




Research Article

Identification and Expression Pattern of EZH2 in Pig Developing Fetuses

Baohua Tan ¹, Linjun Hong,¹ Jiaxin Qiao,¹ Jian Zhou,¹ Pingping Xing,¹ Guanhao Yan,¹ Enqin Zheng,¹ Gengyuan Cai,^{1,2} Sixiu Huang,¹ Zhenfang Wu ^{1,2} and Ting Gu ¹

¹National Engineering Research Center for Breeding Swine Industry, Guangdong Provincial Key Lab of Agro-Animal Genomics and Molecular Breeding, College of Animal Science, South China Agricultural University, Guangzhou 510642, China

²Wens Foodstuff Group Co., Ltd., Yunfu, China

Correspondence should be addressed to Zhenfang Wu; wzf@scau.edu.cn and Ting Gu; tinggu@scau.edu.cn

Received 31 May 2020; Revised 7 August 2020; Accepted 20 August 2020; Published 6 October 2020

Academic Editor: Enrique Gomez

Copyright © 2020 Baohua Tan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The proper methylation status of histones is essential for appropriate cell lineage and organogenesis. EZH2, a methyltransferase catalyzing H3K27me3, has been abundantly studied in human and mouse embryonic development. The pig is an increasing important animal model for molecular study and pharmaceutical research. However, the transcript variant and temporal expression pattern of *EZH2* in the middle and late porcine fetus are still unknown. Here, we identified the coding sequence of the *EZH2* gene and characterized its expression pattern in fetal tissues of Duroc pigs at 65- and 90-day postcoitus (dpc). Our results showed that the coding sequence of *EZH2* was 2241 bp, encoding 746 amino acids. There were 9 amino acid insertions and an amino acid substitution in this transcript compared with the validated reference sequence in NCBI. *EZH2* was ubiquitously expressed in the fetal tissues of two time points with different expression levels. These results validated a different transcript in pigs and characterized its expression profile in fetal tissues of different gestation stages, which indicated that EZH2 played important roles during porcine embryonic development.

1. Introduction

Polycomb Group (PcG) proteins are a family of protein complex that regulate gene expression, especially repressing gene transcription [1]. As one of the two distinct complexes, namely, Polycomb Repressive Complex 1 (PRC1) and PRC2, PRC2 mediates gene silencing by modulating chromatin structure [2]. Enhancer of zeste homolog 2 (EZH2), one of the core components of PRC2, is a methyltransferase possessing the enzymatic activity to generate di/trimethylated lysine 27 in histone H3 [2]. Since identification of EZH2 in the research of protooncogene product Vav [3], studies have shown that EZH2 is highly expressed in tumorigenesis, which regulates the expression of tumor suppressor genes, such as in breast cancer, prostate cancer, and lung cancer [4–6]. In the early embryonic development, abnormal expression of EZH2 impaired embryo growth and pluripotency maintenance [7, 8]. Furthermore, the decreased expres-

sion of the *EZH2* gene is crucial for stem cell differentiation into specific cell lineages involved in myogenesis, adipogenesis, osteogenesis, neurogenesis, and haematopoiesis [9]. In cardiac conditional *Ezh2*-knockout mice, cardiomyocyte proliferation was repressed, and the apoptosis process was induced [10]. Thus, the appropriate expression level of the *EZH2* gene is important for embryonic development. So far, the expression pattern and the function of *EZH2* have been broadly studied in humans [11] and mice [12, 13]. In pigs, several studies have focused on its function in early embryonic development regulation and in SCNT efficiency improvement [8, 14–16]. However, the transcript and expression status of the *EZH2* gene during porcine embryonic development remains unknown.

Alternative splicing of gene can generate multiple transcripts and proteins to regulate tissue and organ development [17]. To date, a lot of *EZH2* variants have been found in various cell and tissue types [18–20]. Considering the varieties of

EZH2 splicing variant, identifying its transcript in porcine fetal tissues is the foundation to study its function. However, only one transcript has been validated in pigs, with several other transcript variants of *EZH2* from digital computational analysis in NCBI that have not been empirically validated. Moreover, none of the *EZH2* transcript variants have yet been identified in porcine fetal tissues.

In this study, we identified a transcript variant of *EZH2* in porcine fetal tissues by cloning and sequencing. Subsequently, we detected the expression of *EZH2* on mRNA level and protein level in two different embryonic development stages (65-dpc and 90-dpc) via qRT-PCR and western blots. The subcellular localization of *EZH2* protein was predicted by using different predictors (CELLO, Euk-mPLoc, WoLF PSORT, and TargetP). Our study is aimed at providing information for understanding the biological function of *EZH2* in porcine embryonic development.

2. Materials and Methods

2.1. Ethics Statement. All experiments were carried out following the rules of the Administration of Affairs Concerning Experimental Animals (Ministry of Science and Technology, China, revised June 2004) [21] and approved by the Animal Care and Use Committee of the South China Agricultural University, Guangzhou, China (approval number # SYXK2014-0136) [22].

2.2. Sample Collection. Six Duroc pig fetuses were used in our study, consisting of 3 fetuses on gestation day 65 and 3 fetuses on gestation day 90. In order to reduce the influence of individual differences on the results, each 3 fetuses on the same gestation day were half sibs. The definition of age was based on the time of fertilization. Each visible tissue including heart, liver, spleen, lung, kidney, and longissimus dorsi muscle was quickly sectioned in a 2 ml centrifuge tube and then stored in a liquid nitrogen tank until laboratory examinations.

