Microbiome Contribute Significantly to Variation in Fat and Growth Traits in Crossbred Pigs?

D. Lu[1], F. Tiezzi[1], C. Schillebeeckx[2], N.P. McNulty[2], C. Schwab[3] & C. Maltecca[1]

[1]North Carolina State University, Raleigh, USA
bdlu@ncsu.edu (Corresponding Author)
[2]Matatu Inc., Saint Louis, MO, USA
[3]The Maschhoffs LLC., Carlyle, IL, USA

Summary

Microbial communities play an important role in biological processes in the host body, and might influence heritabilities (h^2) and accuracy of breeding values (BV) estimated for phenotypic measures on the host. The report herein was made by analyzing bacteria in fecal samples of 1131 pigs at 15 weeks after weaning. Heritability was estimated for back fat thickness at week 22, and average daily gain from weaning to week 22 post weaning, hereafter BF22 and ADG22, respectively. Heritability for BF22 was estimated to be 0.52±0.05 without microbiome information included in the model. This h^2 dropped to 0.31±0.06 when estimated from a model that included microbiome data and interaction between the host genotypes and microbiome. Similar trend was observed in ADG22, where h^2 dropped from 0.27±0.06 to 0.09±0.03. Standard errors of BV for BF22 and ADG22 became smaller when microbiome information was included in the models. The findings from this report might indicate that including information from fecal microbiome in statistical models improves the accuracy of estimated breeding values in pigs.

Keywords: fecal microbiome, genotypes, variance components, pigs

Introduction

The influence of microbial communities living inside an individual has become apparent and documented in the current literature. In pigs, research has shown that bacteria in the pig gut impact the host nutritional, physiological, and immunological processes in various ways (Savage, 1977; Roediger, 1980; Berg, 1996; Backhed *et al.*, 2005; Lee & Mazmanian, 2010; Brestoff & Artis, 2013). Recent advancement in sequencing technology has allowed for large scale sequencing of bacteria to be affordable to the pork industry. Bacteria DNA can be collected from fecal samples, then processed and clustered into operational taxonomic units (hereafter "OTUs"). Studies in the current literature have found OTUs to be associated with phenotypic measures in both humans and animals (e.g. Beaumont et al., 2016; Menni et al., 2017; Yang *et al.*, 2015), and heritable at various levels (e.g. Goodrich et al., 2016). However, most of previous studies have not looked into the relationship between host genetics and its microbiome. In the pork research, there is currently no report on how microbiome affects heritability estimates and accuracy of estimated breeding values.

The report herein was designed to investigate the impact of fecal microbiome on heritability estimates accuracy of breeding values for back fat and average daily gain in pigs.

Materials and Methods

Data

Fecal samples were collected from 1131 pigs (537 female and 594 barrows, average 118.2 ± 1.18 days old), crossed between Duroc sires and Large White x Landrace or Landrace x Large White dams, from a feeding test, which began immediately after weaning (18.6 ± 1.09 days old) and ended when all pigs in a pen achieved an average live weight of 136 kg (approximately 196.4 ± 7.86 day). There were 20 paternal half-siblings of same gender and of similar weaning weight per pen. The experiment was repeated 6 times, each of which comprised of 2 pens (1 pen of female pigs, 1 pen of barrows that are referred to "male" hereafter) from each of the 28 sires. Pigs that came together in 1 replicate were put in

1 contemporary group (hereafter "cg") in analyses that followed. All the pens were located in same house. The pigs were fed standard diets based on sex and live weight. Two phenotypic measures being used in this report were back fat thickness and average daily body weight gain at week 22 post weaning, hereafter BF22 and ADG22, respectively.

