

РАЗВЕДЕНИЕ, СЕЛЕКЦИЯ, ГЕНЕТИКА И БИОТЕХНОЛОГИЯ ЖИВОТНЫХ / BREEDING, SELECTION,
GENETICS AND BIOTECHNOLOGY OF ANIMALS

DOI: <https://doi.org/10.60797/JAE.2024.50.2>

ИСПОЛЬЗОВАНИЕ МЕТОДА ПОЛНОГЕНОМНОГО АССОЦИАТИВНОГО ИССЛЕДОВАНИЯ В СЕЛЕКЦИИ
СВИНЕЙ

Научная статья

Белоус А.А.^{1,*}

¹ ORCID : 0000-0001-7533-4281;

¹ Федеральный исследовательский центр животноводства — ВИЖ имени академика Л.К. Эрнста, Подольск,
Российская Федерация

* Корреспондирующий автор (belousa663[at]gmail.com)

Аннотация

Разработка геномных технологий, направленных на повышение эффекта селекции в животноводстве, напрямую зависит от правильных методических подходов и применения актуальных достоверных программ. Одним из таких направлений является геномная оценка по экономически значимым хозяйственно-полезным признакам. В данной статье описан расчет геномных оценок племенной ценности и проведение полногеномного ассоциативного исследования, на примере среднесуточного прироста свиней породы дюрок. Проанализированы выявленные гены, имеющие биологическую направленность, и приведен анализ их обогащения. Результаты данных исследований можно использовать для написания научных статей, учебных пособий и на курсах повышения квалификации.

Ключевые слова: полногеномное ассоциативное исследование, среднесуточный прирост, аннотация генов, геномная оценка племенной ценности, хряки породы дюрок.

USE OF THE METHOD OF GENOME-WIDE ASSOCIATION STUDY IN PIG BREEDING

Research article

Belous A.A.^{1,*}

¹ ORCID : 0000-0001-7533-4281;

¹ Federal Research Center for Animal Husbandry named after Academy Member L.K. Ernst, Podolsk, Russian Federation

* Corresponding author (belousa663[at]gmail.com)

Abstract

The development of genomic technologies aimed at increasing the effect of breeding in animal breeding directly depends on correct methodological approaches and the use of relevant, reliable programmes. One of such directions is genomic evaluation for economically significant, economically useful traits. This article describes the calculation of genomic evaluations of breeding value and full genomic association study on the example of average daily gain of Duroc pigs. The identified genes with biological orientation are analysed, and their enrichment analysis is presented. The results of these studies can be used for writing scientific articles, training manuals and at advanced training courses.

Keywords: genome-wide association study, average daily gain, gene annotation, genomic evaluation of breeding value, Duroc boars.

Введение

Ускорение генетического улучшения основных хозяйственно-полезных признаков, имеющих значение в секторе экономики, путем применения селекции с помощью маркеров повышает эффективность отбора и разведения свиней. Однако анализ генетической архитектуры этих признаков остается сложной и трудной задачей, и для любого признака у свиней было выявлено несколько определенных причинных вариантов – это миссенс-мутация гена рецептора меланокортина 4 (MC4R), которая участвует в энергетическом гомеостазе и соматическом росте [1], а также в эффективности корма у свиней [2]. Аналогичным образом, миссенс-мутация в гене протеинкиназы В активизируемой некаталитической субъединицы гамма 3 (PRKAG), расположенном на хромосоме 5 *Sus scrofa*, и, как было показано, влияет на уровень гликогена в мышцах у породы гемпшир [3]. Его эффект настолько велик, что данная аллель была устранена из многих коммерческих линий свиней, основанных на гемпшире, с помощью селекции. Было также показано, что мутация сплайсинга в гене каталитической цепи фосфорилазы В киназы гамма (PHKG1), расположенном на хромосоме 3 генома свиней, влияет на содержание гликогена [4]. Чтобы продемонстрировать, что варианты являются причинными, требуются дальнейшие биологические анализы с использованием новых программ. Однако вычислительные методы, такие как полногеномное исследование ассоциаций (GWAS), являются важным первым шагом в идентификации генов-кандидатов. Целью GWAS по одному маркеру является выявление базовой причинной генетической основы признака путем независимого тестирования каждого генетического варианта на статистическую связь с интересующим признаком. Полногеномное секвенирование многих животных является дорогостоящим, и поэтому панели полиморфизма одного нуклеотида (SNP) средней или высокой плотности, в настоящее время, все чаще используются для GWAS у сельскохозяйственных животных. Данный метод основан на неравновесии сцепления (LD) между SNP, присутствующими на панели, и причинными вариантами для идентификации геномных областей, связанных с признаком. Однако высокая LD распространяется на большие

расстояния в популяциях скота, и идентификация причинного варианта из множества вариантов, которые могут находиться в высокой LD с SNP GWAS, остается сложной задачей.

Геномная селекция, основанная на прогнозировании племенной ценности (геномной племенной ценности или GEBV) животных с учетом геномной информации, меняет стратегии и подходы к разведению молочного КРС и ряда других видов животных. Отправными точками для применения геномной селекции у свиней стали разработка первой коммерческой панели однонуклеотидного полиморфизма для высокопроизводительного генотипирования, секвенирование генома свиньи и применение статистических и методологических подходов, впервые разработанных для молочного скота [5], а затем адаптированных к особенностям свиноводческой отрасли [6].