2.3. Cloning and Sequencing. To acquire the coding sequence in fetal Duroc pigs, we extracted total RNA from the heart, liver, spleen, lung, kidney, and longissimus dorsi muscle on gestation 65 according to the instruction of Total RNA Kit I (Omega Bio-Tek, Norcross, USA) and then quantified it with spectrophotometry at 260/280 nm in a NanoDrop 2000 instrument (Thermo Fisher Scientific, Waltham, USA). By using a reverse transcription kit (TAKARA, Dalian, China), we synthesized cDNA in a PCR instrument (Life Technologies, NY, USA). The cDNA products were amplified by Taq DNA polymerase (Genstar, Beijing, China) with specific primer (forward primer: 5-ATGGGCCAGACTGGGAAGAA-3; reverse primer: 5-TCAAGGGATCTCCATTCTCTTTCGAA-3). The primers were designed and produced by Sangon Biotech Co., Ltd. (Shanghai, China) with reference sequence in NCBI (accession no: NM_001244309.1). The PCR amplification steps were as follows: 94°C, 5 min for incubation; 94°C, 30 s, 57.3°C, 30 s; 72°C, 40 s, repeated 35 cycles; and 72°C, 10 min. The PCR product was separated using 2% agarose gel and purified with a Gel

Extraction Kit (Omega Bio-Tek, Norcross, GA, USA). The purified products were ligated into the pMD19T simple vector (TAKARA, Dalian, China), and then, recombinant plasmids were transformed into *Escherichia coli* DH5 α competent cells (TAKARA, Dalian, China). Transformed DH5 α competent cells were smeared onto Luria-Bertani- (LB-) ampicillin plate medium and then incubated at 37°C for 12 h in a constant-temperature incubator. The target colonies were selected and mixed with 200 μ l LB-ampicillin medium. The mixtures were used as a cDNA template to amplify with specific primers as described above, and then, the existence of target gene was confirmed through electrophoresis. Finally, 8 identified colonies containing the target fragment of each tissue were sent to BGI (BGI, Shenzhen, China) for sequencing.

2.4. Bioinformatics Analysis. The cDNA sequence of *EZH2* derived from fetal tissue was deduced via the ProtParam software (<https://web.expasy.org/translate/>). Deduced protein sequence was aligned with the reference sequence (accession no: NP_001231238.1) via the DNAMAN software (Lynnon Biosoft, USA). Prediction of conserved domain was performed in NCBI (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) [23]. The secondary structure was calculated by SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) [24]. Phyre 2.0 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) [25] and Pymol 2.0 (<http://www.pymol.org>) were used to predict tertiary structure. The subcellular localization of the *EZH2* protein was predicted in Euk-mPLoc (<http://www.csbio.sjtu.edu.cn/bioinf/euk-multi-2/>) [26], CELLO (<http://cello.life.nctu.edu.tw/>) [27], WoLF PSORT (<https://wolfsort.hgc.jp/>) [28], and TargetP (<http://www.cbs.dtu.dk/services/TargetP-1.1/index.php>) [29].

2.5. Quantitative Real-Time PCR Analysis. We had performed quantitative real-time PCR (qRT-PCR) to examine the expression pattern of *EZH2* mRNA in different tissues including liver, spleen, lung, kidney, and longissimus dorsi muscle on gestation days 65 and 90. To enable statistical analysis, three fully independent biological replicates and three technical repeats were conducted. The cDNA of each tissue from different individuals was synthesized as described above. Each reaction mixture (10 μ l) contained 1 μ l cDNA solution, 0.4 μ l of each specific primer, 5 μ l SYBR Select Master Mix (Thermo Fisher Scientific, Waltham, USA), and 3.4 μ l ddH₂O. The primers were as follows: forward primer 5-CACGGCAGCCTTGCGACAG-3; reverse primer 5-CGGGAAAGCGGTTCTGACTC-3. The reaction was performed in Quant Studio™ 7 Flex Real-Time PCR System (Thermo Fisher Scientific, Waltham, USA). The specificity of the PCR was confirmed through a single peak in the melting curve. Reaction conditions were as follows: 2 min at 50°C, 10 min at 95°C, then 45 cycles of 15 s at 95°C, 10 s at 60°C, and 15 s at 72°C, followed by melting curve analysis from 60°C to 95°C to evaluate the specificity of the PCR products. Relative expression in a given sample was calculated by normalizing to GAPDH mRNA level.

2.6. Western Blot Analysis. Western blot experiment was manipulated following the previous study [30]. Briefly, the total protein of tissues from different fetuses was extracted according to the direction of RIPA lysis buffer (Beyotime, Beijing, China). The concentration of protein was measured by a BCA protein assay kit (Thermo Fisher Scientific, Waltham, USA) and then adjusted to 30 μg each hole for SDS-PAGE running. In this process, the proteins were subjected to immunoblotting analysis with the following antibodies: anti-EZH2 (1 : 1000, 5246S, Cell Signaling Technology, Danvers, Massachusetts, USA), anti-beta-tubulin (1 : 2000, GB11017, Servicebio, Wuhan, China), and HRP-conjugated Affinipure Goat Anti-Rabbit IgG (H+L) (1 : 5000, SA00001-2, Proteintech, Rosemont, IL, USA). Finally, the membranes were developed with the ECL system (WBULS0500, Millipore Corporation, Billerica, USA).

2.7. Statistical Analysis. The relative expression of EZH2 in the mRNA level and the protein level was visually shown in the figures presenting as mean \pm standard error of the mean (SEM) from three independent individuals of each group. Quantification of transcript levels of the *EZH2* gene was calculated by the comparative method ($2^{-\Delta\Delta\text{Ct}}$) [31]. The band density value of the EZH2 protein was analyzed via ImageJ software [32]. Statistical analyses were performed by independent-sample Student's *t*-test in SPSS 20.0 software (IBM, Armonk, NY, USA). $P < 0.05$ indicates significant difference, and $P < 0.01$ indicates that the difference was extremely significant.

