Multi-breed Genetic Parameters and Genome-wide Association Studies for Mortality Rate at Birth in Pigs

Meijing An
HZAU: Huazhong Agriculture University
Guangliang Zhou
HZAU: Huazhong Agriculture University
Yang Li
HZAU: Huazhong Agriculture University
Tao Xiang
HZAU: Huazhong Agriculture University
Yunlong Ma
HZAU: Huazhong Agriculture University
Xiaolei Liu
HZAU: Huazhong Agriculture University
Xinyun Li
HZAU: Huazhong Agriculture University
Shuhong Zhao
HZAU: Huazhong Agriculture University
Mengjin Zhu (zhumengjin@mail.hzau.edu.cn)
Huazhong Agricultural University: Huazhong Agriculture University https://orcid.org/0000-0001-8931-5022

Research

Keywords: mortality rate at birth, heritability, genetic correlation, GWAS, pig

Posted Date: January 15th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-146253/v1

Title page

Multi-breed genetic parameters and genome-wide association studies for mortality rate at birth in pigs

Meijing An [1], Guangliang Zhou [1], Yang Li [1,] Tao Xiang [1,2], Yunlong Ma [1,2], Xiaolei Liu [1,2], Xinyun Li [1,2], Shuhong Zhao [1,2] and Mengjin Zhu[1,2*]

1 Key Lab of Agricultural Animal Genetics, Breeding, and Reproduction of Ministry of Education, Huazhong Agricultural University, Wuhan 430070, China
2 The Cooperative Innovation Center for Sustainable Pig Production, Huazhong Agricultural University, Wuhan 430070, China
* Correspondence: Mengjin Zhu zhumengjin@mail.hzau.edu.cn

Meijing An	anmeijing521@gmail.com
Guangliang Zhou	glzhou@webmail.hzau.edu.cn
Yang Li	yli@webmail.hzau.edu.cn
Tao Xiang	Tao.Xiang@mail.hzau.edu.cn
Yunlong Ma	Yunlong.Ma@mail.hzau.edu.cn
Xiaolei Liu	xiaoleiliu@mail.hzau.edu.cn
Xinyun Li	xyli@mail.hzau.edu.cn
Shuhong Zhao	shzhao@mail.hzau.edu.cn
Mengjin Zhu	zhumengjin@mail.hzau.edu.cn

RESEARCH

Multi-breed genetic parameters and genome-wide association studies for mortality rate at birth in pigs

Abstract

Background

Piglet mortality is an economically important complex trait that impacts sow prolificacy in the pig industry. The genetic parameters estimations and genome-wide association studies will help us to better understand the genetic fundamentals of piglet mortality. However, compared with other economically important traits, a little breakthrough in the genetic analyses of the trait has been achieved.

Results

In this study, we used multi-breed data sets from Yorkshire, Landrace, and Duroc sows and characterized the genetic and genomic properties of mortality rate at birth by treating each parity as a different trait. The heritability of mortality rate from parity I to III were estimated to be 0.0630, 0.1031, and 0.1140, respectively. The phenotypic and genetic correlations with its component traits were all positive with ranges from 0.0897 to 0.9054, and 0.2388 to 0.9999, respectively. Integrating the results, we identified 21 loci that were detected at least by two tools from standard MLM, FarmCPU, BLINK and mrMLM, and these loci were annotated to 22 genes. The annotations revealed that the gene expressions were associated with the reproductive system, nervous system, digestive system, and embryonic development, which are reasonably related to the piglet mortality.

Conclusions

In brief, the genetic properties of piglet mortality at birth were reported. These findings are expected to provide much information for understanding the genetic and genomic fundamentals of farrowing mortality and also identify candidate molecular markers for breeding practice.

Keywords: mortality rate at birth, heritability, genetic correlation, GWAS, pig

1. Background

Piglet mortality-related traits are a category of economically important traits that provide direct or indirect metrics of piglet deaths and produce heavy economic losses and welfare concerns to the pig industry (1, 2). The majority of piglet mortality traits have been documented to be

complex traits with low heritability ranging from 0.03 to 0.17 (3). In theory, the outcome of mortality is the tri-interactions between piglet, sow and environment, and the phenotypic variation could be caused by diverse systematic and non-systematic factors, including genetic background (breed), parity, season, disease, management, piglet vitality, and sow's behavior such as crushing and starvation (4-6). For the high economic merit, there has been a growing emphasis on reducing piglet losses in pig breeding programs of some countries.

Past experiences from breeders revealed that selection on litter size increases piglet mortality and the intensive couplings with litter size implied that there might exist strong negative linkage disequilibrium (LD) or opposite pleiotropy in the cross-trait genetic architectures (7). However, to date, there have been no more than 10 reports publicly available to describe the genomic fundamentals of piglet mortality-related traits (8, 9). Compared with other economically important traits, little breakthrough in the genetic dissections of piglet mortality-related traits has been achieved. The limited progress cannot underpin a pinpoint understanding of genetic properties of piglet mortality-related traits, and further efforts are needed.

There are usually two time points to measure mortality, including at birth and at weaning (10) . No matter at weaning or at birth the mortality is measured, the metrics are derived from its component traits, i.e. litter size-related traits. When dealing with piglet mortality as well as its components, there remains an important concern for parity. It is still unable to reach a consensus about how to treat this type of data sets in practice. In theory, during the first parities, the reproductive organs of gilt are still undergoing developmental changes, while for higher parity sows the risk of death increases due to oxytocin insufficiency and ruptured umbilical cord (11). Given these, many researchers treated each parity as a different trait in genetic analyses. For example, Roehe & Kennedy (1995) reported that the genetic parameters of litter size were estimated with each parity treated as a different trait (12). More studies revealed that the estimations of genetic parameters varied between different parities in different pig cohorts (13, 14). In addition, researchers also found that the reproductive traits in different parities had a different genetic architecture (15). So, it's quite sound to treat each parity as a different trait.

There were growing studies that used multi-breed data sets for genetic parameter estimation and GWAS. In general, the multi-breed approach has potential advantages, such as enlarging the sample size by putting the multi-breed individuals together, capturing the genetic variants both within and across breeds, and improving the accuracy of genetic evaluation. For example, a simulation study has evaluated the efficiency of the multi-breed approach, and reported that the multi-breed approach could improve the accuracy of genomic estimated breeding values (GEBVs) for the second breed with fewer sizes (16). It was also found that the multi-breed models had a positive effect on the genetic parameter estimations (17). Raven et al. (2014) declared that the multi-breed approach could accurately locate the highly conserved functional mutations because the mixed population had lower levels of long-range LD (18). The multi-breed approach has been widely proven to be feasible in genetic analyses.

Knowledge of genetic property for a trait is involved in many aspects, in which genetic parameters and genomic architecture are two important ones. The aim of this study was to characterize the genetic property of mortality rate at birth using the mixture data sets from

Yorkshire, Landrace, and Duroc sows. In this study, the genetic parameters including breeding value, heritability, and genetic correlation between piglet mortality and its component traits from parity I to III were estimated, and GWAS on piglet mortality was performed to identify the genome-wide variants and putative genes underlying the variability of piglet mortality. This study would accelerate our understanding of the molecular fundamentals of piglet mortality and provides potential markers for pig breeding programs.

2. Methods

2.1.Animals and phenotype collection

Raw data sets were collected from southern China. The breeds in the data sets included Yorkshire, Landrace and Duroc pigs. The raw records were produced from January 2014 to June 2018. Considering that there were too many levels of farrowing dates, we re-formatted them as the labels of seasons. According to the geographical location and weather condition of the local farm, farrowing dates from March to May, from June to August, from September to November, and from December to February of the following year were re-labeled as four seasons. All available fixed factors were preliminarily tested by generalized linear model (GLM), and only the fixed factors that passed the preliminary test were retained for model establishment, which included breed, parity, and re-formatted season. In this study, the mortality rate was derived from its component traits at birth, which was defined as the ratio of the total number dead (TND) over the total number born (TNB). The component traits included the total number born (TNB), the number of stillborn piglets (NS), and the number of mummified at birth (NM). NS was the number of intrapartum deaths during farrowing, and NM was the number of antepartum deaths with tissue degeneration or absorption. The mortality rate at birth was measured using the formula below:

$$\boldsymbol{Mortality} = \frac{\boldsymbol{TND}}{\boldsymbol{TNB}}$$

In the formula, $\boldsymbol{TND} = \boldsymbol{NS} + \boldsymbol{NM}$. Considering the sample size, we analyzed the data set from parity I to III, including 6,073 individuals from parity I, 5,415 individuals from parity II and 4,378 individuals from parity III. In total, there were 35,313 individuals in the pedigree that were used to construct the numerator relationship matrix for estimations of genetic parameters. A more detailed information about the data structure for raw data sets was shown in Additional Table 1.

