AGRICULTURAL BIOLOGY,
2020, V. 55, № 2, pp. 275-284
(SEL'SKOKHOZYAISTVENNAYA BIOLOGIYA)

ISSN 2412-0324 (English ed. Online)
ISSN 0131-6397 (Russian ed. Print)
ISSN 2313-4836 (Russian ed. Online)

UDC 636.4:636.082:575.22:577.21

doi: 10.15389/agrobiology.2020.2.275eng
doi: 10.15389/agrobiology.2020.2.275rus

A SEARCH FOR GENOMIC REGIONS CARRYING THE LETHAL GENETIC VARIANTS IN THE DUROC PIGS

O.V. KOSTYUNINA[1], A.S. ABDELMANOVA[1], E.U. MARTYNOVA[2], N.A. ZINOVIEVA[1]

[1]*Ernst Federal Science Center for Animal Husbandry,* 60, pos. Dubrovitsy, Podolsk District, Moscow Province, 142132 Russia, e-mail kostolan@mail.ru (✉ corresponding author), preevetic@mail.ru, n_zinovieva@mail.ru;
[2]*Center of Life Sciences, Skolkovo Institute of Science and Technology*, 3, ul. Nobelya, Moscow, 143026 Russia, e-mail elenamartynovaster@gmail.com

ORCID:
Kostyunina O.V. orcid.org/0000-0001-8206-3221
Abdelmanova A.S. orcid.org/0000-0003-4752-0727
Martynova E.U. orcid.org/0000-0003-0802-0236
Zinovieva N.A. orcid.org/0000-0003-4017-6863

The authors declare no conflict of interests

Acknowledgements:

The equipment of the Center for Biological Resources and Bioengineering of Farm Animals (Ernst Federal Science Center for Animal Husbandry) was used for the study. Supported financially within the frame of GZ 0445-2019-0026, No. AAAA-A18-118021590138-1

Received October 31, 2019

Abstract

The necessity to address the problem of reducing embryonic losses, which in pigs are estimated at the level up to 30%, is not in doubt. LoF (Loss of Function) mutations, which in the homozygous state can lead to the termination of synthesis or synthesis of non-functional proteins, are considered as one of the genetic factors that cause embryonic mortality. While in cattle, an intensive search for LoF mutations is carried out, in pigs, studies of such mutations are still performed on a smaller scale. Whole-genome analysis using medium- and high-density SNP chips which are uniformly covering the entire genome gives researchers new methodology to identify positional candidates for lethal recessive variants. One approach is to assess the level of linkage disequilibrium (LD) of SNP alleles. In this work, we applied the LD analysis of alleles in SNP loci to detect genome areas with presumptively lethal recessive variants in Duroc pigs (*Sus scrofa*) and for the first time revealed in the genes a series of single nucleotide polymorphisms that significantly affect various physiological processes. Studies were carried out with 715 Duroc boars bred in JSC Top Gen (Voronezh region) in 2017-2019. Whole-genome genotyping was carried out using Porcine GGP HD DNA chips (Neogene/Illumina Inc., USA) containing about 70 thousand SNP. After the quality control, 42981 polymorphic SNP were selected for analysis. Search of reference sequences (rs) and clarification of their localization was carried out using the Ensembl database (http://www.ensembl.org). Functional gene annotations were performed using the GeneCards database (http://www.genecards.org/). Analysis of the maintenance of genetic equilibrium showed the presence of 990 SNPs with the absence of one of the homozygous genotypes (2.30% of the total number of polymorphic SNPs), which were distributed among all pig chromosomes, including 205 SNPs, which were in the linkage disequilibrium (0.48%). Chromosomes SSC9 (0.8 %), SSC5 (0.77 %), SSC7 (0.68%) and SSC2 (0.68%) were characterized by the highest ratio of SNPs in linkage disequilibrium, while chromosomes SSC13 (0.28%), SSC4 (0.29%) and SSC10 (0.30%) were the lowest. For 52 SNPs, of which 25 SNPs were localized within genes, differences in observed and expected heterozygosity frequencies were statistically significant ($p < 0.01$). Among SNPs located in intergenic regions, two SNPs (rs81350198 and rs81337222) are associated with important phenotypes from earlier GWAS studies. For 12 of the 25 identified positional candidate genes (*OR4C45*, *EPHB4*, *EML4*, *SLC4A1AP*, *ZFAT*, *CELSR2*, *NEGR1*, *LRRC32*, *MYOCD*, *HUNK*, *RPH3A*, and *DOCK1*), we obtained the information on their role in various processes in organisms of mammals, including nervous system development, angiogenesis, cardiogenesis, cell differentiation, apoptosis and many others. The integration of DNA markers associated with lethal phenotypes into breeding programs, in addition to DNA markers identified by GWAS studies, will significantly improve the efficiency of marker and genomic breeding programs in pigs.

