**Open Access**
Asian-Australas J Anim Sci
Vol. 31, No. 8:1141-1149 August 2018
https://doi.org/10.5713/ajas.17.0805
pISSN 1011-2367 eISSN 1976-5517

AJAS
Asian-Australasian Journal of Animal Sciences

# Correlation analyses of CpG island methylation of cluster of differentiation 4 protein with gene expression and T lymphocyte subpopulation traits

**Xueyan Zhao[1], Yanping Wang[1], Jianfeng Guo[1], and Jiying Wang[1,*]**

* **Corresponding Author:** Jiying Wang
**Tel:** +86-0531-88611680, **Fax:** +86-0531-88611680,
**E-mail:** jnwangjiying@163.com

[1] Shandong Provincial Key Laboratory of Animal Disease Control and Breeding, Institute of Animal Science and Veterinary Medicine, Shandong Academy of Agricultural Sciences, Jinan 250100, China

**ORCID**
Xueyan Zhao
https://orcid.org/0000-0001-8683-2569
Yanping Wang
https://orcid.org/0000-0003-3828-1293
Jianfeng Guo
https://orcid.org/0000-0001-6492-9351
Jiying Wang
https://orcid.org/0000-0002-6312-1390

Submitted Oct 31, 2017; Revised Jan 3, 2018; Accepted Feb 11, 2018

**Objective:** Cluster of differentiation 4 protein (*CD4*) gene is an important immune related gene which plays a significant role in T cell development and host resistance during viral infection.
**Methods:** In order to unravel the relationship of CpG island methylation level of *CD4* gene with its gene expression and T lymphocyte subpopulation traits, we used one typical Chinese indigenous breed (Dapulian, DP) and one commercial breed (Landrace), then predicted the CpG island of *CD4* gene, determined the methylation status of CpG sites by bisulfite sequencing polymerase chain reaction (BSP), and carried out the correlation analyses of methylation frequencies of CpG sites with mRNA expression and T lymphocyte subpopulation traits.
**Results:** There was one CpG island predicted in the upstream –2 kb region and exon one of porcine *CD4* gene, which located 333 bp upstream from the start site of gene and contained nine CpG sites. The correlation analysis results indicated that the methylation frequency of CpG_2 significantly correlated with *CD4* mRNA expression in the DP and Landrace combined population, though it did not reach significance level in DP and Landrace separately. Additionally, 15 potential binding transcription factors (TFs) were predicted within the CpG island, and one of them (Jumonji) contained CpG_2 site, suggesting that it may influence the *CD4* gene expression through the potential binding TFs. We also found methylation frequency of CpG_2 negatively correlated with T lymphocyte subpopulation traits CD4+CD8-CD3-, CD4-CD8+CD3- and CD4+/CD8+, and positively correlated with CD4-CD8+CD3+ and CD4+CD8+CD3+ (for all correlation, $p<0.01$) in DP and Landrace combined population. Thus, the CpG_2 was a critical methylation site for porcine *CD4* gene expression and T lymphocyte subpopulation traits.
**Conclusion:** We speculated that increased methylation frequency of CpG_2 may lead to the decreased expression of *CD4*, which may have some kind of influence on T lymphocyte subpopulation traits and the immunity of DP population.

**Keywords:** Cluster of Differentiation 4 (*CD4*); Methylation; T Lymphocyte Subpopulations; Correlation Analyses; Dapulian Piglets

## INTRODUCTION

T lymphocytes are central elements of immune system which play a critical role in cell-mediated immunity [1]. During the T lymphocytes development, glycoprotein cluster of differentiation 4 (*CD4*) plays an important role in development of helper T (Th) cells and activates the Th cell maturation process [2,3]. CD4 molecule is expressed not only in T lymphocytes, but also in B cells, macrophages, and granulocytes. *CD4* transcription is under the control of several cis-acting elements including enhancers, silencers and DNA methylation

Copyright © 2018 by Asian-Australasian Journal of Animal Sciences
This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

[4,5]. Zou et al [6] found a silencer element within first intron of *CD4* gene was sufficient for *CD4* transcriptional repression in cells of the cytotoxic lineage, as well as in thymocytes at earlier stages of differentiation. In pigs, the *CD4* can be taken as a candidate gene due to its important function in porcine immunity, especially T lymphocyte subpopulation traits. Xu et al [7] detected the association between polymorphisms of the *CD4* gene and T lymphocyte subpopulations and found the *CD4* gene may influence T lymphocyte subpopulations.

DNA methylation modification is of great importance for genome reprogramming and gene expression which control animal development [8]. In general, DNA methylation occurs most commonly in CpG islands, which are often associated with gene promoters [9,10]. Methylation within CpG islands is involved in repression of transcription, by altering chromatin structure [11,12], DNA conformation [13,14] and regulating transcription factors (TFs) activity [15]. So far, there are a number of methods that have been developed recently to quantify DNA methylation. Bisulfite sequencing polymerase chain reaction (PCR) (BSP) has become one of the most frequently used techniques in this field [16,17]. Target DNA fragment from numerous bacterial clones is sequenced to determine the extent of methylation at each CpG site. Altering *CD4* gene methylation status, the expression was changed which related to resistance to virus infection [18] and inflammatory diseases [19]. In chicken, promoter methylation of *CD4* gene was deemed to be down-regulated after Marek's disease virus infection. By virus-like double-stranded RNA and DNA infection, promoter methylation status of porcine *CD4* gene was changed in kidney epithelial cells [20]. However, the methylation status of porcine *CD4* CpG island in peripheral blood between different breeds with different disease resistance are still unclear.

Breed is one of the most crucial factors that has a direct effect on resistance or susceptibility to various infectious diseases [21,22]. Most of indigenous pig breeds in China are generally better at disease resistance and immunity than modern commercial breeds [23]. Dapulian (DP), an indigenous pig breed distributed in Shandong province of China, exhibits stronger resistance to diseases [23,24]. In previous studies, we found that there were significant differences between DP and Landrace in routine blood parameters, T lymphocyte subpopulation traits and cytokines and receptor mRNA expression in peripheral blood [25,26]. In this study, we explored the methylation status of *CD4* gene in peripheral blood of DP and Landrace piglets, and elucidated the correlation of that with CD4 mRNA expression and T lymphocyte subpopulation traits. Our study will provide crucial information to help understanding molecular mechanisms of indigenous and western pig breeds with different disease resistance.