2.2. Estimations of breeding value, heritability, and genetic correlation

We implemented the pedigree-based best linear unbiased prediction to estimate the genetic parameters, in which the breed and seasons were taken as fixed effects (19, 20). Based on the multi-breed approach, estimations of estimated breeding value (EBV), heritability and genetic correlation were calculated using the HIBLUP software developed by our lab (https://hiblup.github.io). The mixed linear model was formulated as follow:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e}$$

In the model, y was a vector of observations, β is a vector of fixed effects, including the fixed mean, breed, and re-formatted season (19); u ~ N(0, G) was a vector of breeding values, and G ~ $A\sigma_a^2$, in which σ_a^2 was additive genetic variance and A was an additive genetic relationship matrix derived from the pedigree structure; e ~ N (0, $I\sigma_e^2$) represented the residuals, where σ_e^2 was the residual variance. X and Z were design matrices for β and u, respectively. When estimating genetic correlations, the bivariate models had the same components as the univariate models, and the between-trait genetic and residual variance-covariance structures were defined as $\begin{bmatrix} \sigma_{a_x}^2 A & \sigma_{a_x a_y} A \\ \sigma_{a_y a_x} A & \sigma_{a_y}^2 A \end{bmatrix}$ and $\begin{bmatrix} \sigma_{e_x}^2 & \sigma_{e_x e_y} \\ \sigma_{e_y e_x} & \sigma_{e_y}^2 \end{bmatrix}$, in which A was additive genetic relationship matrix, σ_a^2 was additive genetic variance, σ_e^2 was the residual variance, $\sigma_{a_x a_y}$ was the genetic covariance between two traits, $\sigma_{e_x e_y}$ was the residual covariance between two traits, and the subscripts x and y denoted two traits. The average information restricted maximum likelihood (AI-REML) algorithm, i.e., iterative algorithm based on Newton iteration and Fisher score method, was used for (co)variance components estimation.

2.3. DNA isolation, genotyping and quality control

Ear tissue samples were collected and stored in freezers at -20 °C. Genomic DNA was extracted using Tecan Freedom EVO NGS workstation and magnetic animal tissue genomic DNA kit (TIANGEN) according to the manufacturer's protocol. In total, 1,331 individuals were genotyped, of which 1,331 individuals had phenotypic data in parity I, 1,220 individuals in parity II, and 980 individuals in parity III. Genotyping was conducted using the Illumina PorcineSNP60 Bead Chip. All SNPs were mapped to Sus scrofa genome build 11.1 (21). When performing quality control (QC), the SNPs with call rates ≤ 90%, minor allele frequencies ≤ 1% were removed by PLINK (22). The missing genotypes were imputed by Beagle software (23) and the imputed genotype data were also filtered, using the same conditions as the former.

2.4. GWAS and integration of results

To increase the detection credibility, we used a combined approach for GWAS. Concretely speaking, four tools, including the standard MLM, FarmCPU (R package "rMVP") (24), BLINK, and mrMLM, were simultaneously used to perform the GWAS analyses (25-28), and then combined the results from four tools. In the integration, considering that the Bonferroni correction is usually too conservative and may miss putative SNPs with medium effect size in GWAS (29), after normal GWAS analyses, we alternatively used a soft-cutoff to determine the putative SNPs, in which we first sorted the SNPs according to the p-values, and then made the intersections between different tools with following a permutation test procedure to statistically confirm the validity of selection of putative SNPs. It was reported that independent replication in a different cohort provides a gold standard approach for identification of putative SNPs (30), and here, borrowing the similar idea, the SNPs that were repeatedly identified at least by two different tools were suggested as putative SNPs. Simultaneously, considering that the permutation test is computationally intensive, for each tool and each parity, no matter what the p-value was, only top 10 SNPs were selected for combined analyses. For permutation test, with 10,000 random shuffles of real phenotypes, the MLM-based GWAS technique were repeatedly conducted to produce 10,000 pseudo p-values. In the distribution of pseudo p-values of permutation test, the position of raw p-value was referenced to determine the permutated p-values of target SNPs. Here, according to the principle of small probability in statistics, if a p-value derived from the real data is less than the pseudo p-value at quantile 0.05 or 0.01 of the permutated distribution, the p-value is defined to be statistically or high statistically significant.

2.5. Gene annotations

The genes harboring or closely neighboring the identified SNPs were confirmed by mapping into Sus scrofa genome version v11.1. The chromosomal coordinate information was extracted to annotate the candidate genes by selecting the closest gene for each identified SNP based on the Ensembl database (http://uswest.ensembl.org/Sus_scrofa/Info/Index). We used the latest version of the TISSUES database (TISSUES 2.0) to reconstruct the digital expression profiles of putative genes from different tissues, where the top 20 tissues with high confidence score were selected to visualize the gene-tissue expression relationships through the heat map (31).

3. Results

3.1. Distributions of phenotypes and breeding values for piglet mortality

The distributions of phenotypes and EBVs from parity I to III were shown in Figure 1, where the phenotypic distributions were shown in the sub-figures from A to C, and EBVs in the sub-figures from D to F. In this study, the piglet mortality was defined as a ratio trait that is calculated as the ratio of total number dead (TND) over total number born (TNB). In usual, a ratio trait is departure from the normal distribution, and it was found that the phenotypes of piglet mortality from parity I to III followed a heavy skewed distribution. Compared with the heavy skewed distribution of phenotypes, all distribution curves of the EBVs from parity I to III had two tails with a positively skewed distribution, which were relatively closer to the normal distribution. For more detailed information, the descriptive statistics of raw data sets was given in Additional Table 4.

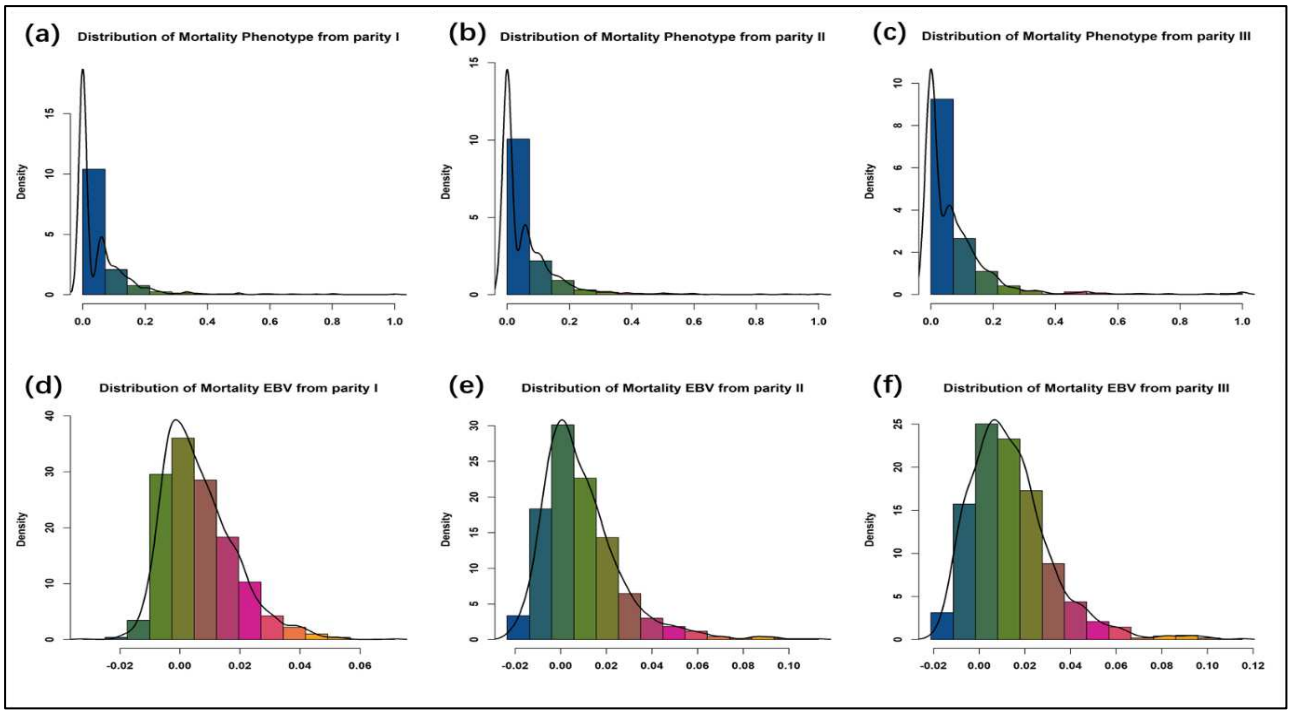

Figure 1. The phenotypic and EBV's distributions of piglet mortality from the three parities. Sub-figures from a to c represented the phenotypic distributions from parity I to parity III, and sub-figures from d to f represented the distributions of EBVs from parity I to III, respectively.

