

# Transgene Expression Is Associated with Copy Number and Cytomegalovirus Promoter Methylation in Transgenic Pigs

Qingran Kong<sup>1</sup>, Meiling Wu<sup>1</sup>, Yanjun Huan<sup>1</sup>, Li Zhang<sup>1</sup>, Haiyan Liu<sup>2\*</sup>, Gerelchimeg Bou<sup>1</sup>, Yibo Luo<sup>1</sup>, Yanshuang Mu<sup>1</sup>, Zhonghua Liu<sup>1</sup>

**1** College of life science, Northeast Agricultural University of China, Harbin, People's Republic of China, **2** College of Medicine and Pharmaceutics, Ocean University of China, Qingdao, People's Republic of China

## Abstract

Transgenic animals have been used for years to study gene function, produce important proteins, and generate models for the study of human diseases. However, inheritance and expression instability of the transgene in transgenic animals is a major limitation. Copy number and promoter methylation are known to regulate gene expression, but no report has systematically examined their effect on transgene expression. In the study, we generated two transgenic pigs by somatic cell nuclear transfer (SCNT) that express green fluorescent protein (GFP) driven by cytomegalovirus (CMV). Absolute quantitative real-time PCR and bisulfite sequencing were performed to determine transgene copy number and promoter methylation level. The correlation of transgene expression with copy number and promoter methylation was analyzed in individual development, fibroblast cells, various tissues, and offspring of the transgenic pigs. Our results demonstrate that transgene expression is associated with copy number and CMV promoter methylation in transgenic pigs.

**Citation:** Kong Q, Wu M, Huan Y, Zhang L, Liu H, et al. (2009) Transgene Expression Is Associated with Copy Number and Cytomegalovirus Promoter Methylation in Transgenic Pigs. PLoS ONE 4(8): e6679. doi:10.1371/journal.pone.0006679

**Editor:** Joanna Mary Bridger, Brunel University, United Kingdom

**Received:** March 6, 2009; **Accepted:** July 14, 2009; **Published:** August 18, 2009

**Copyright:** © 2009 Kong et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by grant from the State Transgenic Research Programme of China (Grant No. 2008ZX08006-002) and Innovative Research Team of NEAU. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: liu086@yahoo.com

## Introduction

Transgenic animals are a powerful tool in the fields of experimental and applied biology. These animals allow study into the function and regulation of genes *in vivo*, the production of important pharmaceutical proteins, and the creation of pathologic models for human disease therapy [1]. Recent progress in animal cloning has provided an attractive alternative to improve transgenic efficiency, through the combination of transfection and somatic cell nuclear transfer (SCNT).

To date, the cloning of pigs has been used successfully to produce transgenic animals expressing enhance green fluorescence protein (eGFP) [2] and omega-3 fatty acids [3], and an  $\alpha$ -1,3-galactosyltransferase deficient pig, which has the potential to be used as an organ donor for xenotransplantation [4]. The production of genetically modified pigs by nuclear transfer has progressed from basic research to practical use. Despite this impressive and growing success, transgenesis still suffers from many limitations. Numerous experiments have shown that the inheritance and expression of the transgene in transgenic animals is predictable only to a limited extent [5–8]. In most cases, transgene expression levels in transfected cells often decline with time [6]. Furthermore, the level of transgene expression appears to correlate inversely with time [9], and the majority of transgenic animals cannot stably pass the transgene to their offspring. Therefore, it is difficult to select founder transgenic animals to establish a line of transgenic animals [7]. Recently, several reports

have demonstrated that induced pluripotent stem cells (iPS) may possess mechanisms to lower the expression level of the four factors (Oct3/4, Sox2, c-Myc, and Klf4) and make them expression silencing after full-reprogramming [10–12].

Mechanisms causing this phenomenon of transgene instability are poorly understood. Generally, transgene copy number and DNA methylation status can influence transgene expression [13]. These are often considered as main factors resulting in incomplete and complete silencing of transgene expression [14]. In most cases, multiple copies of the transgene, arrayed in a head-to-tail manner, are randomly integrated in the host genome, which may cause transcriptional interference that represses expression [15–17]. As long as the transgene is expressed appropriately, calculating the copy number is not usually performed. DNA methylation is the strongest candidate for expression silencing, because it can lead to transcriptional inactivity of certain genes, may be stably inherited through mitosis, and may be transmitted to subsequent generations [18–20]. In particular, promoter methylation has been associated with transgene silencing *in vitro* and *in vivo* [21–23].

A more clear understanding of the factors influencing transgene expression would improve the production of transgenic animals. In order to test the relationship between transgene expression and copy number or promoter methylation, we generated two GFP transgenic pigs by SCNT and analyzed GFP expression, copy number and CMV methylation in regards to individual development, fibroblast cells, various tissues, and the offspring of the transgenic pigs. Our results suggest that transgene expression is

regulated by methylation of the promoter and by transgene copy number.

## Results

### Generation of transgenic pig

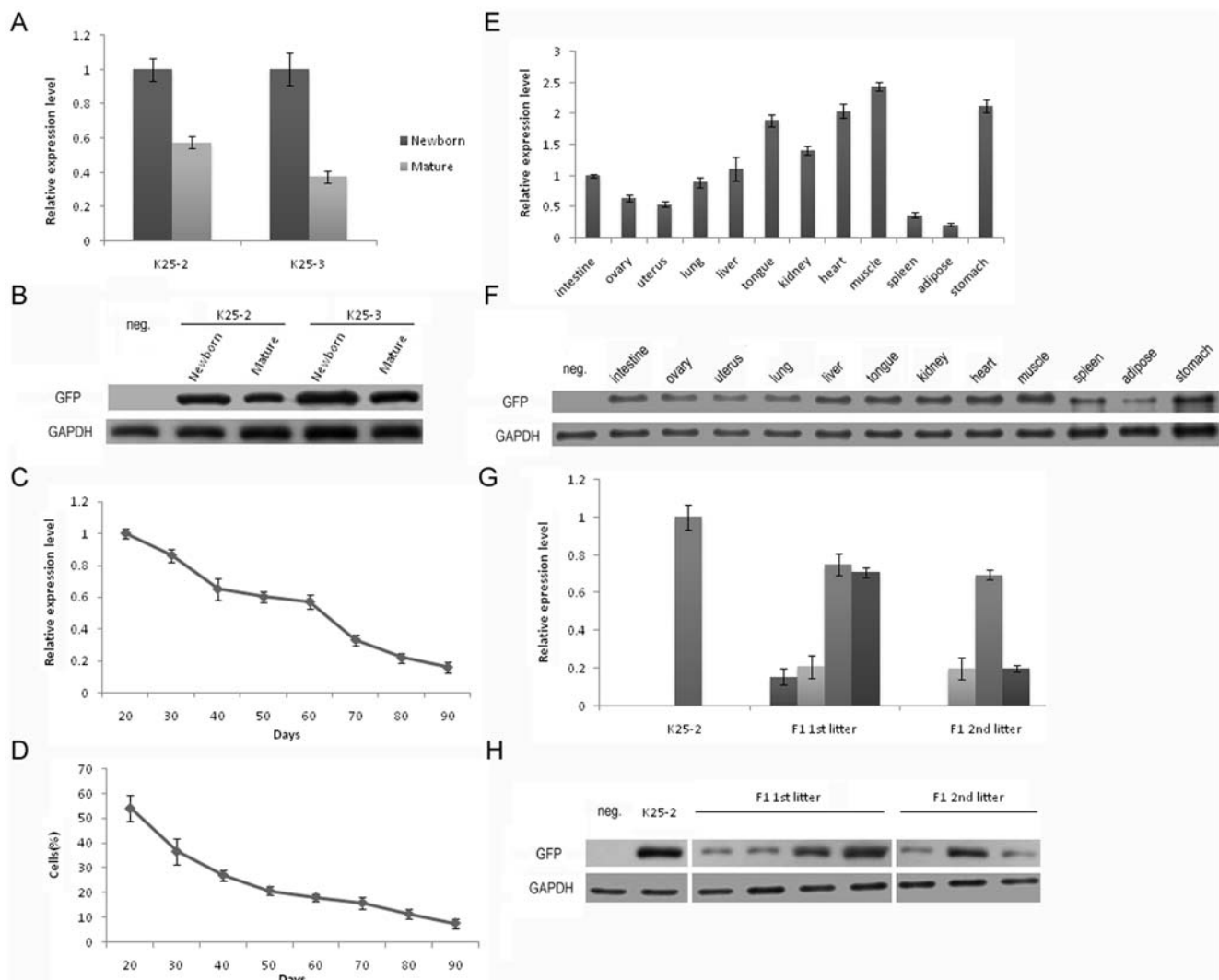
A total of 1978 reconstructed embryos were transferred to ten recipients. Three recipients became pregnant and four founder GFP-positive transgenic pigs were born at full term. Finally, two female pigs, named K25-2 and K25-3, survived to maturity (the others died at birth). After mating with non-transgenic pigs, K25-2 produced seven F1 positive pigs from two litters, but K25-3 died during delivery. Variable factors may decrease the success rate of producing transgenic animals, and an additional difficulty in pigs requires that at least four good embryos are needed to induce and