## MATERIALS AND METHODS

### Blood preparation

We employed 124 DP piglets and 187 Landrace piglets as experimental individuals which were from two stock farms in Jining of China. All piglets born from 13 DP sows and 28 Landrace sows were 35 days old. The whole procedure for collection of blood was performed in strict accordance with guideline (IACC20060101, 1 January 2006) of the Institutional Animal Care and Use Committee of Institute of Animal Science and Veterinary Medicine, Shandong Academy of Agricultural Sciences. About 5 mL peripheral blood was collected per piglet via venipuncture into a vacutainer tube using EDTAK2 as anticoagulant.

### T lymphocyte subpopulation traits measurement

Three-color flow cytometry analyses were performed on blood samples within 24 hours collected to distinguish T cell subpopulations. The three monoclonal antibodies used in the study were purchased from SouthernBiotech (Birmingham, AL, USA). Monoclonal antibodies were labelled three-color surface immunofluorescence, fluorescein isothiocyanate, R-phycoerythrin, and spectral red, for the simultaneous detection of three antigens, CD4, CD8, and CD3 on individual lymphoid cells. The lymphocytes were distinguished into the following subpopulation traits including CD4-CD8-CD3-, CD4+CD8-CD3-, CD4-CD8+CD3-, CD4+CD8+CD3-, CD4-CD8-CD3+, CD4+CD8-CD3+, CD4-CD8+CD3+, and CD4+CD8+CD3+. We summed CD4-CD8-CD3+, CD4+CD8-CD3+, CD4-CD8+CD3+, CD4+CD8+CD3+ as CD3+, and recorded CD4+CD8-CD3+/CD4-CD8+CD3+ as CD4+/CD8+.

### Bioinformatic analyses

Analysis of CpG islands in the upstream –2 kb region and exon one (104 bp) of porcine *CD4* gene transcript (XM_013987331.1) was performed by the online tools Li Lab MethPrimer (http://www.urogene.org/methprimer/). Parameter setting was as follows: Island size >100, guanine-cytosine percent >50.0, Obs/Exp >0.6. MatInspector [27] was used to recognize putative transcription factor binding sites (TFBS) within the CpG island using the following conditions: core similarity was set 1.00; matrix similarity was set 0.90.

### Bisulfite modification and BSP-sequencing analysis

Genomic DNA was extracted from blood samples using phenol-chloroform method, and its quality was checked by agarose gel electrophoresis and NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA, USA). DNA from 45 DP and 48 Landrace were modified with sodium bisulfite according to manufacturer's instructions of EZ DNA Methylation-Gold Kit (ZYMO RESEARCH, Orange County, CA, USA). This procedure converts unmethylated cytosine resi-

dues to uracil that is recognized as thymine by Taq polymerase, whereas the methylated cytosine remains unchanged. The modified DNA was either used immediately as a template for following PCR or stored at –20°C.

The target fragment containing the CpG island of *CD4* gene was amplified by PCR. Primers were also designed by Li Lab MethPrimer [28], and the forward and reverse primers were 5'-GTTTGATTGGAGTTATAGATGTT-3' and 5'-TTAACTC TCAACTCTTAAATACACT-3', respectively. The amplified fragment length was 347 bp. The 50 μL PCR reaction mixture included 100 ng bisulfite-treated DNA, 1×EpiTaq PCR Buffer ($Mg^{2+}$ free, TaKaRa, Dalian, China), 2.5 mM $MgCl_2$, 0.3 mM dNTP mixture, 0.4 μM forward and reverse primers, 1.25 U TaKaRa EpiTaq HS. The following reaction conditions were used: 35 cycles of 98°C for 10 s, 55°C for 30 s, and 72°C for 1 min. The PCR products were subjected to electrophoresis on agarose gels, excised, purified and inserted into the Peasy-T5 Zero Cloning vector (Transgene, Beijing, China). The recombinant clones were used to transform *Escherichia coli* Trans1-T1 cells (Transgene, Beijing, China). Positive recombinant clones were selected on LB agar plates containing 100 μg/mL ampicillin (Tangen, Beijing, China), and confirmed by PCR. Finally, 11-20 positive recombinant clones per individual were selected and sequenced using an ABI3700XL DNA sequencer (Applied Biosystems, Carlsbad, CA, USA).

### Quantitative real-time PCR analysis

Blood samples (0.3 mL) were homogenized in 0.7 mL RNAiso Plus (TaKaRa, Beijing, China) and RNA was extracted for each sample according to manufacturer's instructions. RNA concentrations were measured on Nanodrop 2000 spectrophotometer, and RNA integrity was verified by 0.8% gel electrophoresis. cDNA was synthesized from 500 ng of total RNA (from 104 DP and 171 Landrace) as a template by PrimerScript RT reagent with gDNA Eraser (Takara, China). Real-time quantitative PCR was performed using Roche lightcycler 480 system with LightCycler 480 SYBR Green I Master following manufacturer's instructions (Roche, Basel, Switzerland). Primers were designed with Primer Premier 5.0 software using GenBank accession NM_001001908.2 for porcine *CD4* gene and accession NM_213978.1 for beta-2-microglobulin (*B2M*) gene. Primer sequences were as follows: forward primer 5'-GAGA AGAAGACCTGCCAATG-3' and reverse primer 5'-GAAG CAAGGCCCACTGAA-3' for *CD4* gene; forward primer 5'-TTCACACCGCTCCAGTAG-3' and reverse primer 5'-CCAGATACATAGCAGTTCAGG -3' for *B2M* gene. Real-time PCR amplification was performed in a 20 μL reaction mixture containing 1 μL cDNA, 0.5 μM each forward and reverse primer, 1×SYBR Green I Master. The PCR protocol was as follows: 95°C for 15 s; 40 cycles of 95°C for 10 s, 60°C for 10 s, and 72°C 20 s; followed by 95°C for 5 s and 65°C for 1 min. The expression of *CD4* in each sample was normalized to that of *B2M*. Triplicate PCR amplifications were performed for each sample.

