

1 **INVESTIGATION OF DNA METHYLATION ON PORCINE**
2 **REFERENCE SEQUENCE GENES FOR BOAR TAIN T TRAIT**

3
4 Xiao Wang¹, Haja N. Kadarmideen¹

5 ¹Quantitative Genomics, Bioinformatics and Computational Biology Group, Department of Applied
6 Mathematics and Computer Science, Technical University of Denmark, Richard Petersens Plads,
7 Building 324, 2800, Kongens Lyngby, Denmark

8
9
10
11
12
13
14
15 **Abstract**

16 Boar taint (BT) is an offensive flavor in non-castrated male pigs. However, studies of genome-wide
17 DNA methylation profiles on all reference sequence genes to reveal epigenetic information associated
18 with BT was scarce. Reduced representation bisulfite sequencing (RRBS) is an efficient technology to
19 identify candidate epigenetic biomarkers associated with BT. Three different BT levels were analyzed
20 using RRBS data to calculate differentially methylated genes (DMGs). In this study, we found 411
21 DMGs (Q-value < 0.01) out of 4566 reference genes. Hyper-methylation DMGs were revealed to be
22 mainly enriched in five significant pathways (qvalue < 0.05). These results could contribute to the
23 understanding of methylation levels on reference sequence genes in pigs and the usage of genomic
24 selection for low BT in the breeding programs.

25
26 Keywords: boar taint, DNA methylation, reduced representation bisulfite sequencing, differentially
27 methylated genes

35 INTRODUCTION

36 The offensive odor and/or flavor of boar taint (BT) in porcine meat is primarily caused by the
37 accumulation of skatole (Patterson, 1968) and androstenone (Gower, 1972) in a percentage of non-
38 castrated male pigs (boars). Surgical castration is an effective solution to avoid BT but results in reduced
39 feed conversion efficiency, leaner carcass values and increased fecal and urinary nitrogen, in addition
40 to animal welfare concerns (Bonneau, 1998; Claus *et al.*, 1994). As levels of skatole and androstenone
41 traits exhibit moderate to high heritability (Strathe *et al.*, 2013a, 2013b), selection of low genetic merit
42 of BT can be an effective approach to avoid BT and other disadvantages of surgical castration. Recent
43 studies identified and exploited genetic variation of androstenone and skatole (Rowe *et al.*, 2014) and
44 predicted genomic breeding values for genomic selection (de Campos *et al.*, 2015). In addition, gene
45 expression and expression quantitative trait loci (eQTLs) profiles associated with BT in pigs were
46 revealed by system genomics analyses (Drag *et al.*, 2017, 2018).

47 As an important epigenetic modification, DNA methylation has been examined to be associated with
48 growth (Jin *et al.*, 2014), immune response (Wang *et al.*, 2017) and reproduction traits (Bell *et al.*,
49 2011) in pigs. Reduced representation bisulfite sequencing (RRBS) has been implemented to analyze
50 patterns of DNA methylation by reducing the portion of the genome digestion (Meissner *et al.*, 2005).
51 The RRBS method primarily focuses on the enrichment of CpG-rich regions rather than the non-CpG
52 regions (Meissner *et al.*, 2005). In mammals, DNA methylation almost exclusively occurs at CG
53 dinucleotides with 70-80% throughout the genome (Ehrlich *et al.*, 1982; Law and Jacobsen, 2010).
54 Currently, RRBS analysis of the pigs has been presented using intestinal tissue (Gao *et al.*, 2014),
55 ovaries (Yuan *et al.*, 2016), neocortex, liver, muscle and spleen (Choi *et al.*, 2015) and prepubertal
56 (Chen *et al.*, 2018) and adult (Wang and Kadarmideen, 2019) testis.