maintain pregnancy [3]. In the present study, the overall efficiency of transgenic pig production was 0.69%.

### Expression of GFP

We analyzed the change in GFP expression in relation to aging in ear tissues of transgenic pigs. A significant decline in the mRNA level was observed from newborn to maturity in both K25-2 and K25-3 ( $p < 0.001$ ). The mRNA level decreased about 1-fold and 3-fold in K25-2 and K25-3, respectively (Fig. 1A). These results were consistent with results from Western blot analysis (Fig. 1B). These data indicate that the transgene expression level decreases with aging *in vivo*.

In order to examine whether the expression of the transgene also declined *in vitro*, fibroblast cells from K25-3 were cultured *in vitro* up to 90 days were analyzed for transgene expression. A



**Figure 1. Expression of GFP.** (A) Relative real-time RT-PCR analysis of GFP mRNA expression from newborn to maturity in founder transgenic pigs. The decline was significant in both K25-2 and K25-3 ( $p < 0.001$ ); (B) Western blots analysis of GFP protein from newborn to maturity in founder transgenic pigs; (C) Relative real-time RT-PCR analysis of GFP mRNA expression in transgenic fibroblast cells. The decline of mRNA from 20 to 90 days was significant ( $p < 0.001$ ); (D) Flow cytometry analysis of the percentage of phenotype positive cells. The decline of percentage of positive cells from 20 to 90 days was also significant ( $p < 0.001$ ); (E) Relative real-time RT-PCR analysis of GFP mRNA expression in various tissues of the transgenic pig. Variagation of GFP mRNA expression was shown in different tissues ( $p < 0.001$ ); (F) Western blots analysis of GFP protein in various tissues; (G) Relative real-time RT-PCR analysis of GFP mRNA expression in offspring transgenic pigs. The decline from founder to offspring was significant ( $p < 0.001$ ); (H) Western blots analysis of GFP protein in offspring transgenic pigs. Error bars denote standard deviations. Neg., non-transgenic pig. doi:10.1371/journal.pone.0006679.g001

significant decline (more than 5-fold) in mRNA levels was observed from 20 to 90 days in culture ( $p < 0.001$ ) (Fig. 1C). Flow cytometry analysis also showed a significant decline (almost 8-fold) in the percentage of cells that expressed GFP ( $p < 0.001$ ) (Fig. 1D). A decrease in fluorescence intensity from 20 to 90 days was also found by flow cytometry analysis (data not shown). Previous reports have demonstrated that transgene silencing in transgenic cells may occur through a decline in expression levels rather than in the proportion of expressing cells [6,17,24], but the present data do not clarify this point, and a decrease in the percentage of positive cells was evident.

GFP expression was detected in different tissues of K25-3, namely intestine, ovary, uterus, lung, liver, tongue, kidney, heart, muscle, spleen, adipose and stomach. However, significant difference was found in the mRNA levels observed in these tissues ( $p < 0.001$ ). In tongue, heart, muscle and stomach, mRNA levels were almost 20-fold higher compared to that in spleen, adipose, ovary and uterus (Fig. 1E). The difference was confirmed by Western blot analysis (Fig. 1F). To examine the GFP expression pattern directly, different tissues were observed under UV light (Fig. 2), and a variegation of expression was observed. These results are consistent with previous studies in that transgene expression in transgenic animals may be different between tissues [25,26].

GFP expression was detected in offspring of K25-2, and was varied. The mRNA expression in offspring was significantly lower (more than 5-fold) compared to that in the founder ( $p < 0.001$ ) (Fig. 1G). Western blot analysis also showed a decline between K25-2 and its offspring (Fig. 1H). These findings suggest that the expression of the transgene in transgenic animals does not stably pass to their offspring.

### Copy number of GFP

In order to determine the correlation of transgene expression with copy number, we examined the GFP copy number in ears of newborn and mature transgenic pigs. A decline in copy number was found by absolute quantitative real-time PCR (Fig. 3A) and Southern blot analysis (Fig. 3B). Although a significant decline was not detected in K25-2 ( $p = 0.099$ ), the decline in K25-3 was significant ( $p = 0.016$ ). A more than 4-copy drop was observed in K25-2, and a more than 6-copy drop was observed in K25-3 (table 1). Transgenic animals show a difference in transgene copy number [27,28], but a decline in the copy number with aging is not a common observation.

We also analyzed the GFP copy number in transgenic fibroblast cells cultured *in vitro* for up to 90 days. The data showed that the GFP copy number declined significantly ( $p < 0.001$ ) and the copy number decreased almost 6-fold from 20 to 90 days (Fig. 3C). These results are consistent with the conclusion that transgene copies may be lost in transgenic cells over time *in vitro* [6]. In addition, we observed a significant correlation between GFP expression levels and copy number in transgenic fibroblast cells ( $r = 0.965$ ,  $p < 0.001$ ). These results suggest that the decline of transgene expression may be due to the loss of copies.

The GFP copy number in various tissues of K25-3 was also tested (Fig. 3D). Although the copy number in different tissues was varied, the difference was not significant ( $p = 0.059$ ) and not clearly shown by Southern blots (data not shown). In muscle, heart and lung, the GFP copy number was about 3 copies more than that in adipose, ovary and spleen. In addition, the correlation between GFP expression and copy number in different tissues was significant ( $r = 0.851$ ,  $p < 0.001$ ). In our earlier report, we demonstrated a mosaic genotype in a transgenic pig that died at birth. In that animal, we did not observe GFP sequence in muscle

by PCR analysis [29], but, in the present study, this phenomenon was not observed.

Transgene loss during passage is common [27,30]. In present study, the loss of GFP copies during passage of K25-2 was significant ( $p < 0.001$ ), and the GFP copy number in different offspring varied (Fig. 3E). Moreover, the observed loss of GFP copies was 15 copies at most and 9 copies at least (Table 2). Southern blots confirmed the decline (Fig. 3F). The correlation of GFP expression level with copy number was significant in this case ( $r = 0.864$ ,  $p = 0.006$ ). These results suggest that transgenic animals may transmit the transgene to their offspring only to a limited extent, and this loss of transgene copies is responsible for the expression decline.

### Methylation status of CMV promoter

Transgene expression may be regulated by copy number and repressed by DNA methylation. Therefore, we examined the methylation status of a 278-bp region of the CMV promoter containing one CpG island with 14 CpG sites. The bisulfite sequencing method is able to reveal the methylation status of all the cytosine residues in a DNA region of interest [31].