### Statistical analyses

Methylation sequencing results were processed by BIQ Analyzer software [29] for analysis. An individual's methylation frequency for a CpG site was average percentage of methylated cytosines for CpG dinucleotides of sequenced positive clones in this CpG, and population methylation frequency for a CpG site in DP or Landrace population was average of methylation frequencies of each piglet in this CpG. The real-time PCR results of *CD4* gene were calculated by the $2^{-\Delta\Delta Ct}$ method. The relative expression was represented by ΔCt ($Ct_{CD4}$–$Ct_{B2M}$). Mean value, standard deviation, variable coefficient, maximum and minimum of expression and T lymphocyte subpopulation traits data were calculated within Microsoft Excel (Redmond, WA, USA). Least square mean analyses of *CD4* gene expression and T lymphocyte subpopulation and correlation analyses between CpG sites methylation frequencies of *CD4* gene with mRNA expression level and T lymphocyte subpopulations traits were all carried out with R in 45 DP, 48 Landrace and 93 combined population of these two breeds.

## RESULTS AND DISCUSSION

### DNA methylation profile of porcine *CD4* gene

CpG island analyses by Li Lab MethPrimer showed that there was one CpG island (122 bp) predicted in the upstream –2 kb region and exon one of porcine *CD4* gene, which located 333 bp upstream from the start site of gene. The predicted location of this CpG island in our study was the same as that identified in Wang et al.'s study [20]. The detailed information of the CpG island is presented in Figure 1. Moreover, there were 9 CpG sites found in the CpG island.

To determine the methylation status of CpG sites, a total of 1,716 clones of CpG island-containing fragments were obtained and sequenced using BSP method. And the average clone number for each individual was 15.3 (ranging from 11 to 20), which will ensure the accuracy of methylation status of CpG sites. All sequences were analyzed using the BIQ Analyzer software for quality control and visualization of methylation status. The detailed methylation status for the two breeds is presented in Figure 2A, 2B. Overall, the CpG island population methylation levels were 89.51%±6.17% and 85.73%±9.94% in DP and Landrace piglets, respectively. Except CpG_2 site, all the other CpG sites of *CD4* gene were hypermethylated (Figure 2C). Compared between the two breeds, population methylation frequency of CpG_2 in DP piglets was significantly higher than that in Landrace piglets at 0.01 level (37.20% ±4.03% in DP vs 1.71%±0.60% in Landrace, $p<0.01$), and no significant difference were found for the other eight CpG sites (Figure 2C).

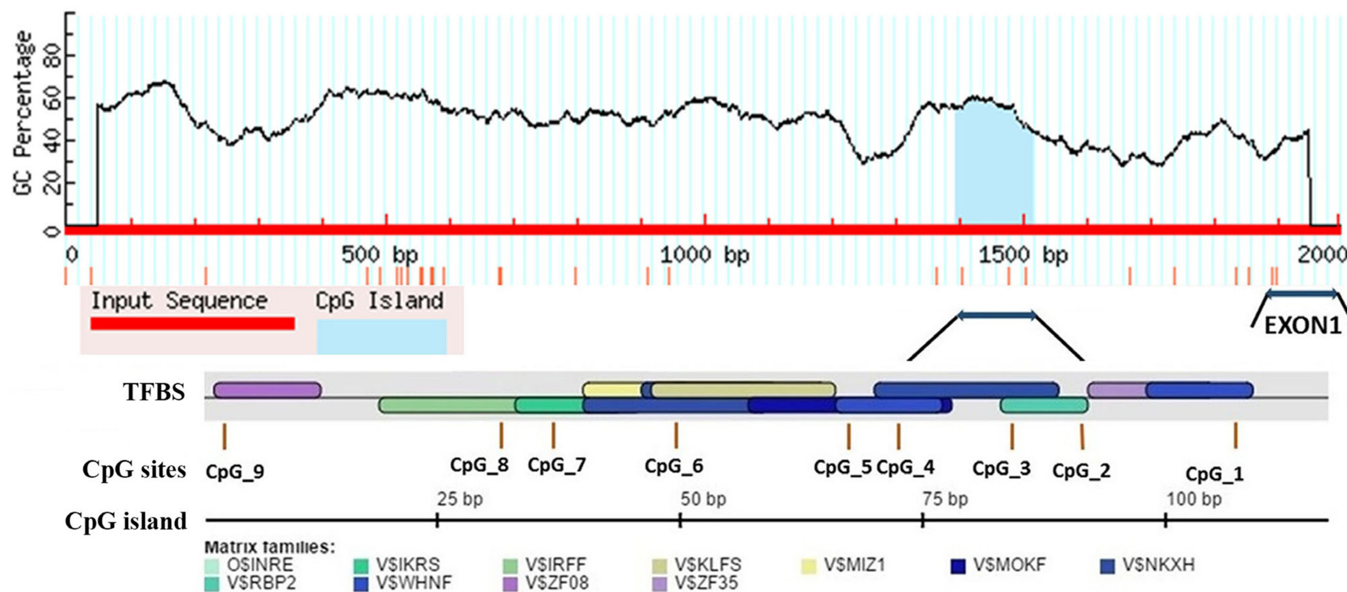


**Figure 1.** Bioinformatic analyses of the CpG island of porcine cluster of differentiation 4 protein (*CD4*) gene upstream –2 kb and exon one region. This figure only demonstrated 11 of 15 TFBS containing CpG sites. TFBS, transcription factor binding sites; Matrix Families: similar and/or functionally related TFBS are grouped into so-called matrix families.

### *CD4* gene expression and T lymphocyte subpopulations traits

The statistical description of *CD4* expression and T lymphocyte subpopulation traits are shown in Table 1. It can be seen that *CD4* gene was expressed at a low level with the ΔCt values ranging between 3.88 and 8.17. Compared between the two breeds, ΔCt value of *CD4* was significantly higher in DP piglets than that in Landrace piglets ($p<0.01$), suggesting that the expression of *CD4* was significantly lower in DP piglets than that in Landrace piglets at 0.01 level. So far, though it is unclear the relationship between *CD4* expression and disease resistance, some studies revealed that overexpression of *CD4* gene may lead to decrease of receptor signaling competence [30] and affect T lymphocytes development [4]. On the other hand, for the T lymphocyte subpopulation traits, except CD4+CD8-CD3-, CD4+CD8+CD3-, and CD3+, the other seven traits were significantly different at 0.05 or 0.01 level between the two breeds.

Additionally, comparing the variation within-population, the variation coefficients in DP (the average is 63.32%) were larger than that in Landrace (the average is 55.96%) for *CD4* expression and most of the T lymphocyte subpopulations traits, which is consistent with the fact that DP is one indigenous breed with less selection pressure compared with Landrace. Finally, we analyzed the sow effects on all the traits detected. The piglets used in the study were sampled from 13 DP sows and 28 Landrace sows. And the results indicated that the sow effects were significant at 0.01 level on all the T lymphocyte traits as well as *CD4* expression.