57 Gene information from Reference Sequence (RefSeq) database (Pruitt *et al.*, 2007) is used as the
58 reference standards for well-characterized genes. To our knowledge, no studies have investigated the
59 genome-wide DNA methylation profiles using RRBS data on all reference sequence genes
60 (RefSeqGenes) to reveal epigenetic information associated with BT. Therefore, the aim of this study
61 is to analyze differentially methylated genes (DMGs) based on RefSeqGene in three different (high,
62 medium and low) BT levels and then reveal candidate DMGs for further epigenetic application in pig
63 breeding programs for low BT levels.

64

65 MATERIALS AND METHODS

66 The study was carried out using three groups of high (n = 3), medium (n = 3) and low (n = 3) BT levels
67 based on the most extreme estimated breeding values (EBVs) within high and low groups and the closest
68 mean values in the medium group. The EBVs of low, medium and high BT were calculated by summing
69 the EBVs of skatole concentration and human nose score (HNS). Nine pigs were slaughtered at a

70 commercial slaughterhouse (Danish Crown, Herning, Denmark), and 150 mg testis tissue samples for
71 each pig were retrieved by punch biopsy for RRBS sequencing.

72 RRBS adapters and reads less than 20 bases long were trimmed by *Trimmomatic* software (Bolger *et*
73 *al.*, 2014). Then, *Bismark* software (Krueger and Andrews, 2011) was applied to map clean reads to
74 the pig reference genome (Sscrofa11.1/susScr11), and then determine the methylation levels for each
75 cytosine site.

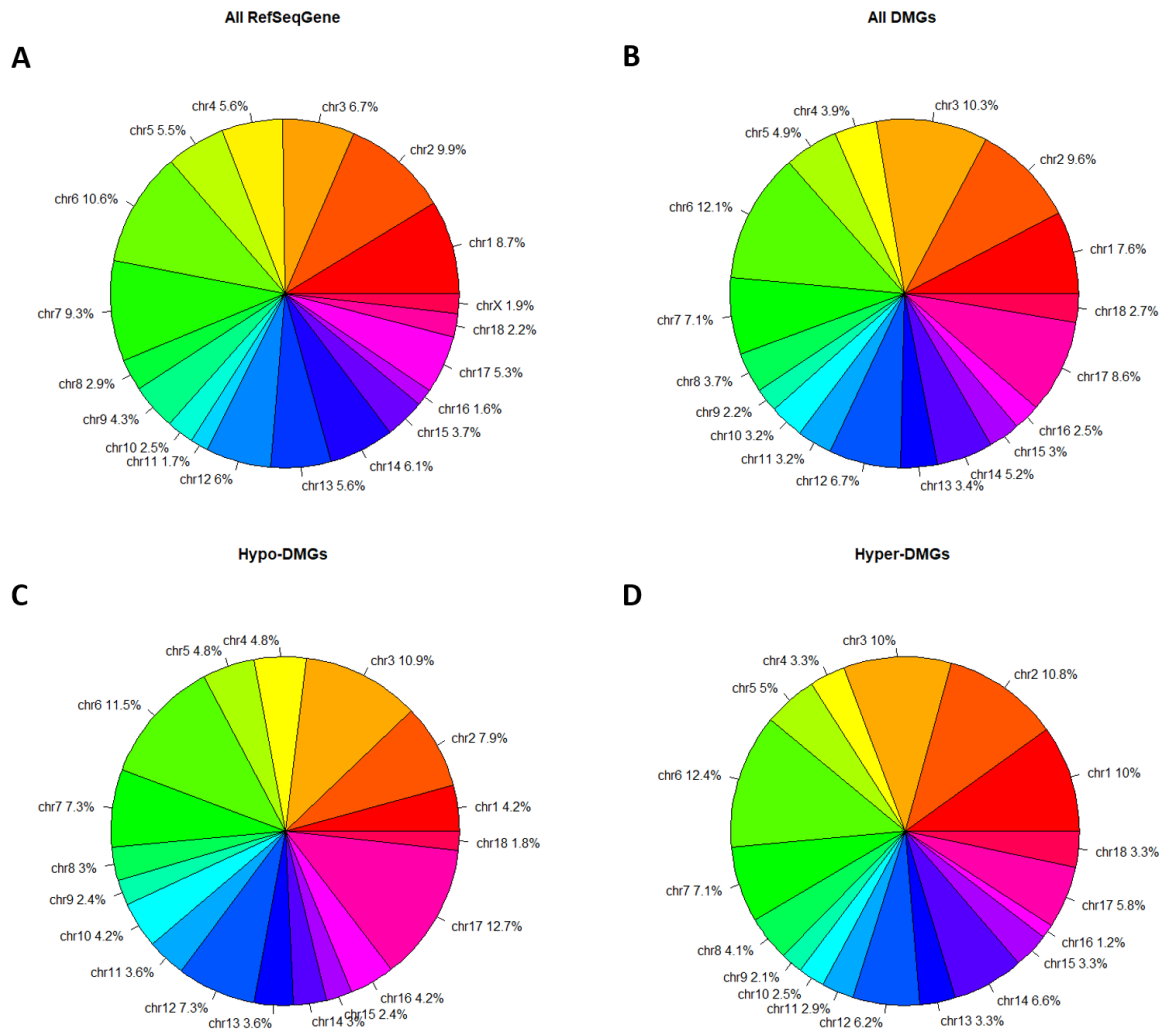
76 The DMGs were defined for significant genes after the comparisons against methylation levels on
77 RefSeqGenes using R package *GeneDMRs* (<https://github.com/xiaowangCN/GeneDMRs>). To account
78 for multiple hypothesis testing, P-values can be adjusted to Q-values by false discovery rate (FDR)
79 method (Hochberg, 1995). RefSeqGene database (Sscrofa11.1/susScr11) used in this study was
80 downloaded from the University of California Santa Cruz (UCSC) website
81 (<http://genome.ucsc.edu/cgi-bin/hgTables>).