Bacterial genomic DNA (gDNA) was extracted from each rectal swab by mechanical disruption in phenol:chloroform. Phased, bi-directional amplification of the V4 region (515–806) of the 16S rRNA gene was employed to generate indexed libraries for Illumina sequencing using the strategy described by Faith et al. (2013). All sequencing was performed at the DNA Sequencing Innovation Lab at the Center for Genome Sciences and Systems Biology at Washington University in St. Louis. Pairs of V4 16S rRNA gene sequences were first merged into a single sequence using FLASh v1.2.11 (Magoc and Salzberg, 2011). Sequences with a mean quality score below Q35 were then filtered out using PRINSEQ v0.20.4 (Schmieder and Edwards, 2011). Sequences with >97% nucleotide sequence identity were then clustered into OTUs. A modified version of GreenGenes (The Greengenes Database Consortium; Schloss and Handelsman, 2006; Ley et al., 2006) was used as the reference database. The most abundant sequence in each cluster was used as the representative sequence for the OTU. Sparse OTUs were filtered out by requiring a minimum total observation count of 1200 for an OTU to be retained, and the OTU table was rarefied to 10,000 counts per sample. There were 1755 OTUs being used for subsequent analyses.

The animals were genotyped with the Illumina PorcineSNP60 Beadchip (Illumina Inc., San Diego, CA USA). All animals had a call rate of at least 0.90. After quality control, there were 53790 SNPs on 18 autosomes, with SNP call rate of at least 0.90 and minor allele frequency of at least 0.01, being used for this report.

Model

Microbiome was used at the OTU level to compute the Jensen-Shannon distance between pairs of samples as follow:

,

in which $D(a,b)$ was the distance between samples a and b; n was the number of OTUs (n=1755); a_i and b_i were the counts of OTU_i in samples a and b, respectively; $m_i=(a_i+b_i)/2$ (Endres, 2003). The resulting square matrix (hereafter "JSD") had zero diagonal, and between 0 and 1 off the diagonal. This JSD matrix was used for clustering the animals *via* partitioning around medoids with the ''pam'' function in the R library ''cluster'' (Maechler et al., 2017). The optimal number of clusters was chosen by maximizing the Calinski–Harabasz index (Calinski and Harabasz, 1974), using ''index.G1'' function in the R library ''clusterSim'' (Walesiak and Dudek, 2017), and the Silhouette index (Rousseeuw, 1987), using the "silhouette" function in the R library ''cluster''. A matrix of microbiome relatedness among the pigs (hereafter "M") was retrieved *via* 1 – JSD.

A genomic relationship matrix (hereafter "G") was generated using the genotypes and VanRaden's method 1 (VanRaden 2008). Three statistical models were used to test the impact of microbiome data on phenotypic variation of BF22 and ADG22. The models were as follows.

[1]
[2]
[3]
[4]

In the 4 models, μ was the overall mean; the fixed effects included sex (), contemporary group (), and age (a, as covariate); the random effects included animal (), microbiome (), pen (), litter (), host genotypes by microbiome interaction was modeled by fitting the interaction between animal (G_k) and

cluster (c_q; n = 2) in models [2] and [4]. Covariance matrices of the animal, microbiome, pen, litter, and residual effects were G,, I, $I$and I, respectively, in which G and M were the genomic relationship matrix and microbiome relatedness matrix as described earlier, and I was the identity matrix. The response in these models was BF22, which was adjusted for liveweight, and ADG22. Potential impact of breeds was tested by including the dam line (Large White x Landrace or Landrace x Large White) as fixed effect in the models, however was not significant for BF22 and ADG22, and thus was removed from the models. The 4 models were tested using ASReml v.4.1 (Gilmour *et al.*, 2015).

Results and Discussion

The animals used in this report were grouped into 2 clusters A (204 females, 237 males) and B (333 females, 357 males), as shown in Figure 1, which had significant impact on BF22 ($P<0.05$) in a separate analysis (more details on BF22 – cluster association and bacterial composition of the clusters are reported by Lu *et al*, 2018). The animal clusters in this report were treated as the environment created by microbiome within the host, and its interaction with the host genetic was explored.

Variance components estimated for BF22 and ADG22 were presented in Tables 1 and 2, respectively. Heritability of BF22 and ADG22 estimated from model 1 was 0.52 and 0.27, respectively. Fitting the interaction between the animal additive genetic effect and the clusters (Model 2) did not significantly improve the model fitness for both BF22 and ADG22, as their final LogLikelihoods changed from 462.05 to 464.25, and from 2294.08 to 2294.74, respectively.