В связи с вышеизложенной актуальностью, целью данного исследования являлся подробное описание методического подхода применения метода GWAS, полученного на рассчитанных геномных оценках племенной ценности среднесуточного прироста свиней коммерческой породы дюрок.

Методы и принципы исследования

Материалом исследования являлись данные геномных оценок племенной ценности показателя среднесуточного прироста 408 хряках породы дюрок 2017-2020 годов рождения, проходивших откорм на автоматических кормовых станциях. В тексте встречается следующее сокращение признака среднесуточного прироста – ADG, г.

Геномный BLUP (GenomicBLUP, GBLUP) также идентичен по применяемой методологии решения уравнения обычному BLUP, однако матрица родства A заменяется матрицей геномного сходства G , учитывающей как генотип по каждому SNP, так и частоту этого генотипа:

$$G = \frac{CC'}{2 \sum p_i(1-p_i)}, \text{ где}$$

C – «центрированная матрица» SNP-маркеров:

$C = M - 2P$, где

M – матрица содержания SNP-файла (закодированная зиготность SNP-маркеров, где 0 – гомозиготность AA, 1 – гетерозиготность AB, 2 – гомозиготность BB), P – матрица частот генотипов, где $P_i = 2(p_i - 0,5)$; p_i – частота встречаемости минорного аллеля (MAF) для i -го SNP.

C' – транспонированная матрица C .

Таким образом, между животными рассчитывается совершенно иной тип родства, позволяющий также учесть, какие сегменты хромосом унаследовали потомки от предков.

Исходя из этого, решение уравнения GBLUP будет иметь вид:

$$\begin{bmatrix} X'X\sigma_e^{-2} & X'Z\sigma_e^{-2} \\ Z'X\sigma_e^{-2} & Z'Z\sigma_e^{-2} + G^{-1}\sigma_a^{-2} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X'y\sigma_e^{-2} \\ Z'y\sigma_e^{-2} \end{bmatrix}, \text{ где}$$

G^{-1} – обратная матрица геномного сходства.

Оценки племенной ценности, получаемые в результате применения GBLUP, называют GEBV (Genomic EBV).

Проведение GWAS основано на определении корреляции между генетическими маркерами средней (несколько десятков тысяч) или высокой (несколько сотен тысяч) плотности и сложными количественными признаками. С этой целью находит применение R-пакет GenABEL, в котором реализована общая линейная смешанная модель. Модель включает в себя случайный полигенный эффект, при котором матрица дисперсии-ковариации пропорциональна единице по всему геному [7]. Для выполнения GWAS предварительно проводят оценку племенной ценности животных, используя смешанные математические модели. В качестве примера можно привести следующую:

$$y = \mu + X + Kw + c + a + e,$$

где y – наблюдаемое значение признака;

μ – популяционная константа X – фиксированный эффект, учитывающий группу содержания, год и сезон;

w – живая масса индивида, рассматриваемая, как коварианта;

c – совокупный эффект SNP;

a – вектор случайных аддитивных генетических эффектов при $\sigma_a^2 \sim N(0, G\sigma_a^2)$, где G – матрица геномного сходства особей, вычисленная на основании данных генотипирования с использованием ДНК-чипов средней или высокой плотности, а σ_a^2 – дисперсия аддитивных полигенных эффектов;

K – коэффициент регрессии оцениваемого признака на живую массу индивидов;

e – остаток модели при $\sigma_e^2 \sim N(0, I\sigma_e^2)$, где I – единичная диагональная матрица, σ_e^2 – остаточная дисперсия.

В данном исследовании GWA-анализ проводили с использованием ДНК-чипа Porcine GGP HD (платформа GeneSeek Genomic Profiler, «Neogene», США), содержащим ~70 тыс. SNP. Контроль качества и фильтрацию данных генотипирования для каждого SNP и каждого образца выполняли с использованием программного пакета PLINK 1.9, применяя следующие фильтры:

- call-rate по всем исследуемым SNP для индивидуального образца не ниже 90% (--mind 0,10);

- call-rate для каждого из исследованных SNP по всем генотипированным образцам не ниже 90% (--geno 0,10);

- частота встречаемости минорных аллелей (MAF) $\geq 0,05$ (--maf 0,05);

- отклонение генотипов по SNP от распределения по Харди-Вайнбергу в совокупности протестированных образцов (--hwe 1e-3).

Достоверность результатов GWAS оценивается по P -значению, которое подвергается коррекции на множественное тестирование. Пороговое значение может быть разным в разных исследованиях, но чаще всего для анализа, в котором рассматривается десятки и сотни тысяч SNP, оно принимается равным 5×10^{-8} . Для корректировки результирующего значения P используются методы проверки нескольких гипотез, такие как Бонферрони, Бенджамина-Хохберга или расчет ожидаемой доли ложных отклонений (FDR) [8].

В результате проведенного анализа, для проведения анализа ассоциаций были отобраны 40 069 SNP, прошедшие фильтрацию по всем параметрам качества.

Для выявления ассоциаций SNP маркеров с изучаемыми признаками применяли регрессионный анализ, реализованный в PLINK 1.90 (--assoc --adjust --qt-means). Для подтверждения достоверного влияния SNP и определения значимых регионов в геноме свиней использовали тест для проверки нулевых гипотез по Бонферрони при пороговом значении $p < 1,25 \times 10^{-5}$, $0,05/40069$. Данные визуализировали в пакете qqman с помощью языка программирования R.