82 The pathway terms were analyzed in hypo-methylated and hyper-methylated categories based on R
83 package *clusterprofiler* (Yu *et al.*, 2012) and *KEGG.db* (version 3.2.3) of *Sus scrofa* organism.
84 According to R package *GeneDMRs*, the hypo-methylated and hyper-methylated DMG was defined
85 when the methylation difference between low and high BT levels was positive and negative,
86 respectively.

87

88 **RESULTS AND DISCUSSION**

89 In this study, a total of 4566 RefSeqGenes were used for analysis. Finally, 2772 genes of them with
90 annotated genomic position and cytosine sites were used for DMG detection. After statistical analysis,
91 411 genes were defined as DMGs with Q-values less than ($<$) 0.01. Compared to the distribution of
92 2772 genes on porcine genome, the percentages of 411 DMGs on chromosome 3 and chromosome 17
93 showed the larger ratios than before, but DMGs on chromosome X were not detected (Figure 1A and
94 Figure 1B). In addition, hypo-DMGs were mostly located on chromosome 17, while hyper-DMGs
95 were mostly located on chromosome 6 (Figure 1C and Figure 1D).



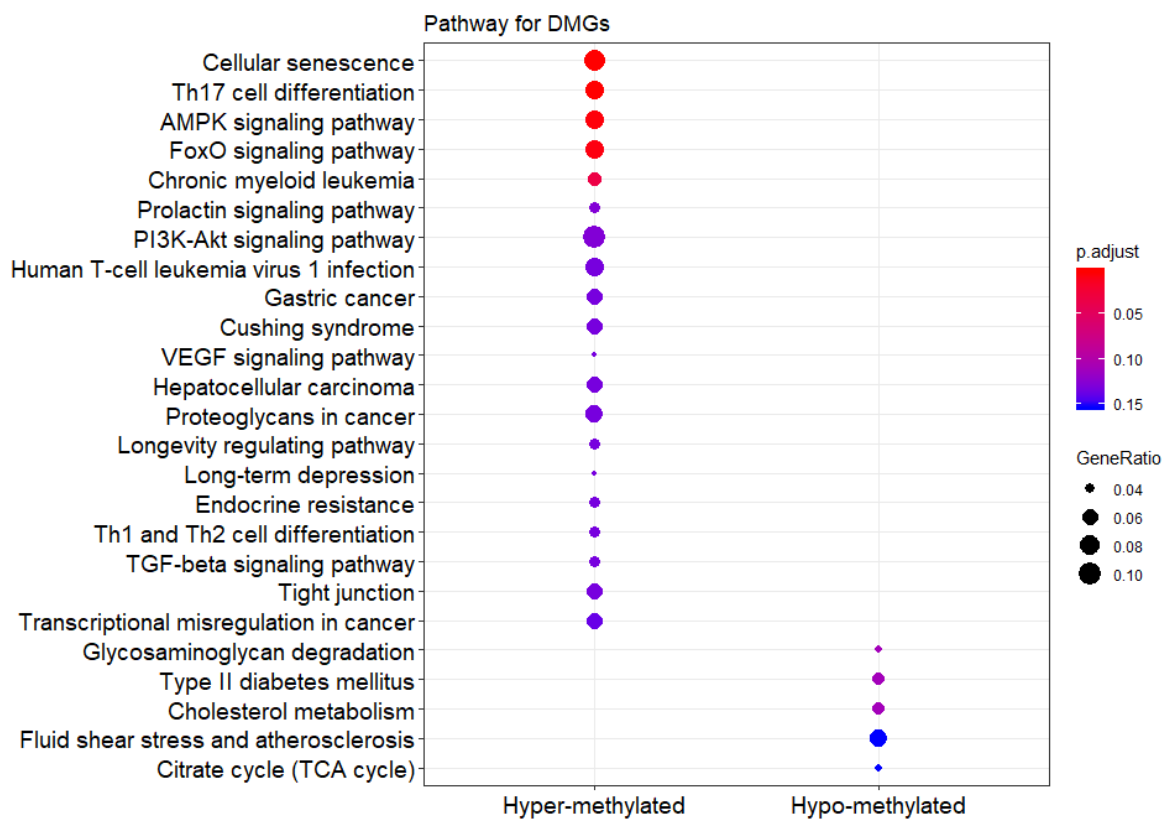
96

97 Figure 1. Percentages of all (A) RefSeqGene, (B) DMGs (Q-value < 0.01), (C) hypo-DMGs (Q-value
 98 < 0.01) and (D) hyper-DMGs (Q-value < 0.01) in different chromosomes.

99

100 Our results revealed that most of the DMGs were enriched in hyper-methylated category and hypo-
 101 methylated DMGs were seldom enriched (Figure 2). Five pathways of them were significant (qvalue
 102 < 0.05) including cellular senescence, Th17 cell differentiation, AMPK signaling pathway, FOXO
 103 signaling pathway and chronic myeloid leukemia (Table 1). In a period of cellular proliferation or a
 104 rapid manner in response to acute stress, senescence usually occurs. The cellular senescence is
 105 activated by normal cells in response to various types of stress including telomere uncapping, DNA
 106 damage, oxidative stress and oncogene activity (Ben-Porath and Weinberg, 2005). In our study, 11
 107 DMGs (NM_001038639, NM_001113452, NM_001123080, NM_001135959, NM_001145750,
 108 NM_001244550, NM_001285967, NM_001287846, NM_214015, NM_214124 and NM_214128)
 109 were involved in the cellular senescence pathway (qvalue = 0.0026). Additionally, T helper 17
 110 cells (Th17), a subset of pro-inflammatory T helper cells, can produce interleukin 17 (IL-17). Th17