Model [3] differed from model [1] by the microbiome effect, which accounted for 47 and 76% of the total variance in BF22 and ADG22, respectively. Estimated heritability for BF22 and ADG22 dropped to 0.33 and 0.12, respectively, which were much less than their estimates obtained from model [1]. Model [3]'s AIC and LRT showed significant improvement ($P<0.001$) in goodness of fit compared to Model [1]. Heritability estimates for BF22 and ADG22 continued to drop to 0.31 and 0.09, respectively, in Model 4, whereas the proportion of total variance due to microbiome remained the same compared to Model [3]. Though Model [4] did not showed significant improvement in goodness of fit compared to Model [3] ($P>0.05$ for both BF22 and ADG22), we observed significant interaction ($P<0.05$) between the animal additive genetic effect and the clusters in ADG22.

Heritability estimates for BF22 and ADG22 from Model [1] in this report appeared to be in the ranges documented in the current literature for backfat, 0.32-0.72 (Kim *et al.*, 2004; Suzuki *et al.*, 2005; Jiao *et al.*, 2014), and ADG, 0.30-0.47 (Suzuki *et al.*, 2005; Miar *et al.*, 2014; Jiao *et al.*, 2014). As we found that the estimates dropped substantially when microbiome information was included in the models, and as suggested by Sandoval-Motta *et al.* (2017), the heritabilities estimated using only kinship information to relate individuals together might have been over-estimated.

Table 1. Proportion of total variance of BF22 explained by random effects.

	Model 1	Model 2	Model 3	Model 4
Var_g (h^2)	0.52±0.05	0.49±0.06	0.33±0.06	0.31±0.06
Var_m	-	-	0.47±0.09	0.47±0.09
Var_{pen}	0.01±0.02	0.01±0.02	0.01±0.01	0.01±0.01
Var_{litter}	0.03±0.03	0.03±0.04	0.02±0.02	0.02±0.02
Var_c	-	0.01±0.01	-	0
$Var_{g.c}$	-	$0.06±0.05^{NS}$	-	$0.03±0.03^{NS}$
AIC	-916.11	-916.49	-951.07	-948.16
LogL	462.05	464.25	480.54	481.08
LRT	-	$P>0.05^{1,2}$	$P<0.001^{1,3}$	$P>0.05^{3,4}$
SEP	19.06e-02±3.21e-03	19.22e-02±3.12e-03	18.47e-02±3.69e-03	18.58e-02±3.58e-03

Models 1, 2, 3, 4 differed by random effects. Random effects in Model 1 were animal additive genetic, pen, litter, and

residual effects; in Model 2 were animal additive genetic, cluster, genotype x cluster, pen, litter, and residual effects; in Model 3 were animal additive genetic, microbiome, pen, litter, and residual effects; in Model 4 were animal additive genetic, cluster, genotype x cluster, microbiome, pen, litter, and residual effects. Var_g: animal additive genetic effect obtained from the genomic relationship matrix. Var_m: microbiome effect obtained from microbial relatedness matrix. Var_{pen}: pen effect; Var_{litter}: litter effect; Var_c: cluster effect; $Var_{g.c}$: interaction between animal additive genetic effect and cluster. AIC: Akaike information criterion. LRT: Likelihood ratio test comparing Models 1 and 2, 1 and 3, as well as 3 and 4. SEP: standard error of prediction. ^{NS}Not significant.

