., Wimmers K., Luther H., Hofer A., Leeb T. A genome-wide association study to detect QTL for commercially important traits in Swiss Large White boars. *PLoS ONE*, 2013, vol. 8, no. 2, p. e55951. <https://doi.org/10.1371/journal.pone.0055951>

10. Wilkinson S., Archibald A. L., Haley C. S., Megens H. J., Crooijmans R. P., Groenen M. A., Wiener P., Ogden R. Development of a genetic tool for product regulation in the diverse British pig breed market. *BMC Genomics*, 2012, vol. 13, art. 580. <https://doi.org/10.1186/1471-2164-13-580>
11. Biffani S., Botti S., Bishop S. C., Stella A., Giuffra E. Using SNP array data to test for host genetic and breed effects on Porcine Reproductive and Respiratory Syndrome Viremia. *BMC Proceedings*, 2011, vol. 5, suppl. 4, art. S28. <https://doi.org/10.1186/1753-6561-5-S4-S28>
12. Huang Y., Bates R. O., Ernst C. W., Fix J. S., Steibel J. P. Estimation of U. S. Yorkshire breed composition using genomic data. *Journal of Animal Science*, 2014, vol. 92, no. 4, pp. 1395-1404. <https://doi.org/10.2527/jas.2013-6907>
13. Uimari P., Sironen A., Sevón-Aimonen M. L. Whole-genome SNP association analysis of reproduction traits in the Finnish Landrace pig breed. *Genetics Selection Evolution*, 2011, vol. 43, art. 42. <https://doi.org/10.1186/1297-9686-43-42>
14. He Y., Li X., Zhang F., Su Y., Hou L., Chen H., Zhang Z., Huang L. Multi-breed genome-wide association study reveals novel loci associated with the weight of internal organs. *Genetics Selection Evolution*, 2015, vol. 47, art. 87. <https://doi.org/10.1186/s12711-015-0168-7>
15. Stratz P., Wimmers K., Meuwissen T. H., Bennewitz J. Investigations on the pattern of linkage disequilibrium and selection signatures in the genomes of German Piétrain pigs. *Journal of Animal Breeding and Genetics*, 2014, vol. 131, no. 6, pp. 473-482. <https://doi.org/10.1111/jbg.12107>
16. Roberts K. S., Lamberson W. R. Relationships among and variation within rare breeds of swine. *Journal of Animal Science*, 2015, vol. 93, no. 8, pp. 3810-3813. <https://doi.org/10.2527/jas.2015-9001>
17. Kipen V. N., Rabcava A. O., Kotava S. A., Zhurina N. V., Handza A. I., Tsybovsky I. S. Polymorphism analysis of MC1R and NR6A1 genes to evaluate the level of introgression of domestic swine (*Sus scrofa domestica*) genes in wild boar (*Sus scrofa scrofa*) population. *Molekulyarnaya i prikladnaya genetika: sbornik nauchnykh trudov* [Molecular and applied genetics: collection of scientific papers]. Minsk, 2019, vol. 26, pp. 83-95 (in Russian).
18. Kolb A., Linz L. Accurate SNP analysis using the IntelliQube® and duplex BHQplus® genotyping assays with a fast PCR protocol. *Annual SLAS conference and exhibition, Washington, DC, Feb. 4–8, 2017*. Available at: <http://info.bio-searchtech.com/hubfs/docs/SLAS%20Poster%20Abstract.pdf> (accessed 31.10.2019).
19. Okonechnikov K., Golosova O., Fursov M. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, 2012, vol. 28, no. 8, pp. 1166-1167. <https://doi.org/10.1093/bioinformatics/bts091>
20. Davoli R., Braglia S., Valastro V., Annaratone C., Comella M., Zambonelli P., Nisi I., Gallo M., Buttazzoni L., Russo V. Analysis of MC4R polymorphism in Italian Large White and Italian Duroc pigs: association with carcass traits. *Meat Science*, 2012, vol. 90, no. 4, pp. 887-892. <https://doi.org/10.1016/j.meatsci.2011.11.025>
21. Choi I., Bates R. O., Raney N. E., Steibel J. P., Ernst C. W. Evaluation of QTL for carcass merit and meat quality traits in a US commercial Duroc population. *Meat Science*, 2012, vol. 92, no. 2, pp. 132-138. <https://doi.org/10.1016/j.meatsci.2012.04.023>
22. Sahana G., Kadlecová V., Hornshøj H., Nielsen B., Christensen O. F. A genome-wide association scan in pig identifies novel regions associated with feed efficiency trait. *Journal of Animal Science*, 2013, vol. 91, no. 3, pp. 1041-1050. <https://doi.org/10.2527/jas.2012-5643>
23. Okumura N., Matsumoto T., Hayashi T., Hirose K., Fukawa K., Itou T., Uenishi H., Mikawa S., Awata T. Genomic regions affecting backfat thickness and cannon bone circumference identified by genome-wide association study in a Duroc pig population. *Animal Genetics*, 2013, vol. 44, no. 4, pp. 454-457. <https://doi.org/10.1111/age.12018>
24. Do D. N., Strathe A. B., Ostersen T., Jensen J., Mark T., Kadarmideen H. N. Genome-wide association study reveals genetic architecture of eating behavior in pigs and its implications for human obesity by comparative mapping. *PLoS One*, 2013, vol. 8, no. 8, p. e71509. <https://doi.org/10.1371/journal.pone.0071509>
25. Jung E. J., Park H. B., Lee J. B., Yoo C. K., Kim B. M., Kim H. I., Kim B. W., Lim H. T. Genome-wide association analysis identifies quantitative trait loci for growth in a Landrace purebred population. *Animal Genetics*, 2014, vol. 45, no. 3, pp. 442-444. <https://doi.org/10.1111/age.12117>
26. Sanchez M. P., Tribout T., Iannuccelli N., Bouffaud M., Servin B., Tenghe A., Dehais P., Muller N., Del Schneider M. P., Mercat M. J., Rogel-Gaillard C., Milan D., Bidanel J. P., Gilbert H. A genome-wide association study of production traits in a commercial population of Large White pigs: evidence of haplotypes affecting meat quality. *Genetics Selection Evolution*, 2014, vol. 46, no. 12, pp. 1-12. <https://doi.org/10.1186/1297-9686-46-12>
27. Do D. N., Ostersen T., Strathe A. B., Mark T., Jensen J., Kadarmideen H. N. Genome-wide association and systems genetic analyses of residual feed intake, daily feed consumption, backfat and weight gain in pigs. *BMC Genomics*, 2014, vol. 15, art. 27. <https://doi.org/10.1186/1471-2156-15-27>
28. Stratz P., Wellmann R., Preuss S., Wimmers K., Bennewitz J. Genome-wide association analysis for growth, muscularity and meat quality in Piétrain pigs. *Animal Genetics*, 2014, vol. 45, no. 3, pp. 350-356. <https://doi.org/10.1111/age.12133>
29. Jiao S., Maltecca C., Gray K. A., Cassidy J. P. Feed intake, average daily gain, feed efficiency, and real-time ultrasound traits in Duroc pigs: II. Genome-wide association. *Journal of Animal Science*, 2014, vol. 92, no. 7, pp. 2846-2860. <https://doi.org/10.2527/jas.2013-7337>
30. Sevillano C. A., Lopes M. S., Harlizius B., Hanenberg E. H., Knol E. F., Bastiaansen J. W. Genome-wide association study using deregressed breeding values for cryptorchidism and scrotal/inguinal hernia in two pig lines. *Genetics Selection Evolution*, 2015, vol. 47, art. 18. <https://doi.org/10.1186/s12711-015-0096-6>
31. Howard J. T., Jiao S., Tiezzi F., Huang Y., Gray K. A., Maltecca C. Genome-wide association study on legendre random regression coefficients for the growth and feed intake trajectory on Duroc Boars. *BMC Genomics*, 2015, vol. 16, art. 59. <https://doi.org/10.1186/s12863-015-0218-8>