111 are related to T regulatory cells because the signals that cause Th17s to differentiate actually inhibit T
 112 regulatory cells differentiation (Hartigan-O'Connor *et al.*, 2011). AMP-activated protein kinase
 113 (AMPK) is one of the central regulators of cellular and organismal metabolism. It has critical roles in
 114 regulating growth, reprogramming metabolism and cellular processes (Mihaylova and Shaw, 2011).
 115 Forkhead box O (FOXO), one subfamily of the fork head transcription factor family, plays important
 116 roles in cell fate decisions and in a wide range of cancers as a tumor suppressor (Farhan *et al.*, 2017).
 117 Chronic myeloid leukaemia (CML) is a clonal myeloproliferative disorder characterized by the
 118 presence of BCR-ABL, while BCR-ABL is an oncogene created by the fusion of *BCR* and *ABL* genes
 119 (Rowley, 1973). The juxtaposition of these genes encodes a novel fusion gene that translates into a
 120 protein with constitutive TK activity when they are in response to genetic mutation (Sinclair *et al.*,
 121 2013).
 122



123

124 Figure 2. Pathways of the DMGs in the hypo-methylated and hyper-methylated categories.

125

126 Table 1. Significant pathways (qvalue < 0.05) for DMGs.

Pathway	Description	pvalue	p.adjust	qvalue	GeneRatio	Genes
ssc04218	Cellular senescence	1.84e-05	0.0031	0.0026	11/119	NM_001038639/NM_001113452/NM_001123080/ NM_001135959/NM_001145750/NM_001244550/

						NM_001285967/NM_001287846/NM_214015/ NM_214124/NM_214128
ssc04659	Th17 cell differentiation	3.02e- 05	0.0031	0.0026	9/119	NM_001038639/NM_001044567/NM_001113452/ NM_001246252/NM_001253352/NM_001315722/ NM_214015/NM_214128/NM_214290
ssc04152	AMPK signaling pathway	7.99e- 05	0.0055	0.0047	9/119	NM_001135959/NM_001145750/NM_001243060/ NM_001243802/NM_001244062/NM_001287846/ NM_214025/NM_214157/NM_214172
ssc04068	FOXO signaling pathway	1.39e- 04	0.0072	0.0061	9/119	NM_001038639/NM_001044588/NM_001123080/ NM_001135959/NM_001145750/NM_001244550/ NM_214015/NM_214124/NM_214172
ssc05220	Chronic myeloid leukemia	8.20e- 04	0.034	0.029	6/119	NM_001038639/NM_001244550/NM_001246252/ NM_001285967/NM_214015/NM_214290

127

128 Conclusion

129 After the investigation of methylation levels on reference genes in pigs, 411 DMGs (Q-value < 0.01)
 130 were identified from the 4566 reference genes. We also found that five significant pathways (qvalue <
 131 0.05) were mainly from the hyper-methylation DMGs. In this study, these results could be used in the
 132 genomic selection for low BT in the breeding programs.

133

134 References

135 Bell, J. T., Pai, A. A., Pickrell, J. K., Gaffney, D. J., Pique-Regi, R., Degner, J. F., et al. (2011). DNA
 136 methylation patterns associate with genetic and gene expression variation in HapMap cell lines.
 137 *Genome Biol.* 12, R10. doi:10.1186/gb-2011-12-1-r10.

138 Ben-Porath, I., and Weinberg, R. A. (2005). The signals and pathways activating cellular senescence.
 139 *Int. J. Biochem. Cell Biol.* doi:10.1016/j.biocel.2004.10.013.

140 Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina
 141 sequence data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170.

142 Bonneau, M. (1998). Use of entire males for pig meat in the European Union. *Meat Sci* 49, Supple,
 143 S257–S272. doi:http://dx.doi.org/10.1016/S0309-1740(98)90053-5.

144 Chen, X., Shen, L. H., Gui, L. X., Yang, F., Li, J., Cao, S. Z., et al. (2018). Genome-wide DNA
 145 methylation profile of prepubertal porcine testis. *Reprod. Fertil. Dev.* 30, 349–358.
 146 doi:10.1071/RD17067.