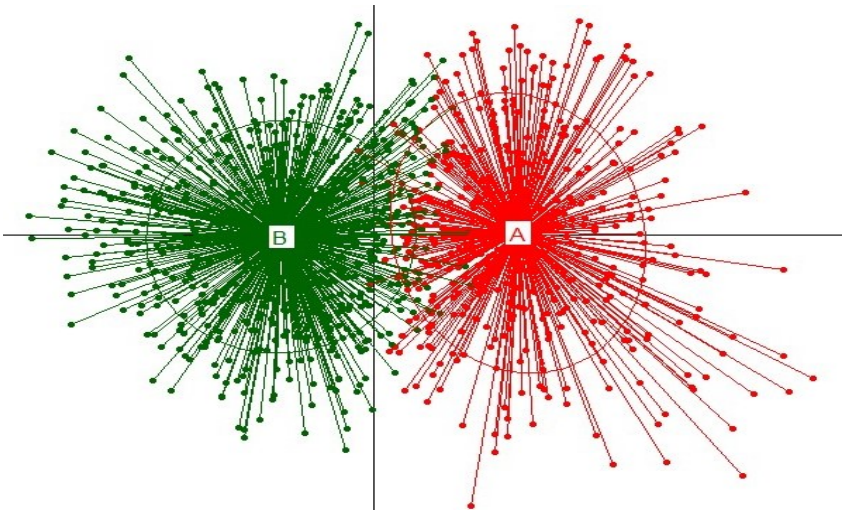
Table 2. Proportion of total variance of ADG22 explained by random effects.

	Model 1	Model 2	Model 3	Model 4
Var_g (h^2)	0.27±0.06	0.22±0.07	0.12±0.03	0.09±0.03
Var_m	-	-	0.76±0.06	0.76±0.06
Var_{pen}	0	0	0	0
Var_{litter}	0.07±0.04	0.07±0.04	0.04±0.02	0.04±0.02
Var_c	-	0	-	0
$Var_{g.c}$	-	0.08±0.07^{NS}	-	0.07±0.03*
AIC	-4580.15	-4577.47	-4644.72	-4644.73
LogL	2294.08	2294.74	2327.36	2329.36
LRT	-	P>0.05^{1,2}	P<0.001^{1,3}	P>0.05^{3,4}
SEP	3.02e-02±5.18e-04	2.91e-02±4.65e-04	2.95e-02±5.72e-04	2.74e-02±4.61e-04

Models 1, 2, 3, 4 differed by random effects. Random effects in Model 1 were animal additive genetic, pen, litter, and residual effects; in Model 2 were animal additive genetic, cluster, genotype x cluster, pen, litter, and residual effects; in Model 3 were animal additive genetic, microbiome, pen, litter, and residual effects; in Model 4 were animal additive genetic, cluster, genotype x cluster, microbiome, pen, litter, and residual effects. Var_g: animal additive genetic effect obtained from the genomic relationship matrix. Var_m: microbiome effect obtained from microbial relatedness matrix. Var_{pen}: pen effect; Var_{litter}: litter effect; Var_c: cluster effect; $Var_{g.c}$: interaction between animal additive genetic effect and cluster. AIC: Akaike information criterion. LRT: Likelihood ratio test comparing Models 1 and 2, 1 and 3, as well as 3 and 4. SEP: standard error of prediction. ^{NS}Not significant. *Significant at P<0.05.