3. Results

3.1. Cloning and Sequence Analysis of Porcine EZH2 Gene. The PCR amplification product of *EZH2* was detected by 2% agarose gel electrophoresis. As expected, a specific band of approximately 2200 bp is shown in Figure 1(a). Sequencing analyses showed that cloned cDNA sequence of *EZH2* is 2241 bp and encoding 746 amino acids (accession no: MN_923188.1). The nucleotide sequences of *EZH2* were aligned with the validated reference sequences in NCBI. The alignment results showed that there was a 27 bp insertion and G \rightarrow A mutation in the cloned coding sequence (Figure 1(b)). This 27 bp insertion was located at the position of g.109412783 to g.109412809 in genomic sequence of the *EZH2* gene (accession no: NC_010451.4). As shown in Figure 1(c), a 9 amino acid insertion and an amino acid substitution appeared in the cloned transcript of porcine fetal tissues.

3.2. Spatial Structures and Conserved Domain Analysis of EZH2 Protein. The secondary structure analysis indicated that the EZH2 protein consisted of alpha helix (31.77%), extended strand (12.47%), beta turn (4.69%), and random coil (51.07%) (Figure 2(a)). For the newly cloned *EZH2* variant, a 9 amino acid insertion was composed of random coils. The amino acid substitution was composed of random coil and consistent with reference transcript. Additionally, the 9 amino acid insertion was not located in the conserved domain including WD-binding domain, CXC, SET, and

Polycomb Repressive Complex 2 Tri-helical domain (PRC2 HTH 1), while the amino acid substitution was located in the PRC2 HTH 1 domain (Figure 2(b)). As shown in Figure 2(c), the tertiary structure result showed the protein encoded by our cloned transcript was mainly composed of alpha helix, beta sheet, and random loop. In particular, the amino acid insertion was mainly composed of alpha helices, while the amino acid substitution did not change tertiary structure.

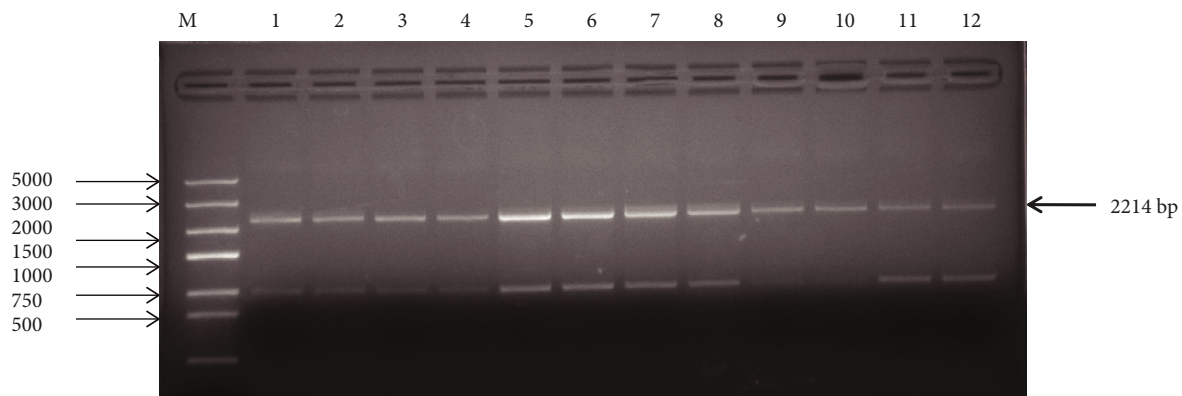
3.3. Subcellular Localization of EZH2 Protein. The deduced protein sequence was applied to predict subcellular localization of EZH2 protein via four well-developed predictors (Euk-mPLOC, CELLO, WoLF PSORT, and TargetP). As shown in Table 1, the TargetP result indicated that no secretory pathway signal peptide and mitochondrial targeting peptide were detected in the EZH2 protein. Euk-mPLOC, CELLO, and WoLF PSORT indicated that the EZH2 protein had the greatest probability to locate in the nucleus. In conclusion, these predicted results showed that the EZH2 protein might be located in the nucleus.

3.4. The Expression Pattern of EZH2 mRNA in Tissues. qRT-PCR was employed to detect the expression pattern of *EZH2* in the mRNA level during porcine embryonic development. As shown in Figure 3, *EZH2* was expressed ubiquitously in all tested tissues, including heart, liver, spleen, lung, kidney, and muscle. Statistical calculation demonstrated that *EZH2* was expressed with a significant difference in the spleen and lung. In the spleen, the mRNA level of *EZH2* significantly increased with the embryonic development ($P < 0.01$) while the opposite tendency of expression level was detected in the lung ($P < 0.05$). No significant difference of *EZH2* appeared in the heart, liver, kidney, and muscle between 65-dpc and 90-dpc. The result of muscle is shown in Supplementary Figure 1(a).

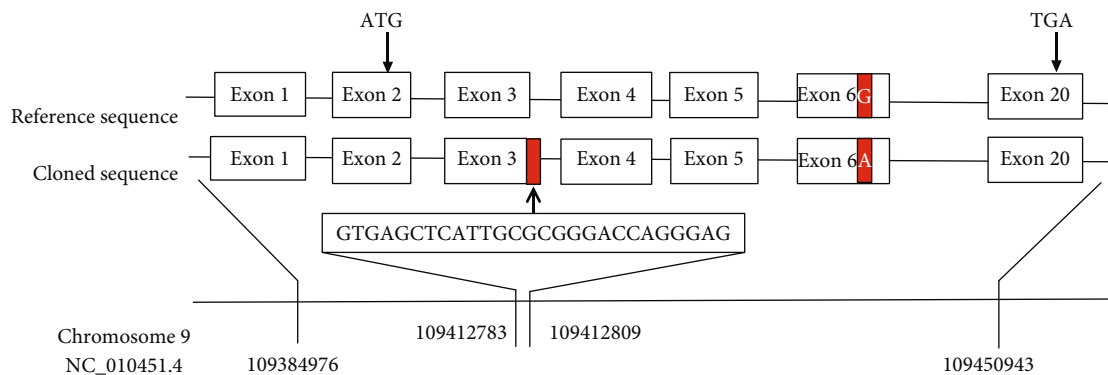
3.5. The Expression Pattern of EZH2 Protein in Tissues. The levels of the EZH2 protein were detected by western blot in various tissues, including heart, liver, spleen, lung, and kidney of Duroc pigs in 65-dpc and 90-dpc (Figure 4). Data indicated that the EZH2 protein was ubiquitously detectable in all tissues of two stages. Consistent with the result of qRT-PCR, there were no significant differences of EZH2 protein level in the heart, liver, muscle, and kidney. The expression of EZH2 showed a significant decrease in the spleen and lung from 65-dpc to 90-dpc ($P < 0.01$ and $P < 0.05$, respectively). The result of muscle is shown in Supplementary Figure 1(b).