The bisulfite sequence data for a representative region of CMV promoter is shown in Fig. 4). On average, 99% of the cytosine residues in the pEGFP-C1 plasmid were converted to thymidine, indicating that the bisulfite conversion reaction on other samples was at least 99% efficient. The converted or unconverted cytosines at CpG sites (on a red background) indicate unmethylated or methylated. Interestingly, there were some unconverted cytosines at non-CpG sites (on a pink background), indicating cytosines that might be methylated. Extensive non-CpG methylation of the CMV promoter has been reported and associated with transgene silencing [22]. However, in this study, these cytosines were not included methylation level analysis because we did not know the mechanism of non-CpG methylation.

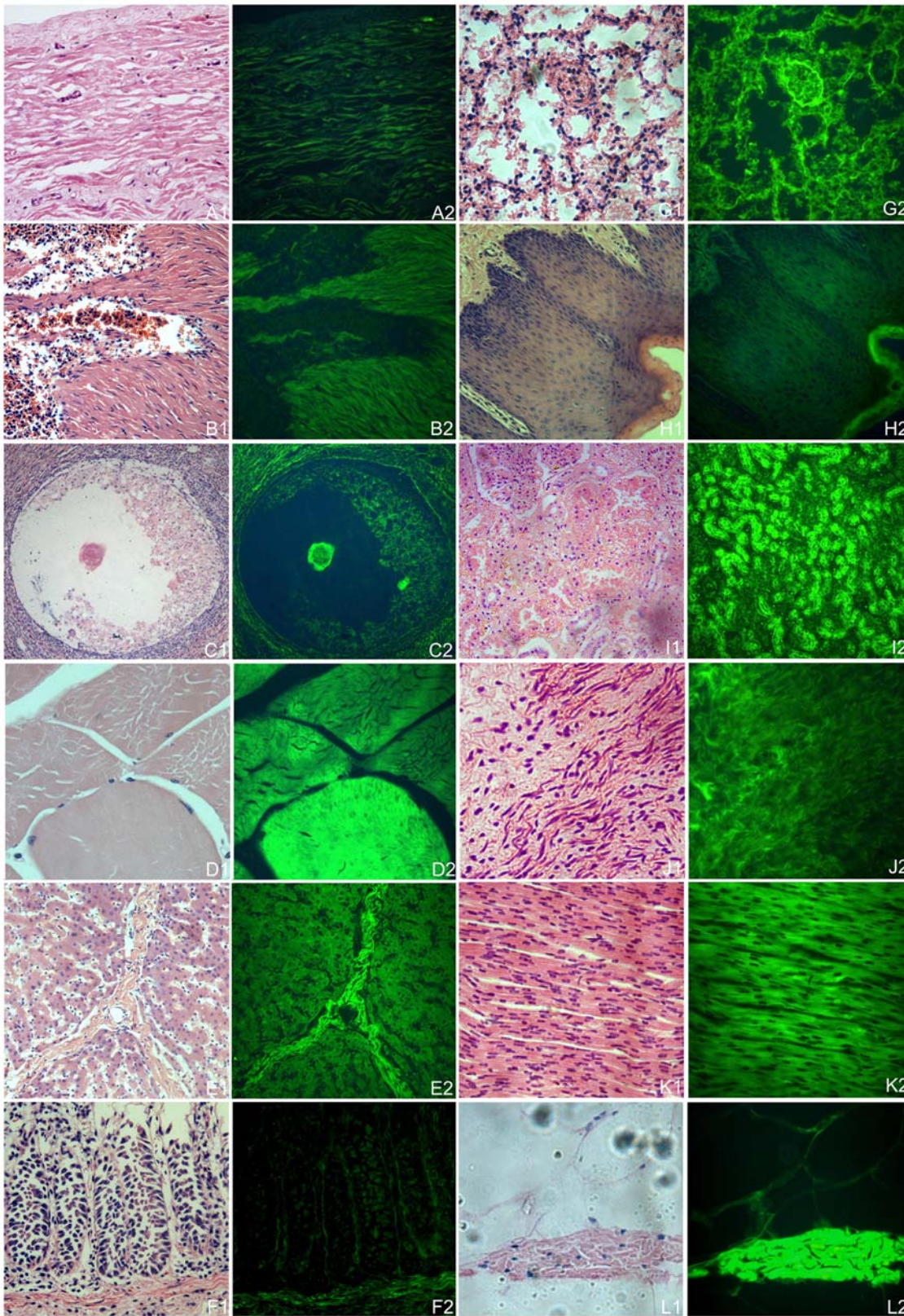
We detected CMV methylation levels in ears of newborn and mature transgenic pigs, and observed an increase from 26% to 40% ( $p = 1.000$ ) and from 19% to 38% ( $p = 0.799$ ) in K25-2 and K25-3, respectively (Fig. 5A).

In the analysis of CMV methylation in transgenic fibroblast cells cultured from 20 to 90 days (Fig. 5B), we found that the level increased more than 3-fold from 26% to 81% ( $p = 0.053$ ), and the increase in CMV methylation was conversely correlated with GFP expression ( $r = -0.967$ ,  $p < 0.001$ ). These results suggest that the methylation level of the transgene promoter appears to increase *in vivo* or *in vitro*, and to repress transgene expression.

A variegation of CMV methylation in different tissues of K25-3 was detected (Fig. 5C). There was a hypermethylation level in uterus (67%), spleen (67%) and adipose (71%), and a hypomethylation level in heart (38%) and muscle (30%). However, the difference in the methylation levels in different tissues was not significant ( $p = 0.153$ ). The correlation of GFP expression with CMV methylation in different tissues was significant ( $r = -0.982$ ,  $p < 0.001$ ), indicating that transgene expression is associated with promoter methylation in transgenic animal tissues, which is consistent with a recent report [23].