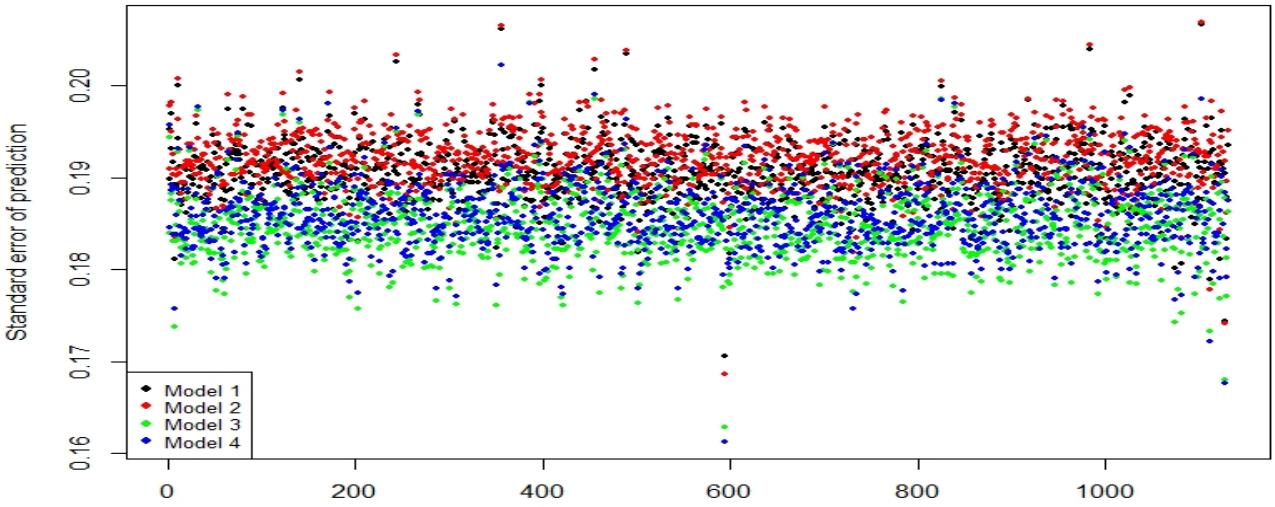
Standard errors of prediction (SEP) for BF22 breeding values and ADG22 breeding values are presented in Tables 1 and 2, as well as Figures 2 and 3, respectively. For BF22, SEP from Models [1] and [2] were around 0.19, decreasing to 0.18 in Models [3] and [4]. Figure 2 shows clear separation between SEPs from Models [1], [2] and those from Models [3], [4]. A clearer decreasing trend in SEP was observed in ADG22, Figure 3, when microbiome information and its interaction with host genetics were added to the model, leading to the smallest SEP of 2.74e-02 from Model [4]. Having smaller SEPs in estimating breeding values is desirable to obtain higher accuracy of prediction.

Figure 1. Identification of clusters in the experimented animals



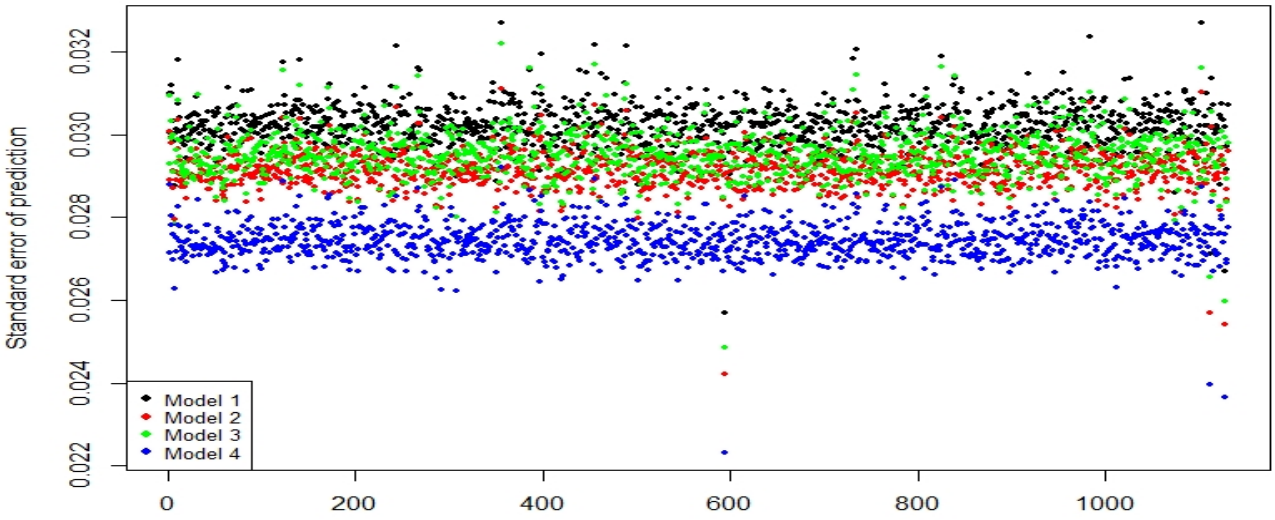
The first two principal coordinates of the Jensen-Shannon distances of the relative abundance profiles. Samples are colored by cluster/enterotype as identified by the partitioning around medoids (PAM) clustering algorithm. Red is cluster A and green is cluster B.