4. Discussion

Genome-wide studies estimated that in 90–95% of human genes exist different levels of alternative splicing, which can generate various mRNA isoforms with different functions [17]. In eukaryotes, alternative splicing usually follows the GT-AG rule by both *cis*-elements and *trans*-acting factor regulation [33]. One of the most common types of alternative pre-mRNA splicing is alternative 5' splice site, which is recognized by base pairing with the end of the U1 small nuclear



(a)



EZH2

(b)

FIGURE 1: Continued.

Reference protein sequence	MGQTGKKEKGPVCRKRVKSEYMRIQLKRRFRRADEVKSMFSSNRQKILERTEILNQEWKQRRIQVHI	70
Cloned EZH2 variant	MGQTGKKEKGPVCRKRVKSEYMRIQLKRRFRRADEVKSMFSSNRQKILERTEILNQEWKQRRIQVHI	70
Consensus	mgqtgkkekqpvcwrkrvkseymrlqlkrfradevksmfssnrqkilerteilngewkqrriqvhi	
Reference protein sequence	LTS.....CSVTSDDLFPFTQVIPLKTLNVAASVPIMYSWSPLQQNFVDETVLHNIPIYMGDEVLD	131
Cloned EZH2 variant	LTSVSSLRGRTRCSVTSDDLFPFTQVIPLKTLNVAASVPIMYSWSPLQQNFVDETVLHNIPIYMGDEVLD	140
Consensus	lts csvtsdldfptqviplktlnavasvpimyswspqqnfvvedetvlhniipyngdevld	
Reference protein sequence	QDGTFFIEELIKNYDGKVGHDRECGFINDEFVELVNALGCYNDIDDDDDGDDPDEREEKQKDLLENREDK	201
Cloned EZH2 variant	QDGTFFIEELIKNYDGKVGHDRECGFINDEFVELVNALGCYNDIDDDDDGDDPDEREEKQKDLLENREDK	210
Consensus	qdggtfieeliknydgkvhgdrecgfindefvelvnalgcynd ddddgdgdpdereekqkdlleenredk	
Reference protein sequence	ETRPPRKFPSDKIFEAISSMFPDRGTAELKEKYKELTEQQLPGALPPECTPNIDGPNKSVQREQSLHS	271
Cloned EZH2 variant	ETRPPRKFPSDKIFEAISSMFPDRGTAELKEKYKELTEQQLPGALPPECTPNIDGPNKSVQREQSLHS	280
Consensus	etrpprkfpsdkifeaissmfpdkgtaelkekykelteqqlpgalppectpnidgpnaksvqreqslhs	
Reference protein sequence	FHTLFCRRCFKYDCLFHPFHATPNTYKRKNTETALDNKPCGPGCYQHLEGAKEFAAALTAERIKTPPPRR	341
Cloned EZH2 variant	FHTLFCRRCFKYDCLFHPFHATPNTYKRKNTETALDNKPCGPGCYQHLEGAKEFAAALTAERIKTPPPRR	350
Consensus	fhtlfcrrcfkydcflhpfhatpntykrkntetaldnkpcgpgqcyqhlegakefaaaltaeriktpprrp	
Reference protein sequence	GGRRRGRLPNNSSRPSTPTINVLESKDTSDREAGTETGGENNDKEEEKKDEASSSSSEANSRCQTFPIKM	411
Cloned EZH2 variant	GGRRRGRLPNNSSRPSTPTINVLESKDTSDREAGTETGGENNDKEEEKKDEASSSSSEANSRCQTFPIKM	420
Consensus	ggrrrgrlpnnsrpstptinvleskdtstdreagtetggenndkeeeekdeasssseansrcqtpikm	
Reference protein sequence	KPNAEPPENVEWSGAEASFRVLIGTYYDNFCAIARLIGTKTCRQVYEFVRKSSVIAFAPTEDVDTPPR	481
Cloned EZH2 variant	KPNAEPPENVEWSGAEASFRVLIGTYYDNFCAIARLIGTKTCRQVYEFVRKSSVIAFAPTEDVDTPPR	490
Consensus	kpnaeppenvewsgaeasmfrvligtyydnfcaiarligtktrqvyefvrkessviafapedvdtppr	
Reference protein sequence	KKKRKHRLWAHCRKRIQLKKGSSNHVYNYQPCDHFRQPCDSSPCVIAQNFCEKFCQCSSECQNRFGC	551
Cloned EZH2 variant	KKKRKHRLWAHCRKRIQLKKGSSNHVYNYQPCDHFRQPCDSSPCVIAQNFCEKFCQCSSECQNRFGC	560
Consensus	kkkrkhrlwaahcrkiqlkkgssnhvynyqpcdhfrqpcdsspcviahqnfcekfqcqcssecqnrfgc	
Reference protein sequence	RCKAQCNTRKQPCYLAVRECDPDLCLTCGAADHWDSKNVSKNCSIQRGSKKHLHLLAPSDVAGWGIFFKD	621
Cloned EZH2 variant	RCKAQCNTRKQPCYLAVRECDPDLCLTCGAADHWDSKNVSKNCSIQRGSKKHLHLLAPSDVAGWGIFFKD	630
Consensus	rckaqcntkqpcyilavrecdpdlcltcgaadhwdsknvskncsiqrgskkhlhllapsdvagwgifkd	
Reference protein sequence	PVQRNEFISEYCGEIIISQDEADRRGKVDYRMCSEFLNLDNDFVVDATRKGKIRFANHSVNPNCYAKVM	691
Cloned EZH2 variant	PVQRNEFISEYCGEIIISQDEADRRGKVDYRMCSEFLNLDNDFVVDATRKGKIRFANHSVNPNCYAKVM	700
Consensus	pvqknefiseycgeiisqdeadrrgkvdyrmcseflnldndfvvdatrkgkirkfanhsvnpncyakvm	
Reference protein sequence	MVNGDHRIGIFAKRAIQTGEELFFDYRYSQADALKYVGIEREMEIP	736
Cloned EZH2 variant	MVNGDHRIGIFAKRAIQTGEELFFDYRYSQADALKYVGIEREMEIP	745
Consensus	mvngdhrigifakraiqtgeelffdyrysquadalkyvgieremeip	