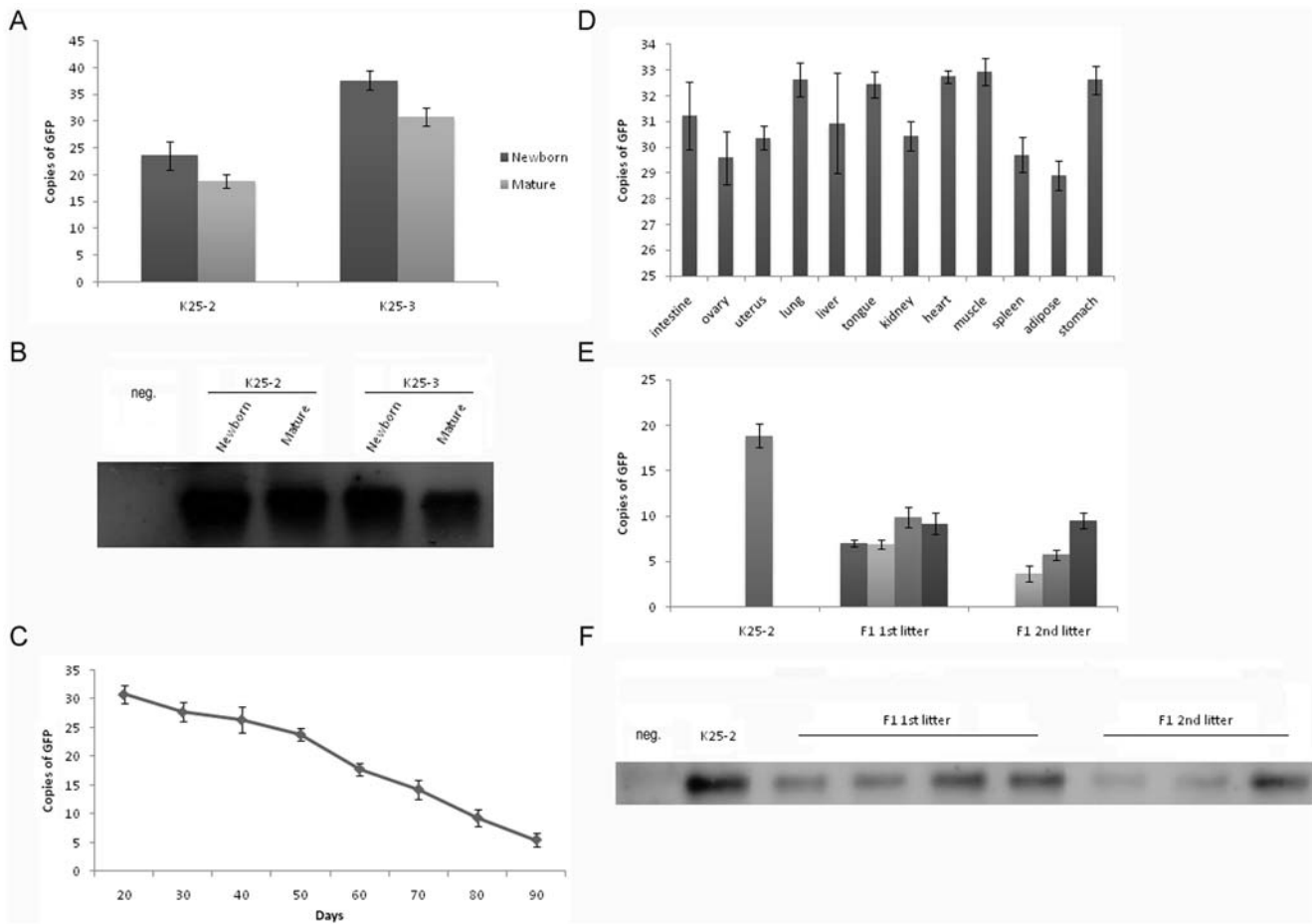
A change in CMV methylation during passage was also observed between K25-2 and its offspring (Fig. 5D). The increase in CMV methylation levels between founder and offspring, from 40% to 58% on average of offspring, was not significant ( $p = 0.537$ ), and there was no significant correlation of GFP expression with CMV methylation ( $r = -0.682$ ,  $p = 0.063$ ). Though a wide range of methylation levels from 50% to 74% among offspring was observed, the difference was not significant ( $p = 0.615$ ). The complex inheritance of transgene promoter





**Figure 2. Variagation of GFP expression in various tissues of the transgenic pig.** A to L: different tissues, namely uterus, spleen, ovary, muscle, liver, intestine, lung, tongue, kidney, stomach, heart and adipose, were visualized by HE staining. A1 to L1: tissues under normal light. A2 to L2: tissues under UV light.

doi:10.1371/journal.pone.0006679.g002



**Figure 3. Copy number of GFP.** (A) Absolute quantitative real-time PCR analysis of GFP copy number from newborn to maturity in founder transgenic pigs. There was no statistically significant decline in K25-2 ( $p=0.099$ ), but the decline in K25-3 was significant ( $p=0.016$ ); (B) Southern blots analysis of GFP copy number in newborn and mature transgenic pigs; (C) Absolute quantitative real-time PCR analysis of GFP copy number in transgenic fibroblast cells. Copy number of GFP declined in cells over time in culture. The decline was significant ( $p<0.001$ ); (D) Variegation of GFP copy number in various tissues of transgenic pig. Variegation of GFP copies was shown in different tissues ( $p=0.059$ ); (E) Absolute quantitative real-time PCR analysis of GFP copy number in offspring transgenic pigs. The decline from founder to offspring was significant ( $p<0.001$ ); (F) Southern blots analysis of GFP copy number in offspring transgenic pigs. Error bars denote standard deviations. Neg., non-transgenic pig. doi:10.1371/journal.pone.0006679.g003

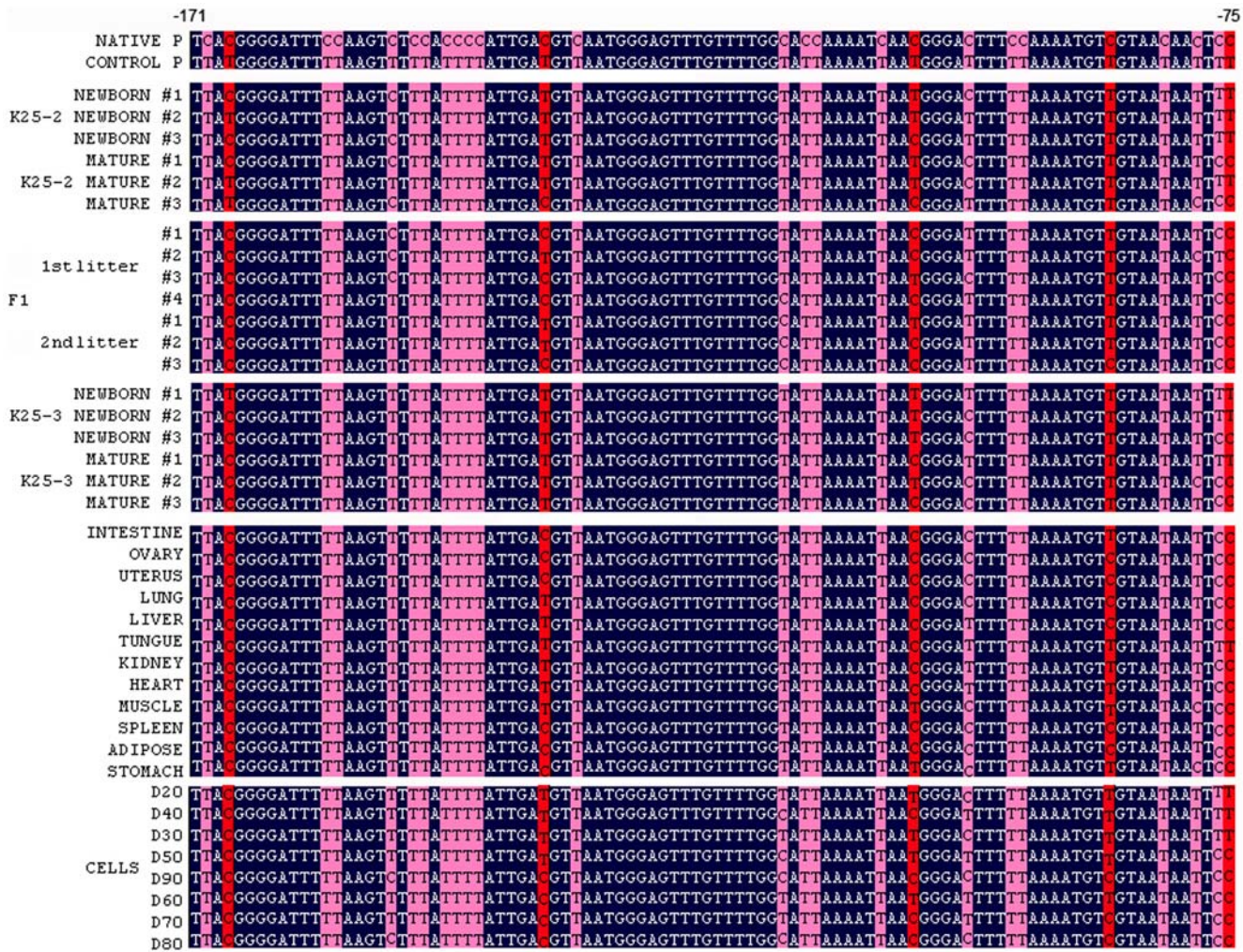
methylation seems to suggest randomly incomplete erasure and reset.