147 Choi, M., Lee, J., Le, M. T., Nguyen, D. T., Park, S., Soundrarajan, N., et al. (2015). Genome-wide
 148 analysis of DNA methylation in pigs using reduced representation bisulfite sequencing. *DNA Res.* 22,

- 149 343–355. doi:10.1093/dnares/dsv017.
- 150 Claus, R., Weiler, U., and Herzog, A. (1994). Physiological aspects of androstenone and skatole
151 formation in the boar-A review with experimental data. *Meat Sci.* 38, 289–305. doi:10.1016/0309-
152 1740(94)90118-X.
- 153 de Campos, C. F., Lopes, M. S., de Silva, F. F., Veroneze, R., Knol, E. F., Sávio Lopes, P., et al.
154 (2015). Genomic selection for boar taint compounds and carcass traits in a commercial pig
155 population. *Livest. Sci.* 174, 10–17. doi:10.1016/j.livsci.2015.01.018.
- 156 Drag, M., Hansen, M. B., and Kadarmideen, H. N. (2018). Systems genomics study reveals
157 expression quantitative trait loci, regulator genes and pathways associated with boar taint in pigs.
158 *PLoS One* 13. doi:10.1371/journal.pone.0192673.
- 159 Drag, M., Skinkytė-Juskienė, R., Do, D. N., Kogelman, L. J. A., and Kadarmideen, H. N. (2017).
160 Differential expression and co-expression gene networks reveal candidate biomarkers of boar taint in
161 non-castrated pigs. *Sci. Rep.* 7. doi:10.1038/s41598-017-11928-0.
- 162 Ehrlich, M., Gama-Sosa, M. A., Huang, L. H., Midgett, R. M., Kuo, K. C., Mccune, R. A., et al.
163 (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or
164 cells. *Nucleic Acids Res.* doi:10.1093/nar/10.8.2709.
- 165 Farhan, M., Wang, H., Gaur, U., Little, P. J., Xu, J., and Zheng, W. (2017). FOXO signaling pathways
166 as therapeutic targets in cancer. *Int. J. Biol. Sci.* doi:10.7150/ijbs.20052.
- 167 Gao, F., Zhang, J., Jiang, P., Gong, D., Wang, J. W., Xia, Y., et al. (2014). Marked methylation
168 changes in intestinal genes during the perinatal period of preterm neonates. *BMC Genomics* 15.
169 doi:10.1186/1471-2164-15-716.
- 170 Gower, D. B. (1972). 16-Unsaturated C19steroids a review of their chemistry, biochemistry and
171 possible physiological role. *J. Steroid Biochem.* 3. doi:10.1016/0022-4731(72)90011-8.
- 172 Hartigan-O'Connor, D. J., Hirao, L. A., McCune, J. M., and Dandekar, S. (2011). Th17 cells and
173 regulatory T cells in elite control over HIV and SIV. *Curr. Opin. HIV AIDS.*
174 doi:10.1097/COH.0b013e32834577b3.
- 175 Hochberg, B. (1995). Controlling the False Discovery Rate: a Practical and Powerful Approach to
176 Multiple Testing. *J. R. Stat. Soc.* doi:10.1017/CBO9781107415324.004.
- 177 Jin, L., Jiang, Z., Xia, Y., Lou, P., Chen, L., Wang, H., et al. (2014). Genome-wide DNA methylation
178 changes in skeletal muscle between young and middle-aged pigs. *BMC Genomics* 15, 653.
179 doi:10.1186/1471-2164-15-653.
- 180 Krueger, F., and Andrews, S. R. (2011). Bismark: A flexible aligner and methylation caller for
181 Bisulfite-Seq applications. *Bioinformatics* 27, 1571–1572. doi:10.1093/bioinformatics/btr167.
- 182 Law, J. A., and Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation
183 patterns in plants and animals. *Nat. Rev. Genet.* 11, 204–220. doi:10.1038/nrg2719.
- 184 Meissner, A., Gnirke, A., Bell, G. W., Ramsahoye, B., Lander, E. S., and Jaenisch, R. (2005).
185 Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation
186 analysis. *Nucleic Acids Res.* doi:10.1093/nar/gki901.
- 187 Mihaylova, M. M., and Shaw, R. J. (2011). The AMPK signalling pathway coordinates cell growth,
188 autophagy and metabolism. *Nat. Cell Biol.* doi:10.1038/ncb2329.