Keywords: pigs, linkage disequilibrium, lethal variants, LoF, loss of function, single nucleotide polymorphisms

Embryonic losses have significant negative impact on the efficiency and profitability of animal husbandry. In pig breeding, embryonic losses make up to 30% [1]. The so-called LoF (loss of function) mutations, which in the homozygous state lead to premature termination of protein synthesis or to the synthesis of non-functional proteins, are considered among the genetic factors associated with embryonic mortality [2]. The most active search for LoF mutations is carried out in various cattle breeds using the appropriate diagnostic test systems [3]. In pig breeding, studies of LoF mutations are less extensive. To identify mutations leading to visible phenotypic changes, an approach based on the analysis of pedigrees is often used [4]. However, it is not suitable for the identification of LoF mutations, since their phenotypic effects consist in the reduction of multiple pregnancies, which may be due to several other reasons.

The development of methods which enable to analyze whole-genome data has opened up new possibilities for the search for genetic factors associated with embryonic mortality. Sufficient levels of linkage disequilibrium (LD) between the marker allele (alleles) and the lethal variant (variants) are required to successfully use DNA markers for identification of lethal recessive variants which cause prenatal mortality [5]. Two alleles at different loci are in linkage disequilibrium, if the frequency of the haplotype which contains both of them is significantly different from the frequency expected in the case of random allele segregation. Haplotype-based approach is used to more accurately identify rare and atypical variants which are generally not included within the single nucleotide polymorphism (SNP) panels used for genotyping. Phasing the genotype data allows more clearly defining haplotype heterogeneity and makes it possible to draw conclusions about the haplotypes of non-genotyped ancestors and animals that were genotyped using lower density panels. This approach was used to identify the lethal recessive haplotype associated with stillborn piglets' number [6].

Derks et al. [7] searched for lethal alleles segregating in the Landrace (n = 28,085) and Duroc (n = 11,255) pig populations using a medium density SNP chip (Illumina, Inc., USA). Using the overlapping sliding window method, the authors have identified a single strong candidate haplotype (DU1) carrying a lethal recessive allele in the Duroc pig population and four candidate haplotypes in the Landrace breed (LA1-4). No homozygotes were detected for the DU1, LA1, and LA3 haplotypes, while their expected number was 26, 126, and 16, respectively. For the LA2 and LA4 haplotypes, genetic equilibrium disturbance was also observed, which may indicate incomplete LD between the haplotypes and lethal recessive mutations. The association between all five haplotypes and a significant decrease in the total number of born piglets (total number born, TNB) and the number of live born piglets (number born alive, NBA) has been shown. At the same time, no significant increase in the number of stillborn or mummified piglets was observed, which indicates that homozygous carriers die at the early stages of gestation [7].

In the presented work, we have for the first time performed a search for genomic regions which may be associated with lethal recessive defects in Duroc pigs using the analysis of the degree of linkage disequilibrium, and revealed a number of highly significant single nucleotide polymorphisms localized within genes and playing an important role in various physiological processes.

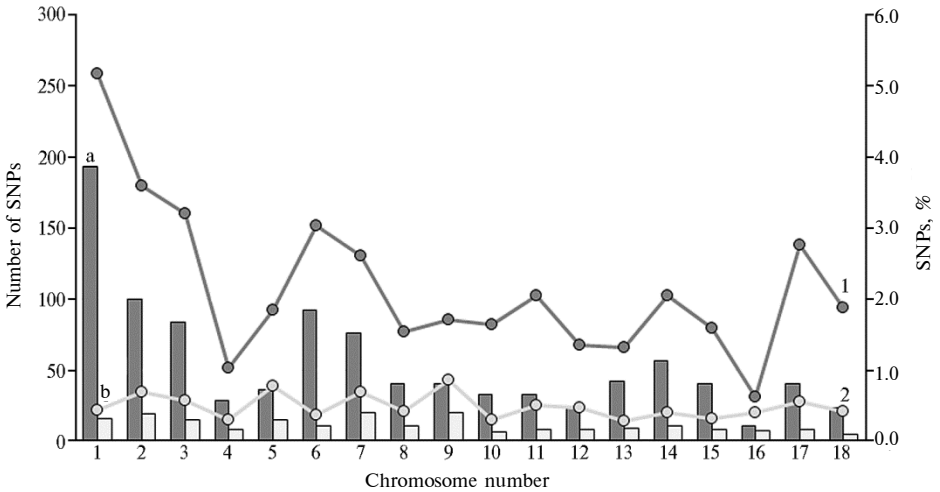
Our goal was to search for the regions bearing presumptively lethal recessive variants in Duroc pigs (*Sus scrofa*) based on the analysis of the degree of linkage disequilibrium between the alleles at SNP loci.

Materials and methods. The study was carried out on 715 boars of the Duroc breed (JSC Breeding-Hybrid Center, Voronezh Region, 2017-2019).

Genomic DNA was extracted from tissue samples (ear plucks) using the DNA Extran-2 kit (Syntol, Russia). DNA quality and concentration were determined using the NanoDrop8000 spectrophotometer (Thermo Fisher Scientific, United States) and the Qubit 2.0 fluorimeter (Invitrogen/Life Technologies, United States), respectively.