In these cases, we did not find significant differences in the methylation level, and the results observed may be related to cellular mosaicism and environment influences on methylation status [20,32].

### Effect of copy number and promoter methylation on transgene expression

In order to determine which factor, copy number or promoter methylation, is more significant in the determination of transgene





**Figure 4. Bisulfite sequence data for a representative region of CMV promoter.** Native P is the native sequence of the CMV promoter. The sequences from clones obtained after bisulfite treatment of DNA samples were aligned to the native sequence. For example, in control P cytosine residues were not methylated and therefore converted to thymidine by the bisulfite treatment. CpG and non-CpG cytosines were highlighted on a red and pink background, respectively.  
doi:10.1371/journal.pone.0006679.g004

expression, a multivariate linear regression model was employed. The linear regression fit has been shown to provide a reasonably close approximation to estimate the correlation of gene expression with copy number and DNA methylation [33,34]. In these models below, E, M and C represents GFP expression level, CMV methylation level and GFP copy number, respectively, and  $\beta_M$  and  $\beta_C$  represents the deterministic force of CMV methylation and GFP copy number, respectively.

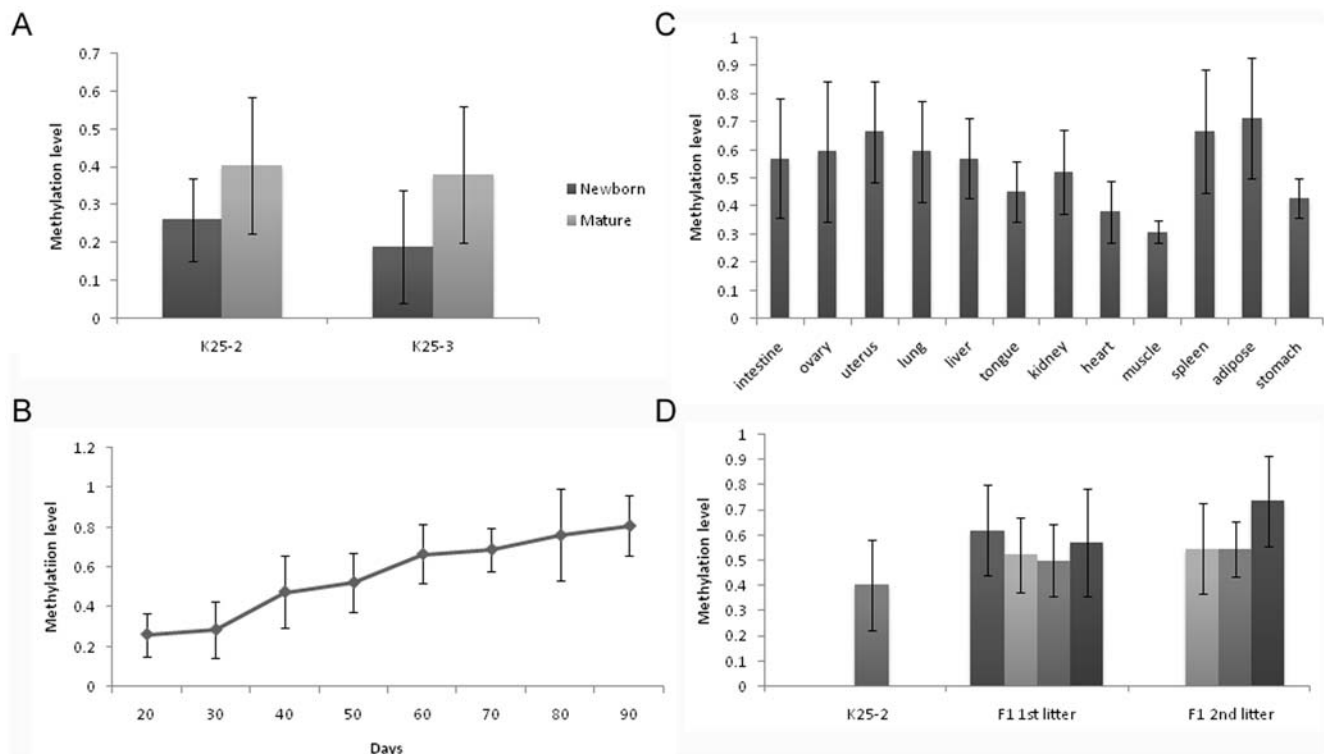
In analysis of transgenic fibroblast cells over a long time course in culture, the model was:  $E = -5.319M + 0.065C + 2.063$  ( $r = 0.984$ ,  $p < 0.001$ ) and  $\beta_M = -0.88$ ,  $\beta_C = 0.124$ . We also analyzed the correlation in different transgenic tissues, and the model was:  $E = -0.74M + 0.015C + 0.667$  ( $r = 0.977$ ,  $p < 0.001$ ) and  $\beta_M = -0.517$ ,  $\beta_C = 0.472$ . These results indicate that although both copy number and promoter methylation can regulate transgene expression levels in fibroblast cells and tissues, promoter methylation appears to be more significant in affecting transgene expression compared to copy number. The model in the passage of the transgenic pig was:  $E = -0.354M + 0.069C - 0.319$  ( $r = 0.865$ ,  $p = 0.032$ ) and  $\beta_M = -0.102$ ,  $\beta_C = 0.947$ , which suggests that copy number is superior to promoter methylation in transgene expression regulation.

**Discussion**

To our knowledge, this is the first study to systematically analyze the effect of copy number and promoter methylation on transgene expression in domestic animals. Our results suggest that transgene expression level is associated with copy number and CMV promoter methylation status in transgenic pigs.

In random transfection, the transgene usually integrates into the host genome in a tandem manner. High copy number tandem integration is thought to lead to transgene silencing [17,35], however, in the present study, the correlation of transgene expression level with copy number was positive ( $\beta_C > 0$ ). This difference may be caused by a position effect, and further experiments are needed to clarify the exact mechanism. CMV is very strong promoter *in vitro* [36]. Several reports have described the CMV promoter as being silent *in vivo*, and described a role of DNA methylation in silencing of the CMV promoter [17,37–39]. A negative correlation between methylation of the promoter and gene expression has been documented previously [40,41], and was confirmed in our study ( $\beta_M < 0$ ).

The decline of transgene expression with time may contribute to the increase in promoter methylation rather than the loss of copy



**Figure 5. Methylation of CMV promoter.** (A) Level of CMV methylation increased in founder transgenic pigs from newborn to maturity. The increases in CMV methylation level from 26% to 40% ( $p = 1.000$ ) and from 19% to 38% ( $p = 0.799$ ) were observed in K25-2 and K25-3, respectively; (B) Level of CMV methylation increased in transgenic fibroblast cells over time in culture. The increase in CMV methylation level from 20 to 90 days was more than 3-fold from 26% to 81% ( $p = 0.053$ ); (C) Variagation of CMV methylation in various tissues of transgenic pig. Hypermethylation and hypomethylation levels of CMV methylation were found in different tissues, but the difference was not significant ( $p = 0.153$ ); (D) Level of CMV methylation in offspring transgenic pigs. The increase in CMV methylation level from 40% to 58% on average of offspring was detected ( $p = 0.537$ ). Error bars denote standard deviations.  
doi:10.1371/journal.pone.0006679.g005

number. Previous reports have shown that transgene expression level declines with time *in vivo* and *in vitro* [6,9]. This is similar with the results observed here, and, in addition, both a loss of copy number and increase in promoter methylation were observed. The loss of a transgene integrated into the host genome has been observed in K562 cells over a long time course *in vitro* [6]. However, we demonstrated the decline of transgene copy number with aging in transgenic animals. The transgene copy number decrease can be understood in two points below: firstly, it is position dependent. A few reports described there were some sites where transgene was inclined to integrate by random transfection, such as LINE elements [6,27]. We believe the random transfection may not result in random integration instead in some hot integration sites and these hot integration sites may have the common character easy for transgene to integrate, but, we propose the hypothesis, it may be also easy for transgene to lose; secondly, it is sequence dependent. The transgene we used is eGFP which is heterogeneous from a kind of medusa, and we propose there may be a sequence dependent mechanism to immune-mediated destruction of transgene, as in the deletion of endogenous virus [42,43]. An increase in transgene promoter methylation was demonstrated during individual development of transgenic pigs and transgenic fibroblast cells. This observation contradicts previous conclusions that established methylation patterns can be maintained and stably transmitted during mitosis [18,20]. A gradual modification of methylation has been observed during individual development of cloned pigs, and it is important to correct the aberrant expression of imprinted genes in cloned

embryos and offspring [44,45]. The mechanism of increase in transgene promoter methylation may be related to a defense system targeted against the transgene [46,47]. In this case, it seems that the decline of transgene expression may be caused by both the loss of copy number and an increase in promoter methylation, but, according to the  $\beta$ -value, the latter likely plays a more important role.