Поскольку стратификация популяции сильно влияет на надежность GWAS, анализ участка квантиль-квантильности (Q-Q) считается эффективным способом определения надежности результатов GWAS. В графике Q-Q горизонтальная ось представляет ожидаемое значение $-\log_{10}P$, а вертикальная ось представляет наблюдаемое значение $-\log_{10}P$. Диагональная линия представляет зависимость вида $y=x$, показывает 95% доверительный интервал, основанный на бета-распределении. Общее отклонение над диагональной линией идентичности обычно указывает на сильную стратификацию популяции. Отклонения от диагональной линии указывают на то, что предполагаемое распределение неверно или что образец содержит значения, возникающие каким-либо другим образом, аналогичные тем, которые порождаются истинным объединением. График Q-Q строится с использованием средств визуализации языка программирования R.

Для поиска генов-кандидатов, локализованных в области идентифицированных SNP, использовали геномный ресурс Sscrofa11.1. Функциональные аннотации генов выполняли с привлечением веб-программы PANTHER 19.0, базы данных DAVID и Pig QTL data base. Чтобы получить представление о функциональном обогащении генов, были впоследствии проведены анализы путей обогащения термина Gene Ontology (GO) [9].

Основные результаты

Впервые в России апробация GWAS анализа по хозяйственно-полезным признакам свиней породы дюрок произошла в 2019 г. [10] и продолжается улучшаться по сей день, включая новые подходы и методы исследования полногеномного анализа [6], [11], [12]. В данной научной работе скомпонован комплекс подходов к получению достоверных генов для дальнейшего их использования в геномной селекции свиней.

По результатам геномных оценок племенной ценности массив животных составил 408 голов, что составляет 47,2% от общего количества изучаемого поголовья. На рисунке 1 представлена диаграмма распределения GEBV по ADG. Наибольшее значение составило +124,5 г, наименьшее – +1,2 г.

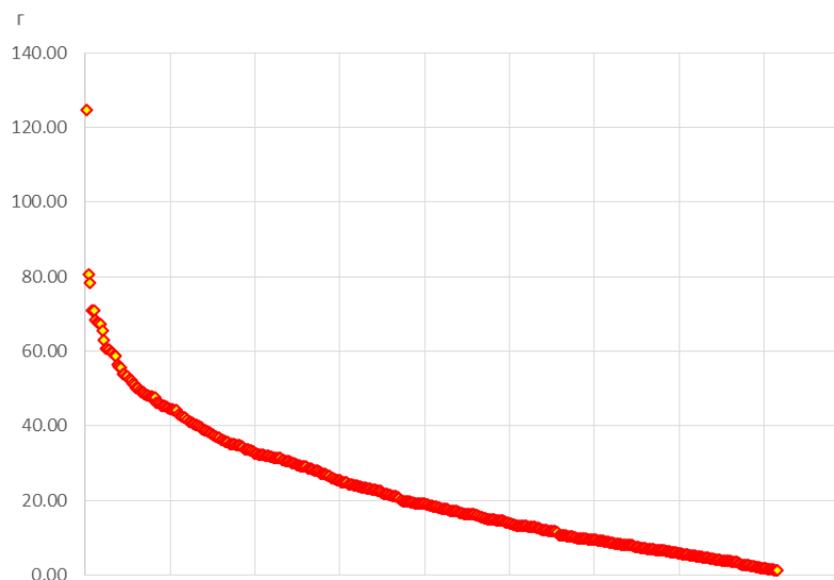


Рисунок 1 - Диаграмма распределения лучших геномных оценок племенной ценности по показателю среднесуточного прироста хряков породы дюрок

DOI: <https://doi.org/10.60797/JAE.2024.50.2.1>

По данным значениям GEBV, описанным выше, на 408 головах свиней породы дюрок, проведено полногеномное ассоциативное исследование, с учетом фильтрации данных, представленных в разделе «Материал и методы исследования», по результатам которых осталось 40 069 SNP и 400 особей. Визуализация Манхэттен плота представлена на рисунке 2. Общее количество достоверно значимых SNP составила 28, из них 3 выходящие за порог полногеномных значений ($p < 1,25 \times 10^{-5}$) и 25 относящихся к диапазону достоверности (от $p < 1,25 \times 10^{-4}$ до $p = 1,25 \times 10^{-5}$).

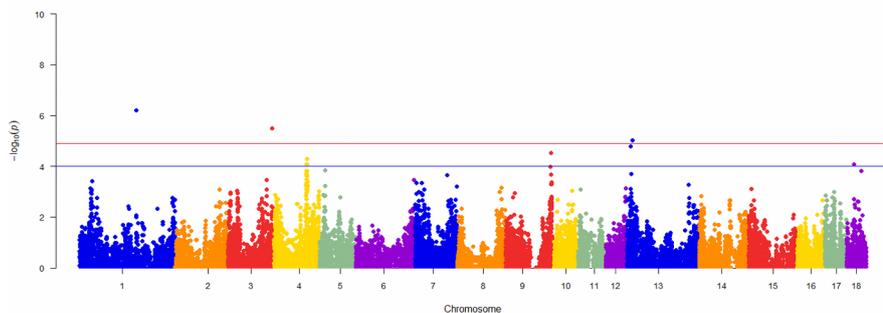


Рисунок 2 - Распределение однонуклеотидных мутаций по хромосомам свиней породы дюрок в связи с уровнем достоверности ($-\log_{10}(p)$) по вероятностному суггестивному значению (синяя линия, $p < 1,25 \times 10^{-4}$) и критерию Бонферрони (красная линия, $p < 1,25 \times 10^{-5}$) для GEBV среднесуточного прироста
DOI: <https://doi.org/10.60797/JAE.2024.50.2.2>

