A genome-wide association study (GWAS) for pH value in the meat of Berkshire pigs

Jun Park#, Sang-Min Lee#, Ja-Yeon Park and Chong-Sam Na*

Department of Animal Biotechnology, Jeonbuk National University, Jeonju 54896, Korea

Abstract

The purpose of this study is to estimate the single nucleotide polymorphism (SNP) effect for pH values affecting Berkshire meat quality. A total of 39,603 SNPs from 1,978 heads after quality control and 882 pH values were used estimate SNP effect by single step genomic best linear unbiased prediction (ssGBLUP) method. The average physical distance between adjacent SNP pairs was 61.7kbp and the number and proportion of SNPs whose minor allele frequency was below 10% were 9,573 and 24.2%, respectively. The average of observed heterozygosity and polymorphic information content was 0.32 ± 0.16 and 0.26 ± 0.11, respectively and the estimate for average linkage disequilibrium was 0.40. The heritability of pH45m and pH24h were 0.10 and 0.15 respectively. SNPs with an absolute value more than 4 standard deviations from the mean were selected as threshold markers, among the selected SNPs, protein-coding genes of pH45m and pH24h were detected in 6 and 4 SNPs, respectively. The distribution of coding genes <RFX8, CREG2, TBC1D8, EXOC6B> were detected at pH45m and <C12orf49, LOC106506010, BICC1, ANK3> were detected at pH24h.

Keywords: Berkshire, pH value, Heritability, Linkage disequilibrium, Genome-wide association study

INTRODUCTION

pH value in the meat of pigs is associated with meat color, drip loss and moisture holding capacity; as pH increases, drip loss and cooking loss decrease, but moisture holding capacity increases, thereby affecting meat quality [1]. Meat pH value in pigs has been a subject of extensive research [2–5] it is well known that changes in pH after slaughter are of great importance in protein denaturation and drip loss in post mortem changes in muscle [6].

Investigating genetic characteristics related to meat pH value in domestic pigs is currently a priority because it involves basic stages in genomic selection using accumulated genomic information. If a reference population that increases selection accuracy can be established in the future, it would be possible to improve the accuracy of estimate breeding values for the various selection traits, by using genomic information and phenotypic data in addition to pedigree information [7].

Through a genome-wide association study (GWAS) among populations using linkage disequilibrium (LD), it is possible to detect the relationships between single nucleotide polymorphisms (SNPs) that affect economic traits, while quantitative trait loci (QTL) can be excavated and tagging
SNPs selected. Since its first used by [8], polymorphic information content (PIC) had become the most widely applied method for genetic studies to measure the information content of molecular markers. The PIC value of marker is defined as the expected fraction of informative offspring from pedigree [9]. Heterozygosity refers to the ratio of heterozygote in the locus, and observed heterozygosity (OHE) was investigated to analyze the actual heterozygosity degree of each marker for the population. Through the single-step genomic best linear unbiased prediction (ssGBLUP) using an algorithm that combines existing pedigree information with genomic information, expected breeding value (EBV) and genomic expected breeding value (GEBV) can be estimated simultaneously, while through the Back solution, SNPs effects can be estimated [10,11].

For the pigs, studies on the excavation of quantitative traits for economic traits have been performed continually, primarily using GWAS [12–14]. The present experiment was conducted to estimate the SNP effects that affect pH value in the meat of Berkshire pigs.

MATERIALS AND METHODS

SNP data and quality control
A total of 2,037 heads were genotyped using Porcine SNP60k v2 Beadchip (Illumina, San Diego, CA, USA) and 61,565 SNPs were collected. To ensure the quality of the genotypic data, the following SNP types were excluded from the analysis: SNP on sex chromosomes (1,458); SNP without information on chromosomes (7,849); SNP with missing rate higher than 10%; SNP without polymorphism (all homo or hetero); SNP with minor allele frequency lower than 1%; SNP with Hardy-Weinberg disequilibrium chi-squared value greater than 23.93 ($p < 10^{-6}$); and animals with SNP missing rate higher than 10%. Therefore, the number of animals and SNPs included after quality control was 1,978 and 39,603, respectively (Table 1).

pH value data
Data on pH values were collected from 882 heads of pigs slaughtered at Namwon Jeil Food between 2015 and 2018. A pH*21K meter (NWK-Binär GmbH., Buchloe, Germany) was inserted into the sirloin muscle and pH values were measured 3 times each at 45 min (pH45m) and at 24
hours (pH24h) after slaughter, and the average values were recorded.