Whole-genome genotyping was carried out using the Porcine GGP BeadChips (Neogene/Illumina, Inc., United States) containing about 70 thousand SNPs. 42,981 polymorphic SNPs were selected for the analysis using the Plink 1.9 software (http://zzz.bwh.harvard.edu/plink/) according to genotyping quality (higher that 90%), minor allele frequency (not higher than 0.5%), and linkage disequilibrium (at 50 kb intervals) [8]. The search for reference sequences (reference sequence, rs) and adjustment of their localization was carried out in the Ensembl database (http://www.ensembl.org, accessed date: August 2019). Functional gene annotation was performed using the GeneCards database (http://www.genecards.org, accessed: August 2019).

Results. Genetic equilibrium analysis revealed the presence of 990 SNPs with the absence of one of the homozygous genotypes, among which 205 SNPs were in linkage disequilibrium. These SNPs constituted 2.30 and 0.48% of the total number of polymorphic SNPs, respectively (Fig.).



Chromosomal distribution of SNPs (single nucleotide polymorphisms) for which the absence of one of the homozygous variants and linkage disequilibrium were observed in the Duroc boars (*Sus scrofa*): a — number of SNPs for which one of the homozygous genotypes is missing, b — number of SNPs in linkage disequilibrium (bar graph); 1 — portion of SNPs for which one of the homozygous genotypes is missing in the total number of polymorphic SNPs on the chromosome, 2 — ratio of SNPs that are in linkage disequilibrium in the total number of polymorphic SNPs on the chromosome (line graph) (n = 715, JSC Breeding-Hybrid Center, Voronezh Province, 2017-2019).

The largest number of SNPs for which the absence of one of the homozygous genotypes was observed was detected on chromosome 1 (*Sus scrofa* chromosome 1, SSC1) — 193 SNPs, and the largest number of SNPs in linkage disequilibrium, on the chromosomes SSC7 and SSC9 (20 SNPs), while the smallest number, on SSC16 (11 SNPs) and SSC18 (5 SNPs), respectively. The largest ratio of SNPs in linkage disequilibrium was demonstrated for the chromosomes SSC9 (0.85%), SSC5 (0.77%), SSC7 (0.68%), and SSC2 (0.68%), the smallest ratio, for SSC13 (0.28%), SSC4 (0.29%), and SSC10 (0.30%).

The most significant polymorphisms for which linkage disequilibrium was observed ($p \leq 0.01$) are presented in Table 1.

1. Most significant single nucleotide polymorphisms in linkage disequilibrium detected in the Duroc boars (*Sus scrofa*) (n = 715, JSC Breeding-Hybrid Center, Voronezh Province, 2017-2019)