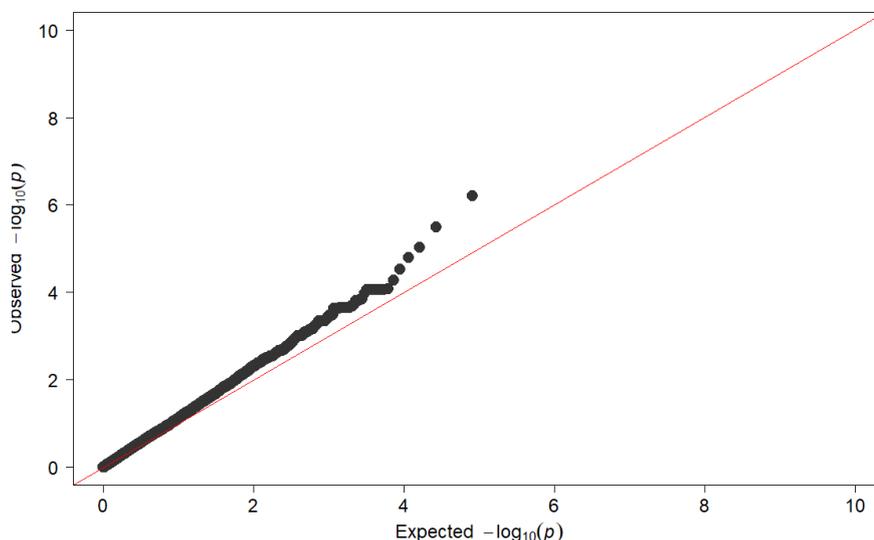


Рисунок 3 - Квартиль вероятностного распределения ожидаемого и наблюдаемого отклонений от нормального распределения для значений достоверности
DOI: <https://doi.org/10.60797/JAE.2024.50.2.3>

Аннотация выявленных SNP по сборке Sscrofa11.1 идентифицировало 79 генов, расположенных на хромосомах 1, 3, 4, 5, 9 и 13 (всего 6 из 18 хромосом). Самая высокая достоверность, превышающая критерий Бонферрони у SNP WU_10.2_1_180284104 - $P = 6,15 \times 10^{-7}$, наименьшая – у SNP DRGA0009891 – $P = 10,6 \times 10^{-4}$. (табл.1).

Наибольшее количество генов расположены на 4 хромосоме ($n=60$).

Таблица 1 - Структурная аннотация достоверно выявленных SNP по признаку среднесуточного прироста

DOI: <https://doi.org/10.60797/JAE.2024.50.2.4>

№хр.	SNP ^{Позиция}	P	Ген ^{Протяженность}
1	WU_10.2_1_18028410 4 ^{162 607 316}	6,15E-07	NEDD4L ^{162364053..162736515}
3	ALGA0116500 ^{127 464 779}	3,17E-06	KIDINS220 ^{127382587..127477753}
4	ASGA0020969 ^{92 646 954}	0,000224	MBOAT2 ^{127243797..127365443}
	ASGA0020983 ^{92 939 481}		ID2 ^{127500759..127503207}
	ASGA0020989 ^{93 062 179}		ETV3 ^{92964808..92979129}
	ASGA0020990 ^{93 081 498}		ETV3L ^{93005259..93011259}