32. Wang K., Liu D., Hernandez-Sanchez J., Chen J., Liu C., Wu Z., Fang M., Li N. Genome wide association analysis reveals new production trait genes in a male Duroc population. *PLoS One*, 2015, vol. 10, no. 9, p. e0139207. <https://doi.org/10.1371/journal.pone.0139207>

33. Abella G., Pena R.N., Nogareda C., Armengol R., Vidal A., Moradell L., Tarancon V., Novell E., Estany J., Fraile L. A WUR SNP is associated with European Porcine Reproductive and Respiratory Virus Syndrome resistance and growth performance in pigs. *Research in Veterinary Science*, 2016, vol. 104, pp. 117-122. <https://doi.org/10.1016/j.rvsc.2015.12.014>

34. Sato S., Uemoto Y., Kikuchi T., Egawa S., Kohira K., Saito T., Sakuma H., Miyashita S., Arata S., Kojima T., Suzuki K. SNP- and haplotype-based genome-wide association studies for growth, carcass, and meat quality traits in a Duroc multigenerational population. *BMC Genetics*, 2016, vol. 17, art. 60. <https://doi.org/10.1186/s12863-016-0368-3>

35. Motsinger A.A., Ritchie M.D. Multifactor dimensionality reduction: An analysis strategy for modelling and detecting gene - gene interactions in human genetics and pharmacogenomics studies. *Human Genomics*, 2006, vol. 2, no. 5, pp. 318-328. <https://doi.org/10.1186/1479-7364-2-5-318>

36. Sheyko I.P., Sheyko R.I., Timoshenko T.N. Belarusian inbred type of pigs in Duroc breed. *Vesti Natsyyanal'nai akademii navuk Belarusi. Seryya agrarnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Agrarian series*, 2016, no. 2, pp. 92-97 (in Russian).