3.2. Estimation of genetic parameters for piglet mortality and its component traits

The results of heritability estimation for piglet mortality were presented in Table 1. The estimations of heritability from parity I to III were 0.0630, 0.1031 and 0.1140, respectively. According to the classification of heritability, the piglet mortality could be considered as a trait with low heritability. Considering the difference of heritabilities between three parities, the same trait from different parities could be taken as different traits. The heritabilities of the component traits of piglet mortality were also listed in Additional Table 2. Table 2 showed the estimations of genetic and phenotypic correlations between piglet mortality and its components traits, including TNB, TND, NS, and NM. All genetic correlations were positive ones ranging from 0.2388 to 0.9999. For the same trait-pair, the estimations of genetic correlations much differed in different parities. The differences of genetic correlation coefficients between three parities also supported taking the same trait in different parities as different traits.

Table 1. Estimations of heritability and standard error (heritability ± se) of piglet mortality from parity I to III.

Trait	Parity I	Parity II	Parity III
Mortality rate	0.0630±0.0219	0.1031± 0.0269	0.1140±0.0285

Table 2. Estimations of genetic and phenotypic correlations between piglet mortality and its component traits from parity I to III.

Parity I					
	TND	TNB	NS	NM	Mortality
TND		0.2306	0.6745	0.8258	0.8854
TNB	0.7042		0.1526	0.1934	0.0442
NS	0.7301	0.9999		0.1405	0.6347
NM	0.8952	0.2388	0.3491		0.7030
Mortality rate	0.9999	0.6549	0.5816	0.9999	
Parity II					
	TND	TNB	NS	NM	Mortality
TND		0.2692	0.6992	0.7748	0.8929
TNB	0.8245		0.2062	0.1922	0.0903
NS	0.9382	0.8577		0.0897	0.6485
NM	0.8679	0.5985	0.6424		0.6745
Mortality rate	0.7747	0.617	0.8596	0.8131	
Parity III					
	TND	TNB	NS	NM	Mortality
TND		0.2939	0.7390	0.7568	0.9054
TNB	0.6131		0.2341	0.2057	0.1055
NS	0.9523	0.4940		0.1189	0.6815
NM	0.7349	0.7471	0.4928		0.6730
Mortality rate	0.9916	0.4896	0.9860	0.7773	

Note: In each parity, the correlation coefficients and standard errors (bracketed) for genetic correlations were in the lower triangle, and phenotypic correlations were in the upper triangle. TND was the total number dead, TNB was the total number born, NS was the number of stillborn piglets, and NM was the number of mummified at birth.

3.3. Results of GWAS analyses for piglet mortality

After quality control, a total of 47,241 SNPs were passed the filtering options. Principal component analysis (PCA) was carried out and the scatterplot of the first two principal components were displayed in Additional Figures S1. Four tools, including standard MLM, FarmCPU, BLINK, and mrMLM were used to run the GWAS analyses. In the GWAS analyses, the target trait from parity I to III have different sample sizes (parity I, n=1331; parity II, n=1220; parity III, n=980). The top ten SNPs identified by each tool for each parity were listed in Additional Table 5. After extracting and summarizing the results of GWAS, Figure 2 showed the circular-Manhattan plots of piglet mortality traits from MLM, FarmCPU, BLINK, mrMLM. In addition, the Venn diagrams were drawn to identify the intersections of the top ten SNPs from four tools (Figure 3).

In total, 21 SNPs were identified, of which, 6 belonged to parity I, 5 belonged to parity II, and 10 belonged to parity III. For the identified SNPs from different parities, no overlapping was observed. All identified SNPs passed the permutation test, and were statistically confirmed. The SNP symbols, the smallest p-values from GWAS, and permutated p-values of the identified SNPs were listed in the Additional Table 3.

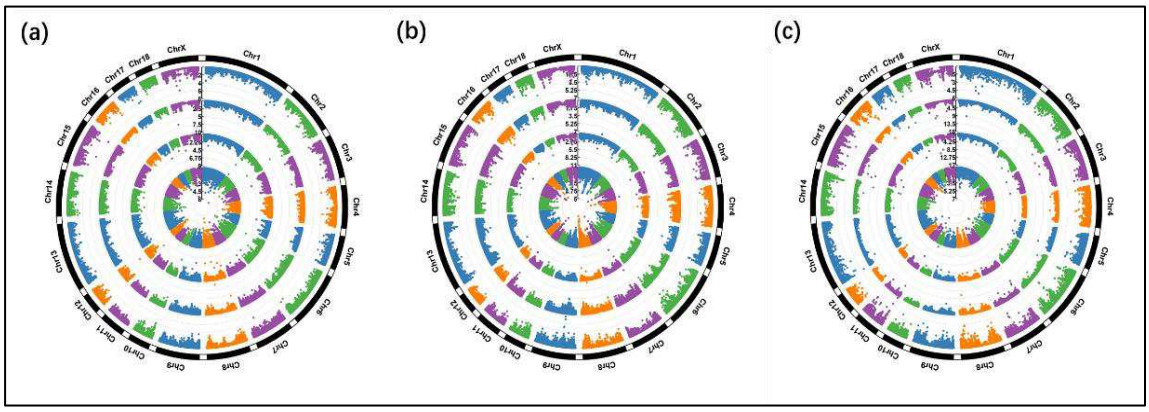

Figure 2. Circular-Manhattan of piglet mortality of different parities. a. Circular-Manhattan plots of MLM, FarmCPU, BLINK, mrMLM for parity I; b. Circular-Manhattan plots of MLM, FarmCPU, BLINK, mrMLM for parity II; c. Circular-Manhattan plots of MLM, FarmCPU, BLINK, mrMLM for parity III; and from inner to outer, they were MLM, FramCPU, BLINK and mrMLM, respectively.

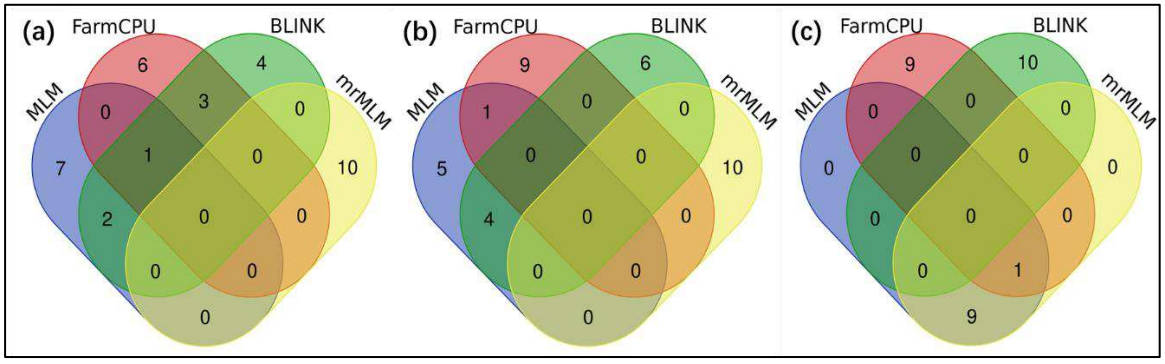

Figure 3. The Venn diagrams of identified SNPs in piglet mortality from parity I, II, III a. Venn diagrams of MLM, FarmCPU, BLINK, mrMLM method for parity I. There were 6 SNPs intersected by at least two tools; b. Venn diagrams of MLM, FarmCPU, BLINK, mrMLM method for parity II. There were 5 SNPs intersected by two tools;

c. Venn diagrams of MLM, FarmCPU, BLINK, mrMLM method for parity III. There were 10 SNPs intersected by two tools.

3.4. Gene annotations

All SNPs that passed the permutation test were further used for gene annotations. In total, we obtained 22 candidate genes that harbor or near the 21 identified SNPs. The position information of 21 SNPs and corresponding 22 genes were shown in Table 3. Among them, the positions of MARC0113660 and DRGA0008818 are located within 5.8kb, and there is only one gene (STPG2) in this region. According to the annotation criterion, there were two genes CDK8 and WASF3 that were both close to ALGA0060358. Three SNPs, including ALGA0036320, H3GA0018655 and ASGA0029165, were annotated to be close to four genes that included CABYR, OSBPL1A, IMPACT, and HRH4. It can be found that, among these SNPs, there were totally eight SNPs clustered on SSC 8. Furthermore, we used the information extracted from TISSUES database (TISSUES 2.0) to visualize the digital tissue expression profiles for target genes. When drawing the heat map, three genes were dropped because there was no expression information for them. At last, the heat map of tissue expressions for 19 annotated genes from different tissues was presented in Figure 3. The heat map revealed that most of these genes have been expressed in reproductive and urinary system, nervous system, and digestive system, and many expressions were detected in fetus. The tissue expression profiles revealed that the identified genes are intuitively related to the physiological processes contributing to piglet mortality, such as embryo development.