	4_101624287 ^{92 954 515}	0,00019	ARHGEF11 ^{93045533..9316075} 6
	ALGA0026835 ^{93 017 369}	0,000224	PEAR1 ^{93178854..93205186}
	ALGA0026871 ^{93 492 857}		NAXE ^{93493234..93495912}
	MARC0025826 ^{92 994 199}		LRRC71 ^{93162453..93176013}
	WU_10.2_4_10241932 ⁷ ^{93 717 288}	0,000152	NTRK1 ^{93219509..93237944}
			INSRR ^{93239678..93258075}
			PRCC ^{93297859..93325499}
			HDGF ^{93341081..93351809}
			METTTL25B ^{93358140..9336616} 9
			ISG20L2 ^{93366188..93372407}
			MRPL24 ^{93352997..93357805}
			SH2D2A ^{93282383..93293214}
			CRABP2 ^{93388198..93393861}
			NES ^{93412660..93421899}
			IQGAP3 ^{93495999..93556915}
			HAPLN2 ^{93462306..93468394}
			BCAN ^{93432331..93450895}
			GPATCH4 ^{93486365..93492837}
	TTC24 ^{93500900..93512039}		
	MEF2D ^{93581563..93616980}		
	ALGA0026931 ⁹⁴⁷⁸⁶⁶⁶⁶	8,55E-05	GBA1 ^{94583905..94606689}
	ALGA0026937 ^{9 4908711}	8,55E-05	MUC1 ^{94626317..94631194}
	M1GA0006136 ⁹⁴⁹²⁰⁶²⁵	8,55E-05	MTX1 ^{94606576..94611439}
	M1GA0006139 ⁹⁴⁹⁵²⁹⁶⁶	8,55E-05	TRIM46 ^{94632073..94642756}
	MARC0047043 ⁹⁴⁹⁹⁵⁷⁵³	5,27E-05	THBS3 ^{94612702..94623905}
	MARC0051061 ⁹⁴³⁸⁵⁰¹⁵	8,55E-05	KRTCAP2 ^{94643300..94646436}
	MARC0093093 ⁹⁴³³⁹³⁴⁷	8,55E-05	SLC50A1 ^{94663057..94665486}
			EFNA1 ^{94666457..94673442}
			DPM3 ^{94661568..94662223}
			EFNA3 ^{94709941..94719350}
			DCST1 ^{94748430..94768322}
			ADAM15 ^{94736639..94748093}
			EFNA4 ^{94730790..94735578}
			PMVK ^{94861500..94913743}
			PBXIP1 ^{94837365..94847652}
			FLAD1 ^{94807280..94813512}
			SHC1 ^{94818226..94831766}
			DCST2 ^{94768533..94781482}
			PYGO2 ^{94832279..94836548}
			LENEP ^{94805409..94807175}
	CKS1B ^{94815769..94819583}		
	DBWU0000960 ⁹⁶²⁰⁷²⁷⁴	0,000156	CHTOP ^{95996030..96006511}
			S100A1 ^{96007026..96012305}
			GATAD2B ^{95766989..95864502}
			INTS3 ^{95904072..95944805}
			SLC27A3 ^{95898780..95903599}
			S100A13 ^{96007280..96025120}
			SNAPIN ^{95987524..95989527}
			S100A16 ^{96030930..96038602}
			NPR1 ^{95957220..95974253}

			ILF2 ^{95980862..95987568}
			S100A14 ^{96027746..96029878}
			S100A5 ^{96093204..96096870}
			S100A2 ^{96073496..96076825}
			S100A4 ^{96088102..96091160}
			S100A3 ^{96084049..96086829}
			S100A6 ^{96098264..96099686}
			S100A7 ^{96176718..96178694}
5	H3GA0015856 ¹⁷¹⁶³⁴⁹⁴	0,000142	SLC4A8 ^{16842667..16921669}
			SCN8A ^{16977468..17173831}
9	ASGA0044888 ¹²⁹⁵³⁵⁹⁸⁸	3,02E-05	PTPN14 ^{129114668..129248898}
	DRGA0009891 ¹²⁸²⁴⁷⁵⁷⁸	0,000106	PLA2G4A ^{127853581..128164825}
	WU_10.2_9_14238534 ₂ ¹²⁹⁶⁸⁸⁷⁰⁷	0,00021	SMYD2 ^{129263524..129316226}
			PROX1 ^{129527704..129583693}
13	ALGA0067932 ¹⁰⁴¹⁵⁸⁹⁴	1,64E-05	UBE2E2 ^{9998097..10380564}
	ASGA0056369 ¹⁵⁵⁸⁶⁷³⁶	9,42E-06	RBMS3 ^{14968921..16458470}
	ALGA0097418 ²¹⁹³⁸⁰⁸⁰	8,36E-05	GRM8 ^{20945049..21756952}
	WU_10.2_18_4622045 ₈ ⁴²¹⁰⁷⁵⁰⁹	0,0001513	ITPRID1 ^{41491840..41768828}
			NEUROD6 ^{41757356..41763837}
			ADCYAP1R1 ^{41915180..41972327}
GHRHR ^{42030505..42046178}			
			AQP1 ^{42063482..42076741}
			MINDY4 ^{42096351..42202511}

Примечание: жирным шрифтом выделены гены, внутри которых локализован выявленный SNP; №хр. – номер хромосомы; SNP – однонуклеотидный полиморфизм; P – достоверность выявленного SNP

Далее все выявленные гены были соотнесены с биологическими функциями в веб-программе PANTHER 19.0 (рис. 3). Идентифицировано 11 групп генов GO, включая метаболический процесс, который из всех представляет наибольший интерес с сопряженностью со среднесуточным приростом.

При этом, распределение по группам было следующее: GO:0009987 – 22%, GO:0051179 – 4,9%, GO:0044419 – 1,1%, GO:0065007 – 18,1%, GO:0050896 – 9,3%, GO:0042592 – 2,2%, GO:0032502 – 7,7%, GO:0048511 – 0,5%, GO:0032501 – 8,2%, GO:0008152 – 7,1%, GO:0002376 – 1,1% и не была присвоена категория – 17,6%.