### Correlation analyses between *CD4* methylation status and mRNA expression

To investigate the correlation between *CD4* gene expression and CpG island methylation level, Pearson correlation analyses were calculated for DP, Landrace and combined population of these two breeds respectively. The results showed that the methylation frequencies of CpG_2, CpG_3, and CpG_7 correlated negatively with CD4 mRNA expression in all three groups (Figure 3). Besides, the correlation coefficient of CpG_2 reached statistical significance in DP and Landrace combined population ($r = -0.28$, $p = 3.1\times10^{-3}$, Figure 4), while it did not reach significant level ($p>0.05$) in DP or Landrace piglets separately. It may be due to the large sample size after combining the two breed samples together. Previous studies demonstrated that, in each CpG island, only a few CpG sites may be critical for gene expression regulation [31,32]. These results suggested that CpG_2 was a critical methylation site influencing *CD4* gene expression.

Methylation within CpG islands regulates gene transcription though a variety of mechanisms, and TFs is essential one of them. In humans and mouse, the TFs of the *CD4* gene, including Myb, Elf, and Ikaros, have been found [4]. To our knowledge, only TF nuclear factor-kappa B has been detected in the promoter region of porcine *CD4* gene in the previous studies [20,33], which indicated that maybe other TFs binding *CD4* promoter region are still not be found. Therefore, we applied TFBS prediction program MatInspector to infer the potential binding TFs. The elaborate results for TFBS are provided in Table 2. Totally, there were 15 putative TFBS identified,

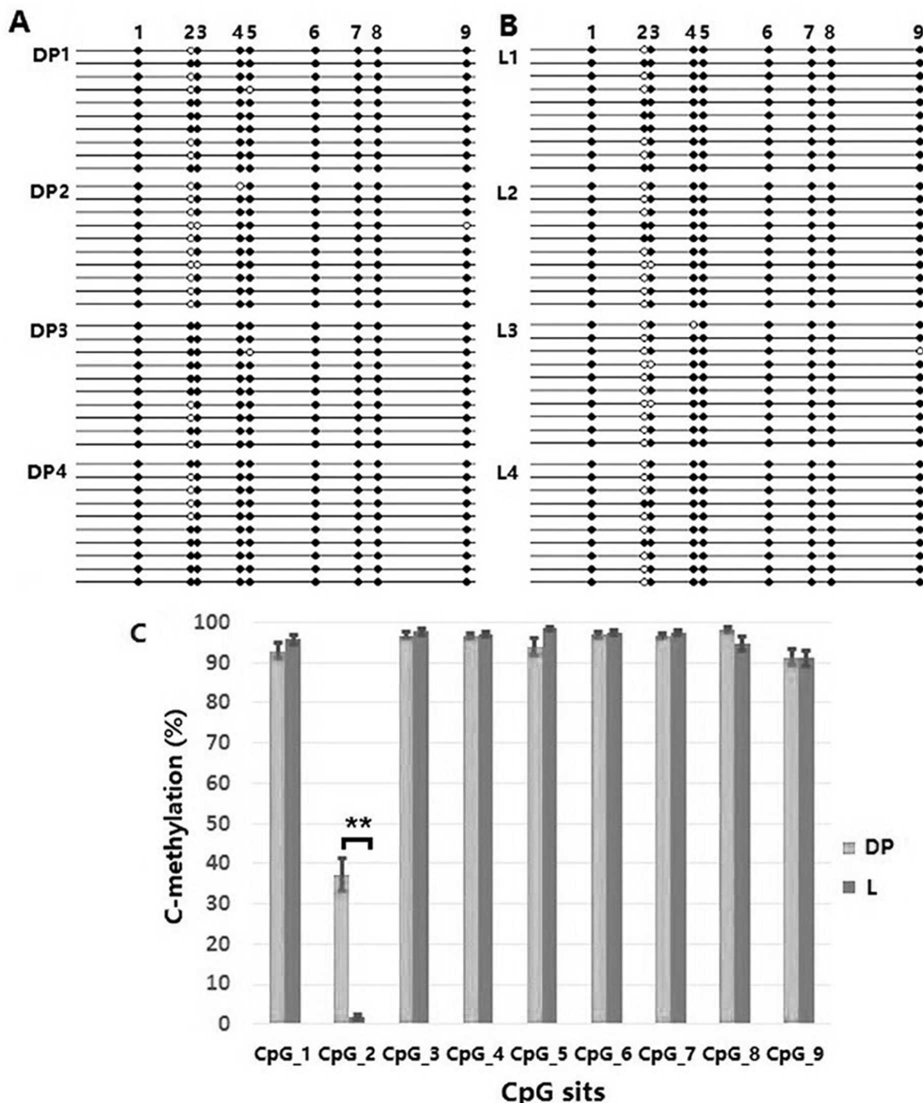


**Figure 2.** Methylation status of porcine cluster of differentiation 4 protein (*CD4*) gene CpG island. (A) CpG island methylation status of *CD4* gene in four random piglets of Dapulian. (B) CpG island methylation status of CD4 gene in four random piglets of Landrace. In picture (A) and (B) blackcircles correspond to methylated CpG sites, and white circles correspond to unmethylated CpG sites. Each line represents an independent clone. DP, Dapulian; L, Landrace. (C) Population methylation frequencies of *CD4* gene CpG island in Dapulian (n = 45) and Landrace (n = 48) groups. Y coordinate is average percentage of methylated cytosines (mean±standard error) for the different CpG dinucleotides in each population. ** represents significant differences of methylation frequencies between two groups in this CpG site at 0.01 level.