Table 3. The position information of 21 SNPs and corresponding 22 genes.

Parity	SNP	Chrom	Position	Gene	Description
I	WU_10.2_2_133608994	2	128365728	***ZNF608***	zinc finger protein 608
	WU_10.2_6_14929389	6	15567222	***ZFHX3***	zinc finger homeobox 3
	ALGA0121819	8	77276189	***GATB***	glutamyl-tRNA amidotransferase subunit B
	WU_10.2_9_47540573	9	42434488	***CADM1***	cell adhesion molecule 1
	WU_10.2_10_4318367	10	2665638	***BRINP3***	BMP/retinoic acid inducible neural specific 3
	WU_10.2_X_7708900	23	7313289	***MID1***	midline 1
II	MARC0052132	3	14585248	***AUTS2***	autism susceptibility candidate 2
	WU_10.2_7_6945588	7	6732442	***MIR9802***	let-7/miR-98 family members are expressed late in mammalian embryonic development
	ALGA0048798	8	99466642	***FAT4***	FAT atypical cadherin 4
	WU_10.2_8_135384225	8	126266286	***GRID2***	glutamate ionotropic receptor delta type subunit 2
	ALGA0090390	16	37352844	***GAPT***	GRB2 binding adaptor protein, transmembrane

III	DRGA0001119	1	67860459	***ASCC3***	activating signal cointegrator 1 complex subunit 3
	Affx-115201707	1	79966051	***HS3ST5***	heparan sulfate-glucosamine 3-sulfotransferase 5
	MARC0113660	8	122083319	***STPG2***	sperm tail PG-rich repeat containing 2
	DRGA0008818	8	122141343		
	ALGA0060358	11	4274441	***CDK8***	cyclin dependent kinase 8
				WASF3	wiskott-Aldrich syndrome protein family member 3
	ASGA0049501	11	6324834	***MTUS2***	microtubule associated tumor suppressor candidate 2
	ASGA0055572	13	1323602	***NEK11***	NIMA (never in mitosis gene a)- related kinase 11
	ALGA0036320	6	108976054	***CABYR***	calcium binding tyrosine phosphorylation regulated
	H3GA0018655	6	109379857	***OSBPL1A***	oxysterol binding protein like 1A
	ASGA0029165	6	109452212	***IMPACT***	impact RWD domain protein
				HRH4	histamine receptor H4

250

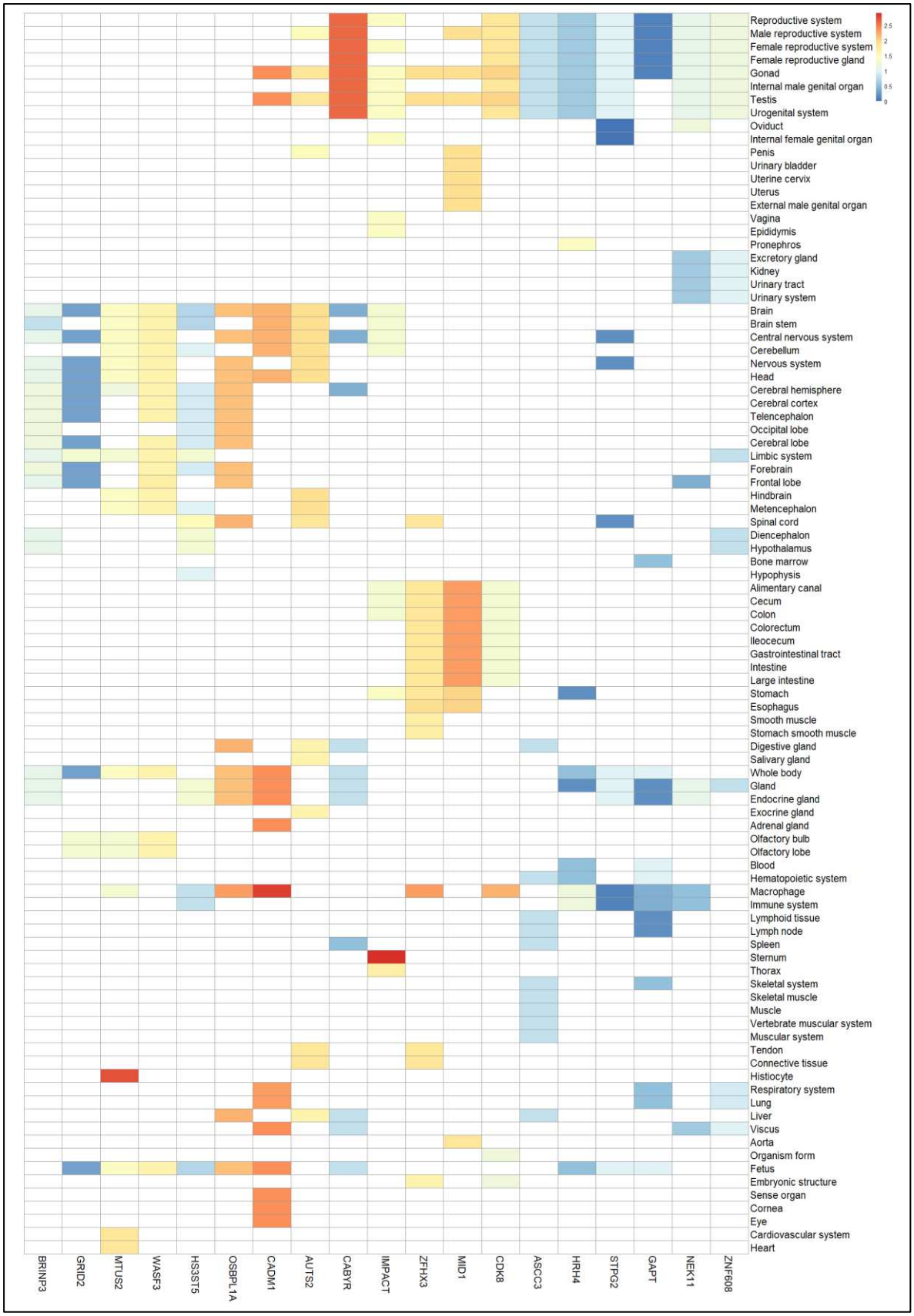

Figure 4. Digital tissue expression profiles of 19 candidate genes.

4. Discussion

In the pig industry, piglet mortality is intensively related to sow prolificacy, which is generally defined as the number of piglets weaned per sow per year (PSY). It is of high importance to characterize the genetic properties of piglet mortality. In this study, the piglet mortality was defined as a ratio trait that was re-constructed by its component traits, which needed simultaneous consideration of multiple component traits. In fact, there are many ratio traits in pigs, such as feed efficiency, lean percentage, and growth rate, which do not follow the normal distribution (32). For piglet mortality, the results displayed a heavily skewed distribution for its phenotypes, and this trait did not follow the normal distribution. Being different from the phenotype, the breeding value reveals the individual's genetic merit while removing environmental effects. Considering the distribution of estimated breeding values exhibited relatively little skewness, it could be inferred that the non-genetic factors might be the main determinant of heavy skewness for phenotypic distribution.

Based on the multi-breed approach, piglet mortality was estimated as a low heritability trait, with the estimations of 0.0630, 0.1031 and 0.1140 from parity I to III, respectively. There was an interesting appearance that with increasing parity, the heritability for piglet mortality also rises. We guess that during the first three parities, there is a growing maturation for reproductive organ of gilt (33), which assumes that the more maturation of reproductive organ, the lesser level of environmental factors disturbing on the fetus is happened, indicating a lower environmental variance component and thus a higher heritability estimation. To our knowledge, this was the first report to reveal the genetic property of piglet mortality at birth. Being different from the piglet mortality at weaning, the mortality at birth has no environmental effects during the lactation period, and more accurately reflects the genetic impact on the mortality of fetus. In similar studies from mortality at weaning, the estimations of heritability of piglet mortality were found to range from 0.03 to 0.17 (3). Although slight differences occurred between the similar literatures that reported the heritability estimation of piglet mortality-related traits (34, 35), besides our estimations, the general conclusion of low heritability for piglet mortality could be supported. For the cross-trait phenotypic and genetic relationships, the results revealed that all phenotypic and genetic correlations between piglet mortality and its component traits were estimated to be positive. Accumulated practical experiences from selection experiments showed that selection on litter size also increases the number born dead, and the increasing magnitude could be up to half of the improvement of litter size (36). Obviously, both phenotypic and genetic relationships supported the empirical conclusion from the practical selection experiments. It indicates that the unfavorable genetic correlation produces a huge challenge for simultaneously improving mortality and its component traits in breeding practice.