**Statistical model**

**Variance Components:** Using the multiple traits animal model, we estimated the variance component and genetic parameters of the pH values, and the equation is as follows:

\[ y_{ijkl} = \mu + YM_i + s_j + a_k + e_{ijkl} \]

Where, \( y_{ijkl} \) is observed values of \( t \) th trait, \( \mu \) is overall mean, \( YM_i \) is the \( i \) th fixed effect of slaughter year-month, \( s_j \) is the \( j \) th fixed effect of sex, \( a_k \) is the \( k \) th additive genetic effect, \( e_{ijkl} \) is the residual random effects. We used VCE ver. 6.0 [15] to estimated variance components with pH45m and pH24h.

**Estimation of genomic breeding values:** The slaughter year-month (YM) and sex were included as fixed effects in a statistical analysis that used the following model:

\[ y_i = X_i b_i + Z_i a_i + e_i \]

Where, \( y_i \) is the vector of observation in \( i \) th trait, \( b_i \) is the vector of fixed effect, \( a_i \) is the vector of additive genetic random effect, \( e_i \) is the vector of residual effect, \( X \) and \( Z \) were known incidence matrix corresponding to \( b \) and \( a \) respectively. Mixed model equation was as follows:

\[
\begin{bmatrix}
X'X & X'Z \\
Z'X & Z'Z + \alpha H^{-1}
\end{bmatrix}
\begin{bmatrix}
b \\
n
\end{bmatrix}
= \begin{bmatrix}
X'y \\
Z'y
\end{bmatrix}
\]

Where, \( \alpha = \sigma^2_a / \sigma^2_e \), \( H^{-1} = A^{-1} + 0.0 \), \( G^{-1} = A^{-1} \), is the inverse matrix of numerator relationship matrix, \( G^{-1} \) is the inverse matrix of genomic relationship matrix, \( A^{-1} \) is the inverse matrix of numerator relationship matrix of pigs with genomic information.

**SNP effect:** The SNP effect of each marker was estimated through reverse operation method of GEBVs and equation was as follows:

\[ \hat{u} = \lambda DZ G^{-1} \hat{a}_g = DZ (ZDZ)^{-1} \hat{a}_g \]

Where, \( \hat{u} \) is the vector of SNP effect, \( \hat{a}_g \) is the vector of GEBV, \( Z \) is the coefficient matrix of SNP, \( D \) is the weighted vector [16,17]. We used BLUPF90 family program [18] to estimated SNP effects with pH45m and pH24h.

**RESULTS AND DISCUSSION**

**Genetic characteristics**

The average physical distance between adjacent SNP pairs was 61.7 kbp (Fig. 1). The number and proportion of SNPs whose minor allele frequency (MAF) was below 10% were 9,573 and 24.2%, respectively, while those for which it was higher than 40% were 7,343 and 18.5%. The number and proportion of SNPs for which the physical distance between adjacent SNP pairs was less than 1 kbp were 415 and 1.0%, respectively, while those for which it was in the range between 10 kbp and
Genome-wide association study for pH value the Berkshire pigs

100 kbp were 30,345 and 76.6%, respectively (Table 2).

OHE was $0.32 \pm 0.16$ on average, and in most chromosomes was within the range of 0.31–0.33; it was highest in chromosome 8 (0.36) and lowest in chromosome 15 (0.25). PIC was $0.26 \pm 0.11$ on average, within the range of 0.22–0.28 (Table 3).

Lee et al. [19] have reported that OHE estimates in Berkshire, Landrace, and Yorkshire were $0.33 \pm 0.15$, $0.36 \pm 0.14$ and $0.36 \pm 0.14$, respectively while estimates for the PIC were on average $0.26 \pm 0.11$, $0.28 \pm 0.10$ and $0.29 \pm 0.10$, respectively. Compared with the results from this experiment with our own, the OHE estimates obtained from our experiment were slightly lower but similar, while PIC estimates were similar.

The estimate for average LD ($r^2$) was 0.40, which was high in the order of chromosomes 1 (0.45), 14 (0.44), 15, and 18 (0.42) (Table 4).