(c)

FIGURE 1: Sequence and molecular characteristics of cloned variant *EZH2*. (a) PCR amplification result of *EZH2* cDNA. M: DL5000 marker; lanes 1-2: heart; lanes 3-4: liver; lanes 5-6: spleen, lanes 7-8: lung; lanes 9-10: kidney; lanes 11-12: muscle. (b) Schematic representation of the genomic and alternative splicing of pig *EZH2* gene. The red boxes represent different bases. (c) Comparison of *EZH2* protein amino acid sequence.

RNA (snRNA) [34]. Previous studies reported that alternative splicing variants of *EZH2* can influence biological effects of PRC2 to function in cell differentiation [18], central nervous system [19], and tumorigenesis [35]. In this study, we identified a transcript variant with a 27 bp insertion and a

missense mutation from several porcine fetal tissues. For our cloned sequence, we found that the 27 bp insertion was located at the position of g.109412783 to g.109412809 in the genomic sequence of the *EZH2* gene, which has been spliced out as part of intron 4 in the reference sequence

TABLE 1: Subcellular localization of the EZH2 protein predicted by four predictors.

Predictor	Predicated localization	Selected parameters
Euk-mPLoc	Nucleus	Default
CELLO ^a	Nucleus (4.750), cytoplasmic (0.133), extracellular (0.052), mitochondrial (0.018), chloroplast (0.013), plasma membrane (0.012), peroxisomal (0.007), ER (0.004), cytoskeletal (0.004), Golgi (0.003), vacuole (0.003), lysosomal (0.002)	Eukaryotes
WoLF PSORT ^b	Nucleus (26.0), nucleus and cytoplasmic (15.5), cytoplasmic (3), mitochondria (2), plastid (1)	Animal
TargetP ^c	mTP (0.139), SP (0.035), other (0.886)	Nonplant No cutoff set

^aOutput results in SVM score. ^b32 nearest neighbors are used for this prediction and 26 of them indicated that EZH2 was located in the nucleus. ^cReliability class from 1 to 5 was used to score the prediction accuracy, while 1 indicates the strongest prediction. mTP: mitochondrial targeting peptide; SP: signal peptide.

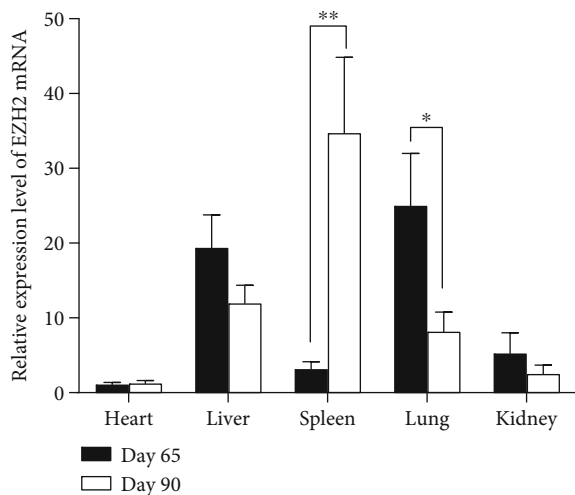


FIGURE 3: The qRT-PCR results of the *EZH2* gene in fetal tissues. The data shown in the figure are means \pm SEM from three half sibs with the same gestation day. * $P < 0.05$ and ** $P < 0.01$.

previously reported. This indicated that it was not caused by DNA variation. Further, the bases of insertion followed the GT-AG rule and were located at the 5' end of introns. Therefore, we drew a primary assumption that the transcript variant might be caused by the alternative 5' splice site. By searching in the NCBI conserved domain database, we found that the insertion was present outside of the conserved domain. The missense G \rightarrow A mutation occurred in the PRC2 HTH 1 domain which participates in the regulation of H3K27me3 as a part of the N-lobe of PRC2 [36]. This discrepancy with the reference sequence was also found in all other transcript variants of the *EZH2* gene in NCBI. We proposed that it may be caused by the inaccuracy of sequencing in the reference sequence. The specific reason needed further investigation. In terms of tertiary structure, we found that the alpha helix proportion was high in the EZH2 protein. Furthermore, amino acid insertion caused additional alpha helices, which might increase the stability of the protein [37]. These results were useful for further study in the alternative splicing, conserved domain, and spatial structure of the EZH2 protein in porcine embryo.

Determining subcellular localization for a protein is significant to investigate its interaction partners, functions, and potential roles in different cells [38, 39]. In recent

decades, many predictors varying on categories of sequence representation methods and classifiers have been proposed to predict protein subcellular localization with higher accuracy [28, 29, 40–43]. For example, WoLF PSORT is based on sorting signals, amino acid composition, and functional motifs to convert protein sequences into numerical localization features, and then, simple k -nearest neighbor classifier was used for prediction [28]. Wang et al. used WoLF PSORT to predict subcellular localization of Bv1 fxre and Bv6 nyuw and subsequently validated the prediction by lab experiments [44]. A similar validation was reported in VP1 protein [45]. Therefore, computational analysis can be applied to predict the subcellular localization of protein. A previous study demonstrated that *EZH2* is highly conserved in many species with about 90% homology [8]. The EZH2 protein was located in the nucleus in mouse embryonic cell and epithelial cells [46, 47] and human renal cell [48, 49]. In this study, the subcellular localization of the EZH2 protein in pigs was predicted by four well-developed predictors (Euk-mPLoc, CELLO, WoLF PSORT, and TargetP), rather than depending on a single predictor. Computational results predicted consistently that the EZH2 protein was located in the nucleus. Based on these findings, we speculated that EZH2 might exert function in the cell nucleus.