**Table 1.** Statistical description and least square mean analyses of *CD4* gene expression and T lymphocyte subpopulation traits in piglets of Dapulian and Landrace

| Traits | Dapulian | | | | | | | Landrace | | | | | | | Least square analyses | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | No. of sample | No. of sow | Mean | Std Dev | Coeff. of Variation | Maximum | Minimum | No. of sample | No. of sow | Mean | Std Dev | Coeff. of Variation | Maximum | Minimum | Breed | Sow |
| CD4 Expression (ΔCt) | 104 | 13 | 5.93 | 0.81 | 13.65 | 8.17 | 3.88 | 171 | 28 | 5.58 | 0.67 | 11.98 | 7.21 | 3.96 | 1.08e-4 | 2.23e-4 |
| CD4-CD8-CD3- | 124 | 13 | 25.92 | 8.40 | 32.42 | 61.35 | 8.60 | 187 | 28 | 24.48 | 6.95 | 28.40 | 46.02 | 9.97 | 0.0381 | < 0.0001 |
| CD4+CD8-CD3- | 124 | 13 | 0.35 | 0.64 | 184.05 | 5.91 | 0.00 | 187 | 28 | 0.38 | 0.44 | 117.35 | 3.40 | 0.00 | 0.5563 | < 0.0001 |
| CD4-CD8+CD3- | 124 | 13 | 6.72 | 4.85 | 72.20 | 29.04 | 0.70 | 187 | 28 | 8.67 | 5.05 | 58.28 | 28.44 | 1.32 | 0.0002 | < 0.0001 |
| CD4+CD8+CD3- | 124 | 13 | 0.14 | 0.22 | 162.40 | 1.38 | 0.00 | 187 | 28 | 0.15 | 0.26 | 173.83 | 2.55 | 0.00 | 0.6928 | 0.0004 |
| CD4-CD8-CD3+ | 124 | 13 | 14.90 | 6.31 | 42.37 | 33.30 | 1.35 | 187 | 28 | 18.86 | 5.94 | 31.49 | 34.16 | 3.08 | < 0.0001 | < 0.0001 |
| CD4+CD8-CD3+ | 124 | 13 | 25.69 | 6.25 | 24.34 | 40.20 | 12.51 | 187 | 28 | 29.85 | 5.96 | 19.98 | 43.81 | 17.56 | < 0.0001 | < 0.0001 |
| CD4-CD8+CD3+ | 124 | 13 | 25.05 | 7.46 | 29.77 | 51.70 | 10.81 | 187 | 28 | 17.08 | 5.22 | 30.59 | 33.45 | 7.94 | < 0.0001 | < 0.0001 |
| CD4+CD8+CD3+ | 124 | 13 | 1.20 | 0.98 | 81.52 | 7.54 | 0.20 | 187 | 28 | 0.53 | 0.47 | 89.27 | 2.83 | 0.02 | < 0.0001 | < 0.0001 |
| CD3+ | 124 | 13 | 66.84 | 8.38 | 12.53 | 82.13 | 35.09 | 187 | 28 | 66.32 | 7.56 | 11.39 | 82.38 | 44.79 | 0.49 | < 0.0001 |
| CD4+/CD8+ | 124 | 13 | 1.12 | 0.47 | 41.32 | 2.36 | 0.29 | 187 | 28 | 1.95 | 0.84 | 42.95 | 5.01 | 0.74 | < 0.0001 | < 0.0001 |

*CD4*, cluster of differentiation 4 protein.

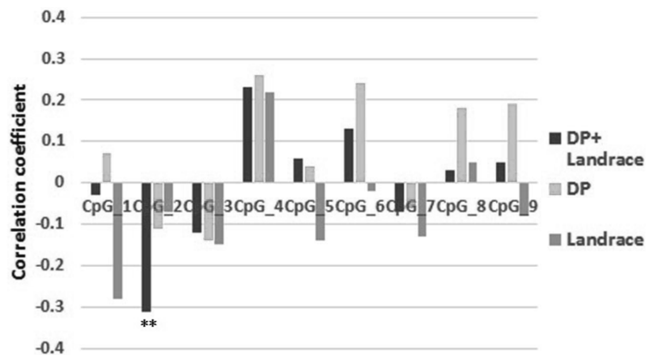


**Figure 3.** Correlation analyses between methylation levels and mRNA expression at different CpG sites of cluster of differentiation 4 protein (*CD4*) gene CpG island. DP, Landrace and DP+Landrace represent Dapulian (n = 45), Landrace (n = 48), and Dapulian and Landrace combined groups (n = 93), respectively. ** represents p value of correlation coefficient reach 0.01.

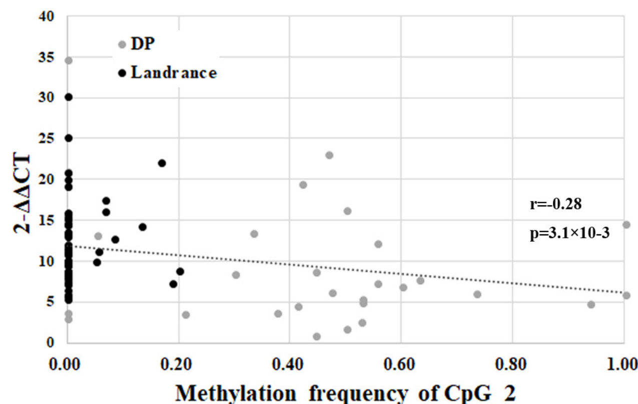


**Figure 4.** CpG_2 methylation significantly correlated with cluster of differentiation 4 protein (*CD4*) gene expression in Dapulian and Landrace combined piglets. Pearson's correlation coefficient with corresponding p values for the correlation between methylation frequency of CpG_2 (x axis) and mRNA expression of *CD4* gene (y axis) in Dapulian and Landrace combined population. Decreased expression of *CD4* gene correlated with CpG_2 methylation frequency ($p<0.01$). Correlation coefficient trend line is shown in broken line. Gray dots represent Dapulian piglets, while black ones represent Landrace piglets.

and 11 of them contained CpG sites (Figure 1). Especially, one TF, Jumonji, contained the CpG_2, suggesting that it may influence the *CD4* gene expression through the potential binding of the predicted TFs. These identified TFBS will provide reference information for further digging out more TFs binding porcine *CD4* promoter.