Clearly markers with greater numbers of alleles tend to have higher PIC values and thus are more informative [9]. In the results of this study, it can be seen that the values of MAF, OHE, and PIC tend to be low. This seems to be a result of high immobilization of Berkshire breed used in the study.

**pH value**

Estimates for pH45m and pH24h values were on average 6.40 (± 0.20) and 5.90 (± 0.12), respectively (Table 5). Ryu et al. [20] have reported that in the Berkshire breed the average pH45m

### Table 2. Minor allele frequency (MAF) and the distance between adjacent SNP pairs (kbp) and percentage (%)

| Criteria | MAF | Frequency (%) | Criteria | Distance between adjacent SNP pairs (kbp) | Frequency (%) |
|----------|-----|---------------|----------|------------------------------------------|---------------|
| MAF < 0.1 | 9,573 (24.2) | ADAM < 1 | 415 (1.0) |
| 0.1 ≤ MAF < 0.2 | 7,787 (19.7) | 1 ≤ ADAM < 10 | 2,599 (6.6) |
| 0.2 ≤ MAF < 0.3 | 7,899 (19.9) | 10 ≤ ADAM < 100 | 3,0345 (76.6) |
| 0.3 ≤ MAF < 0.4 | 7,001 (17.7) | 100 ≤ ADAM < 1,000 | 6,217 (15.7) |
| 0.4 ≤ MAF | 7,343 (18.5) | 1,000 ≤ ADAM | 27 (0.1) |

SNP, single nucleotide polymorphism; ADAM, average distance between adjacent SNP pairs.
Table 3. The number of SNPs, means (± SD) of minor allele frequency (MAF), the observed heterozygosity (OHE) and the polymorphic information content (PIC) by chromosome.

| Chr. No. | SNPs | MAF     | OHE     | PIC     |
|----------|------|---------|---------|---------|
| 1        | 4,552| 0.23 ± 0.14 | 0.32 ± 0.15 | 0.26 ± 0.11 |
| 2        | 2,744| 0.23 ± 0.14 | 0.31 ± 0.16 | 0.25 ± 0.11 |
| 3        | 2,045| 0.23 ± 0.15 | 0.32 ± 0.16 | 0.25 ± 0.11 |
| 4        | 2,657| 0.25 ± 0.16 | 0.33 ± 0.16 | 0.26 ± 0.11 |
| 5        | 1,756| 0.22 ± 0.14 | 0.31 ± 0.16 | 0.25 ± 0.11 |
| 6        | 2,582| 0.26 ± 0.15 | 0.35 ± 0.15 | 0.27 ± 0.11 |
| 7        | 2,535| 0.25 ± 0.15 | 0.34 ± 0.16 | 0.26 ± 0.11 |
| 8        | 2,156| 0.27 ± 0.15 | 0.36 ± 0.16 | 0.28 ± 0.11 |
| 9        | 2,575| 0.23 ± 0.15 | 0.32 ± 0.16 | 0.25 ± 0.11 |
| 10       | 1,426| 0.22 ± 0.15 | 0.31 ± 0.17 | 0.24 ± 0.12 |
| 11       | 1,271| 0.22 ± 0.15 | 0.31 ± 0.17 | 0.24 ± 0.12 |
| 12       | 1,184| 0.24 ± 0.15 | 0.33 ± 0.16 | 0.26 ± 0.11 |
| 13       | 3,186| 0.24 ± 0.15 | 0.33 ± 0.16 | 0.26 ± 0.11 |
| 14       | 3,095| 0.23 ± 0.15 | 0.33 ± 0.16 | 0.25 ± 0.11 |
| 15       | 2,201| 0.20 ± 0.15 | 0.28 ± 0.17 | 0.22 ± 0.12 |
| 16       | 1,378| 0.23 ± 0.14 | 0.33 ± 0.15 | 0.26 ± 0.11 |
| 17       | 1,314| 0.23 ± 0.14 | 0.32 ± 0.15 | 0.25 ± 0.10 |
| 18       | 946  | 0.25 ± 0.15 | 0.34 ± 0.15 | 0.27 ± 0.11 |
| Overall  | 39,603| 0.24 ± 0.15 | 0.32 ± 0.16 | 0.26 ± 0.11 |

SNP: single nucleotide polymorphism.