In this work, we proposed a combined approach to increase the detection credibility in GWAS. In the pipeline, the standard MLM, FarmCPU, BLINK, mrMLM were simultaneously utilized to identify the putative SNPs, and the permutation test was followed to statistically confirm the validity of putative SNPs that were detected at least by two tools. It is known that FarmCPU, BLINK, mrMLM are multi-locus GWAS tools with higher detection power than single-locus scan tools (26-28). There is a growing consensus that the commonly used cutoff for Bonferroni-adjusted p-values is too conservative and stringent (29), and may miss those true SNPs with medium effect size. Alternatively, the combined GWAS approach focuses on intersecting the top SNPs identified by the single tools with following a permutation test procedure, which can decrease the false negatives, and improve the detections of SNPs with medium effect size. It would be reasonable that the combined approach based on the multi-locus tools can provide

more reliable results. Following the pipeline, in total, we identified 21 SNPs that passed the combined test for three parities, and it can be found that the list of the identified SNPs for each parity had no overlapping. The results indicated that, in accordance with the different results of heritability estimation for different parities, piglet mortality in different parities had a different genetic architecture, which was consistent with the study of Onteru et al (15), and it is sound to take the same trait in different parities as different traits.

It was highlighted that several SNPs were identified in a region between the chromosome coordinates 77.2 and 126.4 Mb on SSC 8 for parity III. Recently, two studies reported that in the regions from 107.0 to 113.3 Mb and from 144.9 to 145.5 Mb on SSC 8, there were candidate haplotypes with statistically significant effects on TNB and stillborn (37, 38). Considering TNB and stillborn are the component traits of piglet mortality, the partial region overlap supported the validity of identified SNPs on SSC 8. In addition, the validity of the identification of SNPs in this study could be supported by the digital tissue expression profiles for 19 annotated genes from different tissues in TISSUES database. In the tissue expression profiles, according to a high confidence score, most of the annotated genes were found to be expressed in the reproductive and urinary system, nervous system, digestive system, and fetus. These tissues are intensively related to piglet mortality. For example, Otten at el (2000) reported that the prenatal stress during late gestation could result in high mortality and low birth weights for piglets (39). It can confirm that the digestive system determines the efficiency of nutrition intake during pregnancy (40). The genes expressed in the digestive system are closely involved in fetal development and piglet mortality. Although FAT4, GATB, and MIR9802 were not displayed in the heat map of tissue expression profiles, their functions were also potentially involved in the piglet mortality. For example, FAT4 has been proven to regulate the apical plasma membrane organization in the embryonic cerebral cortex for mammalian, indicating the role in embryonic development (41). It is easy to infer that the identified genes are participating in the reproductive, digestive and nervous regulation and embryo development, and then contribute to piglet mortality (42).

5. Conclusions

In brief, piglet mortality at birth was found a low heritability trait. All phenotypic and genetic correlations between piglet mortality and its component traits were estimated to be positive. Integrating the results from standard MLM, FarmCPU, BLINK, and mrMLM, we identified 21 loci and 22 genes associated with piglet mortality. Most of these genes were annotated to be expressed in the reproductive system, nervous system, digestive system, and embryonic development, which are reasonably related to piglet losses. This study advances our understanding of the genetic and genomic fundamentals of piglet mortality and also provides candidate genes that could be potentially used for pig breeding programs, genomic selection, and further investigations.

6. Abbreviations

GWAS: Genome-wide association study
MLM: mixed linear model
FarmCPU: Fixed and random model Circulating Probability Unification
BLINK: Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway
mrMLM: Multi-Locus Random-SNP-Effect Mixed Linear Model
SNP: Single-nucleotide polymorphism
LD: linkage disequilibrium

GEBVs: genomic estimated breeding values
GLM: generalized linear model
TND: total number dead
TNB: total number born
NS: number of stillborn
NM: number of mummified
EBV: estimated breeding value
AI-REML: The average information restricted maximum likelihood
QC: quality control
PSY: piglets weaned per sow per year

7. Declarations

7.1.Ethics approval and consent to participate

All experiments in this study were performed according to the guidelines of the Key Lab of Agriculture Animal Genetics, Breeding, and Reproduction of Ministry of Education, Animal Care and Use Committee, Wuhan, China (permit HZAUSW2015-0003).

7.2.Consent for publication

All authors have approved the manuscript for submission.

7.3.Availability of data and materials

Authors do not wish to share the data due to the propriety nature of the data.

7.4.Competing interests

no any potential competing interests in the paper

7.5.Funding

This research was funded by the Natural Science Foundation of China (31672392), and Earmarked Fund for China Agriculture Research System (CARS-35).

7.6.Authors' contributions

Writing wrote the codes, did the data analysis and editing, Meijing An; funding acquisition, Mengjin Zhu; writing—review and editing, Tao Xiang, Guangliang Zhou, Yunlong Ma, Xiaolei Liu, and Shuhong Zhao; methodology, Mengjin Zhu.

7.7.Acknowledgements

We gratefully acknowledge that Guanjun Ma, Dengdeng Ye, Qing Zhang, Quanshun Mei, Guangliang Zhou, Yilong Chen, Jiuhong Nan, Mengfen Wang, Ali Liu, Shuailei Yan, Xiao Peng, Haoyuan Zhang, Hailong Liu, Yinlong Liao, Xiaolong Qi, and Binxu Yin for help on collecting sample tissues and extracting DNA.

8. References

1. Crooks A, Hurd H, Dargatz D, Hill G. Economic Cost of Piglet Mortality in the farrowing to weaning phase: A report of the NAHMS National Swine Survey. Proc Ann Mtg AASP. 1992:483-7.
2. Hermesch S, Luxford B, Graser H, editors. Genetic parameters for piglet mortality, within litter variation of birth weight, litter size and litter birth weight. Proc Assoc Advmt Anim Breed Genet; 2001.
3. Varona L, Sorensen D. Joint analysis of binomial and continuous traits with a recursive model: a case study using mortality and litter size of pigs. Genetics. 2014;196(3):643-51.
4. Knecht D, Środoń S, Duziński K. The impact of season, parity and breed on selected reproductive performance parameters of sows. Archives Animal Breeding. 2015;58(1):49-56.
5. Ocepek M, Newberry RC, Andersen IL. Trade-offs between litter size and offspring fitness in domestic pigs subjected to different genetic selection pressures. Appl Anim Behav Sci. 2017;193:7-14.
6. Bakoev S, Getmantseva L, Bakoev F, Kolosova M, Gabova V, Kolosov A, et al. Survey of SNPs Associated with Total Number Born and Total Number Born Alive in Pig. Genes (Basel). 2020;11(5).
7. Andersson H, Lillpers K, Rydhmer L, Forsberg M. Influence of light environment and photoperiod on plasma melatonin and cortisol profiles in young domestic boars, comparing two commercial melatonin assays. Domest Anim Endocrinol. 2000;19(4):261-74.
8. Guo X, Su G, Christensen OF, Janss L, Lund MS. Genome-wide association analyses using a Bayesian approach for litter size and piglet mortality in Danish Landrace and Yorkshire pigs. BMC Genomics. 2016;17:468.
9. Uimari P, Sironen A, Sevon-Aimonen ML. Whole-genome SNP association analysis of reproduction traits in the Finnish Landrace pig breed. Genet Sel Evol. 2011;43:42.
10. Alves K, Schenkel FS, Brito LF, Robinson A. Estimation of direct and maternal genetic parameters for individual birth weight, weaning weight, and probe weight in Yorkshire and Landrace pigs. J Anim Sci. 2018;96(7):2567-78.
11. Langendijk P, Plush K. Parturition and Its Relationship with Stillbirths and Asphyxiated Piglets. Animals (Basel). 2019;9(11).
12. Roehe R, Kennedy B. Estimation of genetic parameters for litter size in Canadian Yorkshire and Landrace swine with each parity of farrowing treated as a different trait. Journal of animal science. 1995;73(10):2959-70.
13. Irgang R, Favero JA, Kennedy BW. Genetic parameters for litter size of different parities in Duroc, Landrace, and large white sows. J Anim Sci. 1994;72(9):2237-46.
14. Suárez M, Hermesch S, Braun JA, Graser H-U, editors. Estimates of genetic parameters for reproductive traits at different parities in Australian hyperprolific Large White sows. Association for the Advancement of Animal Breeding and Genetics; 2005.
15. Onteru SK, Fan B, Du ZQ, Garrick DJ, Stalder KJ, Rothschild MF. A whole-genome association study for pig reproductive traits. Anim Genet. 2012;43(1):18-26.
16. de Roos AP, Hayes BJ, Goddard ME. Reliability of genomic predictions across multiple populations. Genetics. 2009;183(4):1545-53.
17. Vanderick S, Gillon A, Glorieux G, Mayeres P, Mota RR, Gengler N. Usefulness of multi-breed models in genetic evaluation of direct and maternal calving ease in Holstein and Belgian Blue Walloon purebreds and crossbreds. Livest Sci. 2017;198:129-37.
18. Raven LA, Cocks BG, Hayes BJ. Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. BMC Genomics. 2014;15:62.
19. Antunovic B, Baban M, Dobranic V, Margeta V, Mijic P, Njari B, et al. Influence of housing systems on stillbirth and mortality rate in preweaning pigs farrowed by different gilt breeds. Ital J Anim Sci. 2009;8:193-5.
20. Vaillancourt JP, Tubbs RC. Preweaning mortality. Vet Clin North Am Food Anim Pract. 1992;8(3):685-706.