The variegation of transgene expression in different tissues of transgenic animals is more closely correlated with promoter methylation than copy number. The CMV promoter exhibits various expression profiles. Villuendas et al. (2001) [48] and Van den Pol et al. (1998) [35] reported that the CMV promoter is active in neurons, testis and certain other tissues of transgenic mice, but Yang PH et al. (2008) [28] did not detect CMV activity in heart, liver, spleen, lung, kidney, skin and muscle of transgenic cattle. In our study, GFP expression was detected in all the tissues examined, but the expression levels were different between tissues. Varied transgene copies and promoter methylation were found in tissues. A difference in DNA methylation among tissues is well known, and the methylation status of an identical site in different tissues may also be different [49]. Thus, the differential promoter methylation status of the transgene among several tissues is consistent to the previous observations. However, the mechanism leading to the variegation of transgene copies is not clear. Transgene in different tissues of transgenic animals generated by pronuclear microinjection observed a mosaic pattern has been reported [50,51]. But in the case of SCNT animals, all the subsequent offspring and the cells within them originate from a

single donor cell status which basically reflects a single genotype and its epigenetic status. Different cell division rates in different tissues may lead to a difference in the extent of loss of transgene. Nevertheless, according to the data reported here, different transgene copy numbers and the promoter methylation status in different tissues may be responsible for the expression variation.

The decline of transgene expression in offspring seems to result in a loss of the transgene during passage. Founder transgenic pigs mated with non-transgenic pigs may reduce the copy number, consistent with previous works [27,52]. DNA methylation associated with gene silencing is considered to be inherited during mitosis, but cleared during meiosis, enabling the genome to return to the totipotent state. Classic models of *de novo* methylation describe erasure through two germlines, and then a resetting. However, some reports have suggested that DNA methylation is not completely erased, but inherited [53–55]. In this study, we could not observe clearly the inheritance of the methylation status, but an increase in promoter methylation level was obvious in the offspring. Kearns M et al. (2000) [20] reported that a transgene inherited from the mother is completely silenced in some offspring, but, in all the genotype positive transgenic pigs we obtained, the expression of the transgene transmitted by the mother was detected. However, an obvious decline was observed, and, according to multivariate linear regression analysis, the decline is attributed to the loss of transgene copy number.

In the study, we found 1) decline of transgene copy number and increase of promoter methylation level with time *in vivo* or *in vitro*, and during passage, 2) the variegation pattern of transgene copy number and promoter methylation status in various tissues, and 3) all of these are associated with the changes of transgene expression. It is known that the correlation of promoter methylation and copy number with gene expression is not very tight, and position effect, histone modifications or other epigenetic factors can also influence transgene expression [20,23]. However, in conclusion, our results at least demonstrate that transgene copy number and promoter methylation are responsible for the regulation of transgene expression in transgenic pigs.

## Materials and Methods

### Ethics statement

All the treatments of animals in this research followed by the guideline of Northeast Agriculture University and were approved by the committee. All animals (pigs) involved in this research were raised and breed followed the guideline of Animal Husbandry Department of Heilongjiang, P.R.China.

### Establishment of GFP transgenic pigs and passage

Fibroblast cells derived from E32 fetuses were transfected by liposome-mediated plasmid pEGFP-C1 (Clontech) containing eGFP driven by the CMV promoter, which was based on a random insertion of nonhomologous DNA vector into host genome. After G418 selection, surviving cells were used as a nuclear donor, and nuclear transfer was performed as described [29]. After sexual maturity, founder transgenic pigs were mated with non-transgenic pigs to passage. Positive transgenic pigs were identified by PCR using primers P1 and P2. The sequences of the primers were 5'-TGAACCGCATCGAGCTGAAGGG-3' (forward) and 5'-TCCAGCAGGACCATGTGATCGC-3' (reverse), and PCR generated a 308-bp product. All DNA samples were extracted using the Universal Genomic DNA Extraction Kit Ver.3.0 (TaKaRa) according to the manufacturer's instructions.

### Southern blots

Each DNA sample was cleaved with EcoR I and Nhe I (TaKaRa), which can digest the pig genome efficiently and form a 800 bp fragment. The hybridization probe used to detect the GFP transcription unit DNA (753 bp) was synthesized by PCR using primers P3 and P5 and labeled by DIG Oligonucleotide 3'-End Labeling Kit (Roche). The sequences of the primers were 5'-GAGCAAGGGCGAGGAGCTGTTCA-3' (forward) and 5'-TGCAGAATTCGAAGCTTGAGC-3' (reverse).

### Real-time PCR analysis

**Real-time PCR procedure.** Real-time PCR was performed using SYBR Premix Ex Taq™ (TaKaRa) and the 7300 Real-Time PCR System (Applied Biosystems), with the following parameters: 95°C for 10 sec, followed by 40 two steps cycles at 95°C for 5 sec and at 60°C for 31 sec. For RT-PCR, total RNAs were extracted from each sample using the PureLink™ Micro-to-Midi system (Invitrogen) according to the manufacturer's instructions, and reverse transcription was to generate cDNAs using PrimeScript™ RT Reagent Kit (TaKaRa). Primers for the GFP gene were 5'-TGAACCGCATCGAGCTGAAGGG-3' (forward) and 5'-ACCTTGATGCCGTTCTTCTGCTTG-3' (reverse). For absolute quantitative PCR, the TFRC gene was used as a reference gene, and the primers were 5'-GAGACAGAACTTTCGAAGC-3' (forward) and 5'-GAAGTCTGTGGTATCCAATCC-3' (reverse). In relative quantitative RT-PCR, the  $\beta$ -actin gene was used as a reference gene, and the primers were 5'-AGATCGTGGGGACATCAAG-3' (forward) and 5'-GCGGCAGTGGCCATCTC-3' (reverse). The sizes of the amplification products were 110 bp for the GFP gene, 81 bp for the TFRC gene and 93 bp for the  $\beta$ -actin gene. For each DNA and cDNA sample, both target and reference genes were always amplified independently on the same plate and in the same experimental run in triplicate. The melting curve analysis showed that all reactions were free of primer-dimers or other non-specific products (data not shown).  $C_t$  value was calculated by the Sequence Detection System software (Applied Biosystems). In relative quantitative RT-PCR, the amount of target normalized to reference was calculated by:  $2^{-\Delta\Delta C_t}$ .