Figure 2. Standard error of prediction for BF22 breeding values.



Models 1, 2, 3, 4 differed by random effects. Random effects in Model 1 were animal additive genetic, pen, litter, and residual effects; in Model 2 were animal additive genetic, cluster, genotype x cluster, pen, litter, and residual effects; in Model 3 were animal additive genetic, microbiome, pen, litter, and residual effects; in Model 4 were animal additive genetic, cluster, genotype x cluster, microbiome, pen, litter, and residual effects

Figure 3. Standard error of prediction for ADG22 breeding values.



Models 1, 2, 3, 4 differed by random effects. Random effects in Model 1 were animal additive genetic, pen, litter, and residual effects; in Model 2 were animal additive genetic, cluster, genotype x cluster, pen, litter, and residual effects; in Model 3 were animal additive genetic, microbiome, pen, litter, and residual effects; in Model 4 were animal additive genetic, cluster, genotype x cluster, microbiome, pen, litter, and residual effects.

Conclusion

The analyses were completed on 1131 animals with fecal microbiome sampled at week 15 post weaning. The results showed that heritability estimates from BF22 and ADG22 were greatly influenced by the fecal microbiome data. Host genotypes by microbiome interaction was also quantified. Fitting the microbiome effect and the interaction between the host genetic and its microbiome appeared to reduce SEPs of estimated breeding values for BF22 and ADG22.

List of References

Backhed, F., R.E. Ley, J.L.Sonnenburg, D.A. Peterson & J.I. Gordon, 2005. Host-bacterial mutualism in the human intestine. Science 307(5717):1915-1920.

Beaumont, M., J.K. Goodrich, M.A. Jackson, I. Yet, E.R. Davenport, S. Vieira-Silva, J. Debelius, T. Pallister, M. Mangino, J. Raes, R. Knight, A.G. Clark, R.E. Ley, T.D. Spector & J.T. Bell, 2016. Heritable components of the human fecal microbiome are associated with visceral fat. Genome Biology; doi: 10.1186/s13059-016-1052-7.

Berg, R.D., 1996. The indigenous gastrointestinal microflora. Trends in Microbiology 4:430–435.

Brestoff, J.R. & D. Artis, 2013. Commensal bacteria at the interface of host metabolism and the immune system. Nat. Immunol.; doi:10.1038/ni.2640.

Calinski, T. & J. Harabasz, 1974. A dendrite method for cluster analysis. Commun Stat. 3:1–27.

Endres, D.M., 2003. A new metric for probability distributions. IEEE Transactions on Information Theory 49(7):1858–1860.

Faith, J.J, J.L. Guruge, M. Charbonneau, S. Subramanian, H. Seedorf, A.L. Goodman, J.C. Clemente, R. Knight, A.C. Heath, R.L. Leibel, M. Rosenbaum & J.I. Gordon, 2013. The Long-Term Stability of the Human Gut Microbiota. Science; doi: 10.1126/science.1237439.

Gilmour, A.R, B.J. Gogel, B.R. Cullis, S.J. Welham & R. Thompson, 2015. ASReml User Guide Release 4.1 Structural Specification, VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Goodrich, J.K., E.R. Davenport , M. Beaumont , M.A. Jackson , R. Knight , C. Ober , T.D. Spector, J.T. Bell, A.G. Clark & R.E. Ley, 2016. Genetic Determinants of the Gut Microbiome in UK Twins. Cell Host & Microbe; doi: 10.1016/j.chom.2016.04.017.

Jiao, S., C. Maltecca, K.A. Gray & J. P. Cassady, 2014. Feed Intake, Average Daily Gain, Feed Efficiency, and Real-Time Ultrasound Traits in Duroc Pigs: I. Genetic Parameter Estimation and Accuracy of Genomic Prediction. J. Anim. Sci. 92:2377–2386.