21. Ramos AM, Crooijmans RP, Affara NA, Amaral AJ, Archibald AL, Beever JE, et al. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. PLoS One. 2009;4(8):e6524.
22. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7.
23. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am J Hum Genet. 2009;84(2):210-23.
24. Yin L, Zhang H, Tang Z, Xu J, Yin D, Zhang Z, et al. rMVP: A Memory-efficient, Visualization-enhanced, and Parallel-accelerated tool for Genome-Wide Association Study. bioRxiv. 2020:2020.08.20.258491.
25. Yuan X, Zhang J, Wang Y. Probability theory-based SNP association study method for identifying susceptibility loci and genetic disease models in human case-control data. IEEE transactions on nanobioscience. 2010;9(4):232-41.
26. Liu X, Huang M, Fan B, Buckler ES, Zhang Z. Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. PLoS Genet. 2016;12(2):e1005767.
27. Huang M, Liu XL, Zhou Y, Summers RM, Zhang ZW. BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. Gigascience. 2019;8(2).
28. Wang SB, Feng JY, Ren WL, Huang B, Zhou L, Wen YJ, et al. Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. Sci Rep. 2016;6:19444.
29. Kuo KHM. Multiple Testing in the Context of Gene Discovery in Sickle Cell Disease Using Genome-Wide Association Studies. Genomics Insights. 2017;10:1178631017721178.
30. Huffman JE. Examining the current standards for genetic discovery and replication in the era of mega-biobanks. Nat Commun. 2018;9(1):5054.
31. Palasca O, Santos A, Stolte C, Gorodkin J, Jensen LJ. TISSUES 2.0: an integrative web resource on mammalian tissue expression. Database (Oxford). 2018;2018.
32. Yang R, Jin T, Li W. Mapping genome-wide QTL of ratio traits with Bayesian shrinkage analysis for its component traits. Genetica. 2010;138(8):853-60.
33. Huang YH, Yang TS, Lee YP, Roan SW, Liu SH. Effects of sire breed on the subsequent reproductive performances of landrace sows. Asian-Australasian Journal of Animal Sciences. 2003;16(4):489-93.
34. Hellbrugge B, Tolle KH, Bennewitz J, Henze C, Presuhn U, Krieter J. Genetic aspects regarding piglet losses and the maternal behaviour of sows. Part 1. Genetic analysis of piglet mortality and fertility traits in pigs. Animal. 2008;2(9):1273-80.
35. Arango J, Misztal I, Tsuruta S, Culbertson M, Holl JW, Herring W. Genetic study of individual preweaning mortality and birth weight in Large White piglets using threshold-linear models. Livest Sci. 2006;101(1-3):208-18.
36. Varley MA, Varley MA, Varley MA, International CAB. The neonatal pig : development and survival. Wallingford: CAB International; 1995. ix, 342 pages : illustrations ; 25 cm p.
37. Haggman J, Uimari P. Novel harmful recessive haplotypes for reproductive traits in pigs. J Anim Breed Genet. 2017;134(2):129-35.
38. Zhang C, MacNeil MD, Kemp RA, Dyck MK, Plastow GS. Putative Loci Causing Early Embryonic Mortality in Duroc Swine. Front Genet. 2018;9:655.
39. Otten W, Kanitz E, Tuchscherer M. Prenatal stress in pigs: effects on growth, physiological stress reactions and immune function. Arch Tierzucht. 2000;43:159-64.
40. Johnston L, Shurson J, Whitney M. Nutritional effects on fetal imprinting in swine. 2008.
41. Ishiuchi T, Misaki K, Yonemura S, Takeichi M, Tanoue T. Mammalian Fat and Dachsous cadherins regulate

apical membrane organization in the embryonic cerebral cortex. J Cell Biol. 2009;185(6):959-67.
42. Glock XT, Bilkei G. The effect of postparturient urogenital diseases on the lifetime reproductive performance of sows. Can Vet J. 2005;46(12):1103-7.

Additional Files

1. Additional Tables

Additional Table 1. Structure of raw data sets

	spring	summer	autumn	winter	Duroc	Yorkshine	Landrace	Total
Parity I	1801	1267	1288	1717	581	3197	2295	6073
Parity II	1442	1180	1643	1150	323	3132	1960	5415
Parity III	1300	655	1094	1329	180	2626	1572	4378

Additional Table 2. Heritability of NS, NM, TNB and TND in different parities (SE, standard error in brackets).

Traits	Parity I	Parity II	Parity III
NS	0.0720 (0.0175)	0.0880 (0.0191)	0.1534 (0.0269)
NM	0.0436 (0.0192)	0.0417 (0.0213)	0.0324 (0.0191)
TNB	0.2771 (0.0274)	0.2780 (0.0307)	0.3073 (0.0349)
TND	0.0851 (0.0285)	0.1159 (0.0243)	0.1207 (0.0276)

Additional Table 3. The SNP symbols, GWAS p-values, and permutated p-values of the identified SNPs.

Identified SNP	Permutated p-values	GWAS p-values
WU_10.2_2_133608994	0.0001	1.73E-07
WU_10.2_6_14929389	0.0001	1.17E-06
ALGA0121819	0.0001	6.75E-10
WU_10.2_9_47540573	0.0001	7.46E-06
WU_10.2_10_4318367	0.0002	1.10E-05
WU_10.2_X_7708900	0.0001	3.47E-07
MARC0052132	0.0001	2.44E-07
WU_10.2_7_6945588	0.0001	5.67E-07
ALGA0048798	0.0001	1.23E-07
WU_10.2_8_135384225	0.0001	1.20E-05
ALGA0090390	0.0002	7.25E-05
DRGA0001119	0.0001	4.96E-05
Affx-115201707	0.0002	2.00E-05
ALGA0036320	0.0001	8.39E-06
H3GA0018655	0.0001	1.33E-05
ASGA0029165	0.0001	1.50E-05

MARC0113660	0.0001	4.76E-05
DRGA0008818	0.0001	4.76E-05
ALGA0060358	0.0001	1.04E-08
ASGA0049501	0.0001	2.60E-05
ASGA0055572	0.0001	9.25E-06

Note: For the p-values in the table, they were the smallest ones among the p-values from different tools.

Additional Table 4. Descriptive statistics of phenotype (SD, Standard Deviation; C.V, Coefficient of Variation)

	Mean	**Min**	**Max**	**SD**	**C.V**
Parity I	0.079385972	0	1	0.134816603	1.698242139
Parity II	0.073242952	0	1	0.12364918	1.688205833
Parity III	0.068735823	0	1	0.111656107	1.624423818

Additional Table 5. The top ten SNPs identified by the standard MLM, FarmCPU, BLINK, and mrMLM for piglet mortality from Parity I-III

Parity I							
MLM		FarmCPU		BLINK		mrMLM	
SNP	p-value	SNP	p-value	SNP	p-value	SNP	p-value
WU_10.2_2_133608994	1.04E-06	WU_10.2_2_1660824	2.99E-09	ALGA0121819	6.75E-10	WU_10.2_1_303761600	6.48336E-08
ALGA0121819	2.94E-06	WU_10.2_6_14929389	1.17E-06	ASGA0054479	2.06E-08	WU_10.2_15_141565227	1.78E-06
WU_10.2_X_7708900	4.08E-06	MARC0045581	5.24E-06	WU_10.2_2_133608994	1.73E-07	Affx-114978229	4.93E-06
MARC0008576	1.66E-05	ALGA0121819	5.80E-06	WU_10.2_X_7708900	3.47E-07	Affx-114704997	4.93E-06
ALGA0004972	2.15E-05	WU_10.2_X_136142589	6.16E-06	WU_10.2_8_5737649	3.20E-05	Affx-114689732	1.74E-05
MARC0078678	2.74E-05	WU_10.2_9_47540573	7.46E-06	WU_10.2_9_126944287	4.61E-04	MARC0050498	2.51E-05
WU_10.2_X_124874052	3.51E-05	MARC0022036	7.65E-06	M1GA0004763	4.91E-04	DRGA0002303	3.72E-05
Affx-114980032	3.64E-05	ALGA0049751	8.23E-06	WU_10.2_10_4318367	5.06E-04	ALGA0009074	3.72E-05
ALGA0007070	4.16E-05	H3GA0009377	9.83E-06	WU_10.2_9_47540573	5.49E-04	H3GA0049299	5.11E-05
ASGA0063018	5.87E-05	WU_10.2_10_4318367	1.10E-05	WU_10.2_6_14929389	5.59E-04	ASGA0095137	5.49E-05
Parity II							
MLM		FarmCPU		BLINK		mrMLM	
SNP	p-value	SNP	p-value	SNP	p-value	SNP	p-value
WU_10.2_8_135384225	1.20E-05	WU_10.2_3_124682588	6.57E-11	DRGA0004397	1.04E-07	ALGA0052861	1.61E-07
WU_10.2_7_6945588	3.40E-05	WU_10.2_7_8690791	9.78E-08	ALGA0048798	1.23E-07	M1GA0012952	5.93E-07
ALGA0048798	3.53E-05	WU_10.2_15_118972045	2.62E-07	ALGA0006121	2.18E-07	ASGA0080428	1.87E-05