- 189 Patterson, R. L. S. (1968). 5 α -Androst-16-ene-3-one:-Compound responsible for boar taint in boar fat.
190 *J. Sci. Food Agric.* 19, 31–38. doi:10.1002/jsfa.2740190107.
- 191 Pruitt, K. D., Tatusova, T., and Maglott, D. R. (2007). NCBI reference sequences (RefSeq): A curated
192 non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.*
193 doi:10.1093/nar/gkl842.
- 194 Rowe, S. J., Karacaören, B., de Koning, D.-J., Lukic, B., Hastings-Clark, N., Velandar, I., et al.
195 (2014). Analysis of the genetics of boar taint reveals both single SNPs and regional effects. *BMC*
196 *Genomics* 15, 424. doi:10.1186/1471-2164-15-424.
- 197 Rowley, J. D. (1973). A new consistent chromosomal abnormality in chronic myelogenous leukaemia
198 identified by quinacrine fluorescence and Giemsa staining. *Nature*. doi:10.1038/243290a0.
- 199 Sinclair, A., Latif, A. L., and Holyoake, T. L. (2013). Targeting survival pathways in chronic myeloid
200 leukaemia stem cells. *Br. J. Pharmacol.* doi:10.1111/bph.12183.
- 201 Strathe, A. B., Velandar, I. H., Mark, T., and Kadarmideen, H. N. (2013a). Genetic parameters for
202 androstenone and skatole as indicators of boar taint and their relationship to production and litter size
203 traits in Danish Landrace. *J. Anim. Sci.* 91, 2587–2595. doi:10.2527/jas.2012-6107.
- 204 Strathe, A. B., Velandar, I. H., Mark, T., Ostersen, T., Hansen, C., and Kadarmideen, H. N. (2013b).
205 Genetic parameters for male fertility and its relationship to skatole and androstenone in Danish
206 Landrace boars. *J. Anim. Sci.* 91, 4659–4668. doi:10.2527/jas.2013-6454.
- 207 Wang, H., Wang, J., Ning, C., Zheng, X., Fu, J., Wang, A., et al. (2017). Genome-wide DNA
208 methylation and transcriptome analyses reveal genes involved in immune responses of pig peripheral
209 blood mononuclear cells to poly I:C. *Sci. Rep.* 7. doi:10.1038/s41598-017-10648-9.
- 210 Wang, X., and Kadarmideen, H. N. (2019). An Epigenome-Wide DNA Methylation Map of Testis in
211 Pigs for Study of Complex Traits. *Front. Genet.* 10. doi:10.3389/fgene.2019.00405.
- 212 Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing
213 Biological Themes Among Gene Clusters. *Omi. A J. Integr. Biol.* doi:10.1089/omi.2011.0118.
- 214 Yuan, X. L., Gao, N., Xing, Y., Zhang, H. Bin, Zhang, A. L., Liu, J., et al. (2016). Profiling the
215 genome-wide DNA methylation pattern of porcine ovaries using reduced representation bisulfite
216 sequencing. *Sci. Rep.* 6. doi:10.1038/srep22138.

217

218 **Acknowledgment**

219 This study was funded by the AGES project of the Veterinary Department of the Ministry of Food,
220 Agriculture and Fisheries, Denmark. Xiao Wang received Ph.D. stipends from the Technical University
221 of Denmark (DTU Bioinformatics and DTU Compute), Denmark, and the China Scholarship Council,
222 China.