Kim, I.J., G.Y. Sohn, H.J.Jung & I.Y. Park, 2004. Genetic Parameter Estimates for Backfat Thickness at Three Different Sites and Growth Rate in Swine. AsianAust. J. Anim. Sci.17(3):305-308.

Lee, Y.K. & K.S Mazmanian, 2010. Has the Microbiota Played a Critical Role in the Evolution of the Adaptive Immune System? Science; doi:10.1126/ science.1195568.

Ley, E.R., J.P. Turnbaugh, S. Klein & I.J. Gordon, 2006. Microbial ecology: Human gut microbes associated with obesity. Nature; doi:10.1038/4441022a.

Lu, D., F. Tiezzi, C. Schillebeeckx, P.N. McNulty, C. Schwab, C. Shull & C. Maltecca, 2018. Host contributes to longitudinal diversity of fecal microbiota in swine selected for lean growth. Microbiome; doi: 10.1186/s40168-017-0384-1.

Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, K. Hornik, M. Studer, P. Roudier & J. Gonzalez, 2017. Package 'cluster'. https://cran.r-project.org/web/packages/cluster/cluster.pdf. Accessed 16 August 2017.

Magoc, T. & S. Salzberg, 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics; doi:10.1093/bioinformatics/btr507.

Menni, C., M.A. Jackson, T. Pallister, C.J. Steves, T.D. Spector & A.M. Valdes, 2017. Gut microbiome diversity and high-fibre intake are related to lower long-term weight gain. Int'l J. Obesity; doi:10.1038/ijo.2017.66.

Miar, Y., G. Plastow, H. Bruce, S. Moore, G. Manafiazar, R. Kemp, P. Charagu, A.Huisman, B. van Haandel, C. Zhang, R. McKay & Z. Wang., 2014. Genetic and Phenotypic Correlations between Performance Traits with Meat Quality and Carcass Characteristics in Commercial Crossbred Pigs. PLOS One; doi: 10.1371/journal.pone.0110105.

Roediger, W.E., 1980. Role of anaerobic bacteria in the metabolic welfare of the colonic mucosa in man. Gut 21: 793–798.

Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 20:53–65.

Sandoval-Motta, S., M. Aldana, E. Martinez-Romeo & A. Frank, 2017. The Human Microbiome and the Missing Heritability Problem. Frontiers in Genetics; doi:10.3389/fgene.2017.00080.

Savage, D.C., 1977. Microbial ecology of the gastrointestinal tract. Annual Review of Microbiology; doi:10.1146/annurev.mi.31.100177.000543.

Schloss, P.D. & J. Handelsman, 2006. Toward a cencus of bacteria in soil. PLoS Computational Biology; doi:10.1371/journal.pcbi.0020092.

Schmieder, R. & R. Edwards, 2011. Quality control and processing of metagenomic datasets. Bioinformatics; doi: 10.1093/bioinformatics/btr026.

Suzuki, K., M. Irie, H. Kadowaki, T. Shibata, M. Kumagai & A. Nishida, 2005. Genetic parameter estimates of meat quality traits in Duroc pigs selected for average daily gain, longissimus muscle area, backfat thickness, and intramuscular fat content. J. Anim. Sci. 83(9):2058-2065.

The Greengenes Database Consortium. http://greengenes.secondgenome.com/downloads/database/13_5. Accessed 16 August 2017.

VanRaden, M.P., 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91(11):4414-4423.

Walesiak M. & A. Dudek, 2017. clusterSim: Searching for Optimal Clustering Procedure for a Data Set. https://cran.r-project.org/web/packages/clusterSim/index.html. Accessed 16 August 2017.

Yang, H., X. Huang, S. Fang, W. Xin, L. Huang & C. Chen, 2015. Uncovering the composition of microbial community structure and metagenomics among three gut locations in pigs with distinct

fatness. Scientific Report; doi:10.1038/srep27427.