Table 4. Linkage disequilibrium ($r^2$) by chromosome

| Chromosome no | No. of SNPs | No. of SNP pairs | Linkage disequilibrium ($r^2$) |   |   |
|---------------|-------------|------------------|---------------------------------|---|---|
|               |             |                  | All pairs                       | Adjacent pairs |
| 1             | 4,552       | 10,358,076       | 0.04                            | 0.45          |
| 2             | 2,744       | 3,763,396        | 0.03                            | 0.39          |
| 3             | 2,045       | 2,089,990        | 0.03                            | 0.41          |
| 4             | 2,657       | 3,528,496        | 0.03                            | 0.40          |
| 5             | 1,756       | 1,540,890        | 0.03                            | 0.37          |
| 6             | 2,582       | 3,332,071        | 0.03                            | 0.39          |
| 7             | 2,535       | 3,211,845        | 0.03                            | 0.40          |
| 8             | 2,156       | 2,323,090        | 0.04                            | 0.41          |
| 9             | 2,575       | 3,314,025        | 0.03                            | 0.37          |
| 10            | 1,426       | 1,016,025        | 0.02                            | 0.35          |
| 11            | 1,271       | 807,085          | 0.03                            | 0.39          |
| 12            | 1,184       | 700,336          | 0.03                            | 0.41          |
| 13            | 3,186       | 5,073,705        | 0.03                            | 0.40          |
| 14            | 3,095       | 4,787,965        | 0.04                            | 0.44          |
| 15            | 2,201       | 2,421,100        | 0.03                            | 0.42          |
| 16            | 1,378       | 948,753          | 0.03                            | 0.38          |
| 17            | 1,314       | 862,641          | 0.04                            | 0.40          |
| 18            | 946         | 446,985          | 0.04                            | 0.42          |
| Overall       | 39,603      | 50,526,474       | 0.03                            | 0.40          |

SNP: single nucleotide polymorphism.
estimates in females and males were 6.26 (± 0.08) and 6.15 (± 0.05), respectively, while for pH24h they were 5.61 (± 0.02) and 5.69 (± 0.01) for female and male pigs, respectively. The same study also has reported that pH45m estimates for in females and males were 5.80 (± 0.04) and 5.92 (± 0.05), respectively, in Landrace and 6.05 (± 0.02), 5.98 (± 0.03), respectively, in Yorkshire while for pH24h estimates were 5.49 (± 0.01) and 5.51 (± 0.02) in Landrace females and males and 5.56 (± 0.01), and 5.56 (± 0.01) in Yorkshire pigs. In view of these results, average pH values in the Berkshire breed appear to be higher than in other breeds.

Genetic parameters
The genetic parameters of pH45m and pH24h were estimated using VCE6.0 software. Heritability estimates for pH45m and pH24h were 0.10 and 0.15, respectively, while for the phenotypic and genetic correlation between 2 traits they were 0.05 and 0.27, respectively (Table 6). Lee et al. [21] have reported that in the domestic Berkshire breed heritability estimates for pH45m and pH24h were 0.48 and 0.15, respectively. When compare with pH45m and pH24h in this study, the heritability estimates for pH45m obtained from this experiment were lower, while those for the pH24h were similar.

Genome-wide association study
Estimates of SNP effects for pH45m and pH24h were within the ranges of −0.00011536 to 0.00011866 and −0.00009704 to 0.00009562, respectively (Fig. 2). Suitability for normal distribution of SNP effects was tested by 3 methods, but the SNP effects were not normally distributed (Table 7). Absolute values were taken for estimated SNP effects and when expressed as the gamma distribution, most values were near zero, with only a few SNPs having larger effects (Fig. 3).

After standardizing estimates for SNP effects, absolute values were taken and were expressed as a Manhattan plot compared to their relative sizes (Fig. 4).