Understanding the spatial-temporal expression patterns of a gene is helpful to investigate its functions during the development of an organism [50]. Thus, we used qRT-PCR to detect the expression patterns of *EZH2* in different developmental stages of porcine embryo. Results showed that *EZH2* was ubiquitously expressed in tissues. The transcript abundance for *EZH2* varied significantly in the spleen and lung between 65-dpc and 90-dpc. Furthermore, we performed western blots to examine whether the EZH2 protein shared a similar expression pattern with that at the transcript level. Results showed that the EZH2 protein was expressed in all tested tissues, and its expression level appeared to have a decreasing tendency especially in the spleen and lung with development, which was consistent with the results in murine embryo [13, 46]. But in humans, EZH2 was highly expressed in the late stage of human embryonic development than in the middle stage [11], indicating that EZH2 might exert its function with different mechanisms in species. In particular, compared with the EZH2 protein, the expression level of *EZH2* mRNA exhibited an opposite tendency in the spleen from 65-dpc to 90-dpc. This finding suggested

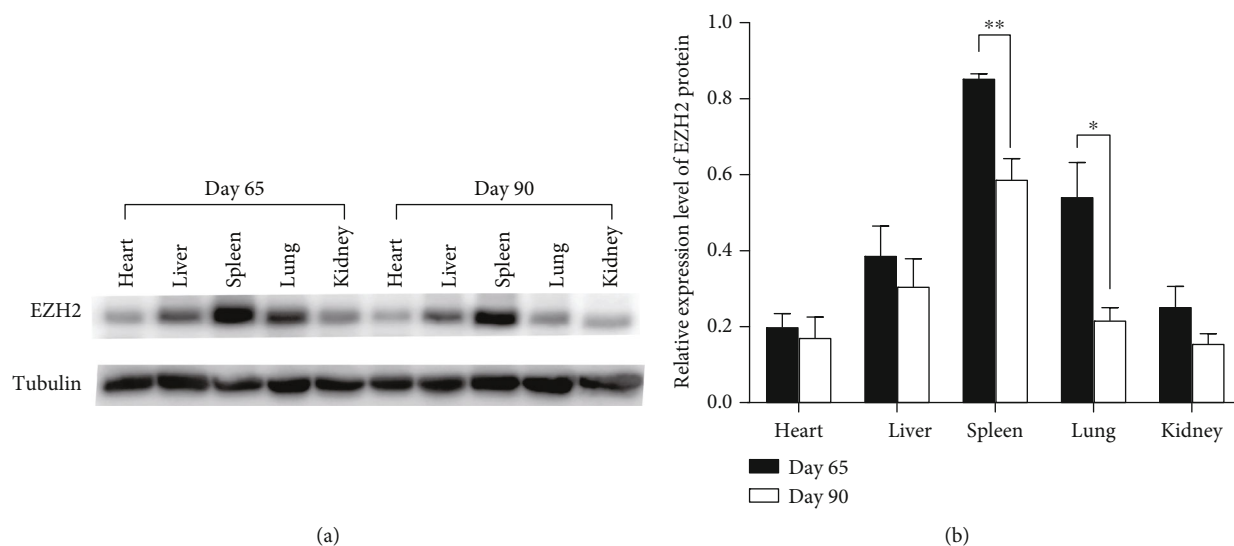


FIGURE 4: The western blot results of EZH2 in fetal tissues. (a) Western blot analysis for EZH2 protein. (b) Relative expression level of EZH2 protein. The data in the figure show means \pm SEM from three half sibs with the same gestation day. * $P < 0.05$ and ** $P < 0.01$.

dynamic posttranscript modification of EZH2 in fetal spleen, as reported in astrocytic tumor cell [51]. Overall, our data provide foundation and clue to investigate the function of EZH2 in embryonic development, and further studies are still needed.

5. Conclusions

Our study validated a transcript variant of *EZH2* by cloning and sequencing in the fetal tissues of Duroc pigs. The sequence was 2241 bp, encoding 746 amino acids, with a 9 amino acid insertion and an amino acid substitution. The insertion caused more alpha helices and was located outside of the conserved domain. The subcellular localization of the EZH2 protein was predicted in the nucleus. Additionally, *EZH2* was ubiquitously expressed in different fetal tissues, and the expression levels varied in different tissues and gestation stages. In conclusion, we provide a comprehensive analysis of sequencing, molecular characterization, and expression pattern of EZH2, which is helpful for further research on regulatory mechanism and function in porcine embryonic development.

Data Availability

All the data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

Baohua Tan and Linjun Hong contributed equally to this work.

Acknowledgments

This study was supported by the Natural Science Foundation of Guangdong Province, China (2017A030310001 and 2018B020203002), and the Natural Science Foundation of China (31802036).