PANTHER GO-Slim Biological Process

Total # Genes: 83 Total # process hits: 182

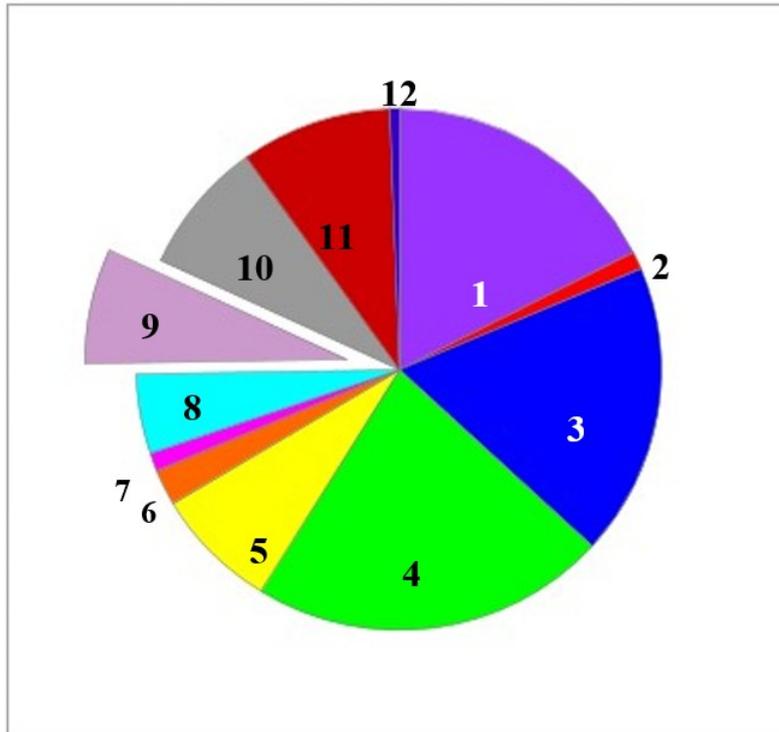


Рисунок 4 - Распределение идентифицированных генов по биологическим функциям (GO):

1 - категория не присвоена; 2 - биологический процесс, участвующий в межвидовом взаимодействии организмов (GO:0044419); 3 - биологическая регуляция (GO:0065007); 4 - клеточный процесс (GO:0009987); 5 - процесс развития (GO:0032502); 6 - гомеостатический процесс (GO:0042592); 7 - процесс иммунной системы (GO:0002376); 8 - локализация (GO:0051179); 9 - метаболический процесс (GO:0008152); 10 - многоклеточный организменный процесс (GO:0032501); 11 - ответ на стимул (GO:0050896); 12 - ритмический процесс (GO:0048511)

DOI: <https://doi.org/10.60797/JAE.2024.50.2.5>

Примечание: выделенная часть в диаграмме – метаболический процесс (GO:0008152)

Цель GO – отразить современное состояние знаний в области биологии, поэтому он постоянно пересматривается и расширяется по мере накопления биологических знаний. По итогу, аннотация генов в с привлечением баз данных DAVID и Pig QTL идентифицировала 30 и 6 генов, соответственно, по которым есть биологический путь и которые зарегистрированы ранее по изучаемому виду животных (табл.2). Интерес представляют гены, внутри которых локализованы SNP, а именно: NEDD4L, KIDINS220, NAXE, S100A1, PLA2G4A и GRM8. На них в дальнейшем можно разработать и апробировать тест-системы, которые смогут детектировать хряков с высоким показателем среднесуточного прироста.

Таблица 2 - Функциональная аннотация генов, ассоциированных с выявленными SNP

DOI: <https://doi.org/10.60797/JAE.2024.50.2.6>

№ хр.	Ген	DAVID (GOTERM_BP)	Pig QTL
1	NEDD4L	нет	QTL:169088
3	KIDINS220	сигнальный путь фактора роста нервов	нет
	MBOAT2	модификация липидов	нет
	ID2	регуляция процессов липидного обмена, двигательный ритм, развитие альвеол молочной	нет

		железы	
4	NAXE	транспорт липидов, регулирование оттока холестерина	нет
	NTRK1	развитие многоклеточного организма, развитие нервной системы	нет
	INSRR	развитие многоклеточного организма, определение мужского пола	нет
	CRABP2	транспорт жирных кислот, эмбриональный морфогенез передних конечностей	нет
	HAPLN2	развитие костной системы, развитие центральной нервной системы	нет
	BCAN	развитие костной системы, центральной нервной системы	нет
	MEF2D	развитие сердца	нет
	GBA1	метаболический процесс, морфогенез мозга, нервно-мышечный процесс	нет
	THBS3	развитие хрящевой пластинки, окостенение	нет
	DCST1	слияние сперматозоида с плазматической мембраной яйцеклетки, участвующей в однократном оплодотворении, распознавание сперматозоидов и яйцеклеток	нет
	ADAM15	врожденный иммунный ответ, переход от эпителия сердца к мезенхиме, иммунный ответ на опухолевую клетку,	нет
	PBXIP1	кроветворение, развитие суставного хряща	нет
	SHC1	ангиогенез, развитие сердца	нет
	PYGO2	развитие почек	нет
S100A1	регуляция сердечных сокращений	нет	
SLC27A3	процесс метаболизма длинноцепочечных жирных кислот	нет	

	S100A14	реакция на липополисахарид	нет
	S100A6	клеточный ответ на вирус	нет
	SCN8A	нет	QTL:160641; QTL:160642
9	PTPN14	нет	нет
	PLA2G4A	процесс метаболизма глицерина, процесс метаболизма арахидоновой кислоты	нет
	PROX1	развитие почек, развитие легких	QTL:179339
13	RBMS3	защитный ответ на опухолевую клетку	QTL:23467
18	GRM8	нет	QTL:19443; QTL:264227
	NEUROD6	развитие нервной системы, развитие органов чувств	нет
	GHRHR	нет	QTL:1186; QTL:1192; QTL:1191

Примечание: жирным шрифтом выделены гены, внутри которых локализован выявленный SNP

Анализ обогащения показал наибольшую локализацию генов по кластеру 1 (коэффициент обогащения = 6,88), принадлежащего к функциям: связывание ионов кальция – 12 генов, $P=4,1 \times 10^{-5}$; кальций-зависимое связывание с белками – 10 генов, $P=2,1 \times 10^{-12}$ и перинуклеарная область цитоплазмы – 8 генов, $P=9,4 \times 10^{-4}$.