MARC0052132	3.92E-05	WU_10.2_18_7356014	3.90E-06	MARC0052132	2.44E-07	MARC0055383	2.42E-05
WU_10.2_8_135464992	6.44E-05	H3GA0046969	1.81E-05	WU_10.2_7_6945588	5.67E-07	MARC0066282	3.80E-05
WU_10.2_4_3739598	6.63E-05	WU_10.2_13_30474727	3.27E-05	WU_10.2_11_82101963	1.10E-04	WU_10.2_6_116413991	6.84E-05
ALGA0090390	7.25E-05	ASGA0007897	3.72E-05	ALGA0090390	2.50E-04	WU_10.2_6_132959422	7.17E-05
WU_10.2_8_104296869	7.77E-05	ASGA0009531	3.75E-05	MARC0010165	3.49E-04	ALGA0001726	1.31E-04
ALGA0114423	7.88E-05	WU_10.2_8_135384225	4.04E-05	ASGA0052141	5.07E-04	ALGA0097061	1.86E-04
MARC0089453	8.09E-05	MARC0070175	5.31E-05	ALGA0052390	5.88E-04	ALGA0097066	1.86E-04

Parity III

MLM		FarmCPU		BLINK		mrMLM	
SNP	p-value	SNP	p-value	SNP	p-value	SNP	p-value
ALGA0060358	9.92E-07	WU_10.2_10_32348084	6.16E-17	ALGA0122208	7.6E-18	ALGA0060358	1.04E-06
ALGA0036320	8.39E-06	WU_10.2_X_7317072	4.47E-10	WU_10.2_X_139666324	3.9E-14	ALGA0036320	9.42E-06
ASGA0055572	9.25E-06	WU_10.2_3_127030944	4.89E-09	ALGA0049681	1.4E-11	ASGA0055572	1.06E-05
H3GA0018655	1.33E-05	ALGA0018083	5.7E-09	WU_10.2_14_139115957	3.36E-11	H3GA0018655	1.46E-05
ASGA0029165	1.5E-05	ALGA0095726	6.96E-09	WU_10.2_15_21804958	1.15E-10	ASGA0029165	1.66E-05
ASGA0049501	2.6E-05	WU_10.2_15_20997846	7.29E-09	WU_10.2_X_34786795	1.11E-09	Affx-115201707	2.00E-05
Affx-11520170 7	3.32E-05	ALGA0060358	1.04E-08	WU_10.2_1_47087424	1.44E-09	ASGA0049501	2.64E-05
MARC0113660	4.76E-05	WU_10.2_X_37383826	1.8E-08	WU_10.2_5_85471520	2.68E-09	DRGA0001119	4.96E-05
DRGA0008818	4.76E-05	H3GA0042609	2.81E-08	WU_10.2_7_4777306	5.96E-09	MARC0113660	5.68E-05
DRGA0001119	6.24E-05	MARC0002720	4.51E-07	ASGA0100851	1.21E-08	DRGA0008818	5.68E-05

Note: For the p-values in the table, they were the smallest ones among the p-values from different tools.

Figures

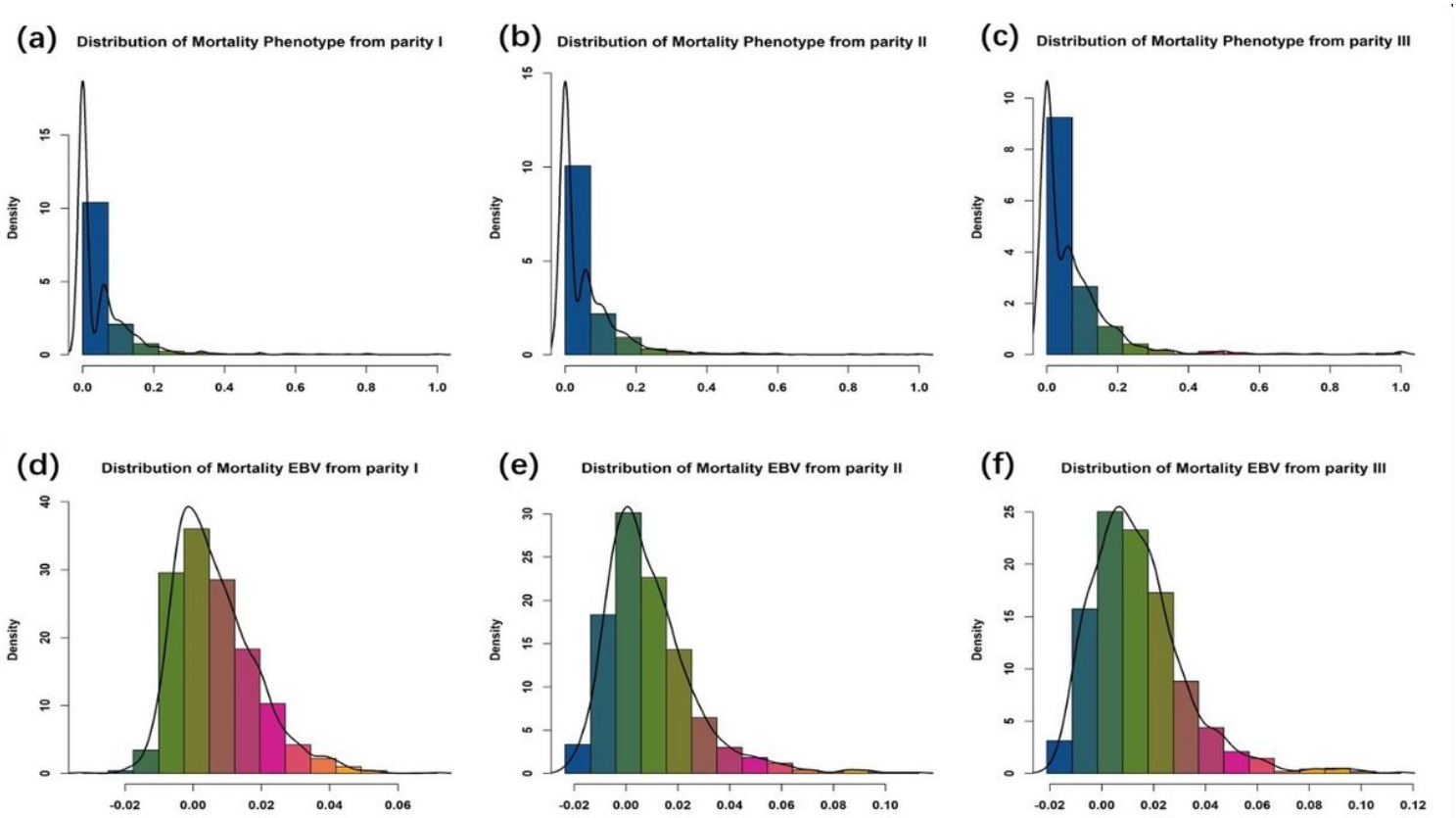

Figure 1

The phenotypic and EBV's distributions of piglet mortality from the three parities. Sub-figures from a to c represented the phenotypic distributions from parity I to parity III, and sub-figures from d to f represented the distributions of EBVs from parity I to III, respectively.