### Establishment of the absolute quantitative standard curve

In order to examine the GFP copy number, generation of the absolute quantitative standard curve was necessary. First, we prepared a series of standard samples containing 1, 2, 4, 8, 16 copies of the GFP gene respectively, by mixing the wild type genome of an E32 pig with plasmid pEGFP-C1. To make a standard sample contain one copy of the GFP gene, the quality of plasmid mixed with genomic DNA was:  $\frac{a \times b \times 0.5}{3 \times 10^9}$  ng ("a" represents the size of plasmid). According to this principle, the standard samples containing 1, 2, 4, 8, and 16 copies of the GFP gene were prepared. The absolute quantitative standard curve was drawn by plotting  $\Delta C_t$  ( $\Delta C_t = C_{tGFP} - C_{tTFRC}$ ) against the log of GFP gene copies of corresponding standard samples. The parameters of the standard curve was:  $\log_2 N = -0.9354\Delta C_t + 3.4116$  ( $R^2 = 0.9974$ ,  $p < 0.001$ ).

### Bisulfite sequencing

Bisulfite modification was performed on 0.3  $\mu$ g of DNA from each sample using the EZ DNA Methylation-Gold™ Kit (Zymo research), according to the instruction manual. PCR primers to amplify the CMV were designed by MethPrimer software on line (<http://www.urogene.org/methprimer/>), which was also used to predict CpG islands and CpG sites in the sequence. The following



PCR primers could efficiently and specifically amplify a 278-bp region containing one CpG island with 14 CpG sites: gps: 5'TGATTTTATGGGATTTTTTTATTTG3' (forward) and gpa: 5'ATTCACATAACCAACTCTACTTATATAAAC3' (reverse). None of these bisulfite dependent residues lie within the first 6 bp of the 3' end of the primers, so if some of these were not bisulfite converted this should have a limited effect upon primer efficiency. The amplification of bisulfited-modified DNA was performed using Hot start Taq<sup>TM</sup> polymerase (TaKaRa), with the following conditions: 94°C for 5 min, followed by 40 three steps cycles at 94°C for 30 sec, 56°C for 30 sec and at 72°C for 1 min. The PCR products were separated on 1% agarose gels and purified, followed by sequencing (Invitrogen). The presence of a cytosine residue after bisulfite treatment shows that the cytosine residue was protected by methylation from bisulfite modification. For each DNA sample, the number of cytosine residues that remained as a cytosine was counted, and converted to a percentage of the 14 CpG cytosines present in the 278-bp region of the CMV that was analyzed. For the control, the pEGFP-C1 plasmid was treated and analyzed. At least five clones were sequenced and analyzed for each sample.

### Western blots

For Western blot analysis, total proteins were isolated from different tissues of K25-3 and the ear of K25-2 and its offspring by homogenization in lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% Triton X-100, 0.25% sodium deoxycholate, and complete protease inhibitor cocktail (Roche)). The concentration of proteins was measured by Bradford reagent (Sigma), separated on 10% SDS-PAGE gels and transferred to Immobilon-P membranes (Millipore). After blocking in 5% low-fat milk in PBST (0.1% Tween 20 in PBS) for 1 h, the membranes were incubated with GFP antibody (1:500, Santa Cruz Biotechnology) or rabbit anti-Gapdh polyclonal antibody (1:2000, Santa Cruz Biotechnology) overnight at 4°C. After washing in PBST, the membranes were incubated in goat anti-rabbit antibody conjugated with horseradish peroxidase (1:5000) for 1 h, followed by

three washes in PBST. The signals were detected by ECL Chemiluminescent kit (Amersham Pharmacia Biotech, Arlington Heights).

### Flow cytometry analysis

Fibroblast cells isolated from the ear of newborn K25-3 were cultured and proliferated in DMEM+20% FBS (Gibco). The fluorescence intensities of fibroblast cells over the time were analyzed in a FACS Calibur (Becton-Dickinson). The argon laser was tuned at 488 nm, and fluorescent cells were evaluated with a 525 nm band-pass filter. To set the parameters for flow cytometry analysis, non-transfected fibroblast cells were used as a negative control.

### Statistical analysis

Statistical analysis was preformed using SPSS 13.0 for MicroSoft<sup>TM</sup> Windows. Data are shown as mean±SD. One-way ANOVA was used to assess differences between groups. Duncan method was employed for pairwise comparison and followed by Bonferroni correction. Pearson correlation coefficient (PCC) and multivariate linear regression analysis were performed to determine independent associations between GFP expression levels and the variables of interest. P<0.05 (two-tailed) was considered statistically significant.

### Acknowledgments

The authors gratefully acknowledge Dr. Jilong Liu for performing the statistical analyses. We also thank our colleagues in the "Lab of Embryo Biotechnology" group for helpful discussions.

### Author Contributions

Conceived and designed the experiments: QK ZL. Performed the experiments: QK MW YH LZ HL YL. Analyzed the data: QK GB YM. Contributed reagents/materials/analysis tools: ZL. Wrote the paper: QK GB ZL.

### References

- Houdebine LM (2002) The methods to generate transgenic animals and to control transgene expression. *Biotechnology* 98: 145–160.
- Park KW, Cheong HT, Lai L, Im GS, Kühholzer B, et al. (2001) Production of nuclear transfer-derived swine that express the enhanced green fluorescent protein. *Anim Biotechnol* 12: 173–181.
- Lai L, Kang J, Li R, Wang J, Witt WT, et al. (2006) Generation of cloned transgenic pigs rich in omega-3 fatty acids. *Nature Biotechnology* 24: 435–436.
- Phelps CJ, Koike C, Vaught TD, Kühholzer B, Samuel M, et al. (2003) A transgenic pig expressing the enhanced green fluorescent protein produced by nuclear transfer using colchicine-treated fibroblasts as donor cells. *Science* 299: 411–414.
- Gubin AN, Reddy B, Njoroge JM, Miller JL (1997) Long-term Retroviral-mediated stable expression of Green Fluorescent Protein in mammalian cells. *Biochem Biophys Res Commun* 236: 347–350.
- Migliaccio AR, Bengra C, Ling J, Pi W, Li C, et al. (2000) Stable and unstable transgene integration sites in the human genome: extinction of the Green Fluorescent Protein transgene in K562 cells. *Gene* 256: 197–214.
- Houdebine LM (2000) Transgenic animal bioreactors. *Transgenic Res* 9: 305–320.
- Yamauchi Y, Doe B, Ajduk A, Ward MA (2007) Genomic DNA damage in mouse transgenesis. *Biol Reprod* 77: 803–812.
- Robertson G, Garrick D, Wilson M, Martin DI, Whitelaw E (1996) Age-dependent silencing of globin transgenes in the mouse. *Nucleic Acids Res* 24: 1465–1471.
- Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126: 663–676.
- Okita K, Ichisaka T, Yamanaka S (2007) Generation of germline-competent induced pluripotent stem cells. *Nature* 448: 313–317.
- Werning M, Meissner A, Foreman R, Brambrink T, Ku M, et al. (2007) In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 448: 318–324.
- Garrick D, Fiering S, Martin DIK, Whitelaw E (1998) Repeat-induced gene silencing in mammals. *Nature Genet* 18: 56–59.
- Martin DIK, Whitelaw E (1996) Variegation of variegating transgenes. *BioEssays* 18: 919–923.
- Eszterhas SK, Bouhassira EE, Martin DIK, Fiering S (2002) Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position. *Mol Cell Biol* 22: 469–479.
- Chiaromonte F, Miller W, Bouhassira EE (2003) Gene length and proximity to neighbors affect genome-wide expression levels. *Genome Res* 13: 2602–2608.
- Strathdee D, Ibbotson H, Grant SGN (2006) Expression of transgene targeted to the *Gt(ROSA)26Sor* locus is orientation dependent. *PLoS ONE* 1(1): e4.
- Balow D (1995) Gametic imprinting in mammals. *Science* 270: 1610–1613.
- Schumacher A, Koetsier PA, Hertz J, Doerfler W (2000) Epigenetic and genotype-specific effects on the stability of *de Novo* imposed methylation patterns in transgenic mice. *Biol Chem* 275: 37915–37921.
- Kearns M, Preis J, McDonald M, Morris C, Whitelaw E (2000) Complex patterns of inheritance of an imprinted murine transgene suggest incomplete germline erasure. *Nucleic Acids Res* 28: 3301–3309.
- Prosch S, Stein J, Staak K, Liebenthal C, Volk HD, et al. (1996) Inactivation of the very strong HCMV immediate early promoter by DNA CpG methylation in vitro. *Biol Chem Hoppe Seyler* 377: 195–201.
- Brooks AR, Harkins RN, Wang P, Qian HS, Liu P, et al. (2004) Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle. *J Gene Med* 6: 395–404.
- Mehta AK, Majumdar SS, Alam P, Gulati N, Brahmachari V (2009) Epigenetic regulation of cytomegalovirus major immediate-early promoter activity in transgenic mice. *Gene* 428: 20–24.
- Liu W, Xiong Y, Gossen M (2006) Stability and homogeneity of transgene expression in isogenic cells. *J Mol Med* 84: 57–64.