In normal distribution, values within 3 standard deviations of the mean account for approximately 99.74% of the data set. Therefore, SNPs with an absolute value more than 4 standard deviations from the mean were selected as threshold markers. Although the number of SNPs with an absolute value of more than 4 standard deviations was 12 for pH45m and 23 for pH24h (Tables 8 and 9), no significant SNP effects were observed. This may be due to the small number of animals used in the analysis in relation to the number of SNPs.

| Trait | $\sigma^2$ | $\sigma^2_e$ | $h^2$ ± SE | Correlation $^1$ | 1 | 2 |
|-------|------------|------------|------------|----------------|-----|-----|
| pH45m | 0.0039 | 0.0351 | 0.10 ± 0.04 | - | 0.05 |
| pH24h | 0.0021 | 0.0017 | 0.15 ± 0.07 | 0.27 | - |

$^1$Upper triangle: phenotypic, lower triangle: genetic correlation

pH45m, pH value after 45 minutes; pH24h, pH value after 24 hours.
Fig. 2. The Manhattan plots of SNP effects for pH45m (top) and pH24h (bottom). SNP, single nucleotide polymorphism; pH45m, pH value after 45 minutes; pH24h, pH value after 24 hours.

Table 7. Goodness-of-fit tests for gamma distribution of estimated single nucleotide polymorphism effects for pH45m and pH24h

| Test method          | pH45m |          | pH24h |          |
|----------------------|-------|----------|-------|----------|
|                      | Statistics | p-value   | Statistics | p-value   |
| Kolmogorov-Smirnov (D) | 0.0438    | < 0.010  | 0.0470    | < 0.010  |
| Cramer-von Mises (W-Sq) | 24.2362    | < 0.005  | 26.4829   | < 0.005  |
| Anderson-Darling (A-Sq) | 131.9259   | < 0.005  | 142.2137  | < 0.005  |

pH45m, pH value after 45 minutes; pH24h, pH value after 24 hours.

Fig. 3. Distribution of the estimated SNP effects for pH45m and pH24h (from left to right). SNP, single nucleotide polymorphism; pH45m, pH value after 45 minutes; pH24h, pH value after 24 hours.
Six markers (MARC0113402, MARC0100326, MARC0065978, ALGA0115738, ASGA0015062, and ASGA0015063) that had similar effects nearby markers with significant effects due to LD were detected in chromosome 3 for pH45m, while for pH24h 9 markers were detected (ASGA0063851, MARC0002354, ALGA0078206, ALGA0078209, H3GA0040626, MARC0035949, ALGA0078332, ASGA0064086, and DRGA0013964) in chromosome 14.
whose SNPs were in the LD blocks.

Among the SNPs with an absolute value of more than 4 standard deviations, for pH45m and pH24h protein-coding genes were detected in 6 and 4 SNPs, respectively.

For distribution of protein coding genes, a total of 4 genes, RFX8 (MARC0113402), CREG2 (MARC0100326), TBC1D8 (MARC0065978), and EXOC6B (ALGA0115738, ASGA0015062, ASGA0015063) were detected in chromosome 3 for pH45m, while for pH24h C12orf49 (ALGA0076917), LOC106506010 (MARC0002354), BICC1 (MARC0035949) and ANK3 (DRGA0013964) were detected in chromosome 14.

Edwards et al. [22] have reported that QTL were discovered for 45-min pH and pH decline on SSC 3 and this QTL region affecting pH on SSC 3 was in a similar location to a pH QTL reported by [23]. When compare with this study, since the markers that we detected in the SSC 3 region are close to or belong to the previously investigated QTL region, it is necessary to study whether they can potentially affect the pH.

### REFERENCES

1. Huff-Lonergan E, Baas TJ, Malek M, Dekkers JCM, Prusa K, Rothschild MF. Correlations among selected pork quality traits. J Anim Sci. 2002;80:617-27. https://doi.org/10.2527/2002.803617x
2. Choi YI, Kim YT, Lee CL, Han IK. Carcass and pork quality characteristics by sex and marketing day. Korean J Anim Sci. 2000;42:933-40.

3. Lee J, Joo Y, Shin W, Cho K, Lee J, Lee J, et al. Comparison of carcass and pork physical characteristics by market weight and gender of Berkshire. Korean J Food Sci Anim Resour. 2004;24:108-14.