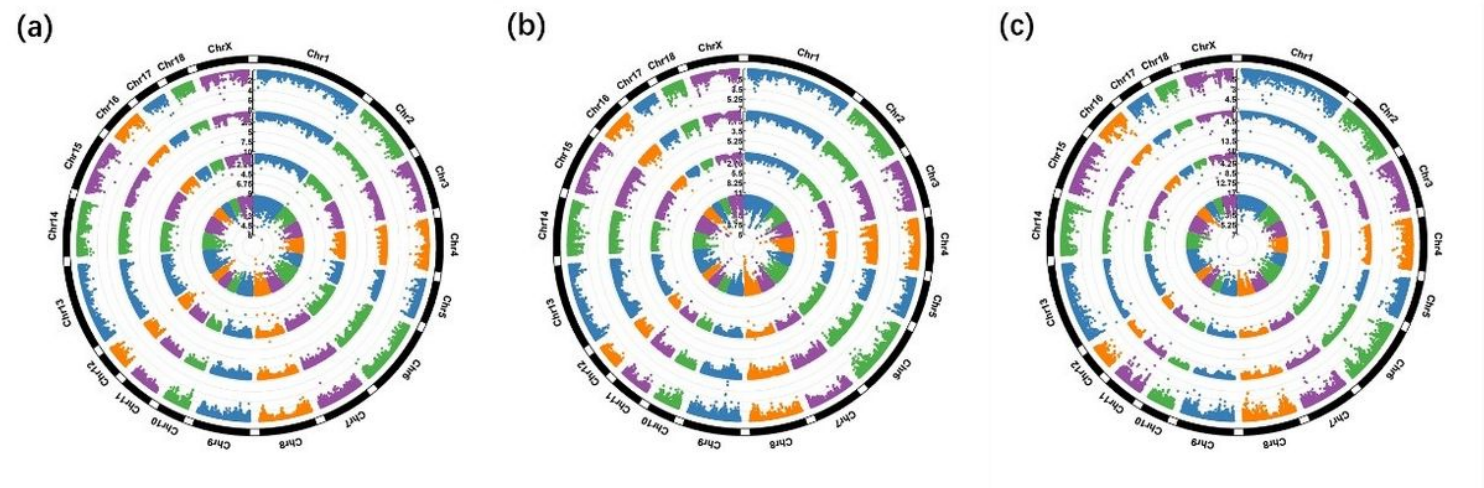

Figure 2

Circular-Manhattan of piglet mortality of different parities. a. Circular-Manhattan plots of MLM, FarmCPU, BLINK, mrMLM for parity I; b. Circular-Manhattan plots of MLM, FarmCPU, BLINK, mrMLM for parity II; c. Circular-Manhattan plots of MLM, FarmCPU, BLINK, mrMLM for parity III; and from inner to outer, they were MLM, FramCPU, BLINK and mrMLM, respectively.

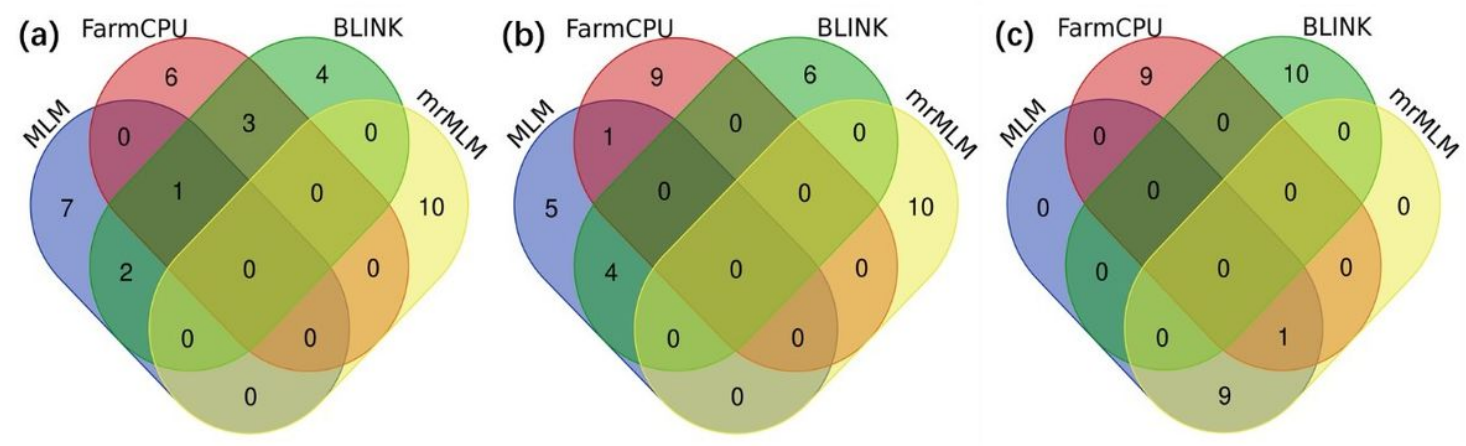

Figure 3

The Venn diagrams of identified SNPs in piglet mortality from parity I, II, III a. Venn diagrams of MLM, FarmCPU, BLINK, mrMLM method for parity I. There were 6 SNPs intersected by at least two tools; b. Venn diagrams of MLM, FarmCPU, BLINK, mrMLM method for parity II. There were 5 SNPs intersected by two tools; c. Venn diagrams of MLM, FarmCPU, BLINK, mrMLM method for parity III. There were 10 SNPs intersected by two tools.

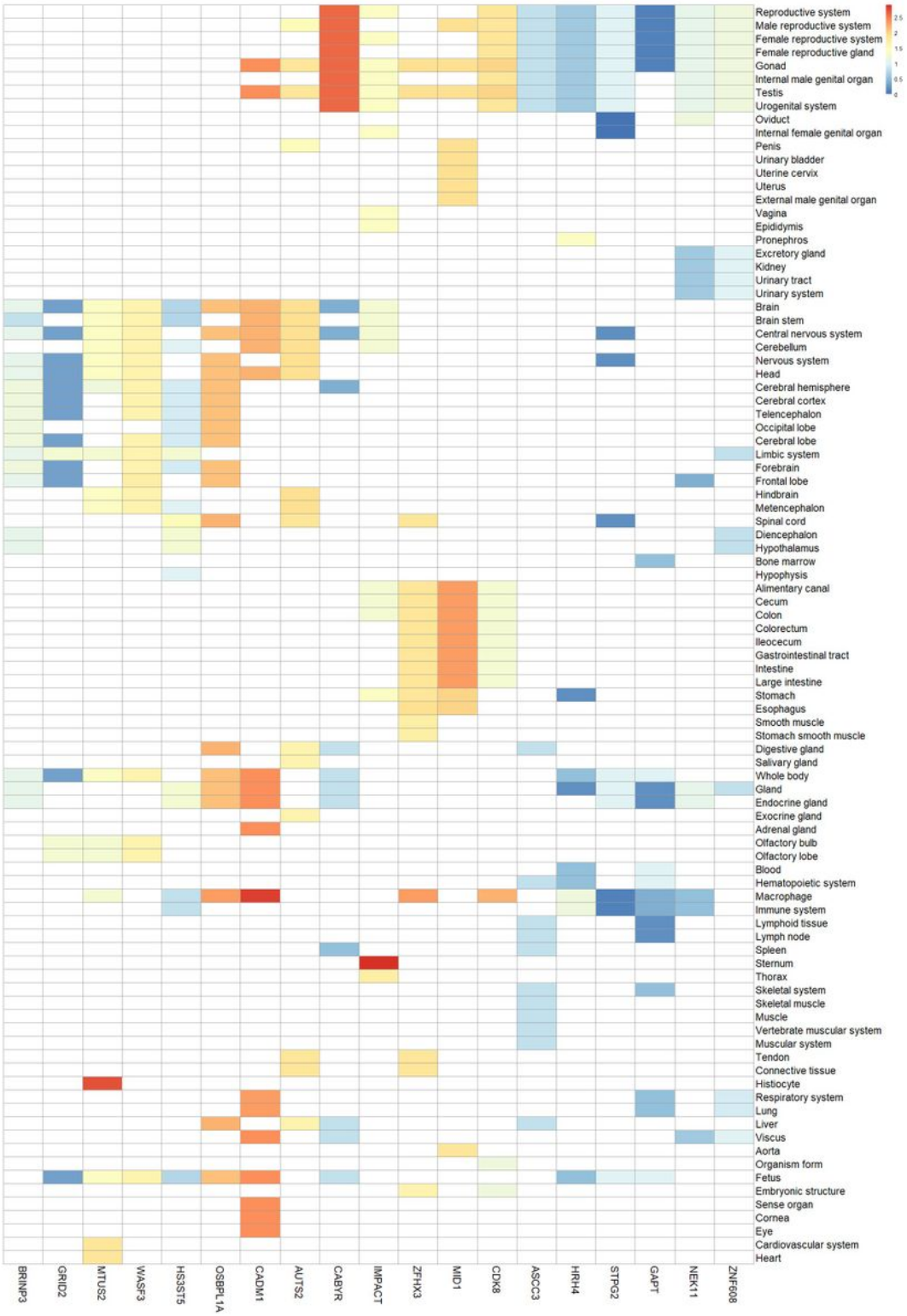

Figure 4

Digital tissue expression profiles of 19 candidate genes.