25. Park KW, Lai L, Cheong HT, Cabot R, Sun QY, et al. (2002) Mosaic gene expression in nuclear transfer-derived embryos and the production of cloned transgenic pigs from ear-derived fibroblasts. *Biol Reprod* 66: 1001–1005.
26. Schulz TJ, Glaubitz M, Kuhlow D, Thierbach R, Birringer M, et al. (2007) Variable expression of Cre recombinase transgenes precludes reliable prediction of tissue-specific gene disruption by tail-biopsy genotyping. *PLoS ONE* 2(10): e1013.
27. Nakanishi T, Kuroiwa A, Yamada S, Isotani A, Yamashita A, et al. (2002) FISH analysis of 142 EGFP transgene integration sites into the mouse genome. *Genomics* 80: 564–574.
28. Yang P, Wang J, Gong G, Sun X, Zhang R, et al. (2008) Cattle mammary bioreactor generated by a novel procedure of transgenic cloning for large-scale production of functional human lactoferrin. *PLoS ONE* 3(10): e3453.
29. Liu ZH, Song J, Wang ZK, Tian JT, Kong QR, et al. (2008) Green fluorescent protein (GFP) transgenic pig produced by somatic cell nuclear transfer. *Chinese Science Bulletin* 53: 1035–1039.
30. Eduardo OM, Aurea MOC, Mauricio MF, Rumpf R (2007) Animal transgenesis: state of the art and applications. *J Appl Genet* 48: 47–61.
31. Grunau C, Clark SJ, Rosenthal A (2001) Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res* 29: e65.
32. Whitelaw NC, Whitelaw E (2006) How lifetimes shape epigenotype within and across generations. *Hum Mol Genet* 15: 131–137.
33. Balko JM, Potti A, Saunders C, Stromberg A, Haura EB, et al. (2006) Gene expression patterns that predict sensitivity to epidermal growth factor receptor tyrosine kinase inhibitors in lung cancer cell lines and human lung tumors. *BMC Genomics* 7: 289.
34. Bönsch D, Lenz B, Fiszler R, Frieling H, Kornhuber J, et al. (2006) Lowered DNA methyltransferase (DNMT-3b) mRNA expression is associated with genomic DNA hypermethylation in patients with chronic alcoholism. *J Neural Transm* 113: 1299–304.
35. Tang W, Newton RJ, Weidner DA (2007) Genetic transformation and gene silencing mediated by multiple copies of a transgene in eastern white pine. *J Exp Bot* 58: 545–554.
36. Van den Pol AN, Ghosh PK (1998) Selective neuronal expression of green fluorescent protein with cytomegalovirus promoter reveals entire neuronal arbor in transgenic mice. *J Neurosci* 18: 10640–10651.
37. Grassi G, Maccaroni P, Meyer R, Kaiser H, D'Ambrosio E, Pascale E, et al. (2003) Inhibitors of DNA methylation and histone deacetylation activate cytomegalovirus promoter-controlled reporter gene expression in human glioblastoma cell line U87. *Carcinogenesis* 24: 1625–1635.
38. Meier JL (2001) Reactivation of the human cytomegalovirus major immediate-early regulatory region and viral replication in embryonal NTera2 cells: role of trichostatin A, retinoic acid, and deletion of 21-base-pair repeats and modulator. *J Virol* 75: 1581–1593.
39. Choi KH, Basma H, Singh J, Cheng PW (2005) Activation of CMV promoter controlled glycosyltransferase and  $\beta$ -galactosidase glycozymes by butyrate, trichostatin A, and 5-Aza-2'-deoxycytidine. *Glycoconjugate J* 22: 63–69.
40. Toth M, Lichtenberg U, Doerfler W (1989) Interindividual concordance of methylation profiles in human genes for tumor necrosis factors alpha and beta. *Proc Natl Acad Sci U S A* 86: 3728–3732.
41. Muller U, Doerfler W (1987) Fixation of the unmethylated or the 5'-CCGG-3' methylated adenovirus late E2A promoter-cat gene construct in the genome of hamster cells: gene expression and stability of methylation patterns. *J Virol* 61: 3710–3720.
42. Bacharach E, Gonsky J, Lim D, Goff SP (2000) Deletion of a short, untranslated region adjacent to the polypurine tract in Moloney murine leukemia virus leads to formation of aberrant 5' plus-strand DNA ends in vivo. *J Virol* 74: 4755–64.
43. Pupecka M, Pacak A (2006) The innate antiretroviral defense of human cells, based on the DNA editing. *Postepy Biochem* 52: 247–52.
44. Kang YK, Koo DB, Park JS, Choi YH, Chung AS, et al. (2001) Aberrant methylation of donor genome in cloned bovine embryos. *Nat Genet* 28: 173–177.
45. Han DW, Young BI, Teong TD, Gupta MK, Uhm SJ, et al. (2008) Methylation Status of Putative Differentially Methylated Regions of Porcine IGF2 and H19. *Mol Reprod Dev* 75: 777–784.
46. Doerfler W (1992) DNA methylation: eukaryotic defense against the transcription of foreign genes? *Microb Pathog* 12: 1–8.
47. Yoder JA, Walsh CP, Bestor TH (1997) Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet* 13: 335–340.
48. Villuendas G, Gutierrez-Adan A, Jimenez A, Rojo C, Roldan ER, et al. (2001) CMV-driven expression of green fluorescent protein (GFP) in male germ cells of transgenic mice and its effect on fertility. *Int J Androl* 24: 300–305.
49. Oakes CC, La sale S, Smiraglia DJ, Robaire B, Trasler JM, et al. (2007) A unique configuration of genome-wide DNA methylation patterns in the testis. *Proc Natl Acad Sci U S A* 104: 228–233.
50. Burdon TG, Wall RJ (1992) Fate of microinjected genes in preimplantation mouse embryos. *Mol Reprod Dev* 33: 436–442.
51. Chan AW, Kukulj G, Skalka AM, Bremel RD (1999) Timing of DNA integration, transgenic mosaicism, and pronuclear microinjection. *Mol Reprod Dev* 33: 436–442.
52. Pillai MM, Venkataraman GM, Kosak S, Torok-Storb B (2008) Integration-site analysis in transgenic mice by thermal asymmetric interlaced (TAIL)-PCR: segregating multiple-integrant founder lines and determining zygosity. *Transgenic Res* 17(4): 749–54.
53. Sutherland HGE, Kearns M, Morgan HD, Headley AP, Morris C, et al. (2000) Reactivation of heritably silenced gene expression in mice. *Mamm Genome* 11: 347–355.
54. Morgan HD, Sutherland HGE, Martin DIK, Whitelaw E (1999) Epigenetic inheritance at the agouti locus in the mouse. *Nature Genet* 23: 314–318.
55. Gavalli G, Paro R (1998) The Drosophila Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis. *Cell* 93: 505–518.