4. Hah KH, Kim IS, Jin SK, Nam YW, Cho JH. Proximate composition and physico-chemical characteristics of Berkshire pork by gender. Korean J Food Sci Anim Res. 2007;27:137-41. https://doi.org/10.5851/kosfa.2007.27.2.137

5. Park BY, Cho SH, Kim JH, Seong PN, Kang GH, Jeong D, et al. Comparison of pork quality by different Berkshire line. Korean J Food Sci Anim Res. 2010;30:867-71. https://doi.org/10.5851/kosfa.2010.30.5.867

6. Bendall J. Postmortem changes in muscle. In: Bourne GH, editor. The structure and function of muscle: structure. New York, NY: Academic Press; 1973. p. 243-309.

7. Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME. Invited review: genomic selection in dairy cattle: progress and challenges. J Dairy Sci. 2009;92:433-43. https://doi.org/10.3168/jds.2008-1646

8. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet. 1980;32:314-31.

9. Hildebrand CE, David C, Torney C, Wagner RP. Informativeness of polymorphic DNA markers. Los Alamos Sci. 1994;20:100-2.

10. Legarra A, Aguilar I, Misztal I. A relationship matrix including full pedigree and genomic information. J Dairy Sci. 2009;92:4656-63. https://doi.org/10.3168/jds.2009-2061

11. VanRaden PM. Efficient methods to compute genomic predictions. J Dairy Sci. 2008;91:4414-23. https://doi.org/10.3168/jds.2007-0980

12. Knol EF, Nielsen B, Knap PW. Genomic selection in commercial pig breeding. Anim Front. 2016;6:15-22. https://doi.org/10.2527/af.2016-0003

13. Puig-Oliveras A, Folch A, Ballester D. Genetic dissection of growth and meat quality traits in pigs. Barcelona, Spain: Universitat Autònoma de Barcelona; 2015.

14. Qiao R, Gao J, Zhang Z, Li L, Xie X, Fan Y, et al. Genome-wide association analyses reveal significant loci and strong candidate genes for growth and fatness traits in two pig populations. Genet Sel Evol. 2015;47:17. https://doi.org/10.1186/s12711-015-0089-5

15. Groeneveld E, Kovac M, Mielenz N. VCE user’s guide and reference manual, ver. 6.0. Neustadt, Germany: Institute of Farm Animal Genetics, Friedrich Loeffler Institute (FLI); 2008.

16. Christensen OF, Madsen P, Nielsen B, Ostersen T, Su G. Single-step methods for genomic evaluation in pigs. Animal. 2012;6:1565-71. https://doi.org/10.1017/s1751731112000742

17. Wang HY, Misztal I, Aguilar I, Legarra A, Fernando RL, Vitezica Z, et al. Genome-wide association mapping including phenotypes from relatives without genotypes in a single-step (ssGWAS) for 6-week body weight in broiler chickens. Front Genet. 2014;5:134. https://doi.org/10.3389/fgene.2014.00134

18. Misztal I, Tsuruta S, Lourenco D, Aguilar I, Legarra A, Vitezica Z. Manual for BLUPF90 family of programs. Athens, GA: University of Georgia; 2014.

19. Lee SM, Oh JD, Park KD, Do KT. Analysis of genetic characteristics of pig breeds using information on single nucleotide polymorphisms. Asian-Australas J Anim Sci. 2019;32:485-93. https://doi.org/10.5713/ajas.18.0304

20. Ryu YC, Choi YM, Lee SH, Shin HG, Choe JH, Kim JM, et al. Comparing the histochemical characteristics and meat quality traits of different pig breeds. Meat Sci. 2008;80:363-9. https://
21. Lee JH, Song KD, Lee HK, Cho KH, Park HC, Park KD. Genetic parameters of reproductive and meat quality traits in Korean berkshire pigs. Asian–Australas J Anim Sci. 2015;28:1388-93. https://doi.org/10.5713/ajas.15.0097

22. Edwards DB, Ernst CW, Raney NE, Doumit ME, Hoge MD, Bates RO. Quantitative trait locus mapping in an F2 Duroc × Pietrain resource population: II. carcass and meat quality traits. J Anim Sci. 2008;86:254-66. https://doi.org/10.2527/jas.2006-626

23. Beeckmann P, Schröffel J, Moser G, Bartenschlager H, Reiner G, Geldermann H. Linkage and QTL mapping for Sus scrofa chromosome 3. J Anim Breed Genet. 2003;120:20-7. https://doi.org/10.1046/j.0931-2668.2003.00420.x