Supplementary Materials

Supplementary Figure 1: EZH2 expression in muscle. (a) The expression of EZH2 mRNA in muscle. (b) The expression of EZH2 protein in muscle. (*Supplementary Materials*)

References

- [1] Y. Qu, Y. Wang, and J. Qiao, "PHF1 is required for chromosome alignment and asymmetric division during mouse meiotic oocyte maturation," *Cell Cycle*, vol. 17, no. 21-22, pp. 2447-2459, 2018.
- [2] C. Han Li and Y. Chen, "Targeting EZH2 for cancer therapy: progress and perspective," *Current Protein & Peptide Science*, vol. 16, no. 6, pp. 559-570, 2015.
- [3] O. Hobert, B. Jallal, and A. Ullrich, "Interaction of Vav with ENX-1, a putative transcriptional regulator of homeobox gene expression," *Molecular and Cellular Biology*, vol. 16, no. 6, pp. 3066-3073, 1996.
- [4] Y. Gong, L. Huo, P. Liu et al., "Polycomb group protein EZH2 is frequently expressed in inflammatory breast cancer and is predictive of worse clinical outcome," *Cancer*, vol. 117, no. 24, pp. 5476-5484, 2011.
- [5] S. Varambally, S. M. Dhanasekaran, M. Zhou et al., "The polycomb group protein EZH2 is involved in progression of prostate cancer," *Nature*, vol. 419, no. 6907, pp. 624-629, 2002.
- [6] R. H. Breuer, P. J. Snijders, E. F. Smit et al., "Increased expression of the *EZH2* polycomb group gene in *BMI-1*-positive neoplastic cells during bronchial carcinogenesis," *Neoplasia*, vol. 6, no. 6, pp. 736-743, 2004.

- [7] X.-J. Huang, X. Wang, X. Ma et al., "EZH2 is essential for development of mouse preimplantation embryos," *Reproduction, Fertility, and Development*, vol. 26, no. 8, pp. 1166–1175, 2013.
- [8] Q. Cai, H. Niu, B. Zhang et al., "Effect of EZH2 knockdown on preimplantation development of porcine parthenogenetic embryos," *Theriogenology*, vol. 132, pp. 95–105, 2019.
- [9] R.-H. Chou, Y.-L. Yu, and M.-C. Hung, "The roles of EZH2 in cell lineage commitment," *American Journal of Translational Research*, vol. 3, no. 3, pp. 243–250, 2011.
- [10] L. Chen, Y. Ma, E. Y. Kim et al., "Conditional ablation of Ezh2 in murine hearts reveals its essential roles in endocardial cushion formation, cardiomyocyte proliferation and survival," *PLoS One*, vol. 7, no. 2, article e31005, 2012.
- [11] N. Z. Shuyan Liu and H. Chen, "Expression pattern of EZH2 in organogenesis of human fetuses," *Carcinogenesis, Teratogenesis & Mutagenesis*, vol. 26, pp. 247–253, 2014.
- [12] O. Hobert, I. Sures, T. Ciossek, M. Fuchs, and A. Ullrich, "Isolation and developmental expression analysis of *Enx-1*, a novel mouse Polycomb group gene," *Mechanisms of Development*, vol. 55, no. 2, pp. 171–184, 1996.
- [13] D. O'Carroll, S. Erhardt, M. Pagani, S. C. Barton, M. A. Surani, and T. Jenuwein, "The *polycomb*-group gene *Ezh2* is required for early mouse development," *Molecular and Cellular Biology*, vol. 21, no. 13, pp. 4330–4336, 2001.
- [14] K.-E. Park, L. Magnani, and R. A. Cabot, "Differential remodeling of mono- and trimethylated H3K27 during porcine embryo development," *Molecular Reproduction and Development Molecular Reproduction and Development*, vol. 76, no. 11, pp. 1033–1042, 2009.
- [15] Y. Gao, P. Hyttel, and V. J. Hall, "Regulation of H3K27me3 and H3K4me3 during early porcine embryonic development," *Molecular Reproduction and Development Molecular Reproduction and Development*, vol. 77, no. 6, pp. 540–549, 2010.
- [16] B. Xie, H. Zhang, R. Wei et al., "Histone H3 lysine 27 trimethylation acts as an epigenetic barrier in porcine nuclear reprogramming," *Reproduction (Cambridge, England)*, vol. 151, no. 1, pp. 9–16, 2016.
- [17] F. E. Baralle and J. Giudice, "Alternative splicing as a regulator of development and tissue identity," *Molecular cell biology*, vol. 18, no. 7, pp. 437–451, 2017.
- [18] W. Mu, J. Starmer, D. Yee, and T. Magnuson, "EZH2 variants differentially regulate polycomb repressive complex 2 in histone methylation and cell differentiation," *Epigenetics & Chromatin*, vol. 11, no. 1, 2018.
- [19] D. Li, H. L. Wang, X. Huang, X. Gu, W. Xue, and Y. Xu, "Identification and functional characterization of a new splicing variant of EZH2 in the central nervous system," *International Journal of Biological Sciences*, vol. 15, no. 1, pp. 69–80, 2019.
- [20] M. Shirahata-Adachi, C. Iriyama, A. Tomita, Y. Suzuki, K. Shimada, and H. Kiyoi, "Altered EZH2 splicing and expression is associated with impaired histone H3 lysine 27 trimethylation in myelodysplastic syndrome," *Leukemia Research*, vol. 63, pp. 90–97, 2017.
- [21] S. Diao, S. Huang, Z. Xu et al., "Genetic diversity of indigenous pigs from South China area revealed by SNP array," *Animals*, vol. 9, no. 6, 2019.
- [22] T. Gu, J. Shi, L. Luo et al., "Study on hematological and biochemical characters of cloned Duroc pigs and their progeny," *Animals : an open access journal from MDPI*, vol. 9, no. 11, 2019.
- [23] S. Lu, J. Wang, F. Chitsaz et al., "CDD/SPARCLE: the conserved domain database in 2020," *Nucleic Acids Research*, vol. 48, no. D1, pp. D265–D268, 2020.
- [24] C. Geourjon and G. Deléage, "SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments," *Computer Applications in the Biosciences*, vol. 11, no. 6, pp. 681–684, 1995.
- [25] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, and M. J. E. Sternberg, "The Phyre2 web portal for protein modeling, prediction and analysis," *Nature Protocols*, vol. 10, no. 6, pp. 845–858, 2015.
- [26] K.-C. Chou and H. B. Shen, "A new method for predicting the subcellular localization of eukaryotic proteins with both single and multiple sites: Euk-mPLoc 2.0," *PLoS One*, vol. 5, no. 4, article e9931, 2010.
- [27] C.-S. Yu, C.-J. Lin, and J.-K. Hwang, "Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on *n*-peptide compositions," *Protein Science*, vol. 13, no. 5, pp. 1402–1406, 2004.
- [28] P. Horton, K.-J. Park, T. Obayashi et al., "WoLF PSORT: protein localization predictor," *Nucleic Acids Research*, vol. 35, supplement 2, pp. W585–W587, 2007.
- [29] O. Emanuelsson, H. Nielsen, S. Brunak, and G. von Heijne, "Predicting subcellular localization of proteins based on their N-terminal amino acid sequence," *Journal of Molecular Biology*, vol. 300, no. 4, pp. 1005–1016, 2000.
- [30] T. Gu, G. Xu, C. Jiang, L. Hou, Z. Wu, and C. Wang, "PRDM16 represses the pig white lipogenesis through promoting lipolysis activity," *BioMed Research International*, vol. 2019, Article ID 1969413, 7 pages, 2019.
- [31] K. J. Livak and T. D. Schmittgen, "Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta C_t}$ Method," *Methods*, vol. 25, no. 4, pp. 402–408, 2001.
- [32] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH image to ImageJ: 25 years of image analysis," *Nature Methods*, vol. 9, no. 7, pp. 671–675, 2012.
- [33] Y. Li and Y. Yuan, "Alternative RNA splicing and gastric cancer," *Mutation Research*, vol. 773, pp. 263–273, 2017.
- [34] X. Roca, A. R. Krainer, and I. C. Eperon, "Pick one, but be quick: 5' splice sites and the problems of too many choices," *Genes & Development*, vol. 27, no. 2, pp. 129–144, 2013.
- [35] K. Chen, H. Xiao, J. Zeng et al., "Alternative splicing of EZH2 pre-mRNA by SF3B3 contributes to the tumorigenic potential of renal cancer," *Clinical Cancer Research*, vol. 23, no. 13, pp. 3428–3441, 2017.
- [36] N. Justin, Y. Zhang, C. Tarricone et al., "Structural basis of oncogenic histone H3K27M inhibition of human polycomb repressive complex 2," *Nature Communications*, vol. 7, no. 1, article 11316, 2016.
- [37] S. Kumar, C. J. Tsai, and R. Nussinov, "Factors enhancing protein thermostability," *Protein Engineering*, vol. 13, no. 3, pp. 179–191, 2000.
- [38] D. Yu, X. Wu, H. Shen et al., "Enhancing membrane protein subcellular localization prediction by parallel fusion of multi-view features," *IEEE Transactions on Nanobioscience*, vol. 11, no. 4, pp. 375–385, 2012.
- [39] K.-C. Chou and Y.-D. Cai, "Using functional domain composition and support vector machines for prediction of protein subcellular location," *The Journal of Biological Chemistry*, vol. 277, no. 48, pp. 45765–45769, 2002.