## Correlation analyses between *CD4* methylation status and T lymphocyte subpopulations

To investigate whether CpG sites methylation status affecting T lymphocyte subpopulations, we implemented correlation analyses between CpG sites methylation frequencies and T lymphocyte subpopulation in DP, Landrace and combined population of these two breeds. The detailed correlation analyses results are provided in Table 3. The results showed that each CpG site methylation frequency was significant correlation with one or more traits of T lymphocyte subpopulations, except CpG_6 and CpG_9. Among these CpG sites, we founded CpG_2 methylation frequency was significantly positive or negative correlation with T lymphocytes subpopulations at CD4+CD8-CD3- ($r = -0.29$), CD4-CD8+CD3- ($r = -0.27$), CD4-CD8+CD3+ ($r = 0.36$), CD4+CD8+CD3+ ($r = 0.46$),

**Table 2.** Information of transcription factor binding sites

| Matrix family[1)] | Detailed matrix information | Start position[2)] | End position[2)] | Matrix similarity[3)] |
|---|---|---|---|---|
| V$ZF08 | KRAB-zinc finger protein synten (KID3) | 2 | 12 | 0.941 |
| V$ZFHX | AREB6 (Atp1a1 regulatory element binding factor 6) | 20 | 32 | 0.974 |
| V$IRFF | Interferon regulatory factor 4 | 19 | 43 | 0.936 |
| V$IKRS | Ikaros 2, potential regulator of lymphocyte differentiation | 33 | 45 | 0.921 |
| V$MIZ1 | Myc-interacting Zn finger protein 1, zinc finger and BTB domain containing 17 (ZBTB17) | 40 | 50 | 0.975 |
| V$NKXH | Homeodomain factor Nkx-2.5/Csx | 40 | 58 | 0.919 |
| O$INRE | Drosophila initiator motifs | 49 | 59 | 0.906 |
| V$NKXH | Homeodomain protein NKX3.2 (BAPX1, NKX3B, Bagpipe homolog) | 46 | 64 | 0.948 |
| V$KLFS | Krueppel-like factor 12 (AP-2rep) | 47 | 65 | 0.922 |
| V$MOKF | Ribonucleoprotein associated zinc finger protein MOK-2 (human) | 57 | 77 | 0.97 |
| V$WHNF | Winged helix protein, involved in hair keratinization and thymus epithelium differentiation | 66 | 76 | 0.907 |
| V$NKXH | Homeodomain protein NKX3.2 (BAPX1, NKX3B, Bagpipe homolog) | 70 | 88 | 0.903 |
| V$RBP2 | Jumonji, AT rich interactive domain 1B | 83 | 91 | 0.935 |
| V$ZF35 | Human zinc finger protein ZNF35 | 92 | 104 | 0.922 |
| V$WHNF | Winged helix protein, involved in hair keratinization and thymus epithelium differentiation | 98 | 108 | 0.941 |

1) Similar and/or functionally related transcription factor binding sites are grouped into so-called matrix families.
2) Start and end position were the positions where TFBS located at in *CD4* CpG island.
3) The matrix similarity is calculated as described in the MatInspector papers, a perfect match to the matrix gets a score of 1.00, a "good" match to the matrix usually has a similarity of >0.80.
In this examination, the matrix similarity was set at 0.90.

**Table 3.** CpG sites significantly correlated with T lymphocyte subpopulation traits in Dapulian and Landrace piglets[1)]

| CpG sites | Breed | CD3- | | | | CD3+ | | | | CD3+ | CD4+/CD8+ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | CD4-CD8- | CD4+CD8- | CD4-CD8+ | CD4+CD8+ | CD4-CD8- | CD4+CD8- | CD4-CD8+ | CD4+CD8+ | | |
| CpG_1 | L | 0.00 | –0.15 | 0.06 | 0.06 | –0.04 | 0.09 | –0.13 | –0.26* | –0.03 | 0.17 |
| CpG_2 | DP+L | –0.05 | –0.29** | –0.27** | –0.06 | –0.02 | –0.10 | 0.36** | 0.46** | 0.21* | –0.28** |
| | DP | –0.09 | –0.33* | –0.14 | –0.07 | 0.13 | –0.03 | 0.04 | 0.17 | 0.17 | –0.04 |
| | L | 0.01 | 0.11 | –0.15 | 0.28* | 0.17 | 0.05 | –0.17 | 0.06 | 0.08 | 0.18 |
| CpG_3 | DP+L | 0.07 | 0.18 | 0.03 | –0.12 | 0.00 | 0.03 | –0.12 | –0.2* | –0.08 | 0.09 |
| | DP | 0.21 | 0.19 | 0.01 | –0.40** | –0.03 | –0.01 | –0.15 | –0.31* | –0.20 | 0.18 |
| CpG_4 | DP+L | 0.07 | –0.08 | 0.04 | –0.17 | 0.01 | –0.09 | –0.02 | –0.10 | –0.08 | –0.04 |
| | L | 0.13 | –0.15 | 0.02 | –0.33* | –0.05 | –0.09 | –0.01 | –0.07 | –0.12 | –0.04 |
| CpG_5 | DP+L | –0.20* | –0.23* | 0.13 | 0.03 | 0.07 | 0.01 | 0.06 | 0.03 | 0.11 | –0.08 |
| | DP | –0.34* | –0.40** | 0.14 | 0.09 | 0.05 | –0.01 | 0.20 | 0.25 | 0.26 | –0.28* |
| | L | –0.06 | –0.43** | 0 | –0.15 | 0.00 | –0.1 | 0.30* | 0.04 | 0.07 | –0.28* |
| CpG_7 | L | 0.00 | –0.27* | –0.08 | –0.01 | 0.09 | –0.15 | 0.23 | 0.12 | 0.06 | –0.24 |
| CpG_8 | DP+L | 0.03 | –0.29** | 0.06 | 0.05 | –0.01 | –0.12 | 0.05 | 0.03 | –0.05 | –0.10 |
| | L | 0.04 | –0.32* | 0.14 | 0.09 | 0.01 | –0.15 | 0.02 | –0.05 | –0.10 | –0.08 |

L, Landrace; DP, Dapulian; DP+L, combined population of Dapulian and Landrace.
[1)] The value in Table 3 above represented correlation coefficient between CpG sites methylation frequencies and T lymphocyte subpopulation traits.
* represents p value of correlation coefficient reaches 0.05; ** represents p value of correlation coefficient reaches 0.01.