В группу связывания ионов кальция (GO:0005509) входят гены: S100A2, S100A16, S100A1, S100A6, S100A13, PLA2G4A, S100A5, S100A4, S100A14, S100A3, S100A7, THBS3. В кальций-зависимое связывание с белками (GO:0048306) – гены S100A2, S100A16, S100A1, S100A6, S100A13, S100A5, S100A4, S100A14, S100A3, S100A7. И в группу, отвечающую за перинуклеарную область цитоплазмы (GO:0048471) – гены SNAPIN, S100A16, S100A1, S100A6, S100A13, S100A4, S100A14, THBS3.

Заключение

Данное научно-практическое исследование можно использовать как методические аспекты структуризации расчета геномной оценки в области свиноводства, применяя актуальные программы для получения достоверного результата.

По выявленным генам, внутри которых локализованы SNP, а именно, NEDD4L, KIDINS220, NAXE, S100A1, PLA2G4A и GRM8, в дальнейшем будут разработаны тест-системы и проведен массовый скрининг на остальных коммерческих породах свиней.

Результаты данных исследований можно использовать для написания научных статей, учебных пособий и на курсах повышения квалификации.

Финансирование

Исследования выполнены в рамках Государственного задания Минобрнауки России № FGGN-2022-0007.

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Funding

The research was carried out within the State Assignment of the Ministry of Education and Science of Russia No. FGGN-2022-0007.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы / References

1. Zeng H. PigGTEx Consortium. PigBiobank: a valuable resource for understanding genetic and biological mechanisms of diverse complex traits in pigs / H. Zeng, W. Zhang, Q. Lin [et al.] // Nucleic Acids Res. — 2024. — № 52 (D1). — P. D980-D989. — DOI: 10.1093/nar/gkad1080.

2. Huang L. Cloning, distribution, and effects of growth regulation of MC3R and MC4R in red crucian carp (*Carassius auratus red var.*) / L. Huang, X. Deng, X. Yang [et al.] // *Front Endocrinol (Lausanne)*. — 2024. — № 14. — P. 1310000. — DOI: 10.3389/fendo.2023.1310000.
3. Milan D. A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle / D. Milan, J.T. Jeon, C. Looft // *Science*. — 2000. — № 288 (5469). — P. 1248-1251. — DOI: 10.1126/science.288.5469.1248.
4. Farhangi S. Expression genome-wide association study identifies key regulatory variants enriched with metabolic and immune functions in four porcine tissues / S. Farhangi, M. Gòdia, M.F. Derks [et al.] // *BMC Genomics*. — Vol. 25. — № 684. — DOI: 10.1186/s12864-024-10583-w.
5. Сермягин А.А. Валидация геномного прогноза племенной ценности быков-производителей по признакам молочной продуктивности дочерей на примере популяции черно-пестрого и голштинского скота / А.А. Сермягин, А.А. Белоус, А.Ф. Контэ [и др.] // *Сельскохозяйственная биология*. — 2017. — Т. 52. — № 6. — С. 1148–1156. — DOI: 10.15389/agrobiology.2017.6.1148rus.
6. Белоус А.А. Генетическая архитектура оценок племенной ценности методом полногеномного анализа ассоциаций свиней породы дюрок / А.А. Белоус, Д.М. Рудиянов, А.А. Решетникова. — DOI: 10.60797/JAE.2024.48.1.
7. Ding R. Genome-wide association analysis reveals genetic loci and candidate genes for feeding behavior and eating efficiency in Duroc boars / R. Ding, J. Quan, M. Yang [et al.] // *PLoS ONE*. — 2017. — Vol. 12(8). — DOI: 10.1371/journal.pone.0183244.
8. Banerjee S. Bayesian multiple logistic regression for case-control GWAS / S. Banerjee, L. Zeng, H. Schunkert [et al.] // *PLoS Genet*. — 2018. — Vol. 14(12). — DOI: 10.1371/journal.pgen.1007856.
9. Ashburner M. Gene ontology: tool for the unification of biology / M. Ashburner, C.A. Ball, J.A. Blake [et al.] // *The Gene Ontology Consortium. Nat Genet*. — 2000. — № 25(1). — P. 25–29. — DOI: 10.1038/75556.
10. Белоус А.А. Изучение генетической архитектуры конверсии корма у хряков (*sus scrofa*) породы дюрок на основе полногеномного анализа SNP / А.А. Белоус, А.А. Сермягин, О.В. Костюнина [и др.] // *Сельскохозяйственная биология*. — 2019. — Т. 54. — № 4. — С. 705–712. — DOI: 10.15389/agrobiology.2019.4.705rus.
11. Траспов А.А. Полногеномные ассоциативные исследования распространения пороков развития и других селекционно значимых качественных признаков у потомства хряков крупной белой породы российской селекции / А.А. Траспов, О.В. Костюнина, А.А. Белоус [и др.] // *Вавиловский журнал генетики и селекции*. — 2020. — № 24 (2). — С. 185–190. — DOI: 10.18699/VJ20.612.
12. Белоус А.А. Зиновьева генетическая характеристика конверсии корма и кормового поведения у свиней породы ландрас / А.А. Белоус, А.Ф. Контэ, П.И. Отрадных [и др.] // *Свиноводство*. — 2022. — № 3. — С. 23–27. — DOI: :10.37925/0039-713X -2022-3-23-27.