SSC	RS	Position on chromosome (assembly v.10.2)	A1A1	A1A2	A2A2	H_o	H_e	p
1	rs81350198	226 188 042	0	163	547	0.2296	0.2032	5.1×10^{-5}
1	rs80795638	277 854 630	0	227	440	0.3403	0.2824	1.7×10^{-10}
1	rs81002425	296 461 073	0	700	9	0.9873	0.4999	5.4×10^{-195}
1	rs334911415	312 050 746	0	662	45	0.9364	0.4980	1.2×10^{-152}
1	rs342062641	312 083 991	0	645	63	0.9110	0.4960	1.2×10^{-137}
2	rs81362641	12 082 068	0	618	92	0.8704	0.4916	1.7×10^{-117}
2	rs81255095	14 056 188	0	345	309	0.5275	0.3884	5.2×10^{-28}
2	rs319913462	14 759 005	0	522	126	0.8056	0.4811	4.0×10^{-84}
2	rs323641934	15 384 171	0	537	172	0.7574	0.4706	1.1×10^{-76}
2	rs343381067	15 551 096	0	665	45	0.9366	0.4980	2.0×10^{-153}
2	rs80911461	160 111 983	0	186	506	0.2688	0.2327	7.8×10^{-7}
3	rs344115015	8 107 291	0	139	570	0.1961	0.1768	9.7×10^{-4}
3	rs323044318	104 257 589	0	120	590	0.1690	0.1547	6.4×10^{-3}
3	rs81375606	116 445 052	0	124	586	0.1746	0.1594	4.1×10^{-3}
3	rs81375903	118 615 471	0	142	568	0.2000	0.1800	6.0×10^{-4}
3	rs80828678	132 719 179	0	348	305	0.5329	0.3909	1.9×10^{-28}
3	rs327044542	141265 677	0	170	540	0.2394	0.2108	1.8×10^{-5}
4	rs331053365	4 353 353	0	156	554	0.2197	0.1956	1.4×10^{-4}
4	rs323787335	7 185 799	0	591	114	0.8383	0.4869	1.8×10^{-103}
4	rs343205058	109 273 333	0	604	106	0.8507	0.4889	5.1×10^{-109}
4	rs80949619	121 334 654	0	123	587	0.1732	0.1582	6.7×10^{-3}
5	rs340620949	1 453 085	0	163	546	0.2299	0.2035	5.1×10^{-5}
5	rs81323749	18 046 364	0	184	526	0.2592	0.2256	2.4×10^{-6}
5	rs80875559	99 245 749	0	119	591	0.1676	0.1536	6.5×10^{-3}
6	rs81337222	13 141	0	709	1	0.9986	0.5000	3.3×10^{-210}
6	rs81476539	67 022 132	0	237	471	0.3347	0.2787	6.2×10^{-11}
6	rs337799081	130 798 722	0	130	580	0.1831	0.1663	2.5×10^{-3}
7	rs319008071	27 213 610	0	237	451	0.3445	0.2851	3.0×10^{-11}
7	rs80944793	129 041 116	0	120	590	0.1690	0.1547	6.4×10^{-3}
7	rs331172717	131 889 604	0	614	79	0.8860	0.4935	4.5×10^{-122}
8	rs81399201	31 445 286	0	134	576	0.1887	0.1709	1.6×10^{-3}
8	rs322099448	78 592 741	0	580	127	0.8204	0.4839	5.5×10^{-97}
9	rs343201786	11 913 668	0	137	573	0.1930	0.1743	9.8×10^{-4}
9	rs346413844	12 946 073	0	638	3	0.9953	0.5000	2.7×10^{-185}
9	rs81337172	15 049 063	0	125	585	0.1761	0.1606	4.0×10^{-3}
10	rs81305281	76 905 575	0	520	189	0.7334	0.4645	4.7×10^{-70}
11	rs80816476	939 424	0	131	579	0.1845	0.1675	2.5×10^{-3}
11	rs325221950	21 018 670	0	145	561	0.2054	0.1843	3.7×10^{-4}
11	rs329067201	21 371 677	0	148	562	0.2085	0.1867	3.7×10^{-4}
12	rs81478101	26 712 700	0	697	13	0.9817	0.4998	2.4×10^{-189}
12	rs81436301	50 907 615	0	342	368	0.4817	0.3657	1.9×10^{-24}
12	rs81228589	59 880 454	0	137	573	0.1930	0.1743	9.8×10^{-4}
13	rs329645817	5 325 304	0	320	390	0.4507	0.3491	5.3×10^{-21}
13	rs322958990	205 932 444	0	276	404	0.4059	0.3235	7.0×10^{-16}
13	rs328137225	218 478 357	0	135	575	0.1901	0.1721	1.6×10^{-3}
14	rs80958173	41 518 669	0	125	585	0.1761	0.1606	4.0×10^{-3}
14	rs80862470	95 881 035	0	624	86	0.8789	0.4927	1.6×10^{-121}
14	rs80993446	148 199 735	0	128	582	0.1803	0.1640	4.3×10^{-3}
15	rs337254355	2 154 617	0	138	570	0.1949	0.1759	9.6×10^{-4}
16	rs334615079	81 560 836	0	381	321	0.5427	0.3955	5.0×10^{-32}
17	rs80988530	36 703 557	0	587	68	0.8962	0.4946	4.3×10^{-120}
17	rs345268841	66 549 094	0	258	450	0.3644	0.2980	5.09×10^{-13}

N o t e. SSC — chromosome number, RS — reference sequence, A1 — allele 1, A2 — allele 2, H_o — observed heterozygosity, H_e — expected heterozygosity, and p — level of statistical significance.

A total of 52 significant polymorphisms were identified, for which the differences in the observed and expected heterozygosity were statistically significant. These SNPs were distributed between 17 chromosomes (except for SSC18), including 5 SNPs on SSC1, 6 on SSC2, 6 on SSC3, 4 on SSC4, 3 on SSC5, 3 on SSC6, 3 on SSC7, 2 on SSC8, 3 on SSC9, 1 on SSC10, 3 on SSC11, 3 on SSC12, 3 on SSC13, 3 on SSC14, 1 on SSC15, 1 on SSC16, and 2 on SSC17. Two SNPs (rs81350198 on SSC1 and rs81337222 on SSC6) were identified

as the DNA markers associated with economically-important phenotypes based on the results of the previous GWAS (genome-wide association study) analysis [9, 10]. The rs81350198 polymorphism is associated with the taste of meat in non-castrated boars due to the accumulation of skatol and androstenone during puberty [9]. For the rs81337222 polymorphism, a moderate association was found ($p = 2.4\times10^{-5}$) with the development of umbilical hernia [10]. Although no quantitative trait loci that are associated with umbilical hernia have been identified in the immediate vicinity of rs81337222, and the identified genes for which the association with this phenotypic trait was confirmed have not been localized, its presumable QTL may be located in the upstream region of SSC6 (6:3 814 021-3 870 534) identified by the CNV (copy number variation) analysis [11]. Therefore, at the start of SSC6 there are regulatory regions that require additional research.