CD4+/CD8+ (r = –0.28) (for all correlation, p<0.01) in DP and Landrace combined population (Figure 5). In peripheral blood, CD4+CD8-CD3- cells represent double positive cells lacking CD8. CD4+CD8-CD3- should develop with comparable kinetics as the CD4+CD8+ double positive cells [34]. CD4-CD8+CD3- cells represent NK cells, which protect the body against infections by killing target cells and secreting inflammatory cytokine [35]. The CD4+/CD8+ ratio is the most useful marker of disease. CpG_2 methylation frequency was significantly negative correlation with CD4+CD8-CD3-, CD4-CD8+CD3-, and CD4+/CD8+. The results revealed that hypomethylation of CpG_2 site may lead to the increase proportion of CD4+CD8-CD3-, CD4-CD8+CD3-, and CD4+/CD8+. Meanwhile, CD4-CD8+ CD3+ and CD4+CD8-CD3+

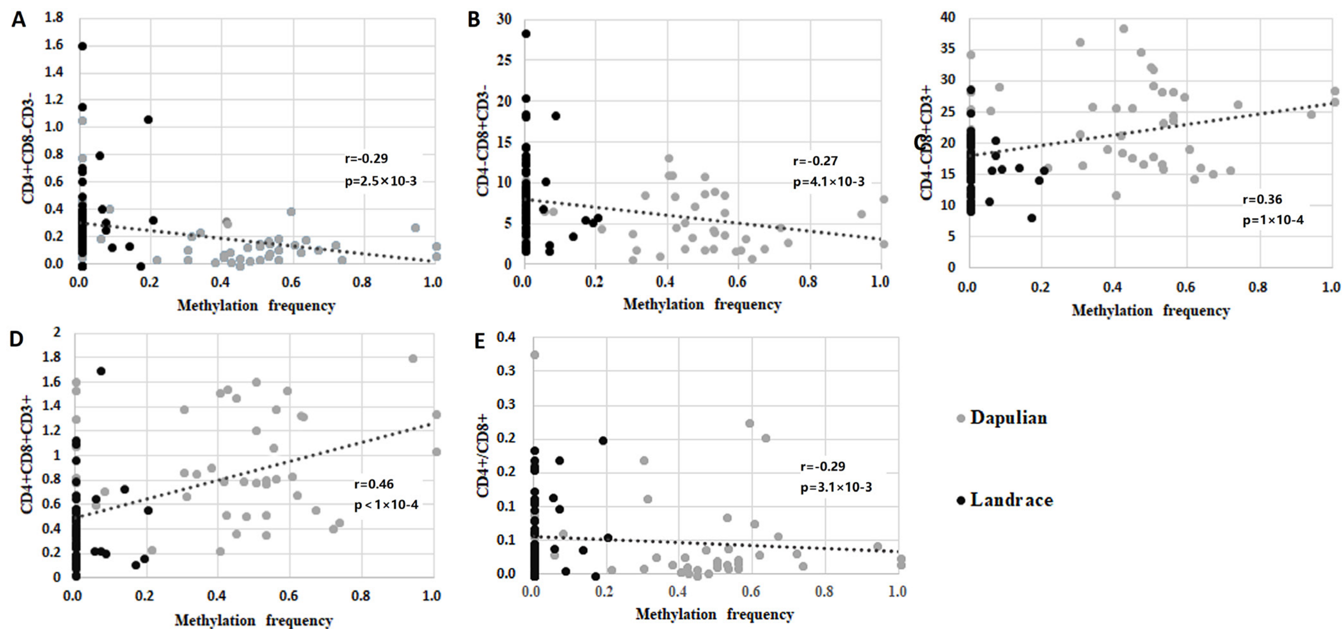


**Figure 5.** CpG_2 methylation significantly correlated with T lymphocyte subpopulation traits in Dapulian and Landrace combined piglets. Pearson's r correlation coefficient with corresponding p values for the correlation between methylation frequency of CpG_2 (x axis) and CD4+CD8-CD3- (y axis, A), CD4-CD8+CD3- (y axis, B), CD4-CD8+CD3+ (y axis, C), CD4+CD8+CD3+ (y axis, D), and CD4+/CD8+ (y axis, E) in Dapulian and Landrace combined population. Decreased of T-lymphocyte subpopulation traits CD4+CD8-CD3-, CD4-CD8+CD3-, and CD4+/CD8+ significantly correlated with methylation frequency of CpG_2 (p<0.01). Increase of traits CD4-CD8+CD3+ and CD4+CD8+CD3+ significantly correlated with methylation frequency of CpG_2 (p<0.01). Correlation coefficient trend line is shown in broken line. Gray dots represent Dapulian piglets. Black ones represent Landrace piglets.

represent cytotoxic T lymphocytes (CTLs) and helper T lymphocytes (Th), respectively. CTLs are responsible for killing antigen-bearing target cells, such as virus-infected cells, which are often dependent on 'help' from Th cells [36]. The significantly positively correlation between CpG_2 methylation frequency with CD4-CD8+CD3+ and CD4+CD8+CD3+ implied that CpG_2 methylation may lead to the decrease number of CD4-CD8+CD3+ and CD4+CD8+CD3+.

The negative correlation of CpG_5 methylation frequency with CD4+CD8-CD3- were consistent in the three groups, DP ($p<0.05$), Landrace ($p<0.01$) and combined population of these two breeds ($p<0.01$) (Table 3). Although the methylation difference at CpG_5 sites did not reach significant level ($p>0.05$), CpG_5 site can be used as a methylation marker for CD4+CD8-CD3- of T lymphocyte subpopulations.

## CONCLUSION

In this study, we determined porcine *CD4* gene CpG island methylation level and conducted correlation analyses of CpG sites methylation frequencies with the gene expression and T lymphocyte subpopulations. We found that CpG_2 site correlated negatively with CD4 mRNA expression, which may influence the *CD4* gene expression through the potential binding predicted TFs. We also found that CpG_2 methylation frequency was significantly positive or negative correlated with several T lymphocytes subpopulation traits. Thus, the CpG_2 was a critical methylation site for porcine *CD4* gene expression and T lymphocyte subpopulation traits. We speculated that increased methylation frequency of CpG_2 may lead to the decreased expression of *CD4*, which may have some kind of influence on T lymphocyte subpopulation traits and the immunity of DP population.

## AUTHOR CONTRIBUTIONS

JW conceived and designed the experiments. XZ and YW carried out experiment and computational analysis. XZ and JW wrote the manuscript. JG contributed to the sample collecting and interpretation of data. All authors read and approved the final manuscript.

## CONFLICT OF INTEREST

We certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

## ACKNOWLEDGMENTS

This work was financed by the National Major Special Project of China on New Varieties Cultivation for Transgenic Organisms (2014ZX0800945B), National Natural Science Foundations of China (31372293), Shandong Swine Industry Technology System Innovation (SDAIT-08-03), Natural Science Foundations of Shandong (ZR2016CB15).