Информация об авторах

Кипень Вячеслав Николаевич – кандидат биологических наук, научный сотрудник лаборатории генетической и клеточной инженерии, Институт генетики и цитологии Национальной академии наук Беларуси (ул. Академическая, 27, 220072 Минск, Республика Беларусь). E-mail: v.kipen@igc.by. <http://orcid.org/0000-0002-7822-0746>

Михайлова Мария Егоровна – кандидат биологических наук, зав. лабораторией генетики животных, Институт генетики и цитологии Национальной академии наук Беларуси (ул. Академическая, 27, 220072 Минск, Республика Беларусь). E-mail: M. Mikhailova@igc.by <http://orcid.org/0000-0001-6087-5069>

Снытков Евгений Владимирович – младший научный сотрудник лаборатории генетики животных, Институт генетики и цитологии Национальной академии наук Беларуси (ул. Академическая, 27, 220072 Минск, Республика Беларусь). E-mail: evsnytkov@gmail.com <http://orcid.org/0000-0001-7961-1952>

Романишко Елена Леонидовна – научный сотрудник лаборатории генетики животных, Институт генетики и цитологии Национальной академии наук Беларуси (ул. Академическая, 27, 220072 Минск, Республика Беларусь). E-mail: lenaramanishko@mail.ru <http://orcid.org/0000-0002-1455-3621>

Иванова Екатерина Вадимовна – стажер младшего научного сотрудника лабораторий генетической и клеточной инженерии. Институт генетики и цитологии Национальной академии наук Беларуси (ул. Академическая, 27, 220072 Минск, Республика Беларусь). E-mail: katya-girl119@mail.ru. <http://orcid.org/0000-0002-4559-6704>

Шейко Руслан Иванович – член-корреспондент НАН Беларуси, доктор сельскохозяйственных наук, профессор, директор Института генетики и цитологии Национальной академии наук Беларуси (ул. Академическая, 27, 220072 Минск, Республика Беларусь). E-mail: R. I. Sheyko@igc.by. <http://orcid.org/0000-0001-5442-566X>

Information about the authors

Viachaslau N. Kipen - Ph. D. (Biological). Institute of Genetics and Cytology of the National Academy of Sciences of Belarus (27, Academicheskaya Str., Minsk 220072, Republic of Belarus). E-mail: v.kipen@igc.by. <http://orcid.org/0000-0002-7822-0746>

Mariya E. Mikhailova - Ph. D. (Biology), head of laboratory of animal genetics. Institute of Genetics and Cytology of the National Academy of Sciences of Belarus (27, Academicheskaya Str., Minsk 220072, Republic of Belarus). E-mail: M. Mikhailova@igc.by. <http://orcid.org/0000-0001-6087-5069>

Evgenij V. Snytkov - Institute of Genetics and Cytology of the National Academy of Sciences of Belarus (27, Academicheskaya Str., Minsk 220072, Republic of Belarus). E-mail: evsnytkov@gmail.com. <http://orcid.org/0000-0001-7961-1952>

Elena L. Romanishko - Institute of Genetics and Cytology of the National Academy of Sciences of Belarus (27, Academicheskaya Str., Minsk 220072, Republic of Belarus). E-mail: lenaramanishko@mail.ru. <http://orcid.org/0000-0002-1455-3621>

Ekaterina V. Ivanova - Institute of Genetics and Cytology of the National Academy of Sciences of Belarus (27, Academicheskaya Str., Minsk 220072, Republic of Belarus). E-mail: katya-girl119@mail.ru. <http://orcid.org/0000-0002-4559-6704>

Ruslan I. Sheyko - Corresponding Member of the NAS of Belarus, D.Sc. (Agricultural), Professor. Institute of Genetics and Cytology of the National Academy of Sciences of Belarus (27, Academicheskaya Str., Minsk 220072, Republic of Belarus). E-mail: R. I. Sheyko@igc.by. <http://orcid.org/0000-0001-5442-566X>