2. Single nucleotide polymorphisms localized within the genes and presumptively associated with lethal recessive variants in Duroc pigs (*Sus scrofa*) ($n = 715$, JSC Breeding-Hybrid Center, Voronezh Province, 2017-2019)

RS (p)	Gene name (Ensembl)	Candidate gene	Mutation type
rs343381067 (2.0×10^{-153})	ENSSSCG00000031436	*OR4C45*	Substitution within intron
rs344115015 (9.7×10^{-4})	ENSSSCG00000007675	*EPHB4*	Substitution within intron
rs323044318 (6.4×10^{-3})	ENSSSCG00000008467	*EML4*	Substitution within intron
rs81375606 (4.1×10^{-3})	ENSSSCG00000008533	–	Substitution within intron
rs81375903 (6.0×10^{-4})	ENSSSCG00000008549	*SLC4A1AP*	Substitution at the 3′-end of the gene
rs327044542 (1.8×10^{-5})	ENSSSCG00000049737	–	Substitution within intron
rs323787335 (1.8×10^{-103})	ENSSSCG00000030947	*ZFAT*	Mutation in 3′-UTR
rs343205058 (5.1×10^{-109})	ENSSSCG00000006694	–	Substitution within intron
rs80949619 (6.7×10^{-3})	ENSSSCG00000034360	*CELSR2*	Substitution within intron
rs340620949 (5.1×10^{-5})	ENSSSCG00000024474	–	Substitution at the 3′-end of the gene
rs81323749 (2.4×10^{-6})	ENSSSCG00000024474	–	Substitution at the 3′-end of the gene
rs81476539 (6.2×10^{-11})	ENSSSCG00000003444	–	Substitution at the 3′-end of the gene
rs337799081 (2.5×10^{-3})	ENSSSCG00000025085	*NEGR1*	Substitution within intron
rs319008071 (3.0×10^{-11})	ENSSSCG00000001395	–	Synonymous mutation
rs343201786 (9.8×10^{-4})	ENSSSCG00000014869	*LRRC32*	Mutation in 3′-UTR
rs325221950 (3.7×10^{-4})	ENSSSCG00000045677	–	Substitution at the 3′-end of the gene
rs81436301 (1.9×10^{-24})	ENSSSCG00000017853	–	Substitution at the 3′-end of the gene
rs81228589 (9.8×10^{-4})	ENSSSCG00000031988	*MYOCD*	Substitution within intron
rs322958990 (7.0×10^{-16})	ENSSSCG00000029392	*HUNK*	Substitution within intron
rs80958173 (4.0×10^{-3})	ENSSSCG00000009883	*RPH3A*	Substitution within intron
rs80862470 (1.6×10^{-121})	ENSSSCG00000043778	–	Substitution within intron
rs80993446 (4.3×10^{-3})	ENSSSCG00000035045	*DOCK1*	Substitution within intron
rs337254355 (9.6×10^{-4})	ENSSSCG00000044919	–	Substitution within intron
rs80988530 (4.3×10^{-120})	ENSSSCG00000007155	*C20orf194*	Substitution within intron
rs345268841 (5.09×10^{-13})	ENSSSCG00000007525	–	Missense mutation

N o t e. *OR4C45* — olfactory receptor, family 4, subfamily C, member 45; *EPHB4* — EPH receptor B4; *EML4* — protein associated with excessively expressed proliferation; *SLC4A1AP* — kanadaptin; *ZFAT* — zinc finger protein; *CELSR2* — Cadherin EGF LAG Seven-Pass G-Type Receptor 2; *NEGR1* — Neuronal Growth Regulator 1; *LRRC32* — Leucine-Rich Repeat-Containing protein 32; *MYOCD* — myocardin; *HUNK* — Hormonally Up-Regulated Neu-Associated Kinase; *RPH3A* — rabphilin 3A; *DOCK1* — Dedicator of Cytokinesis. RS — reference sequence, p — level of statistical significance, UTR — untranslated region of the gene. The type of mutation is given as per Borisevich et al. [12]. Dashes indicate the absence of data.

The analysis of the genomic regions in which we found the detected SNPs showed that 25 SNPs were located within genes. By the type of localization, most SNPs were mutations in introns (a total of 15 mutations). Nucleotide substitutions were also found in the 3′-terminal sequences of genes (6 mutations) and in the 3′-untranslated region (2 mutations), along with one missense mutation and one synonymous mutation (Table 2).

For 12 out of the 25 identified positional candidate genes, the data are available on their role in various processes in the mammalian organism. For example, the *OR4C45* gene encodes an olfactory receptor (OR) protein, which is important for maintaining intestinal homeostasis. OR is expressed in the enterochromaffin cells of the mucous membrane. Odorant ligands through OR present in enterochromaffin cells cause serotonin release, which controls motility and intestinal secretion and is involved in the pathological conditions such as vomiting and diarrhea [13]. The role for the OR ligand in the regulation of epithelial permeability and secretion of electrogenic anions in human colon has been reported [14]. *EPHB4* plays a special role in various biological processes, such as neuronal development, bone homeostasis, and angiogenesis [15]. Genetically modified mouse embryos homozygous for the *EphB4taulacZ* allele had cardiovascular defects and were characterized by embryonic mortality with very high penetrance. In such embryos, growth retardation, lack of blood flow, and cardiac development arrest were observed [16]. EML4 is a poorly characterized microtubule-associated protein. It is assumed that its natural function is to stabilize microtubules in the axons and dendrites of neuronal cells. Chimeric EML4-ALK causes the development of lung cancer in humans [17].