Список литературы на английском языке / References in English

1. Zeng H. PigGTE Consortium. PigBiobank: a valuable resource for understanding genetic and biological mechanisms of diverse complex traits in pigs / H. Zeng, W. Zhang, Q. Lin [et al.] // *Nucleic Acids Res*. — 2024. — № 52 (D1). — P. D980-D989. — DOI: 10.1093/nar/gkad1080.
2. Huang L. Cloning, distribution, and effects of growth regulation of MC3R and MC4R in red crucian carp (*Carassius auratus red var.*) / L. Huang, X. Deng, X. Yang [et al.] // *Front Endocrinol (Lausanne)*. — 2024. — № 14. — P. 1310000. — DOI: 10.3389/fendo.2023.1310000.
3. Milan D. A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle / D. Milan, J.T. Jeon, C. Looft // *Science*. — 2000. — № 288 (5469). — P. 1248-1251. — DOI: 10.1126/science.288.5469.1248.
4. Farhangi S. Expression genome-wide association study identifies key regulatory variants enriched with metabolic and immune functions in four porcine tissues / S. Farhangi, M. Gòdia, M.F. Derks [et al.] // *BMC Genomics*. — Vol. 25. — № 684. — DOI: 10.1186/s12864-024-10583-w.
5. Sermjagin A.A. Validacija genomnogo prognoza plemennoj cennosti bykov-proizvoditelej po priznakam molochnoj produktivnosti docherej na primere populjicii cherno-pestrogo i golshitinskogo skota [Validation of genomic prediction of breeding value of bulls-producers by traits of milk productivity of daughters on the example of the population of black-breed and Holstein cattle] / A.A. Sermjagin, A.A. Belous, A.F. Kontje [et al.] // *Sel'skhozjajstvennaja biologija [Agricultural Biology]*. — 2017. — Vol. 52. — № 6. — P. 1148–1156. — DOI: 10.15389/agrobiology.2017.6.1148rus. [in Russian]
6. Belous A.A. Geneticheskaja arhitektura ocenok plemennoj cennosti metodom polnogenomnogo analiza asociacij svinej porody djurok [Genetic architecture of breeding value estimates by whole-genome association analysis of Duroc pigs] / A.A. Belous, D.M. Rudijanov, A.A. Reshetnikova. — DOI: 10.60797/JAE.2024.48.1. [in Russian]
7. Ding R. Genome-wide association analysis reveals genetic loci and candidate genes for feeding behavior and eating efficiency in Duroc boars / R. Ding, J. Quan, M. Yang [et al.] // *PLoS ONE*. — 2017. — Vol. 12(8). — DOI: 10.1371/journal.pone.0183244.
8. Banerjee S. Bayesian multiple logistic regression for case-control GWAS / S. Banerjee, L. Zeng, H. Schunkert [et al.] // *PLoS Genet*. — 2018. — Vol. 14(12). — DOI: 10.1371/journal.pgen.1007856.
9. Ashburner M. Gene ontology: tool for the unification of biology / M. Ashburner, C.A. Ball, J.A. Blake [et al.] // *The Gene Ontology Consortium. Nat Genet*. — 2000. — № 25(1). — P. 25–29. — DOI: 10.1038/75556.
10. Belous A.A. Izuchenie geneticheskoi arhitektury konversii korma u hrjakov (*sus scrofa*) porody djurok na osnove polnogenomnogo analiza SNP [Study of genetic architecture of feed conversion in boars (*sus scrofa*) of the Duroc breed on the basis of full genome SNP analysis] / A.A. Belous, A.A. Sermjagin, O.V. Kostjunina [et al.] // *Sel'skhozjajstvennaja biologija [Agricultural Biology]*. — 2019. — Vol. 54. — № 4. — P. 705–712. — DOI: 10.15389/agrobiology.2019.4.705rus. [in Russian]

11. Traspov A.A. Polnogenomnye asociativnye issledovanija rasprostraneniya porokov razvitija i drugih selekcionno znachimyh kachestvennyh priznakov u potomstva hrjakov krupnoj beloju porody rossijskoj selekcii [Full genome association studies of the distribution of developmental defects and other breeding significant qualitative traits in the offspring of boars of large white breed of Russian selection] / A.A. Traspov, O.V. Kostjunina, A.A. Belous [et al.] // Vavilovskij zhurnal genetiki i selekcii [Vavilov Journal of Genetics and Breeding]. — 2020. — № 24 (2). — P. 185–190. — DOI: 10.18699/VJ20.612. [in Russian]

12. Belous A.A. Zinov'eva geneticheskaja harakteristika konversii korma i kormovogo povedenija u svinej porody landras [Zinoviev genetic characteristics of feed conversion and feeding behaviour in Landrace pigs] / A.A. Belous, A.F. Kontje, P.I. Otradnov [et al.] // Svinovodstvo [Pig Breeding]. — 2022. — № 3. — P. 23-27. — DOI: :10.37925/0039-713X -2022-3-23-27. [in Russian]