- [40] K.-C. Chou and H.-B. Shen, "Cell-PLoc: a package of web servers for predicting subcellular localization of proteins in various organisms," *Nature Protocols*, vol. 3, no. 2, pp. 153–162, 2008.
- [41] K. C. Chou, "Prediction of protein subcellular locations by incorporating quasi-sequence-order effect," *Biochemical and Biophysical Research Communications*, vol. 278, no. 2, pp. 477–483, 2000.
- [42] Y.-Q. Shen and G. Burger, "TESTLoc: protein subcellular localization prediction from EST data," *BMC Bioinformatics*, vol. 11, no. 1, 2010.
- [43] S. Wan, Y. Duan, and Q. Zou, "HPSLPred: an ensemble multi-label classifier for human protein subcellular location prediction with imbalanced source," *Proteomics*, vol. 17, no. 17-18, 2017.
- [44] W. Wang, Y.-Q. Sun, G.-L. Li, and S.-Y. Zhang, "Genome-wide identification, characterization, and expression patterns of the BZR transcription factor family in sugar beet (*Beta vulgaris L.*)," *BMC Plant Biology*, vol. 19, no. 1, 2019.
- [45] J.-H. Cheng, G.-H. Lai, Y.-Y. Lien et al., "Identification of nuclear localization signal and nuclear export signal of VP1 from the chicken anemia virus and effects on VP2 shuttling in cells," *Virology Journal*, vol. 16, no. 1, 2019.
- [46] Z. Ningxia, L. Shuyan, S. Zhongjing et al., "The expression pattern of polycomb group protein Ezh2 during mouse embryogenesis," *The Anatomical Record*, vol. 294, no. 7, pp. 1150–1157, 2011.
- [47] Y. Chen, W. Li, W. Li, R. Chai, and H. Li, "Spatiotemporal expression of Ezh2 in the developing mouse cochlear sensory epithelium," *Frontiers of medicine*, vol. 10, no. 3, pp. 330–335, 2016.
- [48] S. Hinz, S. Weikert, A. Magheli et al., "Expression profile of the polycomb group protein enhancer of Zeste homologue 2 and its prognostic relevance in renal cell carcinoma," *The Journal of Urology*, vol. 182, no. 6, pp. 2920–2925, 2009.
- [49] Y. Wang, Y. Chen, H. Geng, C. Qi, Y. Liu, and D. Yue, "Overexpression of YB1 and EZH2 are associated with cancer metastasis and poor prognosis in renal cell carcinomas," *Tumour Biology*, vol. 36, no. 9, pp. 7159–7166, 2015.
- [50] T. Li, H. Zhang, X. Wang et al., "Cloning, molecular characterization and expression patterns of *DMRTC2* implicated in germ cell development of male Tibetan sheep," *International Journal of Molecular Sciences*, vol. 21, no. 7, 2020.
- [51] V. Sharma, S. Purkait, S. Takkar et al., "Analysis of EZH2: micro-RNA network in low and high grade astrocytic tumors," *Brain Tumor Pathology*, vol. 33, no. 2, pp. 117–128, 2016.