Kanadaptin (SLC4A1AP) is a nuclear protein with unknown function that is widely expressed in mammalian tissues. The ubiquitous distribution of kanadaptin in mammals suggests that it should play an important physiological role [18]. ZFAT is involved in the development and peripheral homeostasis of T cells. There is evidence that a deletion in the *Zfat* gene in mice leads to embryonic death and disrupts primitive hematopoiesis in the blood islands of the yolk sac [19, 20]. In pigs, ZFAT is associated with susceptibility to enterotoxin infection caused by *Escherichia coli* [21]. CELSR2 is expressed in all brain areas and regulates the maintenance and growth of dendrites. Mice homozygous for the *CELSR2* mutation develop hydrocephaly due to a decrease in the number, size, and orientation of ependymal cilia [22]. NEGR1 is involved in the regulation of neurite proliferation in the developing brain [23]. As a result of the search for QTL associated with obesity in humans and pigs, three most probable genes have been identified, including *NEGR1*, which is responsible for genetic predisposition to the common obesity types, in particular, for the thickness of subcutaneous fat [24]. LRRC32 functions as a receptor for latent transforming growth factor molecules; it has been detected in regulatory T cells [25]. Its important role in immune regulation has been noted. GWAS studies revealed an association between rs11236909 located approximately 58 kb upstream from the *LRRC32* gene and certain parameters of human sperm motility [26].

MYOCD contributes to heart development and differentiation of cardiomyocytes. It was noted that mutant mice with the knockout of the *MYOCD* gene developed dilated cardiomyopathy, which was accompanied by the disturbance of the structural organization of cardiomyocytes and severe depression of systolic function [27]. The functions of the *HUNK* gene are still not clear. Probably, it is involved in the transfer of phosphorus-containing groups and possesses transferase and protein tyrosine kinase activities. RPH3A plays an important role in the adhesion of neutrophils to endothelial cells during inflammatory reactions [28]. DOCK1 regulates phagocytosis, fusion of myoblasts and cell migration, is involved in embryonic development. The detected underdevelopment of all skeletal muscle tissues in *Dock1*- knockout embryos allowed to identify DOCK1 as an important regulator of the fusion stage in the myogenesis in mammals [29]. No information on the roles of the *ENSSSCG00000024474* gene has been found in open databases; however, GWAS of DNA methylation in lard, lean, and miniature pig breeds, identified it as a differentially methylated region [30]. The *C20orf194 gene* (194th

open reading frame on the 20th chromosome) encodes an uncharacterized protein with the C-terminal coiled coil. The gene is located on the 20p13 chromosome in the 1.8 Mb region associated with the spinocerebellar ataxia phenotype in humans. The work of Ponsuksili et al. [31] concerned with the description of the regions associated with behavioural reactions in Landrace pigs revealed the presence of the rs80988530 SNP, which we detected in the present work in the region with high degree of linkage disequilibrium, within this gene.

To summarize, the genome-wide study which was conducted using the GGP Porcine HD beadchips allowed us to identify the regions mutations in which may cause lethal effects. The most significant single-nucleotide polymorphisms which are in linkage disequilibrium in the boars of the Duroc breed are localized in the following genes: *OR4C45* (olfactory receptor, family 4, subfamily C, member 45); *EPHB4* (EPH receptor B4), *EML4* (protein associated with excessively expressed proliferation), *SLC4A1AP* (kanadaptin), *ZFAT* (zinc finger protein), *CELSR2* (Cadherin EGF LAG Seven-Pass G-Type Receptor 2), *NEGR1* (Neuronal Growth Regulator 1), *LRRC32* (Leucine-Rich Repeat-Containing 32), *MYOCD* (myocardin), *HUNK* (Hormonally Up-Regulated Neu-Associated Kinase), *RPH3A* (rabphilin 3A), and *DOCK1* (Dedicator of Cytokinesis). For almost each identified candidate gene, an important role in various processes, including the development of the nervous system, angiogenesis, cardiogenesis, cell differentiation, apoptosis, etc., has been demonstrated by now. In humans, many of these genes are associated with various organ and tissue disorders, hence, their participation in the occurrence of lethal effects in pigs cannot be ruled out. Understanding the processes which take place during the growth and development of embryos and using this knowledge in the analysis of the actual zootechnical data will expand the arsenal of tools which would allow to propose approaches for the genetic improvement of breeding products in proper time. The integration of DNA markers associated with lethal phenotypes into breeding programs along with the DNA markers identified based on the results of GWAS studies, will significantly increase the efficiency of marker and genomic selection programs in pig breeding.