## REFERENCES

1. Fabbri M, Smart C, Pardi R. T lymphocytes. Int J Biochem Cell Biol 2003;35:1004-8.
2. Davis CB, Littman DR. Thymocyte lineage commitment: is it instructed or stochastic? Curr Opin Immunol 1994;6:266-72.
3. Killeen N, Davis CB, Chu K, et al. CD4 function in thymocyte differentiation and T cell activation. Philos Trans R Soc Lond B Biol Sci 1993;342(1299):25-34.
4. Ellmeier W, Sawada S, Dan R. The regulation of CD4 and CD8 coreceptor gene expression during T cell development. Annu Rev Immunol 2003;17:523-54.
5. Adlam M, Siu G. Hierarchical interactions control *CD4* gene expression during thymocyte development. Immunity 2003; 18:173-84.
6. Zou Y, Sunshine M, Taniuchi I, et al. Epigenetic silencing of CD4 in T cells committed to the cytotoxic lineage. Nat Genet 2001;29:332-6.
7. Xu J, Liu Y, Fu W, et al. Association of the porcine cluster of differentiation 4 gene with T lymphocyte subpopulations and its expression in immune tissues. Asian-Australas J Anim Sci 2013;26:463-9.
8. Jones PA, Takai D. The role of DNA methylation in mammalian epigenetics. Science 2001;293(5532):1068-70.
9. Ioshikhes IP, Zhang MQ. Large-scale human promoter mapping using CpG islands. Nat Genet 2000;26:61-3.
10. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci USA 2006; 103:1412-7.
11. Wachter E, Quante T, Merusi C, et al. Synthetic CpG islands reveal DNA sequence determinants of chromatin structure. Elife 2014;3:e03397.
12. Jursch T, Miskey C, Izsvák Z, Ivics Z. Regulation of DNA transposition by CpG methylation and chromatin structure in human cells. Mob DNA 2013;4:15.
13. Choy JS, Wei S, Lee JY, et al. DNA methylation increases nucleosome compaction and rigidity. J Am Chem Soc 2010;132: 1782-3.
14. Fang J, Cheng J, Wang J, et al. Hemi-methylated DNA opens a closed conformation of UHRF1 to facilitate its histone recognition. Nat Commun 2016;7:11197.
15. Banovich NE, Lan X, McVicker G, et al. Methylation QTLs are associated with coordinated changes in transcription factor binding, histone modifications, and gene expression levels. PLoS Genet 2014;10(9):e1004663.

16. Dai C, Sun L, Xia R, et al. Correlation between the methylation of the FUT1 promoter region and FUT1 expression in the duodenum of piglets from newborn to weaning. 3 Biotech 2017;7:247.
17. Yan J, Fu H, Shen J, et al. Application of bisulfite sequencing PCR in detecting the abnormal methylation of suppressor gene of wnt signaling pathway in acute promyelocytic leukemia. Zhongguo Shi Yan Xue Ye Xue Za Zhi 2016;24:1299-304.
18. Luo J, Ying Y, Zhang H, et al. Down-regulation of promoter methylation level of *CD4* gene after MDV infection in MD-susceptible chicken line. BMC Proc 2011;Suppl 4:S7.
19. Wang XS, Zhang Y, He YH, et al. Aberrant promoter methylation of the CD4 gene in peripheral blood cells of mastitic dairy cows. Genet Mol Res 2013;12:6228-39.
20. Wang X, Ao H, Zhai L, et al. Transcriptional regulation of CD4 gene expression in porcine kidney epithelial cells by virus-like double-stranded RNA and DNA methyltransferase inhibitor. Genet Mol Res 2014;13:3346-55.
21. Oppmann B, Lesley R, Blom B, et al. Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with biological activities similar as well as distinct from IL-12. Immunity 2000;13:715-25.
22. Lunney JK, Chen H. Genetic control of host resistance to porcine reproductive and respiratory syndrome virus (PRRSV) infection. Virus Res 2010;154:161-9.
23. Yang H. Livestock development in China: animal production, consumption and genetic resources. J Anim Breed Genet 2013;130:249-51.
24. Jiang C, Xing F, Xing J, Jiang Y, Zhou E. Different expression patterns of PRRSV mediator genes in the lung tissues of PRRSV resistant and susceptible pigs. Dev Comp Immunol 2013;39:127-31.
25. Wang HZ, Lin S, Wang YP, et al. Analyses of breeds and litter effect on routine blood traits and T-lymphocyte subpopulation traits in piglets. Chinese J Anim Vet Sci 2014;22:724-30.
26. Wang JY, Wang YP, Guo JF, et al. Selection of reference genes and determination of cytokines and receptor mRNA expression in peripheral blood of piglets. Scientia Agricultura Sinica 2015;48:1437-44.
27. Cartharius K, Frech K, Grote K, et al. MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics 2005;21:2933-42.
28. Li L-C, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics 2002;18:1427-31.
29. Bock C, Reither S, Mikeska T, et al. BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. Bioinformatics 2005;21:4067-8.
30. Nakayama T, Wiest DL, Abraham KM, et al. Decreased signaling competence as a result of receptor overexpression: overexpression of CD4 reduces its ability to activate p56lck tyrosine kinase and to regulate T-cell antigen receptor expression in immature CD4+CD8+ thymocytes. Proc Natl Acad Sci USA 1993;90:10534-8.
31. Jones PA, Takai D. The role of DNA methylation in mammalian epigenetics. Science 2001;293(5532):1068-70.
32. Meng H, Murrelle EL, Li G. Identification of a small optimal subset of CpG sites as bio-markers from high-throughput DNA methylation profiles. BMC Bioinformatics 2008;9:457.
33. Bensaude E, Turner JL, Wakeley PR, et al. Classical swine fever virus induces proinflammatory cytokines and tissue factor expression and inhibits apoptosis and interferon synthesis during the establishment of long-term infection of porcine vascular endothelial cells. J Gen Virol 2004;85:1029-37.
34. Chi T H, Wan M, Zhao K, et al. Reciprocal regulation of CD4/CD8 expression by SWI/SNF-like BAF complexes. Nature 2002;418:195-9.
35. Liu KY, Comstock SS, Shunk JM, Monaco MH, Donovan SM. Natural killer cell populations and cytotoxic activity in pigs fed mother's milk, formula, or formula supplemented with bovine lactoferrin. Pediatr Res 2013;74:402-7.
36. Bennett SRM, Carbone FR, Karamalis F, et al. Induction of a $CD8^+$ cytotoxic T lymphocyte response by cross-priming requires cognate $CD4^+$ help. J Exp Med 1997;186:65-70.