REFERENCES

1. Pope W.F. Embryonic mortality in swine. In: *Embryonic mortality in domestic species.* M.T. Zavy, R.D. Geisert (eds.). CRC Press, Boca Raton, 1994: 53-77.
2. Bickhart D.M., Hou Y., Schroeder S.G., Alkan C., Cardone M.F., Matukumalli L.K., Song J., Schnabel R.D., Ventura M., Taylor J.F., Garcia J.F., Van Tassell C.P., Sonstegard T.S., Eichler E.E., Liu G.E. Copy number variation of individual cattle genomes using next-generation sequencing. *Genome Res.*, 2012, 22(4): 778-790 (doi: 10.1101/gr.133967.111).
3. Zinov'eva N.A. Haplotypes affecting fertility in Holstein cattle. *Sel'skokhozyaistvennaya biologiya* [*Agricultural Biology*], 2016, 51(4): 423-435 (doi: 10.15389/agrobiology.2016.4.423eng).
4. Gorbach D., Mote B., Totir L., Fernando R., Rothschild M. Polydactyl inheritance in the pig. *Journal of Heredity*, 2010, 101(4): 469-475 (doi: 10.1093/jhered/esq037).
5. Howard D.M., Pong-Wong R., Knap P.W., Woolliams J.A. Use of haplotypes to identify regions harbouring lethal recessive variants in pigs. *Genetics Selection Evolution*, 2017, 49: 57 (doi: 10.1186/s12711-017-0332-3).
6. Häggman J., Uimari P. Novel harmful recessive haplotypes for reproductive traits in pigs. *J. Anim. Breed. Genet.*, 2017, 134(2): 129-135 (doi: 10.1111/jbg.12240).
7. Derks M.F.L., Gjuvsland A.B., Bosse M., Lopes M.S., van Son M., Harlizius B., Tan B.F., Hamland H., Grindflek E., Groenen M.A.M., Megens H.-J. Loss of function mutations in essential genes cause embryonic lethality in pigs. *PLoS Genet.*, 2019, 15(3): e1008055 (doi: 10.1371/journal.pgen.1008055).
8. Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., Sklar P., de Bakker P.I.W., Daly M.J., Sham P.C. PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 2007, 81(3): 559-575 (doi: 10.1086/519795).
9. Drag M., Hansen M.B., Kadarmideen H.N. Systems genomics study reveals expression

quantitative trait loci, regulator genes and pathways associated with boar taint in pigs. *PLoS ONE*, 2018, 13(2): e0192673 (doi: 10.1371/journal.pone.0192673).
10. Fernandes L.T., Ono R.K., Ibelli A.M.G., Lagos E.B., Morés M.A.Z., Cantão M.E., Lorenzetti W.R., Peixoto J. de O., Pedrosa V.B., Ledur M.C. Novel putative candidate genes associated with umbilical hernia in pigs. *Proc. World Congress on Genetics Applied to Livestock Production. Electronic Poster Session. Species. Porcine 2.* Auckland, New Zealand, 2018: 743.
11. Long Y., Su Y., Ai H., Zhang Z., Yang B., Ruan G., Xiao S., Liao X., Ren J., Huang L., Ding N. A genome-wide association study of copy number variations with umbilical hernia in swine. *Anim. Genet.*, 2016, 47(3): 298-305 (doi: 10.1111/age.12402).
12. Borisevich D.I., Shatalova L.V., Korostin D.O., Il'inskii V.V. *Vestnik Rossiiskogo gosudarstvennogo meditsinskogo universiteta*, 2016, 1: 20-24 (in Russ.).
13. Braun T., Voland P., Kunz L., Prinz C., Gratzl M. Enterochromaffin cells of the human gut: sensors for spices and odorants. *Gastroenterology*, 2007, 132(5): 1890-1901 (doi: 10.1053/j.gastro.2007.02.036).
14. Kaji I., Karaki S., Kuwahara A. Effects of luminal thymol on epithelial transport in human and rat colon. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 2011, 300(6): G1132-G1143 (doi: 10.1152/ajpgi.00503.2010).
15. Becerikli M., Merwart B., Lam M.C., Suppelna P., Rittig A., Mirmohammedsadegh A., Stricker I., Theiss C., Singer B.B., Jacobsen F., Steinstraesser L. EPHB4 tyrosine-kinase receptor expression and biological significance in soft tissue sarcoma. *Int. J. Cancer.*, 2015, 136(8): 1781-1791 (doi: 10.1002/ijc.29244).
16. Gerety S.S., Wang H.U., Chen Z.F., Anderson D.J. Symmetrical mutant phenotypes of the receptor EphB4 and its specific transmembrane ligand ephrin-B2 in cardiovascular development. *Molecular Cell*, 1999, 4(3): 403-414 (doi: 10.1016/s1097-2765(00)80342-1).
17. Fry A.M., O'Regan L., Montgomery J., Adib R., Bayliss R. EML proteins in microtubule regulation and human disease. *Biochemical Society Transactions*, 2016, 44(5): 1281-1288 (doi: 10.1042/BST20160125).
18. Xu Q., Deller M.C., Nielsen T.K., Grant J.C., Lesley S.A., Elsliger M.A., Deacon A.M., Wilson I.A. Structural insights into the recognition of phosphopeptide by the FHA domain of kanadaptin. *PLoS ONE*, 2014, 9(9): e107309 (doi: 10.1371/journal.pone.0107309).
19. Tsunoda T., Doi K., Ishikura S., Luo H., Nishi K., Matsuzaki H., Koyanagi M., Tanaka Y., Okamura T., Shirasawa S. Zfat expression in ZsGreen reporter gene knock in mice: Implications for a novel function of Zfat in definitive erythropoiesis. *International Journal of Molecular Medicine*, 2018, 42(5): 2595-2603 (doi: 10.3892/ijmm.2018.3806).
20. Tsunoda T., Takashima Y., Tanaka Y., Fujimoto T., Doi K., Hirose Y., Koyanagi M., Yoshida Y., Okamura T., Kuroki M., Sasazuki T., Shirasawa S. Immune-related zinc finger gene ZFAT is an essential transcriptional regulator for hematopoietic differentiation in blood islands. *Proceedings of the National Academy of Sciences*, 2010, 107(32): 14199-14204 (doi: 10.1073/pnas.1002494107).
21. Ji H.Y., Yang B., Zhang Z.Y., Ouyang J., Yang M., Zhang X.F., Zhang W.C., Su Y., Zhao K.W., Xiao S.J., Yan X.M., Ren J., Huang L.S. A genome-wide association analysis for susceptibility of pigs to enterotoxigenic *Escherichia coli* F41. *Animal*, 2016, 10(10): 1602-1608 (doi: 10.1017/S1751731116000306).
22. Tissir F., Qu Y., Montcouquiol M., Zhou L., Komatsu K., Shi D., Fujimori T., Labeau J., Tyteca D., Courtoy P., Poumay Y., Uemura T., Goffinet A.M. Lack of cadherins Celsr2 and Celsr3 impairs ependymal ciliogenesis, leading to fatal hydrocephalus. *Nat. Neurosci.*, 2010, 13: 700-707 (doi: 10.1038/nn.2555).
23. Schäfer M., Bräuer A.U., Savaskan N.E., Rathjen F.G., Brümmendorf T. Neurotractin/kilon promotes neurite outgrowth and is expressed on reactive astrocytes after entorhinal cortex lesion. *Molecular and Cellular Neuroscience*, 2005, 29(4): 580-590 (doi: 10.1016/j.mcn.2005.04.010).
24. Lee K.-T., Byun M.-J., Kang K.-S., Park E.-W., Lee S.-H., Cho S., Kim H.Y., Kim K.-W., Lee T.H., Park J.-E., Park W.C., Shin D.H., Park H.-S., Jeon J.-T., Choi B.-H., Jang G.-W., Choi S.-H., Kim D.-W., Lim D., Park H.-S., Park M.-R., Ott J., Schook L. B., Kim T.-H., Kim H. Neuronal genes for subcutaneous fat thickness in human and pig are identified by local genomic sequencing and combined SNP association study. *PLoS ONE*, 2011, 6(2): e16356 (doi: 10.1371/journal.pone.0016356).
25. Stockis J., Colau D., Coulie P.G., Lucas S. Membrane protein GARP is a receptor for latent TGF-β on the surface of activated human Treg. *Eur. J. Immunol.*, 2009, 39(12): 3315-3322 (doi: 10.1002/eji.200939684).
26. Kosova G., Scott N.M., Niederberger C., Prins G.S., Ober C. Genome-wide association study identifies candidate genes for male fertility traits in humans. *The American Journal of Human Genetics*, 2012, 90(6): 950-961 (doi: 10.1016/j.ajhg.2012.04.016).
27. Hoofnagle M.H., Neppl R.L., Berzin E.L., Teg Pipes G.C., Olson E.N., Wamhoff B.W., Somlyo A.V., Owens G.K. Myocardin is differentially required for the development of smooth muscle cells and cardiomyocytes. *American Journal of Physiology-Heart and Circulatory Physiology*, 2011, 300(5): H1707-H1721 (doi: 10.1152/ajpheart.01192.2010).

28. Lin C.-C., Huang C.-C., Lin K.-H., Cheng K.-H., Yang D.-M., Tsai Y.-S., Ong R.-Y., Huang Y.N., Kao L-.S. Visualization of Rab3A dissociation during exocytosis: a study by total internal reflection microscopy. *J. Cell. Physiol.*, 2007, 211(2): 316-326 (doi: 10.1002/jcp.20938).
29. Laurin M., Fradet N., Blangy A., Hall A., Vuori K., Côté J.-F. The atypical Rac activator Dock180 (Dock1) regulates myoblast fusion in vivo. *Proceedings of the National Academy of Sciences*, 2008, 105(40): 15446-15451 (doi: 10.1073/pnas.0805546105).
30. Yang Y., Zhou R., Mu Y., Hou X., Tang Z., Li K. Genome-wide analysis of DNA methylation in obese, lean, and miniature pig breeds. *Scientific Reports*, 2016, 6: 30160 (doi: 10.1038/srep30160).
31. Ponsuksili S., Zebunke M., Murani E., Trakooljul N., Krieter J., Puppe B., Schwerin M., Wimmers K. Integrated Genome-wide association and hypothalamus eQTL studies indicate a link between the circadian rhythm-related gene PER1 and coping behavior. *Scientific Reports*, 2015, 5: 16264 (doi: 10.1038/srep16264).
32. Belous A.A., Sermyagin A.A., Kostyunina O.V., Brem G., Zinov'eva N.A. Study of genetic architecture of feed conversion rate in duroc young boars (*Sus scrofa*) based on the genome-wide SNP analysis. *Sel'skokhozyaistvennaya biologiya* [*Agricultural Biology*], 2019, 54(4): 705-712 (doi: 10.15389/agrobiology.2019.4.705rus).